Summer 8-15-2019

Understanding the Transcriptional Mechanisms Underlying Dendritic Cell Development

Prachi Bagadia
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Division of Biology and Biomedical Sciences
Immunology

Dissertation Examination Committee:
Kenneth M. Murphy, Chair
Marco Colonna
Brian Edelson
Takeshi Egawa
Chyi-Song Hsieh

Understanding the Transcriptional Mechanisms Underlying Dendritic Cell Development

by

Prachi Bagadia

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

August 2019
St. Louis, Missouri

Table of Contents

List of Figures

Acknowledgments

I owe my deepest gratitude to my mentors, colleagues, family, and friends for the completion of this thesis. First, I would like to thank my thesis advisor, Dr. Kenneth Murphy. His mentorship taught me how to think and speak clearly, how to ask questions, and how to approach a problem. These skills have applications beyond science, and for that I am grateful. I would also like to thank Theresa Murphy, who was instrumental in helping me develop a voice of my own. To both Drs. Murphy, thank you for having confidence in me. To members of my thesis committee, Drs. Colonna, Edelson, Egawa, and Hsieh: thank you for your guidance, advice, and approachability.

I also need to thank the past and present members of the Murphy lab. To Nicole Kretzer, Gary Grajales-Reyes, Xiaodi Wu, Carlos Briseño, and Vivek Durai: your guidance and deep friendship are invaluable to me. To Derek Theisen, Xiao Huang, Renèe Wu, Tiantian Liu, Sunny Kim, Jesse Davison IV, and Steve Ferris: I am so excited to see what you will accomplish. Also, special thanks to Teresa Page, who was both a friend and the reason so many things got done.

I would also like to thank my friends, who provided comfort, solidarity, and humor throughout my life. In particular, thank you to Brittany Gamble, Masha Sitnova, Priyanka Gokhale, Rachel Wong, and Siddarth Krishmamurthy. Without you, my life would not be so rich. Most importantly, I would like to thank my father and mother, Gopal and Mamta, and my brother, Vishal. The unwavering support and unconditional love I receive from you allow me to conquer the hardest obstacles, and for that, and many other things, I can never thank you enough.

Prachi Bagadia

Washington University in St. Louis

August 2019

ABSTRACT OF THE DISSERTATION

Understanding the Transcriptional Mechanisms Underlying Dendritic Cell Development

by

Prachi Bagadia

Doctor of Philosophy in Biology and Biomedical Sciences

Immunology

Washington University in St. Louis, 2019

Professor Kenneth M. Murphy, Thesis Advisor

Dendritic cells (DCs) comprise an important immune lineage that plays a critical role in initiating and sustaining the proper immune response. They can be divided into two distinct branches, classical/conventional DCs (cDCs) or plasmacytoid DCs (pDCs). cDCs can further be classified as cDC1 or cDC2. Each DC subset exerts unique functions *in vivo* and are necessary for a complete immune response.

The precise transcriptional programs underlying DC specification and commitment remain unclear. cDC1, cDC2, and pDC all arise from the common DC progenitor (CDP) in the bone marrow. How the CDP gives rise to all three DC subsets in an important outstanding question in the field. Several transcription factors have been shown to be important for the development of certain subsets. The transcription factors *Irf8*, *Batf3*, *Id2*, *Nfil3*, and *Bcl6* are required for the cDC1 lineage, while the transcription factors *Klf4* and *Notch2* are necessary for specific cDC2 subsets. pDCs rely on the transcription factors *Tcf4* and *Zeb2* for their development. Despite knowing that these factors influence DC development, the interactions between these factors and their timing of action are unclear.

Recently, understanding of how the CDP specifies benefited from identifying cDC progenitors (pre-cDCs) that were found to include clonogenic populations separately committed to cDC1 or cDC2 lineages. Two *Irf8* enhancers were found to affect cDC1 development in different stages: an E-protein dependent enhancer located 41 kilobases downstream of the transcription start site of IRF8 (+41 kb *Irf8* enhancer) is required for the specification of the pre-cDC1, and a BATF-dependent +32 kb *Irf8* enhancer required for the maturation to the cDC1. To understand the switch in *Irf8* enhancer usage during cDC1 specification, we used single-cell RNA-sequencing of the CDP and identified a cluster of cells that expressed transcription factors that influence cDC1 development, such as *Nfil3, Id2,* and *Zeb2***.** We then performed 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.

CHAPTER 1:

Innate Immune Cell Development

Contents of this chapter will be published in *Annual Reviews of Cell Development and Biology.*

Bagadia, P., Huang, X., Liu, T., and Murphy KM. Shared transcriptional control of innate lymphoid cell and dendritic cell development. *Annu. Rev. Cell Dev. Bio.* 2019. 35: 5.1 – 5:26.

1.1 Abstract

Dendritic cells (DCs) are cells of the innate immune system that are required for the potent immune response. The development of this cell lineage is not fully understood, as it contains three distinct cell types that all exert unique functions *in vivo*. Recent developments in identifying clonogenic progenitors to specific DC subsets has helped to delineate the stages of specification, commitment, and maturation, but the precise transcriptional mechanisms that govern these processes is not well known. In this dissertation, we will first identify the similarities between ILCs and DCs, two innate immune cell lineages that help to promote a specialized immune response. Both cell lineages rely on similar transcription factors and transcription factor families for their developments. We attempt to recognize their shared transcriptional requirements. cDC1 specification relies on a switch in *Irf8* enhancer usage. To understand this switch, we used single-cell RNA-sequencing of the CDP and identified a cluster of cells that expressed transcription factors that influence cDC1 development, such as *Nfil3, Id2,* and *Zeb2***.** We performed genetic epistasis to determine the functional hierarchy of transcription factors involved in cDC1 specification and organized a transcriptional circuit that explains the switch in *Irf8* expression. 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. This new

understanding of cDC1 specification could be applied to cDC2 and pDC specification, as well as innate immune cell specification, lineages that rely on similar transcription factors.

1.2 Innate cell lineage function in an immune response

A potent immune response requires crosstalk and collaboration between the innate and adaptive immune systems, both of which contain highly specialized immune lineages. Innate immune responses involve many different types of innate cells, such as neutrophils, monocytes, and specifically discussed in this review, innate lymphoid cells (ILCs) and dendritic cells (DCs). Innate cells recognize pathogens through their germline-encoded receptors and first initiate a proinflammatory response that aims to contain and rapidly clear the infection. Importantly, the innate cells direct the specific type of adaptive immune response that is most effective at clearing the particular type of infection by secreting cytokines and chemokines to alert the adaptive immune response. The adaptive immune response can be divided into three types of immunity, specific to the type of pathogen that evokes the response. Type I immunity is in response to intracellular microbes, such as bacteria and viruses, while type II immunity protects against helminthes and environmental substances. Type III immunity is involved in protection against extracellular bacteria and fungi. Each of these immune responses corresponds to a specific type of T cell, cells which express antigen-specific receptors, and to specific innate cells, ILCs and DCs, that together effectively responds to a pathogen.

ILCs are often considered innate T cells, as both cell types share functional and developmental similarities. Like T cells, ILCs can be divided into several distinct subsets that correspond to the three types of immunity the cells elicit. However, unlike T cells which have antigen-specific receptors, undergo clonal selection, and expand when stimulated, ILCs do not

have antigen-specific receptors or the ability to undergo clonal selection. They instead rapidly respond to the pathogen and secrete cytokines to control the infection. Their developmental similarities have been discussed in detail in other reviews, and will be discussed only briefly in this review[1,2].

DCs are characterized as professional antigen presenting cells and are responsible for priming T cells for potent T cell activation. DCs encode a number of receptors that allow for antigen uptake, antigen processing, and antigen presentation to MHC molecules on T cells. Initially thought of as one family, DCs are now recognized for their heterogeneity in both location and function, with distinct subsets specialized for specific response. Two sets of classical/conventional DCs (cDCs), cDC1 and cDC2, as well as plasmacytoid DCs (pDCs) and epidermal Langerhans cells (LCs) have been identified[3-6].

There are similarities between T cells and their innate counterparts, ILCs and DCs, in promoting a specialized immune response. We present a model in which these three immune cell types are distinguished by the major immune effector module they promote and then discuss recent progress in understanding the development of ILC and DC subsets. Both cell lineages share unrecognized transcriptional network similarity that deepens our understanding of innate cell lineage communication with adaptive immunity.

1.3 Common effector modules between T cells, ILCs, and DCs

As discussed earlier, there are three broad types of immune responses, types I, II and III, where each type of immunity corresponds to specific ILC and DC subsets to elicit innate immunity, which in turn leads to a coordinated effort to prime a T-cell subset for the proper adaptive immune response. Type I immunity protects against intracellular pathogens, and can be

divided into two types of responses: a cytotoxic response led by NK cells, pDCs, and $CD8^+$ T cells, and an intracellular defense module held by ILC1s, cDC1s, and T_H1 cells. Type II immunity protects against helminths and environmental substances and is elicited by ILC2, a subset of cDC2, and T_H2 cells. Finally, Type III immunity is involved in protection against extracellular bacteria and fungi and involves ILC3s, a subset of cDC2, and T_H17 cells. A schematic of the types of immunity of each immune module, and the cell types involved, is presented in Figure 1.1.

Type I immunity involves cytotoxic and intracellular defense modules

NK cells, pDCs and $CD8^+$ T cells comprise the cytotoxic module to protect against viruses

NK cells and pDCs are among the first cells to initiate and mount a response to viral infection. NK cells express activating or inhibitory receptors that can recognize ligands on infected cells and can directly lyse the infected cells through granules containing pore-forming enzymes called perforin or serine proteases called granzymes. NK cells also express receptors for the pro-inflammatory cytokines IL-12, IL-15, IL-18, and type I interferons (IFNs), all of which are produced by other lineages during viral infection and are important for NK cell activation. Importantly, NK cells produce large amounts of interferon-gamma (IFN-γ), which activates macrophages for phagocytosis, and prevents viral spread.

pDCs sense viruses through surface receptors such as toll-like receptors 7 and 9 (TLRs 7 and 9) and cytosolic sensors such as RIG-I and MDA5. Their main effector function is to secrete large amounts of type I interferon to rapidly clear infected cells. Several studies have selectively ablated pDCs *in vivo* to demonstrate their function in regulating the antiviral immune response. Conditionally deleting the transcription factor *Tcf4* (E2-2) using an *Itgax*-cre that allows for

deletion in CD11c$^+$ immune cells results in an impaired innate response to mouse hepatitis virus (MHV) and an impaired adaptive response to lymphocytic choriomeningitis virus (LCMV)[7]. Specifically, pDCs are vital in sustaining a cytotoxic T lymphocyte (CTL) response against chronic viral infection[7]. Another study in which pDC-specific expression of diphtheria toxin receptor (DTR) was achieved using a human BDCA2 promoter indicated that lacking pDCs at the early stage of response against murine cytomegalovirus (MCMV) infection leads to lowered type I IFN production and attenuated NK cell activation[8]. pDCs may recruit and activate NK cells through their cytokine production[9].

CD8$^+$ T cells are essential for antiviral immunity, as they can both directly lyse cells through perforin and granzyme and provide memory to the immune system for protection against re-infection. In corroboration with earlier studies, Brewitz and Kastenmuller showed that pDCs promote CD8$^+$ T cell help in response to viral infection, and moreover that pDCs are recruited to CD8$^+$ T cell priming sites through the chemokines CCL3 and CCL4[10]. Type I IFN produced by pDCs bolstered cDC1 maturation and cross-presentation efficacy[10].

Development of both NK cells and CD8$^+$ T cells depends on the transcription factors T-bet (encoded by *Tbx21*) and Eomesodermin (Eomes, encoded by *Eomes*)[11,12]. Both transcription factors belong to the T-box transcription factor family[13]. Like other members in the family, these transcription factors share a highly similar T-box domain, suggesting that they bind to the same DNA motifs[14]. However, they possess divergent amino- and carboxy-termini, suggesting that they interact with distinct partners for activity[15]. In NK and CD8$^+$ T cells function, they play different roles, but can only partially compensate for each other [15,16]. In T cell development, T-bet appears to act before Eomes[17]. While T-bet modulates the expression of IFN-γ and IL-12Rβ2, Eomes non-redundantly drives perforin and granzyme B production[15,18]. Additionally, studies

have suggested that T-bet and Eomes mutually regulate each other to maintain a balance between NK cells and ILC1 populations[19,20]. The regulation of both transcription factors remains largely unclear and studies have indicated that exogenous factors, such as TGF-β, might modulate Eomes and T-bet expression[21].

ILC1s, cDC1s and T_H1 cells belong to the intracellular defense module

ILC1s and cDC1s provide inflammatory signals to activate T_H1 cells and promote an immune response against intracellular pathogens, such as *Toxoplasma gondii* (*T. gondii*). ILC1s are found in nearly all tissues and secrete large amounts of IFN-γ[22]. While both NK cells and ILC1s produce IFN-γ, ILC1s produce IFN-γ more quickly compared to NK cells in response to several viruses, such as MCMV, SeV and PR8 influenza virus, in part due to the crosstalk with cDC1[23-25]. cDC1s are the main and non-redundant producers of IL-12, a cytokine responsible for activating ILC1s, driving IFN-γ production, and instigating a T_H1 response[26-33].

cDC1s are required for protection against intracellular pathogens and for rejection of tumors. Studies demonstrating the importance of this cDC subset have been done with *Batf3*$^{-/-}$ mice, which specifically lack cDC1[34]. Infection of these mice with *T. gondii* showed that cDC1 produce IL-12, and furthermore that IL-12 can restore cDC1 in infected *Batf3*$^{-/-}$ mice[26,35]. Mice lacking all Batf family members (*Batf*$^{-/-}$*Batf2*$^{-/-}$ *Batf3*$^{-/-}$) could not restore the cDC1 population after administration of IL-12, suggesting that Batf family members may be able to compensate for the lack of Batf3 in cases of infection[26,35]. cDC1s also promote production of IFN-γ by activating invariant NK T cells (iNKT), as iNKT cells can produce IFN-γ after interacting with cDC1 pulsed with α -galactosylceramide (α-GalCer), a glycolipid antigen presented on CD1d molecules which are highly expressed on activated cDC1s[36-39]. cDC1s are also uniquely capable

of cross-presentation, which is crucial for anti-tumor immunity[34,40,41]. Recent work has indicated a necessary role for the protein Wdfy4 in cross-presentation, but the precise mechanism by which this protein, or other proteins play in the antigen-presentation process, remains unclear[42].

T-bet is also important for the development of ILC1 and T_H1 cells[19,43]. These cells are characterized by their shared function in producing IFN-γ and TNF-α and their reliance on IL-12, IL-15 and IL-18 for activation. IL-12 signaling through STAT4 activation results in T_H1 differentiation *in vivo* but is dispensable for ILC1 polarization[44,45]. In T cell development, T-bet directly represses T_H2 specification, but it is unclear whether an analogy can be drawn in the ILC population where T-bet represses ILC2 polarization[46,47]. T_H1s and ILC1s primarily participate in immune response against viral infection and intracellular pathogens but are also associated with certain inflammatory bowel diseases and Type I diabetes[2]. T-bet drives IFN-γ production by directly binding to the regulatory elements of IFNG gene and inducing the expression of Runx3, another transcription factor that drives IFN-γ production. Recent studies have indicated that Eomes is induced upon $T_H1$cell activation and that it is expressed at steady state in certain ILC1 subsets[48-50]. These data suggest that Eomes might play a role in intracellular defense module under certain circumstances, such as infection.

Type II immunity protects against helminthes and environmental substances

ILC2s, *Klf4*-dependent cDC2 and T_H2 cells contribute to maintain barrier immunity

The third immune module is characterized by immunity to helminthes infection and maintenance of barrier function. ILC2s and *Klf4*-dependent cDC2 are important innate cells for this type of immunity and T_H2 cells provide the adaptive help. This type of immunity is modulated by the cytokines IL-5, IL-9 and IL-13. ILC2s are the major innate sources of IL-5 and IL-13[51]. These

cytokines bolster the production of IL-4, which is required for T_H2 differentiation. This group of cells is further characterized by their expression of the transcription factor GATA3. GATA3 is essential for the development for both lineages and drives the production of IL-5 and IL-13 by binding directly to the promoter region of these genes, and conditional deletion of GATA3 results in the reduction of IL-5 and IL-13 production[52-58]. Upon induction, GATA3 needs to overcome a repressive threshold maintained by FOG-1 expression to stabilize the polarization of both lineages via autoactivation[59,60]. Such polarization can also be achieved via cytokine production. In particular, ILC2s in mouse small intestine produce IL-13 to actively drive differentiation of and IL-25 production by Tuft cells. IL-25 in turn further drives IL-13 production by ILC2s and maintains lineage stability[61].

While the cDC1 population appears to be homogeneous, the cDC2 population seems to be heterogeneous. Conditional deletion of Kruppel-like factor 4 (*Klf4*) results in the loss of $CD24^+$ $CD172^+$ cDC2s in the lung and lymph node, as well as a decrease in a progenitor population in the bone marrow[62]. Moreover, *Klf4*-deficient mice showed impaired protection against *Schistosoma mansoni* infection but not *herpes simplex virus*, *T. gondii*, or *Citrobacter rodentium* (*C. rodentium*) infections, indicating a specific defect in Type II but not CTL, Type I, or Type III responses[62]. Klf4 has been shown to be transcriptional activator or repressor and modulates the development of multiple lineages in epithelial tissues such as skin, lung and intestine[63-73]. However, the specific function and target of KLF4 in cDC2 remains unclear. Several studies argue that cDC2s might modulate Th2 responses to house dust mite (HDM) antigen[74,75]. Upon HDM challenge, cDC2s are rapidly recruited to lung airways and will migrate to the lymph node to induce Type II immunity[76]. Also, IL-13 produced by ILC-2s induce

CCL17 production by lung and dermal cDC2s to attract memory T_H2 cells in response to allergen[77].

Type III immunity protects against extracellular bacteria and fungi

ILC3, *Notch2*-dependent cDC2, and T_H17 cells protect against extracellular pathogens and fungi

ILC3s and *Notch2*-dependent cDC2 comprise the innate lineages for the fourth immune module. These cells are associated with defense against extracellular pathogens and fungi and can also contribute to tissue homeostasis. Deregulation of these cells often results in autoimmune diseases such as multiple sclerosis, psoriasis and Crohn's disease. STAT3 is required for proper response to IL-23 in ILC3s[78,79], whereas it is wholly required for IL-23 signaling and subsequent *in vivo* T_H17 polarization.

Conditional deletion of *Notch2* in $CD11c^+$ cells revealed that cDC2 non-redundantly produce IL-23 in response to the extracellular bacteria *C. rodentium*, a mouse model for enteropathogenic *Escherichia coli*[80]. Notch 2 is a member of Notch family transcription factors that has 4 members in mammals, Notch 1-4. This family of transcription factors all function through ligand mediate activation. Upon binding of ligand such as Delta-like family proteins, sequential proteolytic cleavages release the Notch intracellular domain (NICD). NICD then enters the nucleus and drives the expression of target genes in cooperation with several co-factors, including RBPJ and Mam. CX3Cr1 and ESAM expression can distinguish two subsets within the cDC2 population, and Notch2 deficiency result in the specific loss of the $CX3CR1^{lo}$ $ESAM^{hi}$ subset in the spleen[76,81]. Mice with conditional deletion of *Notch2* in cDCs using *Itgax*-cre have shown a decrease in T_H17 numbers in the MLN[81], and loss of Irf4 results in impaired T_H17 differentiation induced by small intestine $CD103^+CD11b^+$ cDC2s[82,83]. The same

phenomenon occurs with the specific deletion of small intestine lamina propria $CD103^{+}CD11b^{+}$ cDC2s using a human Langerin-DTA transgenic mouse model which also leads to a decrease in T_H17 cell numbers[84]. Although $ESAM^{hi}$ cDC2s are required for resistance to *C. rodentium*, mice lacking expression of Irf4 or CCR7and thereby having DCs with impaired migration capacity, do not exhibit significant susceptibility to *C. rodentium*, suggesting IL-23 production by lamina propria resident cDC2s is sufficient for the effective control of this pathogen[80,85]. As mentioned above, IL-23 is required for ILC3 secretion of IL-22 and therefore, cDC2s modulate intestine Type III immunity by targeting both T_H17 and ILC3s.

The master regulator for this module is transcription factor RORγt and a defect in this factor results in the complete absence of both lineages in vivo[86,87]. They also share a reliance on the transcription factor AHR, a receptor that binds various ligand with high risk of exposure in daily life including dietary metabolites and pollutants[88]. AHR is also essential for IL-22 production by this module[89-91].

1.4 ILC, DC, and T cell plasticity between immune models

Although all three cell types reviewed here can be divided into subsets each contributing to a specific type of immunity, recent work has shown that considerable plasticity can exist between the subsets. This imparts an important role for the local microenvironment in shaping an immune response. Plasticity has also been observed in both mouse and human.

The ability for some T cell subsets to convert to other subsets has been extensively studied[92,93]. IL-2 and IL-4, the instructive signals for T_H1 and T_H2 polarization respectively, drive lineage conversion. Recently polarized T_H1 cells can start to produce IL-4, cease IFN-γ production and convert to T_H2-like cells upon IL-4 treatment *in vitro* or helminth infection *in*

vivo[94,95]. T_H2 cells are relatively more stable than T_H1s, partially because of GATA3 auto-activation and the mutual exclusion between GATA3 and T-bet[60]. However, under circumstances where IL-12Rβ expression is restored in T_H2 polarized cells by Type I IFNs, IL-12 treatment can result in conversion to T_H1 phenotype[96]. Plasticity has also been demonstrated in T_H17 lineage, which may be the most plastic member of the T helper family[97]. Another study demonstrated that IL-12 treatment induces Th17 to downregulate RoRγt and IL-17 expression and start to express T-bet and IFN-γ and retains a Th1 like gene expression profile[98,99]. In humans, a population of cells expressing both RoRγt and T-bet that could produce both Type I and Type III cytokines was identified in patients with Crohn disease[100]..

ILC plasticity largely mirrors T cell plasticity, with most ILC plasticity in ILC3s[92,101]. *In vitro* treatment of ILC3 with IL-2 or IL-5 can transform ILC3s into IFN-γ producing ILC1-like cells[50,102]. In mice, increase of T-bet expression and Notch signaling together with a decrease in RoRγt expression converts $CCR6^-$ $NKp46^+$ ILC3s into $NK1.1^+$ ILC1s that can produce IFN-γ[103]. This conversion was later shown to be T-bet dependent[104,105]. ILC2s also show some degree of plasticity. ILC2s can produce IL-17 and be converted into ILC3s by injection of IL-25 or exposure to Notch ligand[106,107]. These converted ILC2s also induce RORγt expression[107]. Multiple groups have also shown that ILC2 secrete IFN-γ in response to IL-12 and IL-1B and convert to ILC1s[108-110].

Many of the studies of DC plasticity have focused on functional plasticity without investigating changes in transcriptional profile[111]. In this sense, plasticity refers to altered ability to stimulate T cell response. For example, DCs cultured with IFN-γ can induce T_H1 responses[112], while thymic stromal lymphopoieitin (TSLP), can strongly activate and modulate DCs to stimulate T_H2 responses in an OX40-L dependent manner[113]. However, these studies failed to

determine whether the differential ability to induce T cell response is simply because the same subset of DCs is responding to different extracellular signaling, or if there is actually subset conversion driven by differential expression of lineage defining transcription factors such as *Irf8*, *Batf3*, and *Klf4* induced by environmental cues.

1.5 Transcriptional basis of early ILC and DC development

We will now focus on the transcriptional networks governing ILC and DC development. Models of ILC and DC development can be divided into three distinct stages, as discussed for ILCs in a recent review by Serafini and colleagues[24]. Briefly, stage 1 is the specification of common precursors from a multipotent progenitor that has not excluded other cell fates. Stage 2 is the commitment of the precursors to their mature counterparts. Both stage 1 and 2 normally occur in the bone marrow. Stage 3 is the maintenance and regulation of the mature cell subsets in tissues. Both ILC and DC development are shown in Figure 1.2.

ILC development

All subsets of ILCs are found in nearly all organs and tissues in the body, but ILC progenitors develop in the fetal liver and bone marrow. In the fetal liver, ILC progenitors that are phenotypically similar to LTis arrive on day E 12.5-13.5 and subsequently express lymphotoxins to support lymphoid structure development. ILC progenitors in the bone marrow arise from the all lymphoid progenitor, or ALP, which are defined as Ly6D^{-} common lymphoid progenitor (CLP), and the IL-7Ra^{+} lymphoid-primed multipotent progenitor (LMPP)[114-121]. Another faction of the CLP, designated by positive expression of the integrin $\alpha_4\beta_7$ is thought to include the first uncommitted ILC progenitor, but further studies are needed to identify this progenitor[122]. Briefly, stage 1 is the specification of common ILC precursors from the CLP, which include the αLP,

CHILP, and the ILCP. These progenitors have largely excluded B and T cell potentials, marking them distinct from the CLP, but may not represent fully committed ILC precursors. Within the αLP progenitor, only the $CXCR6^+$ subset excludes T cell potential, but does give rise to conventional NK (cNK), non-NK ILC1, ILC2, and ILC3 cell types. However, the $CXCR6^-$ subset can develop into T cells as well[123]. This αLP might be the most uncommitted ILC progenitor found thus far, but further studies are required to dissect this population. The next progenitor, the CHILP or a common progenitor to the "helper-like" ILC lineages was discovered in 2014. The ChILP only gives rise to ILC2, ILC3, and to non-NK cells ILC1[119]. The last progenitor, the common precursor to ILCs, or ILCP, was also discovered in 2014 and expresses high levels of the transcription factor PLZF[124]. This progenitor is committed to all three ILC lineages, but excludes LTi and NK cell potentials[124]. Early innate lymphoid progenitors, or EILPs, were identified in 2015 using a reporter mouse specific for the transcription factor TCF-1, which is expressed by all ILC progenitors[125]. Recently, they were established as intermediate progenitors between ALPs and ILCPs because certain transcription factors, such as PLZF and GATA-3, were expressed at intermediate levels between ALPs and ILCPs[126]. Importantly, EILPs are a functionally distinct cell type from ALPs and ILCPs because they are specified, but not committed to the ILC lineage[126]. Progenitors identified in stage 1 represent specified cells to the ILC lineage, but because they can give rise to other cell lineages under *in vivo* or *in vitro* conditions, do not represent committed progenitors.

Commitment to specific ILC subsets occurs in stage 2 of ILC development. In this stage, identification of progenitors specific to NK cells, non-NK ILC1, ILC2, and ILC3 are aided by knowledge of T helper subsets, as ILCs and T helper cells are similar in both development and function. Recent studies have attempted to find a committed progenitor to the ILC1 lineage that

is distinct from a proposed progenitor to cNK cells. Both lineages express T-bet and Eomes.T-bet is important for the development of ILC1 and cNK cells, as *T-bet*$^{-/-}$ do not have liver or intestinal ILC1[12,18,19,119]. ILCs express T-bet, but only conventional NK cells express Eomes[20]. Additionally, while a committed progenitor for the non-NK cell ILC1 is yet to be defined, fate mapping studies of the transcription factor PLZF suggest that PLZF governs the divergence between ILC1 and NK cells[127].

A precursor committed to the ILC2 lineage, termed ILC2p, has been identified by high expression of the transcription factor GATA3[53]. While GATA3 is also expressed in the CHILP, continuous expression of GATA3 is required for ILC2 maturation and function, as described with GATA3 reporter mice[53]. GATA3$^{-/-}$ mice lack ILC1, ILC2, and a subset of ILC3, so while GATA3 might not be a commitment factor for ILC2, it is critical for the maintenance of this ILC subset[58]. Other transcription factors necessary for ILC2 development are TCF-1 (encoded by *Tcf7*) and RORα. *Tcf7*$^{-/-}$ mice lack ILC2 in the lung and lack immature ILC2 in the bone marrow[128,129]. These mice also have reduced numbers of RORγt^{+} ILCs, suggesting that TCF-1 is required for the full development of more than one ILC subset. However, TCF-1 is absolutely required for ILC2 development in a cell intrinsic manner[128,129]. Likewise, RORα is a transcription factor required for ILC2 development in a cell intrinsic manner[130,131]. The latter two transcription factors may depend on Notch signaling, similar to their dependence in the T cell lineage, but that has not been explored completely.

ILC3 development depends on the transcription factor RORγt, which is analogous to the requirement of this transcription factor inT_H17 development. RORγt is necessary for LTi cells and for NKp46^{+} ILC3s[103,118,132]. The transcription factor TOX is also necessary for full

development of LTi cells and for complete differentiation into NK cells[133]. AHR deficiency also affects the same populations as RORγt and TOX deficiency[89,90,131,134,135].

Stage 3 of ILC development is about ILC maintenance and regulation in peripheral tissues. We will not go into much detail in this review, but studies have suggested that cytokines and some transcription factors are required for control of ILC populations and can contribute to ILC plasticity as described earlier in this review.

DC development

In this review, we draw a comparison between ILC and DC development and believe that like ILC development, DC development can also be divided into three distinct stages. Furthermore, both ILC and DC development rely on some of the same transcription factors and we explore where those similarities occur and what role the transcription factor plays in either lineage.

Stage 1 begins with the specification of dendritic cell precursors from progenitor cells that are multipotent for myeloid potential called the common myeloid progenitor (CMP), to cells that are multipotent for macrophage and dendritic cell potentials, termed the MDP, to progenitors that retain only dendritic cell potential, or the CDP. DCs comprise three subsets – pDCs, cDC1, and cDC2 – and the CDP gives rise to all three subsets. Stage 2 is defined by the commitment of the specified progenitor to specific DC subsets. Recent work has defined progenitors that are committed to either cDC1 or cDC2 fate, and newer work has aimed to elucidate a progenitor that is committed to pDC fate. Stage 3 is the maintenance of DCs in peripheral tissues. Stage 1 is the specification of multipotent progenitors that have lymphoid, granulocyte and myeloid potentials to progenitors that retain only dendritic cell potential. Early progenitors that can give rise to DCs

in vivo are the CMP and the LMPP[136-139]. The next progenitor thought to arise from the CMP and still retain dendritic cell potential is the granulocyte/macrophage progenitor (GMP), but recent studies have shown that the GMP cannot develop into dendritic cells[140]. The macrophage/dendritic cell progenitor (MDP) arises from the CMP and only produces macrophages and dendritic cells both *in vitro* and *in vivo*[140-143]. The exact transcriptional mechanisms that cause the divergence and the exclusion from neutrophil fate between these progenitors are not known. However, the transcription factors PU.1, IRF8, and members of the CEBP family are thought to influence the development of these cell lineages. The MDP is thought to give rise to the CDP, but how the MDP gives rise to the CDP and how macrophage potential is lost are unanswered questions. The LMPP can give rise to other lymphoid progenitors, such as the CLP, which has the potential to give rise to pDCs[144-146]. pDCs have been thought to arise from lymphoid cells, as they can be traced with IL7R and recent work has shown that the majority of pDCs come from the CLP rather than the CDP[146].

Stage 2 of DC development is the commitment of the CDP to clonogenic progenitors that give rise to cDC1s, cDC2, or pDCs. Clonogenic progenitors for cDC1s and cDC2, namely pre-cDC1s and pre-cDC2s, were identified in 2015[147,148], and a progenitor for pDCs was elucidated in 2018[146]. Many transcription factors have been identified as important factors for cDC1, cDC2, and pDC development, and recent work has identified how transcription factors interact in the cDC1 lineage.

cDC1 development depends on expression of the transcription factors IRF8, Batf3, Nfil3, and Id2 and suppression of the transcription factor Zeb2[34,147,149-153]. IRF8 is a lineage-defining factor for cDC1 and *Irf8*$^{-/-}$ lack CDPs, pre-cDC1, and cDC1. Progenitors deficient in IRF8 diverted toward the granulocyte lineage and produced more neutrophils, indicating a role for

IRF8 in regulating myeloid/granulocyte potential[154]. Batf3, a transcription factor belonging to the Batf family, has a leucine zipper domain that heterodimerizes with JUN and IRF factors[155]. Studies aiming to understand the relationship between IRF8 and Batf3 began in 2015 when the pre-cDC1 was identified as a Lineage^{-}CD117^{int}CD135^{+}MHC-II$^{low\text{-}int}$CD11c^{+}SiglecH^{-} cell that was either CD24^{+} or *Zbtb46*$^{gfp+}$[147]. Zbtb46 is a transcription factor belonging to the Broad Complex, Tramtrack, Bric-a-Brac, and Zinc Finger family, that is selectively expressed in cDCs and their progenitors, but is not required for the development of them[156,157]. The pre-cDC1 is present in *Batf3*$^{-/-}$ mice, but not *Irf8*$^{-/-}$ mice, indicating that specification of the pre-cDC1 could occur in the absence of Batf3. *Batf3*$^{-/-}$ pre-cDC1 fail to maintain IRF8 expression, causing it to divert into the cDC2 lineage[147]. High expression of IRF8 is necessary for the cDC1 lineage, and it was discovered that Batf3 is required to maintain IRF8 autoactivation following specification to cDC1 fate[147]. ChIP-seq analysis identified a +32 kb *Irf8* enhancer containing several AP1-IRF composite elements (AICEs) that binds IRF8 and BATF3 in cDC1s *in vivo*. Recently, CRISPR-mediated deletion of the +32 kb *Irf8* enhancer in mice (*Irf8 +32*$^{-/-}$) suggests that *Batf3* supports IRF8 autoactivation using this enhancer (Durai et al, accepted). 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. 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.

Recent work from our lab has organized the transcription factors Nfil3, Id2, and Zeb2 into a transcriptional network that promotes cDC1 fate (Bagadia et al, submitted). Nfil3, a basic leucine zipper (bZIP) transcriptional repressor[158,159], is expressed in, and required for cDC1, but not cDC2 or pDC, development[160,161]. Id2 is a known inhibitor of E proteins and is expressed in cDC1 and cDC2, but not in pDCs, and is required only for cDC1 development[162,163]. Current models propose that Id2 excludes the pDC fate in DC progenitors by blocking activity of E proteins, particularly E2-2 (*Tcf4*), required for pDCs[164-166]. The transcriptional repressor Zeb2 is expressed in pDCs and cDC2s, but not cDC1s. It acts to suppress cDC1 development and is required for pDC development, perhaps through inhibition of *Id2* transcription[152,153].

We found that 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 of the CDP identified a fraction of the CDP that is already specified to cDC1 fate, in a stage earlier than the pre-cDC1. This fraction already expressed Id2, Batf3, and Zbtb46, and excluded Zeb2. 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 (Bagadia et al, submitted).

Transcriptional mechanisms governing cDC2 and pDC development are less known, but progenitors for each lineage have been identified. The pre-cDC2 was identified in 2015 as a Lineage^{-}CD117^{low}CD135^{+}CD115^{+}MHC-II^{-}CD11c^{+}*Zbtb46*$^{gfp+}$ cell in the BM[147]. Although the pre-cDC2 expresses IRF8, mature cDC2 only express IRF4.However, *Irf4*$^{-/-}$ mice do not lack cDC2, although the cDC2 that develop exhibit defective migration[85]. As discussed earlier in the

review, two transcription factors, Klf4 and Notch2, have selectively ablated specific cDC2 populations, but how heterogeneity in the cDC2 lineage occurs, and how the pre-cDC2 specifies to each cDC2 population remains unclear. Additionally, how the pre-cDC2 loses IRF8 to become IRF4-dependent is unknown.

pDC development depends on the transcription factors Tcf4, Zeb2, and Bcl11a[153,164,167,168]. pDCs also express high levels of IRF8, but are present in *Irf8*$^{-/-}$ mice with altered phenotype and functionality[151]. The basis for lineage divergence between pDCs and cDCs from the CDP is not known, but analysis of the +41 kb *Irf8* enhancer might suggest that a shared progenitor between pDCs and cDC1s exist. Alternatively, a pDCs might arise completely separately from the CLP, as has been suggested in a recent work that identified a pre-pDC[146]. This study characterized a pre-pDC as a $Lin^{-}CD16/32^{-}B220^{-}Ly6C^{-}CD117^{int/lo}CD135^{+}CD115^{-}CD127^{+}SiglecH^{+}Ly6D^{+}$ cell[146]. Pre-pDCs express high levels of IRF8 and once matured, express Tcf4.

Stage 3 of DC development concerns maintenance and regulation in peripheral tissues. Studies have suggested that cytokines and some transcription factors are required for control of DC populations and may regulate DC plasticity in tissues as discussed earlier in the review.

1.6 Analogies between ILC and DC development

ILC and DC development depends on some of the same transcription factor families, and often, the same transcription factors. Here, we will discuss the similarities between the development of these lineages by studying E proteins/Id proteins, Nfil3, and Zeb2.

E proteins and Id protein family member interactions are often portrayed as "transcriptional switches" in the development or function of specific immune subsets. Both types

of proteins belong to the basic-helix-loop-helix (bHLH) family of transcription factors and exert both transcriptional activation and transcriptional repressive roles in the immune system. bHLH family members contain two protein domains that are highly conserved but functionally different. The amino-terminal of the proteins contains the basic region, which allows binding to DNA at a specific sequence, known as an E box. The carboxy-terminal of the proteins contains the HLH domain which allows hetero- or homo-dimeric binding to other protein subunits. bHLH proteins also contain two activation domains, ADI and ADII, which map to regions that are distinct from the zipper[169]. These activation domains were identified in the N-terminal half of E2A and are conserved in the E protein sub-family. These activation domains can function independently of bHLH and employ different roles. ADI is active in many cells types[169] and can recruit the SAGA chromatin-remodeling complex[170]. ADII is thought to direct transcriptional activation, as site-directed mutagenesis at this region decreased trans-activation potential[169].

E proteins are a member of the class I bHLH family and canonically bind to the DNA sequence CAnnTG. There are three known E proteins in mouse: E2A, HEB, and E2-2. The *E2a* gene encodes for two proteins by alternative splicing, E12 and E47. E47 can homodimerize, while E12 can bind to other members of the bHLH family. *HEB* can also be alternatively spliced to produce HEBAlt and HEBCan.

Inhibitor of differentiation (Id) proteins are a member of the class V bHLH family and lack the basic DNA binding domain present in other bHLH family members. There are four known Id proteins in mouse: Id1, Id2, Id3, and Id4. Their primary function is to inhibit the activity of E proteins by sequestering E proteins and acting as dominant-negative inhibitors of E protein function. The HLH domain of Id proteins can heterodimerize with the bHLH domain of E proteins, thus causing nonfunctional heterodimers.

Id and E proteins are required in the initial stages of ILC development. Id2 is required in the ILC lineage to extinguish T cell potential and is required for all ILC subsets. The discovery of the ChILP as a lineage$^-$IL-7R$^+$CD135$^-\alpha_4\beta_7^+$CD25$^-$Id2^{high} progenitor cell elucidated the fate-determining role of Id and E proteins in ILC/T cell lineage split. Id2 is required for total ILC development[171][172], but Id2 is only necessary for proper CD4/CD8 T cell development. Id2 is also required specifically for the differentiation of ILC22s and type 2 ILCs and for the induction of $\alpha_4\beta_7$[173]. Overexpression of Id2, on the other hand, prevents the development of T and B cells, as well as pDCS, and promotes NK and ILC cell differentiation[174]. Sequestration of E47, a protein subunit of E2a, by Id2 promotes mature NK and LTi cell development[174] but loss of both E2a and Id2 in doubly-deficient mice can restore mature NK cells in the bone marrow and LTi development[175]. These particular results suggested that Id2 does have a function in mature NK cell development, but additionally has other requirements in the bone marrow and thymus.

In DCs, Id2 is absolutely required for cDC1 development, and as discussed earlier, required for early cDC1 specification. Work done in our lab has elucidated a newfound role for E proteins in the cDC1 lineage. E proteins are required for the activity of the +41 kb *Irf8* enhancer, but *Id2* expression in the cDC1-specified fraction of the CDP blocks activity at this enhancer and allows for cDC1 fate through the activation of the +32 kb *Irf8* enhancer. In summary, for both ILC and DC lineages, E proteins and Id proteins act as switches to specify one subtype over the other.

Nfil3 (Nuclear factor Interleukin 3 regulated, also known as E4bp4) is a bZIP transcriptional regulator and regulates many diverse biological processes[176].

The N-terminal portion of its bZIP domain contains a basic motif, which directly binds to DNA. The C-terminal portion of the bZIP domain contains the leucine zipper region, which is responsible for its homo-dimerization. The Nfil3 protein also contains a unique transcriptional repression domain which is transferable, since its fusion with the GAL4 DNA binding domain leads to transcriptional repression in reporter assays.

In the hematopoietic system, Nfil3 is essential for NK cell development. Nfil3 was the first transcription factor shown to be selectively and critically required for NK cell development, as *Nfil3*$^{-/-}$ mice lack those populations but not B cells, T cells or NKT cells[177,178]. The defect in NK cells is intrinsic in nature, which leads to a failure to eliminate major histocompatibility complex class I-deficient target cells and produce IFN- γ. However, further studies suggest that the NK cell population is far more complex and may have several different origins. Firth et al. have shown that the MCMV/recombinant virus express the viral m157 glycoprotein could induce a Ly49H^{+} NK cell population in the *Nfil3*$^{-/-}$ mice[179]. The viral induced Ly49H^{+} NK cells are fully functional with respect to IFN- γ production and cytotoxicity, and could comparably produce long-lived memory NK cells. Even at steady state, the development of several tissue resident NK cells, particularly in mucosal sites turns out to be Nfil3-independent, like the salivary gland NK cells[180,181], the kidney tissue resident NK cells[182] and the uterine NK cells[183]. Nfil3 is required for the formation of Eomes-expressing NK cells, whereas Eomes^{-} NK cells develop independently of Nfil3[184,185].

Nfil3 acts early in NK cell specification, as ablation of Nfil3 in the immature NK cells in bone marrow, or mature peripheral NK cells, through the use of Ub- or Nkp46-cre lines in combination with *Nfil3* floxed mice, does not influence NK lineage maintenance or homeostasis[179]. A later study examined the different stages of NK cell progenitors in *Nfil3*$^{-/-}$

mice, and shows that Nfil3 is required at the NK lineage commitment point when NK progenitors develop from common lymphoid progenitors (CLPs)[186]. Both studies conclude that Nfil3 acts early in NK cell specification.

How Nfil3 regulates NK cell development is an open question. Male et al. have shown that *Eomes*, *T-bet* and *Id2* can rescue NK production from *Nfil3*$^{-/-}$ progenitors because Nfil3 binds directly to the regulatory regions of both Eomes and Id2, promoting their transcription[186]. Nandakumar et al. further demonstrate that the histone H2A deubiquitinase MYSM1 interacts with Nfil3 and recruits Nfil3 to the *Id2* locus. They observed that MYSM1 is involved in maintaining an active chromatin at the *Id2* locus to promote NK cell development[187]. In addition, Brady and colleagues have shown that Notch1 is another novel Nfil3 target gene. While abrogation of Notch signaling impedes NK cell production, Notch peptide ligands could rescue NK cell development from *Nfil3*$^{-/-}$ progenitors[188].

Besides NK cells, Nfil3 later has been shown to be essential for the development of nearly all innate lymphoid cell subtypes. All ILC subsets exhibit high Nfil3 expression and thus, Nfil3 deficiency leads to a compromised development of all ILC subsets in a cell-intrinsic manner[184,189,190]. The only known ILC subtype that does not require Nfil3 for its development is the uterine ILC3 which participate in maintaining tissue homeostasis and barrier immunity during pregnancy[191].

Nfil3 directs the differentiation of a committed ILC progenitor, and acts transiently to enforce ILC lineage commitment Yu et al. showed that Nfil3 is required for the development of the earliest ILC lineage progenitors, Id2^{+} CHILP and α4β7^{hi}PLZF^{+} ILC progenitor[123,192]. In the meantime, Geiger et al. generated the *Nfil3*$^{fl/fl}$ X *Nkp46*^{iCre} mice and showed that they have

normal numbers of ILC3, which indicate that Nfil3 is not required for the lineage maintenance[190]. Id2 has been reported to be a target of Nfil3 for NK cell development. Indeed, Nfil3 also directly binds to *Id2* locus, promotes Id2 expression in the CHILP, and orchestrates their emergence from CLPs. Ectopic Id2 expression in *Nfil3*$^{-/-}$ progenitors also rescues all ILC lineage development. Recently, Belz and colleagues generated the Nfil3$^{fl/fl}$Id2-CreERT2$^{+/T}$ mice to spatiotemporally delete Nfil3 in Id2-expressing cells. Their results show that all ILC lineages develop normally from those mice[122]. Thus, Nfil3 is a key factor for ILC lineage commitment, but its expression is only transiently required before Id2 expression.

For DC lineages, Nfil3 has been shown to be specifically required for cDC1 development at steady state. And *Nfil3*$^{-/-}$ mice display impaired cross presentation to CD8^{+} T cells against cell-associated antigens[160]. However, a later study suggests that cDC1 can be induced in an Nfil3-independent manner in short-term bone marrow reconstitution[161]. Mechanically, *Nfil3*$^{-/-}$ mice have been shown to have normal numbers of pre-cDC progenitors, which had significantly reduced Batf3 level[160]. Our recent work with Nfil3 in cDC1 commitment shows that Nfil3 is required for early specification, perhaps initiation, of the cDC1 lineage (Bagadia et al, submitted). As such, its role in both ILC and DC lineages suggests that Nfil3 acts early for and in lineage specification.

Zeb2 is a zinc-finger transcriptional repressor that was first shown to be a regulator of epithelial-mesenchymal transition (EMT) via interaction with Smad family proteins[193-196]. Zeb2 has two clusters of zinc fingers for DNA binding, one at each terminus. Both clusters binds to CACCT sequence and are necessary for repression[194,195]. Zeb2 can exert its repressive function by directly interacting with C-terminal binding proteins (CtBP), a known co-repressor family, via its CtBP interaction domain (CID). Most notably, Zeb2 is a known partner of Smad family

proteins and can bind with Smads 1, 2, 3, 5, and 8, known as R-smads[196]. Germline deletion of Zeb2 leads to embryonic lethality in mice[197,198]. Zeb2 performs a wide range of functions in multiple systems, ranging from dysregulation in several cancers[199] to modulating myelination of oligodendrocytes[200]. In 2011, the role of Zeb2 in the hematopoietic system was first demonstrated, as Zeb2 deletion using Tie-2 cre or Vav-cre results in defect in HSC differentiation and homing to bone marrow[201].

Zeb2 is required for NK cell terminal maturation, as shown with a NK cell-specific deletion of Zeb2 ($Ncr1^{icre}$)[202]. NK cell-specific Zeb2 deletion results in reduced survival for mature NK cells, defect in their exit from BM, and increased susceptibility to B16F10 melanomas[202]. T-bet was shown to be necessary and sufficient to induce Zeb2 expression in NK cells, and Zeb2-deficient mature NK cells phenocopies their T-bet deficient counterparts.

Several studies had associated Zeb2 with T cell terminal differentiation and memory formation. In response to LCMV infection, Zeb2 is upregulated by $KLRG1^{hi}$ effector $CD8^+$ T cells, loss of Zeb2 expression in these cells results in the loss of antigen-specific $CD8^+$ effector cells and the impairment of generation of effector memory cells, while the formation of central memory T cells was accelerated[203204]. Later studies further demonstrate that coordinated expression of Zeb2 and its family member, Zeb1, is critical for $CD8^+$ T cell fate decision. Zeb2 promoted terminal T cell differentiation, whereas ZEB1 was critical for memory T cell survival and function, with the TGF-b signaling selectively induce Zeb1 and repress Zeb2[205].

Within the DC compartment, Zeb2 was shown to be required for pDC development[152,153]. One study argues that Zeb2 is also required for cDC2 development[152], while another indicated that Zeb2 is dispensable for cDC2 development and is instead required to actively repress

generation of cDC1 progenitors[153]. Our recent work shows that Zeb2 forms a mutually repressive loop with Id2, and is repressed by Nfil3 in the CDP to allow overall expression of Id2 for cDC1 specification (Bagadia et al, submitted). The mechanisms by which *Zeb2* influences ILC and DC lineage fates are less clear than the mechanisms by which *Id2* and *Nfil3* might influence fate, but work is currently being done in both fields to understand where and when *Zeb2* acts.

1.7 Conclusions

Major questions remain regarding how ILC and DC subsets exert distinct effector functions and how transcription factors necessary for the development of each subset function at molecular and genetic levels. It is clear that ILC and DC subsets share transcription factors and it is possible that ILC specification and DC specification are similar. This could be applied to many cell lineages, as most understanding of cell specification and development are incomplete.

In this dissertation, we attempt to understand the transcriptional mechanisms governing cDC1 development from the CDP. We found that cDC1 specification relies on two different *Irf8* enhancers: +41 kb *Irf8* enhancer is required for the transition between the CDP and the pre-cDC1, while the +32 kb *Irf8* enhancer is required for subsequent maturation to the cDC1. Both enhancers are dependent on different transcription factor families for their activity, and it was unclear why the +41 kb enhancer relies on E proteins, while the +32 kb enhancer relies on BATFs. To understand the switch in *Irf8* enhancer usage during cDC1 specification, we used single-cell RNA-sequencing of the CDP and identified a cluster of cells that expressed transcription factors that influence cDC1 development, such as *Nfil3, Id2,* and *Zeb2***.** We then performed genetic epistasis to determine the functional hierarchy of transcription factors involved in cDC1 specification. We found that 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.

1.8 Author Contributions

P.B., X.H., and T.L wrote the review article.

1.9 References

1. Robinette, M. L. and Colonna, M. Immune modules shared by innate lymphoid cells and T cells. *J Allergy Clin.Immunol* **138**, 1243-1251 (2016).

2. Eberl, G. et al., Innate lymphoid cells. Innate lymphoid cells: a new paradigm in immunology. *Science* **348**, aaa6566-(2015).

3. Steinman, R. M. and Witmer, M. D. Lymphoid dendritic cells are potent stimulators of the primary mixed leukocyte reaction in mice. *Proc.Natl.Acad.Sci U.S A* **75**, 5132-5136 (1978).

4. Schuler, G., Romani, N., and Steinman, R. M. A comparison of murine epidermal Langerhans cells with spleen dendritic cells. *J Invest Dermatol.* **85**, 99s-106s (1985).

5. Siegal, F. P. et al., The nature of the principal type 1 interferon-producing cells in human blood. *Science* **284**, 1835-1837 (1999).

6. 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).

7. Cervantes-Barragan, L. et al., Plasmacytoid dendritic cells control T-cell response to chronic viral infection. *Proc.Natl.Acad.Sci.U.S A* **109**, 3012-3017 (2012).

8. Takagi, H. et al., Plasmacytoid dendritic cells are crucial for the initiation of inflammation and T cell immunity in vivo. *Immunity.* **35**, 958-971 (2011).

9. Guillerey, C. et al., Pivotal role of plasmacytoid dendritic cells in inflammation and NK-cell responses after TLR9 triggering in mice. *Blood* **120**, 90-99 (2012).

10. Brewitz, A. et al., CD8+ T Cells Orchestrate pDC-XCR1+ Dendritic Cell Spatial and Functional Cooperativity to Optimize Priming. *Immunity* **46**, 205-219 (2017).

11. Intlekofer, A. M. et al., Effector and memory CD8+ T cell fate coupled by T-bet and eomesodermin. *Nat Immunol* **6**, 1236-1244 (2005).

12. Gordon, S. M. et al., The transcription factors T-bet and Eomes control key checkpoints of natural killer cell maturation. *Immunity* **36**, 55-67 (2012).

13. Papaioannou, V. E. The T-box gene family: emerging roles in development, stem cells and cancer. *Development* **141**, 3819-3833 (2014).

14. Zhang, J. et al., T-bet and Eomes govern differentiation and function of mouse and human NK cells and ILC1. *Eur.J Immunol* **48**, 738-750 (2018).

15. Pearce, E. L. et al., Control of effector CD8+ T cell function by the transcription factor Eomesodermin. *Science* **302**, 1041-1043 (2003).

16. Kaech, S. M. and Cui, W. Transcriptional control of effector and memory CD8+ T cell differentiation. *Nat Rev Immunol* **12**, 749-761 (2012).

17. Cruz-Guilloty, F. et al., Runx3 and T-box proteins cooperate to establish the transcriptional program of effector CTLs. *J Exp.Med* **206**, 51-59 (2009).

18. Townsend, M. J. et al., T-bet regulates the terminal maturation and homeostasis of NK and Valpha14i NKT cells. *Immunity* **20**, 477-494 (2004).

19. Daussy, C. et al., T-bet and Eomes instruct the development of two distinct natural killer cell lineages in the liver and in the bone marrow. *J Exp.Med* **211**, 563-577 (2014).

20. Pikovskaya, O. et al., Cutting Edge: Eomesodermin Is Sufficient To Direct Type 1 Innate Lymphocyte Development into the Conventional NK Lineage. *The Journal of Immunology* **196**, 1449-1454 (2016).

21. Cortez, V. S. et al., Transforming Growth Factor-beta Signaling Guides the Differentiation of Innate Lymphoid Cells in Salivary Glands. *Immunity* **44**, 1127-1139 (2016).

22. Sojka, D. K. et al., Tissue-resident natural killer (NK) cells are cell lineages distinct from thymic and conventional splenic NK cells. *Elife.* **3**, e01659-(2014).

23. Artis, D. and Spits, H. The biology of innate lymphoid cells. *Nature* **517**, 293-301 (2015).

24. Serafini, N., Vosshenrich, C. A., and Di Santo, J. P. Transcriptional regulation of innate lymphoid cell fate. *Nat Rev Immunol* **15**, 415-428 (2015).

25. Weizman, O. E. et al., ILC1 Confer Early Host Protection at Initial Sites of Viral Infection. *Cell* **171**, 795-808 (2017).

26. Mashayekhi, M. et al., CD8a+ Dendritic Cells Are the Critical Source of Interleukin-12 that Controls Acute Infection by Toxoplasma gondii Tachyzoites. *Immunity* **35**, 249-259 (2011).

27. Hou, B. et al., Critical coordination of innate immune defense against Toxoplasma gondii by dendritic cells responding via their Toll-like receptors. *Proc.Natl.Acad.Sci U.S A* **108**, 278-283 (2011).

28. Gazzinelli, R. T. et al., Parasite-induced IL-12 stimulates early IFN-gamma synthesis and resistance during acute infection with Toxoplasma gondii. *The Journal of Immunology* **153**, 2533-2543 (1994).

29. Reis e Sousa, C. et al., In vivo microbial stimulation induces rapid CD40 ligand-independent production of interleukin 12 by dendritic cells and their redistribution to T cell areas [see comments]. *Journal of Experimental Medicine* **186**, 1819-1829 (1997).

30. Bliss, S. K., Zhang, Y., and Denkers, E. Y. Murine neutrophil stimulation by Toxoplasma gondii antigen drives high level production of IFN-gamma-independent IL-12. *The Journal of Immunology* **163**, 2081-2088 (1999).

31. Bliss, S. K., Butcher, B. A., and Denkers, E. Y. Rapid recruitment of neutrophils containing prestored IL-12 during microbial infection. *The Journal of Immunology* **165**, 4515-4521 (2000).

32. Pepper, M. et al., Plasmacytoid dendritic cells are activated by Toxoplasma gondii to present antigen and produce cytokines. *The Journal of Immunology* **180**, 6229-6236 (2008).

33. Liu, C. H. et al., Cutting edge: dendritic cells are essential for in vivo IL-12 production and development of resistance against Toxoplasma gondii infection in mice

1. *The Journal of Immunology* **177**, 31-35 (2006).

34. Hildner, K. et al., Batf3 deficiency reveals a critical role for CD8alpha+ dendritic cells in cytotoxic T cell immunity. *Science* **322**, 1097-1100 (2008).

35. Tussiwand, R. et al., Compensatory dendritic cell development mediated by BATF-IRF interactions. *Nature* **490**, 502-507 (2012).

36. Arora, P. et al., A single subset of dendritic cells controls the cytokine bias of natural killer T cell responses to diverse glycolipid antigens. *Immunity* **40**, 105-116 (2014).

37. Fujii, S. et al., The linkage of innate to adaptive immunity via maturing dendritic cells in vivo requires CD40 ligation in addition to antigen presentation and CD80/86 costimulation. *Journal of Experimental Medicine* **199**, 1607-1618 (2004).

38. Fujii, S. et al., Prolonged IFN-gamma-producing NKT response induced with alpha-galactosylceramide-loaded DCs. *Nat.Immunol.* **3**, 867-874 (2002).

39. Kawano, T. et al., CD1d-restricted and TCR-mediated activation of valpha14 NKT cells by glycosylceramides. *Science* **278**, 1626-1629 (1997).

40. Fuertes, M. B. et al., Host type I IFN signals are required for antitumor CD8+ T cell responses through CD8{alpha}+ dendritic cells. *Journal of Experimental Medicine* **208**, 2005-2016 (2011).

41. Diamond, M. S. et al., Type I interferon is selectively required by dendritic cells for immune rejection of tumors. *Journal of Experimental Medicine* **208**, 1989-2003 (2011).

42. Theisen, D. J. et al., WDFY4 is required for cross-presentation in response to viral and tumor antigens. *Science* **362**, 694-699 (2018).

43. Robinette, M. L. et al., Transcriptional programs define molecular characteristics of innate lymphoid cell classes and subsets. *Nat Immunol* **16**, 306-317 (2015).

44. Hsieh, C. S. et al., Pillars article: development of TH1 CD4+ T cells through IL-12 produced by Listeria-induced macrophages. 1993. Science 260(5107): 547-549. *The Journal of Immunology* **181**, 4437-4439 (2008).

45. O'Sullivan, T. E. et al., Adipose-Resident Group 1 Innate Lymphoid Cells Promote Obesity-Associated Insulin Resistance. *Immunity* **45**, 428-441 (2016).

46. Hwang, E. S. et al., T helper cell fate specified by kinase-mediated interaction of T-bet with GATA-3. *Science* **307**, 430-433 (2005).

47. Zhu, J. et al., The transcription factor T-bet is induced by multiple pathways and prevents an endogenous Th2 cell program during Th1 cell responses. *Immunity* **37**, 660-673 (2012).

48. Knox, J. J. et al., Characterization of T-bet and eomes in peripheral human immune cells. *Front Immunol* **5**, 217-(2014).

49. Lupar, E. et al., Eomesodermin Expression in CD4+ T Cells Restricts Peripheral Foxp3 Induction. *The Journal of Immunology* **195**, 4742-4752 (2015).

50. Bernink, J. H. et al., Interleukin-12 and -23 Control Plasticity of CD127(+) Group 1 and Group 3 Innate Lymphoid Cells in the Intestinal Lamina Propria. *Immunity* **43**, 146-160 (2015).

51. von Moltke, J. and Locksley, R. M. I-L-C-2 it: type 2 immunity and group 2 innate lymphoid cells in homeostasis. *Curr.Opin.Immunol* **31**, 58-65 (2014).

52. Mjosberg, J. et al., The transcription factor GATA3 is essential for the function of human type 2 innate lymphoid cells. *Immunity* **37**, 649-659 (2012).

53. Hoyler, T. et al., The transcription factor GATA-3 controls cell fate and maintenance of type 2 innate lymphoid cells. *Immunity* **37**, 634-648 (2012).

54. Zheng, W. and Flavell, R. A. The transcription factor GATA-3 is necessary and sufficient for Th2 cytokine gene expression in CD4 T cells. *Cell* **89**, 587-596 (1997).

55. Zhu, J. et al., GATA-3 promotes Th2 responses through three different mechanisms: induction of Th2 cytokine production, selective growth of Th2 cells and inhibition of Th1 cell-specific factors. *Cell Res.* **16**, 3-10 (2006).

56. Yagi, R., Zhu, J., and Paul, W. E. An updated view on transcription factor GATA3-mediated regulation of Th1 and Th2 cell differentiation. *Int Immunol* **23**, 415-420 (2011).

57. Zhu, J. et al., Conditional deletion of Gata3 shows its essential function in T(H)1-T(H)2 responses. *Nat.Immunol.* **5**, 1157-1165 (2004).

58. Yagi, R. et al., The transcription factor GATA3 is critical for the development of all IL-7Ralpha-expressing innate lymphoid cells. *Immunity* **40**, 378-388 (2014).

59. Zhou, M. et al., Friend of GATA-1 represses GATA-3-dependent activity in CD4+ T cells. *Journal of Experimental Medicine* **194**, 1461-1471 (2001).

60. Ouyang, W. et al., Stat6-independent GATA-3 autoactivation directs IL-4-independent Th2 development and commitment. *Immunity* **12**, 27-37 (2000).

61. von Moltke, J. et al., Tuft-cell-derived IL-25 regulates an intestinal ILC2-epithelial response circuit. *Nature* **529**, 221-225 (2016).

62. Tussiwand, R. et al., Klf4 expression in conventional dendritic cells is required for T helper 2 cell responses. *Immunity* **42**, 916-928 (2015).

63. Segre, J. A., Bauer, C., and Fuchs, E. Klf4 is a transcription factor required for establishing the barrier function of the skin. *Nat.Genet.* **22**, 356-360 (1999).

64. Dang, D. T., Pevsner, J., and Yang, V. W. The biology of the mammalian Kruppel-like family of transcription factors. *Int.J.Biochem.Cell Biol.* **32**, 1103-1121 (2000).

65. Katz, J. P. et al., The zinc-finger transcription factor Klf4 is required for terminal differentiation of goblet cells in the colon. *Development* **129**, 2619-2628 (2002).

66. Ghaleb, A. M. et al., Kruppel-like factors 4 and 5: the yin and yang regulators of cellular proliferation. *Cell Res.* **15**, 92-96 (2005).

67. Feinberg, M. W. et al., The Kruppel-like factor KLF4 is a critical regulator of monocyte differentiation. *The EMBO Journal* **26**, 4138-4148 (2007).

68. Alder, J. K. et al., Kruppel-like factor 4 is essential for inflammatory monocyte differentiation in vivo. *Journal of Immunology* **180**, 5645-5652 (2008).

69. Zheng, H. et al., KLF4 gene expression is inhibited by the notch signaling pathway that controls goblet cell differentiation in mouse gastrointestinal tract. *Am.J.Physiol Gastrointest.Liver Physiol* **296**, G490-G498 (2009).

70. McConnell, B. B. and Yang, V. W. Mammalian Kruppel-like factors in health and diseases. *Physiol Rev.* **90**, 1337-1381 (2010).

71. Yamanaka, S. Induction of pluripotent stem cells from mouse fibroblasts by four transcription factors. *Cell Prolif.* **41 Suppl 1**, 51-56 (2008).

72. Yoshida, T. and Hayashi, M. Role of Kruppel-like factor 4 and its binding proteins in vascular disease. *J.Atheroscler.Thromb.* **21**, 402-413 (2014).

73. Kurotaki, D. et al., Essential role of the IRF8-KLF4 transcription factor cascade in murine monocyte differentiation. *Blood* **121**, 1839-1849 (2013).

74. Hammad, H. et al., Inflammatory dendritic cells--not basophils--are necessary and sufficient for induction of Th2 immunity to inhaled house dust mite allergen. *J Exp.Med.* **207**, 2097-2111 (2010).

75. Williams, J. W. et al., Transcription factor IRF4 drives dendritic cells to promote Th2 differentiation. *Nat.Commun.* **4**, 2990-(2013).

76. Mesnil, C. et al., Resident CD11b(+)Ly6C(-) lung dendritic cells are responsible for allergic airway sensitization to house dust mite in mice. *PLoS One* **7**, e53242-(2012).

77. Halim, T. Y. et al., Group 2 innate lymphoid cells license dendritic cells to potentiate memory TH2 cell responses. *Nat Immunol* **17**, 57-64 (2016).

78. Guo, X. et al., Induction of innate lymphoid cell-derived interleukin-22 by the transcription factor STAT3 mediates protection against intestinal infection. *Immunity* **40**, 25-39 (2014).

79. Rankin, L. C. et al., Complementarity and redundancy of IL-22-producing innate lymphoid cells. *Nat Immunol* **17**, 179-186 (2016).

80. Satpathy, A. T. et al., Notch2-dependent classical dendritic cells orchestrate intestinal immunity to attaching-and-effacing bacterial pathogens. *Nat.Immunol.* **14**, 937-948 (2013).

81. Lewis, K. L. et al., Notch2 receptor signaling controls functional differentiation of dendritic cells in the spleen and intestine. *Immunity* **35**, 780-791 (2011).

82. Schlitzer, A. et al., IRF4 Transcription Factor-Dependent CD11b(+) Dendritic Cells in Human and Mouse Control Mucosal IL-17 Cytokine Responses. *Immunity* **38**, 970-983 (2013).

83. Persson, E. K. et al., IRF4 Transcription-Factor-Dependent CD103(+)CD11b(+) Dendritic Cells Drive Mucosal T Helper 17 Cell Differentiation. *Immunity* **38**, 958-969 (2013).

84. Welty, N. E. et al., Intestinal lamina propria dendritic cells maintain T cell homeostasis but do not affect commensalism. *Journal of Experimental Medicine*(2013).

85. Bajana, S. et al., IRF4 promotes cutaneous dendritic cell migration to lymph nodes during homeostasis and inflammation. *Journal of Immunology* **189**, 3368-3377 (2012).

86. Sawa, S. et al., Lineage relationship analysis of RORgammat+ innate lymphoid cells. *Science* **330**, 665-669 (2010).

87. Ivanov, S. et al., CCR7 and IRF4-dependent dendritic cells regulate lymphatic collecting vessel permeability. *J Clin.Invest* **126**, 1581-1591 (2016).

88. Cella, M. and Colonna, M. Aryl hydrocarbon receptor: Linking environment to immunity. *Semin Immunol* **27**, 310-314 (2015).

89. Kiss, E. A. et al., Natural aryl hydrocarbon receptor ligands control organogenesis of intestinal lymphoid follicles. *Science* **334**, 1561-1565 (2011).

90. Qiu, J. et al., The aryl hydrocarbon receptor regulates gut immunity through modulation of innate lymphoid cells. *Immunity* **36**, 92-104 (2012).

91. Basu, R. et al., Th22 cells are an important source of IL-22 for host protection against enteropathogenic bacteria. *Immunity* **37**, 1061-1075 (2012).

92. Murphy, K. M. and Stockinger, B. Effector T cell plasticity: flexibility in the face of changing circumstances. *Nat Immunol* **11**, 674-680 (2010).

93. DuPage, M. and Bluestone, J. A. Harnessing the plasticity of CD4(+) T cells to treat immune-mediated disease. *Nat Rev Immunol* **16**, 149-163 (2016).

94. Szabo, S. J. et al., Developmental commitment to the Th2 lineage by extinction of IL-12 signaling. *Immunity* **2**, 665-675 (1995).

95. Panzer, M. et al., Rapid in vivo conversion of effector T cells into Th2 cells during helminth infection. *The Journal of Immunology* **188**, 615-623 (2012).

96. Hegazy, A. N. et al., Interferons direct Th2 cell reprogramming to generate a stable GATA-3(+)T-bet(+) cell subset with combined Th2 and Th1 cell functions. *Immunity* **32**, 116-128 (2010).

97. Stockinger, B. and Omenetti, S. The dichotomous nature of T helper 17 cells. *Nat Rev Immunol* **17**, 535-544 (2017).

98. Lee, Y. K. et al., Late developmental plasticity in the T helper 17 lineage. *Immunity* **30**, 92-107 (2009).

99. Lee, Y. K. et al., Developmental plasticity of Th17 and Treg cells. *Curr.Opin.Immunol* **21**, 274-280 (2009).

100. Annunziato, F. et al., Phenotypic and functional features of human Th17 cells. *J Exp.Med* **204**, 1849-1861 (2007).

101. Cherrier, D. E., Serafini, N., and Di Santo, J. P. Innate Lymphoid Cell Development: A T Cell Perspective. *Immunity* **48**, 1091-1103 (2018).

102. Cella, M., Otero, K., and Colonna, M. Expansion of human NK-22 cells with IL-7, IL-2, and IL-1beta reveals intrinsic functional plasticity. *Proc.Natl Acad.Sci.U S A* **107**, 10961-10966 (2010).

103. Vonarbourg, C. et al., Regulated expression of nuclear receptor RORgammat confers distinct functional fates to NK cell receptor-expressing RORgammat(+) innate lymphocytes. *Immunity* **33**, 736-751 (2010).

104. Klose, C. S. et al., A T-bet gradient controls the fate and function of CCR6-RORgammat+ innate lymphoid cells. *Nature* **494**, 261-265 (2013).

105. Rankin, L. C. et al., The transcription factor T-bet is essential for the development of NKp46+ innate lymphocytes via the Notch pathway. *Nat.Immunol.* **14**, 389-395 (2013).

106. Huang, Y. et al., IL-25-responsive, lineage-negative KLRG1(hi) cells are multipotential 'inflammatory' type 2 innate lymphoid cells. *Nat Immunol* **16**, 161-169 (2015).

107. Zhang, K. et al., Cutting Edge: Notch Signaling Promotes the Plasticity of Group-2 Innate Lymphoid Cells. *The Journal of Immunology* **198**, 1798-1803 (2017).

108. Lim, A. I. et al., IL-12 drives functional plasticity of human group 2 innate lymphoid cells. *J Exp.Med* **213**, 569-583 (2016).

109. Ohne, Y. et al., IL-1 is a critical regulator of group 2 innate lymphoid cell function and plasticity. *Nat Immunol* **17**, 646-655 (2016).

110. Silver, J. S. et al., Inflammatory triggers associated with exacerbations of COPD orchestrate plasticity of group 2 innate lymphoid cells in the lungs. *Nat Immunol* **17**, 626-635 (2016).

111. Pulendran, B., Tang, H., and Denning, T. L. Division of labor, plasticity, and crosstalk between dendritic cell subsets

4. *Curr.Opin.Immunol* **20**, 61-67 (2008).

112. Ochsenbein, A. F. et al., Roles of tumour localization, second signals and cross priming in cytotoxic T-cell induction. *Nature* **411**, 1058-1064 (2001).

113. Ito, T. et al., TSLP-activated dendritic cells induce an inflammatory T helper type 2 cell response through OX40 ligand. *Journal of Experimental Medicine* **202**, 1213-1223 (2005).

114. Inlay, M. A. et al., Ly6d marks the earliest stage of B-cell specification and identifies the branchpoint between B-cell and T-cell development. *Genes & Development* **23**, 2376-2381 (2009).

115. Moro, K. et al., Innate production of T(H)2 cytokines by adipose tissue-associated c-Kit(+)Sca-1(+) lymphoid cells. *Nature* **463**, 540-544 (2010).

116. Possot, C. et al., Notch signaling is necessary for adult, but not fetal, development of RORgammat(+) innate lymphoid cells. *Nat.Immunol.* **12**, 949-958 (2011).

117. Yang, Q. et al., Cutting edge: Natural helper cells derive from lymphoid progenitors. *The Journal of Immunology* **187**, 5505-5509 (2011).

118. Cherrier, M. and Eberl, G. The development of LTi cells. *Curr.Opin.Immunol* **24**, 178-183 (2012).

119. Klose, C. S. et al., Differentiation of Type 1 ILCs from a Common Progenitor to All Helper-like Innate Lymphoid Cell Lineages. *Cell* **157**, 340-356 (2014).

120. Ghaedi, M. et al., Common-Lymphoid-Progenitor-Independent Pathways of Innate and T Lymphocyte Development. *Cell Rep.* **15**, 471-480 (2016).

121. Ishizuka, I. E. et al., The Innate Lymphoid Cell Precursor. *Annu.Rev Immunol* **34**, 299-316 (2016).

122. Seillet, C. et al., Deciphering the Innate Lymphoid Cell Transcriptional Program. *Cell Rep.* **17**, 436-447 (2016).

123. Yu, X. et al., The basic leucine zipper transcription factor NFIL3 directs the development of a common innate lymphoid cell precursor. *Elife.* **3**, (2014).

124. Constantinides, M. G. et al., A committed precursor to innate lymphoid cells. *Nature*(2014).

125. Yang, M. et al., PDK1 orchestrates early NK cell development through induction of E4BP4 expression and maintenance of IL-15 responsiveness. *J Exp.Med* **212**, 253-265 (2015).

126. Harly, C. et al., Development and differentiation of early innate lymphoid progenitors. *J Exp.Med* **215**, 249-262 (2018).

127. Constantinides, M. G. et al., PLZF expression maps the early stages of ILC1 lineage development. *Proc.Natl Acad.Sci.U S A* **112**, 5123-5128 (2015).

128. Yang, Q. et al., T cell factor 1 is required for group 2 innate lymphoid cell generation. *Immunity* **38**, 694-704 (2013).

129. Mielke, L. A. et al., TCF-1 controls ILC2 and NKp46+RORgammat+ innate lymphocyte differentiation and protection in intestinal inflammation. *The Journal of Immunology* **191**, 4383-4391 (2013).

130. Wong, S. H. et al., Transcription factor RORalpha is critical for nuocyte development. *Nat.Immunol.* **13**, 229-236 (2012).

131. Halim, T. Y. et al., Retinoic-acid-receptor-related orphan nuclear receptor alpha is required for natural helper cell development and allergic inflammation. *Immunity* **37**, 463-474 (2012).

132. Eberl, G. et al., An essential function for the nuclear receptor RORgamma(t) in the generation of fetal lymphoid tissue inducer cells. *Nat Immunol* **5**, 64-73 (2004).

133. Aliahmad, P., de la, Torre B., and Kaye, J. Shared dependence on the DNA-binding factor TOX for the development of lymphoid tissue-inducer cell and NK cell lineages. *Nat.Immunol.* **11**, 945-952 (2010).

134. Lee, J. S., Cella, M., and Colonna, M. AHR and the Transcriptional Regulation of Type-17/22 ILC. *Front Immunol.* **3**, 10-(2012).

135. Lee, J. S. et al., AHR drives the development of gut ILC22 cells and postnatal lymphoid tissues via pathways dependent on and independent of Notch. *Nat.Immunol.* **13**, 144-151 (2012).

136. Traver, D. et al., Development of CD8alpha-positive dendritic cells from a common myeloid progenitor. *Science* **290**, 2152-2154 (2000).

137. 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).

138. Schlenner, S. M. et al., Fate mapping reveals separate origins of T cells and myeloid lineages in the thymus. *Immunity* **32**, 426-436 (2010).

139. Naik, S. H. et al., Diverse and heritable lineage imprinting of early haematopoietic progenitors. *Nature* **496**, 229-232 (2013).

140. Yanez, A. et al., Granulocyte-Monocyte Progenitors and Monocyte-Dendritic Cell Progenitors Independently Produce Functionally Distinct Monocytes. *Immunity* **47**, 890-902 (2017).

141. Fogg, D. K. et al., A clonogenic bone marrow progenitor specific for macrophages and dendritic cells. *Science* **311**, 83-87 (2006).

142. Auffray, C. et al., CX3CR1+ CD115+ CD135+ common macrophage/DC precursors and the role of CX3CR1 in their response to inflammation. *J Exp.Med.* **206**, 595-606 (2009).

143. Auffray, C., Sieweke, M. H., and Geissmann, F. Blood monocytes: development, heterogeneity, and relationship with dendritic cells. *Annu.Rev Immunol* **27**, 669-692 (2009).

144. Yang, L. et al., Identification of Lin(-)Sca1(+)kit(+)CD34(+)Flt3- short-term hematopoietic stem cells capable of rapidly reconstituting and rescuing myeloablated transplant recipients. *Blood* **105**, 2717-2723 (2005).

145. Sathe, P. et al., Lymphoid tissue and plasmacytoid dendritic cells and macrophages do not share a common macrophage-dendritic cell-restricted progenitor. *Immunity* **41**, 104-115 (2014).

146. Rodrigues, P. F. et al., Distinct progenitor lineages contribute to the heterogeneity of plasmacytoid dendritic cells. *Nat Immunol* **19**, 711-722 (2018).

147. Grajales-Reyes, G. E. et al., Batf3 maintains autoactivation of Irf8 for commitment of a CD8alpha(+) conventional DC clonogenic progenitor. *Nat Immunol* **16**, 708-717 (2015).

148. 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).

149. 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).

150. 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).

151. Sichien, D. et al., IRF8 Transcription Factor Controls Survival and Function of Terminally Differentiated Conventional and Plasmacytoid Dendritic Cells, Respectively. *Immunity* **45**, 626-640 (2016).

152. 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).

153. 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).

154. Schonheit, J. et al., PU.1 level-directed chromatin structure remodeling at the Irf8 gene drives dendritic cell commitment. *Cell Rep.* **3**, 1617-1628 (2013).

155. Vinson, C. et al., Classification of human B-ZIP proteins based on dimerization properties. *Molecular & Cellular Biology* **22**, 6321-6335 (2002).

156. Satpathy, A. T. et al., Zbtb46 expression distinguishes classical dendritic cells and their committed progenitors from other immune lineages. *Journal of Experimental Medicine* **209**, 1135-1152 (2012).

157. Satpathy, A. T. et al., Re(de)fining the dendritic cell lineage. *Nat.Immunol.* **13**, 1145-1154 (2012).

158. Cowell, I. G., Skinner, A., and Hurst, H. C. Transcriptional repression by a novel member of the bZIP family of transcription factors. *Mol Cell Biol* **12**, 3070-3077 (1992).

159. 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).

160. Kashiwada, M. et al., NFIL3/E4BP4 is a key transcription factor for CD8{alpha}+ dendritic cell development. *Blood* **117**, 6193-6197 (2011).

161. Seillet, C. et al., CD8alpha+ DCs can be induced in the absence of transcription factors Id2, Nfil3, and Batf3. *Blood* **121**, 1574-1583 (2013).

162. 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).

163. Hacker, C. et al., Transcriptional profiling identifies Id2 function in dendritic cell development. *Nat Immunol* **4**, 380-386 (2003).

164. 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).

165. Watowich, S. S. and Liu, Y. J. Mechanisms regulating dendritic cell specification and development. *Immunol.Rev.* **238**, 76-92 (2010).

166. Grajkowska, L. T. et al., Isoform-Specific Expression and Feedback Regulation of E Protein TCF4 Control Dendritic Cell Lineage Specification. *Immunity* **46**, 65-77 (2017).

167. Cisse, B. et al., Transcription factor E2-2 is an essential and specific regulator of plasmacytoid dendritic cell development. *Cell* **135**, 37-48 (2008).

168. 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).

169. Aronheim, A. et al., The E2A gene product contains two separable and functionally distinct transcription activation domains. *Proc.Natl Acad.Sci.U S A* **90**, 8063-8067 (1993).

170. Massari, M. E. et al., A conserved motif present in a class of helix-loop-helix proteins activates transcription by direct recruitment of the SAGA complex. *Mol Cell* **4**, 63-73 (1999).

171. Verykokakis, M., Zook, E. C., and Kee, B. L. ID'ing innate and innate-like lymphoid cells. *Immunol Rev* **261**, 177-197 (2014).

172. Yokota, Y. et al., Development of peripheral lymphoid organs and natural killer cells depends on the helix-loop-helix inhibitor Id2. *Nature* **397**, 702-706 (1999).

173. Hwang, Y. Y. and McKenzie, A. N. Innate lymphoid cells in immunity and disease. *Adv.Exp.Med Biol* **785**, 9-26 (2013).

174. Mjosberg, J. et al., Transcriptional control of innate lymphoid cells. *Eur.J Immunol* **42**, 1916-1923 (2012).

175. Boos, M. D. et al., Mature natural killer cell and lymphoid tissue-inducing cell development requires Id2-mediated suppression of E protein activity. *Journal of Experimental Medicine* **204**, 1119-1130 (2007).

176. Keniry, M. et al., Survival factor NFIL3 restricts FOXO-induced gene expression in cancer. *Genes & Development* **27**, 916-927 (2013).

177. Gascoyne, D. M. et al., The basic leucine zipper transcription factor E4BP4 is essential for natural killer cell development. *Nat.Immunol.* **10**, 1118-1124 (2009).

178. 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).

179. Firth, M. A. et al., Nfil3-independent lineage maintenance and antiviral response of natural killer cells. *J Exp.Med* **210**, 2981-2990 (2013).

180. Cortez, V. S. et al., Cutting edge: Salivary gland NK cells develop independently of Nfil3 in steady-state. *The Journal of Immunology* **192**, 4487-4491 (2014).

181. Erick, T. K. et al., NFIL3 Expression Distinguishes Tissue-Resident NK Cells and Conventional NK-like Cells in the Mouse Submandibular Glands. *The Journal of Immunology* **197**, 2485-2491 (2016).

182. Victorino, F. et al., Tissue-Resident NK Cells Mediate Ischemic Kidney Injury and Are Not Depleted by Anti-Asialo-GM1 Antibody. *The Journal of Immunology* **195**, 4973-4985 (2015).

183. Redhead, M. L. et al., The Transcription Factor NFIL3 Is Essential for Normal Placental and Embryonic Development but Not for Uterine Natural Killer (UNK) Cell Differentiation in Mice. *Biol Reprod.* **94**, 101-(2016).

184. Seillet, C. et al., Nfil3 is required for the development of all innate lymphoid cell subsets. *J Exp.Med* **211**, 1733-1740 (2014).

185. Crotta, S. et al., The transcription factor E4BP4 is not required for extramedullary pathways of NK cell development. *The Journal of Immunology* **192**, 2677-2688 (2014).

186. Male, V. et al., The transcription factor E4bp4/Nfil3 controls commitment to the NK lineage and directly regulates Eomes and Id2 expression. *J Exp.Med* **211**, 635-642 (2014).

187. Nandakumar, V. et al., Epigenetic control of natural killer cell maturation by histone H2A deubiquitinase, MYSM1. *Proc.Natl Acad.Sci.U S A* **110**, E3927-E3936 (2013).

188. Kostrzewski, T. et al., Multiple Levels of Control Determine How E4bp4/Nfil3 Regulates NK Cell Development. *The Journal of Immunology* **200**, 1370-1381 (2018).

189. Seillet, C. et al., Differential requirement for Nfil3 during NK cell development. *The Journal of Immunology* **192**, 2667-2676 (2014).

190. Geiger, T. L. et al., Nfil3 is crucial for development of innate lymphoid cells and host protection against intestinal pathogens. *J Exp.Med* **211**, 1723-1731 (2014).

191. Doisne, J. M. et al., Composition, Development, and Function of Uterine Innate Lymphoid Cells. *The Journal of Immunology* **195**, 3937-3945 (2015).

192. Xu, W. et al., NFIL3 orchestrates the emergence of common helper innate lymphoid cell precursors. *Cell Rep.* **10**, 2043-2054 (2015).

193. Verschueren, K. et al., SIP1, a novel zinc finger/homeodomain repressor, interacts with Smad proteins and binds to 5'-CACCT sequences in candidate target genes. *J Biol Chem.* **274**, 20489-20498 (1999).

194. Remacle, J. E. et al., New mode of DNA binding of multi-zinc finger transcription factors: deltaEF1 family members bind with two hands to two target sites. *EMBO J* **18**, 5073-5084 (1999).

195. Comijn, J. et al., The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion. *Mol Cell* **7**, 1267-1278 (2001).

196. van Grunsven, L. A. et al., Interaction between Smad-interacting protein-1 and the corepressor C-terminal binding protein is dispensable for transcriptional repression of E-cadherin. *J Biol Chem.* **278**, 26135-26145 (2003).

197. 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).

198. Van de Putte, T. et al., Mice lacking ZFHX1B, the gene that codes for Smad-interacting protein-1, reveal a role for multiple neural crest cell defects in the etiology of Hirschsprung disease-mental retardation syndrome. *Am.J Hum.Genet.* **72**, 465-470 (2003).

199. Vandewalle, C., Van Roy, F., and Berx, G. The role of the ZEB family of transcription factors in development and disease. *Cell Mol Life Sci.* **66**, 773-787 (2009).

200. Weng, Q. et al., Dual-mode modulation of Smad signaling by Smad-interacting protein Sip1 is required for myelination in the central nervous system. *Neuron* **73**, 713-728 (2012).

201. Goossens, S. et al., The EMT regulator Zeb2/Sip1 is essential for murine embryonic hematopoietic stem/progenitor cell differentiation and mobilization. *Blood* **117**, 5620-5630 (2011).

202. van Helden, M. J. et al., Terminal NK cell maturation is controlled by concerted actions of T-bet and Zeb2 and is essential for melanoma rejection. *J Exp.Med* **212**, 2015-2025 (2015).

203. Dominguez, C. X. et al., The transcription factors ZEB2 and T-bet cooperate to program cytotoxic T cell terminal differentiation in response to LCMV viral infection. *J Exp.Med* **212**, 2041-2056 (2015).

204. Omilusik, K. D. et al., Transcriptional repressor ZEB2 promotes terminal differentiation of CD8+ effector and memory T cell populations during infection. *J Exp.Med* **212**, 2027-2039 (2015).

205. Guan, T. et al., ZEB1, ZEB2, and the miR-200 family form a counterregulatory network to regulate CD8(+) T cell fates. *J Exp.Med* **215**, 1153-1168 (2018).

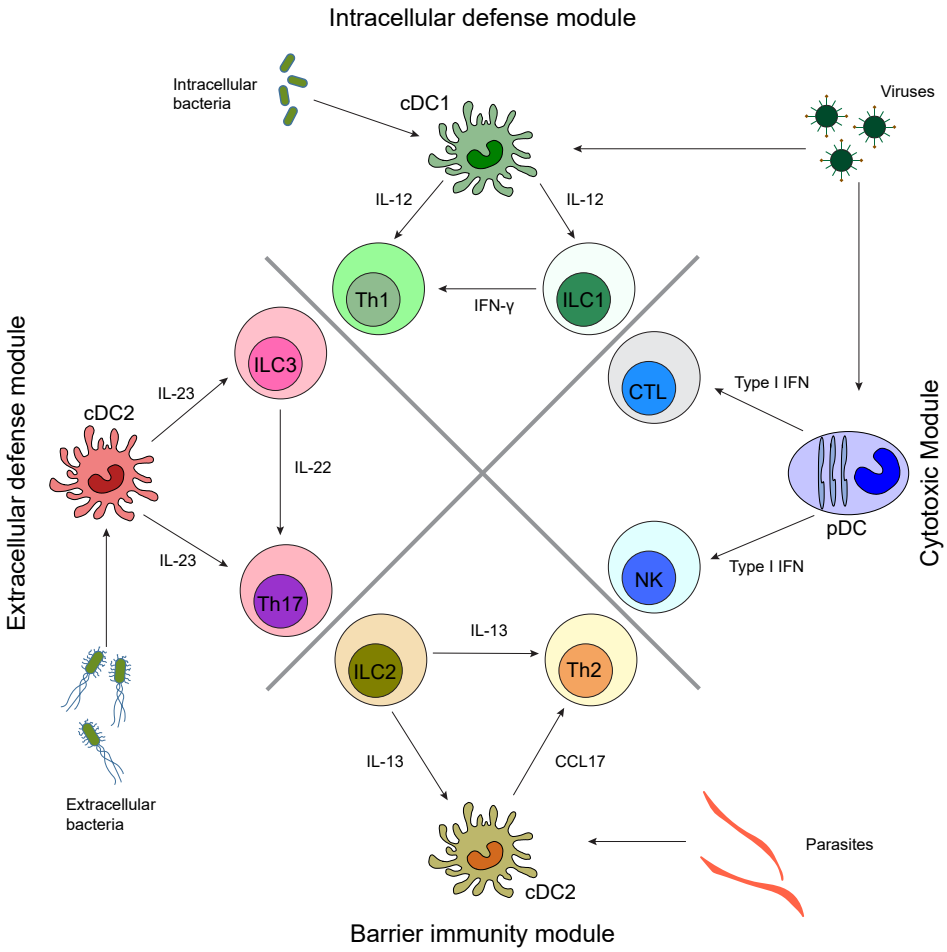

Figure 1.1 Four core immune modules shared between innate lymphoid cells (ILCs), dendritic cells (DCs) and T cells. Four core immune modules shared between ILCs, DCs, and T cells. An immune response specific to a particular pathogen involves cross talk and collaboration between two innate cell lineages, ILCs and DCs, and an adaptive cell lineage, T cells, by secretion of cytokines and chemokines. Each cell type in this module is governed by specific transcription factors that influence the effector functions. Abbreviations: cDC, classical/conventional dendritic cell; CTL, cytotoxic T lymphocyte; DC, dendritic cell; IFN, interferon; ILC, innate lymphoid cell; pDC, plasmacytoid dendritic cell.

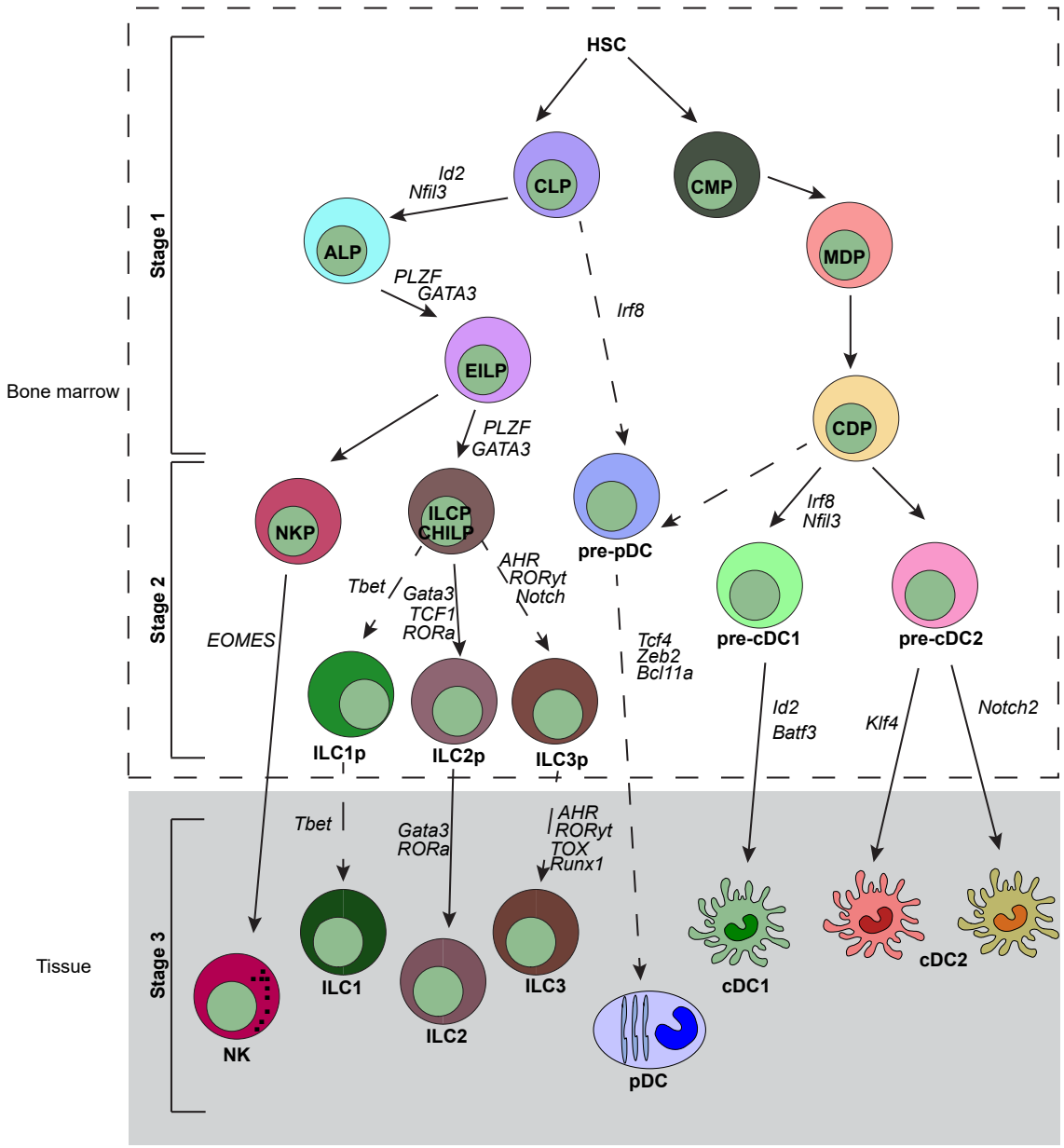

Figure 1.2 ILC and DC development can be divided into three stages. Stage 1 refers to specification of common precursors from multipotent progenitors that have not yet excluded other cell lineage fates. Stage 2 is the commitment of those common precursors to the mature cell. Stage 3 is the maintenance of those cells in tissues. Many transcription factors influence either specification or commitment, and the precise roles for those factors are still unknown. Abbreviations: ALP, all-lymphoid progenitor; cDC, classical/conventional dendritic cell; CDP, common dendritic progenitor; CHILP, common helper-like ILC progenitor; CLP, common lymphoid progenitor; CMP, common myeloid progenitor; EILP, early innate lymphoid

progenitor; HSC, hematopoietic stem cell; ILC, innate lymphoid cell; ILCP, ILC progenitor; MDP, macrophage/DC progenitor; NKP, NK progenitor; pDC, plasmacytoid dendritic cell.

CHAPTER 2:

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

The contents of this chapter will be published in *Nature Immunology* in 2019.

Bagadia, P., Huang, X., Liu, T., Durai, V., Grajales-Reyes, GE, Nitschke, M., Modrusan, Z., Granja, JM, Sathpathy, AT, Briseño CG, Gargaro M, Iwata, A, Kim, S, Chang, HY, Shaw, AS, Murphy TL, Murphy KM. An Nfil3–Zeb2–Id2 pathway imposes Irf8 enhancer switching during cDC1 development. *Nat. Immunol*.

2.1 Abstract

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.

2.2 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 (Durai, V., accepted). 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 (Durai, V., accepted). 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.

2.3 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 (Figure 2.1a)[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} (Figure 2.1b). 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 (Figure 2.1c, Figure 2.3a).

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 (Figure 2.1d,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 (Figure 2.1d,e). *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 (Figure 2.1e). 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$^-$ (Figure 2.2a) 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 (Figure 2.2b,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 Figure 2.3b). 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* (Figure 2.2d,e). Cluster 3 also showed reduced *Csf1r* expression (Figure 2.2d), consistent with lower CD115 expression in pre-cDC1 and incongruent with the higher CD115 expression in the bulk CDP (Figure 2.1a). As expected, *Flt3* and *Irf8* were uniformly and highly expressed (Figure 2.2d, 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 (Figure 2.2c,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 (Figure 2.2d, Figure 2.3b). In addition, the CDP appeared homogenous with respect to markers of proliferation (Figure 2.2f). 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 (Figure 2.5a,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 (Figure 2.4a). 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 (Figure 2.4b). 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%) (Figure 2.4c,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 (Figure 2.4d,e, Figure 2.5c,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%) (Figure 2.1c, 2.4e). In all three cases, pDCs developed exclusively from CDPs

that were either *Zbtb46*-GFP^{neg}, ZEB2-EGFP^{hi}, or *Id2*-GFP^{lo} (Figure 2.5e-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 (Figure 2.4f-i). We considered genes whose expression differed more than 5-fold between the pre-cDC1 and either ZEB2-EGFP^{hi} CDPs (Figure 3f,g) or *Id2*-GFP^{lo} CDPs (Figure 2.4h,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 (Figure 2.4f,g). 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 (Figure 2.4h,i). 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* (Figure 2.4f-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 (Figure 2.6a,b). However, these cells are absent in *Nfil3*$^{-/-}$*Zbtb46*$^{gfp/+}$ mice, but do develop normally in *Batf3*$^{-/-}$*Zbtb46*$^{gfp/+}$ mice as previously described (Figure 2.7a)[9]. Within the CDP, cDC1-specified cells can be identified as *Zbtb46*-GFP^{pos} cells that comprise 5% of the CDP (Figure 2.6a,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 (Figure 2.6c,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 (Figure 2.6c,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 (Figure 2.6e,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 (Figure 2.6e,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 (Figure 2.8). First, *Zeb2*$^{-/-}$ mice have more than a 2-fold increase in splenic cDC1s compared with wildtype mice (Figure 2.8a,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 (Figure 2.8c,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 (Figure 2.8e,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*$^{Cre\text{-}ERT2}$ 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 (Figure 2.9a-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 (Figure 2.8). *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 (Figure 2.9a-d). Moreover, similar results were obtained from *in vitro* Flt3L cultures of BM cells from these mice (2.10a,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 (Figure 2.9e, Supplementary Figure 2.10c). 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 (Figure 2.11a). We also showed that $Id2^{-/-}Zbtb46^{gfp/gfp}$ cDC2s are transcriptionally essentially identical to $Id2^{+/+}Zbtb46^{gfp/+}$cDC2s (Figure 2.11b). 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 (Figure 2.9f), and transcriptional profile (Figure 2.9g, Figure 2.11c). 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 (Figure 2.9g and Figure 2.11c). 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 (Figure 2.11d,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 (Durai, V., accepted). 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](Durai, V., accepted). This 454 bp region contains six E-box motifs that are conserved between human and murine *Irf8* loci (Figure 2.12a) and is known to bind E2-2 in human pDCs (Figure 2.13)[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 (Figure 2.12b,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 (Figure 2.12b,c). Mutation of both E-boxes 1 and 2 in the 454 bp enhancer significantly reduced enhancer activity in pDCs (Figure 2.14a,b). Within segment A, mutation of either E-box alone reduced overall activity, while mutation of both E-boxes together completely extinguished activity (Figure 2.12d, Figure 2.14c). The most active segment B was also E-box dependent, showing reduced overall activity upon mutation of E-boxes 3 and 4 (Figure 2.13e, Figure 2.14d). 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 (Figure 2.12f).

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 (Figure 2.12g, red dashed line). This peak was located at +32 kb of the *Irf8* TSS and was shown to be BATF3-dependent[9] (Durai, V., accepted). The induction of *Id2*, and the subsequent repression of *Zeb2*, thus forces a new requirement for *Batf3* in maintaining *Irf8* expression during cDC1 development.

2.4 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. It is possible that similar networks like this exist in ILC development, but that will require further study.

2.5 Materials and Methods

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 Figure 2.9f,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*$^{Cre\text{-}ERT2}$ (*Zeb2*$^{-/-}$), *Id2*$^{f/f}$ *Rosa26*$^{Cre\text{-}ERT2}$ (*Id2*$^{-/-}$) and *Zeb2*$^{f/f}$ *Id2*$^{f/f}$ *Rosa26*$^{Cre\text{-}ERT2}$ (*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-II$^-$CD11c$^+$; pre-cDC1s are Lin$^-$CD117^{int}CD135$^+$CD115$^-$MHC-II$^{lo\text{-}int}$CD11c$^+$CD24$^+$Siglec-H$^-$ or as Lin$^-$CD117^{int}CD135$^+$MHC-II$^{lo\text{-}int}$CD11c$^+$Siglec-H$^-$Zbtb46-GFP$^+$, and pre-cDC2s as Lin$^-$CD117^{lo}CD135$^+$CD115$^+$MHC-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-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-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 mE3/4*:

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

2.6 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.

2.7 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. and K.M.M. wrote the manuscript with advice from all authors.

2.8 References

1. Saxena, M. and Bhardwaj, N. Re-Emergence of Dendritic Cell Vaccines for Cancer Treatment. *Trends Cancer* **4**, 119-137 (2018).
2. Steinman, R. M. and 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 [see comments]. *Nature Medicine* **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. 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).

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. Grajales-Reyes, G. E. et al., Batf3 maintains autoactivation of Irf8 for commitment of a CD8alpha(+) conventional DC clonogenic progenitor. *Nat Immunol* **16**, 708-717 (2015).

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 cells (DCs) are derived from distinct circulating precursors that are precommitted to become CD1c+ or CD141+ 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. 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).

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).

15. Kashiwada, M. et al., NFIL3/E4BP4 is a key transcription factor for CD8{alpha}+ dendritic cell development. *Blood* **117**, 6193-6197 (2011).

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).

17. Hacker, C. et al., Transcriptional profiling identifies Id2 function in dendritic cell development. *Nat Immunol* **4**, 380-386 (2003).

18. 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).

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).

20. Cowell, I. G., Skinner, A., and Hurst, H. C. Transcriptional repression by a novel member of the bZIP family of transcription factors. *Mol Cell Biol* **12**, 3070-3077 (1992).

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).

22. 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).

23. Watowich, S. S. and Liu, Y. J. Mechanisms regulating dendritic cell specification and development. *Immunol.Rev.* **238**, 76-92 (2010).

24. Grajkowska, L. T. et al., Isoform-Specific Expression and Feedback Regulation of E Protein TCF4 Control Dendritic Cell Lineage Specification. *Immunity* **46**, 65-77 (2017).

25. Satpathy, A. T. et al., Zbtb46 expression distinguishes classical dendritic cells and their committed progenitors from other immune lineages. *Journal of Experimental Medicine* **209**, 1135-1152 (2012).

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).

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).

31. Tussiwand, R. et al., Klf4 expression in conventional dendritic cells is required for T helper 2 cell responses. *Immunity* **42**, 916-928 (2015).

32. Briseno, C. G. et al., Notch2-dependent DC2s mediate splenic germinal center responses. *Proc.Natl Acad.Sci.U S A* **115**, 10726-10731 (2018).

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).

34. Jackson, J. T. 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).

35. Calero-Nieto, F. J. et al., Key regulators control distinct transcriptional programmes in blood progenitor and mast cells. *EMBO J* **33**, 1212-1226 (2014).

36. Corces, M. R. et al., An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat Methods* **14**, 959-962 (2017).

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).

38. Cowell, I. G. and Hurst, H. C. Transcriptional repression by the human bZIP factor E4BP4: definition of a minimal repression domain. *Nucleic Acids Res.* **22**, 59-65 (1994).

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).

40. Ishizuka, I. E. et al., The Innate Lymphoid Cell Precursor. *Annu.Rev Immunol* **34**, 299-316 (2016).

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).

42. Sathe, P. et al., Convergent differentiation: myeloid and lymphoid pathways to murine plasmacytoid dendritic cells. *Blood* **121**, 11-19 (2013).

43. Rodrigues, P. F. et al., Distinct progenitor lineages contribute to the heterogeneity of plasmacytoid dendritic cells. *Nat Immunol* **19**, 711-722 (2018).

44. Herman, J. S., Sagar, and Grun, D. FateID infers cell fate bias in multipotent progenitors from single-cell RNA-seq data. *Nat Methods* **15**, 379-386 (2018).

45. Ghosh, H. S. et al., ETO family protein Mtg16 regulates the balance of dendritic cell subsets by repressing Id2. *Journal of Experimental Medicine* **211**, 1623-1635 (2014).

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).

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).

48. Rawlins, E. L. et al., The Id2+ distal tip lung epithelium contains individual multipotent embryonic progenitor cells. *Development* **136**, 3741-3745 (2009).

49. Butler, A. et al., Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol.* **36**, 411-420 (2018).

50. Estrada, J. et al., SiteOut: An Online Tool to Design Binding Site-Free DNA Sequences. *PLoS One* **11**, e0151740-(2016).

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).

52. Darrah, P. A. et al., Multifunctional TH1 cells define a correlate of vaccine-mediated protection against Leishmania major. *Nat.Med.* **13**, 843-850 (2007).

53. Bailey, T. L. et al., MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* **37**, W202-W208 (2009).

54. Heng, T. S., Painter, M. W., and Immunological Genome Project Consortium The Immunological Genome Project: networks of gene expression in immune cells. *Nat Immunol* **9**, 1091-1094 (2008).

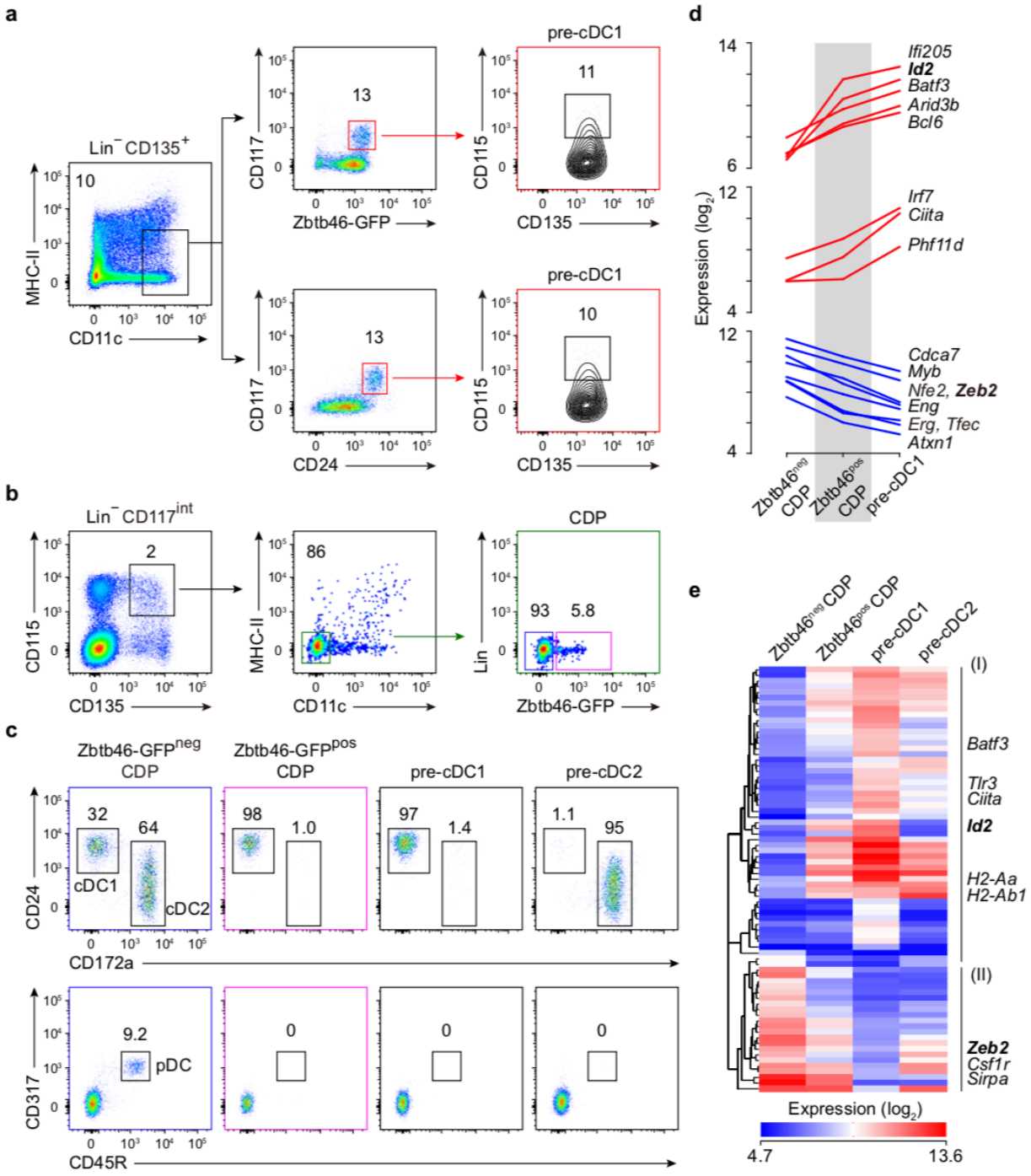

Figure 2.1 *Zbtb46*-GFP Expression in CDPs Identifies the Earliest Committed cDC1 Progenitor.

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).

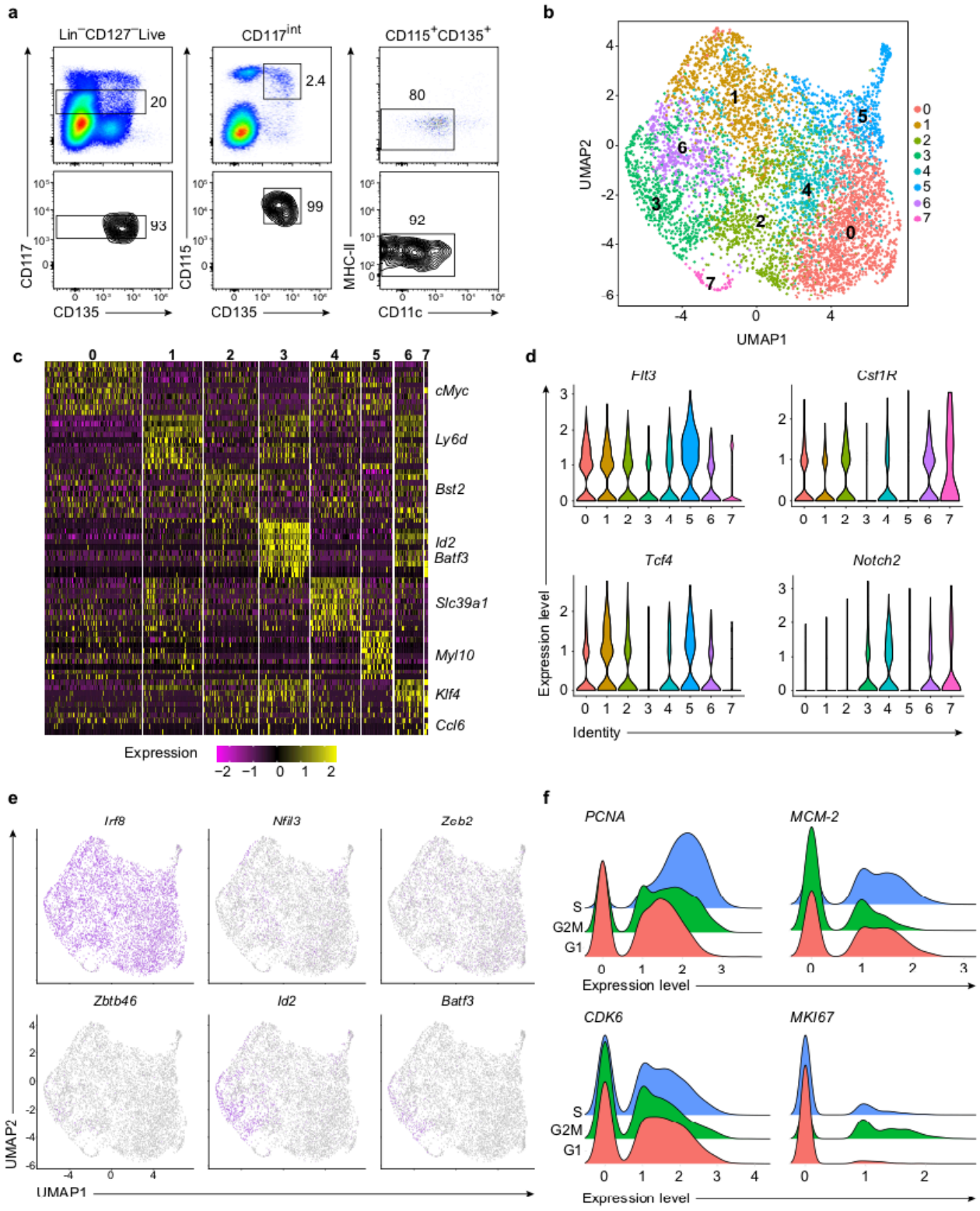

Figure 2.2 Single-cell RNA Transcriptome Analysis of CDPs.
a, CDPs gated as Live,[CD105, CD3, CD19, Ly6G, Ter119]$^{-}$CD127^{-} CD117^{int}CD115^{+}CD135^{+}MHC-II^{-}CD11c^{-} cells were purified by sorting from C57BL/6J mice.

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.

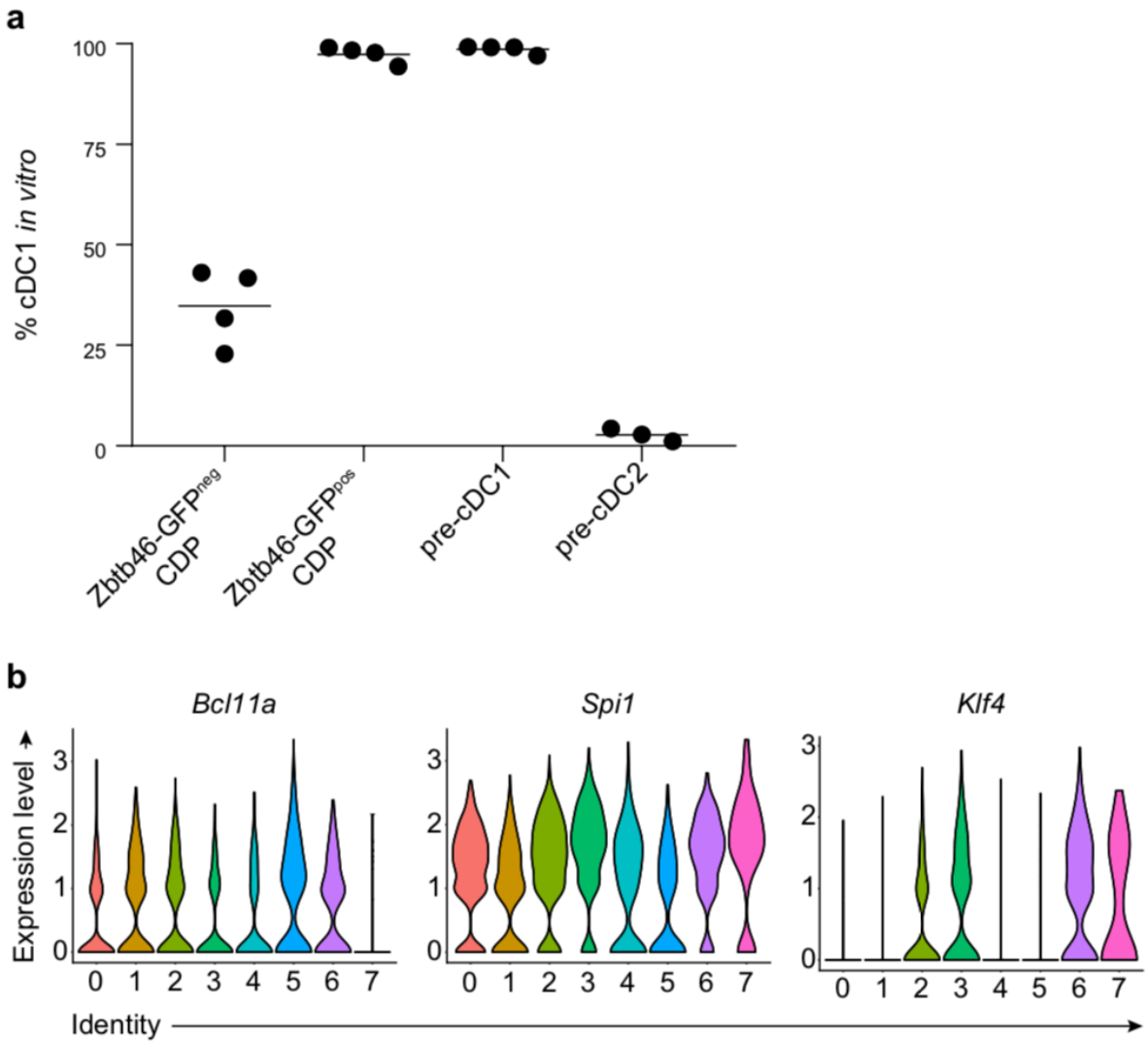

Figure 2.3 cDC1 specification occurs in the CDP
a, *Zbtb46*-GFP^{pos} CDPs, *Zbtb46*-GFP^{neg} CDPs, pre-cDC1 and pre-cDC2 were sort purified from *Zbtb46*$^{gfp/+}$ mice, cultured for 5 d in Flt3L, and analyzed by flow cytometry for development of pDCs and cDC1 (data presented for three independent experiments, n = 4 for *Zbtb46*-GFP^{pos}, *Zbtb46*-GFP^{neg} CDPs, pre-cDC1 and n = 3 for pre-cDC2). Small horizontal lines indicate the mean. **b,** Violin plots depicting cluster identity and expression level for the indicated genes.

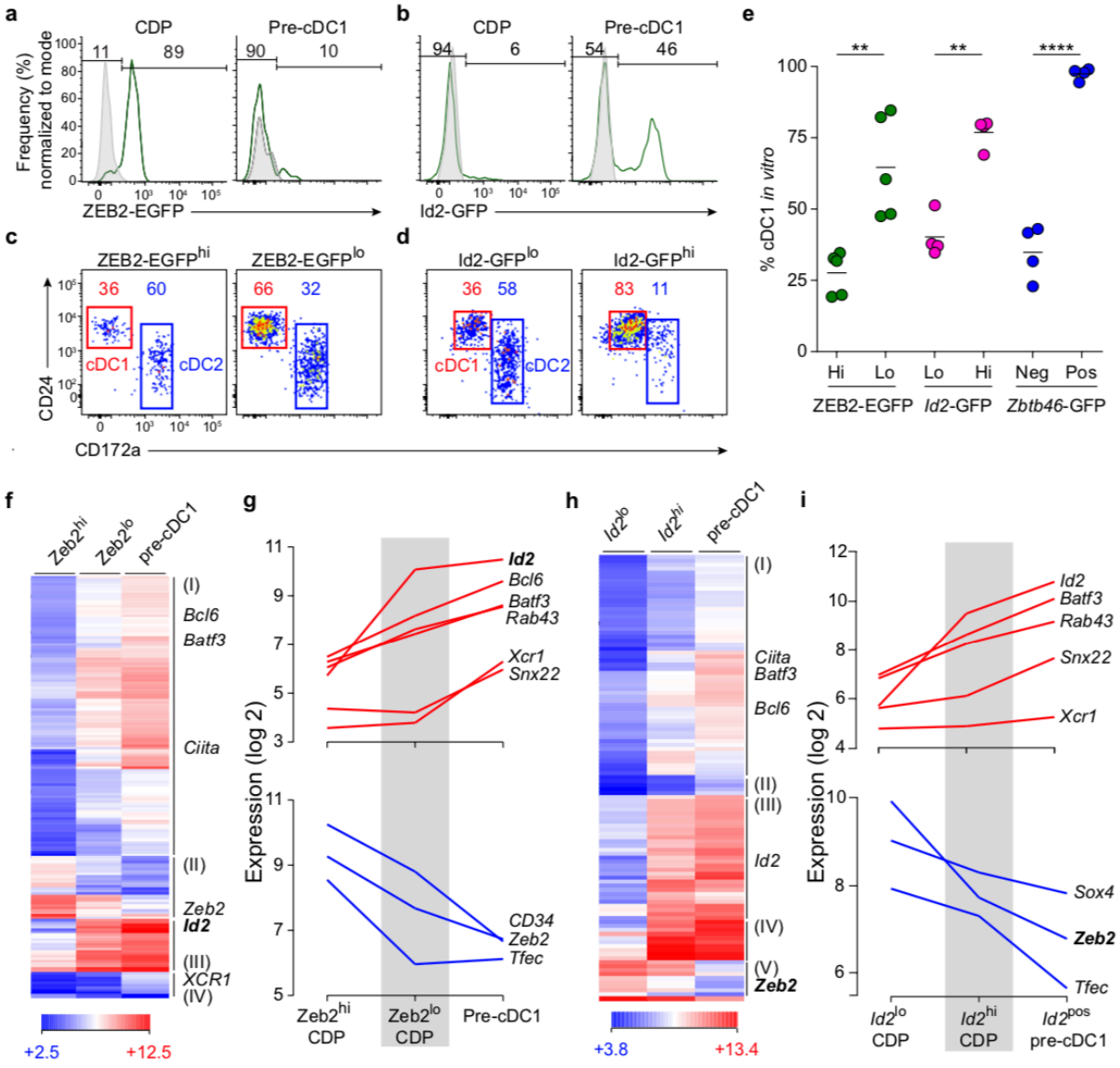

Figure 2.4 *Zeb2* and *Id2* Heterogeneity Identifies cDC1 Specification in CDPs.
BM from *Zeb2*$^{egfp/egfp}$ (**a**) and *Id2*^{gfp} (**b**) mice were analyzed by flow cytometry to identify GFP expression in CDPs and pre-cDC1s. WT mice (*Zeb2*$^{+/+}$ and *Id2*$^{+/+}$) are shown as gray histograms. 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 Figure 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. *$p < 0.05$, **$p < 0.01$, ****$p < 0.0001$.

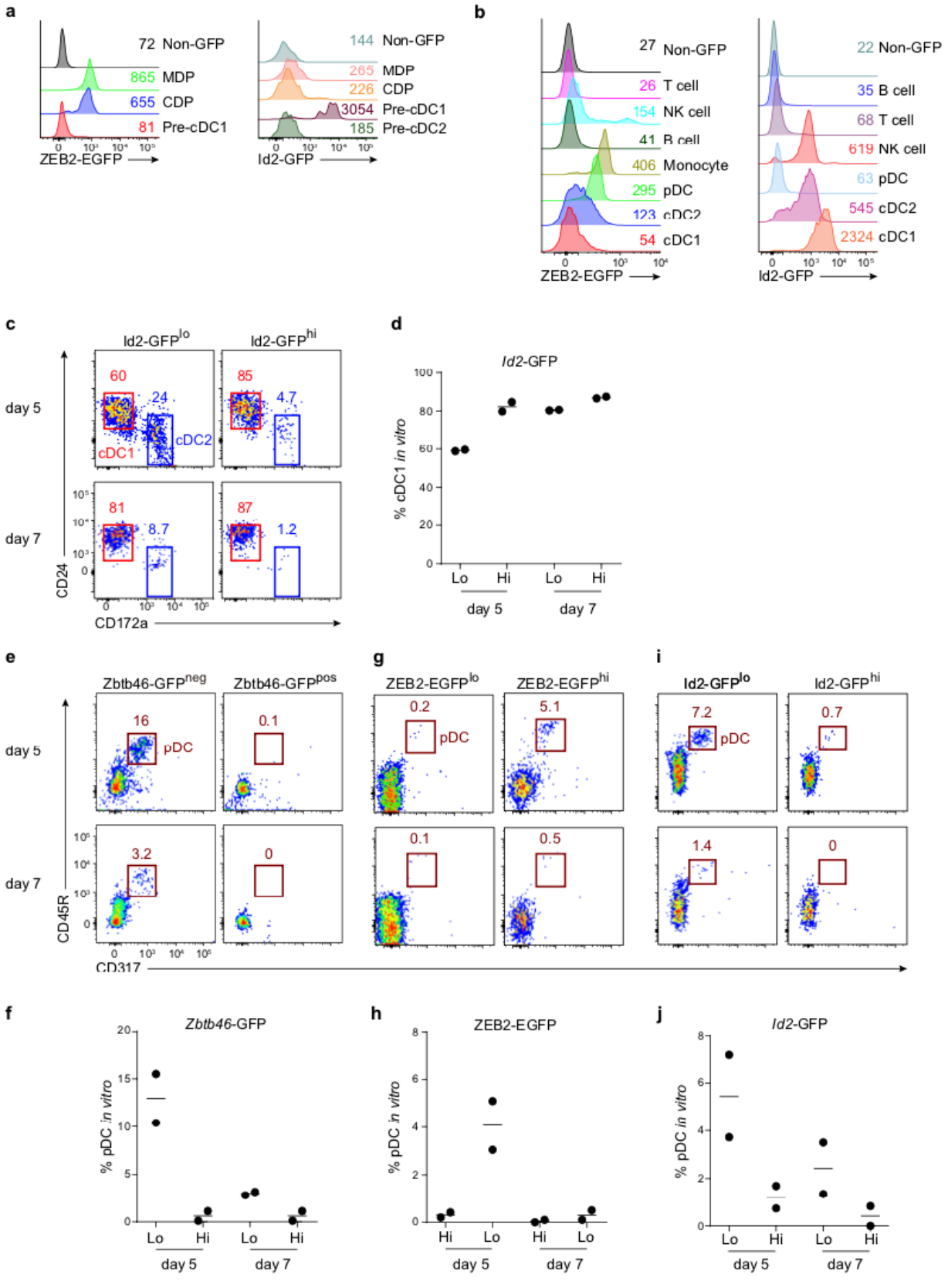
a
72 Non-GFP
865 MDP
655 CDP
81 Pre-cDC1
ZEB2-EGFP
144 Non-GFP
265 MDP
226 CDP
3054 Pre-cDC1
185 Pre-cDC2
Id2-GFP
b
27 Non-GFP
26 T cell
154 NK cell
41 B cell
406 Monocyte
295 pDC
123 cDC2
54 cDC1
ZEB2-EGFP
22 Non-GFP
35 B cell
68 T cell
619 NK cell
63 pDC
545 cDC2
2324 cDC1
Id2-GFP
c
Id2-GFP^lo
Id2-GFP^hi
day 5
day 7
60
24
cDC1
cDC2
85
4.7
81
8.7
87
1.2
CD24
CD172a
d
Id2-GFP
% cDC1 in vitro
Lo
Hi
Lo
Hi
day 5
day 7
e
Zbtb46-GFP^neg
Zbtb46-GFP^pos
day 5
day 7
16
pDC
0.1
3.2
0
CD45R
CD317
g
ZEB2-EGFP^lo
ZEB2-EGFP^hi
0.2
pDC
5.1
0.1
0.5
i
Id2-GFP^lo
Id2-GFP^hi
7.2
pDC
0.7
1.4
0
f
Zbtb46-GFP
% pDC in vitro
Lo
Hi
Lo
Hi
day 5
day 7
h
ZEB2-EGFP
% pDC in vitro
Hi
Lo
Hi
Lo
day 5
day 7
j
Id2-GFP
% pDC in vitro
Lo
Hi
Lo
Hi
day 5
day 7

Figure 2.5 ZEB2-EGFP and *Id2*-GFP Expression in BM and Spleen
a, BM from *Zeb2*$^{egfp/egfp}$ (**left**) and *Id2*$^{gfp/+}$ (**right**) mice were analyzed by FACS to identify GFP expression in the indicated progenitors. Numbers are the MFI (data representative of three independent experiments, n = 3 mice). **b,** Spleen from *Zeb2*$^{gfp/gfp}$ (**left**) and *Id2*$^{gfp/+}$ (**right**) mice was analyzed by FACS for GFP expression in T cells ($CD3^+CD45R^-$), B cells ($CD45R^+CD3^-$), NK cells ($NK1.1^+CD3^-$), monocytes ($Ly6C^+CD115^+$), and DCs as gated in the Methods. Numbers are the MFI (data representative of three independent experiments, n =3 mice). **c,** *Id2*-GFP^{lo} or *Id2*-GFP^{hi} CDPs were sort purified from *Id2*$^{gfp/+}$ mice, cultured for 5 or 7 d in Flt3L, and analyzed by FACS for development of cDC1 (**red**) or cDC2 (**blue**). Number indicates percentage of cells in the gate (data representative of two independent experiments, n = 2 mice). **d,** Percentages of cDC1s from *in vitro* cultures as described in (**c**). Small horizontal lines indicate the mean. **e,** *Zbtb46*-GFP^{pos} or *Zbtb46*-GFP^{neg} CDPs were sort purified from *Zbtb46*$^{gfp/+}$ mice, cultured for 5 or 7 d in Flt3L, and analyzed by FACS for development of pDCs (**brown**). Number indicate percentage of cells in the gate (data representative of two independent experiments, n = 2 mice). **f,** Percentages of pDCs from *in vitro* cultures as described in (**e**). Small horizontal lines indicate the mean. **g,** ZEB2-EGFP^{lo} or ZEB2-EGFP^{hi} CDPs were sort purified from *Zeb2*$^{egfp/egfp}$ mice, cultured for 5 or 7 d in Flt3L, and analyzed by FACS for development of pDCs (**brown**). Number indicates percentage of cells in gate (data representative of two independent experiments, n = 2 mice). **h,** Percentages of pDCs from *in vitro* cultures as described in (**g**). Small horizontal lines indicate the mean. **i,** *Id2*-GFP^{lo} or *Id2*-GFP^{hi} CDPs were sort purified from *Id2*$^{gfp/+}$ mice, cultured for 5 or 7 d in Flt3L, and analyzed by FACS for development of pDCs (**brown**). Number indicates percentage of cells in the gate (data representative of two independent experiments, n = 2 mice). **j,** Percentages of pDCs from *in vitro* cultures as described in (**i**). Small horizontal lines indicate the mean.

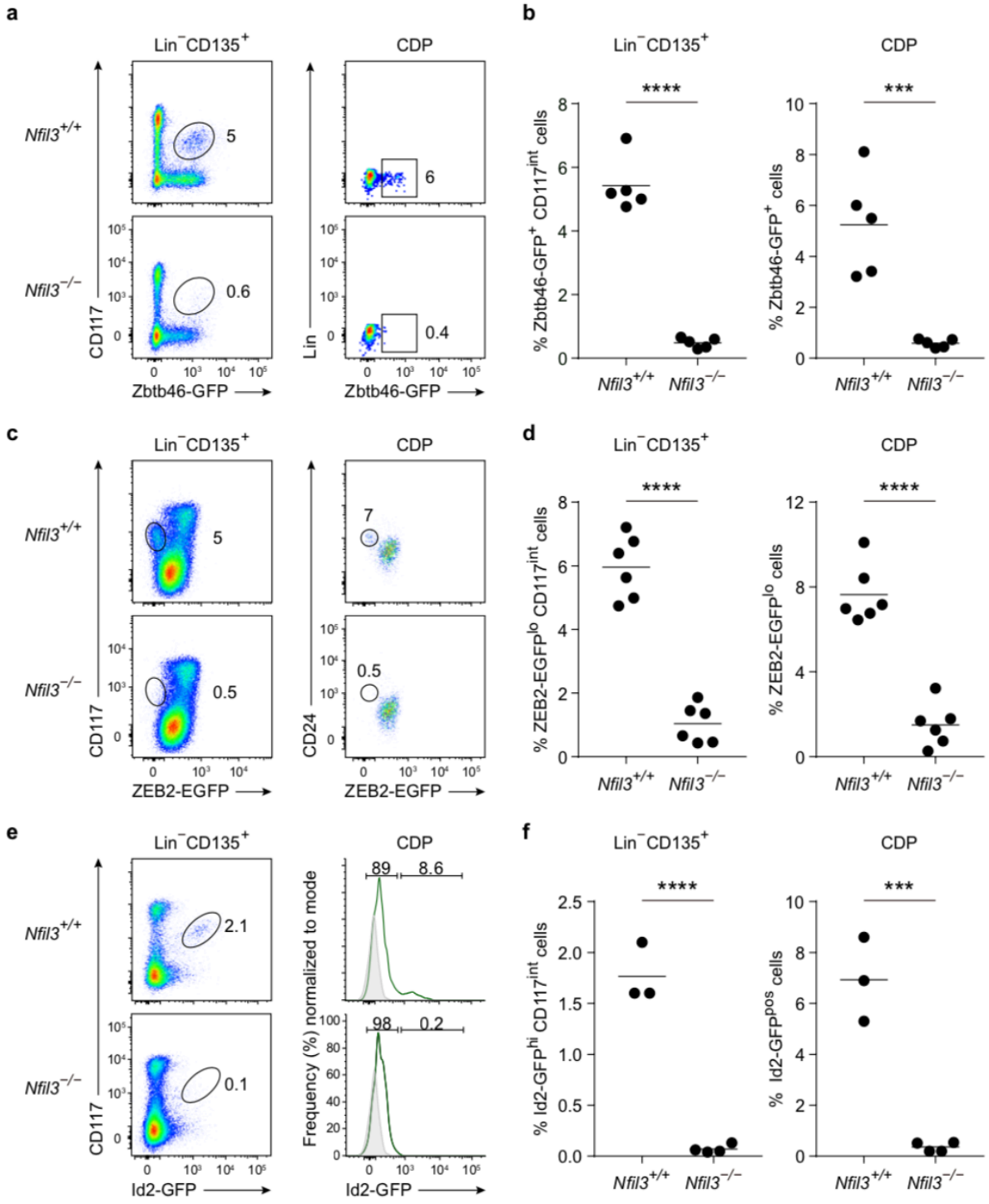

Figure 2.6 *Nfil3* is Required for cDC1 Specification.
a, BM from *Nfil3*$^{+/+}$*Zbtb46*$^{gfp/+}$ and *Nfil3*$^{-/-}$*Zbtb46*$^{gfp/+}$ mice was analyzed by flow cytometry for Lin^{-}CD135^{+}CD117^{int} *Zbtb46*-GFP^{pos} cells (**left**) or *Zbtb46*-GFP^{pos} CDPs (**right**). Numbers indicate the 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.

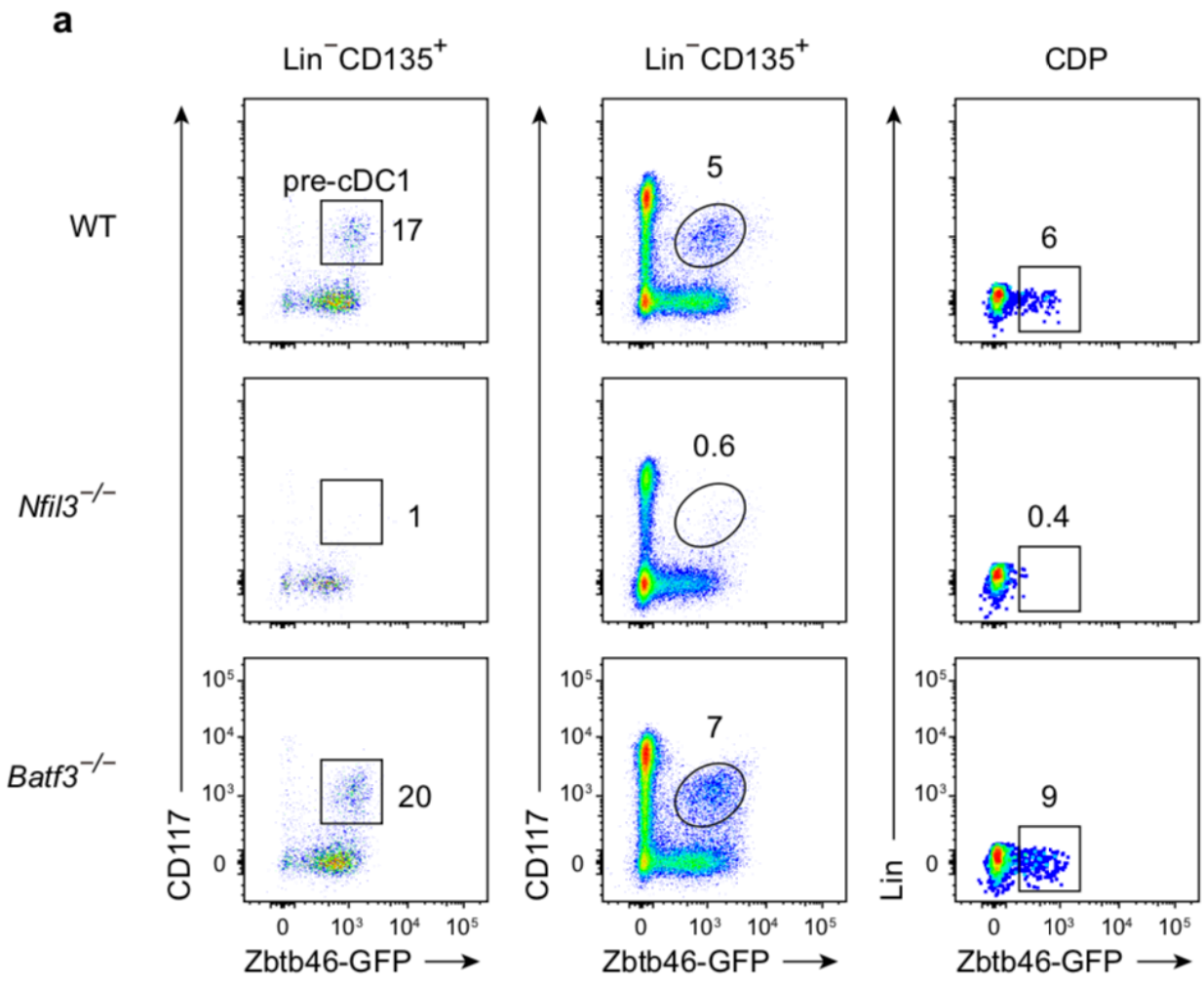

Figure 2.7 *Nfil3* is Required for cDC1 Specification
a, BM from *Zbtb46*$^{gfp/+}$ (WT), *Nfil3*$^{-/-}$*Zbtb46*$^{gfp/+}$ and *Batf3*$^{-/-}$*Zbtb46*$^{gfp/+}$ mice was analyzed by FACS for development of pre-cDC1 (**left**), Lin^{-}CD135^{+}CD117^{int}*Zbtb46*-GFP^{pos} (**middle**) and Zbtb46-GFP^{pos} CDPs (**right**) (data representative of three independent experiments, n = 3 mice).

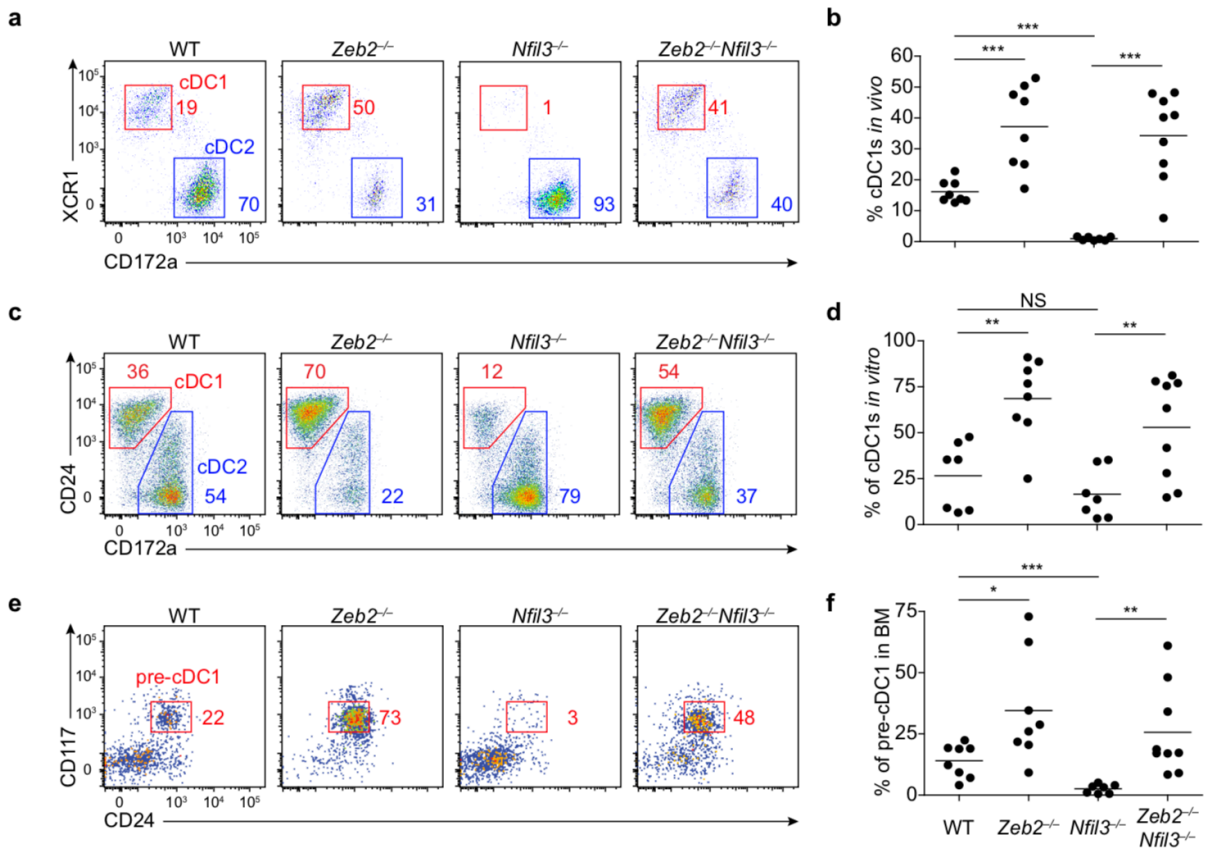

Figure 2.8 *Zeb2* is Downstream of *Nfil3* in cDC1 Development.
a, Splenic cDCs from *Nfil3*$^{+/+}$*Zeb2*$^{f/f}$ *Mx1-cre*$^{-}$ (WT), *Zeb2*$^{f/f}$ *Mx1-cre*$^{+}$ (*Zeb2*$^{-/-}$), *Nfil3*$^{-/-}$ (*Nfil3*$^{-/-}$), and *Nfil3*$^{-/-}$ *Zeb2*$^{f/f}$ *Mx1-cre*$^{+}$ (*Nfil3*$^{-/-}$ *Zeb2*$^{-/-}$) mice 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.

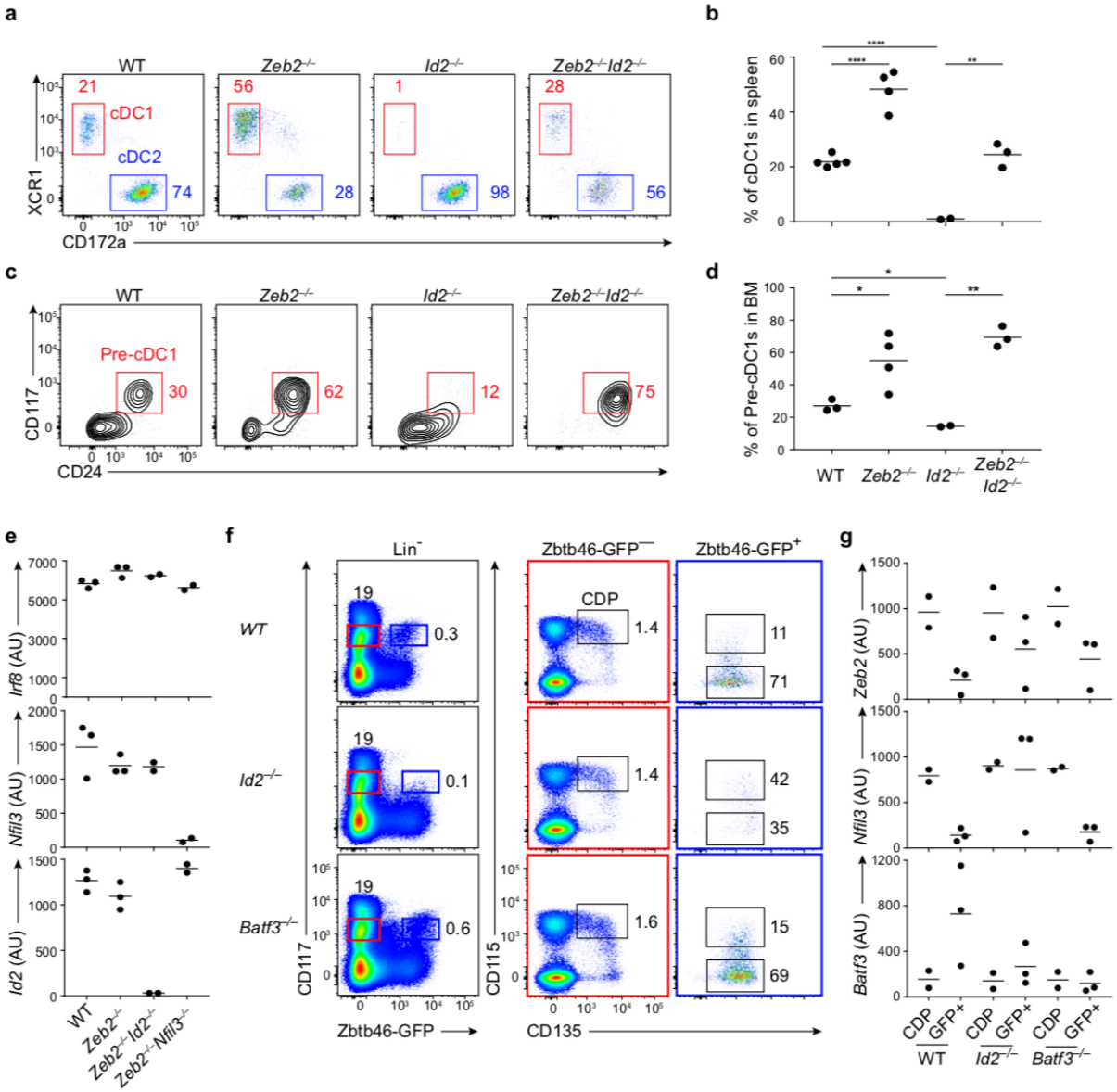

Figure 2.9 Expression of *Id2* and *Zeb2* is Mutually Repressive in the CDP.
a, Splenic cDCs harvested from WT, *Zeb2*$^{f/f}$ *Rosa26* $^{(cre\text{-}ERT2/+)}$ (*Zeb2*$^{-/-}$), *Id2*$^{f/f}$ *Rosa26* $^{(cre\text{-}ERT2/+)}$ (*Id2*$^{-/-}$), and *Id2*$^{f/f}$ *Zeb2*$^{f/f}$ *Rosa26* $^{(cre\text{-}ERT2/cre\text{-}ERT2)}$ (*Zeb2*$^{-/-}$ *Id2*$^{-/-}$) were analyzed for cDC1 (**red**) and cDC2 (**blue**) frequency. 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.

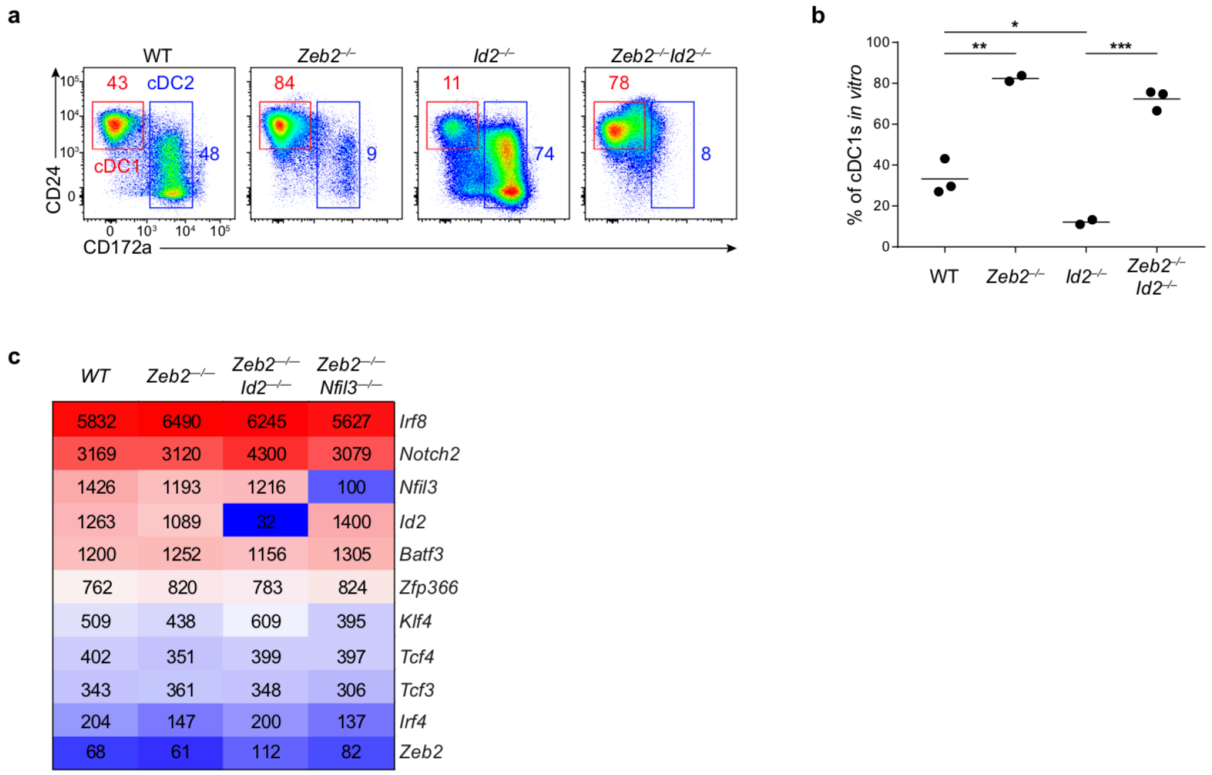

Figure 2.10 *Id2-Zeb2* Loop Regulates cDC1 Fate

a, cDCs derived *in vitro* from Flt3L-treated BM cultures from WT, *Zeb2*$^{-/-}$, *Id2*$^{-/-}$, and *Zeb2*$^{-/-}$ *Id2*$^{-/-}$ mice were analyzed for cDC1 (**red**) and cDC2 (**blue**) frequency. 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,** cDC1 frequency is shown for individual mice in (**a**) as a percentage of total cDCs. Small horizontal lines indicate the mean. **c,** Splenic cDC1 from WT, *Zeb2*$^{-/-}$, *Id2*$^{-/-}$, and *Zeb2*$^{-/-}$ *Id2*$^{-/-}$ mice were purified by sorting and analyzed by gene expression microarrays. Shown are expression for the indicated genes for each genotype. Numbers are the average gene expression of three biological replicates.

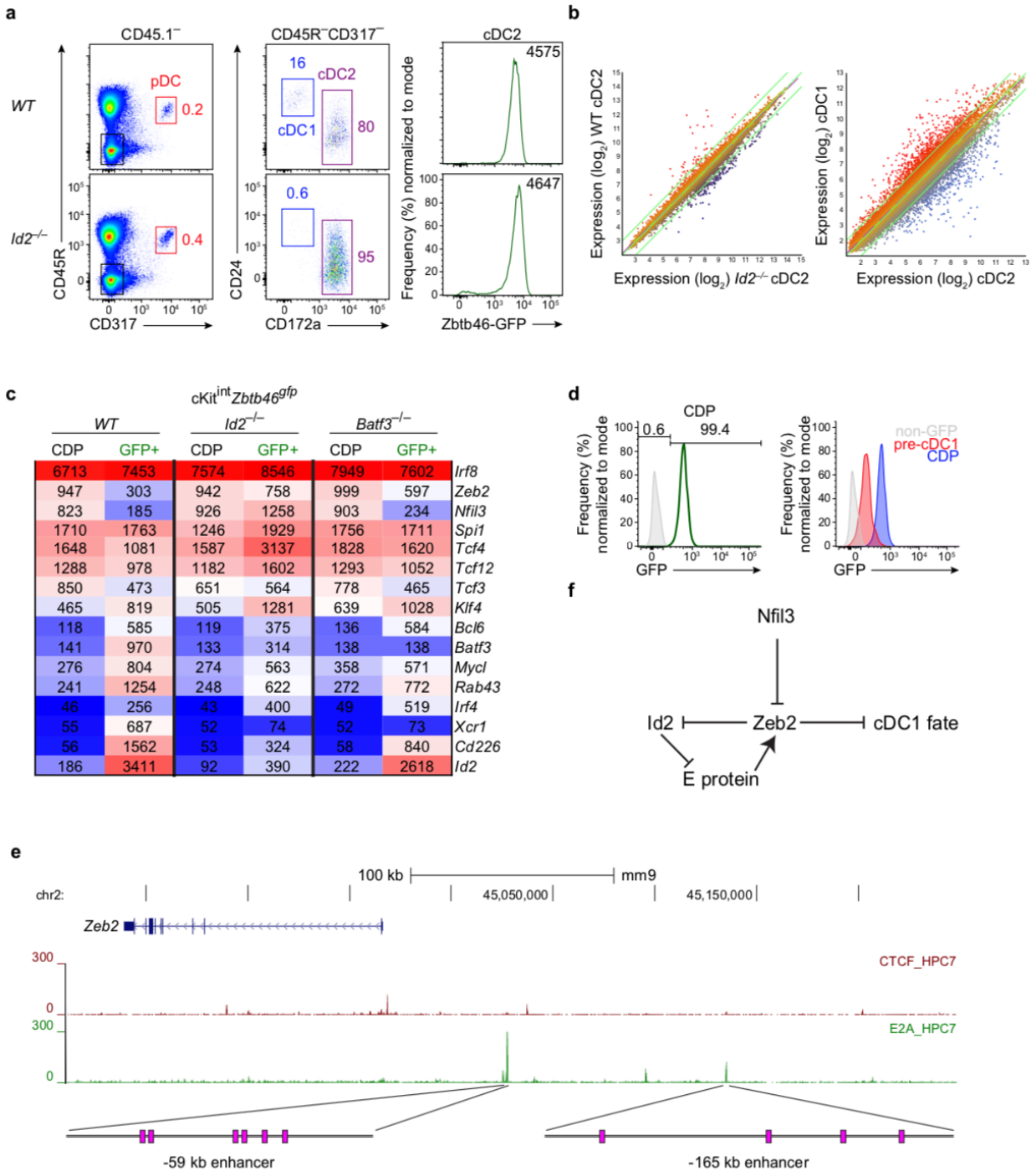

Figure 2.11 *Id2* Does Not Regulate cDC2 Development or Transcriptome, But May Indirectly Repress *Zeb2* through E proteins

a, Spleens from radiation chimeras receiving either *Zbtb46*$^{gfp+}$ (WT) or *Id2*$^{-/-}$*Zbtb46*$^{gfp/gfp}$ (*Id2*$^{-/-}$) BM were analyzed by FACS for the frequency of pDC (**left**), and cDC1 and cDC2 (**middle**) populations, and for *Zbtb46*-GFP expression on cells gated on cDC2 populations (**right**) (data representative of three independent experiments, n = 3 mice) **b,** Microarray analysis of *Id2*$^{+/+}$ and *Id2*$^{-/-}$ cDC2. Gene expression comparing *Id2*$^{+/+}$ cDC2 to *Id2*$^{-/-}$cDC2 described in (**a**) are shown (**left**), or WT cDC2 compared to WT cDC1 (**right**) (data pooled from two independent

experiments, n = 2 mice). 255 genes were at least 2-fold differentially expressed between *Id2*$^{+/+}$ and *Id2*$^{-/-}$ cDC2. Green lines indicate two-fold change threshold **c,** Gene expression microarray was performed on two CDP populations, sort purified separately as either Lin- CD117^{int}*Zbtb46*-GFP^{neg} (GFP-), or Lin- CD117^{int}*Zbtb46*-GFP^{pos} (GFP+) cells, from each of three genotypes, *Zbtb46*$^{gfp/+}$ (WT), *Id2*$^{-/-}$*Zbtb46*$^{gfp/gfp}$ (*Id2*$^{-/-}$), *and Batf3*$^{-/-}$*Zbtb46*$^{gfp/gfp}$ (*Batf3*$^{-/-}$) mice. Shown are averages of duplicate or triplicate gene expression values of the indicated genes in the indicated populations. **d,** BM from E2A-GFP reporter mice (*Tcf3*$^{gfp/+}$) or WT (gray histograms) were analyzed for GFP expression in CDPs (**left**), or as an overlay (**right**) for pre-cDC1 (**red**) and CDP (**blue**) (data representative of four independent experiments, n = 5 mice). **e,** Shown are E2-A and CTCF peaks identified by ChIP-seq in HPC-7 cells at the *Zeb2* locus (mm9)[35]. Squares represent E-box motifs (CANNTG) within the indicated enhancer regions. **f,** Proposed epistatic model for *Nfil3, Id2, Zeb2*, E protein, and cDC1 fate.

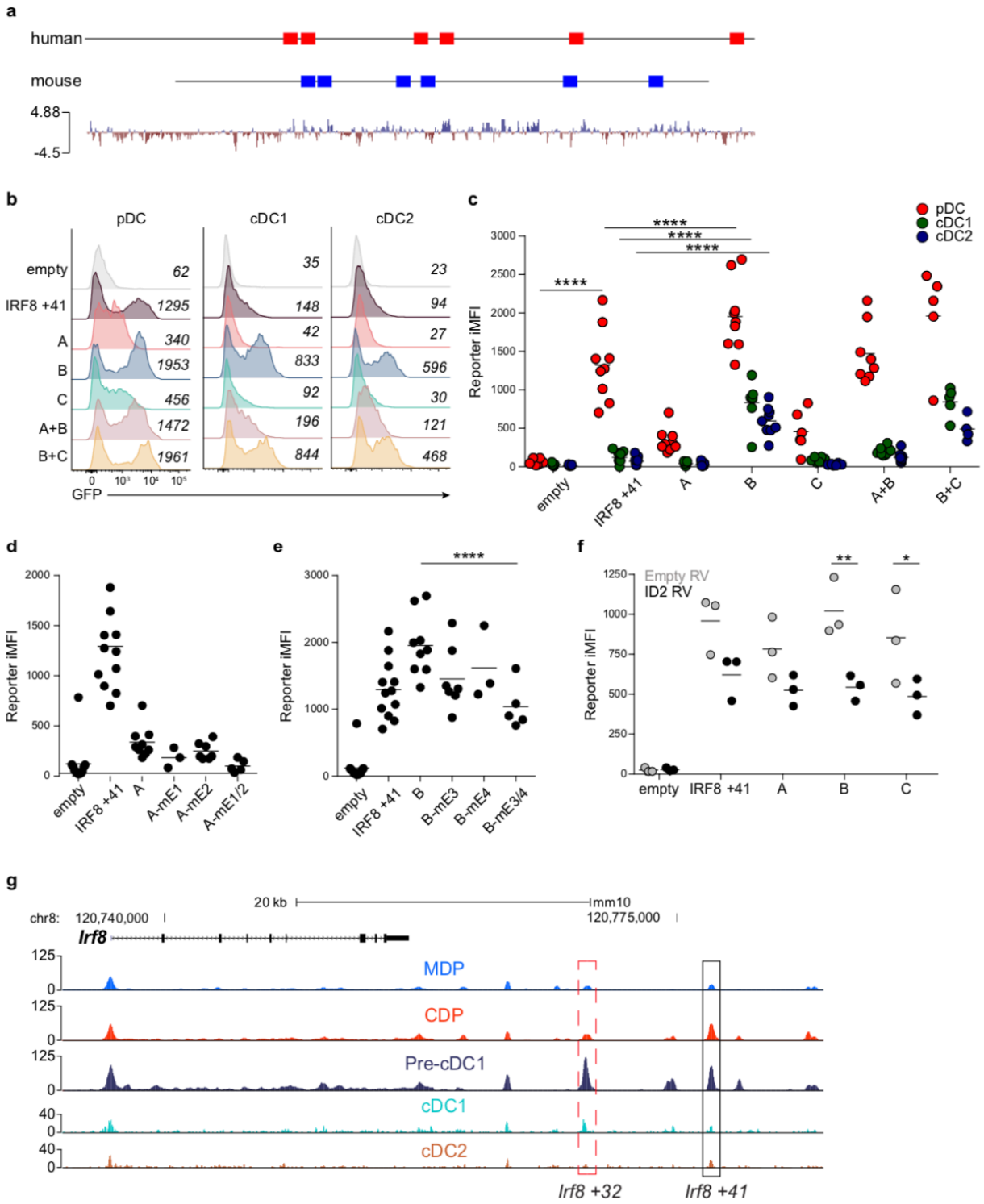

Figure 2.12 *Id2* imposes a switch from the +41 kb *Irf8* enhancer to the +32 kb *Irf8* enhancer by Reducing E protein activity,

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. *$p < 0.05$, **$p < 0.01$, ****$p < 0.0001$.

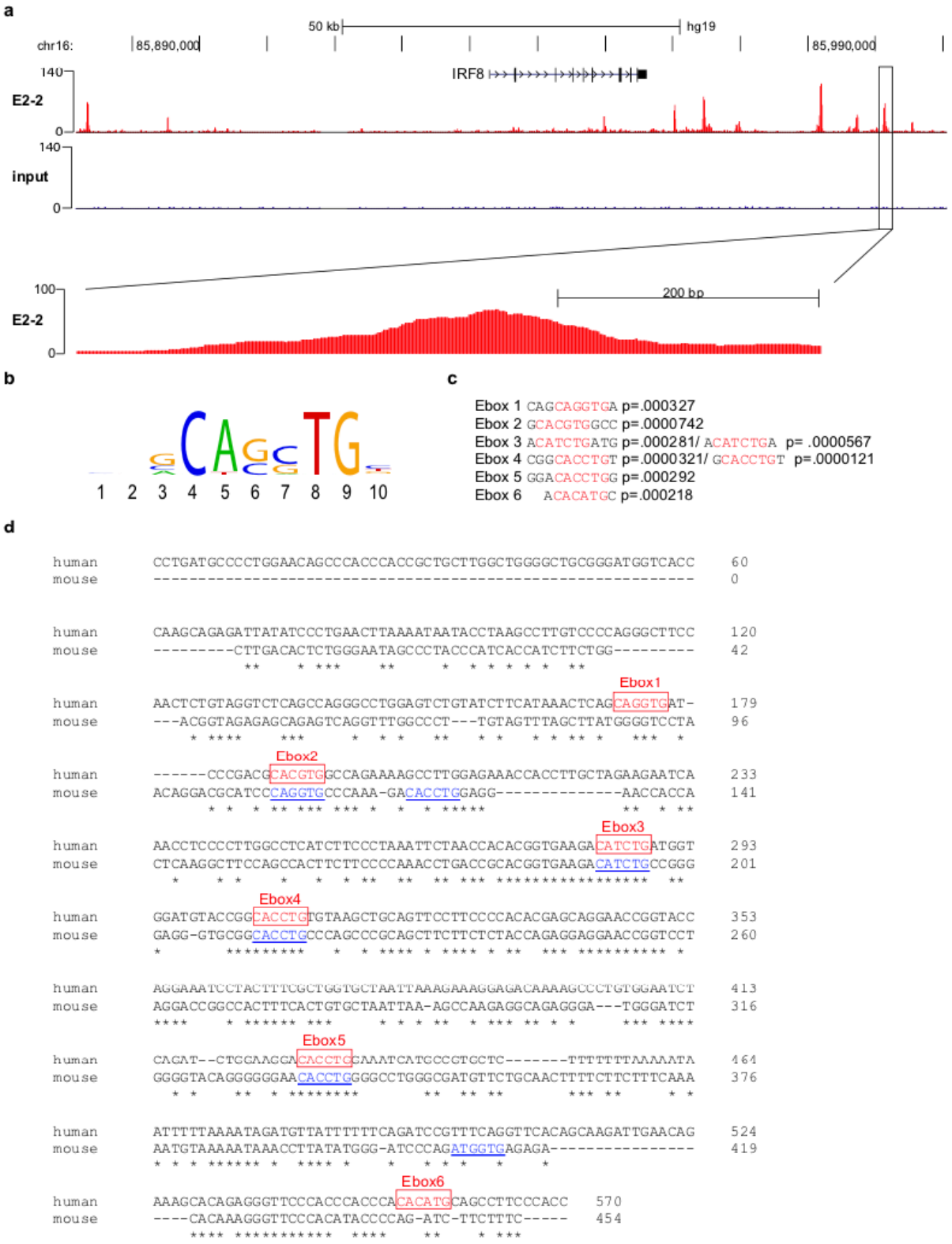

Figure 2.13 Conservation of +41 kb Enhancer Between Human and Mice
a, ChIP-seq peaks for E2-2 (E2-2) or control (input) for human *IRF8* locus[37]. Numbers are chromosomal coordinates, human chromosome 16, draft genome hg19. Box indicates E2-2

binding peak corresponding to mouse +41 kb *Irf8* enhancer. **b,** Consensus human E-box motif from E2-2 track. **c,** FIMO analysis depicting p-values of predicted E-boxes in human *IRF8* chr16:85991064-85991633 (+58 kb from *IRF8* TSS) **d,** Alignment of human, genome draft hg19, and mouse, genome draft mm10, for the +41 kb *IRF8* enhancer regions. Conserved human (**red box**) and mouse (**blue underlined**) E-box motifs are indicated.

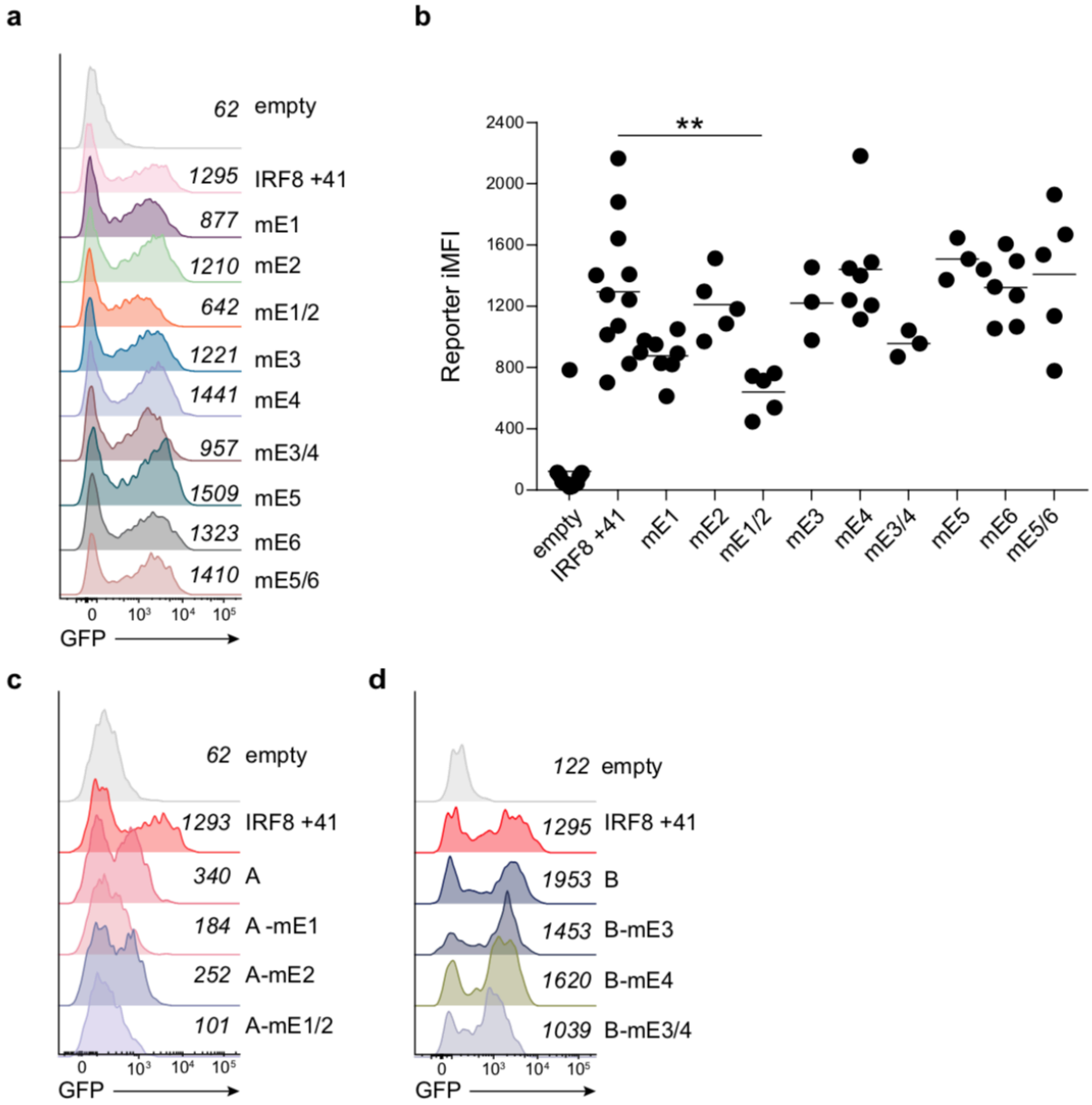

Figure 2.14 Activity of +41 kb *Irf8* Enhancer is E-box Dependent

a, GFP expression in pDCs, from RV reporters without (empty) or with the entire 454 bp +41 kb enhancer region (IRF8 +41), or with single or double mutations of the indicated E-boxes (**left**) (data representative of at least three independent experiments, n > 3). **b,** Data shown in (**a**) shown as integrated MFI (data pooled from at least three independent experiments, n > 3). Small horizontal lines indicate the mean. **c,** 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/m1), shown as histograms (data representative of at least three independent experiments, n > 3). **d,** 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 histograms (data representative of at least three independent experiments, n > 3). Data are

presented as mean and one-way ANOVA was used to compare groups. *$p < 0.05$, **$p < 0.01$, ****$p < 0.0001$.

CHAPTER 3:

Discussion and Future Directions

Portions of this chapter will be published in *Nature Immunology* in 2019.

Bagadia, P., Huang, X., Liu, T., Durai, V., Grajales-Reyes, GE, Nitschke, M., Modrusan, Z., Granja, JM, Sathpathy, AT, Briseño CG, Gargaro M, Iwata, A, Kim, S, Chang, HY, Shaw, AS, Murphy TL, Murphy KM. An Nfil3–Zeb2–Id2 pathway imposes Irf8 enhancer switching during cDC1 development. *Nat. Immunol*.

3.1 Abstract

In this dissertation, we identified the similarities between ILCs and DCs, two innate immune cell lineages that help promote a specialized immune response. Both cell lineages rely on similar transcription factors and transcription factor families for their developments. However, an incomplete understanding of their shared transcriptional requirements remained, and we attempted to resolve for cDC1 specification. We found that cDC1 specification relies on a switch in *Irf8* enhancer usage and that the switch occurs because of a complex mechanism relying on transcriptional repressors. We used single-cell RNA-sequencing of the CDP, the multipotent DC progenitor, and identified a cluster of cells that expressed transcription factors that influence cDC1 development, such as *Nfil3, Id2,* and *Zeb2*. We performed genetic epistasis to determine the functional hierarchy of transcription factors involved in cDC1 specification and organized a transcriptional circuit that explains the switch in *Irf8* expression. 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. This new understanding of cDC1 specification could be applied to cDC2 and pDC specification, as well as innate immune cell specification, lineages that rely on similar transcription factors.

3.2 Transcriptional networks in cDC1 development

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[1,2], but *Id2*$^{-/-}$ mice lacked only cDC1, and did not show the expected loss of all cDCs[3]. 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[4]. 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 [5,6],but may activate transcription in contexts[7]. Likewise, *Zeb2* has been suggested to directly repress *Id2* expression[8,9], although this has not been rigorously tested. *Nfil3*, *Zeb2*, and *Id2* have also been shown to regulate ILC development[10], but the mechanisms by which these transcription factors act in these cells has not been studied.

3.3 Future Directions

DC specification and commitment are complicated processes that we have attempted to clarify in this dissertation. However, we have possibly uncovered a need to revise the DC developmental 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[11]. cDC1 specification may occur even earlier than this dissertation 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. This study could extend to other transcription factors, because we do not yet understand how expression levels of transcription factors relate to transcription factor activity in all contexts. This is particularly important in the case of transcriptional repressors, such as the ones studied in this dissertation.

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[12]. However, two recent studies indicated that late pDC progenitors emerge from the common lymphoid progenitor and a "pre-pDC" was described[13,14]. 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[15]. 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. Understanding the precise relationships between different progenitor populations will clarify the transcriptional mechanisms required for certain cell subsets to develop, but more genetic tools and lineage models will be required.

3.4 References

1. 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).
2. Watowich, S. S. and Liu, Y. J. Mechanisms regulating dendritic cell specification and development. *Immunol.Rev.* **238**, 76-92 (2010).
3. 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).
4. Grajales-Reyes, G. E. et al., Batf3 maintains autoactivation of Irf8 for commitment of a CD8alpha(+) conventional DC clonogenic progenitor. *Nat Immunol* **16**, 708-717 (2015).
5. Cowell, I. G. and Hurst, H. C. Transcriptional repression by the human bZIP factor E4BP4: definition of a minimal repression domain. *Nucleic Acids Res.* **22**, 59-65 (1994).

6. Cowell, I. G., Skinner, A., and Hurst, H. C. Transcriptional repression by a novel member of the bZIP family of transcription factors. *Mol Cell Biol* **12**, 3070-3077 (1992).

7. 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).

8. 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).

9. 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).

10. Ishizuka, I. E. et al., The Innate Lymphoid Cell Precursor. *Annu.Rev Immunol* **34**, 299-316 (2016).

11. 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).

12. Sathe, P. et al., Convergent differentiation: myeloid and lymphoid pathways to murine plasmacytoid dendritic cells. *Blood* **121**, 11-19 (2013).

13. Rodrigues, P. F. et al., Distinct progenitor lineages contribute to the heterogeneity of plasmacytoid dendritic cells. *Nat Immunol* **19**, 711-722 (2018).

14. Herman, J. S., Sagar, and Grun, D. FateID infers cell fate bias in multipotent progenitors from single-cell RNA-seq data. *Nat Methods* **15**, 379-386 (2018).

15. Ghosh, H. S. et al., ETO family protein Mtg16 regulates the balance of dendritic cell subsets by repressing Id2. *Journal of Experimental Medicine* **211**, 1623-1635 (2014).