The transcription factor TCF-1 enforces commitment to the innate lymphoid cell lineage

Christelle Harly1,2,3, Devin Kenney1, Gang Ren4, Binbin Lai4, Tobias Raabe5, Qi Yang6, Margaret C. Cam7, Hai-Hui Xue8,9, Keji Zhao4 and Avinash Bhandoola1*.

Innate lymphoid cells (ILCs) play important functions in immunity and tissue homeostasis, but their development is poorly understood. Through the use of single-cell approaches, we examined the transcriptional and functional heterogeneity of ILC progenitors, and studied the precursor-product relationships that link the subsets identified. This analysis identified two successive stages of ILC development within T cell factor 1-positive (TCF-1+) early innate lymphoid progenitors (EILPs), which we named ‘specified EILPs’ and ‘committed EILPs’. Specified EILPs generated dendritic cells, whereas this potential was greatly decreased in committed EILPs. TCF-1 was dispensable for the generation of specified EILPs, but required for the generation of committed EILPs. TCF-1 used a pre-existing regulatory landscape established in upstream lymphoid precursors to bind chromatin in EILPs. Our results provide insight into the mechanisms by which TCF-1 promotes developmental progression of ILC precursors, while constraining their dendritic cell lineage potential and enforcing commitment to ILC fate.

Results
EILPs are transcriptionally heterogeneous. We performed single-cell RNA sequencing (scRNA-seq) on Ly-6D–B220–CD19–Mac-1–Gr-1–CD11c–Ter119–NK1.1–CD3ε–CD8α–CD8β–CD4–TCRβ–TCRγδ+ (Lin+) Kit+2B4α4β7–Tcf7–GFP+ cells isolated from the BM of Tcf7−GFP mice, which express enhanced green fluorescent protein (EGFP) under the control of Tcf7 regulatory elements. These Lin–+Kit–2B4α4β7–Tcf7–GFP+ cells included EILPs, ILC precursors (ILCPs) and early ILC2 progenitors (ILC2Ps)1–3. We also sequenced Lin–+Kit–2B4α4β7+Flt3–IL-7Rα–all-lymphoid progenitors (ALPs), which are upstream of ILC progenitors and are Tcf7−GFP− (Fig. 1a,b). After clustering, ALPs formed one cluster (cluster 1), whereas Tcf7−GFP+ cells comprised five clusters (clusters 2–6; Fig. 1c), which we identified on the basis of the expression of key transcription factors1–3,4. The Zbtb16loLmo4hi cluster 4 corresponded to ILCPs; the Zbtb16hiBcl11bhi cluster 5 corresponded to ILC2Ps; and clusters 2, 3, and 6, which were Spi1−Irf8−Nfil3− Zbtb16hiId2lo, were putative EILP clusters (Fig. 1c,d). Consistently, clusters 2, 3, and 6 had very low expression of the cytokine receptor II7r compared with all other ILC progenitor cells1–3.

Next, we examined the developmental relationship between clusters using pseudotime reconstruction (Fig. 1e). A main developmental progression linked ALPs, EILPs, ILCPs and ILC2Ps (clusters 1–5; Fig. 1e). Analysis of differentially expressed transcription factors between clusters showed progressive upregulation or downregulation of factors such as Irf8, Spi1, Tcf7 or Tox along the pseudotime (Fig. 1f). In an alternative developmental progression, cluster 6 EILPs arose from cluster 2 EILPs (Fig. 1e). Cluster 6 EILPs had high expression of the transcription factors Batf3, Irf8, Spi1, Nfil3 and Id2 (Fig. 1f), which are associated with conventional
type 1 DC (cDC1) development, as well as the DC structural gene Cd74 and major histocompatibility complex class II (MHC-II) molecules (Fig. 1f and Supplementary Table 1). Cluster 6 EILPs appeared transcriptionally similar to cDC1-committed pre-cDC1 (refs. 11,12), although these cDC1 progenitors are CD11c+ and were excluded by the LinILC cocktail. Zbtb46, which is upregulated on committed cDC BM precursors11,12, was not detectably expressed in cluster 6 EILPs (Supplementary Table 1), suggesting that these cells represent an early stage of cDC1 priming. ILC-specific transcription factors, such as Tlox, Tcf7 and Runx3, which are expressed in EILPs, have no described function in DCs and are not expressed in DC progenitors13, were detectable in cluster 6 (Fig. 1f), and their expression was downregulated along the pseudotemporal progression (Fig. 1f). On the basis of these results, we designated cluster 2 and cluster 6 as specified EILPs (sEILPs), comprising sEILP1s (cluster 2) and sEILP2s (cluster 6). sEILPs and sEILP2s highly expressed the cDC genes Flt3, Il7r, Nfil3 and Spi1. Cluster 3 EILPs, which had downregulated the expression of Flt3, Il7r, Nfil3 and Spi1 (Fig. 1f), were designated as committed EILPs (cEILPs). Our analysis placed sEILP1s at a branch point between the ILC and cDC1 lineages, indicating that ILC progenitors might access cDC1 lineage fate in vivo (Fig. 1g).

sEILPs, but not cEILPs, have DC lineage potential. Next, we characterized the DC potential of EILPs. After 7 d of culture in the presence of OP9 stromal cells and the cytokines stem cell factor, Flt3L and interleukin-7 (IL-7) (hereafter, SF7 conditions), together with granulocyte–macrophage colony-stimulating factor and IL-3, which support DC survival and expansion (hereafter, SF7-GM3 conditions), Lin−Kit+2B4−α4β7+Tcf7−GFP−Thy1−EILPs, but not Lin−Kit+2B4−α4β7+Tcf7−GFP+Thy1−EILPs, gave rise to Mac-1−CD11c+MHC-II+ cDCs (Fig. 2a). Seven days after transfer into irradiated mice, EILPs, but not cEILPs, differentiated into splenic CD11c+MHC-II+ cDCs (Fig. 2b). To address whether the DC potential of EILPs was due to a small subset of DC precursor cells that co-purified with EILPs, we used Il7r-iCre R26-stop-YFP mice, in which cells that express Il7r permanently express yellow fluorescent protein (YFP). In these mice, most EILPs and ~15% of cDCs are Il7r−iCre R26-stop-YFP+. We isolated EILPs and LinILC−Kit+α4β1−Flt3+2B4+α4β7−IL-7R− lymphoid-prone multipotential progenitors (LMPPs) from the BM of Il7r−iCre R26-stop-YFP Tcf7−GFPm mice and cultured them for 5 d under SF7-GM3 conditions. About 80% of EILPs and EILP-derived Mac-1+ cells were YFP+ (Fig. 2c). In contrast, only 10% of IL-7Rα+ LMPPs and IL-7Rα− LMPP-derived Mac-1+ cells were YFP+ (Fig. 2c), suggesting that cDCs develop from EILPs.

To quantify the frequency of DC-competent precursors within EILPs, we sorted single EILPs on OP9 cells in 96-well plates and cultured them for 10 d under SF7-GM3 conditions, or with the addition of IL-7Rα− LMPPs and IL-7Rα− LMPP-derived Mac-1+ cells (hereafter, SF7-GM3-MG6 conditions)14. Under both sets of conditions, CD45+ EILP-derived colonies contained Mac-1+CD11c+MHC-II+ cDCs (Fig. 2d). Mac-1+ cells were ILC lineage cells (Supplementary Fig. 2a–c)1,2. In the SF7-GM3-MG6 cultures, >50% of CD45+ wells contained only Mac-1+ cDCs, ~40% contained only Mac-1− ILCs, and <5% contained both ILCs and DCs (Fig. 2d,e). The frequency of Mac-1+ colonies was decreased under SF7-GM3 conditions compared with SF7-GM3-MG6 conditions, but the size of these colonies was comparable under both conditions (Fig. 2e and Supplementary Fig. 2d), suggesting that MG6 cytokines affected the survival of some EILPs. Furthermore, >50% of DC-competent EILPs gave rise to both CD24+ cDC1 and CD172a+ cDC2 under SF7-GM3-MG6 and SF7-GM3 conditions (Supplementary Fig. 2f)1,11. Together, these results indicate that a large fraction of EILPs possessed DC potential. However, 40% of EILPs exclusively generated ILCs.

Next, we examined whether transcriptional heterogeneity identified by scRNA-Seq in EILPs correlated with heterogeneity identified in differentiation assays. Because the expression of Flt3 and promyelocytic leukemia zinc finger (PLZF) protein (encoded by Zbtb16) inversely correlates in EILPs, we tested whether Flt3 and PLZF could be used to distinguish the Flt3hiZbtb16hi EILPs and Flt3loZbtb16lo EILPs identified by scRNA-Seq. We crossed newly generated Tcf7YFP reporter mice, which express YFP downstream of the Tcf7 gene (Supplementary Fig. 3), with Zbtb16hi/YFP reporter mice. We used index sorting, which allows a posteriori analyses linking flow cytometric measurements with developmental fate, to sort 1,176 single EILPs from Tcf7hiZbtb16hi mice into 96-well plates, then cultured them under either SF7-GM3 or SF7-GM3-MG6 conditions. Analysis of heterogeneously expressed factors (Zbtb16-GFP, Tcf7−YFP and Flt3) and cell fate (generation of DCs or ILCs) for single EILPs indicated that almost all DC progenitor cells were Flt3hiZbtb16-GFP+, whereas ILC progenitors were mostly Flt3hiZbtb16-GFP− (Fig. 2f). Using Flt3 and Zbtb16-GFP, we defined gates that matched the EILP fate (Fig. 2f). These gates identified a Flt3hiZbtb16-GFP+ population that gave rise almost exclusively to Mac-1+ ILCs under SF7-GM3 and SF7-GM3-MG6 conditions (Fig. 2g,h), and a Flt3loZbtb16-GFP− subset that predominantly gave rise to Mac-1− DCs under SF7-GM3-MG6 conditions, but predominantly gave rise to ILC under SF7-GM3 conditions (Fig. 2g,h).

Bulk RNA-Seq of Flt3hiZbtb16-GFP+ and Flt3loZbtb16-GFP− EILP subsets (Supplementary Fig. 4a) indicated that they were transcriptionally similar to sEILPs and cEILPs, respectively (Fig. 3a). We examined expression of the transcription factors GATA binding protein 3 (GATA-3), thymocyte selection-associated high mobility group box (TOX), Nfil3, purine-rich binding protein PU.1 and interferon regulatory factor 8 (IRF-8) in Flt3hiPLZFlo sEILPs and Flt3hiPLZFlow cEILPs. Consistent with the bulk RNA-Seq and scRNA-Seq, GATA-3 expression was upregulated from sEILPs to cEILPs; TOX expression was high in sEILPs and cEILPs; and PU.1 (encoded by Spi1), IRF-8 and Nfil3 expression was downregulated from sEILPs to cEILPs (Fig. 3b). Thus, Flt3hiZbtb16-GFP+ and Flt3loZbtb16-GFP− EILP populations correspond, respectively, to the sEILP and cEILP populations identified by scRNA-Seq.
sEILPs differentiate into cEILPs. Next, we examined the developmental relationship between sEILPs and cEILPs. We examined ILC-positive colonies derived from single sEILPs and cEILPs cultured for 10 d under SF7-GM3 conditions. Mac1-1ICOS-NK1.1- Tcf7-YFP+ cells with the phenotype of ILC progenitors (Supplementary Fig. 2) were more abundant in sEILP-derived colonies (70%) than cEILP-derived colonies (50%; Supplementary Fig. 4b), suggesting that sEILPs were more immature than cEILPs. ILC progeny (Tcf7+NK1.1+ ILC1s or NK cells, Tcf7+ICOS+α4β7+ ILC2s and Tcf7+ICOS+α4β7+ ILC3s) were similar between sEILP- and cEILP-derived colonies (Supplementary Fig. 4c). However, sEILP-derived ILC colonies were larger than cEILP-derived ILC colonies (Fig. 3c). Furthermore, sEILPs were larger than cEILPs (Supplementary Fig. 4d), and contained a greater fraction of cycling cells (27% of DAPI+ sEILPs versus 5% of
cEILPs; Supplementary Fig. 4e), consistent with a larger proliferative capacity of more upstream progenitors.\(^{16}\)

To test whether sEILPs were upstream of cEILPs, we isolated Flt3\(^+\)Zbtb16-GFP\(^+\) sEILPs and Flt3\(^+\)Zbtb16-GFP\(^+\) cEILPs from Tcf7\(^{GFP}\)Zbtb16\(^{GFP}\) mice and cultured them for 2 d under SF7 conditions. Most sEILPs downregulated Tcf7\(^{YFP}\) and upregulated Mac-1 (Fig. 3d), consistent with differentiation into DCs. Most Tcf7\(^{YFP}\) sEILP progeny were similar to Zbtb16-GFP\(^+\)Thy1\(^{−}\) cEILPs and Zbtb16-GFP\(^+\)Thy1\(^{+}\) ILCPs (Fig. 3d). sEILP-derived PLZF\(^{−}\)Thy1\(^{−}\) sEILPs, PLZF\(^{−}\)Thy1\(^{−}\) cEILPs and PLZF\(^{−}\)Thy1\(^{+}\) ILCPs showed transcription factor expression consistent with corresponding ex vivo populations (GATA-3, TOX, NFIL3, PU.1 and IRF-8; Supplementary Fig. 4f). Most cEILP-derived cells were Tcf7\(^{YFP}\)Zbtb16-GFP\(^+\)Thy1\(^{+}\) ILCPs, and had higher expression of Thy1 than sEILP-derived ILCPs (Fig. 3d). To establish that sEILP-derived cEILPs were similar to ex vivo cEILPs, we isolated Tcf7\(^{YFP}\)Zbtb16-GFP\(^+\)Thy1\(^{−}\) cEILPs from sEILP SF7 cultures after 2 d, and cultured them for 4 d under SF7 conditions. Only 10% of Mac-1\(^{+}\) DCs were
Articles Nature Immunology

−expression was similar on NRP-1 sEILPs cultured under the same conditions. However, in vitro Flt3−Gr-1−Ter119−NK1.1−CD3−Tcf7−sEILPs and cEILPs isolated from t0.01). -test with Welsh correction was performed to determine the significance (**%

A two-tailed unpaired Student’s t-test with Welsh correction was performed to determine the significance (**%

To determine whether sEILP1s and sEILP2s represented two success- tages of development towards DCs, we isolated Flt3+Zbb16-GFP+NRP-1− sEILP1s and Flt3+Zbb16-GFP+NRP-1−MHC-II− sEILP2s from Tcf7+Zbb16GFP mice and examined their fate after 2 d of culture under SF7 conditions. In total, 16% of sEILP1-derived cells were phenotypically similar to Tcf7−YFP+NRP-1− sEILP2s (Fig. 4c). The remaining cells were Tcf7−YFP+NRP-1− (Fig. 4c) and had upregulated the expression of the DC markers Mac-1, MHC-II and CD11c, which were mostly DCs (Fig. 4c and Supplementary Fig. 4i). More Tcf7− sEILP2 progeny expressed Mac-1, MHC-II and CD11c compared with Tcf7− sEILP1 progeny, and at higher levels (Fig. 4c and Supplementary Fig. 4i). Both sEILP1s and sEILP2s gave rise to Mac-1+ DCs after culture for 10 d under SF7-GM3 and SF7-GM3-MG6 conditions; however, sEILP1-derived DC colonies were larger than sEILP2-derived DC colonies (Supplementary Fig. 4j,k), consistent with a larger proliferative capacity of more upstream progenitors. These results indicate that sEILP1s transitioned through the sEILP2 stage before DC differentiation.

generated under these cultures, similar to ex vivo Flt3+Zbb16+ sEILPs cultured under the same conditions. However, in vitro-generated cEILPs and ex vivo cEILPs had similar potential to generate NK1.1+ or ICOS+ ILCs (Fig. 3e). Our results support a linear developmental relationship linking sEILPs, cEILPs and ILCPs.

eSILPs develop into cDC1 precursors in vivo. Using scRNA-Seq, we identified Nr1p as encoding a surface marker that could resolve sEILP1s and sEILP2s (Fig. 1f). Neurilin-1 (NRP-1) was homogeneously expressed by Lin+(B220−CD19−Thy1−Mac-1−Gr-1−Ter119−NK1.1−CD3−CD8−CD4−Tcrβ−Tcrγδ−) Flt3−CD11c−MHC-II− cEILPs, whereas pre-DCs, as well as a fraction of sEILPs that corresponded to sEILP2s (Fig. 4a), C-C chemokine receptor type 2 (Ccr2) and MHC-II+, which were upregulated from sEILP1 to sEILP2 at the RNA level (Fig. 1f), were upregulated on NRP-1+ sEILPs compared with NRP-1− sEILP1s (Fig. 4b). Expression of Tcf7+YFP, PLZF and IRF-8 on NRP-1+ sEILP2s was intermediate between NRP-1− sEILP1s and pre-DCs, whereas TOX expression was similar on NRP-1− sEILP1s and NRP-1+ sEILP2s (Supplementary Fig. 4g,h).

Fig. 3 Characterization of sEILP and cEILP populations. a, RNA-Seq analysis of ALPs, sEILPs, cEILPs and ILCPs, showing the heat map of expression of transcription factors from Fig. 1f. Data are averaged from two samples for sEILPs and ILCPs, or three samples for ALPs and cEILPs, as shown in Supplementary Fig. 4a, 4b and 4c. b, Flow cytometric analysis of Lin−Kit+Thy1+2B4+7-KitintFlt3hi cells (gray), sEILPs (black), cEILPs (orange) and ILCPs (purple). c, Quantification of absolute numbers of ILCs per ILC-positive colony from Fig. 2h. Data are presented as means ± s.d. for n=78 sEILP- and n=143 cEILP-derived colonies. d, Flow cytometric analysis of cultures from Thy1−2B4−Sca1−KitintFlt3hi cells (gray), sEILPs (black), cEILPs (orange) and ILCPs (purple). These cells were phenotypically similar to Tcf7−YFP+NRP-1− sEILP2s (Fig. 4c). The remaining cells were Tcf7−YFP+NRP-1− (Fig. 4c) and had upregulated the expression of the DC markers Mac-1, MHC-II and CD11c (Supplementary Fig. 4i), consistent with DC differentiation. Conversely, 82% of sEILP2 progeny were NRP-1− and were mostly DCs (Fig. 4c and Supplementary Fig. 4i). More Tcf7− sEILP2 progeny expressed Mac-1, MHC-II and CD11c compared with Tcf7− sEILP1 progeny, and at higher levels (Fig. 4c and Supplementary Fig. 4i). Both sEILP1s and sEILP2s gave rise to Mac-1+ DCs after culture for 10 d under SF7-GM3 and SF7-GM3-MG6 conditions; however, sEILP1-derived DC colonies were larger than sEILP2-derived DC colonies (Supplementary Fig. 4j,k), consistent with a larger proliferative capacity of more upstream progenitors. These results indicate that sEILP1s transitioned through the sEILP2 stage before DC differentiation.
To determine whether EILPs differentiated into DCs in vivo at steady state, we analyzed Lin<sup>−</sup>C<sup>+</sup>-K<sup>+</sup>-Thy1<sup>−</sup>2B4<sup>+</sup>-α4<sup>+</sup>-TCF-1<sup>−</sup>-YFP<sup>+</sup>-EILPs for CD11c expression (Fig. 4d). This analysis identified a population of CD11c<sup>+</sup> cells phenotypically similar to EILPs (CD11c<sup>+</sup> EILPs; Fig. 4d). Expression of NRP-1 (Fig. 4d), MHC-II and CCR2 proteins (Fig. 4e) was low on sEILPs1, intermediate on sEILPs2, and high on CD11c<sup>+</sup> EILPs. Tcf7<sup>−</sup>-GFP was lower on CD11c<sup>+</sup> EILPs (red) compared with pre-DCs (gray). Further characterization of CD11c<sup>+</sup> EILPs using surface markers described on pre-DC populations (CD24, CCR2, Ly-6D, Ly-6C and Siglec-H) showed that CD11c<sup>+</sup> EILPs were phenotypically similar to pre-cDC1s (Supplementary Fig. 5c).

To test whether CD11c<sup>+</sup> EILPs were derived from EILPs, we analyzed these cells in Tcf7<sup>−</sup>-iCre R26-stop-YFP mice crossed with Tcf7<sup>−</sup>-GFP mice. More than 80% of CD11c<sup>+</sup> EILPs were Tcf7<sup>−</sup>-iCre R26-stop-YFP<sup>+</sup>, similar to sEILPs (Fig. 4f). In contrast, only ~10% of DC precursor populations (pre-cDC1 and pre-cDC2), and even fewer Lin<sup>−</sup>Kit<sup>−</sup> upstream hematopoietic progenitor cells, were Tcf7<sup>−</sup>-iCre R26-stop-YFP<sup>+</sup> (Fig. 4f). In Tcf7<sup>−</sup>-Cre<sup>−</sup> mice, CD11c<sup>+</sup> EILP and sEILP absolute numbers were 3–4-fold reduced compared with Tcf7<sup>−</sup>-Cre<sup>+</sup> littermates (Fig. 4g), whereas canonical DC precursors such as pre-cDC1s and pre-cDC2s were present in normal numbers (Fig. 4g). These results indicate that CD11c<sup>+</sup> EILPs derive from EILPs.

Although EILPs are tenfold less abundant than pre-cDC1 and pre-cDC2 (Fig. 4g), we examined whether ILC progenitors contribute to DC generation in vivo. In Tcf7<sup>−</sup>-GFP/Ill<sup>−</sup>-iCre R26-stop-YFP mice, Tcf7<sup>−</sup>-GFP/Ill<sup>−</sup>-iCre R26-stop-YFP<sup>+</sup> cells represented ~5% of pre-cDC1 in vivo, and were almost undetectable within other DC precursor populations (Supplementary Fig. 5e). These cells likely include CD11c<sup>+</sup> EILPs, and other putative EILP-derived DC lineage cells. Furthermore, Tcf7<sup>−</sup>-Lin<sup>−</sup>Il7r-iCre Sca-1<sup>+</sup> hematopoietic progenitor cells, which are deficient in EILP generation<sup>11</sup>, generated CD8α<sup>−</sup>Mac-1<sup>−</sup> cDC1s normally in long-term competitive BM chimeras (Supplementary Fig. 5f), suggesting that EILPs were unlikely to contribute significantly to DC development in vivo. Our results establish that EILPs generate cDC1 precursors in vivo at steady state, and support a linear developmental relationship linking sEILPs1, sEILPs2 and the pre-cDC1 subset CD11c<sup>+</sup> EILPs.

**TCF-1-deficient SEILPs do not generate cEILPs.** Next, we investigated whether TCF-1 regulates the transition from sEILPs1 to cEILPs. Deletion of exon 2 of Tcf7 and the surrounding floxed

---

**Fig. 4 | Characterization of EILP-derived DC precursor populations.**

- **a.** Flow cytometric analysis of pre-DCs (gray), sEILPs (black) and cEILPs (orange).
- **b.** Flow cytometric analysis of Lin<sup>−</sup>C<sup>+</sup>-K<sup>+</sup>-Thy1<sup>−</sup>2B4<sup>+</sup>-α4<sup>+</sup>-TCF-1<sup>−</sup>-YFP<sup>+</sup>-EILPs defining sEILPs1 (green) and sEILPs2 (blue).
- **c.** Flow cytometric analysis of sEILPs and sEILP1s isolated as shown in b (left) and cultured for 2 d under SF7 conditions (right).
- **d.** Flow cytometric analysis of Lin<sup>−</sup>C<sup>+</sup>-K<sup>+</sup>-Thy1<sup>−</sup>2B4<sup>+</sup>-α4<sup>+</sup>-TCF-1<sup>−</sup>-YFP<sup>+</sup>-EILPs showing sEILPs1 (green), sEILPs2 (blue) and CD11c<sup>+</sup> EILPs (red). Arrows show successive gating.
- **e.** Flow cytometric analysis of the sEILP1s (green), sEILP2s (blue) and CD11c<sup>+</sup> EILPs (red) defined in d, compared with pre-DCs (gray).
- **f.** Flow cytometric analysis of the indicated populations in Tcf7<sup>−</sup>-Cre<sup>−</sup> mice (black histograms) compared with Tcf7<sup>−</sup>-Cre<sup>+</sup> control mice (gray histograms).
- **g.** Quantification by flow cytometry of the indicated populations in Tcf7<sup>−</sup>-Cre<sup>−</sup> mice and Tcf7<sup>−</sup>-Cre<sup>+</sup> littermates. Data are presented as means ± s.e.m. for n = 7 mice per group, pooled from three independent experiments. A two-tailed unpaired Student’s t-test was performed to determine significance (***P < 0.001; **P < 0.01; *P < 0.05).
region in Tcf7null mice generated mice with a Tcf7null allele (Supplementary Fig. 6a), which were crossed to Tcf7null mice to obtain Tcf7null mice. Tcf7null mice reported Tcf7 expression like Tcf7null mice, but lacked TCF-1 protein (Supplementary Fig. 6a–d). LinIlC– TCF7– GFP+ cells were detectable in Tcf7null BM (Supplementary Fig. 6e) and were δβ7+2B4+ (Supplementary Fig. 6f). Numbers of EILPs in Tcf7null mice were reduced twofold compared with wild-type mice (Fig. 5a), and Tcf7null EILPs were Flt3hi (Supplementary Fig. 6f), suggesting that ILC development was arrested at the EILP stage. In the absence of TCF-1 protein, we used TOX expression to quantify the number of EILPs in Tcf7null mice (Supplementary Fig. 6g,h), and used LinIlC-depleted BM cells, which allowed reliable detection of TOX by intracellular staining. Absolute numbers of LinIlC–Kit+ were similar in Tcf7null and Tcf7null⋅Kit+ BM (Supplementary Fig. 6h); thus, the frequency of LinIlC–Kit+ cells was unaffected, whereas the numbers of EILPs were greatly reduced in Tcf7null mice compared with TCF7– Kit+ mice (Fig. 5a). Both strategies showed that the numbers of cEILPs were not significantly affected, whereas the numbers of cEILPs were greatly reduced in Tcf7null mice compared with TCF7– cEILPs (Fig. 5a). The defect was similar when ILC progenitors developed in the presence of competitor wild-type cells in lethally irradiated CD45.1 mice to Tcf7null mice strain, and Supplementary Table 2 for additional RNA-Seq data.

**Fig. 5 | Developmental arrest at the sEILP stage in the absence of TCF-1.** a, Flow cytometric analysis of LinIlC–Kit+⋅2B4+δβ7+ BM cells from Tcf7null and Tcf7null⋅Kit+ littermate mice (left), and quantification of EILP numbers (right). b, Left, flow cytometric analysis of TOX+ EILPs defined in Supplementary Fig. 6g, in Tcf7null and Tcf7null littermate mice. sEILP and cEILP gates are shown in black and orange, respectively. Right, quantification of the frequencies of sEILPs and cEILPs within LinIlC–Kit+ BM cells. c, Flow cytometric analysis of LinIlC–Kit+⋅2B4+δβ7+ BM cells, which allowed reliable detection of TOX by intracellular staining. Absolute numbers of LinIlC–Kit+ were similar in Tcf7null and Tcf7null⋅Kit+ BM (Supplementary Fig. 6h); thus, the frequency of LinIlC–Kit+ cells was unaffected, whereas the numbers of EILPs were greatly reduced in Tcf7null mice compared with TCF7– Kit+ mice (Fig. 5a). Both strategies showed that the numbers of cEILPs were not significantly affected, whereas the numbers of cEILPs were greatly reduced in Tcf7null mice compared with TCF7– cEILPs (Fig. 5a). The defect was similar when ILC progenitors developed in the presence of competitor wild-type cells in lethally irradiated CD45.1 mice reconstituted with a mix of CD45.2+ LinIlC– BM cells from Tcf7null− or wild-type littermate controls, and CD45.1+ LinIlC– BM cells (Fig. 5d),
confirming that TCF-1 was required cell-autonomously for the generation of cEILPs.

TCF-1−GFPnull− EILPs were transcriptionally similar to wild-type sEILPs (Fig. 5e), consistent with developmental arrest of Tcf7−GFPnull− EILPs at the sEILP stage. Genes dynamically regulated between ALPs and cEILPs were similarly regulated between ALP and Tcf7−GFPnull− EILPs (Fig. 5f, Supplementary Fig. 6i and Supplementary Table 2), indicating that early ILC specification was largely unaffected by loss of TCF-1. Consistently, Tcf7−GFPnull− EILPs cultured for 4 d under SF7 conditions were still able to generate DCs (Fig. 5g,h). These observations indicate that TCF-1 was dispensable for initial ILC specification, but was required for progression towards the ILc lineage.

TCF-1-deficient EILPs are diverted towards the DC lineage. Although developmental arrest at the sEILP stage in TCF-1-deficient mice predicted a twofold higher frequency of DC-competent sEILPs in Tcf7−GFPnull− EILPs, Tcf7−GFPnull− EILPs cultured for 4 d under SF7 conditions gave rise to fewer DCs than wild-type EILPs (Fig. 5g,h). Expression of Batf3 and Id2, which was higher on sEILPs compared with sEILPs (Fig. 1f), was twofold higher on Tcf7−GFPnull− EILPs compared with wild-type EILPs (Fig. 5g,h). A two-sided Wilcoxon rank-sum test was used to determine the significance of gene expression differences between Tcf7−GFPnull− and Tcf7−GFPnull− cells for a given subset (***P < 0.005). See also Supplementary Fig. 7 and Supplementary Tables 3 and 4.

**Fig. 6** | Commitment failure and lineage diversion in the absence of TCF-1. **a**, Left, flow cytometric analysis of Lin−ILC−Kit+2B4−α4β7− Tcf7−GFP+Flt3+ sEILPs from Tcf7−GFPnull+ and Tcf7−GFPnull− littermate mice, showing sEILPs (green) and sEILPs (blue). Right, quantification of the percentages of sEILPs and sEILPs within Lin−ILC−Kit+ BM cells. Data are representative of three independent experiments and are presented as means ± s.e.m. for n = 3 mice per group. A two-tailed unpaired Student’s t-test was performed to determine the significance (*P < 0.05). **b,c**, scRNA-Seq analysis of Tcf7−GFP+ BM progenitors isolated from Tcf7−GFPnull− mice and compared with wild-type ALPs and Tcf7−GFP+ BM progenitors from Fig. 1. t-SNE plots showing: **b** ALP (n = 786 cells), Tcf7−GFPnull− (n = 1,799 cells) and Tcf7−GFPnull− (n = 594 cells) samples; and **c** the similarity of Tcf7−GFPnull− Tcf7−GFP+ BM progenitors (black) with the wild-type ALP and Tcf7−GFP+ BM progenitor subsets colored by cluster, as defined in Fig. 1c. **d**, Quantification of Tcf7−GFPnull− and Tcf7−GFPnull− cEILP numbers in each cluster from c, calculated as the percentage of sEILPs. The arrows show the developmental relationships linking clusters. Sizes of circles symbolizing each cluster are relative to cell numbers. **e**, Expression of the indicated genes by n = 752 ALPs (red), n = 270 Tcf7−GFPnull− sEILPs and n = 276 Tcf7−GFPnull− sEILPs (black), and n = 615 Tcf7−GFPnull− sEILPs and n = 187 Tcf7−GFPnull− sEILPs (pink) from scRNA-Seq analysis. A two-sided Wilcoxon rank-sum test was used to determine the significance of gene expression differences between Tcf7−GFPnull− and Tcf7−GFPnull− cells for a given subset (***P < 0.005). See also Supplementary Fig. 7 and Supplementary Tables 3 and 4.

NATURE IMMUNOLOGY | VOL 20 | SEPTEMBER 2019 | 1150–1160 | www.nature.com/natureimmunology

1157
Articles Nature Immunology

Tcf7<sup>EGFP<sup>+</sup></sup> Lin<sup>−</sup>Kit<sup>−</sup>ILC<sup>−</sup>Kit<sup>+</sup>2B4<sup>−</sup>α4β7<sup>−</sup> BM progenitors and ALP scRNA-Seq data (Fig. 6b). In t-distributed stochastic neighbor embedding (t-SNE) analysis, only ~5% Tcf7<sup>EGFP<sup>−</sup></sup> cells overlapped with wild-type cEILPs, ILCPs or ILC2Ps (Fig. 6c). Relative quantification of Tcf7<sup>EGFP<sup>−</sup></sup> and Tcf7<sup>EGFP<sup>+</sup></sup> cells in each cluster showed that 51% of Tcf7<sup>EGFP<sup>−</sup></sup>sEILPs corresponded to sEILP2, whereas only 23% of wild-type sEILPs were sEILP2s (Fig. 6d). The transcription factors Irf8, Id2, Nfil3, Spi1 and Mef2c were upregulated, and Zbtb16 was downregulated in Tcf7<sup>EGFP<sup>−</sup></sup> sEILP1s compared with wild-type sEILP1s (Fig. 6e and Supplementary Table 3). These altered expression patterns were most obvious for sEILP1s on the main progression (Supplementary Fig. 7c,d), suggesting that they preceded diversion towards the DC lineage (Supplementary Fig. 7e and Supplementary Table 4). These observations indicate that, in the absence of TCF-1, sEILP1s adopted a cDC1 progenitor-like transcriptional profile and were diverted towards the DC fate.

Fig. 7 | Identification of TCF-1 gene targets during early ILC development. a, Left, TCF-1 ChIC-Seq analysis in EILPs, and DNase-Seq analysis in ALPs, EILPs, ILCPs. The heat map is centered on TCF-1 binding sites in EILPs (±2 kb) (n = 9,549 peaks). Right, quantification of DNase I hypersensitivity enrichment. b, Known transcription factor motif analysis (HOMER) of the TCF-1 binding sites from a (n = 9,549 peaks). Chosen enriched motifs and associated P values were calculated based on the cumulative binomial distribution. c, RNA-Seq analysis averaged from n = 3 ALP samples, n = 2 sEILP samples, n = 3 cEILP samples and n = 2 ILCP samples, showing the expression of TCF-1 gene targets. d, Biological process enrichment analysis on genes upregulated (n = 31 genes; red), or downregulated by TCF-1 (n = 28 genes; blue). The most significant biological processes and associated P values were calculated based on the accumulative hypergeometric distribution. e, Traces of DNase-Seq in EILPs (black), IgG or TCF-1 ChIC-Seq in EILPs (red), and IgG or TCF-1 ChIP-Seq in thymocytes (blue) 19. Previously described enhancers are indicated by black boxes, along with their distance relative to the transcription start site 20,21. See also Supplementary Fig. 7 and Supplementary Tables 5–7.

TCF-1 is a homeodomain transcription factor that plays a crucial role in the development of early lymphoid progenitors. Its expression is essential for the differentiation of thymic T cells, but its role in other lymphoid lineages is less well understood. In this study, the authors investigated the role of TCF-1 in the development of early lymphoid progenitors (ILPs) and their differentiation into various lymphoid lineages. They used a combination of ChIP-seq, DNase-seq, and scRNA-seq to analyze the transcriptional landscape of TCF-1 in early ILPs. They found that TCF-1 binds to specific motifs in the promoters of genes involved in lymphoid development, and that its expression is enriched in early ILPs, particularly in sEILPs. These findings suggest that TCF-1 plays a key role in the specification of early lymphoid progenitors and their commitment to specific lineages.

The authors also used scRNA-seq to identify the transcriptional profiles of different ILP subsets and their relationship to the DC lineage. They found that Tcf7<sup>EGFP<sup>−</sup></sup> sEILPs resembled wild-type sEILP2s, whereas Tcf7<sup>EGFP<sup>+</sup></sup> sEILPs showed a more diverse transcriptional profile. These findings suggest that the absence of TCF-1 leads to a more promiscuous developmental fate, with sEILP1s adopting a cDC1 progenitor-like transcriptional profile.

Overall, these results provide a detailed understanding of the role of TCF-1 in the development of early lymphoid progenitors and their commitment to specific lineages. They also highlight the importance of TCF-1 in the regulation of gene expression in early lymphoid progenitors and the potential for its therapeutic manipulation in the treatment of lymphoid disorders.
TCF-1 enhances expression of ILC genes and represses DC genes. To identify TCF-1 gene targets during ILC development, we first inferred regulatory interactions between genes using the lag-based expression association for pseudotime series (LEAP) algorithm\(^1\). We used the main progression to identify genes whose expression was correlated during ILC development, allowing a gap in correlation to take into account delays between the expression of a controller and its targets (Supplementary Fig. 7f and Supplementary Table 5). To determine which of the putative TCF-1 target genes identified were directly regulated by TCF-1, we characterized TCF-1 binding in EILPs using chromatin immunocleavage sequencing (ChiC-Seq)\(^3\), and identified open chromatin regions in ALPs, EILPs and ILCPs using DNase I hypersensitive sites sequencing (DNase-Seq). We identified 9,649 TCF-1 peaks genome wide in EILPs (Supplementary Table 6). Of these, 99% were located in regions of open chromatin in ALPs (Fig. 7a and Supplementary Fig. 7g). Regions bound by TCF-1 showed significant enrichment for TCF motifs, as well as runt-related transcription factor (RUNX) and PU.1 motifs (Fig. 7b).

To identify TCF-1 target genes, we identified genes predicted to be regulated by TCF-1 (Supplementary Table 5), and showed TCF-1 binding in open chromatin regions in their vicinity (Supplementary Table 6). We excluded genes that were properly regulated in the absence of TCF-1 in scRNA-Seq or bulk RNA-Seq (Supplementary Tables 2 and 3). This analysis identified 59 genes that received direct regulatory inputs from TCF-1 in EILPs (Fig. 7c and Supplementary Table 7). Downregulated genes were enriched for genes involved in myeloid cell differentiation (Fig. 7d), some of which (Flt3, Nfil3, Mef2c, Spi1 and Ifi38) were upregulated in Tcf7\(^{−/−}\) or sEILP1 cells in the main trajectory compared with wild-type sEILP1s (Supplementary Fig. 7h and Supplementary Table 4), and were downregulated from sEILP1s to eILPs in wild-type ILC progenitors (Fig. 11). In contrast, genes upregulated by TCF-1 were enriched for genes involved in lymphocyte differentiation and activation (Fig. 7d), such as Id2 or Gata3, which play important functions in EILPs\(^3\) (Fig. 7c), indicating that TCF-1 promoted ILC differentiation. Genes important at later stages of development and ILC maturation, such as Il17r, Ets1 or Pdcd1 (encoding PD-1) were also predicted TCF-1 targets (Fig. 7d). A few genes predicted to be positively regulated by TCF-1 (Zbtb16, Cd5 and Rora) were upregulated between ALPs and sEILP1s (Supplementary Table 1) and had lower expression in Tcf7\(^{−/−}\) sEILP1s compared with wild-type sEILP1s (Supplementary Fig. 7e,h).

Comparison of TCF-1 binding in EILPs and T lineage cells\(^19\) revealed that several TCF-1 binding regions were shared, including the downstream Gata3 enhancer important for early T cell development (Tce1)\(^29\) and the Pdcd1 enhancer active in exhausted T cells\(^21\) (Fig. 7e). TCF-1 contributes to the regulation of these enhancers in T cells\(^20,21\), indicating that TCF-1 can regulate similar loci in T cells and ILCs. Additional genes expressed in both T cells and ILCs, such as Id2, Socs1, Sla2 and Tsk had similar patterns of TCF-1 binding in both lineages, whereas other genes such as Ifi8, Lmo2 and Flt3 showed distinct patterns (Fig. 7e). Our analysis indicated that TCF-1 promoted ILC development by positively regulating ILC-specific genes, and enforced ILC commitment through the repression of genes shared by sEILP1s and DCs (Supplementary Fig. 7i).

Discussion

In this study, we used population-level and minimally biased single-cell approaches to identify transcriptional and functional heterogeneity of ILC progenitor cells, and to understand developmental relationships between them. Our work confirmed a linear developmental path between ALPs, EILPs, ILCPs and ILC2Ps\(^8\). We identified two successive stages of ILC development within EILPs, which we called sEILP1 and cEILP, and delineated a commitment checkpoint during early ILC development. We also identified a developmental bifurcation towards the DC fate at the sEILP1 stage at steady state in vivo. We further defined a role for the transcription factor TCF-1 in controlling these developmental transitions. We found that TCF-1 was dispensable for initial ILC specification and generation of sEILP1s, but required for the development of cEILPs. Mechanistically, TCF-1 enforced commitment to the ILC lineage by providing positive regulatory inputs to key ILC genes, and by repressing the expression of genes key for DC development that were expressed at the early steps of ILC specification and down-regulated during ILC commitment.

Many genes bound and upregulated by TCF-1 during early ILC development also have important functions in T cells, which raises the possibility that TCF-1 regulates these genes in both lineages. Consistently, key enhancers described to receive regulatory inputs from TCF-1 in T cells, including the distal Gata3 and Pdcd1 enhancer, were bound by TCF-1 in ILC progenitors. We found that chromatin bound by TCF-1 in EILPs was almost entirely accessible in upstream ALPs. This finding indicates that TCF-1 binding is largely dictated by the regulatory landscape established before T cell and ILC specification, in precursors of these lineages. TCF-1 binding sites showed significant enrichment for binding motifs for RUNX and PU.1 transcription factors. Hence, RUNX1 and PU.1 may mediate chromatin opening to allow TCF-1 binding\(^22,23\), or these factors may serve to recruit TCF-1 to its binding sites during T cell and ILC specification\(^11\). Our findings raise the possibility that TCF-1 regulates a shared set of core genes in T cells and ILCs, which coordinates the acquisition of functional similarities between the two lineages.

Although the observation that ILC progenitors had the ability to develop into DCs was similar to DC lineage potentials reported for other uncommitted lymphoid progenitors, such as early T cell precursors\(^4\) and ALPs\(^8\), our study provides evidence that some lymphoid precursors actually fulfilled this potential in vivo at steady state, and that DC development occurred from non-canonical DC precursors. Because sEILPs and other lymphoid precursors do not appear to contribute significantly to DC generation, nor do they generate distinct lineages of DCs\(^24\), such a pathway may not be important in physiological conditions. We speculate that the residual DC development described here is the result of the imprecision of molecular mechanisms that underly early lymphoid specification, and a consequence of the transcriptional similarities between EILPs and DC progenitors\(^20\). More generally, because of the complexity of specification mechanisms, other developmental processes may be similarly prone to imprecisions\(^20,20\).

In summary, our work revealed that EILPs are heterogeneous, and that innate cell lineage specification before commitment occurred with a degree of imprecision and co-opted transcriptional programs permissive for DC development. Our work revealed that a residual DC development accompanied early ILC specification in vivo, and identified mechanisms by which TCF-1 restrained this alternative potential and mediated commitment to the ILC lineage.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at https://doi.org/10.1038/s41590-019-0445-7.

Received: 8 August 2018; Accepted: 12 June 2019; Published online: 29 July 2019

References

1. Harly, C., Cam, M., Kaye, J. & Bhandoola, A. Development and differentiation of early innate lymphoid progenitors. J. Exp. Med. 215, 249–262 (2018).
2. Jeevan-Raj, B. et al. The transcription factor Tcfl contributes to normal NK cell development and function by limiting the expression of granzymes. Cell Rep. 3, 613–626 (2017).

1. Harly, C., Cam, M., Kaye, J. & Bhandoola, A. Development and differentiation of early innate lymphoid progenitors. J. Exp. Med. 215, 249–262 (2018).
2. Jeevan-Raj, B. et al. The transcription factor Tcfl contributes to normal NK cell development and function by limiting the expression of granzymes. Cell Rep. 3, 613–626 (2017).
3. Weber, B. N. et al. A critical role for TCF-1 in T-lineage specification and differentiation. *Nature* **476**, 63–68 (2011).
4. Yang, Q. et al. TCF-1 upregulation identifies early innate lymphoid progenitors in the bone marrow. *Nat. Immunol.* **16**, 1044–1050 (2015).
5. 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).
6. Klose, C. S. et al. Differentiation of type 1 ILCs from a common progenitor to all helper-innate lymphoid cell lineages. *Cell* **157**, 340–356 (2014).
7. Constantinides, M. G., McDonald, B. D., Verhoef, P. A. & Bendelac, A. A committed precursor to innate lymphoid cells. *Nature* **508**, 397–401 (2014).
8. Yu, Y. et al. Single-cell RNA-Seq identifies a PD-1³ ILC progenitor and defines its development pathway. *Nature* **539**, 102–106 (2016).
9. Ji, Z. & Ji, H. TSCAN: pseudo-time reconstruction and evaluation in single-cell RNA-Seq analysis. *Nucleic Acids Res.* **44**, e117 (2016).
10. Murphy, T. L. et al. Transcriptional control of dendritic cell development. *Annu. Rev. Immunol.* **34**, 93–119 (2016).
11. Grajales-Reyes, G. E. et al. Batf3 maintains autoactivation of Ifi30 for commitment of a CD8α+ conventional DC clonogenic progenitor. *Nat. Immunol.* **16**, 708–717 (2015).
12. Schlützer, 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).
13. Heng, T. S., Painter, M. W. and The Immunological Genome Project Consortium. The Immunological Genome Project: networks of gene expression in immune cells. *Nat. Immunol.* **9**, 1091–1094 (2008).
14. Schellen, S. M. et al. Fate mapping reveals separate origins of T cells and myeloid lineages in the thymus. *Immunity* **32**, 426–436 (2010).
15. Bell, J. J. & Bhandoola, A. The earliest thymic progenitors for T cells possess myeloid lineage potential. *Nature* **452**, 764–767 (2008).
16. Morrison, S. J., Wandyzc, A. M., Hemmati, H. D., Wright, D. E. & Weissman, I. L. Identification of a lineage of multipotent hematopoietic progenitors. *Development* **124**, 1929–1939 (1997).
17. Specht, A. T. & Li, J. LEAP: constructing gene co-expression networks for single-cell RNA-sequencing data using pseudotime ordering. *Bioinformatics* **33**, 764–766 (2017).
18. Ku, W. L. et al. Single-cell chromatin immunonecavage sequencing (scChIC-seq) to profile histone modification. *Nat. Methods* **16**, 323–325 (2019).
19. Li, Y. et al. A far downstream enhancer for murine Bcl11b controls its T-cell specific expression. *Blood* **122**, 902–911 (2013).
20. Ohmura, S. et al. Lineage-affiliated transcription factors bind the Gata3 Tc1 enhancer to mediate lineage-specific programs. *J. Clin. Invest.* **126**, 865–878 (2016).
21. Sen, D. R. et al. The epigenetic landscape of T cell exhaustion. *Science* **354**, 1165–1169 (2016).
22. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
23. Wang, D. et al. The transcription factor Runx3 establishes chromatin accessibility of cis-regulatory landscapes that drive memory cytotoxic T lymphocyte formation. *Immunity* **48**, 659–674.e6 (2018).
24. Steinke, F. C. et al. TCF-1 and LEF-1 act upstream of Th-POK to promote the CD4⁺ T cell fate and interact with Runx3 to silence Cd4 in CD8⁺ T cells. *Nat. Immunol.* **15**, 646–656 (2014).
25. Feyerabend, T. B. et al. Deletion of Notch1 converts pro-T cells to dendritic cells and promotes thymic B cells by cell-extrinsic and cell-intrinsic mechanisms. *Immunity* **30**, 67–79 (2009).
26. Wu, L. et al. Development of thymic and splenic dendritic cell populations from different hemopoietic precursors. *Blood* **98**, 3376–3382 (2001).
27. Shigematsu, H. et al. Plasmacytoid dendritic cells activate lymphoid-specific genetic programs irrespective of their cellular origin. *Immunity* **21**, 43–53 (2004).
28. Ishikawa, F. et al. The developmental program of human dendritic cells is operated independently of conventional myeloid and lymphoid pathways. *Blood* **110**, 3591–3660 (2007).
29. Rothenberg, E. V., Kueh, H. Y., Yui, M. A. & Zhang, J. A. Hematopoiesis and T-cell specification as a model developmental system. *Immunol. Rev.* **271**, 72–97 (2016).
30. Farrell, J. A. et al. Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science* **360**, eaar3131 (2018).

**Acknowledgements**
We thank H.-R. Rodewald for sharing Il7r-iCre mice and J. Richa from the Transgenic and Chimeric Mouse Facility, University of Pennsylvania, for injection of Tg(Prdm1)
embryonic stem cells into mouse blastocysts. We thank T. Giucci, R. Bosselut, J. Chen, the CCR Sequencing Facility, the CCR Flow Cytometry Core Facility, and the DNA Sequencing Facility of the University of Pennsylvania for technical support. This work utilized the computational resources of the NIH high-performance computing Biowulf cluster (http://bpc.nih.gov). This research was supported by the Intramural Research Program of the NIH, National Cancer Institute and Center for Cancer Research, and by grants from the NIH (AI121080 and AI139874 to H.-H.X.), Veteran Affairs BLR&D Merit Review Program (BX002903A to H.-H.X.), and Foundation pour la Recherche Médicale (DEQ20170839118 to C.H.) and National Research Agency Investissements d'Avenir via the program LabEX IGO (ANR-11-LABX-0016-01 to C.H.).

**Author contributions**
C.H. designed the research and performed most of the experiments, alongside D.K. and G.R. H.C., B.L., M.C.C. and A.B. analyzed the data. C.H., M.C.C., T.R. and A.B. produced the figures. C.H., T.R., Q.Y. and H.-H.X. designed and generated the new mouse models. C.H., K.Z. and A.B. directed and oversaw the experiments. C.H. and A.B. wrote the paper. All authors helped to design the research, and read and commented on the manuscript.

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

**Additional information**
Supplementary information is available for this paper at https://doi.org/10.1038/s41590-019-0445-7.
Reprints and permissions information is available at www.nature.com/reprints.
Correspondence and requests for materials should be addressed to A.B.
Peer review information: Ioana Visan was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

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

Mice. B6-Ly5.2 (CD45.1) and CMV-Cre mice were obtained from the Jackson Laboratory. Tcf7loxP2AEGFPnull mice were generated as described. Tcf7loxP2AEGFPnull mice were used as experimental samples. As LinILC constructs were generated as described, the targeting vector was generated by bacterial artificial chromosome (BAC) recombineering.

To begin with, a 1,538-nucleotide-long mouse genomic Tcf7 fragment starting with ATGCTATATGTTGGCCCTTCCT... and ending with CAGGCTTGAAAGCTGGGTGT was inserted into the BAC recombinatorial vector PL253 (NCI at Frederick) by BAC recombineering. Next, a synthetic DNA construct was designed. This started with ACAAGAAGCTTCTCCTTG... of Tcf7 intron 9, followed by the beginning of exon 10, which was fused at the carboxy terminus of its ...PLPMTVL amino acid sequence to the P2A ribosomal skipping peptide. This was followed by the YFP sequence and the rest of Tcf7 exon 10, starting at TAGCTCTCTGCGGCTCGCAAGC... and the synthetic construct ended with the RCAGAAGCTTCTCCTTG... of Tcf7 intron 9, which was fused to the P2A ribosomal skipping peptide inserted after the ...PLPMTVL sequence (a total of 310 nucleotides), the YFP sequence (720 nucleotides), the additional downstream part of exon 10 (1,159 nucleotides) and an additional 4,039-nucleotide intronic sequence (the total downstream arm is 5,198 nucleotides). C57Bl6 ES cells (strain EAP6; developed by T.R., R. Richa, K. Kaestner and E. Pierce as members of the Penn Knockout Mouse Project) were electroporated with this vector and isolated by G418 (geneticin), and 196 resulting embryonic stem cell clones were screened by Southern blotting, resulting in several positive clones. After ascertaining the correct karyotype by chromosome counting, independent positive embryonic stem cell clones were injected into C57/BL6 blastocysts at the Penn Transgenic and Chimeric Mouse Facility; yielding 100% C57BL6 background male chimeras, which were bred to 86 females. Tcf7loxP2AEGFPnull mice are described in Supplementary Fig. 3. Tcf7loxP2AEGFPnull mice were generated by breeding the Tcf7loxP2AEGFPnull mice with CMV-Cre mice. Antibodies and flow cytometry. BM cell suspensions were incubated with a mix of purified rat, mouse and hamster immunoglobulin G (IgG) before the addition of specific antibodies. Antibodies specific for Ly-6D (49H4), B220 (RA3-6B3), CD19 (1D3), Mac-1 (M170), Gr-1 (8C5), CD11c (N418), Ter119 (1A8), NK1.1 (PK136), Zbtb16 (CD4), CD45.1 (A20), CD45.2 (205yekta), Kit (2B8), Sca-1 (D7), Thy1.2 (53–2.1), αβ7 (DATK32), IL-7Rβ (53–6.72), CD8α (53–6.72), and CD8β were added for CSF7-GM3-MG6 cytokine conditions. These antibodies were added for SF7-GM3-MG6 cytokine conditions. The Tcf7loxP2AEGFPnull mice were separated using NRP-1 and MHC-II expression on sEILPs, as shown in Fig. 3b. Alternatively, sEILPs could be defined, showing the following: α±, β±, Tcf7+±, with a pre-cDC1s and Kit−Ly-6C−Ly-6C− pre-cDC2s (Supplementary Fig. 4). The LMPPs were defined by staining with Gr-1, Ter119, NK1.1, CD3e, CD11c and CD3e, CD11c and CD3e. All cytokines were purchased from PeproTech. CD45.2+ cells were considered for analysis of hematopoietic progeny.

scRNA-Seq and analysis. B cells isolated from spleen were added to each of the scRNA-Seq samples as an internal control. scRNA-Seq libraries were prepared using Chromium Single Cell 3' kits, according to the manufacturer’s instructions (version 2 chemistry; 10x Genomics). The obtained libraries were sequenced with a NextSeq system (version 2 chemistry; Illumina). Primary analysis was performed with Cell Ranger version 2.0.1 software using the default parameters. Median numbers of unique molecular identifications (UMIs) were 44,908 ± 4,073 (95% confidence interval) for 2,894 and 5,309 per cell. Cells with low UMI counts were determined by the 10x Genomics Cell Ranger Algorithm (https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/algorithms/overview). After estimating the maximum total UMI count (or m) as the 99th percentile of the top 3,000 barcodes, those whose UMI count exceeded m/10 were called cells.

Seurat’s FindClusters function was used to identify the main progenitor populations on t-SNE. B cells and contaminant cells were further separated in Supplementary Fig. 6. The mice used were 6–10 weeks old and are described in Supplementary Fig. 6. The data were normalized and scaled, and cell cycle scores and mitochondrial percentages were regressed using the Seurat package. Seurat’s PCAFast function was run and the first 13 principal components were used in subsequent analysis. Cells were clustered using Seurat’s FindClusters function. To visualize the data, t-SNE plots were generated using Seurat’s RunTSNE function. B cells spiked in all samples formed their own cluster and overlapped on the t-SNE plot, thus confirming the validity of the normalization across samples. Small contaminant populations were identified on the basis of signature gene expression of mature populations from BM. Each clustered cells formed their own clusters, and appeared distant from the main progenitor populations on t-SNE. B cells and contaminant cells that represented 5% of the data were electronically removed before subsequent analysis. Seurat’s FindClusters function was used to cluster the remaining data using default perplexity. A range of resolutions were tested, and a resolution of 0.65 was selected to capture known ILC progenitor groups and new subgroups without adding subdivisions within relatively homogeneous ALPs. Differentially expressed genes between clusters were identified using Seurat’s FindMarkers function. Only genes showing expression in at least 10% of cells in one of the groups compared were considered. A two-sided Wilcoxon rank-sum test was used to determine significance. Differentially significant clusters between cells are described in Supplementary Table 1. Pseudotime reconstruction analysis was performed on the first five principal components using the TSCAN package. Violin plots showing Seurat’s scores were made using Seurat’s t-SNE and cell cycle scores. Correlation network analysis along the scRNA-Seq pseudotime was performed using the LEAP package (Applie et al., 2014) and the top 500 variable genes expressed by at least 10% of the cells (1,094 genes) was lag of one-third of the pseudotime was allowed. A lag of one-third of the pseudotime was allowed was used to determine significance. Correlations of ≥0.2 were considered for further analysis (false discovery rate: <5 × 10−5) and are shown in Supplementary Table 5. Network visualization was done using Cytocepte (Supplementary Fig. 7). Additional analysis and visualization was done using R. Bulk RNA-Seq and analysis. RNA was extracted using an RNeasy Plus Micro Kit (Qiagen) according to the manufacturer’s instructions. Quality control was performed by bioanalyzer (Agilent) and RNA samples with a RNA Integrity number of ≥9 were subsequently used. Messenger RNA-Seq libraries were prepared using a SMARTer Ultra Low Input RNA Kit version 3 (Clontech) and Nextera XT DNA Library Preparation Kit (Illumina). Pair-end sequence reads of 128 base pairs were generated with a HiSeq 2500 sequencer (Illumina). The raw RNA-Seq FASTQ reads were aligned to the mouse genome (mm10) using STAR (2.5.2b) on two-pass mode with mouse genome (release M1) gff. GFF3 files were subsequently counted using Rsubread and further analyzed for gene expression changes and statistics using limma-voom with quantile normalization and batch correction using Combat. The gene and sample-specific normalization factors were then used to correct counts and to generate bigwig files. Visualization was done using R. DNase-Seq. ChIC-seq and analysis. DNase-Seq. DNase-Seq assays were performed as described. Briefly, 300 cells from each cell type were collected by fluorescence-activated cell sorting. Then, 0.3 μm DNase I (04716728001, Roche) was added to each cell type and incubated at 37°C for 5 min. Reactions were stopped by
adding 80 μl of stop buffer (10 mM Tris–HCl (pH 7.5), 10 mM NaCl, 0.15% SDS and 10 mM EDTA) containing 1 μl of 20 mg ml⁻¹ protease K. Samples were incubated at 55 °C overnight, and DNA was purified by phenol–chloroform extraction followed by precipitation with ethanol in the presence of 20 μg glycogen. DNA was further processed for library production.

Small cell number ChIC-Seq. ChIC-Seq was performed as described⁴. Cells were fixed by adding a 1:15 volume of 16% w/v formaldehyde solution (Thermo Fisher Scientific) and incubating at room temperature for 10 min. The reaction was terminated by adding a 1:10 volume of 1.25 M glycine and incubating on ice for 5 min. The fixed cells were washed with ice-cold phosphate buffered saline twice, and collected by centrifugation. Then, 1 ml RIPA buffer (10 mM Tris–Cl, 1 mM EDTA, 0.1% sodium deoxycholate, 0.2% sodium dodecyl sulfate and 1% Triton X-100) was added to each of the sorted 500 ALP, EILP and ILCP cells, and incubated at room temperature for 30 min. Cells were rinsed with 500 μl binding buffer (10 mM Tris–Cl, 1 mM EDTA, 150 mM sodium chloride and 0.1% Triton X-100) twice and resuspended in 50 μl binding buffer. To prepare anti-TCF-1-bound protein A–micrococcal nuclease (pA-MNase; antibody+PA-MNase), anti-TCF-1 (C46C7; Cell Signaling) and the PA-MNase at a molecular ratio of 1:2 were pre-incubated at 4 °C for 30 min in 50 μl binding buffer. TCF-1 + PA-MNase were added to 50 μl binding buffer-resuspended cells and incubated for 1 h at 4 °C with rotation. Cells were washed using 200 μl wash buffer (10 mM Tris–Cl, 1 mM EDTA, 150 mM sodium chloride, 0.1% w/v SDS, 0.1% w/v sodium deoxycholate and 1% w/v Triton X-100) three times and pelleted by centrifugation at 600g for 2 min. Next, cells were rinsed using 200 μl rinsing buffer (10 mM Tris–Cl, 10 mM sodium chloride and 0.1% w/v Triton X-100). The MNase digestion was initiated by resuspending the rinsed cells in 40 μl reaction solution buffer (10 mM Tris–Cl, 10 mM sodium chloride, 0.1% w/v Triton X-100 and 2 mM CaCl₂) and incubating at 37 °C for 3 min. The reaction was stopped by adding 80 μl stop buffer (20 mM Tris–Cl (pH 8.0), 10 mM ethylenedioxo-bis-(ethylenenitrilo)-tetraacetic acid, 20 mM sodium chloride and 0.2% w/v SDS) and 1 μl protease K (Sigma–Aldrich), then incubating at 65 °C overnight. DNA was purified using phenol–chloroform extraction and ethanol precipitation. The purified DNA was further processed for library production.

Library preparation. Libraries were prepared according to Illumina’s instructions. Briefly, DNA was end-repaired using a combination of T4 DNA polymerase, E. coli DNA Pol I large fragment and T4 polyethylene glycolase kinase. The blunt, phosphorylated ends were treated with Klenow fragment (3′ to 5′ exo minus) and dATP to yield a protruding 3′- A base for ligation of Illumina’s adapters, which have a single ‘T’ base overhang at the 3′ end. The ligation reaction was performed by adding 2 μM Illumina’s adapters to each sample and incubating at room temperature for 1 h. After adapter ligation, DNA was PCR amplified with Illumina primers for 16 cycles, and library fragments from 180–300 base pairs were isolated from an agarose gel. The purified DNA was captured on an Illumina flow cell for cluster generation. Libraries were sequenced on a Hi-Seq 3000 following the manufacturer’s protocols.

Analysis. Sequencing reads were aligned against the mm9 reference genome using Bowtie 2 (ref. ³) with default parameters. Reads from technical replicates and biological replicates were combined for peak calling. Duplicated reads were removed from further analysis. TCF-1 peaks were called using MACS2 (ref. ⁴) with default parameters. Motif discovery (Fig. 7h), peak file annotation (Supplementary Table 6) and additional analysis were done with HOMER (version 4.10.1)⁵. Biological process enrichment and visualization was done with Metascape (Fig. 7d)⁶. Additional visualization was done using R⁷.

Statistics. Statistical analysis was performed on groups with limited variance using Excel or Prism. Differences between groups of mice or wells were determined by two-tailed unpaired Student’s t-test. A Welch correction was applied for groups of unequal s.d. P < 0.05 was considered significant. Sample sizes were empirically determined. No samples or animals were excluded from the analysis, and no randomization or blinding was used.

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

Data availability
The accession number for the raw data of the RNA-Seq is GSE113767. The accession number for the raw data of the DNase-Seq and ChIC-Seq is GSE128483. All other relevant data are available from the corresponding author on request.

References
31. Verbeek, S. et al. An HMG-box-containing T-cell factor required for thymocyte differentiation. Nature 374, 70–74 (1995).
32. Aliiahmad, P. & Kaye, J. Development of all CD4 T lineages requires nuclear factor TOX. J. Exp. Med. 205, 254–258 (2008).
33. 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).
34. Onai, N. et al. Identification of clonogenic common Flt3-M-CSFR⁺ plasmacytoid and conventional dendritic cell progenitors in mouse bone marrow. Nat. Immunol. 8, 1207–1216 (2007).
35. Satija R., Butler A. & Hoffman P. Seurat: Tools for single cell genomics. R package version 2.3.0 https://cran.r-project.org/web/packages/Seurat/index.html (2018).
36. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13, 2498–2504 (2003).
37. R Development Core Team R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, 2014).
38. Dobin, A. et al. STAR: ultrafast universal RNA-Seq aligner. Bioinformatics 29, 15–21 (2013).
39. Liao, Y., Smyth, G. K. & Shi, W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-extend. Nucleic Acids Res. 41, e108 (2013).
40. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 43, e47 (2015).
41. Leek, J. T. et al. sva: Surrogate variable analysis. R package version 3.26.0 (2017).
42. Cooper, J., Ding, Y., Song, J. & Zhao, K. Genome-wide mapping of DNase I hypersensitive sites in rare cell populations using single-cell DNase sequencing. Nat. Protoc. 12, 2342–2354 (2017).
43. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359 (2012).
44. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 9, R137 (2008).
45. Tripathi, S. et al. Meta- and orthogonal integration of influenza "OMICs" data defines a role for UBR4 in virus budding. Cell Host Microbe 18, 723–735 (2015).
Reporting Summary

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

Statistics

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

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

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

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

Data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

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

RNA-seq data are deposited under the public GEO accession number GSE113767. ChiC-seq and DNase-seq data are deposited under the public GEO accession number GSE128483. All other data are available in the manuscript.
Field-specific reporting

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

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

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

Life sciences study design

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

- Sample size: Sample sizes were empirically determined
- Data exclusions: No samples or animals were excluded from the analysis
- Replication: Experiments were independently repeated to ensure the reproducibility of the findings. Every finding was validated using independent approaches to remove possible experimental bias
- Randomization: No randomization was used
- Blinding: No blinding was used

Reporting for specific materials, systems and methods

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

| Materials & experimental systems | Methods |
|----------------------------------|---------|
| n/a involved in the study        | n/a involved in the study |
| Antibodies                       | ChIP-seq |
| Eukaryotic cell lines            | Flow cytometry |
| Palaeontology                    | MRI-based neuroimaging |
| Animals and other organisms      |         |
| Human research participants      |         |
| Clinical data                    |         |

Antibodies

All antibodies used are commercially available. Antibodies specific for Ly-6D (49H4), B220 (RA3-6B3), CD19 (1D3), Mac-1 (M1/70), Gr-1 (8C-5), CD11c (N418), Ter119 (TER119), NK1.1 (PK136), CD3e (2C11), CD8a (53-6.72), CD8b (H35-17.2), CD4 (GK1.5), TCRa (H5/5), TCRgd (GL-3), MHC-II (MB2/11.14.5-2), CD205 (205yekta), Kt (288), Sca-1 (D7), Thy-1.2 (53-2-1), 4b7 (DART12), IL-7Ra (7R34), ICOS (C398.4A), CD8 (eBio244F4), CD24 (M1/69), CD172a (2F1), Ly-6C (HK1.4), Siglec-H (eBio440c), CD25 (PC61.5), CD45.1 (A20), CD45.2 (104), TOX (TXR10), PLZF (Mags.1F7), GATA-3 (TWAJ), Nfil3 (S2M-E19), IRF-8 (V3GYWCH) were from eBioscience, anti-NRP-1 (3E12), CD122 (TM-b1), and Pu-1 (7C2C34) were from Biolegend, anti-Flt3 (A2F10) was from BD, anti-CCR2 (475301) was from R&D, and anti-TCF-1 (C63D9 and C46C7) were from Cell Signaling.

Validation

TCF-1 antibodies (C63D9 and C46C7) are validated for flow cytometry on mouse cells by the provider Cell Signaling. These two antibodies were specifically validated for visualization of mouse EILP by flow cytometry in our previous study (Figure S1C-D, Harly et al., 2018, JEM) and on thymocytes in this study (Figure S6H).

Animals and other organisms

Policy information about studies involving animals: ARRIVE guidelines recommended for reporting animal research.

| Laboratory animals | Mice used were on the C57 background, 6-10 weeks old and of either sex. All strains used (CD45.1, CMV-Cre, Tcf7EGFP, Zbtb16GFP, Tcf7l2−/−, Tox−/−, Il7r-iCre R26-stop-YFP, Tcf7YFP, Tcf7EGFPRnull) are referenced in the material and method section of the manuscript. |
|---------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Wild animals        | The study does not involve wild animals.                                                                                                                                                           |
| Field-collected samples | the study does not involve samples collected from the field.                                                                                                                                         |
Ethics oversight: Animal procedures were approved by the NIH Animal Care and Use Committee (ACUCC).

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

**ChIP-seq**

**Data deposition**

- Confirm that both raw and final processed data have been deposited in a public database such as GEO.
- Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

**Data access links**

For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.

**Files in database submission**

Provide a list of all files available in the database submission.

**Genome browser session**

Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.

**Methodology**

**Replicates**

3-4 reciprocates were sequenced for each type of sample, as described in the GEO submission.

**Sequencing depth**

Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.

**Antibodies**

anti-TCF-1 (#46C7, Cell Signaling)

**Peak calling parameters**

Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.

**Data quality**

Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.

**Software**

Sequencing reads were aligned against mm9 reference genome using Bowtie 2 with default parameters. Reads from technical replicates and biological replicates were combined for peak calling. Duplicated reads were removed from further analysis. TCF1 peaks were called using MACS2 with default parameters. Motif discovery, peak file annotation and additional analysis was done with HOMER (version 4.10.1).

**Flow Cytometry**

**Plots**

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a ‘group’ is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

**Methodology**

**Sample preparation**

BM single cell suspension were obtained by flushing and pipetting, and red cells were lysed before staining.

**Instrument**

sorts were performed using a FACS-Aria II (BD Bioscience), and analysis was done with a FORTESSA (BD Bioscience).

**Software**

data were acquired using Diva (v8.0.1) and analyzed in Flowjo (v9.9.6 for all experiments except index sort that was analyzed with v10.4.2).

**Cell population abundance**

post-sort sample purity was examined by flow cytometry. samples containing at least 95% of desired population were used for further analysis.

**Gating strategy**

all plots are gated as singlets (using both SSC-H/SSC-W and FSC-H/FSC-W).
a live dead gate is then applied, and FSC/SSC.
a representative gating strategy is provided in Figure S1.
all additional gating strategies are described in the section "BM progenitor definition, isolation and culture" of the material and method, referencing figures for all the gating strategies used. additional details are provided in the figure legends.

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