A shared Runx1-bound Zbtb16 enhancer directs innate and innate-like lymphoid lineage development

Ai-Ping Mao1,2, Isabel E. Ishizuka1,2, Darshan N. Kasal1,2, Malay Mandal1,3 & Albert Bendelac1,2

Zbtb16-encoded PLZF is a signature transcription factor (TF) that directs the acquisition of T-helper effector programs during the development of multiple innate lymphocyte lineages, including natural killer T cell, innate lymphoid cell, mucosal-associated invariant T cell and γδ lineages. PLZF is also essential in osteoblast and spermatogonial development. How Zbtb16 itself is regulated in different lineages is incompletely understood. Here, by systematic CRISPR/Cas9-assisted deletions of chromatin accessible regions within the Zbtb16 locus in mouse, we identify a critical enhancer controlling PLZF expression exclusively in innate lymphoid lineages. Multiple sites within this enhancer express canonical motifs for the TF Runx1, which is essential for the development of these lineages. Notably, some regulatory sites control the kinetic rather than the overall level of PLZF expression. Thus, our comprehensive, unbiased analysis of regulatory elements in vivo reveals critical mechanisms of Zbtb16 regulation shared between innate and innate-like lymphoid lineages.
A hallmark of innate and innate-like lymphocytes is their acquisition of T-helper effector programs during development, independent of exogenous stimuli or exposure to infection\textsuperscript{1–4}. In the bone marrow, a common innate lymphoid cell precursor (ILCP) gives rise to ILC1, ILC2 and ILC3 lineages expressing T-helper 1, Th2 and Th17 programs directed by T-bet, GATA3 and ROR\textgamma, respectively. Likewise, in the thymus, a common CD1d-restricted V\textalpha14-J\alpha18 expressing thymic natural...
killer T cell (NKT) precursor gives rise to NKT1, NKT2 and NKT17 subsets. The acquisition of these innate helper effecter programs depends on the specific induction of the signature transcription factor (TF) PLZF encoded by Zbtb16, which directly binds and regulates key effector genes involved in lymphocyte migration, adhesion, costimulation, cytokine signaling as well as genes encoding master TFs controlling the T-helper programs.

PLZF is not expressed in conventional lineages, whether resting or activated, indicating tight lineage-specific control of expression. In the ILC lineages, PLZF is transiently expressed at the ILCP stage, capturing a partially defined sequence of TFs that are critically involved in early ILC development from a common lymphoid precursor (CLP) precursor, including Nfil3, Tof, Tcf7 and Id2. How these TFs are induced in the first place and their potential interconnections remain poorly understood. Furthermore, which of these or other TFs contribute to Zbtb16 induction has not been defined.

In the NKT lineage, PLZF induction is dependent on Vα14-Jα18 TCR signaling elicited by recognition of CD1d-lipid ligands on thymocytes in the context of homotypic Slamf receptor engagement. These signals result in elevated and sustained level of NFAT-induced Egr2, which directly activates the Zbtb16 promoter. Other innate-like lymphoid lineages such as MR1-restricted Vα19-Jα33 mucosal-associated invariant T cell (MAIT) cells and Vγ1/Vδ8 γδT cells also depend on the specific expression of PLZF during their development, likely through similar TCR-induced signals which remain to be studied.

Thus, PLZF appears to be broadly involved in the development of innate and innate-like cells and their divergence from conventional lymphocytes. Furthermore, PLZF is also critically involved in some non-haematopoietic developmental processes, including limb patterning and spermatogonial self-renewal, raising intriguing issues regarding the mechanisms underlying this regulation of PLZF in multiple lineages. Cis-regulatory elements located in the non-coding regions of gene loci (termed enhancers) represent a major component of lineage-specific gene regulation, but the potential function of enhancers in controlling PLZF expression has not been investigated.

To characterize the regulatory elements controlling PLZF expression during ILC and NKT precursor development, here we employ a systematic approach based on ATAC-seq to identify chromatin accessible regions in the Zbtb16 locus of these lineages, and CRISPR/Cas9 deletions to evaluate their individual impact on gene expression. By computational motif analysis, chromatin immunoprecipitation and conditional gene deletion, we identify the TFs bound to these regulatory elements. Our study reveals a shared lymphoid-specific enhancer essential for the induction of PLZF in both innate and innate-like lymphoid lineages and suggests a critical function of Runx1 in activating this enhancer.

**Results**

**Chromatin accessibility at the Zbtb16 locus.** To define regions of chromatin accessibility, we performed ATAC-seq analysis of freshly isolated NKT thymocytes and bone marrow ILCP. Numerous ATAC-seq peaks were called by magnetic-activated cell sorting (MACS) software analysis along a ~200 kb segment encompassing the Zbtb16 locus (Fig. 1a). The peaks were particularly abundant in intron 2 and often grouped in clusters. They were mostly present in thymic NKT or bone marrow ILCP, both of which express PLZF, and were less frequent in control CD4 single positive thymocytes (CD4SP) or CLP, which do not express PLZF.

**CRISPR/Cas9-mediated deletions of ATAC-seq peaks.** For a first round of CRISPR/Cas9 deletions, the specific ATAC-seq peaks were grouped into eight transposase hypersensitivity regions that together covered most of the individual ATAC peaks, as indicated in Fig. 1a. These regions ranged in size between 1 and 15 kb, and were designated by their position with respect to the transcription start site, including +185/191, +120/124, +115/116, +75/81, +40/43, +18/32, +5/16 and −18/19. Each of these eight regions was deleted by the CRISPR/Cas9 method in B6 mouse zygotes microinjected with flanking pairs of single-guide RNAs. Mice carrying homozygous deletions were examined after at least five backcrosses to B6 mice if the deletion was associated with a developmental phenotype, and at least two backcrosses in the absence of apparent phenotype. To assess phenotypic alterations in NKT and ILC lineage development, we directly examined the level of expression of PLZF and associated changes in precursor frequencies by flow cytometry of thymus and bone marrow cells, respectively. For skeletal phenotypes, we scored the presence of hindlimb autopod and zygodont abnormalities, namely ectopic or homeotically transformed digits and fibulation of tibia, as described.

Mice lacking segment +5/16 exhibited morphological defects in hind limbs, including extra or homeotically transformed digits (Fig. 1b, top row) and fibulation of tibia (Supplementary Fig. 1), that are characteristic of previously described coding mutations of Zbtb16. In contrast, PLZF expression and cell frequencies were unaltered in bone marrow ILCP, identified by their Lin−IL7Rα+41B7+PD1+ phenotype, and in thymic and splenic NKT, identified by CD1d-αGalCer tetramer staining (Fig. 1b, bottom rows).

In mice carrying the adjacent +18/32 deletion, the opposite phenotype was observed. These mutant mice lacked skeletal abnormalities, but had complete abrogation of PLZF expression, both in bone marrow ILCP and in thymic and splenic NKT (Fig. 1b). As a consequence, NKT thymocytes and splenocytes...
were massively decreased and ILCP were modestly decreased both in frequency and in absolute numbers, similar to Zbtb16−/− mice. Interestingly, expression of PLZF by γδ and MAIT thymocytes was also abolished (Supplementary Fig. 2). Another region, + 40/43, was associated with decreased NKT cell frequency, although the defect was less severe than + 18/32. Mice carrying the other genomic deletions, including + 185/191, + 120/124, + 115/116, + 75/81 and −18/19 failed to exhibit lymphoid or skeletal developmental phenotypes (Fig. 1b).
Altogether, this systematic deletion analysis of the chromatin accessible regions specifically associated with bone marrow ILCP or NKT thymocytes indicates the presence of a limited number of regulatory regions controlling Zbtb16 expression during the development of innate lymphoid lineages, and their clear separation from regions involved in the regulation of other developmental processes, such as limb patterning.

NKT phenotypes in the + 18/32 and + 40/43 deletion mutants. Because PLZF levels vary considerably during different stages of normal NKT cell development, we performed a detailed examination of NKT thymocytes in mutant mice carrying the + 18/32 deletion, where the NKT cell defect is complete, and in the + 40/43 deleted mice, where the NKT cell defect is partial. After MACS-enrichment using CD1d-αGalCer tetramers, we could zoom in on the rare residual NKT thymocytes of + 18/32Δ/Δ mice and show that they were arrested at the CD44− NK1.1− CD24− stage 1 (Fig. 2a), similar to the previously described defect in Zbtb16Δ/Δ mice.6 We also examined the earliest CD44− NK1.1− CD24hi developmental stage 0 cells,22 which are normally present at very low numbers of around 300 cells per thymus and represent the stage where PLZF is first induced. Whereas, in WT littermates, ~30% of stage 0 cells expressed a high level of PLZF, expression was totally abrogated after deletion of + 18/32. Thus, the NKT developmental defect associated with the + 18/32 deletion is identical to the reported defect of Zbtb16Δ/− mice.

In + 40/43ΔΔ mice, a modest (~20–40%) but statistically significant defect in PLZF mean fluorescence intensity was observed in both stage 0 (P = 0.015; Student’s t-test) and stage 1 (P = 0.0084; Student’s t-test) cells, while later stages 2 and 3 NKT cells expressed similar levels of PLZF as wild type (Fig. 2b). This altered pattern of expression resulted in a partial block of NKT development, particularly visible at the stages 2 to 3 transition. These findings, which were confirmed in multiple independent experiments (Supplementary Fig. 3), indicate that NKT levels are exquisite sensitive to modest changes in PLZF expression at early stages of development. This conclusion is in line with the similar NKT developmental defect observed in Zbtb16Δ/Δ heterozygous mice (Fig. 2c and Supplementary Fig. 3). In contrast with the NKT defect, the ILCP of + 40/43ΔΔ mice did not have alterations of PLZF expression (Fig. 1b), suggesting that + 40/43 specifically regulates NKT but not ILC development.

Dissection of the + 18/32 region. To further characterize the precise regulatory elements regulating ILCP and NKT development within the + 18/32 region, we generated five deletion mutants spanning this entire segment (Fig. 3a). Only one region, + 21/23, appeared to exert strong control on PLZF expression in ILCP and NKT cells (Fig. 3b). Indeed, mice carrying a deletion of + 21/23 exhibited undetectable PLZF expression in ILCP, similar to the larger + 18/32 deletion examined previously. In NKT thymocytes and splenocytes, the defect associated with the + 21/23 deletion was partial and did not fully recapitulate the + 18/32 deletion, suggesting the presence of additional ‘subthreshold’ regulatory elements that complement + 21/23. One such regulatory element may be encoded by + 29/31 whose deletion was not associated with alterations of PLZF expression (Fig. 1b), suggesting that this regulatory element may be encoded by + 29/31 whose deletion was partial and did not fully recapitulate the + 18/32 deletion.

Furthermore, a deletion centered around the 431/438 motif induced a statistically significant decrease in NKT that did not reach significance (Fig. 5e, f). Mice lacking another Runx1 motif-containing sequence, 199 had modest but significant decrease in PLZF expression by ILCP (P < 0.0001; analysis of variance (ANOVA)) as well as decreased NKT thymocytes (P = 0.0027; ANOVA) (Fig. 5d, f).

Impact of PLZF defect on NKT sublineages. The development of NKT thymocytes can also be described by their acquisition of T helper types 1, 2 or 17 programs, with NKT1, NKT2 and NKT17 characterized by their PLZFlow T-bethigh, PLZFhigh ROrgtlow and PLZFhigh ROrgthigh profiles, respectively. We found that NKT1 and NKT17 were predominantly affected while NKT2 were relatively preserved in the thymus and peripheral tissues of + 21/23Δ/Δ and in + 40/43ΔΔ mice, in which the decrease in PLZF expression is limited to the early developmental stages (Fig. 4 and Supplementary Fig. 4). In contrast, + 18/32ΔΔ mice lacked all NKT sublineages. Thus, early, transient defects in PLZF expression have long-range effects on terminally differentiated NKT sublineages.

Sequence-level dissection of the + 21/23 region. Region + 21/23 contains ATAC-seq reads over a 730 base-pair interval (Fig. 5a). The precise deletion of this 730 bp fragment reproduced the phenotype of the larger + 21/23 deletion, as expected, with complete abrogation of PLZF expression in ILCP and impaired PLZF expression in stage 0 NKT thymocytes (Fig. 5b). We subdivided 1/730 with individual deletions of 1/338 and 339/730 segments. Both deletions reproduced the ILCP and NKT phenotypes of the larger 1/730 deletion, suggesting the presence of essential regulatory elements in both fragments. Since both fragments contained consensus DNA sequence motifs for Runx factor binding (Fig. 5a), we generated new deletions of three of these sequences. Excision of the entire 307/317 motif, as well as partial deletions of this motif (309/314 and 311/317) induced a significant decrease in PLZF expression by ILCP (P < 0.0001; analysis of variance (ANOVA)) as well as decreased NKT thymocytes (P = 0.0027; ANOVA) (Fig. 5d, f).

Figure 2. NKT phenotype of + 18/32 and + 41/43 deletions. a FACS analysis of CD1d-αGalCer-tetramer+ MACS-enriched NKT thymocytes in mice carrying the + 18/32 deletion compared with WT littermates (LM) mice. Stage 0 NKT (gated as Tetramer+ CD24high) are analyzed for PLZF expression (upper panels). Tetramer+ CD24low NKT are stained for CD44 and NK1.1 to identify stages 1–3 and determine their percentages and PLZF mean fluorescence intensity (MFI) (bottom row). Summary data are combined from two independent experiments, with a total of three to four mice in each group. b FACS analysis of NKT developmental stages and PLZF levels in mice carrying the + 40/43 deletion compared with WT littermate (LM). Data are combined from two independent experiments, with a total of three to five mice in each group. c FACS analysis of NKT thymocytes in Zbtb16Δ/Δ mice. Two-tailed Student’s t-test was performed for statistical analysis. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.
Fig. 3 Dissection of the +18/32 region. a ATAC-seq tracks. b FACS analysis of BM ILCP (upper row) and NKT thymocytes (middle row) and splenocytes (bottom row) in mice carrying the indicated deletions. Summary data in the right column are compiled from six independent experiments, with a total of 4–13 mice in each group. One-way analysis of variance for multiple comparisons was used for statistical analysis. c FACS analysis of NKT developmental stages and PLZF expression in +21/23Δ/Δ mice and littermate controls (LM). Data are combined from two independent experiments with a total of four mice per group. Two-tailed Student’s t-test was performed. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001
Runx1-specific antibody (Supplementary Fig. 5), we demonstrated prominent binding of Runx1 to several regions of chromatin accessibility in the Zbtb16 locus (Fig. 6a). Thus, Runx1 ChIP-seq peaks were found within the +21/23, +29/31, +40/43, +115/116 and +120/124 regions in NKT thymocytes. Due to lower cell input, the ChIP analysis of ILCP was performed using ChIP-seq peaks were found within the +21/23, +29/31, +40/43, +115/116 and +120/124 regions in NKT thymocytes. Due to lower cell input, the ChIP analysis of ILCP was performed using one-way ANOVA for multiple comparisons to WT LM. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001

PLZF expression in mice with conditional Runx1 deletion. To further evaluate the role of Runx1 on NKT cell development and PLZF expression in vivo, we used a Cd4-Cre Runx1fl/− strain. As previously reported, Cd4-Cre Runx1fl/− mice exhibited a profound and selective defect in thymic NKT numbers compared with Cd4-littermates, while CD4 and CD8 SP cells appeared normal (Fig. 7a, b). Furthermore, Cd4-Cre Runx1fl/− heterozygous mice also exhibited a defect, albeit less severe than homozygous littermates, demonstrating haploinsufficiency. After MACS-enrichment of rare residual NKT thymocytes using CD1d-tGal-Cer tetramers, we determined that the residual Runx1-deficient

Zbtb16 expression in NKT thymocytes, including +185/191, +120/124, +115/116 and +5/16, indicating that Runx1-binding at a particular site is not necessarily associated with active regulation of Zbtb16 gene expression.
Fig. 5 Sequence-level analysis of the +21/23 region. a ATAC-seq profiles of the +21/23 region in indicated lymphoid populations, magnified from Fig 1a, reveals that all reads are contained within a 730 bp located between mm10 chr9:48,814,346 and 48,815,095. b FACS analysis of BM ILCP (top row), and thymic NKT (second row) in mice carrying indicated deletions, with summary plots on the right column. Bottom two rows shows the FACS analysis of NKT (third row), and thymic NKT cells (bottom row) in mice carrying indicated deletions, with summary plots on the right column. For ILCP analysis, Runx-motif mutant mice had the second allele carrying the +18/32 deletion. Summary data are pooled from four independent experiments, with 3–10 mice in each group. c e Schematic representation of CRISPR/Cas9 targeting of different Runx1 motif-containing sequences. c Runx1 motif (189–199) in mice carrying indicated deletions, with summary plots on the right column. FACS analysis of BM ILCP cells (top row), and thymic NKT cells (bottom row) in mice carrying indicated deletions, with summary plots on the right column. For ILCP analysis, Runx-motif mutant mice had the second allele carrying the +18/32 deletion. Summary data are pooled from four independent experiments, with 3–10 mice in each group. One-way ANOVA with multiple comparisons compared to WT were used for statistical analysis. *, P < 0.05; **, P < 0.01; ***, P < 0.001; ****, P < 0.0001.
NKT thymocytes were arrested at stage 1 and that they failed to express detectable PLZF at stage 0 and stage 1 (Fig. 7c), a phenotype identical to that of +18/32 deletion mutants shown in Fig. 2a and to that of Zbtb16−/− mice. Moreover, C4d-Cre Runx1fl/fl heterozygous mice exhibited defects in PLZF expression in stage 0 but not stage 1 NKT thymocytes, with a partial defect in total NKT thymocyte frequency (Fig. 7c). Thus, conditional deletion of Runx1 alters PLZF expression in NKT precursors in a dose-dependent manner. Furthermore, crossing C4d-Zbtb16 transgenic mice to C4d-Cre Runx1fl/fl heterozygous mice partially but significantly rescued the defect in NKT percentage (P = 0.0025; Student’s t-test) and cell number (P = 0.017; Student’s t-test) (Fig. 7d), supporting the conclusion that Runx1 promotes NKT cell development in part by upregulating PLZF expression. In addition, Runx1 bound and partially regulated the expression of Il7ra, Il2rb and Egr2, suggesting other levels of regulation of NKT development (Fig. 7e–h), perhaps accounting for the incomplete rescue of Runx1-deficient NKT development by the PLZF transgene. Although Egr2 could directly regulate PLZF13, its modest downregulation in Runx1-deficient NKT thymocytes was insufficient to account for the complete loss of PLZF.

To evaluate the role of Runx1 on ILCP development, we used an Il7ra-Cre Runx1fl/fl strain to direct deletion of Runx1 after the CLP stage. We observed modest but significant decrease in ILCP cell number (P = 0.0002; Student’s t-test) and a trend toward decreased PLZF expression compared with Il7ra-Cre controls, consistent with a role of Runx1 in ILCP development (Fig. 8a). There was only partial decrease of Runx1 among BM Lin−/− IL7Rα+/− α4β7+ ILCP precursors, however, perhaps due to a delayed kinetics of Il7ra-Cre-mediated excision (Fig. 8a, bottom). In addition, the known redundancy of Runx1 with other Runx proteins, particularly Runx3 which is expressed at the ILCP stage, might explain the modest phenotype. We therefore examined Il7ra-Cre Cbfbfl/fl mice lacking the common partner to all Runx proteins. Fig. 8b demonstrated a statistically significant defect in both ILCP numbers (P < 0.0001; Student’s t-test) and in their expression of PLZF (P = 0.0002; Student’s t-test). Thus, in contrast with NKT where Runx1 and Runx3 have little overlap (Supplementary Fig. 6) and deletion of...
Runx1 alone is associated with a complete developmental block, Runx3 might partially compensate for the loss of Runx1 in ILCP. Altogether, these results demonstrated the critical role of Runx1 and the shared intronic enhancer bound by Runx1 in regulating PLZF expression during ILC and NKT development.

**Discussion**

The TF PLZF is tightly regulated in a lineage-specific as well as temporal-specific manner to direct diverse lymphoid, skeletal and spermatogonial development programs, a pattern highly suggestive of enhancer-mediated regulation. By combining ATAC-seq and CRISPR/Cas9 analysis, our study identified a critical regulatory element within the intronic region +21/+23 that was critical for PLZF expression in both NKT and ILC lineages, as well as in MAIT and γδ cells. This region carried a high density of H3K27Ac and H3K4me1 marks and was heavily bound by Runx1. Although our study did not provide formal proof that Runx1 binding directly activated the enhancer, this region contained several Runx motifs whose individual deletions impaired...
PLZF expression in both ILC and NKT precursors. Furthermore, conditional deletion of Runx1 led to defective PLZF expression and impaired lineage development, which was partially rescued by transgenic expression of PLZF. Runx1 also exerted modest regulation of other factors involved in development, such as IL7Rα, IL2Rβ and Egr2, which might explain why the Runx1 deficiency phenotype was only partially rescued by transgenic PLZF. Altogether, these findings demonstrate that Runx1 is a critical TF and suggest that it directly activates PLZF at a key enhancer shared by ILC and NKT lineages. They reveal the existence of a Runx1-PLZF axis as a central component of the enhancer shared by ILC and NKT lineages. They reveal the critical TF and suggest that it directly activates PLZF at a key stage 1 NKT.

Furthermore, Runx1 was previously reported to be essential for NKT progression and the unexpected impact of relatively modest alterations related to the unexpected impact of relatively modest alterations of the level or the kinetics of early PLZF expression on the stage progression and the final frequency of NKT cells. Thus, a transient defect limited to the very early stage of NKT cell development was associated with long-range impact in the development of terminally differentiated NKT sublineages, particularly NKT1 and NKT17. Examples of these alterations were found in +21/23ΔΔ, +41/43ΔΔ and Cd4-Cre Runx1Δ/Δ mice, as well as in Zbtb16ΔΔ mice. As broad variations in NKT cell frequencies have been reported between different mouse strains and in humans, our findings suggest that relatively minor alterations of PLZF expression at early thymic stages, which would go undetected if not specifically investigated, could account for significant variations in NKT development.

Our study also identifies an adjacent but distinct skeletal specific enhancer region controlling limb patterning. The non-overlapping pattern of enhancers governing PLZF expression in lymphoid and in osteoblast precursors is consistent with emerging knowledge about gene regulation in different mammalian tissues. Furthermore, as osteoblast differentiation is characteristically triggered by Runx2 expression in mesenchymal cells, before expression of PLZF, it is intriguing to speculate that an analogous Runx2-PLZF axis may govern bone formation as well.

While traditional approaches to dissect the regulation of gene expression begin with a candidate TF and search for potential target sites, our study demonstrates the value of combining genome wide methods such as ATAC-seq and CRISPR/Cas9 to identify the genomic regulatory element first and then search for candidate TFs based on motif analysis and transcriptional context. This combination of approaches has the advantage of being systematic and unbiased, and is readily applicable to a wide range of systems. In addition, more refined epigenetic studies to identify poised or active enhancers, which were not available in our study due to low cell input, can potentially narrow down the number of targets for CRISPR/Cas9 analysis. While this approach is extremely powerful to identify key enhancers and focus the investigations on well-defined regulatory elements, there remain limitations in dissecting the full complement of factors regulating the enhancers, for example in the absence of obvious consensus binding motif. In that regard, it is possible that factors other than Runx1 may also regulate the common NKT/ILC enhancer.

Another advantage of this ATAC-seq/CRISPR/Cas9 combination approach is that valuable enhancer mutant model strains can be generated with selective defects of some but not other PLZF functions. In the current study, for example, +21/23ΔΔ mice exhibited PLZF defects across the innate lymphoid system, without associated defects in tissues such as bone and sperm which have confounded or complicated the studies of PLZF-deficient mice.

![Fig. 7](https://www.nature.com/naturecommunications) Conditional ablation of Runx1 in NKT. a Intraocular flow cytometry for Runx1 expression in thymic CD4SP and NKT cells. (LM controls (LM), solid histogram; Cd4-Cre Runx1Δ/Δ (Het), dashed histogram; Cd4-Cre Runx1Δ/Δ (KO), gray shaded.) b NKT cells in the thymus of Het and KO compared to their LM. The frequency and absolute cell number of NKT are summarized on the right panels. Data are representative of five to eight mice from four independent experiments. c Individual thymus of indicated strains were MACS-enriched using Cd1d-aGalCer tetramers before FACS analysis of NKT development subsets and PLZF expression. Data are summarized from four independent experiments, with 4-12 mice in each group. d Rescue of NKT defect in Het by PLZF transgene expression. Data are representative of four to 11 mice from three independent experiments. e Representative FACS analysis of IL-7Rα and IL-2Rβ expression in stage 1 NKT thymocytes from LM and KO; mean ± S.E.M. of mean fluorescence intensity (MFI) is 5535 ± 331 (n = 4) vs 1501 ± 40 (n = 3) for IL-7Rα; and 5721 ± 87 (n = 4) vs 3040 ± 163 (n = 3) for IL-2Rβ, respectively. Leftmost histogram represents uninfected control. f Runx1 ChIP-seq tracks at the Il7ra and Il2rb loci. g FACS analysis of Egr2 expression in CD4 SP, stage 0 and stage 1 NKT thymocytes from LM, Het and KO. Mean ± S.E.M. of Egr2 MFI: 2090 ± 84 (LM, n = 5), 1927 ± 45 (Het, n = 4) and 1506 ± 18 (KO, n = 6) in CD4 SP; 9688 ± 461 (LM, n = 5), 8872 ± 389 (Het, n = 4) and 5722 ± 316 (KO, n = 6) in stage 0 NKT; 4174 ± 287 (LM, n = 5), 4310 ± 257 (Het, n = 4), and 3892 ± 216 (KO, n = 6) in stage 1 NKT. h Runx1 ChIP-seq track Egr2 locus. Statistical analysis was performed using one-way ANOVA with multiple comparisons. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.
In summary, we report the successful combination of ATAC-seq and CRISPR/Cas9 methods to systematically dissect the extended Zbtb16 locus in vivo, revealing a shared lymphoid enhancer and the critical role of Runx1 in PLZF expression and the development of innate and innate-like lineages. These findings further advance our understanding of the architecture of the common transcriptional regulatory network that governs a fundamental lineage decision between innate and adaptive lymphocyte development.

Methods

Mice. C57BL/6 (stock no. 000664), B6.129P2-Ifg-7ab/J (stock no. 028550) and Cd4-Cre (B6.Cg-Tg(CD4-cre)1Cwi/J (stock no. 008772), Cd4-Cre both were a gift from Dr. P. P. Pandolfi (Beth Israel Deaconess Medical Center, Boston, MA) and were backcrossed to C57BL/6 for at least nine generations18. The Zbtb16IRESGFPcre (B6.129S7-Ifg-7ab/J) (stock no. 024529) and the C57BL/6 Cd4-Zbtb16 transgenic (stock no. 014644) mice were generated,33 and maintained in our laboratory. C57BL/6 Il7ra-Cre was obtained from Dr Rodeck34. All the CRISPR/Cas9-mediated knockout mice were generated by micro-injection of C57BL6 zygotes in the transgenic core service of the University of Chicago. The age of mice was between 4- and 8-week old for all the experiments performed. All mice were raised in a specific pathogen-free environment, and killed by CO2. The permission was granted to perform all mice experiments by the Institutional Animal Care and Use Committee of the University of Chicago.

Flow cytometry and MACS enrichment. Bone marrow was isolated by gently crushing femurs and tibiae before filtrations with 70 μm filters. Cell suspensions were incubated with purified anti-CD11b/32 (clone 93) for 10 min on ice to block Fc receptors. The following fluorochrome-labeled monoclonal antibodies were used: α4β7 (DAKJ32; 1:200), B220 (RA3-6B2; 1:400), CD3ε (17A2; 1:400), CD4 (L3T4; 1:400), CD8α (53-6.7; 1:400), CD11b (M1/70; 1:400), CD11c (N418; 1:400), CD19 (6D5; 1:400), CD25 (PC61; 1:200), CD45.2 (104; 1:100), Kit (288; 1:100), Il3 (A2F10; 1:100), Gr-1 (RB6-8C5; 1:400), Il2Rα/CD122 (31H4; 1:50), Il-7Rα/CD127 (A7R34; 1:20), Nk1.1 (PK136; 1:200), Pdi (133; 1:50), Sca-1 (D7; 1:500), Tcrβ (H57-597; 1:400) and Ter-119 (TER-119; 1:400). All the antibodies were purchased from eBioscience, BD Biosciences or BioLegend. For intracellular staining, cells were fixed and permeabilized using the eBioscience Foxp3 Staining Buffer Set and then stained with fluorochrome-labeled Egr2 (12-6691-82; eBioscience; 1:50), PLZF (563490 or 564850; BD Biosciences; 1:200 or 1:1000), RORγt (563490 or 564850; BD Biosciences; 1:200) or CD1d (564723; BD Biosciences; 1:1000, RORγt; 564723; BD Biosciences; 1:1000, CD1d; 564723; BD Biosciences; 1:1000, CD1d; 564723; BD Biosciences; 1:1000). ATAC-seq was performed as described20. Nuclei were prepared from NIH tetramer facility. For NKT or MAIT cell MACS enrichment, samples were labeled with APC-conjugated anti-PLZF antibody, bound to anti-APC microbeads (Miltenyi Biotec), and then subjected to positive selection on autoMACS (Miltenyi Biotec). ILCP were sorted as Lin-α4β7- cells from eBioscience, BD Biosciences or BioLegend.
of 10–12 cycles. The libraries were then purified using a Qiagen PCR MinElute kit and size selected in the 150–650 bp range. The size-selected libraries were quantified using the Agilent Bioanalyzer and by qPCR using the KAPA Library Quantification Kit. Libraries were sequenced on the Illumina HiSeq2000 system to generate 30–50 million reads. ATAC-seq raw reads of 50 bp were aligned to the mm10 mouse genome using bowtie allowing for two mismatches. The bedgraphs of ChIP-seq were generated using hypergeometric optimization of motif enrichment (HOMER), where the total number of aligned reads was normalized to 10 million. For each cell type, two or three biological replicates were performed.

**Microinjection into mouse zygotes.** Vectors pT3Ts-nCas9n (46757) and pT7-gRNA (46759) were obtained from Addgene. A pair of oligos for each targeting site (Supplementary Table 1) was annealed and ligated to linearized pT7-gRNA vector. LINEARIZE pT3Ts-nCas9n was purified and used as the template for in vitro transcription (IVT) using mMESSAGE mMACHINE T3 Transcription Kit (AM1348; ThermoFisher Scientific). The ligated sgRNA constructs were linearized and then used as the template for IVT using MEGAscript T7 Transcription Kit (AM1354; ThermoFisher Scientific). Both the Cas9 mRNA and the sgRNAs were purified using MEGAClear (AM1908; ThermoFisher Scientific) and eluted in RNase-free water.

C57BL/6 females at 9–15 weeks of age were screened for non-estrous state and were superovulated with 5–5.5 IU PMSG (Millipore). 46.5 h later mice were given 5–5.5 IU HCG (Sigma) and mated 1.5 h later to C57BL/6 males. The next morning the females were checked for the presence of a vaginal plug, and those having one were harvested for fertilized eggs. Injection of the CRISPR reagents into the pronucleus was performed by using an Eppendorf FemptoJet. Surviving eggs were transplanted in CD-1 pseudopregnant females. The founders were screened with PCR followed by sequencing. The sgRNAs and corresponding deleteme are listed in Supplementary Table 1.

**ChIP-seq and ChIP-qPCR.** For Runx1 ChIP-seq in NKT cells, 20 million MACS-enriched NKT cells from Vv14-tg18 mice were fixed for 15 min at room temperature with 10% formaldehyde. Glycine was then added to a final concentration of 0.125 M to quench the reaction. Cells were washed twice with ice-cold phosphate-buffered saline, lysed and followed by sonication to fragment DNA to 100–200 bp. The chromatin was then incubated overnight with Runx1 antibody (ab23980; Abcam), followed by 2 h incubation with Protein G Dynabeads (10003D; Thermo Fisher Scientific). The precipitated ChIP and input DNA were washed, reverse crosslinked and digested with proteinase K and RNase A. The DNA was then purified with phenol/chloroform extraction or Qiagen PCR MinElute kit. Sequencing libraries were prepared by standard end-repair, adenylation, adaptor-ligation, PCR amplification and size-selection procedure. For Runx1 ChIP-qPCR in BM Lin+ IL7Rα+ αβ+γδ+ cells, BM cells were MACS-enriched with anti-CD3ε before staining with other antibodies and sorting. About 25,000–50,000 cells from 10 to 20 mice were used for each biological replicate. The primers used for ChIP-qPCR are listed in Supplementary Table 2.

**Immunoprecipitation and western blot analysis.** 293 cells (CRL-1573; ATCC) were transfected with Flag-mRUNX (construct 14585; Addgene). Cells were lysed with RIPA buffer 24 h post transfection and immunoprecipitated with IgG control, were transfected with Flag-mRunx1 construct (AM1348; ThermoFisher Scientific). The ligated sgRNA constructs were linearized and then used as the template for IVT using MEGAscript T7 Transcription Kit (AM1354; ThermoFisher Scientific). Both the Cas9 mRNA and the sgRNAs were purified using MEGAClear (AM1908; ThermoFisher Scientific) and eluted in RNase-free water.
Grants NIH R01 AI038339 and AI108643 (A.B.) and the University of Chicago Digestive Diseases Research Core Center P30DK042086.

**Author contributions**
A.-P.M. and A.B. designed research. A.-P.M., I.E.I. and D.N.K. performed research. M.M. helped with ATAC-seq. A.-P.M. performed bioinformatics analysis. A.-P.M. and A.B. analyzed the data and wrote the manuscript.

**Additional information**
Supplementary Information accompanies this paper at doi:10.1038/s41467-017-00882-0.

**Competing interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at http://npg.nature.com/reprintsandpermissions/

**Change history:** A correction to this article has been published and is linked from the HTML version of this paper.

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