



# Multiscale 3D genome organization underlies ILC2 ontogenesis and allergic airway inflammation

Michaël F. Michieletto<sup>1,2,8</sup>, John J. Tello-Cajiao<sup>1,2,3,4,8</sup>, Walter K. Mowel<sup>1,2</sup>, Aditi Chandra<sup>2,3,4</sup>, Sora Yoon<sup>2,3,4</sup>, Leonel Joannas<sup>1,2</sup>, Megan L. Clark<sup>1,2</sup>, Monica T. Jimenez<sup>1,2</sup>, Jasmine M. Wright<sup>1,2</sup>, Patrick Lundgren<sup>1,2,5</sup>, Adam Williams<sup>6</sup>, Christoph A. Thaiss<sup>1,2,5</sup>, Golnaz Vahedi<sup>1,2,3,4</sup>✉ and Jorge Henao-Mejia<sup>1,2,7</sup>✉

**Innate lymphoid cells (ILCs) are well-characterized immune cells that play key roles in host defense and tissue homeostasis. Yet, how the three-dimensional (3D) genome organization underlies the development and functions of ILCs is unknown. Herein, we carried out an integrative analysis of the 3D genome structure, chromatin accessibility and gene expression in mature ILCs. Our results revealed that the local 3D configuration of the genome is rewired specifically at loci associated with ILC biology to promote their development and functional differentiation. Importantly, we demonstrated that the ontogenesis of ILC2s and the progression of allergic airway inflammation are determined by a unique local 3D configuration of the region containing the ILC-lineage-defining factor *Id2*, which is characterized by multiple interactions between the *Id2* promoter and distal regulatory elements bound by the transcription factors GATA-3 and ROR $\alpha$ , unveiling the mechanism whereby the *Id2* expression is specifically controlled in group 2 ILCs.**

ILCs are effector innate cell populations that contribute to host defense against pathogens, metabolic homeostasis and tissue repair, but also play key roles in diseases such as cancer, asthma and colitis<sup>1</sup>. ILCs are known to be the innate counterpart of T lymphocytes and have been classified into three groups based on the transcription factors and cytokines they express<sup>2</sup>. Group 1 comprises natural killer (NK) cells and ILC1s, which are considered the innate counterparts of cytotoxic CD8<sup>+</sup> T cells and CD4<sup>+</sup> T helper 1 (T<sub>H</sub>1) cells, respectively. Group 2 comprises ILC2s, which, similar to T<sub>H</sub>2 cells, produce interleukin-5 (IL-5) and IL-13 and express the transcription factor GATA-3. Group 3 cells, similar to T<sub>H</sub>17 cells, require the transcription factor ROR $\gamma$ t for their development and functions. Extensive work over the last decade has led to the identification of the transcription-factor networks that control the development and functions of each group of ILCs<sup>2–4</sup>. However, how the 3D organization of the genome contributes to the regulation of their development, cell identity and functions remains unknown.

At large scales, chromosomes are segregated into regions of chromatin enriched in active genes and regions associated with repressive chromatin markers where genes tend to be silenced<sup>5</sup>. This higher level of chromatin compaction is referred to as compartment organization<sup>6</sup>. Topologically associating domains (TADs) represent the next layer of genome organization, in which sub-megabase chromatin regions show a noteworthy preference for internal interactions (intra-TAD) as compared with interactions with loci outside the given region (extra-TAD)<sup>7</sup>. In this manner, regulation of one or several genes can be coordinated by allowing different sets of promoter–enhancer interactions to occur. Finally, short-range

and long-range interactions between distal genomic regulatory elements (REs), such as loops and stripes, constitute the finest layer of spatial organization of the genome, with critical roles in chromatin compaction, folding and gene expression regulation. Despite our increased understanding of the 3D genome conformation, how large-scale 3D configurations of the genome or the structure of local interactions at specific genomic loci determine ILC development and functions is yet to be elucidated.

## Results

**The large-scale 3D genome structure of ILCs.** To establish how the 3D topology of the genome contributes to the development and functional specification of each ILC group, we coupled RNA sequencing (RNA-seq) and assay for transposable-accessible chromatin using sequencing (ATAC-seq) analysis to high-resolution chromosome conformation capture (Hi-C). As high-quality Hi-C datasets require high numbers of cells with >90% viability, we selected the following representative cell types of each group of ILCs to perform ultra-deep Hi-C and RNA-seq analysis: mouse splenic NK cells (group 1), in vitro expanded bone marrow ILC2s (group 2) and the ILC3-like cell line MNK-3 (group 3) because ILC3 numbers are limited and cannot be expanded in vitro<sup>8</sup>. Previous works demonstrated the high degree of similarity between primary ILC3 and MNK-3 cells at both the transcriptional and the chromatin accessibility levels<sup>9,10</sup>. Indeed, we estimated the similarity between MNK-3 cells and primary ILC3s using principal component analysis (PCA) and differential accessibility analysis of ATAC-seq profiles, and determined that 92% of accessible DNA regions overlap between

<sup>1</sup>Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. <sup>2</sup>Institute for Immunology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. <sup>3</sup>Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. <sup>4</sup>Penn Epigenetics Institute, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. <sup>5</sup>Department of Microbiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. <sup>6</sup>Division of Allergy and Immunology, Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA. <sup>7</sup>Division of Protective Immunity, Department of Pathology and Laboratory Medicine, The Children's Hospital of Philadelphia, Philadelphia, PA, USA. <sup>8</sup>These authors contributed equally: Michaël F. Michieletto, John J. Tello-Cajiao. ✉e-mail: [vahedi@penntmedicine.upenn.edu](mailto:vahedi@penntmedicine.upenn.edu); [jhena@penntmedicine.upenn.edu](mailto:jhena@penntmedicine.upenn.edu)

primary ILC3s and MNK-3 cells (Fig. 1a and Extended Data Fig. 1a–d). Moreover, MNK-3 cells express classical functional markers associated with ILC3 functions in culture (Extended Data Fig. 1e), indicating that these cells are a suitable experimental system to query the 3D genome architecture of ILC3s.

At large scales, chromatin segregates into A compartments that often contain active gene-rich euchromatin, and B compartments associated with repressive histone marks mostly delineating inactive heterochromatin regions<sup>6,11</sup>. Thus, we first assessed the distribution of A/B compartments in each ILC group and in common lymphoid progenitors (CLPs) to establish the extent to which the compartment structure of each ILC subset changes in relation to their earliest progenitor<sup>12</sup> (Extended Data Fig. 2a,b and Supplementary Table 1). Overall, ILCs and CLPs show similar distributions of compartment states, as most of the genome (71.8%) does not change compartment status during the transition from CLPs to mature ILCs (Fig. 1b). Importantly, we observed that key transcription factors driving the commitment to the ILC lineage (*Id2*, *Nfil3*, *Tox*, *Tcf7* and *Gata3*) already reside in conserved A compartment regions at the CLP stage, suggesting that they might be primed to be activated in these multipotent lymphoid precursors (Fig. 1c and Extended Data Fig. 2c). Nevertheless, we also observed that around 28% of genomic regions flipped compartment status in at least one ILC subset and that these regions display gene expression patterns concordant with the compartment organization in each ILC group (Extended Data Fig. 2d–g and Supplementary Table 2). Altogether, these results indicate that most of the large-scale 3D genome organization of ILCs is already pre-established at the CLP stage; thus, ILC specification and functional differentiation might be instead dependent on local chromatin remodeling at the megabase and sub-megabase scale.

TADs are fundamental units of 3D genome organization that span up to a few megabases in size. TADs are demarcated by two genomic elements usually enriched in CTCF binding motifs that act as boundaries<sup>7,13–15</sup>. The preference for local associations within TADs facilitates gene regulation by insulating chromatin regions containing promoter–enhancer interactions. Therefore, to investigate the influence of the 3D genome organization at the sub-megabase scale in ILCs, we next examined the TAD structure and its contribution to gene expression programs in each ILC group.

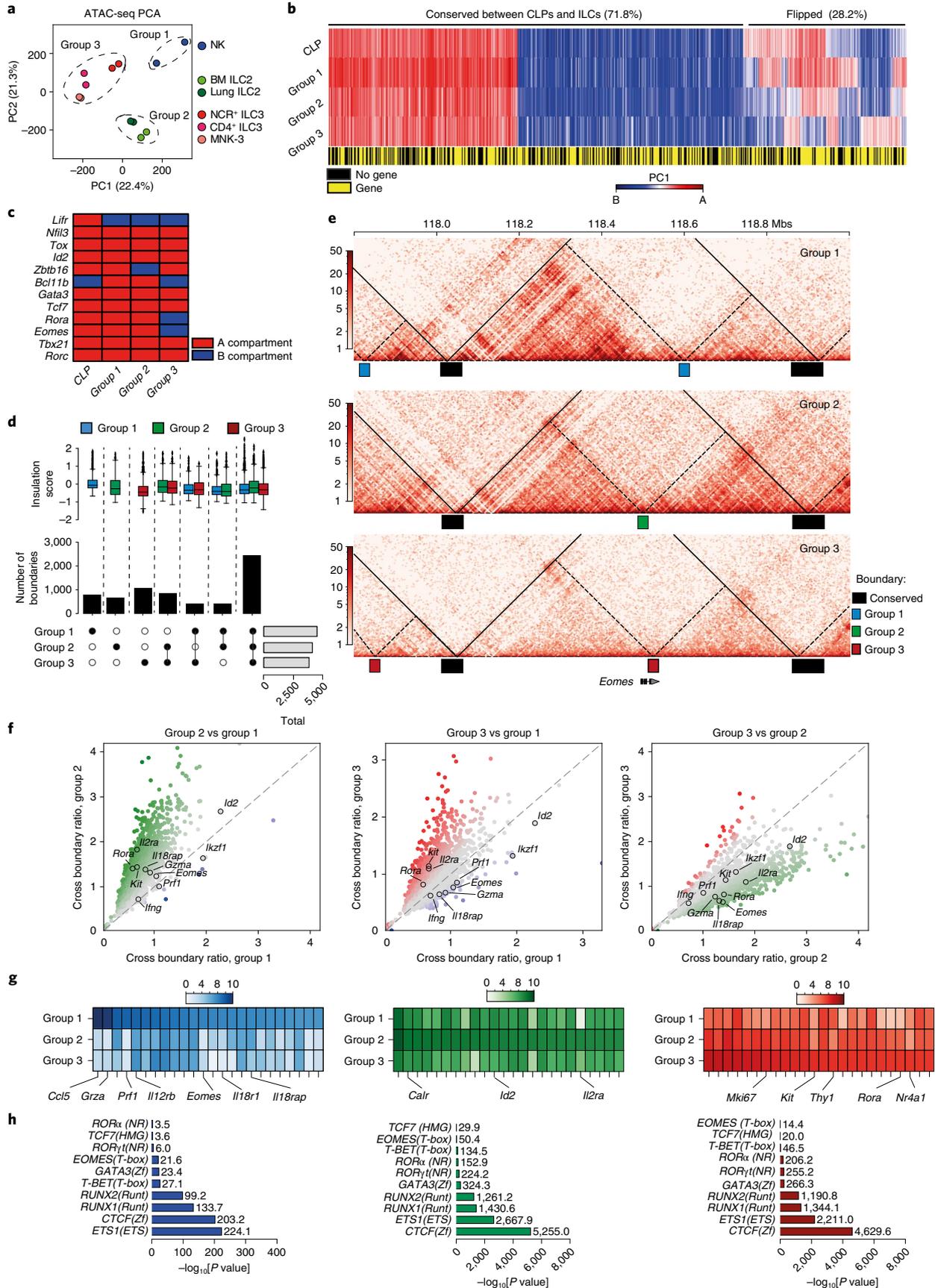
First, we looked at the distribution of boundary elements in ILCs and observed that approximately 36% of the boundaries were specific to one ILC subset. The remaining boundaries (64%) were therefore shared by at least two ILC groups, and more than half of these were conserved in all ILC groups (Fig. 1d). These observations indicate an important degree of conservation in the 3D genome organization at the megabase level across ILC groups. We then hypothesized that TADs containing genes associated with a specific ILC-subset program should have higher frequencies of intra-TAD

interactions than the other two groups of ILCs. To perform a direct comparison, we sought to calculate a cross-boundary ratio, which expresses the relative odds of loci within a TAD to interact with one another or with elements outside the TAD boundaries<sup>12,16</sup>. To facilitate this comparison, we defined TADs as the genomic region between two conserved boundaries (Fig. 1e and Extended Data Fig. 3a–c). Interestingly, we found a shift toward higher cross-boundary ratios in TADs from group 2 ILCs (Extended Data Fig. 3b and Supplementary Table 3). More importantly, we observed that the TADs with comparatively higher cross-boundary ratios in each group of ILCs ( $\log[\text{fold change}] > 0.25$ ) contain genes with prominent biological roles in the given ILC program such as *Eomes* and *Ikzf1* in group 1 ILCs, *Il2ra* and *Rora* in group 2 ILCs and *Kit* in group 3 ILCs (Fig. 1f). Hereafter, we will refer to these TADs as ILC-subset-associated TADs. In concordance, among the top 25 most differentially expressed genes within ILC-subset-associated TADs, we identified genes known to be critical for the development or functions of each ILC group (Fig. 1g). Moreover, each set of subset-associated TADs also displays enrichment in motifs of transcription factors with canonical roles in the development or functions of the corresponding ILC group (Fig. 1h). Altogether, these results indicate that, at the sub-megabase scale, the 3D chromatin organization contributes to the biology of each ILC group through selective enrichment of intra-TAD interactions.

**The local 3D genome structure of ILCs.** Short-range and long-range interactions between promoters and enhancers via looping represent the finest scale of spatial arrangement that directly relates to gene activity. As such, specific loop interactions and their combination determine cell identity and functions. Therefore, we examined how local interactions within ILC-subset-associated TADs determine ILC-subset-specific gene expression programs.

First, we identified DNA loops that link or contain gene bodies (hereafter referred to as gene loops) within each set of ILC-subset-associated TADs, and then we ranked those genes based on the number of loops in which they are involved (Fig. 2a,b and Supplementary Table 4). Importantly, we observed that critical genes for the development or functions of each group of ILCs reside in highly DNA-interacting neighborhoods as reflected by the number of loops around them (Fig. 2a,b and Supplementary Tables 5–7). Next, we sought to compare the average strength of each gene loop across the three groups of ILCs. To do so, we computed the fold change of its average contact frequency between every pair of cell types. Our comparisons show that genes specific to an ILC program have stronger (average) loop interactions in the corresponding ILC group (Fig. 2c) and that the strength of those loops correlates with their expression level, indicating that loop strength is critical for establishing gene expression programs in ILCs (Extended Data Fig. 3d). We also observed that regions that form stripes, which

**Fig. 1 | Large-scale 3D genome organization in ILCs.** **a**, PCA of chromatin accessibility profiles of NK cells (group 1), bone marrow and lung ILC2s (group 2), small intestine CD4<sup>+</sup> and NCR<sup>+</sup> ILC3s, and MNK-3 cells (group 3). **b**, Genome-wide compartment distribution in CLPs and ILCs. Regions in the A and B compartments correspond to positive (red) and negative (blue) values in the PC1 bar. The black and yellow bar indicates regions that overlap or do not overlap with genes. **c**, A/B compartment distribution of genes known to play critical roles in CLPs and ILCs. **d**, ‘Upset’ plot of the distribution of boundary elements. Gray bars represent the total number of boundaries called in group 1, 2 and 3 ILCs. The number of boundaries shared by one, two or three ILC groups are represented by black bars. The top panel shows insulation scores of boundaries in each ILC group per intersection set as defined in the bottom panel. Boxes show dataset quartiles, and whiskers show the distribution range. Dots represent outliers determined by the interquartile range. **e**, TAD containing the ILC1-associated gene *Eomes*. TADs are delineated by boundaries conserved in all groups of ILCs (black rectangles), whereas nonconserved boundaries define sub-TADs (colored rectangles). Solid and dashed lines represent TADs and sub-TADs, respectively. The scale bar represents normalized contact frequency. **f**, Scatterplot of the cross-boundary ratios of conserved TADs. The color gradient indicates the bias in the fold change of the TAD cross-boundary ratio in the corresponding comparison. Gray dots represent TADs with absolute  $\log[\text{fold change}] < 0.25$ . TADs containing genes involved in ILC biology are highlighted. **g**, Heatmap of the top 25 differentially expressed genes in TADs with the highest comparative cross-boundary ratios in group 1 (left), group 2 (middle) and group 3 (right) ILCs. Genes with roles in each group of ILCs are denoted. **h**, Transcription factor binding motif analysis over chromatin-accessible regions located within TADs with the highest ( $\log[\text{fold change}] > 0.25$ ) comparative cross-boundary ratios in group 1 (left), group 2 (middle) and group 3 (right) ILCs. Numbers indicate the  $-\log_{10}[P\text{value}]$  of denoted binding motifs (two-sided binomial test).



represent genomic loci that interact with entire contiguous regions<sup>17,18</sup>, contain genes involved in ILC development and maintenance such as *Id2*, *Gata3*, *Eomes* or *Rorc* (Fig. 2d, Extended Data Fig. 3e and Supplementary Tables 8–10). Such stripes are thought to form via loop extrusion and are likely to constitute an additional mechanism of gene regulation at the local scale (Fig. 2e and Extended Data Fig. 3e). Taken together, these results indicate that the loop and stripe structures of key genes critical for ILC development and functions correlate with their transcriptional activity, and in this way, local 3D interactions at the sub-TAD scale appear to contribute to ILC-subset-specific programs.

### 3D organization of the *Id2* locus underlies ILC2 ontogenesis.

Commitment to the ILC lineage and maintenance of its identity is determined by the transcriptional repressor *Id2*<sup>19</sup>. ID2 binds to and functionally inactivates a set of transcriptional activators such as E2A, E2-2 and HEB, which are important for adaptive lymphocyte development. As such, *Id2* expression is now considered a hallmark of all ILC subsets in mice and humans.

Our integrative analysis revealed that the local spatial organization of the region containing the *Id2* locus, as well as its expression level, is unique in each group of ILCs, indicating that *Id2* expression is precisely controlled by the interplay between ILC-subset-specific transcription factors and the remodeling of the 3D structure of the chromatin during development<sup>20</sup>. Interestingly, group 2 ILCs expressed higher levels of *Id2* than other ILC subsets (Extended Data Fig. 3d). Moreover, the 3D architecture of the *Id2* locus in group 2 ILCs shows multiple unique characteristics that correlate with the high expression of this transcriptional regulator in these cells. First, the cross-boundary ratio of the TAD containing *Id2* was higher in ILC2s than in group 1 and 3 ILCs. Second, *Id2* was among the top 25 most significantly expressed genes located only within TADs associated with ILC2s (Fig. 1g). Third, the mean strength of DNA loops that link or contain the *Id2* gene body was stronger in ILC2s than in other ILC groups (Fig. 2c). Finally, although shorter than in group 1 ILCs, the stripe at the *Id2* locus was stronger in group 2 ILCs as reflected by its higher stripiness score (Fig. 2d). Indeed, *Id2* was among the top ten genes by stripiness in ILC2s (Fig. 2d). Altogether, these observations indicate that the frequency and strength of interactions between the *Id2* gene body and downstream distal elements within the TAD are higher in group 2 ILCs than in other ILC subsets. However, how the 3D folding of the *Id2* locus underlies ILC2 ontogenesis by promoting high levels of *Id2* expression remains unknown. Therefore, we hypothesized that a combination of long-distance DNA interactions between the *Id2* promoter and *cis*-REs bound by transcription factors specifically required for ILC2 development control *Id2* expression in these cells.

To establish which unique features of the 3D folding of the *Id2* locus determine the expression of this lineage-defining factor in ILC2s, we examined chromatin accessibility, deposition of histone marks associated with active enhancers (H3K27ac), and loop

interactions involving the *Id2* locus and distal DNA elements in the three groups of ILCs (Extended Data Fig. 4a). Moreover, we used previously generated chromatin immunoprecipitation followed by sequencing (ChIP-seq) datasets of ILC2s to identify DNA elements bound by ROR $\alpha$  and GATA-3, two key transcription factors required for early ILC2 development<sup>4,21–26</sup> (Fig. 2f). Using this approach, we identified a region located approximately 125 kilobases (kb) downstream of *Id2* containing two segments highly enriched in H3K27ac, bound by ROR $\alpha$  and GATA-3, which interacted with the *Id2* promoter through DNA loops, specifically in group 2 ILCs (hereafter referred to as locus control region 1, *LCR1*) (Fig. 2f). We also identified a second region located 65 kb downstream of *Id2* (hereafter, *LCR2*) characterized by an enhanced deposition of H3K27ac and a highly accessible region that interacted with *Id2* in group 2 ILCs, but not in other ILCs (Fig. 2f and Extended Data Fig. 4a). In addition, motif analysis showed that binding sites for key transcription factors controlling ILC2 development and functions, such as ETS1, GATA-3, RUNX1, RAR:RXR and ROR $\alpha$ , are present within *LCR1* and *LCR2* (Extended Data Fig. 5a and Supplementary Table 11). Altogether, these analyses suggest that distal DNA interactions between the *Id2* promoter and *cis*-REs within *LCR1* or *LCR2* might be required for the control of *Id2* expression in ILC2s.

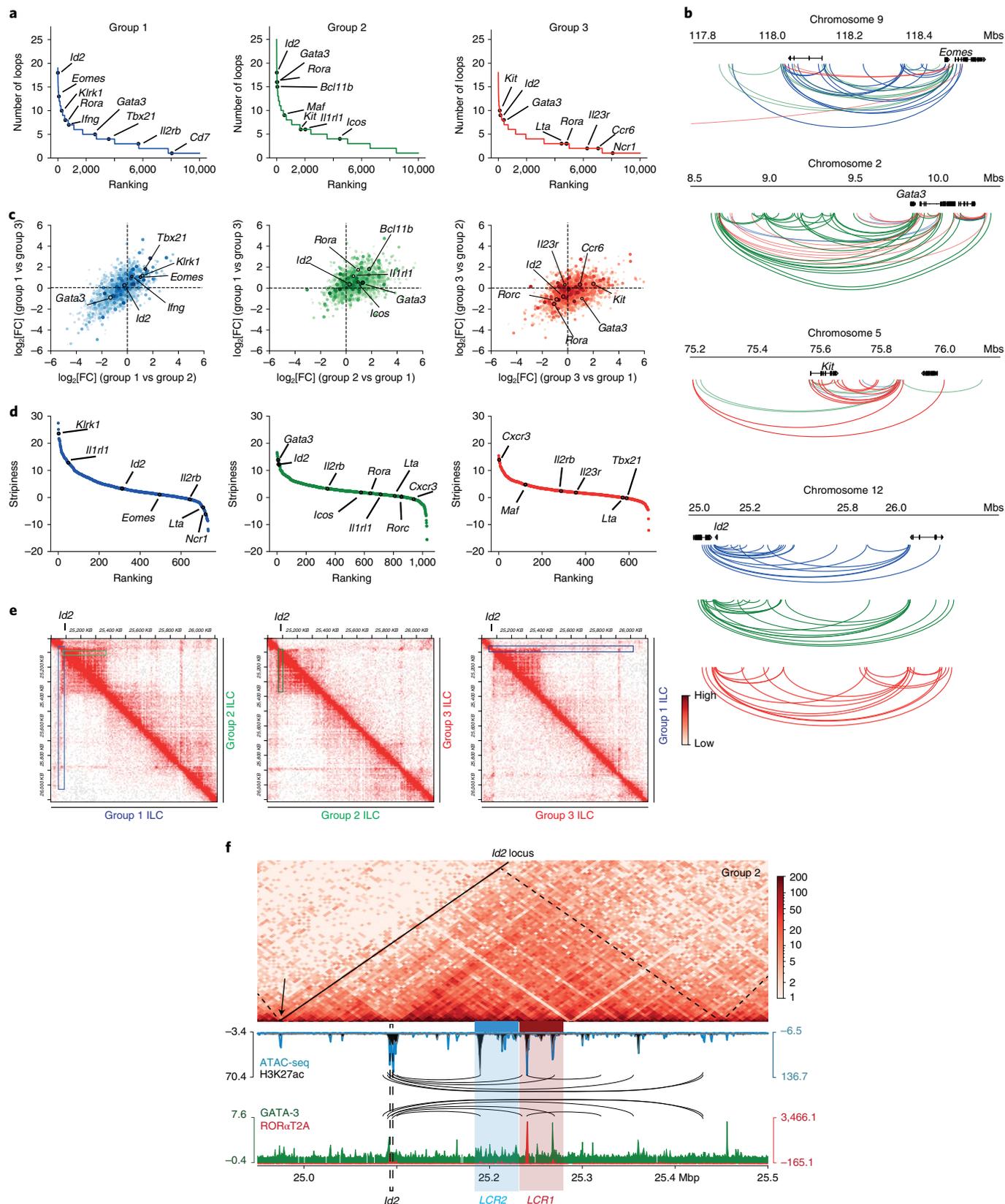
To establish whether *LCR1* or *LCR2* contain *cis*-REs necessary for the development, homeostasis or functions of ILC2s, we generated mice in which we individually deleted these loci using the CRISPR/Cas9 system. First, we determined the frequencies and numbers of each ILC group in their canonical tissue locations, in *LCR1*-deficient and *LCR2*-deficient mice. Strikingly, ILC2s were dramatically reduced in the tissues examined such as lungs, visceral adipose tissue, skin, small intestine lamina propria and bone marrow of *LCR1*<sup>-/-</sup> mice, but not in *LCR2*<sup>-/-</sup> mice (Fig. 3a,b). In contrast, group 1 ILCs (NK cells and ILC1s), ILC3s or any other *Id2*-expressing myeloid or lymphoid cell remained unaltered in both mouse strains, indicating that *LCR1* might be specifically required for ILC2 development (Fig. 3c and Extended Data Fig. 5b–d). To confirm the ILC2-specific defect in *LCR1*-deficient mice, we crossed *LCR1*<sup>-/-</sup> mice with *Arg1*-YFP reporter mice, which labels mature and developing ILC2s<sup>27,28</sup>. As expected, ILC2s were absent in the lungs and bone marrow of *LCR1*<sup>-/-</sup> *Arg1*-YFP mice (Extended Data Fig. 5e). In concordance with previous reports<sup>29,30</sup>, this defect was associated with decreased eosinophil frequencies and numbers, as well as reduced levels of IL-5 in the serum (Extended Data Fig. 5f,g). Altogether, our results indicate that distal DNA interactions between the *Id2* promoter and *cis*-REs within *LCR1* are specifically required for the development or homeostasis of ILC2s.

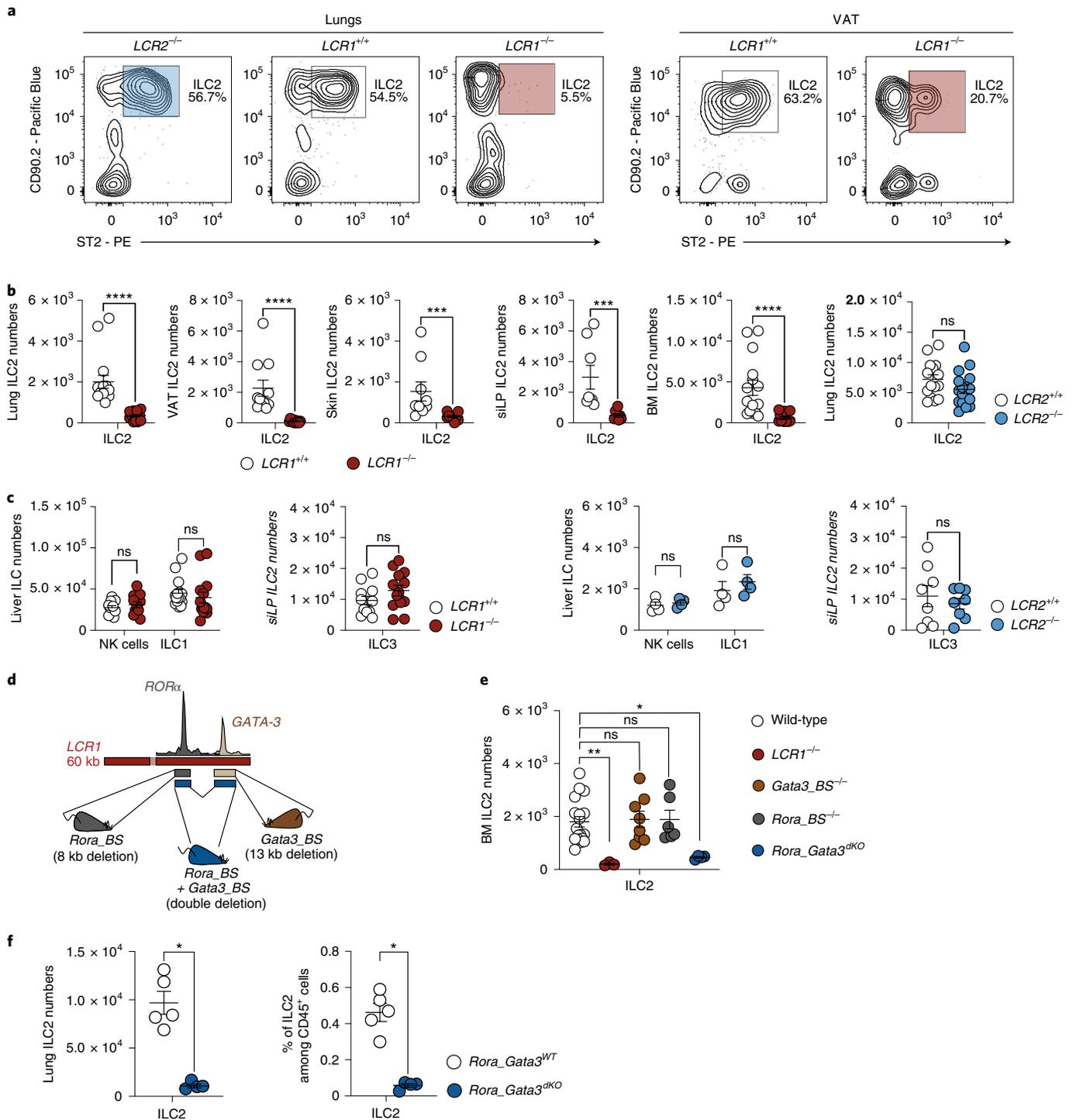
*LCR1* contains two distinct *cis*-REs bound by the transcription factors ROR $\alpha$  and GATA-3, which directly interact with the *Id2* promoter in ILC2s. Thus, to determine the individual contribution of each *cis*-RE to the *LCR1*<sup>-/-</sup> phenotype, we deleted the ROR $\alpha$  and GATA-3 binding sites individually from the mouse genome and determined ILC2 numbers in the bone marrow (Fig. 3d).

**Fig. 2 | Differential local 3D structure at key loci correlates with the gene expression programs of each ILC subset.** **a**, Rank plot of the number of loops that link or contain gene bodies in group 1, group 2 and group 3 ILCs. Genes with known roles in each group of ILCs are denoted. **b**, Examples of loop structures around ILC-subset-specific genes and *Id2* in all ILC groups. Blue, green and red arcs refer to loops present in group 1, group 2 and group 3 ILCs, respectively. **c**, Pairwise comparisons of log[fold change] of average gene loop strength in each ILC group vs the others. Dots represent genes expressed in group 1 (left), group 2 (middle) and group 3 (right) ILCs. Color intensity indicates low-to-high gene expression levels in log[RPKM + 1] units. FC, fold change. **d**, Rank plot of the stripiness of stripes detected in group 1 (left), group 2 (middle) and group 3 (right) ILCs. Genes with known roles in each group of ILCs are highlighted. **e**, Example of the differences in stripe structure at the *Id2* locus between ILCs. Pairwise visualization between ILCs of the contact heatmap of the 1-megabase region around the *Id2* locus. Color boxes highlight the significant ( $P < 0.05$ ) stripe imputed in the group 1 (blue), group 2 (green) or group 3 (none) ILC contact frequency map. The scale bar in the heatmap represents normalized contact frequency. **f**, Visualization of the contact heatmap at the *Id2* locus of bone marrow ILC2s overlapped with chromatin-accessible regions determined by ATAC-seq, H3K27ac deposition (gray and blue overlapped tracks, respectively), and GATA-3 and ROR $\alpha$  binding determined by ChIP-seq (green and red overlapped tracks, respectively). Black arcs represent loop interactions detected in the heatmap. Shaded boxes delineate *LCR1* and *LCR2*. Dashed lines represent sub-TADs, and solid lines represent TADs. Arrowheads indicate the position of conserved boundaries. The scale bar in the heatmap represents normalized contact frequency.

Surprisingly, ILC2 development was unaffected in both mouse strains when compared with wild-type and *LCR1*<sup>-/-</sup> mice (Fig. 3e). However, deletion of both the ROR $\alpha$  and GATA-3 binding sites in *cis* from the mouse genome (referred to as *Rora\_Gata3*<sup>dKO</sup>) led

to a complete absence of ILC2s in the bone marrow and lungs (Fig. 3f). Altogether, our results indicate that specific distal interactions between the *Id2* promoter and unique *cis*-REs are required for the development of group 2 ILCs.





**cis-REs within *LCR1* regulate the early development of ILC2s.** ILC2s develop in the bone marrow and fetal liver from CLPs, which subsequently differentiate into intermediate progenitors that eventually give rise to mature ILCs<sup>31–33</sup>. More specifically, CLPs give rise to  $\alpha_4\beta_7^+$  lymphoid progenitors ( $\alpha$ -LPs) that differentiate into common helper-like ILC progenitors (ChILPs)<sup>34–37</sup>. ChILPs are the precursors of a heterogeneous population of ILC precursors (ILCp) that includes ILC2 precursors (ILC2p), which is the last developmental stage before mature ILC2s arise. To determine when during development the *LCR1* is required for ILC2 generation, we performed single-cell RNA sequencing (scRNA-seq) on 2,309 wild-type cells and 3,729 *LCR1*<sup>-/-</sup> CD45<sup>+</sup> Lin<sup>-</sup> CD127<sup>+</sup> bone marrow cells, which

contain ILC2s and their progenitors. Using graph-based analyses, we identified 15 different clusters expressing *Il7r* (Fig. 4a and Extended Data Fig. 6a). In contrast, only clusters 3–7, 9, 11 and 12 expressed *Id2* and other genes critical for early ILC development and specification (Fig. 4b and Extended Data Fig. 6a). Next, we annotated each cluster based on a combination of genes that delineate early (*Flt3*, *Notch1* and *Bcl11a*) and common ILC progenitors (*Zbtb16*, *Tcf7* and *Tox*) (Fig. 4b). In addition, committed ILC subsets were assigned by the expression of transcription factors and surface molecules characteristic of each ILC subset such as *Tbx21*, *Gata3* and *Rorc* for group 1, 2 and 3 ILCs, respectively (Fig. 4b,c). Indeed, hierarchical clustering of this dataset revealed two major

**Fig. 3 | Multiple long-distance interactions between the *Id2* promoter and *LCR1* specifically control ILC2 homeostasis or development.** **a**, Representative flow cytometry plots for lung and visceral adipose tissue (VAT) ILC2s. **b**, Quantification of ILC2 numbers in the lung, VAT, skin, small intestine lamina propria (siLP) and bone marrow (BM) of wild-type, *LCR1*<sup>-/-</sup> and *LCR2*<sup>-/-</sup> mice. siLP and skin are a pool of two independent experiments (skin, *n* = 9 vs 9; siLP, *n* = 8 vs 8). VAT and lungs are a pool of at least three independent experiments (VAT, *n* = 11 vs 10; lungs, *n* = 15 vs 16 and 15 vs 17). BM is a pool of four experiments (*n* = 15 vs 14). All experiments were repeated at least three times. ns, not significant; \*\*\**P* ≤ 0.0005; \*\*\*\**P* ≤ 0.0001 (Mann-Whitney *U* two-tailed test). **c**, Quantification of ILC1s and NK cells in the liver, and siLP ILC3 numbers in wild-type, *LCR1*<sup>-/-</sup> and *LCR2*<sup>-/-</sup> mice. Each histogram is a pool of two independent experiments, except for liver ILC numbers in *LCR2*<sup>-/-</sup> mice (one experiment). Experiments were repeated at least two times (liver, *n* = 11 vs 11 and 4 vs 4; ILC3s, *n* = 12 vs 13 and 8 vs 8). ns, not significant (ILC3, Mann-Whitney *U* two-tailed test; liver ILCs, two-way ANOVA with multiple comparisons and Bonferroni correction). **d**, Schematic representation of the RORα and GATA-3 binding site deletions in the mouse genome. **e**, Quantification of BM ILC2 numbers in wild-type, *LCR1*<sup>-/-</sup>, *Gata3*<sub>BS</sub><sup>-/-</sup>, *Rora*<sub>BS</sub><sup>-/-</sup> and *Rora*<sub>Gata3</sub><sup>dKO</sup> mice. Histogram is a pool of at least two independent experiments. Quantification was repeated at least two times (*n* = 18 wild-type, 3 *LCR1*<sup>-/-</sup>, 8 *Gata3*<sub>BS</sub><sup>-/-</sup>, 6 *Rora*<sub>BS</sub><sup>-/-</sup> and 4 *Rora*<sub>Gata3</sub><sup>dKO</sup> mice). ns, not significant; \**P* ≤ 0.05; \*\**P* ≤ 0.01 (one-way ANOVA with multiple comparisons and Bonferroni correction). **f**, Quantification of lung ILC2s in *Rora*<sub>Gata3</sub><sup>WT</sup> and *Rora*<sub>Gata3</sub><sup>dKO</sup> mice (*n* = 5 *Rora*<sub>Gata3</sub><sup>WT</sup> and 4 *Rora*<sub>Gata3</sub><sup>dKO</sup> mice). Data are a pool of two independent experiments. \**P* ≤ 0.05 (Mann-Whitney *U* two-tailed test). BS, binding site; dKO, double knockout; WT, wild-type. All error bars indicate s.e.m.

branches separating *Id2*-expressing clusters from clusters 1, 8, 13, 14 and 15, which consist of early lymphoid progenitors (Extended Data Fig. 6a,b).

Strikingly, by comparing the relative frequency of the different *Id2*-expressing clusters, we observed a drastic reduction in the proportions of cells constituting ILC2s (cluster 3), and immature ILC2s (cluster 5) in *LCR1*-deficient mice when compared with wild-type mice (Fig. 4d). Moreover, transitioning ILC2s (cluster 12) were present only in *LCR1*-deficient mice. Surprisingly, we also observed a tenfold increase in the relative frequency of ILC3p (cluster 9) in *LCR1*<sup>-/-</sup> mice (Fig. 4d and Extended Data Fig. 6c). In concordance, reconstruction of early ILC development in wild-type and *LCR1*-deficient mice through pseudotime analysis confirmed the developmental trajectory from CLPs (cluster 8) to ILC2s (cluster 3) (Fig. 4e). Moreover, it demonstrated a developmental blockade at the ILC2p stage, as well as a skewing toward cells with gene signatures associated with ILC3p (cluster 9) and ILC1p (cluster 11) in *LCR1*<sup>-/-</sup> mice (Fig. 4a,e).

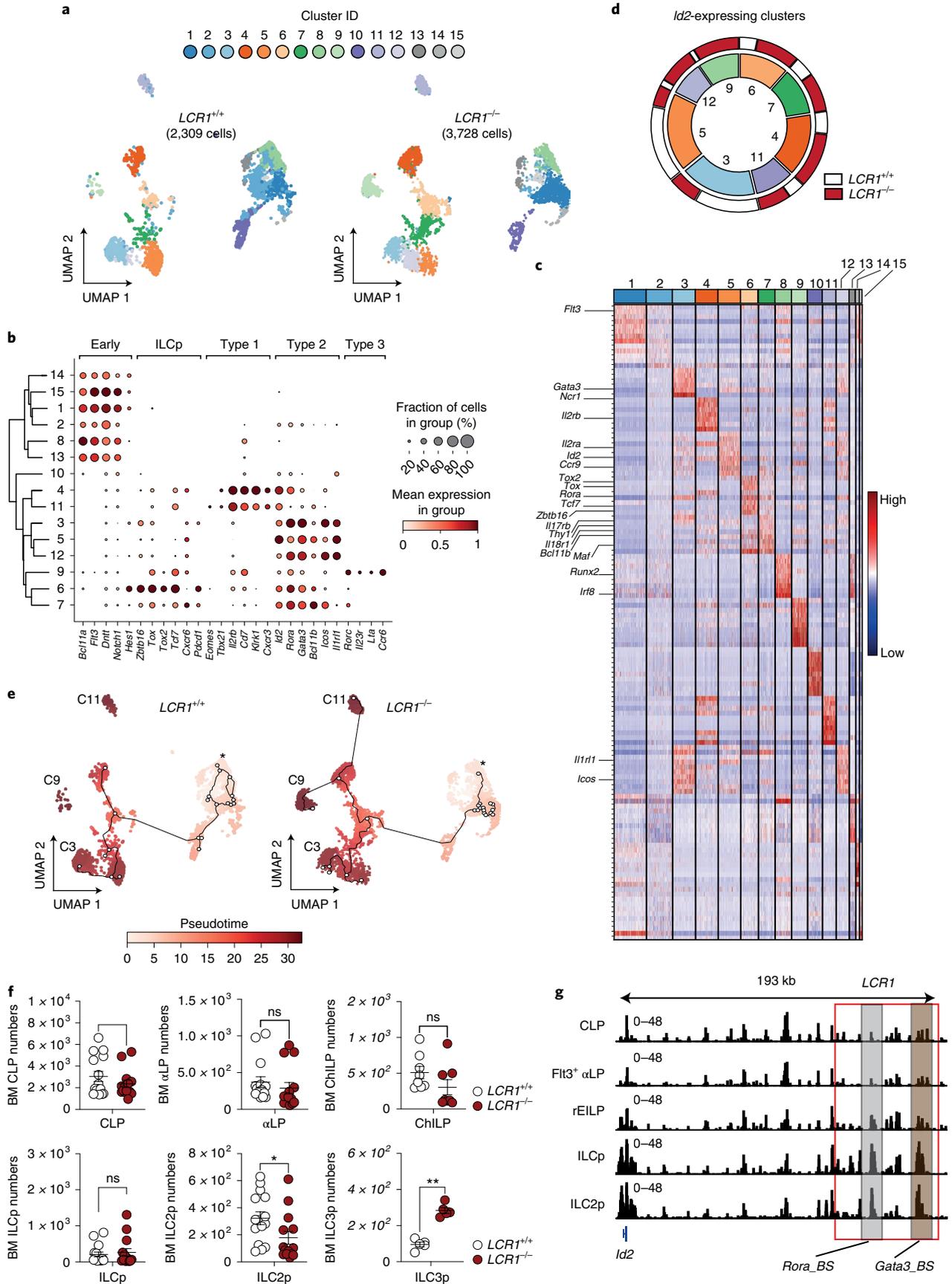
To corroborate our scRNA-seq results, we quantified ILC progenitors and ILC2 populations in the bone marrow of wild-type and *LCR1*-deficient mice using flow cytometry (Fig. 4f and Extended Data Fig. 6d). Indeed, we observed a significant reduction in the frequency and number of ILC2p and almost complete absence of bone marrow ILC2s in *LCR1*<sup>-/-</sup> mice, whereas the ILCp population remained intact (Figs. 3c and 4f). As ILC development starts during embryogenesis, we also quantified ILC2 progenitors in *LCR1*<sup>+/+</sup> and *LCR1*<sup>-/-</sup> E15.5 embryos<sup>38</sup>. Similarly, ILC2p were drastically reduced in fetal liver, fetal intestine and fetal lungs of *LCR1*<sup>-/-</sup> E15.5 embryos (Extended Data Fig. 6e). Moreover, the number of ILC3p was significantly increased in the bone marrow of *LCR1*<sup>-/-</sup> mice (Fig. 4f and Extended Data Fig. 6c). We reasoned that the *cis*-REs bound by RORα and GATA-3 within *LCR1* should become accessible during early ILC development. Indeed, we observed that the RORα and GATA-3 binding sites become fully accessible only in ILCp

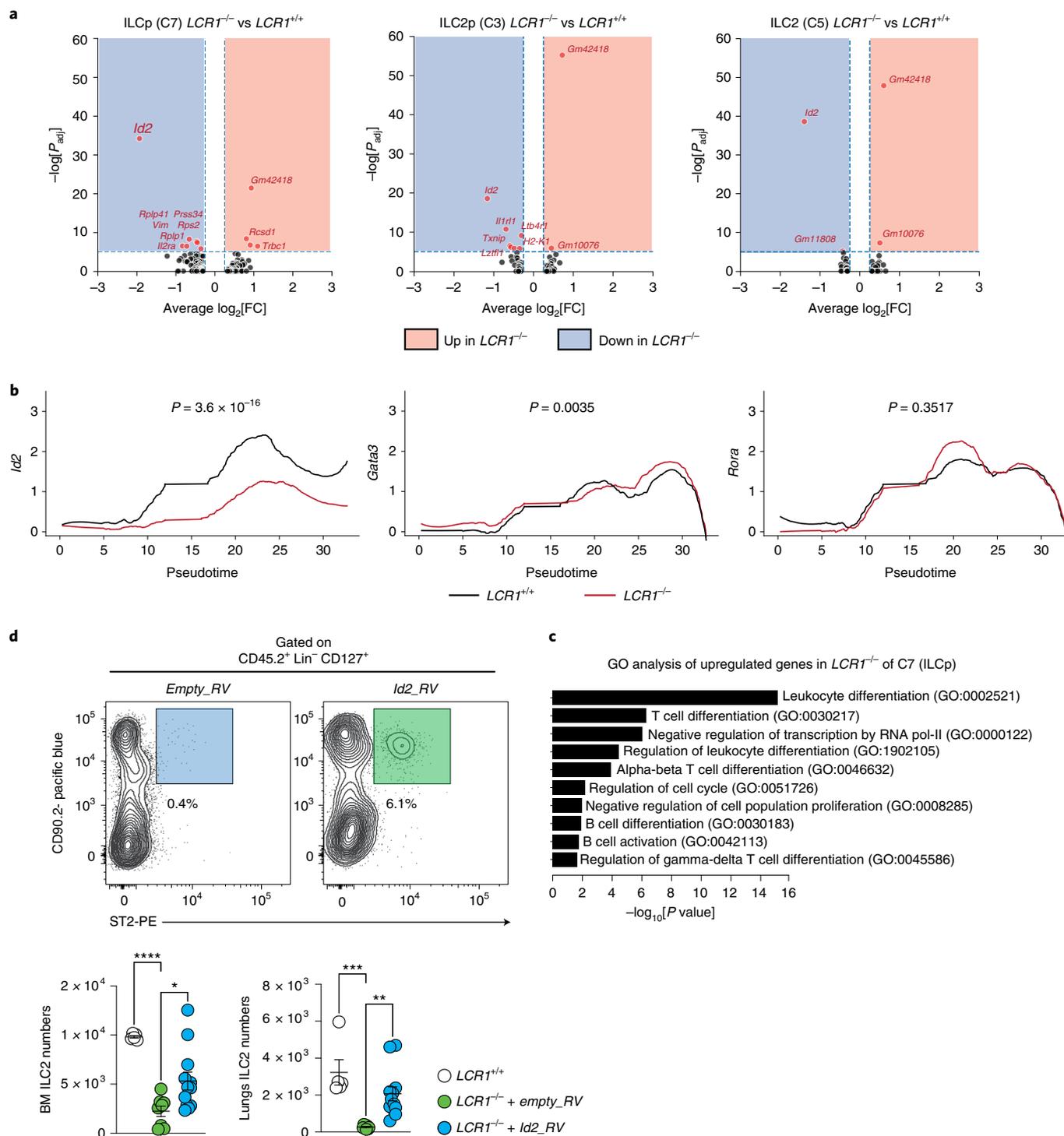
(Fig. 4g). Altogether, our data strongly indicate that two long-range DNA interactions between the *Id2* promoter and *cis*-REs within *LCR1* upon commitment to the ILC lineage are required for early ILC2 development.

As ILC2s are absent in *LCR1*<sup>-/-</sup> mice, we hypothesized that the *cis*-REs within *LCR1* might regulate *Id2* expression during early ILC2 development. To address this possibility, we performed differential gene expression analysis in the *Id2*-expressing clusters. Indeed, *Id2* was among the top downregulated genes in *LCR1*<sup>-/-</sup> ILCp when compared with wild-type (Fig. 5a, Extended Data Fig. 6f and Supplementary Table 12). Interestingly, among all of the transcription factors known to be required for ILC2 development (*Id2*, *Bcl11b*, *Rora*, *Gata3*, *Zbtb16*, *Tox*, *Ets1* and *Gfi1*), only *Id2* was downregulated throughout the ILC developmental trajectory in *LCR1*-deficient bone marrow cells when compared with wild-type (Fig. 5b and Extended Data Fig. 6g). Moreover, gene ontology (GO) analysis showed that decreased levels of *Id2* expression in *LCR1*-deficient developing ILCs were associated with pathways that positively regulate adaptive lymphocyte development and cell death (Fig. 5c). Of note, *Id2* expression was slightly decreased in CLPs and throughout ILC2 development in *LCR2*-deficient mice (Extended Data Fig. 6h). These results indicate that the activity of *cis*-REs within *LCR1* is necessary for adequate *Id2* expression upon commitment to the ILC2 lineage, and that *Id2* does not seem to control the expression of other key transcription factors involved in ILC2 development.

To confirm that *Id2* downregulation is the central defect in *LCR1*-deficient developing ILC2s, we transduced CD45.2<sup>+</sup>*LCR1*<sup>-/-</sup> bone marrow progenitor cells with retroviral particles encoding *Id2* and green fluorescent protein (*Id2*<sub>RV</sub>), or an empty vector (*empty*<sub>RV</sub>), which were then transferred to lethally irradiated congenic CD45.1<sup>+</sup> mice. Eight weeks after engraftment, we observed that mice transplanted with *LCR1*<sup>-/-</sup> bone marrow cells transduced with the *Id2*-GFP retroviral vector gave rise to ILC2s in the bone

**Fig. 4 | Multiple long-distance interactions between the *Id2* promoter and *LCR1* determine early ILC2 development.** **a**, Uniform manifold approximation and projection (UMAP) of the scRNA-seq data of 2,309 *LCR1*<sup>+/+</sup> and 3,728 *LCR1*<sup>-/-</sup> LiveCD45<sup>+</sup> Lineage<sup>-</sup> CD127<sup>+</sup> bone marrow (BM) cells. **b**, DotPlot visualization of genes that characterize early ILCs, unspecified ILCp, or clusters with signatures associated with ILC1s, ILC2s and ILC3s (denoted as Type 1, Type 2 and Type 3, respectively). Dot size represents the fraction of cells in a given cluster that express the listed gene. The color gradient represents the mean expression of the listed gene in a given cluster. **c**, Heatmap displaying normalized-scaled expression levels of marker genes of the clusters identified in **a**. Genes included correspond to the pool of top ten markers expressed by each cluster. Genes with known roles in ILC biology are highlighted. **d**, Pie chart of the distribution and size and composition of clusters in **a** that express *Id2* (clusters (C) 3–7, 9, 11 and 12). The inner ring represents relative cluster size; the outer ring represents the fraction of *LCR1*<sup>+/+</sup> and *LCR1*<sup>-/-</sup> cells in each cluster. **e**, Pseudotime plot illustrating the developmental trajectories of *LCR1*<sup>+/+</sup> and *LCR1*<sup>-/-</sup> CD45<sup>+</sup> Lineage<sup>-</sup> CD127<sup>+</sup> BM cells. The asterisk (\*) represents the selected starting point at the earliest progenitor identified, cluster 8. **f**, Quantification of CLP, α-LP, ChILP, ILCp, ILC2p, ILC3p and ILC2s from the BM of *LCR1*<sup>+/+</sup> and *LCR1*<sup>-/-</sup> mice. Data represent the pool of at least three independent experiments, except for BM ILC3p numbers (one experiment). (CLP, α-LP, ILCp and ILC2p, *n* = 15 vs 14; ChILP, *n* = 8 vs 8; ILC3p, *n* = 5 vs 5). Error bars indicate s.e.m. ns, not significant; \**P* ≤ 0.05; \*\**P* ≤ 0.01 (Mann-Whitney *U* two-tailed test). **g**, ATAC-seq peaks at the *Id2* locus during ILC2 development. The red box demarcates *LCR1*, which contains the RORα and GATA-3 binding sites. rEILP, redefined early innate lymphoid progenitor.





**Fig. 5 | *LCR1* specifically controls group 2 ILC development through *Id2* expression regulation.** **a**, Volcano plot displaying upregulated (red rectangle) and downregulated (blue rectangle) genes in  $LCR1^{+/+}$  vs  $LCR1^{-/-}$  for clusters from Fig. 4a representing ILCp (C7), ILC2p (C3) and ILC2 (C5) populations. **b**, Savitzky-Golay smoothing curve of the normalized expression levels of *Id2*, *Gata3* and *Rora* along the pseudotime axis determined in Fig. 4e. *P* value corresponds to Mann-Whitney *U* two-tailed test between smoothed signals from  $LCR1^{+/+}$  and  $LCR1^{-/-}$  mice. **c**, GO terms associated with differentially upregulated genes in cluster 7 (ILCp). For each term, the corresponding false discovery rate (FDR) is less than 0.05 (two-sided binomial test). **d**, Top, Representative flow cytometry gating strategy of bone marrow (BM) ILC2s from lethally irradiated  $CD45.1^+$  wild-type mice (host) transplanted with  $CD45.2^+$   $LCR1^{-/-}$  BM transduced with retroviral particles encoding *Id2* and the GFP (*Id2\_RV*) or an empty vector expressing only the GFP (*empty\_RV*). Bottom, Quantification of lung and BM ILC2s generated after BM transplantation of  $LCR1^{-/-}$  ( $CD45.2^+$ ) transduced cells with an empty retroviral vector (*empty\_RV*, green dots), or the *Id2*-encoding retroviral vector (*Id2\_RV*, blue dots) into a  $CD45.1^+$  host. White dots represent BM ILC2 numbers in untransduced ( $CD45.2^+$ )  $LCR1^{+/+}$  mice. Data are a pool of two independent experiments that was repeated five times. ( $n = 5$   $LCR1^{+/+}$ , 8  $LCR1^{-/-}$  *empty\_RV* and 12  $LCR1^{-/-}$  *Id2\_RV*). Error bars indicate s.e.m. \* $P \leq 0.05$ ; \*\* $P \leq 0.01$ ; \*\*\* $P \leq 0.0005$ ; \*\*\*\* $P \leq 0.0001$  (one-way ANOVA with multiple comparisons and Bonferroni correction).

marrow and lungs. However, mice that received *LCR1*-deficient bone marrow cells transduced with the empty retroviral vector failed to rescue the ILC2 developmental blockade (Fig. 5d and Extended Data Fig. 6i). These results indicate that ILC2 deficiency in *LCR1*-deficient mice occurs as a consequence of insufficient *Id2* expression in ILC progenitors.

***cis*-REs within *LCR1* promote allergic airway inflammation.** Allergic airway inflammation (AAI) is characterized by airway hyperactivation induced by eosinophilic infiltration, type 2 cytokines, and mucus hyperproduction<sup>38–40</sup>. Early responses during AAI are driven by local activation of ILC2s upon allergen exposure, which in turn promotes CD4<sup>+</sup> T cell recruitment and T<sub>H</sub>2 polarization. Indeed, recent novel genetic tools have shown a nonredundant role for ILC2s during AAI<sup>26,41,42</sup>. In addition, it has been reported that lung ILC2s are composed of cells that seed tissues at steady state and cells that expand or are de novo generated during inflammatory responses<sup>43</sup>. However, whether the development and functions of ILC2s during inflammation are determined by a similar or different 3D architecture of the *Id2* locus is unknown.

To address the role of *LCR1* during AAI, we challenged *LCR1*<sup>+/+</sup> and *LCR1*<sup>-/-</sup> mice with house dust mite (HDM) for 5 consecutive days after an initial challenge 6 days before, and quantified immune cell infiltration to the lungs, type 2 cytokine production, and lung histopathology (Fig. 6a). Importantly, except for ILC2s, we did not detect significant changes in the frequencies or total numbers of major lymphoid populations, indicating that the selective ablation of ILC2s does not impair the colonization or homeostasis of these cells at steady state (Extended Data Fig. 7a–c). Interestingly, 16 days after HDM challenge, lung ILC2s were still drastically reduced in *LCR1*-deficient mice (Fig. 6b), a phenomenon that we also observed in the context of acute allergic challenges with IL-33 and papain, indicating a nonredundant role for *LCR1* in ILC2 development at steady state and during AAI (Extended Data Fig. 8a–e).

Notably, ILC2 deficiency was associated with lower CD45<sup>+</sup> cell infiltration in the lung parenchyma and the bronchoalveolar lavage fluid (BALF) in *LCR1*-deficient mice when compared with wild-type mice (Fig. 6c). More specifically, activated CD4<sup>+</sup> T cells, T<sub>H</sub>2 cells, CD8<sup>+</sup> T cells and regulatory T (T<sub>reg</sub>) cells in the lungs, as well as IL-5 levels in serum, were significantly decreased in HDM-challenged *LCR1*<sup>-/-</sup> mice (Fig. 6d–f). Moreover, we observed a decrease, albeit nonstatistically significant, of eosinophils in ILC2-deficient mice, whereas other CD45<sup>+</sup> immune cell populations remained unchanged (Fig. 6d,e and Extended Data Fig. 9a). Importantly, histological analysis also revealed a drastic decrease in overall cell infiltration and mucus production in the lungs of *LCR1*<sup>-/-</sup> mice (Fig. 6g and Extended Data Fig. 9b). Of note, in concordance with the well-described role of ILC2s in the context of acute AAI models<sup>44–46</sup>, eosinophils were also significantly decreased in the lung parenchyma of *LCR1*<sup>-/-</sup> mice when compared with wild-type

mice in response to intranasal administration of IL-33 and papain (Extended Data Fig. 8c,e).

To exclude cell-intrinsic effects on CD4<sup>+</sup> T<sub>H</sub>2 cell differentiation or functions as a consequence of *LCR1* ablation, we assessed type 2 cytokine production in T<sub>H</sub>2 cells 16 days after the initial HDM challenge, as well as in *in vitro* polarized CD4<sup>+</sup> T<sub>H</sub>2 cells. We did not observe any differences in activation, proliferation, or type 2 cytokine production in *ex vivo* restimulated lung infiltrating T<sub>H</sub>2 or *in vitro* polarized T<sub>H</sub>2 cells (Extended Data Fig. 9c,d), indicating that a CD4<sup>+</sup> T cell-intrinsic defect in *LCR1*<sup>-/-</sup> mice is likely not responsible for decreased numbers of T<sub>H</sub>2 cells in the lungs of these mice during HDM-induced AAI. Although we cannot fully exclude the possibility that *LCR1* has additional and yet-to-be-discovered roles during AAI, these results strongly indicate that disruption of the 3D architecture of the *Id2* locus impairs ILC2 numbers irreversibly and, consequently, HDM-induced AAI.

## Discussion

Herein, we report that the local spatial configuration of the genome is significantly rewired specifically at loci associated with ILC biology to promote their development and functional differentiation. Importantly, we demonstrated that the ontogenesis of ILC2s and the progression of AAI are controlled by a unique local 3D configuration characterized by multiple long-distance interactions between the *Id2* gene body and distal *cis*-REs bound by the ILC2-associated transcription factors GATA-3 and ROR $\alpha$ , unveiling the mechanism whereby the *Id2* expression is specifically controlled in group 2 ILCs through a dynamic remodeling of the 3D architecture of its locus early in development.

We observed that approximately one-third of the genome compartments are repositioned in the transition from CLP to mature ILC. Furthermore, we found that compartment flipping correlated with the gene expression patterns detected in ILCs at these flipping regions. These results are in agreement with studies in developing T and B cells showing that transcriptional activity is affected by compartment status positioning. Interestingly, transcription factors required for the priming of the T and B lineages (*Bcl11b* and *Ebfl*, respectively) are located within B compartments in early progenitors and repositioned to A compartments upon lineage commitment<sup>12,47,48</sup>. Surprisingly, our results show that the ILC-lineage-defining transcriptional regulator ID2, as well as other key transcription factors required for early ILC development, are already within A compartments at the CLP stage. These findings suggest that the default 3D chromatin conformation during lymphocyte development might be that of innate lymphocytes, and that commitment to the adaptive lymphocyte lineage requires an active process of compartment repositioning at loci containing lineage-defining transcription factors.

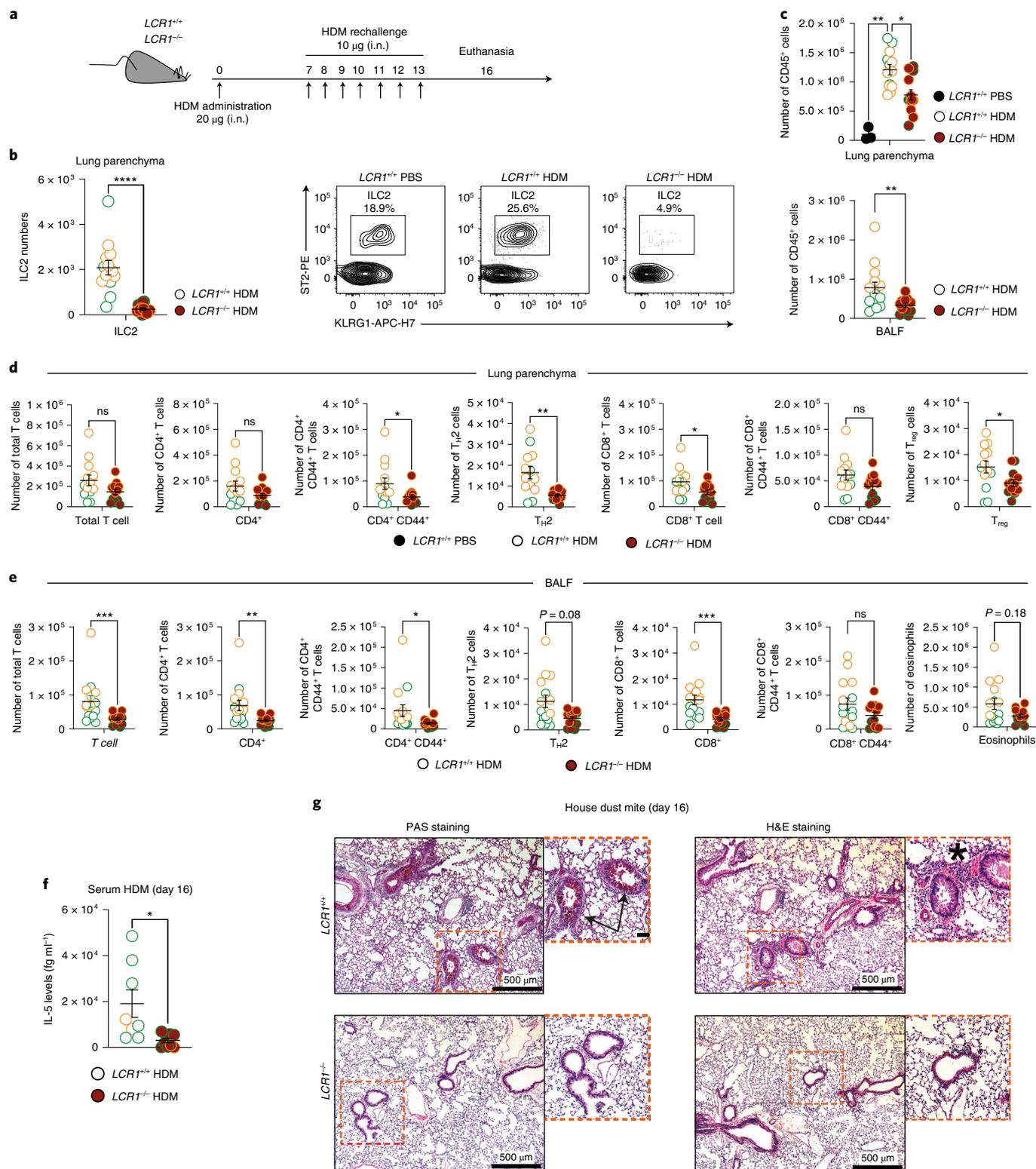
At the megabase scale, we observed higher yields of intra-TAD interactions in regions containing key genes with specific roles in

**Fig. 6 | *LCR1* determines the progression of AAI.** **a**, Schematic representation of the HDM challenge. i.n., intranasal. **b**, Left, Quantification of lung parenchyma ILC2s at day 16 after the initial HDM challenge. Data are a pool of two age- and sex-matched independent experiments ( $n = 13$  vs  $13$ ). The experiment was repeated three times. \*\*\*\* $P \leq 0.0001$  (Mann-Whitney  $U$  two-tailed test). Right, Representative flow cytometry plots of ILC2s in PBS-treated and HDM-treated mice. **c**, Quantification of CD45<sup>+</sup> cells at day 16 after the initial HDM challenge in the lung parenchyma and BALF of *LCR1*<sup>+/+</sup> and *LCR1*<sup>-/-</sup> mice. Data are a pool of two age- and sex-matched independent experiments for the lung parenchyma and two age-matched pooled experiments for the BALF (lung parenchyma,  $n = 13$  vs  $13$ ; BALF,  $n = 15$  vs  $14$ ). The experiments were repeated three times. \* $P \leq 0.05$ ; \*\* $P \leq 0.01$  (Mann-Whitney  $U$  two-tailed test). **d,e**, Quantification of total T cells, CD4<sup>+</sup> and CD8<sup>+</sup> T cells, activated CD4<sup>+</sup> and CD8<sup>+</sup> T cells, CD4<sup>+</sup> T<sub>H</sub>2 cells, T<sub>reg</sub> cells and eosinophils in the lung parenchyma (**d**) and BALF (**e**) of *LCR1*<sup>+/+</sup> and *LCR1*<sup>-/-</sup> mice 16 days after the initial HDM challenge. Data are a pool of two age- and sex-matched independent experiments for the lung parenchyma and two age-matched pooled independent experiments for the BALF (lung parenchyma,  $n = 13$  vs  $13$ ; BALF,  $n = 15$  vs  $14$ ). The experiments were repeated three times. ns, not significant; \* $P \leq 0.05$ ; \*\* $P \leq 0.01$  (Mann-Whitney  $U$  two-tailed test). **f**, Quantification of serum IL-5 in *LCR1*<sup>+/+</sup> and *LCR1*<sup>-/-</sup> mice 16 days after the initial HDM challenge. Data are a pool of two age-matched independent experiments ( $n = 8$  vs  $9$ ). \* $P \leq 0.05$  (Mann-Whitney  $U$  two-tailed test). **g**, Representative histological sections of lungs stained with Periodic Acid Schiff (PAS) and Hematoxylin and Eosin (H&E) of two *LCR1*<sup>+/+</sup> mice and two *LCR1*<sup>-/-</sup> mice at day 16. Black arrows indicate areas of mucus production; asterisk indicates cell infiltration. All error bars indicate s.e.m. Green-bordered dots denote males; orange-bordered dots denote females.

each group of ILCs. Interestingly, these regions were also enriched in motifs of transcription factors required for the commitment and differentiation of each ILC subset (Eomes, GATA-3, ROR $\alpha$  or ROR $\gamma$ t), and of architectural proteins such as CTCF and ETS1<sup>49</sup>. It has been previously suggested that architectural proteins play active roles in cell fate choices. Thus, it is possible that differential chromatin connectivity at the TAD scale may contribute to ILC development through the interplay of ILC-subset-specific transcription

factors and subclasses of architectural proteins. Future work should be aimed at establishing how the expression of transcription factors such as Eomes, GATA-3, ROR $\alpha$  and ROR $\gamma$ t contribute to the 3D architecture of early ILCp.

Examination of the TAD containing *Id2* revealed that its 3D organization in group 2 ILCs has multiple unique characteristics. First, the intra-TAD interactions and the mean strength of DNA loops around the *Id2* gene body are stronger in ILC2 than in other ILCs.



Second, *Id2* is the anchor of a stripe only in group 1 and 2 ILCs. However, the stripiness score was much higher in group 2 ILCs. Third, we identified *LCR1* located 125 kb downstream of *Id2* containing two chromatin-accessible elements highly enriched in H3K27ac, bound by GATA-3 and ROR $\alpha$ , which interact with the *Id2* promoter specifically in group 2 ILCs. Importantly, our analyses revealed that both loop interactions promote *Id2* expression in group 2 ILCs, whereas *LCR2* and the *Rroid* locus are dispensable for this process. These results, along with the previous report showing that *Id2* expression in group 1 ILCs is regulated by *cis*-REs within the lncRNA *Rroid* locus<sup>20</sup>, indicate that the activity of multiple promoter–enhancer interactions bound by different transcription factors is required for the tight regulation of this lineage-defining transcription factor in the different ILC subsets. How GATA-3, ROR $\alpha$  and other architectural proteins work in concert to establish a unique 3D configuration of the *Id2* locus remains to be elucidated. In addition, whether these transcription factors bind directly to DNA or are recruited through protein–protein interactions to their corresponding target regions, and the signals that induce their expression in group 2 ILCs are key questions to be addressed in future work.

Using scRNA-seq and flow cytometry, we established that ablation of *LCR1* blocks ILC2 development at the ILC2p stage in the bone marrow and fetal liver, which was accompanied by a skewing toward a developmental trajectory that gives rise to ILC3p and ILC1p-like cells. Moreover, using ATAC-seq, we showed that the ROR $\alpha$  and GATA-3 binding sites within *LCR1* are closed in CLPs and Flt3<sup>+</sup>  $\alpha$ -LPs, and become fully accessible at the ILCp stage. These results suggest that precursors for each group of ILCs might acquire unique 3D configurations at the *Id2* locus characterized by multiple distal promoter–enhancer interactions that are required for their development and maturation. Hence, similarly to group 1 and 2 ILCs, it is plausible to infer that another set of distal *cis*-REs that directly interact with the *Id2* promoter is necessary for group 3 ILC development. Addressing this question in future work will shed new light on early ILC development and the mechanisms that determine commitment to the ILC3 lineage.

We showed that disruption of unique promoter–enhancer interactions specifically blocks group 2 ILC development, and consequently, progression of AAI is blunted. These results highlight the importance of understanding the precise contribution of the 3D genome configuration to the transcriptional programs of specific immune cell subsets. The knowledge obtained from these studies can be harnessed for developing genetic tools to probe the functions of specific cell subsets in the context of an otherwise intact immune system<sup>50</sup>. This can be a particularly valuable tool in the study of CD4<sup>+</sup> T helper cells and ILCs, as the core components of their transcriptional programs are shared. More importantly, how alterations in the spatial organization of the genome contribute to the development of inflammatory disorders remains poorly understood. We propose that an integrative analysis of chromatin accessibility patterns, 3D genome architecture, and single nucleotide polymorphisms associated with immunological diseases in specific immune cells will shed new light on the etiology of autoimmune and inflammatory pathologies.

### Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41590-022-01295-y>.

Received: 3 December 2021; Accepted: 18 July 2022;  
Published online: 01 September 2022

### References

- Artis, D. & Spits, H. The biology of innate lymphoid cells. *Nature* **517**, 293–301 (2015).
- Robinette, M. L. et al. Transcriptional programs define molecular characteristics of innate lymphoid cell classes and subsets. *Nat. Immunol.* **16**, 306–317 (2015).
- Gury-BenAri, M. et al. The spectrum and regulatory landscape of intestinal innate lymphoid cells are shaped by the microbiome. *Cell* **166**, 1231–1246.e13 (2016).
- Shih, H.-Y. et al. Developmental acquisition of regulomes underlies innate lymphoid cell functionality. *Cell* **165**, 1120–1133 (2016).
- Dixon, J. R. et al. Chromatin architecture reorganization during stem cell differentiation. *Nature* **518**, 331–336 (2015).
- Fortin, J.-P. & Hansen, K. D. Reconstructing A/B compartments as revealed by Hi-C using long-range correlations in epigenetic data. *Genome Biol.* **16**, 180 (2015).
- Dixon, J. R. et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
- Allan, D. S. J. et al. An in vitro model of innate lymphoid cell function and differentiation. *Mucosal Immunol.* **8**, 340–351 (2015).
- Cella, M. et al. Subsets of ILC3–ILC1-like cells generate a diversity spectrum of innate lymphoid cells in human mucosal tissues. *Nat. Immunol.* **20**, 980–991 (2019).
- Parker, M. E. et al. c-Maf regulates the plasticity of group 3 innate lymphoid cells by restraining the type 1 program. *J. Exp. Med.* **217**, e20191030 (2020).
- Lieberman-Aiden, E. et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293 (2009).
- Hu, G. et al. Transformation of accessible chromatin and 3D nucleome underlies lineage commitment of early T cells. *Immunity* **48**, 227–242.e8 (2018).
- Gibcus, J. H. & Dekker, J. The hierarchy of the 3D genome. *Mol. Cell* **49**, 773–782 (2013).
- Fudenberg, G. et al. Formation of chromosomal domains by loop extrusion. *Cell Rep.* **15**, 2038–2049 (2016).
- Rowley, M. J. & Corces, V. G. Organizational principles of 3D genome architecture. *Nat. Rev. Genet.* **19**, 789–800 (2018).
- Chandra, T. et al. Global reorganization of the nuclear landscape in senescent cells. *Cell Rep.* **10**, 471–483 (2015).
- Vian, L. et al. The energetics and physiological impact of cohesin extrusion. *Cell* **173**, 1165–1178.e20 (2018).
- Yoon, S., Chandra, A. & Vahedi, G. Stripenn detects architectural stripes from chromatin conformation data using computer vision. *Nat. Commun.* **13**, 1602 (2022).
- Engel, I. & Murre, C. The function of E- and Id proteins in lymphocyte development. *Nat. Rev. Immunol.* **1**, 193–199 (2001).
- Mowel, W. K. et al. Group 1 innate lymphoid cell lineage identity is determined by a *cis*-regulatory element marked by a long non-coding RNA. *Immunity* **47**, 435–449.e8 (2017).
- Yagi, R. et al. The transcription factor GATA3 is critical for the development of all IL-7R $\alpha$ -expressing innate lymphoid cells. *Immunity* **40**, 378–388 (2014).
- Ferreira, A. C. F. et al. ROR $\alpha$  is a critical checkpoint for T cell and ILC2 commitment in the embryonic thymus. *Nat. Immunol.* **22**, 166–178 (2021).
- Miyamoto, C. et al. Runx/Cbfb complexes protect group 2 innate lymphoid cells from exhausted-like hyporesponsiveness during allergic airway inflammation. *Nat. Commun.* **10**, 447 (2019).
- Wong, S. H. et al. Transcription factor ROR $\alpha$  is critical for nuocyte development. *Nat. Immunol.* **13**, 229–236 (2012).
- Mjösberg, J. et al. The transcription factor GATA3 is essential for the function of human type 2 innate lymphoid cells. *Immunity* **37**, 649–659 (2012).
- KleinJan, A. et al. Enforced expression of Gata3 in T cells and group 2 innate lymphoid cells increases susceptibility to allergic airway inflammation in mice. *J. Immunol.* **192**, 1385–1394 (2014).
- Bando, J. K., Nussbaum, J. C., Liang, H.-E. & Locksley, R. M. Type 2 innate lymphoid cells constitutively express arginase-I in the naïve and inflamed lung. *J. Leukoc. Biol.* **94**, 877–884 (2013).
- Bando, J. K., Liang, H.-E. & Locksley, R. M. Identification and distribution of developing innate lymphoid cells in the fetal mouse intestine. *Nat. Immunol.* **16**, 153–160 (2015).
- Nussbaum, J. C. et al. Type 2 innate lymphoid cells control eosinophil homeostasis. *Nature* **502**, 245–248 (2013).
- Carlens, J. et al. Common  $\gamma$ -chain-dependent signals confer selective survival of eosinophils in the murine small intestine. *J. Immunol.* **183**, 5600–5607 (2009).
- Zook, E. C. & Kee, B. L. Development of innate lymphoid cells. *Nat. Immunol.* **17**, 775–782 (2016).
- Michieletto, M. F. & Henao-Mejia, J. Ontogeny and heterogeneity of innate lymphoid cells and the noncoding genome. *Immunol. Rev.* **300**, 152–166 (2021).

33. Martinez-Gonzalez, I., Steer, C. A. & Takei, F. Lung ILC2s link innate and adaptive responses in allergic inflammation. *Trends Immunol.* **36**, 189–195 (2015).
34. Seillet, C. et al. Deciphering the innate lymphoid cell transcriptional program. *Cell Rep.* **17**, 436–447 (2016).
35. Xu, W. et al. An Id2<sup>REP</sup>-reporter mouse redefines innate lymphoid cell precursor potentials. *Immunity* **50**, 1054–1068.e3 (2019).
36. Walker, J. A. et al. Polychromic reporter mice reveal unappreciated innate lymphoid cell progenitor heterogeneity and elusive ILC3 progenitors in bone marrow. *Immunity* **51**, 104–118.e7 (2019).
37. Constantinides, M. G., McDonald, B. D., Verhoef, P. A. & Bendelac, A. A committed precursor to innate lymphoid cells. *Nature* **508**, 397–401 (2014).
38. Hammad, H. & Lambrecht, B. N. The basic immunology of asthma. *Cell* **184**, 1469–1485 (2021).
39. Barlow, J. L. et al. IL-33 is more potent than IL-25 in provoking IL-13-producing nuocytes (type 2 innate lymphoid cells) and airway contraction. *J. Allergy Clin. Immunol.* **132**, 933–941 (2013).
40. Barlow, J. L. et al. Innate IL-13-producing nuocytes arise during allergic lung inflammation and contribute to airways hyperreactivity. *J. Allergy Clin. Immunol.* **129**, 191–198.e4 (2012).
41. Kasal, D. N. et al. A *Gata3* enhancer necessary for ILC2 development and function. *Proc. Natl Acad. Sci. USA* **118**, e2106311118 (2021).
42. Halim, T. Y. F. et al. Group 2 innate lymphoid cells are critical for the initiation of adaptive T helper 2 cell-mediated allergic lung inflammation. *Immunity* **40**, 425–435 (2014).
43. Huang, Y. et al. IL-25-responsive, lineage-negative KLRG1<sup>hi</sup> cells are multipotential ‘inflammatory’ type 2 innate lymphoid cells. *Nat. Immunol.* **16**, 161–169 (2015).
44. Xiao, Q. et al. PPAR $\gamma$  enhances ILC2 function during allergic airway inflammation via transcription regulation of ST2. *Mucosal Immunol.* **14**, 468–478 (2021).
45. Yang, Q. et al. T cell factor 1 is required for group 2 innate lymphoid cell generation. *Immunity* **38**, 694–704 (2013).
46. Toki, S. et al. TSLP and IL-33 reciprocally promote each other’s lung protein expression and ILC2 receptor expression to enhance innate type-2 airway inflammation. *Allergy* **75**, 1606–1617 (2020).
47. Lin, Y. C. et al. Global changes in the nuclear positioning of genes and intra- and interdomain genomic interactions that orchestrate B cell fate. *Nat. Immunol.* **13**, 1196–1204 (2012).
48. van Schoonhoven, A., Huylebroeck, D., Hendriks, R. W. & Stadhouders, R. 3D genome organization during lymphocyte development and activation. *Briefings Funct. Genomics* **19**, 71–82 (2020).
49. Phillips-Cremins, J. E. et al. Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* **153**, 1281–1295 (2013).
50. Claringbould, A. & Zaugg, J. B. Enhancers in disease: molecular basis and emerging treatment strategies. *Trends Mol. Med.* **27**, 1060–1073 (2021).

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

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2022

## Methods

**Mice.** The *LCR1*<sup>-/-</sup>, *LCR2*<sup>-/-</sup>, *Rora*<sub>BS</sub><sup>-/-</sup>, *Gata3*<sub>BS</sub><sup>-/-</sup> and *Rora*<sub>Gata3</sub><sup>flKO</sup> mice were generated using the CRISPR/Cas9 system. The single guide RNA (sgRNA) sequences used for the generation of knockout mice are listed in Supplementary Table 13. Genomic deletion of each locus was confirmed by PCR and Sanger sequencing. All strains were backcrossed onto the C57BL/6 background for at least three generations to control potential off-target effects. B6.SJL-*Ptprc*<sup>c</sup> *Pep*<sup>cd/</sup> Boy (CD45.1<sup>+</sup>, strain 002014) and B6.129S4-Arg1tm1.1Lky/J (Arg1-YFP, strain 015857) mice were purchased from the Jackson Laboratory. All mice were bred and maintained under pathogen-free conditions at an animal facility at the University of Pennsylvania accredited by the American Association for the Accreditation of Laboratory Animal Care. Mice were housed in accordance with the procedures outlined in the Guide for the Care and Use of Laboratory Animals under the animal study protocol 805188 approved by the Institutional Animal Care and Use Committee at the University of Pennsylvania. All experiments were performed using mice 6–19 weeks old that were age and sex matched, and all experiments were performed in both males and females for this study. *Rora*<sub>Gata3</sub><sup>flKO</sup> mice were backcrossed for two generations to control potential off-target effects and were used at an age of 4–5 weeks. Mice were housed at 21 ± 2 °C and 55% humidity (±10%) with a 12 h light/dark cycle in 7–7 individually ventilated caging (IVC) with environmental enrichment of plastic houses plus paper bedding.

**Cell isolation.** Cells from spleen and lymph node were isolated by physical dissociation and filtered through a 70 µm cell strainer. For bone marrow cell isolation, femurs were collected and crushed using a mortar and pestle. All red blood cells were lysed using ACK lysis buffer (Gibco).

**Lung cell isolation.** Lungs were isolated, minced with scissors and digested in PBS containing FBS (2%), collagenase D (1 mg ml<sup>-1</sup>) and DNase I (0.2 mg ml<sup>-1</sup>) for 35 min at 37 °C with shaking at 200 rpm using a MaxQ 4450 (Thermo Scientific). The digested lungs were then passed through a 70 µm cell strainer.

**Liver cell isolation.** Mice were perfused with 10 ml of PBS and then transferred to DMEM on ice. The liver was then removed from media and mechanically dissociated using a tissue grinder, then filtered through a 100 µm cell strainer. To pellet hepatocytes, the digested livers were centrifuged at 20g for 5 min at 4 °C. Leukocytes were then purified over an 80/40% Percoll gradient.

**Small intestine lamina propria digestion.** Small intestines were first flushed with PBS to remove feces, and Peyer's patches were removed using scissors. The small intestines were then opened lengthwise, and tissues were shaken in a Petri dish containing cold PBS to remove any remaining feces. Mucus and epithelial cells were removed by first incubating 2 × 15 min at 180 rpm and 37 °C in PBS containing FBS (2%), HEPES (20 mM, pH 7.2–7.5) and EDTA (10 mM, pH 8) using a MaxQ 4450. Intestines were thoroughly washed twice with ice-cold PBS and minced into 1 cm pieces using scissors. The minced small intestines were then digested in PBS containing FBS (2%), HEPES (20 mM, pH 7.2–7.5), collagenase D (1 mg ml<sup>-1</sup>), DNase I (0.2 mg ml<sup>-1</sup>) and dispase (0.1 U ml<sup>-1</sup>). The digested intestines were filtered through a 100 µm cell strainer and purified over an 80/40% Percoll gradient.

**Adipose tissue digestion.** Visceral adipose tissue was finely dissected with scissors and digested in PBS containing BSA (2%), collagenase D (0.5 mg ml<sup>-1</sup>) and DNase I (0.2 mg ml<sup>-1</sup>) at 37 °C for 30 min with gentle agitation. Digests were filtered through a 70 µm cell strainer and centrifuged at 450g for 5 min at 4 °C to pellet cells.

**Skin digestion.** Mice were first shaved using a hair trimmer, and then 2 cm<sup>2</sup> of skin was used for digestion. Skin was first minced into small pieces prior to digestion. Skin pieces were digested in PBS containing FBS (2%), collagenase IX (1 mg ml<sup>-1</sup>), DNase I (0.2 mg ml<sup>-1</sup>) and hyaluronidase (0.5 mg ml<sup>-1</sup>) for 90 min at 37 °C with gentle shaking. Digested tissues were then filtered through a 70 µm cell strainer and purified over a 67/44% Percoll gradient at 970g for 20 min at 25 °C. Cells were collected at the 67/44% interphase.

**IL-5 serum dosage.** For serum quantification, mice were anesthetized and blood was collected by intracardial puncture. Blood was then transferred to a BD Microtainer Serum Separator Tube and left at 20 to 25 °C for at least 30 min, followed by 2 min of centrifugation at 11,000 RCF. IL-5 from serum of naive or HDM-challenged mice was measured using an Enhanced Sensitivity Flex Set with an Enhanced Sensitivity Cytometric Bead Array Kit (BD Biosciences) following the manufacturer's recommendations.

**HDM-induced allergic asthma.** HDM extracts (*Dermatophagoides pteronyssinus* extracts; Greer Laboratories, lot numbers 361863 and 385930) were used to induce AAI<sup>31</sup>. In brief, mice were sensitized intranasally with 20 µg HDM extracts on day 0 and subsequently challenged with 10 µg per mouse per day on days 7–13. Three days after the last challenge, mice were anesthetized and used for immune cell quantification in the lung parenchyma or BALF, as well as for lung histology and measurement of serum IL-5. BALF was collected by administration of 2 × 1 ml of

FACS buffer (PBS, 2% FBS, 2 mM EDTA, pH 8). For lung parenchyma experiments, lungs were isolated, minced with scissors and digested in PBS containing FBS (2%), collagenase D (1 mg ml<sup>-1</sup>) and DNase I (0.2 mg ml<sup>-1</sup>) for 35 min at 37 °C with shaking at 200 rpm using a MaxQ 4450. The digested lungs were then passed through a 70 µm cell strainer. Cells were then stained and analyzed by flow cytometry.

Eight HDM challenge experiments were performed containing a total of 58 male and 35 female HDM-challenged mice, as ILC2s are absent in both sexes. Out of these eight experiments, three experiments were used to measure immune cell infiltration in the lung parenchyma, three experiments were used to measure immune cell infiltration in the BALF and for histopathological analysis, and two experiments were used to measure IL-5 levels in serum. For the three experiments for measuring immune infiltration to the lung parenchyma, we challenged a total of 20 male and 16 female age- and sex-matched mice (6.9–17.2 weeks old), which were part of two independent cohorts of males and one independent cohort of females. Two representative experimental cohorts containing 10 male and 16 female mice (6.9–12.9 weeks old) were pooled for statistical analysis and data representation. For the three experiments for measuring immune infiltration to the BALF, we challenged a total of 21 male and 15 female age-matched mice (6.9–19.5 weeks old), which were part of three independent cohorts that contained both males and females. Two representative experimental cohorts containing 14 male and 15 female mice (6.9–8.0 weeks old) were pooled for statistical analysis and data representation. For the two experiments for measuring IL-5 levels in serum, we challenged a total of 13 male and 4 female age-matched mice (7.5–8.5 weeks old), which were part of one independent cohort of males and one independent cohort that contained both males and females. These two experiments were pooled for statistical analysis and data representation.

**Lung histology.** Lung tissues from five males (three *LCR1*<sup>+/+</sup> and two *LCR1*<sup>-/-</sup>) were fixed in 10% formalin for 24 h and then placed in 70% ethanol before embedding in paraffin. Lung sections were performed by the Molecular Pathology and Imaging Core at the Center for Molecular Studies in Digestive and Liver Diseases (P30DK050306) and were stained with periodic acid–Schiff (PAS) stain or hematoxylin and eosin (H&E).

**IL-33 injection.** Mice were anesthetized with isoflurane and treated with 200 ng of recombinant murine IL-33 (rIL-33; PeproTech, 210-33) diluted in 20 µl of PBS every 24 h for 4 days. Mice were euthanized 24 h after the last administration, and lungs were collected to assess lung ILC2 and eosinophil numbers. Two rIL-33 challenge experiments were performed containing a total of 18 age-matched males of age 9.4–9.7 weeks.

**Papain challenge.** Mice were anesthetized with isoflurane, followed by intranasal administration of 30 µg of papain (Sigma-Aldrich, 5125-50GM) diluted in 20 µl of PBS once per day for 4 days. Mice were euthanized 24 h after the last administration, and lungs were collected to assess lung ILC2 and eosinophil numbers. One experiment was performed containing a total of 12 age-matched males of age 11.4–12.2 weeks.

**Retroviral particle production.** To generate retroviral particles for *Id2* overexpression, HEK-293T cells were purchased from the American Type Culture Collection (ATCC). In brief, HEK-293T cells were maintained in high-glucose DMEM medium 1X with L-glutamine (Invitrogen), supplemented with 100 U ml<sup>-1</sup> penicillin and 100 mg ml<sup>-1</sup> streptomycin (Gibco) with 10% FBS. Retroviral vectors were packaged in HEK-293T cells. In brief, 4 × 10<sup>6</sup> HEK-293T cells were plated in 5 ml of media in 10 cm dishes on the day prior to transfection. During transfection, 15 µg of MSCV-IRES-GFP or MSCV-GFP (control) plasmid was co-transfected with packaging plasmid, 15 µg of pCL-Eco, using Lipofectamine 3000 (Invitrogen). The cells were then returned to the incubator for 6 h. Subsequently, the media were changed to fresh media. Virions were collected 24 h and 48 h after transfection, snap-frozen and stored at -80 °C for future use. Viral particles were titrated in NIH-3T3 cells (ATCC) using serial dilution of HEK-derived supernatant.

**Bone marrow chimera and retroviral bone marrow transduction.** Mice were first treated 4 days before takedown with 5 mg of 5-fluorouracil. Bone marrow from CD45.2<sup>+</sup> *LCR1*<sup>-/-</sup> mice was isolated by crushing bones with a mortar and pestle. Following red blood cell lysis, cells were counted and plated at 2.5 × 10<sup>6</sup> cells ml<sup>-1</sup> in a six-well plate in IMDM containing FBS (15%) + penicillin/streptomycin (1X) and 10 ng ml<sup>-1</sup> of IL-3, 5 ng ml<sup>-1</sup> of IL-6 and 100 µg ml<sup>-1</sup> of SCF (BioLegend) and left in a 5% CO<sub>2</sub> incubator at 37 °C overnight. The next day, cells were counted and readjusted to 5 × 10<sup>6</sup> cells ml<sup>-1</sup> containing the previous cytokine mix, and retroviral supernatants supplemented with 8 µg ml<sup>-1</sup> Polybrene (EMD Millipore) were added to the appropriate wells. Plates were then spinfected at 1,800g for 90 min at 32 °C and incubated for 4 h at 32 °C prior to an overnight incubation at 37 °C. Transduced bone marrow was then washed two times with PBS and transferred to lethally irradiated CD45.1<sup>+</sup> hosts via retro-orbital injection (down to 0.2 × 10<sup>6</sup> cells per mouse). Congenic recipients were irradiated with 1,100 rad, and cells were injected later the same day. Mice were bled in order to check bone marrow reconstitution after 4 weeks, and mice were euthanized and analyzed 8 weeks after engraftment.

**Antibodies, flow cytometry and cell sorting.** All antibodies used are listed in the Nature Portfolio and were diluted in FACS buffer (PBS, 2% FBS, 2 mM EDTA, pH 8) and used to stain single cell suspensions for 30 min at 4°C. Cells were washed with FACS buffer and either fixed for intracellular staining using the Foxp3 staining buffer (eBioscience) or were fixed with 2% PFA then subjected to flow cytometry. Dead cells were eliminated by incubation of cell suspension in Viability Dye (eFluor 780 or eFluor 506) diluted in PBS for 10 min at 4°C. Cell counting was performed using 123count eBeads (Thermo Fisher Scientific, 01-1234) following manufacturer's recommendations.

**Cell sorting.** For sorts, splenic NK cells (Live, CD45<sup>+</sup>, CD3e<sup>-</sup>, NK1.1<sup>+</sup>, NKp46<sup>+</sup>) and bone marrow ILC2s (Live, CD45<sup>+</sup>, Lin<sup>-</sup>, CD90.2<sup>+</sup>, ST2<sup>+</sup>) were sorted using a BD FACSAria II SORP. Lineage (Lin) markers included CD3, CD5, CD11b, CD11c, Gr1, TCRγ/δ, NK1.1, B220, CD19 and Ter119 if not stated otherwise. For scRNA-seq, bone marrow cells were first enriched using the biotinylated Lineage antibodies. Remaining cells (progenitors) were then stained and sorted as follows: Live, CD45<sup>+</sup>, Lin<sup>-</sup>, NK1.1<sup>-</sup>CD127<sup>+</sup>.

**MNK-3 in vitro culture.** MNK-3 cells were cultured in DMEM containing FBS (10%), glutamine (1X), sodium pyruvate (1X), 2-mercaptoethanol (1X), HEPES (10 mM, pH 7.2–7.5) and penicillin/streptomycin (1X). MNK-3 cells were maintained in 10 ng ml<sup>-1</sup> mouse IL-7 (BioLegend) and 10 ng ml<sup>-1</sup> mouse IL-15 (BioLegend). To maintain high cell viability, the medium was changed every 3 days, and cells were passaged when reaching confluency. For characterization of MNK-3 cell markers, cultured cells were stimulated with 10 ng ml<sup>-1</sup> IL-2, IL-23 and IL-1β, and expression of defined functional markers was addressed by flow cytometry.

**ILC2 in vitro culture.** A total of 50,000 ILC2s were sorted from enriched bone marrow of *Arg1*-YFP mice by sorting LiveCD45<sup>+</sup>Lin<sup>-</sup>ST2<sup>+</sup>YFP<sup>+</sup> cells and expanded in vitro in complete RPMI supplemented with IL-33 (10 ng ml<sup>-1</sup>), IL-25 (10 ng ml<sup>-1</sup>), IL-7 (10 ng ml<sup>-1</sup>) and IL-2 (10 ng ml<sup>-1</sup>) with medium replacement every 2 days. Viability and proliferation were calculated using a hemocytometer and Trypan blue incorporation.

**Hi-C analysis.** Hi-C libraries were generated on up to 10<sup>6</sup> cells using an Arima-HiC Kit (Arima Genomics) and an Accel-NGS Plus DNA Library Kit (Swift Biosciences, 21024) according to the manufacturers' recommendations. Libraries were validated for quality and size distribution using a Qubit dsDNA HS Assay Kit (Invitrogen, Q32851) and a TapeStation 2200 (Agilent). Libraries were paired-end sequenced (43 bp + 43 bp) on NextSeq 550 (Illumina). The CLP Hi-C dataset comes from a previously published study under the Gene Expression Omnibus (GEO) accession number GSE79422<sup>12</sup>. All samples were aligned using the HiC-Pro pipeline as recommended by the authors<sup>52</sup>. Next, Hi-C matrices were converted into the cooler format and normalized by matrix balancing with default parameters<sup>53</sup>.

**Compartment analysis.** Compartments were called using the cooltools command-line interface (CLI) on each Hi-C dataset at 50 kb resolution. We used the GC content file generated by cooltools as the reference track. A weighted compartment score was assigned to genes that span multiple regions with different compartment scores (weighting factor, number of base pairs overlapping each region).

**TAD analysis.** TADs were called using the HiCExplorer tools<sup>54–56</sup>. The function *hicFindTADs* was applied to group 1, 2 and 3 ILC Hi-C datasets at 25 kb resolution, with default parameters. Conserved boundaries were determined by pooling all of the TAD boundaries initially reported by *hicFindTADs* and selecting the ones that appeared in all datasets of interest. Two boundaries were considered the same if they differed by no more than one bin (25 kb). Elements within a TAD can interact with other elements that are also within the TAD (intra-TAD interaction) or outside the TAD (extra-TAD interaction). For each TAD, we divided the sum of all intra-TAD interactions by the sum of all extra-TAD interactions in each ILC group. This quantity is referred to as the cross-boundary ratio<sup>16</sup>. Next, the cross-boundary ratios of a given TAD were compared across cell types by computing corresponding pairwise  $\log[\text{fold change}]$  between ILCs. Figure 1f was constructed by selecting, for each cell type, TADs with cross-boundary ratios that had  $\log[\text{fold change}] > 0.25$  when compared with the remaining cell types (termed type 1, type 2 and type 3 TADs). Next, we selected genes within these TADs that had reads per kilobase of transcript per million mapped reads (RPKM)  $\log[\text{fold change}] > 0.25$  with respect to the remaining cell types. Motif analysis of these regions was carried out by intersecting the coordinates of each TAD with the ATAC-seq peaks of the corresponding cell type, and then calling the *findMotifsGenome.pl* function of the HOMER package with default parameters.

**Cross-boundary ratio formula.** Let  $M_{ij}^{\text{ILC}_k}$  be the contact frequency between loci *i* and *j* in ILC<sub>*k*</sub> as determined by Hi-C, and let *T* be a TAD detected ILC<sub>*k*</sub> in. Thus, the cross-boundary ratio of *T* is given by

$$\text{cb}_T = \frac{\sum_{i,j} T M_{ij}^{\text{ILC}_k}}{\sum_{i,T,j} T M_{ij}^{\text{ILC}_k}}$$

The numerator represents the total count of interactions between elements belonging to the TAD, and the denominator represents the sum of interactions between the loci in the TAD with loci outside the TAD.

**Loop analysis.** We called loops at three different resolutions (5 kb, 10 kb and 25 kb) using the Mustache software with parameters -pt 0.1<sup>57</sup>. Next, for each dataset, we pooled together loops called at different resolutions by examining their anchors (overlapping of both anchors defined equivalent loops) using the bedtools suite<sup>58</sup>.

**Stripe analysis.** Stripes were detected using the software Stripenn (<https://github.com/ysora/stripenn>) with default parameters<sup>18</sup>. Stripenn uses a Canny edge detector algorithm borrowed from image processing analysis to find vertical and horizontal patterns in the Hi-C maps that correspond to highly interactive regions of chromatin. These regions are then classified and scored to identify significant stripes.

**Library preparation of bulk RNA-seq.** NK cells, in vitro expanded ILC2s and the MNK-3 cell line were washed once with 1X PBS before resuspending the pellet in 350 μl of Buffer RLT Plus (Qiagen) with 1% 2-mercaptoethanol (Sigma), vortexed briefly and stored at -80°C. Subsequently, total RNA was isolated using an RNeasy Plus Micro Kit (Qiagen). RNA integrity numbers (RINs) were determined using the TapeStation 2200, and all samples used for RNA-seq library preparation had RINs greater than 9. Libraries were prepared using the SMARTer Stranded Total RNA-seq Kit v2-Pico Input Mammalian kit (Takara Bio). Two technical replicates were generated for each experiment. Libraries were validated for quality and size distribution using the TapeStation 2200. Libraries were paired-end sequenced (38 bp + 38 bp) on the NextSeq 550. Bulk RNA data were aligned to the mm10 reference genome using STAR\_2.7.8a with `-outFilterMultimapNmax 1 -outFilterScoreMinOverRead 0 -outFilterMatchNminOverRead 0 -outSAMtype BAM SortedByCoordinate -alignEndsType Local -outReadsUnmapped Fastx`. Additionally, we counted reads using HTSeq-count with `-fbam -r name -s no -t exon -i gene_id -m intersection-nonempty` parameters<sup>59</sup>. Next, we transformed read counts into RPKM for subsequent analyses.

**Preparation of ATAC-seq samples and library.** ATAC-seq was performed as previously described, with minor modifications<sup>60,61</sup>. A total of 50,000 cells were pelleted at 550g and washed with 50 μl of ice-cold 1X PBS, followed by treatment with 50 μl of lysis buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM MgCl<sub>2</sub>, 0.1% IGEPAL CA-630). Nuclei pellets were resuspended in a 50 μl transposition reaction containing 2.5 μl of Tn5 transposase (Illumina, FC-121-1030). The reaction was incubated in a 37°C heat block for 45 min. Tagmented DNA was purified using a MinElute Reaction Cleanup Kit (Qiagen) and amplified with varying cycles, depending on the side reaction results. Libraries were purified using a QIAquick PCR Purification Kit (Qiagen). Libraries were validated for quality and size distribution using the TapeStation 2200. Libraries were paired-end sequenced (38 bp + 38 bp) on the NextSeq 550. ATAC-seq fastq files were aligned to the mm10 reference genome using the *mem* function of the software *bwa*-0.7.17 with the additional parameters `-M`. ATAC-seq peaks were subsequently called using MACS2 software with the additional parameters `-nomodel -f BAM -B -keep-dup all -broad -broad-cutoff 0.1 -q 0.1`. ATAC-seq profiles between primary ILC3s and the MNK-3 cell line were compared by first creating a reference peak catalog<sup>23,63</sup>. The reference peak catalog consisted of the pool of peaks called in each cell type with the corresponding raw counts in each sample. Next, we performed differential analysis using DESeq2<sup>64</sup>. Additionally, we extracted the normalized counts computed by DESeq2 to perform the PCA between different ATAC-seq profiles shown in Fig. 1a. Motif analysis over the *Id2* region containing *LCR1* and *LCR2* was performed using HOMER.

**Intracellular cytokine staining.** Indicated cells were plated in 96-well round-bottom plates in RPMI (Gibco) containing FBS (10%), HEPES (20 mM, pH 7.2–7.5), sodium pyruvate (1X), glutamine (1X), penicillin/streptomycin (1X) and 2-mercaptoethanol (1X), and held at 37°C for 4 h. Media containing PMA (Sigma; final concentration, 100 ng ml<sup>-1</sup>) and Ionomycin (Sigma; final concentration, 10 ng ml<sup>-1</sup>) were then added to the appropriate wells, in addition to GolgiPlug (1X; BD Biosciences). Cytokines and transcription factor expression were measured by intracellular staining using the Foxp3 staining buffer (eBioscience).

**CD4<sup>+</sup> T cell purification, stimulation and flow cytometry.** Naive CD62L<sup>+</sup>CD4<sup>+</sup> T cells were obtained by negative selection using a Naive CD4<sup>+</sup> T cell purification kit (STEMCELL Technologies, 19765). Naive CD4<sup>+</sup> T cells were stimulated with plate-bound anti-CD3 antibody (BioLegend) with soluble anti-CD28 (BioLegend) under T<sub>H</sub>2 cell-polarizing medium containing 10 μg ml<sup>-1</sup> anti-IFN-γ, 50 ng ml<sup>-1</sup> IL-4 and 1 ng ml<sup>-1</sup> IL-2. Cytokines and transcription factor expression were measured by intracellular staining using the Foxp3 staining buffer. Antibodies were all purchased from eBioscience, BD Biosciences or BioLegend. All samples were acquired and analyzed with the LSR II flow cytometer (Becton Dickinson) and analyzed using FlowJo software (TreeStar).

**scRNA-seq.** Bone marrow progenitors (Live, CD45<sup>+</sup>, Lin<sup>-</sup>, NK1.1<sup>-</sup>, CD127<sup>+</sup>) were obtained by pooling *LCR1*<sup>+/+</sup> or *LCR1*<sup>-/-</sup> mice (equal number of males and females, between 8 and 12 weeks old). After cell sorting, cells were washed two

times with PBS containing 0.01% BSA (Sigma, A8806). Cell viability and cell counts were assessed using a hemacytometer and Trypan blue staining. Libraries were constructed using Chromium Controller and Chromium Single Cell 3' Library & Gel Bead Kit v3.1 (10x Genomics) according to the manufacturer's protocol for 10,000 cells recovery. In brief, cellular suspension was added to the master mix containing RT Reagent B, Template Switch Oligo, Reducing Agent B and RT Enzyme C. A total of 70 ml of master mix with cell suspension was then loaded to row 1, and rows 2 and 3 were filled with 50 ml of gel beads and 45 ml of partitioning oil, respectively. After run completion, 100 ml of GEM mix was used for reverse transcription and was performed using a C1000 Touch Thermal Cycler (Bio-Rad) at the following conditions: 53 °C for 45 min, 85 °C for 5 min, and held at 4 °C. Post-GEM-RT cleanup was performed with Dynabeads MyOne Silane beads (Thermo Fisher Scientific). Complementary DNA was amplified using the C1000 Touch Thermal Cycler at the following conditions: 98 °C for 3 min; 12 cycles of 90 °C for 15 s, 67 °C for 20 s and 72 °C for 1 min; 72 °C for 1 min; and held 4 °C. Amplified cDNA was cleaned with a SPRIselect Reagent Kit (Beckman Coulter), and quality was assessed using a TapeStation. Libraries were constructed following the manufacturer's protocol and sequenced with a targeted sequencing depth of 20,000 read pairs per cell on pair-end mode on either a NextSeq 500/550 or a NovaSeq with the recommended loading concentration (1.8 pM and 300 pM for NextSeq and NovaSeq, respectively).

**scRNA-seq normalization and clustering.** First, we filtered out low-quality or dead cells by excluding any cell that had less than 1,100 unique molecular identifiers (UMIs) and more than 8% of mitochondrial genes; we also excluded genes expressed in fewer than ten cells. Next, we processed the data following the steps recommended in the SCTransform tutorial<sup>65</sup>. In brief, we computed the percentage of mitochondrial genes (percent.mt) and the cell cycling score (S.score, g2m.score) of each cell, and then called the SCTransform function of the Seurat package<sup>66</sup>. For the SCTransform function we selected the parameters `method='GlmGamPoi'`, `var.to.regress=( 'percent.mt','S.score','g2m.score' )`.

Subsequently, we applied dimensional reduction using the RunPCA, RunUMAP and RunTSNE functions, and created a Shared Nearest Neighbors (SNN) graph using the FindNeighbors function with default parameters. Finally, we called the function FindClusters with the parameters `resolution=0.8`, `algorithm=4`, `method='igraph'` to obtain the clusters reported in this work.

**Differential expression analysis.** We created a catalog of differentially expressed genes for several comparisons between clusters and between conditions. To do so, we found marker genes of each cluster using the FindAllMarkers command in Seurat with the parameters `logfc.threshold=log(2)`, `only.pos=TRUE`. For differentially expressed genes both within clusters (between treat and control condition) and between clusters, we used the FindMarkers function with the parameters `logfc.threshold=log(2)`, `only.pos=TRUE`. Genes with adjusted  $P$  values  $< 10^{-5}$  and average  $\log_2$ [fold change]  $< 0.25$  were considered as differentially expressed in subsequent analysis. Pseudotime trajectories were calculated using the R package Monocle3 with default parameters<sup>67</sup>. We computed pseudotime trajectories independently for  $LCRI^{+/+}$  and  $LCRI^{-/-}$  scRNA-seq data. For this analysis, we selected a random cell in cluster 8 (a cluster enriched in markers of early progenitor cell populations) as the starting point of the trajectory. Additionally, to create the curves in Fig. 5b and Extended Data Fig. 6g, we used the `scipy.signals` package. We divided the pseudotime axis in 30 bins, computed mean expression levels per bin for each gene and applied a Savitzky–Golay smoothing over the resulting signal (`window_length=81`, `polyorder=3`).

**Gene ontology.** GO enrichment analysis was performed using the online tool PANTHER<sup>68</sup>. For this analysis, we compared  $LCRI^{+/+}$  vs  $LCRI^{-/-}$  expression profiles in cluster 7 (Fig. 5c). To do so, we divided the list of differentially expressed genes into upregulated and downregulated genes based on the sign of their fold change. Next, we passed each list separately to PANTHER (<http://geneontology.org>).

**Statistical analysis.** No statistical methods were used to predetermine sample sizes, but our sample sizes are similar to those reported in our previous publications, and data collection was appropriately blocked. For all experiments, data distribution was assumed to be normal, but this was not formally tested. Two-tailed Mann–Whitney  $U$  unpaired  $t$ -test was used for comparisons of two conditions within one group. Two-tailed one-way analysis of variance (ANOVA) was used to compare more than two conditions for one group. Two-tailed two-way ANOVA was used to compare two conditions across multiple groups. All graphs show the mean and s.e.m. One-way and two-way ANOVA were corrected for multiple comparison using Bonferroni correction. GraphPad Prism 9 (v9.4.0 or earlier versions) was used to calculate  $P$  values.

**Data collection and sample exclusion.** Data collection and analysis were not performed blind to the conditions of the experiments. We excluded two mice that did not show an increase of CD45<sup>+</sup> cells in the lung parenchyma of HDM-treated mice when compared with the PBS-treated mice. We corroborated our exclusion criteria by performing an Iglewicz and Hoaglin's outlier test with modified  $z$ -scores, which confirmed our exclusion rationale (modified  $z$ -score  $> 2$ ,  $P < 0.05$ ).

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

## Data availability

All bulk RNA-seq, ATAC-seq, Hi-C and scRNA-seq datasets have been uploaded to the GEO repository (accession no. GSE191312). The following published datasets were used: GEO GSE111871<sup>23</sup>, GEO GSE77695<sup>4</sup>, GEO GSE130775<sup>5</sup>, GEO GSE169542<sup>31</sup>, GEO GSE146745<sup>32</sup> and GEO GSE79422<sup>12</sup>. Source data are provided with this paper.

## References

- Laffont, S. et al. Androgen signaling negatively controls group 2 innate lymphoid cells. *J. Exp. Med.* **214**, 1581–1592 (2017).
- Servant, N. et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol.* **16**, 259 (2015).
- Abdennur, N. & Mirny, L. A. Cooler: scalable storage for Hi-C data and other genomically labeled arrays. *Bioinformatics* **36**, 311–316 (2020).
- Wolff, J. et al. Galaxy HiCExplorer 3: a web server for reproducible Hi-C, capture Hi-C and single-cell Hi-C data analysis, quality control and visualization. *Nucleic Acids Res.* **48**, W177–W184 (2020).
- Wolff, J. et al. Galaxy HiCExplorer: a web server for reproducible Hi-C data analysis, quality control and visualization. *Nucleic Acids Res.* **46**, W11–W16 (2018).
- Ramírez, F. et al. High-resolution TADs reveal DNA sequences underlying genome organization in flies. *Nat. Commun.* **9**, 189 (2018).
- Roayaei Ardakany, A., Gezer, H. T., Lonardi, S. & Ay, F. Mustache: multi-scale detection of chromatin loops from Hi-C and Micro-C maps using scale-space representation. *Genome Biol.* **21**, 256 (2020).
- Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
- Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
- Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013).
- Fasolino, M. et al. Genetic variation in type 1 diabetes reconfigures the 3D chromatin organization of T cells and alters gene expression. *Immunity* **52**, 257–274.e11 (2020).
- Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
- Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
- Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
- Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol.* **20**, 296 (2019).
- Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015).
- Trapnell, C. et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).
- Mi, H. et al. Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). *Nat. Protoc.* **14**, 703–721 (2019).

## Acknowledgements

The work in this manuscript was supported by funds from CHOP and the NIH grant R01HL136572 (J. H.-M.); the Burroughs Wellcome Fund, the Chang Zuckerberg Initiative Award, and the NIH grant R01AI168240 (J. H.-M. and G.V.); the PEW Charitable Trust (J. H.-M. and J.J. T.-C.); the NIH grants UC4 DK112217, U01DK11221702A1, R01HL145754, U01DK127768, U01DA052715 (G.V.); and F31DK122677 (M.T.J.).

## Author contributions

These studies were conceptualized by M.F.M., W.K.M., G.V. and J.H.-M. All experiments were designed, performed and analyzed by M.F.M. or W.K.M. with some help from M.L.C., M.T.J. and J.M.W. J.J.T.-C. performed all of the bioinformatic analysis with the help of S.Y. for the stripe analysis and of P.L., A.W. and C.A.T. for scRNA-seq analysis. A.C. generated Hi-C, RNA-seq and ATAC-seq libraries and quality controls. L.J. generated the novel mouse strains. J.H.-M. and G.V. supervised the biological and computational parts of this project, respectively. M.F.M., J.J.T.-C. and J.H.-M. wrote the manuscript. G.V. edited the manuscript.

## Competing interests

The authors declare no competing interests.

**Additional information**

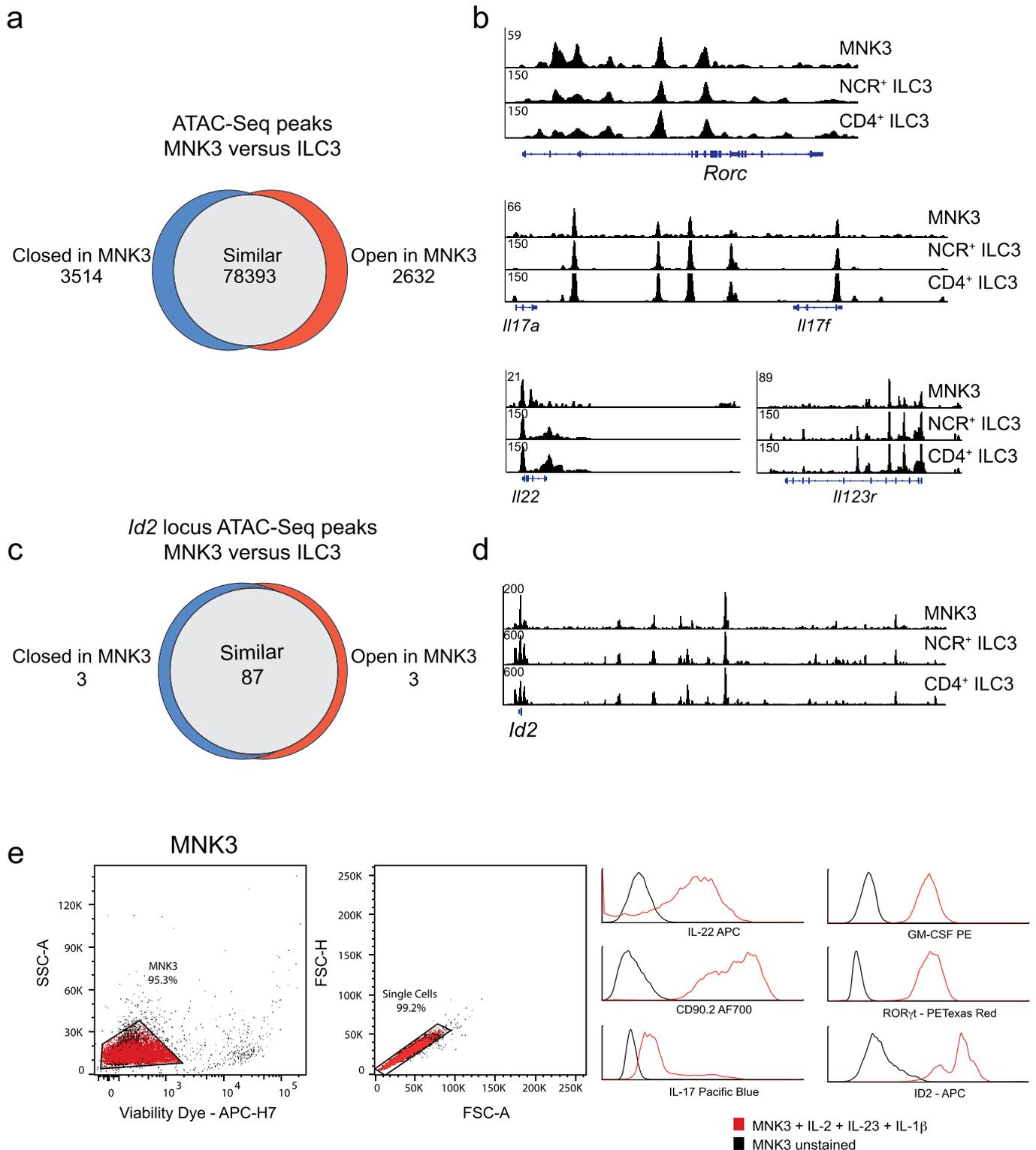
**Extended data** is available for this paper at <https://doi.org/10.1038/s41590-022-01295-y>.

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41590-022-01295-y>.

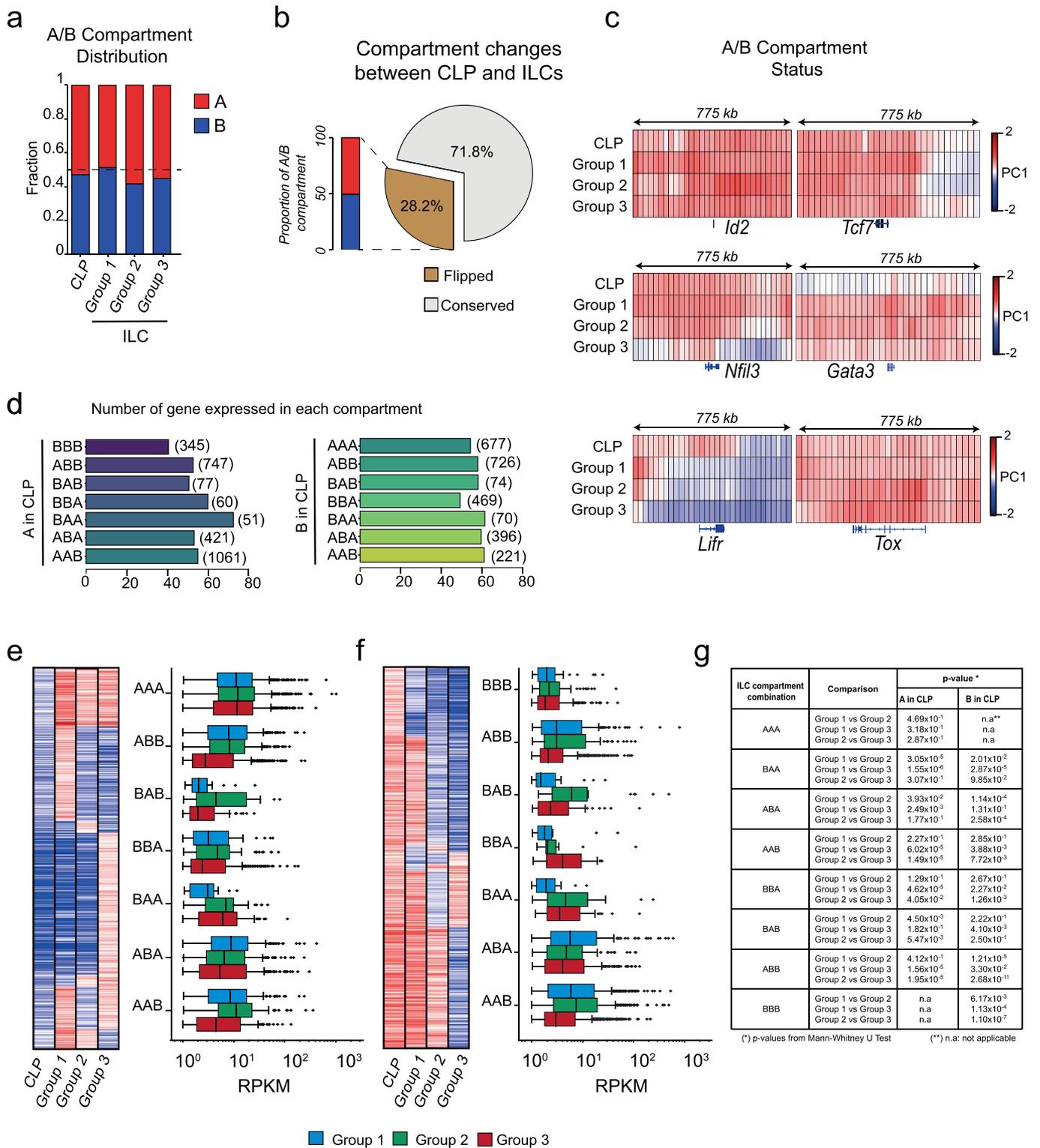
**Correspondence and requests for materials** should be addressed to Golnaz Vahedi or Jorge Henao-Mejia.

**Peer review information** *Nature Immunology* thanks Shigeo Koyasu and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available. Primary Handling Editor: L.A. Dempsey, in collaboration with the *Nature Immunology* team.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

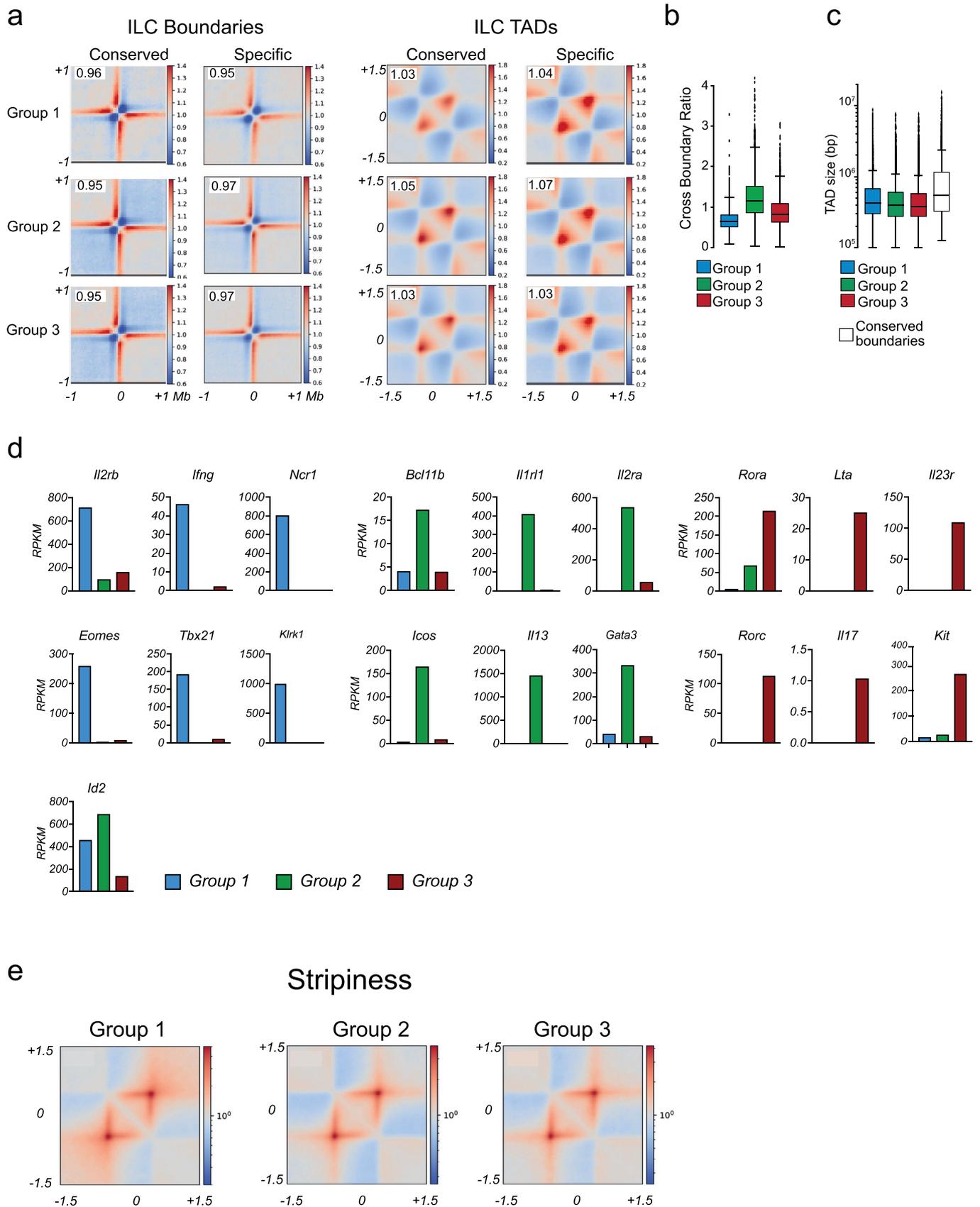


**Extended Data Fig. 1 | Comparison of accessibility between the MNK3 cell line and primary ILC3.** **a**, Venn diagram representation of the differential chromatin accessibility between ATAC-seq peaks of MNK3 and small intestine (SI) CD4<sup>+</sup> ILC3, and SI NCR<sup>+</sup> ILC3 cells. Absolute numbers of peaks differentially open (red) or closed (blue) in MNK3, or with similar accessibility (gray) in all samples, are represented. **b**, View of the chromatin accessibility (ATAC-seq signal) at the *Rorc*, *Il17a*, *Il17f*, *Il22* and *Il23r* loci in small intestine (SI) CD4<sup>+</sup> ILC3, SI NCR<sup>+</sup> ILC3, and MNK3 cells. **c**, Venn diagram representation of the differential accessibility analysis results in the region displayed in the right panel. **d**, View of the chromatin accessibility (ATAC-seq signal) at the *Id2* locus in SI CD4<sup>+</sup> ILC3, SI NCR<sup>+</sup> ILC3, and MNK3 cells. Absolute numbers of peaks differentially open (red) or closed (blue) in MNK3, or with similar accessibility (gray) in all samples, are shown. **e**, Representative flow cytometry plots showing cell viability and the expression of ILC3-specific markers (IL-22, GM-CSF, CD90.2, RORyt, IL-17 and ID2) by MNK3 cells after stimulation with IL-2, IL-23, and IL-1 $\beta$ . The red line represents activated MNK3 cells and the black line represents unstained MNK3 cells.



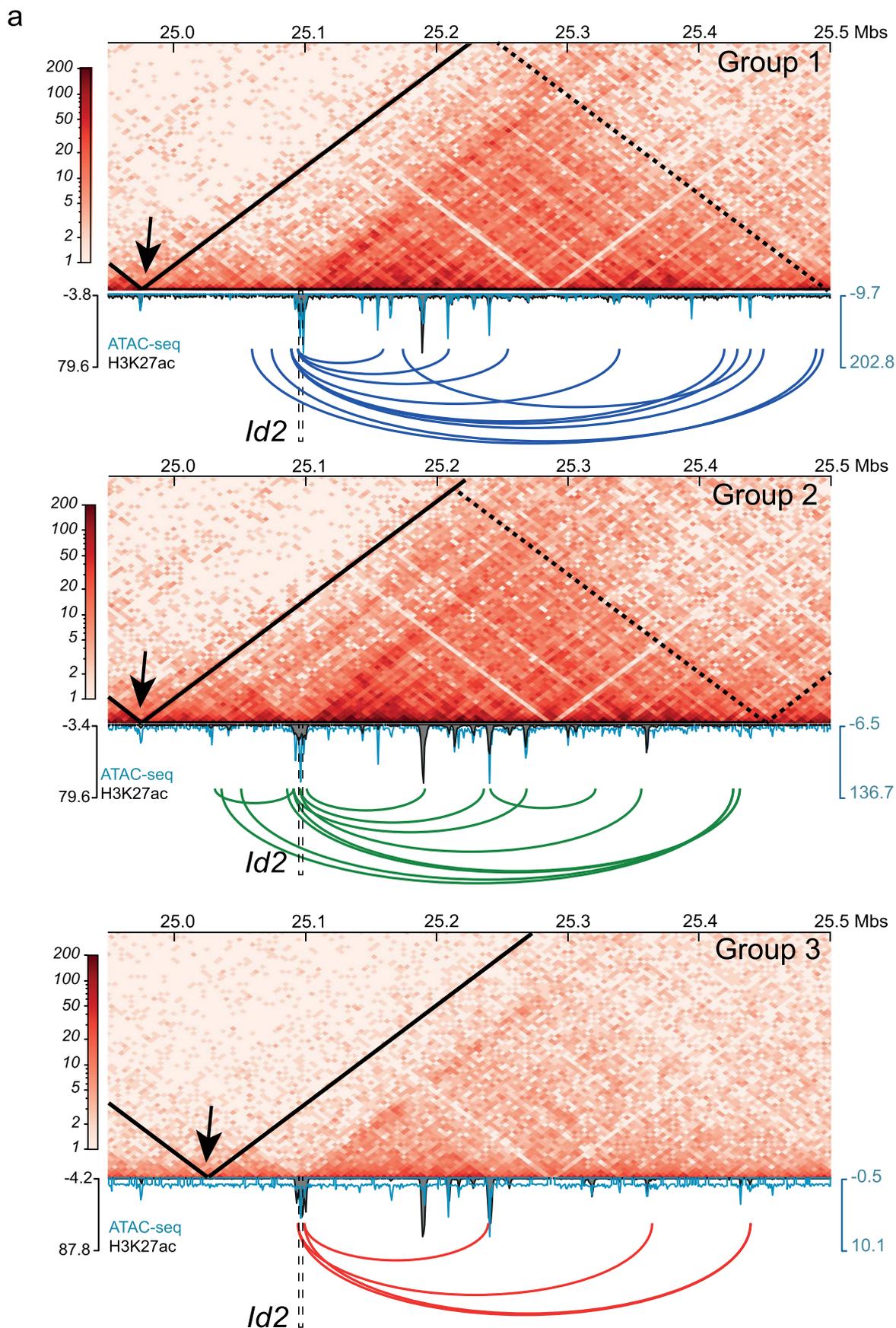
Extended Data Fig. 2 | See next page for caption.

**Extended Data Fig. 2 | Changes in compartment distribution in ILCs are accompanied by changes in expression patterns concordant with euchromatin/heterochromatin state.** **a**, Fraction of regions in the A or B compartments in common lymphoid progenitors (CLP), group 1 (splenic NK cells), group 2 (bone marrow ILC2s), and group 3 (MINK3) ILC. **b**, Proportion of regions in CLPs that flipped compartment status in at least one ILC group (brown) or that remained unchanged in all ILC groups (gray). Sidebar indicates the proportion of flipped regions that were in the A (red) or B (blue) compartment at the CLP stage. **c**, Examples of compartment state distribution in the vicinity of genes with known roles in ILC biology such as *Id2*, *Tcf7*, *Nfil3*, *Gata3*, *Lifr* and *Tox*. Regions in the A and B compartment correspond to positive (red) and negative (blue) values in the PC1 bar, respectively. Each square represents 25kb. **d**, Percentage of expressed genes among genes within compartments that underwent similar flipping from CLP. The Y-axis represents the different combinations of final compartment states in ILC for flipping regions originally in the status indicated on the left. Parenthetical numbers indicate the total number of genes overlapping the corresponding group of regions. Expressed genes correspond to genes with average RPKM > 1. RPKM: Reads per kilobase of transcripts per million mapped reads. Letter triplets represent the compartment status in group 1, group 2 and group 3 ILC, respectively. **e**, **(left)** Heatmap of the compartment status in ILCs of flipping regions that were in the B compartment in CLPs. Regions in the A and B compartment correspond to positive (red) and negative (blue) values in the PC1 bar. **(right)** Comparison of gene expression levels between group 1 (blue boxes), group 2 (green boxes) and group 3 (red boxes) ILCs of groups of genes undergoing similar compartment flipping from CLP. The Y-axis represents the different combinations of final compartment states in ILC. RPKM: Reads per kilobase of transcripts per million mapped reads. Letter triplets represent the compartment status in group 1, group 2 and group 3 ILC, respectively. Box shows dataset quartiles and whiskers the distribution range. Dots represent outliers as determined by the inter-quartile range. **f**, **(left)** Heatmap of the compartment status in ILCs of flipping regions that were in the A compartment in CLP. Regions in the A and B compartment correspond to positive (red) and negative (blue) values in the PC1 bar. **(right)** Comparison of gene expression levels between group 1 (blue boxes), group 2 (green boxes) and group 3 (red boxes) ILCs of groups of genes undergoing similar compartment flipping from CLP. The Y-axis represents the different combinations of final compartment states in ILC. RPKM: Reads per kilobase of transcripts per million mapped reads. Letter triplets represent the compartment status in group 1, group 2 and group 3 ILC, respectively. Box shows dataset quartiles and whiskers the distribution range. Dots represent outliers as determined by the inter-quartile range. **g**, Table with the p-values (Mann-Whitney U two-tailed test) for Fig. 1 **e,f** of the pairwise comparison between ILC of expression levels of genes that underwent similar compartment flippings from the A or B compartment in CLP. Letter triplets represent the compartment status in group 1, group 2 and group 3 ILC, respectively.



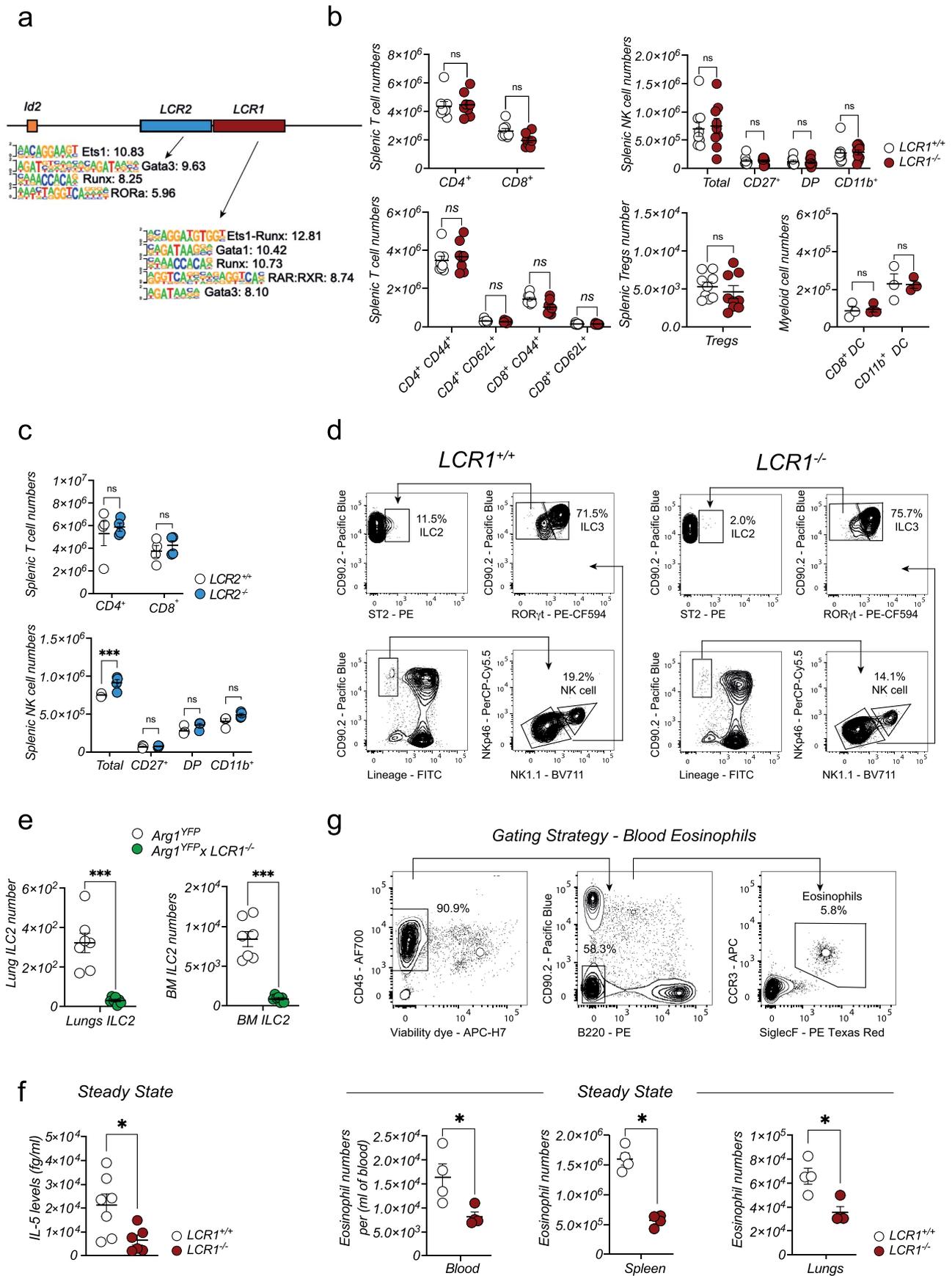
Extended Data Fig. 3 | See next page for caption.

**Extended Data Fig. 3 | Local structures of the 3D genome architecture are comparable at the sub-megabase scale and correlate with gene expression in each ILC group.** **a**, Comparison of the pileup plots computed over conserved vs ILC-specific (**left**) boundary elements and (**right**) Topological Associated Domains (TADs) in each group of ILCs. Numbers in the corner indicate average signal strength. **b**, Boxplot of cross-boundary ratios of TADs for group 1 (blue), group 2 (green) and group 3 (red) ILCs. Box shows dataset quartiles and whiskers the distribution range. Dots represent outliers as determined by the inter-quartile range. **c**, Boxplot of the size distribution of TADs defined by boundaries detected in each ILC (blue; group 1 ILC, green; group 2 ILC, red; group 3 ILC) and by boundaries conserved across ILCs (white). Box shows dataset quartiles and whiskers the distribution range. Dots represent outliers as determined by the inter-quartile range. **d**, Expression levels of genes with critical roles in ILC biology in group 1 (blue bars), group 2 (green bars) and group 3 (red bars) ILCs obtained by RNA-sequencing. RPKM: Reads per kilobase of transcripts per million mapped reads. **e**, Pileup plots of stripes detected in group 1 (**left**), group 2 (**middle**) and group 3 ILC (**right**). Axes in units of mean feature (stripes) size as computed for each group ILC.



Extended Data Fig. 4 | See next page for caption.

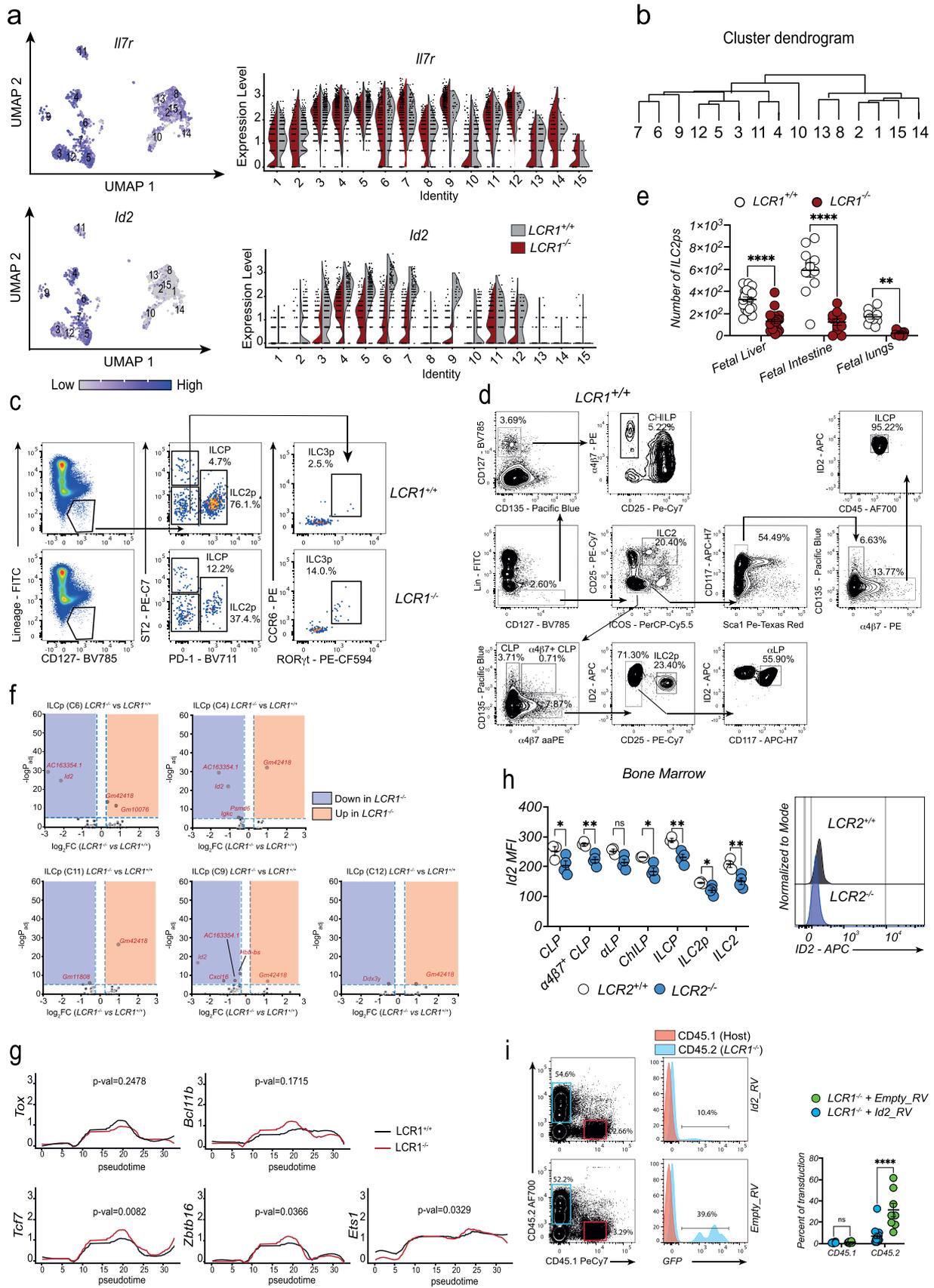
**Extended Data Fig. 4 | 3D architecture, chromatin accessibility, and enhancer activity patterns at the *Id2* locus in each ILC group. a.** Visualization of the contact heatmap at the *Id2* locus in (**top**) group 1, (**middle**) group 2, and (**bottom**) group 3 ILCs. Chromatin accessible regions determined by ATAC-seq (gray track), and H3K27ac deposition (blue track) determined by ChIP-seq (group 1 and group 2) or Cut & Run (group 3) in each group of ILCs. Colored arcs represent loops detected in the corresponding contact map (blue; group 1 ILC, green; group 2 ILC, red; group 3 ILC). The location of the *Id2* promoter is indicated with an extended vertical dashed line. Dotted and solid lines represent sub-TAD and TAD boundaries, respectively. Arrowheads indicate the position of boundaries conserved across all ILCs. Scale bar in heatmap represents normalized contact frequency.



Extended Data Fig. 5 | See next page for caption.

**Extended Data Fig. 5 | The locus control region 1 (LCR1), but not LCR2, is specifically required for the development of all tissue resident ILC2s.**

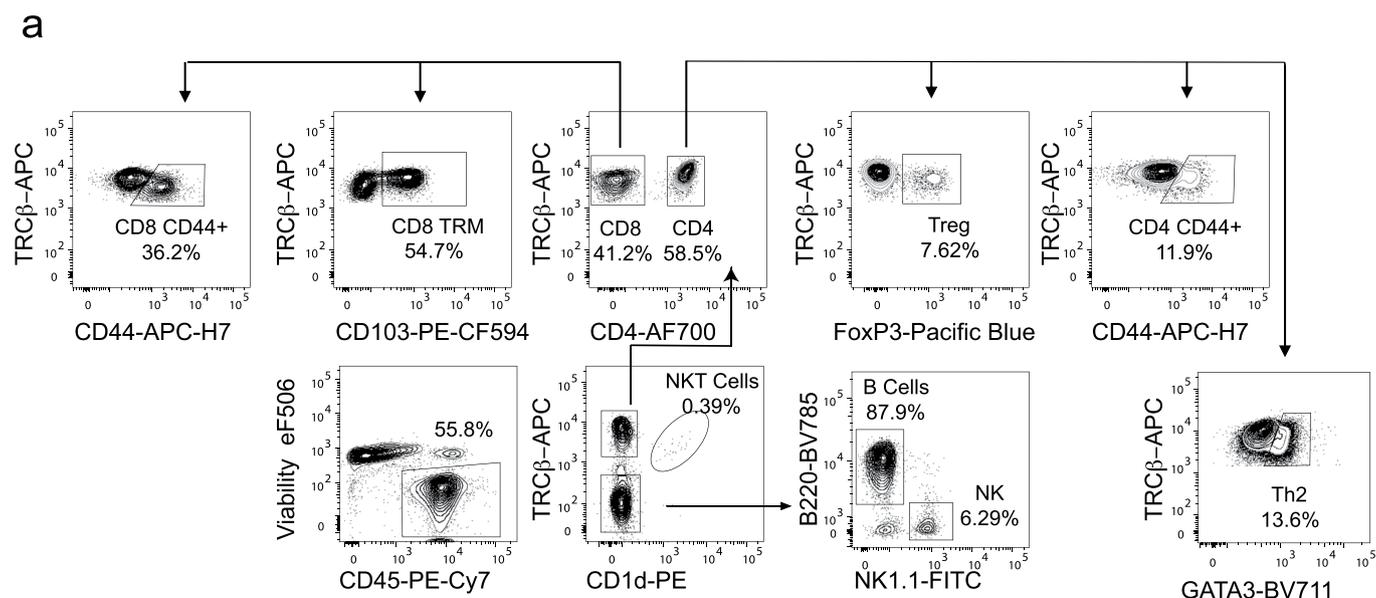
**a.** Representative transcription factor binding motifs enriched in ATAC-seq peaks (peaks from all ILCs pooled) within *LCR1* (red section) and *LCR2* (blue section). Numbers represent log odds detection scores as computed by HOMER. **b. (top)** Quantification of absolute numbers of splenic CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells, and NK cells in their different maturation states in *LCR1*<sup>+/+</sup> and *LCR1*<sup>-/-</sup> mice. **(bottom)** Quantification of absolute numbers of splenic naive (CD62L<sup>+</sup>, CD44<sup>-</sup>) and effector (CD62L<sup>-</sup>, CD44<sup>+</sup>) CD4<sup>+</sup> and CD8<sup>+</sup> T cell populations, as well as CD8<sup>+</sup> and CD11b<sup>+</sup> dendritic cells. Splenic T cell numbers are one representative experiment that was repeated six times (n = 8 *LCR1*<sup>+/+</sup> and 8 *LCR1*<sup>-/-</sup>). Splenic NK cell numbers are a pool of two independent experiments. This experiment was repeated four (n = 9 *LCR1*<sup>+/+</sup> and 9 *LCR1*<sup>-/-</sup>) and was repeated three times. Tregs represents a pool of two independent experiments (n = 9 *LCR1*<sup>+/+</sup> and 9 *LCR1*<sup>-/-</sup>) and was repeated three times. Myeloid cell numbers represent one experiment (n = 3 *LCR1*<sup>+/+</sup> and 3 *LCR1*<sup>-/-</sup>). Error bars = SEM; and p-values: ns = not significant (Treg panel: Mann-Whitney U two-tailed test; Splenic T cells NK cell and Myeloid cell: Two-way ANOVA with multiple comparison and Bonferroni correction). **c.** Quantification of absolute numbers of splenic CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells, and NK cells in their different maturation state in *LCR2*<sup>+/+</sup> and *LCR2*<sup>-/-</sup> mice (n = 4 *LCR2*<sup>+/+</sup> and 4 *LCR2*<sup>-/-</sup>). Data represents one experiment that was repeated two times. Error bars = SEM; and p-values: ns = not significant, \* = p ≤ 0.05, \*\* = p ≤ 0.01, \*\*\* = p ≤ 0.0005 (Two-way ANOVA with multiple comparison and Bonferroni correction). **d.** Representative flow cytometry gating strategy used for the identification of group 1 (NK cell), group 2 (ILC2s), and group 3 (ILC3s) ILC in small intestine lamina propria (siLP) of *LCR1*<sup>+/+</sup> and *LCR1*<sup>-/-</sup> animals. **e.** Quantification of lungs and bone marrow (BM) ILC2s numbers of Arginase1 reporter mouse (*Arg1*<sup>YFP</sup>) and *Arg1*<sup>YFP</sup>; *LCR1*<sup>-/-</sup> mice (Live, CD45<sup>+</sup>, Lineage<sup>-</sup>, CD127<sup>+</sup>, CD90.2<sup>+</sup>, ST2<sup>+</sup>, YFP<sup>+</sup>) at steady state. Data is a pool of two independent experiments and were repeated two times. Each dot represents an individual mouse (n = 7 *Arg1*<sup>YFP</sup>; *LCR1*<sup>+/+</sup> and 7 *Arg1*<sup>YFP</sup>; *LCR1*<sup>-/-</sup>). Error bars = SEM; and p-values: ns = not significant, \* = p ≤ 0.05, \*\* = p ≤ 0.01, \*\*\* = p ≤ 0.0005 (Mann-Whitney U two-tailed test). **f.** Quantification of IL-5 in the serum of *LCR1*<sup>+/+</sup> and *LCR1*<sup>-/-</sup> mice at steady state by cytometry bead assay (CBA). Data is representative of two experiments. Each dot represents an individual mouse, (n = 7 *LCR1*<sup>+/+</sup> and 6 *LCR1*<sup>-/-</sup>). Error bars = SEM; and p-values: ns = not significant, \* = p ≤ 0.05 (Mann-Whitney U two-tailed test). **g. (top)** Representative flow cytometry gating strategy used for the identification eosinophils in the blood, spleen, and lungs of *LCR1*<sup>+/+</sup> and *LCR1*<sup>-/-</sup> mice at steady state and **(bottom)** their subsequent quantification. Data represents one experiment. Each dot represents an individual mouse, (n = 4 *LCR1*<sup>+/+</sup> and 4 *LCR1*<sup>-/-</sup>). Error bars = SEM; and p-values: ns = not significant, \* = p ≤ 0.05 (Mann-Whitney U two-tailed test).



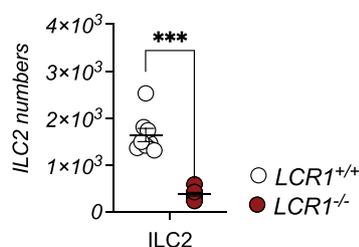
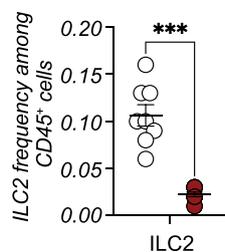
Extended Data Fig. 6 | See next page for caption.

**Extended Data Fig. 6 | The *LCR1* impairs ILC2 development through reduction of *Id2* expression and skewing toward ILC1- and ILC3-like ILC progenitors.**

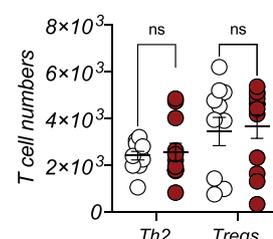
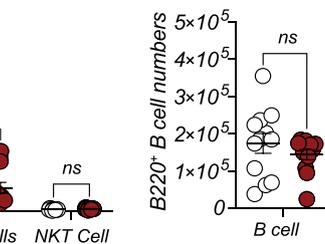
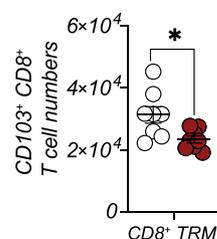
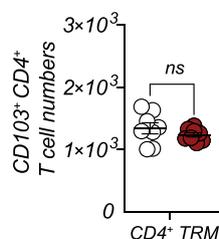
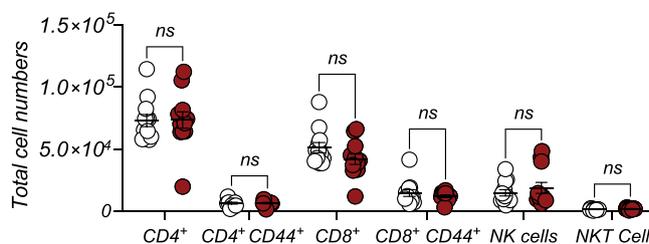
**a, (left)** Expression levels of *Il7r* and *Id2* over the Uniform Manifold Approximation and Projection of the scRNA-seq data set described in Fig. 4a. **(right)**, Violin plot representation of the expression levels of *Id2* and *Il7r* across all identified clusters in Fig. 4a. For each cluster (horizontal axis), the expression levels in *LCR1*<sup>+/+</sup> (gray) and *LCR1*<sup>-/-</sup> (red) mice is shown. **b**, Hierarchical representation of similarity between the clusters identified in Fig. 4a. Similarity was measured using complete linkage over the Pearson correlation matrix between clusters. **c**, Gating strategy used to identify putative ILC3p in the bone marrow of *LCR1*<sup>+/+</sup> and *LCR1*<sup>-/-</sup> animals. Lin<sup>-</sup>, CD127<sup>+</sup> cells were pre-gated on Singlets, Live, CD45<sup>+</sup> cells. **d**, Gating strategy for the identification of CLP: Common Lymphoid Progenitors (Live, CD45<sup>+</sup>, Lin<sup>-</sup>, CD127<sup>+</sup>, CD25<sup>-</sup>, ICOS<sup>-</sup>, CD135<sup>+</sup>, α4β7<sup>-</sup>); α-LP: alpha-Lymphoid Progenitors (Live, CD45<sup>+</sup>, Lin<sup>-</sup>, CD127<sup>+</sup>, ICOS<sup>-</sup>, CD135<sup>+</sup>, α4β7<sup>+</sup>, Id2<sup>+</sup>, CD25<sup>-</sup>, CD117<sup>+</sup>); ChILP: Common helper Innate Lymphoid Progenitors (Live, CD45<sup>+</sup>, Lin<sup>-</sup>, CD127<sup>+</sup>, CD135<sup>-</sup>, α4β7<sup>+</sup>, CD25<sup>-</sup>); ILCp: Innate Lymphoid Cell precursors (Live, CD45<sup>+</sup>, Lin<sup>-</sup>, CD127<sup>+</sup>, ICOS<sup>-</sup>, CD25<sup>-</sup>, CD117<sup>+</sup>, Sca1<sup>-</sup>, α4β7<sup>+</sup>, ID2<sup>+</sup>); ILC2p: group 2 Innate Lymphoid Cell precursors (Live, CD45<sup>+</sup>, Lin<sup>-</sup>, CD127<sup>+</sup>, CD90.2<sup>+</sup>, ST2<sup>+</sup> or CD25<sup>+</sup>, ICOS<sup>+</sup>). **e**, Quantification of ILC2 progenitors in the fetal liver (FL), fetal intestine (FI) or fetal lungs (FLu) of *LCR1*<sup>+/+</sup> and *LCR1*<sup>-/-</sup> E15.5 embryos. Data represents the pool of two independent experiments. Each dot represents an individual embryo (numbers: FL: = 21 *LCR1*<sup>+/+</sup> and 20 *LCR1*<sup>-/-</sup>; FI and FLu: 11 *LCR1*<sup>+/+</sup> and 10 *LCR1*<sup>-/-</sup>). Error bars = SEM; and p-values: ns = not significant, \* = p ≤ 0.05, \*\* = p ≤ 0.01, \*\*\* = p ≤ 0.0005, \*\*\*\* = p ≤ 0.0001 (Mann-Whitney U two-tailed test). **f**, Volcano plot displaying upregulated (red rectangle) and downregulated (blue rectangle) genes in *LCR1*<sup>+/+</sup> when compared to *LCR1*<sup>-/-</sup> cells for clusters in Fig. 5a representing ILCp populations. **g**, Savitzky-Golay smoothing curve of the normalized expression levels of *Tox*, *Bcl11b*, *Ets1*, *Tcf7*, and *Zbtb16* along the pseudotime axis determined in Fig. 5e. P-value corresponds to Mann-Whitney U test between smoothed signals from *LCR1*<sup>+/+</sup> (black curve) and *LCR1*<sup>-/-</sup> (red curve) mice. **h, (left)** Quantification of ID2 mean fluorescence intensity (MFI) in *LCR2*<sup>+/+</sup> and *LCR2*<sup>-/-</sup> bone marrow (BM) CLP, a4b7<sup>+</sup> CLP, aLP, ChILP, ILCp, ILC2p gated as shown in Extended Data Fig. 6d. **(right)** Representative histogram of ID2 expression in *LCR2*<sup>+/+</sup> and *LCR2*<sup>-/-</sup> BM. Each dot represents an individual mouse (n = 3 *LCR2*<sup>+/+</sup> and 4 *LCR2*<sup>-/-</sup>) and was observed in 2 independent experiments. Error bars = SEM; and p-values: ns = not significant, \* = p ≤ 0.05, \*\* = p ≤ 0.01 (Two-way ANOVA with multiple comparison and Bonferroni correction). **i**, Representative FACS plots and histograms of GFP expression in wild-type (CD45.1<sup>+</sup>) or *LCR1*<sup>-/-</sup> CD45.2 BM previously transduced with retroviral particles encoding an empty vector expressing the GFP (*empty\_RV* - green) or retroviral particles encoding *Id2* and the GFP (*Id2\_RV* - blue), at least eight weeks post engraftment in CD45.1<sup>+</sup> hosts. Each dot represents an individual mouse (n = 8 *LCR1*<sup>-/-</sup> *empty\_RV*; 12 *LCR1*<sup>-/-</sup> *Id2\_RV*). Error bars = SEM; and p-values: ns = not significant, \* = p ≤ 0.05, \*\* = p ≤ 0.01 (Two-way ANOVA with multiple comparison and Bonferroni correction).



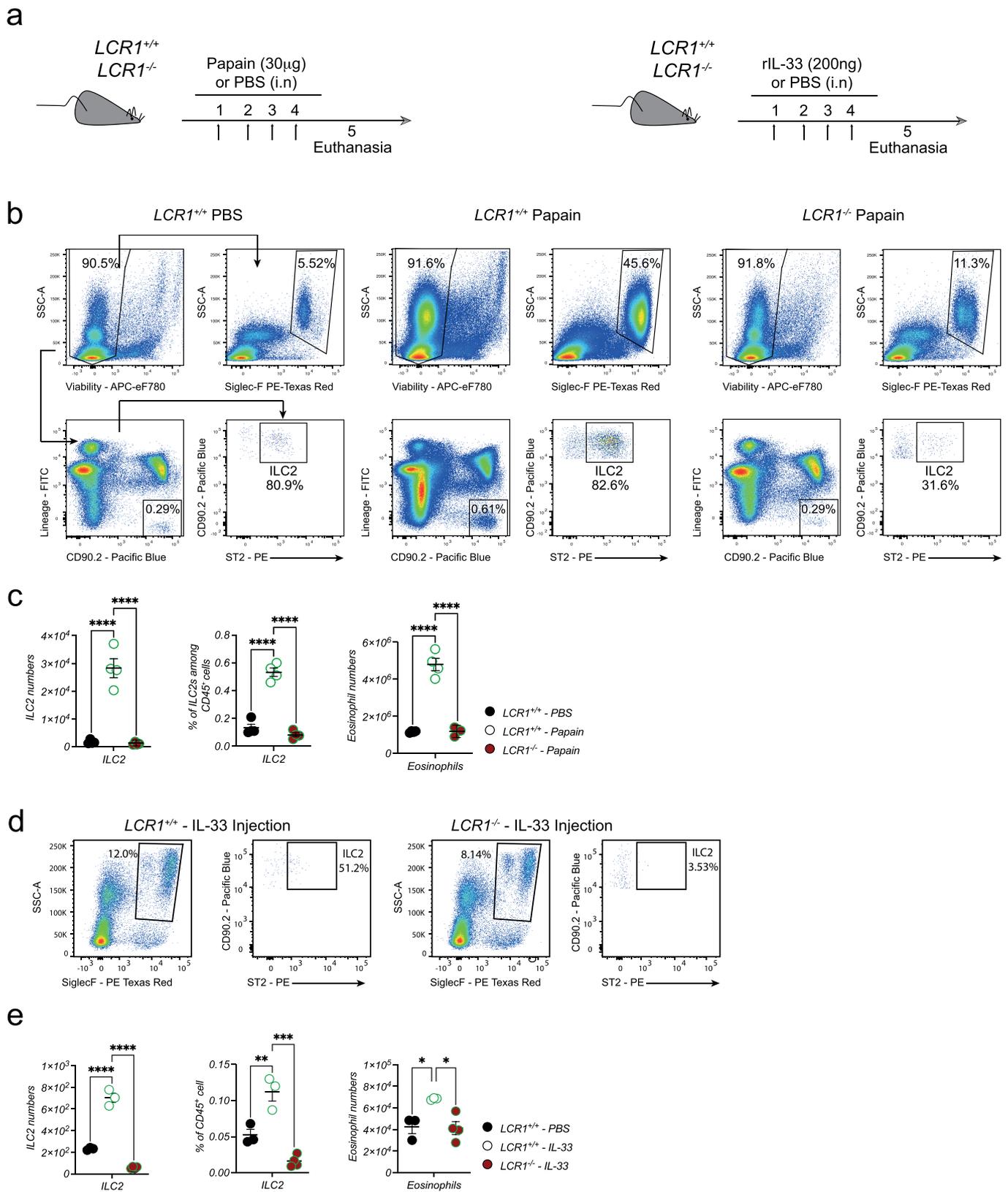
**b** Lung parenchyma steady state



**c** Lung parenchyma steady state

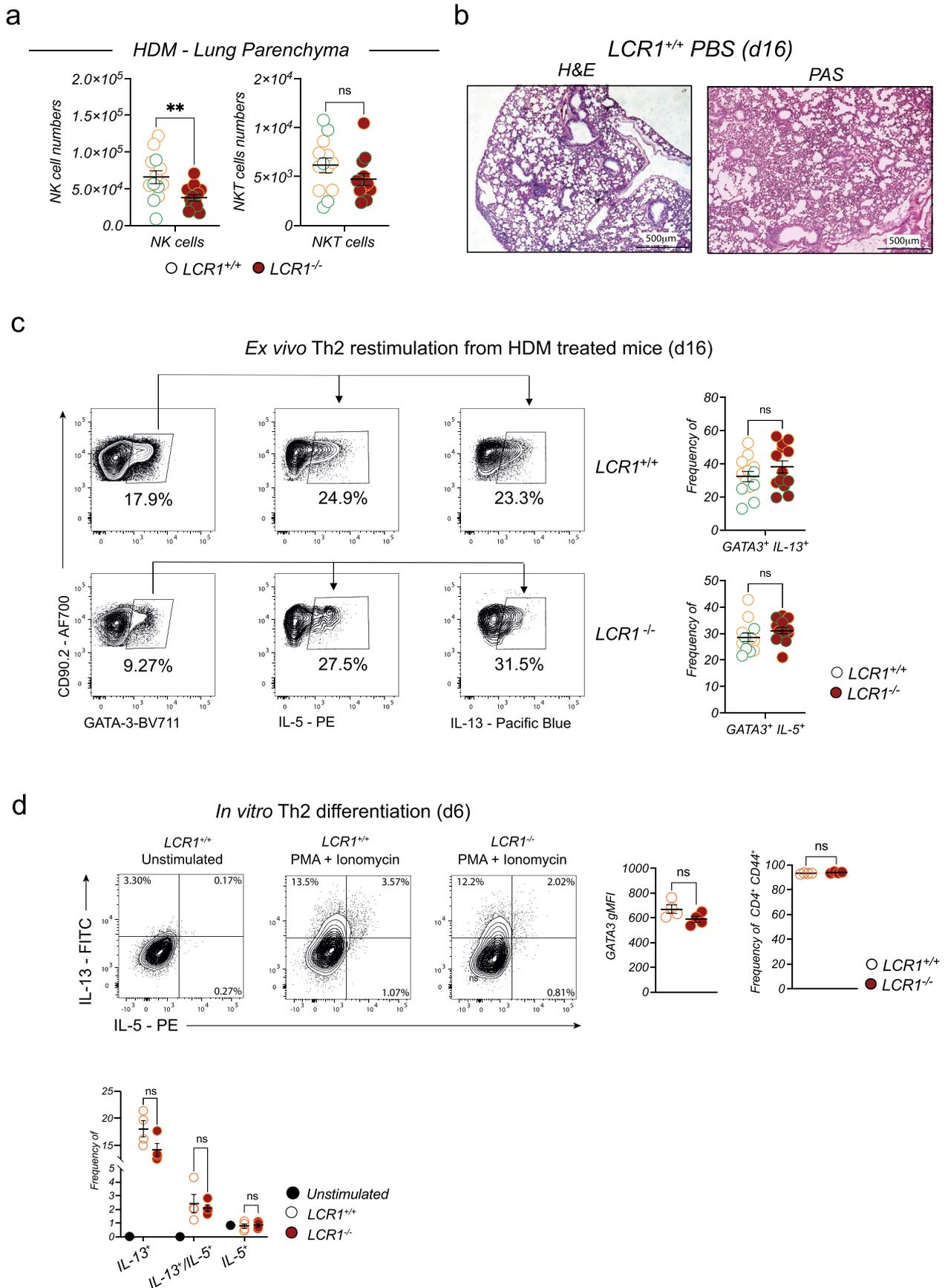


**Extended Data Fig. 7 | Disruption of the 3D architecture of the *Id2* locus reduces ILC2 numbers in the lung but does not drastically change other immune cell populations at steady state.** **a**, Representative flow cytometry gating strategy used for the identification of immune cell populations in digested lungs or in bronchoalveolar lavage fluid (BALF) at steady state or after 16 days of house dust mite (HDM) challenge in *LCR1*<sup>+/+</sup> and *LCR1*<sup>-/-</sup> animals. **b**, Quantification of frequencies among CD45<sup>+</sup> cells and absolute numbers of lung parenchyma ILC2s in *LCR1*<sup>+/+</sup> and *LCR1*<sup>-/-</sup> mice at steady state, (n=8 *LCR1*<sup>+/+</sup> and 8 *LCR1*<sup>-/-</sup>). Data are representative of one experiment that was repeated five times. Error bars = SEM; and p-values: ns = not significant, \* = p ≤ 0.05, \*\* = p ≤ 0.01, \*\*\* = p ≤ 0.0005 (Mann-Whitney U two-tailed test). **c**, Quantification of total number of cells of the indicated immune cell populations in Extended Data Fig. 7a in the lung parenchyma of *LCR1*<sup>+/+</sup> and *LCR1*<sup>-/-</sup> mice at steady state, (n=12 *LCR1*<sup>+/+</sup> and 12 *LCR1*<sup>-/-</sup>). Data are a pool of two independent experiments, except CD4<sup>+</sup> and CD8<sup>+</sup> TRM which is a representative experiment. (n=8 *LCR1*<sup>+/+</sup> and 8 *LCR1*<sup>-/-</sup>). Error bars = SEM; and p-values: ns = not significant, \* = p ≤ 0.05 (Mann-Whitney U two-tailed test or Two-way ANOVA with multiple comparison and Bonferroni correction).



Extended Data Fig. 8 | See next page for caption.

**Extended Data Fig. 8 | Acute intranasal papain and IL-33 challenges in LCR1-deficient mice.** **a**, Schematic representation of (**left**) the experimental acute papain or (**right**) recombinant IL-33 (rIL-33) challenges. Mice were challenged intranasally with either 30mg of papain (in 20mL of PBS) or with 200ng of rIL-33 (in 20ml of PBS) for four consecutive days. Euthanasia was carried out 24 hours after the last challenge. Lung parenchyma ILC2s and eosinophil numbers were assessed by flow cytometry. **b**, Flow cytometry gating strategy used for the identification of lung parenchyma ILC2s and eosinophils in *LCR1*<sup>+/+</sup> and *LCR1*<sup>-/-</sup> mice challenged with papain or *LCR1*<sup>+/+</sup> mice treated with PBS (20mL). Cells were pre-gated on SSC/FSC and Singlets. **c**, Quantification of ILC2 frequency among CD45<sup>+</sup> cells, ILC2 and eosinophil numbers in *LCR1*<sup>+/+</sup> and *LCR1*<sup>-/-</sup> mice challenged with papain, or *LCR1*<sup>+/+</sup> mice treated with PBS (20mL). Data represents one experiment with males (green dots; n = 4 *LCR1*<sup>+/+</sup> PBS, 4 *LCR1*<sup>+/+</sup>, and 4 *LCR1*<sup>-/-</sup> papain - 11.4-12.2-week-old). Error bars = SEM; and p-values: ns = not significant, \* = p ≤ 0.05, \*\* = p ≤ 0.01, \*\*\* = p ≤ 0.0005, \*\*\*\* = p ≤ 0.0001 (One-way ANOVA with multiple comparisons and Bonferroni correction). **d**, Flow cytometry gating strategy used for the identification of lung parenchyma ILC2s and eosinophils in *LCR1*<sup>+/+</sup> and *LCR1*<sup>-/-</sup> mice challenged with rIL-33 (200ng) or *LCR1*<sup>+/+</sup> mice treated with PBS (20mL). Cells were pre-gated on SSC/FSC and Singlets. **e**, Quantification of ILC2 frequency among CD45<sup>+</sup> cells, ILC2 and Eosinophil numbers in *LCR1*<sup>+/+</sup> and *LCR1*<sup>-/-</sup> mice challenged with rIL-33 (200ng), or *LCR1*<sup>+/+</sup> mice treated with PBS (20mL). Data is representative of two independent experiments (green dots; n = 3 *LCR1*<sup>+/+</sup> PBS, 3 *LCR1*<sup>+/+</sup> rIL-33, and 4 *LCR1*<sup>-/-</sup> rIL-33 - 9.4-9.9-week-old) and was repeated two times. Error bars = SEM; and p-values: ns = not significant, \* = p ≤ 0.05, \*\* = p ≤ 0.01, \*\*\* = p ≤ 0.0005, \*\*\*\* = p ≤ 0.0001 (One-way ANOVA with multiple comparisons and Bonferroni correction).



Extended Data Fig. 9 | See next page for caption.

**Extended Data Fig. 9 | Disruption of the 3D architecture of the *Id2* locus in ILC2 reduces HDM-induced allergic airway inflammation progression.**

**a**, Quantification of NK and NKT cells numbers in the lungs of *LCR1<sup>+/+</sup>* and *LCR1<sup>-/-</sup>* mice 16 days after the initial HDM challenge. Data is a pool of two age and sex-matched independent experiments containing males (green dots;  $n = 5$  *LCR1<sup>+/+</sup>* and 5 *LCR1<sup>-/-</sup>* - age 12.9-week-old) and females (orange dots,  $n = 8$  *LCR1<sup>+/+</sup>* and 8 *LCR1<sup>-/-</sup>* - age 6.9-week-old). The experiments were repeated three times. Error bars = SEM; and p-values: ns = not significant, \* =  $p \leq 0.05$ , \*\* =  $p \leq 0.01$  (Mann-Whitney U two-tailed test). **b**, Representative histological sections of lungs stained with hematoxylin and eosin (**H&E - top**) or Periodic Shift Acid (**PAS - bottom**) of 1 *LCR1<sup>+/+</sup>* mice treated with PBS. The scale represents a length of 500 $\mu$ m and the magnification used is 40x. **c**, Representative flow cytometry plot (**left**) and quantification (**right**) of IL-5<sup>+</sup> and IL-13<sup>+</sup> producing CD4<sup>+</sup> Th2 cells (CD4<sup>+</sup>, GATA3<sup>+</sup>) from lung parenchyma. Cells were *ex vivo* stimulated with PMA and Ionomycin for 4 hours in presence of Golgi inhibitor and identified as Live, CD45<sup>+</sup>, TCR $\beta$ <sup>+</sup>, CD4<sup>+</sup>, GATA-3<sup>+</sup> cells. Data is a pool of two age and sex-matched independent experiments containing males (green dots;  $n = 5$  *LCR1<sup>+/+</sup>* and 5 *LCR1<sup>-/-</sup>* - age 12.9-week-old) and females (orange dots,  $n = 8$  *LCR1<sup>+/+</sup>* and 8 *LCR1<sup>-/-</sup>* - age 6.9 week-old). The experiment was repeated three times. Error bars = SEM; and p-values: ns = not significant (Mann-Whitney U two-tailed test). **d**, (**top left**) Representative flow cytometry plots of IL-13<sup>+</sup> and IL-5<sup>+</sup> producing naive CD4<sup>+</sup> T cells from *LCR1<sup>+/+</sup>* and *LCR1<sup>-/-</sup>* mice polarized under Th2 conditions for 6 days. (**Bottom left**) Proportion of IL-13<sup>+</sup>, IL-5<sup>+</sup> and IL-13<sup>+</sup>/IL-5<sup>+</sup> producing Th2 cells. (**Top right**), Quantification of the geometric mean fluorescence intensity (gMFI) of GATA-3, and frequency of activated (CD44<sup>+</sup>) CD4<sup>+</sup> Th2 cells at day 6 after polarization. Data is representative of three independent experiments. Each dot represents an individual mouse, (Females;  $n = 4$  *LCR1<sup>+/+</sup>* and 4 *LCR1<sup>-/-</sup>*). Error bars = SEM; and p-values: ns = not significant (Mann-Whitney U two-tailed test).

## Reporting Summary

Nature Portfolio 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 Portfolio policies, see our [Editorial Policies](#) 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.

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

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

#### Data collection

Hi-C analysis: Hi-C libraries were generated on up to 106 cells using Arima-Hi-C kit (Arima Genomics) and Accel-NGS Plus DNA Library kit (21024 Swift Biosciences), according to the manufacturer's recommendations. Libraries were validated for quality and size distribution using Qubit dsDNA HS Assay Kit (Invitrogen, Q32851) and TapeStation 2200 (Agilent). Libraries were paired-end sequenced (43bp+43bp) on NextSeq 550 (Illumina). The common lymphoid progenitor Hi-C dataset comes from previously published study under the GEO accession number GSE7942212

RNA-sequencing : Libraries were paired-end sequenced (38bp+38bp) on a NextSeq 550 (Illumina).

ATAC-sequencing : Libraries were paired-end sequenced (38bp+38bp) on a NextSeq 550 (Illumina).

All samples were acquired and analyzed with the LSR II flow cytometer (Becton Dickinson) and analyzed using FlowJo software (TreeStar).

scRNA-sequencing : Libraries were constructed following the 10X Genomic manufacturer's protocol and sequenced with a targeted sequencing depth of 20,000 read pairs per cell on pair-end mode on either a NextSeq 500/550 or a NovaSeq with recommended loading concentration (1.8pM and 300pM for NextSeq and NovaSeq respectively).

#### Data analysis

Hi-C analysis: All samples were aligned using the HiC-Pro. Next, Hi-C matrices were converted into the cooler format and normalized by matrix balancing with default parameters

Compartment analysis: Compartments were called using the cooltools command line interface (CLI) on each Hi-C dataset at 50kb resolution. We used the GC content file generated by cooltools as the reference track. A weighted compartment score was assigned to genes that span multiple regions with different compartment scores (weighting factor = Number of bps overlapping each region).

TAD analysis: TADs were called using the HiCExplorer tools. The function hicFindTADs was applied to group 1, group 2 and group 3 ILCs Hi-C

datasets at 25kb resolution, with default parameters. Conserved boundaries were determined by pooling all the TAD boundaries initially reported by hicFindTADs and selecting the ones that appeared in all datasets of interest. Two boundaries were considered the same if they differed by no more than one bin (25kb). Elements within a TAD can interact with other elements that are also within the TAD (intra-TAD interaction) or outside the TAD (extra-TAD interaction). For each TAD, we divided the sum of all intra-TAD interactions by the sum of all extra-TAD interactions in each group ILC. This quantity is referred to as the cross-boundary ratio. Next, the cross-boundary ratios of a given TAD, were compared across cell types by computing corresponding pairwise log fold changes between ILCs. Fig. 1f was constructed by selecting, for each cell type, TADs with cross-boundary ratios whose log fold change was greater than 0.25 when compared to the remaining cell types (termed type 1, type 2, type 3 TADs correspondingly).

Next, we selected genes within these TADs, whose RPKM log fold change was greater than 0.25 with respect to the remaining cell types. Motif analysis of these regions was carried out by intersecting the coordinates of each TAD with the ATAC-seq peaks of the corresponding cell type, and then, calling the findMotifsGenome.pl function of the HOMER package with default parameters.

Loop analysis: We called loops at three different resolutions (5, 10, 25kb) using the Mustache software with parameters; -pt 0.1. Next, for each dataset, we pooled together loops called at different resolutions by examining their anchors (overlapping of both anchors defined equivalent loops) using the bedtools suite.

Stripe analysis: Stripes were detected using the software Stripenn (<https://github.com/ysora/stripenn>) with default parameters.

RNA Sequencing: Bulk RNA data was aligned to the mm10 reference genome using STAR\_2.7.8a with --outFilterMultimapNmax 1 --outFilterScoreMinOverLread 0 --outFilterMatchNminOverLread 0 --outSAMtype BAM SortedByCoordinate --alignEndsType Local --outReadsUnmapped Fastx. Additionally, we counted reads using HTseq-count with -f bam -r name -s no -t exon -i gene\_id -m intersection-nonempty parameters 59. Next, we transformed read counts into RPKM for subsequent analyses.

ATAC-seq fastq files were aligned to the mm10 reference genome using the mem function of the software bwa-0.7.17 with additional parameters; -M. ATAC-seq peaks were subsequently called using MACS2 software with additional parameters; -nomodel -f BAM -B --keep-dup all --broad --broad-cutoff 0.1 -q 0.1. ATAC-seq profiles between primary ILC3s and the MNK3 cell line were compared by first creating a reference peak catalog. The reference peak catalog consisted of the pool of peaks called in each cell type with the corresponding raw counts in each sample. Next, we performed differential analysis using DESeq2. Additionally, we extracted the normalized counts computed by DESeq2 to perform the Principal Component Analysis between different ATAC-seq profiles shown in Fig. 1a. Motif analysis over the Id2 region containing LCR1, LCR2 and Rroid, was performed using HOMER.

All samples were acquired and analyzed with the LSR II flow cytometer (Becton Dickinson) and analyzed using FlowJo software (TreeStar).

scRNA-sequencing normalization and clustering. First, we filtered out low quality or dead cells by excluding any cell that had less than 1100 UMI and more than 8% of mitochondrial genes, also, we excluded genes expressed in less than 10 cells. Next, we processed the data following the steps recommended in the SCTransform tutorial. Briefly, we computed the percentage of mitochondrial genes (percent.mt) in addition to the cell cycling score (S.score, g2m.score) of each cell, and then called the SCTransform function of the Seurat package. For the SCTransform function we selected the following parameters:

```
method='GlmGamPoi', var.to.regress=('percent.mt','S.score','g2m.score')
```

Subsequently, we applied dimensional reduction using the RunPCA, RunUMAP, RunTSNE functions and created a Shared Nearest Neighbors (SNN) graph using the FindNeighbors function with default parameters. Finally, we called the function FindClusters with parameters: resolution=0.8, algorithm=4, method='igraph', to obtain the clusters reported in this work.

Differential Expression Analysis: We created a catalog of differentially expressed genes for several comparisons between clusters and between conditions. To do so, we found marker genes of each cluster using the FindAllMarkers command in Seurat with parameters; logfc.threshold=log(2), only.pos=TRUE. For differentially expressed genes both within clusters (between treat and control condition) and between clusters, we used the FindMarkers function with parameters; logfc.threshold=log(2), only.pos=TRUE. Genes with adjusted p-values smaller than 10<sup>-5</sup> and average log<sub>2</sub> fold change less than 0.25 were considered as differentially expressed in subsequent analysis. Pseudotime trajectories were calculated using the R package Monocle3 with default parameters. We computed pseudotime trajectories independently for LCR1+/- and LCR1-/- scRNA-seq data. For this analysis, we selected a random cell in cluster 8 (a cluster enriched in markers of early progenitor cell populations) as the starting point of the trajectory.

Fig. 5b and Extended Data Fig. 6g we used the scipy.signals package. We divided the pseudotime axis in 30 bins, computed mean expression levels per bin for each gene, and applied a Savitzky-Golay smoothing over the resulting signal (window\_lenght=81, polyorder=3).

Gene Ontology. GO Enrichment Analysis was performed using the online tool PANTHER (<http://geneontology.org/>).

Prism 10 GraphPad Software (version 9.4.0 or earlier versions) was used to calculate p-values. The difference between two groups was calculated using two-tailed unpaired t-test (Mann and Whitney) Prism 10 GraphPad Software. ANOVA and Bonferroni test were used for multiple comparisons (ns = not significant, \* = p ≤ 0.05, \*\* = p ≤ 0.01, \*\*\* = p ≤ 0.0005, \*\*\*\* = p ≤ 0.0001). The Mann and Whitney unpaired t-test was used for comparisons of one group between two conditions. One-way ANOVA was used to compare one group between more than two conditions. The Two-way ANOVA was used to compare multiple groups with two conditions. All graphs show the mean and the standard error of the mean (SEM). One- and Two-way ANOVA were corrected for multiple comparison using Bonferroni correction.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The mouse reference genome mm10 has been used for RNA-seq, ATAC-seq, Hi-C and scRNA-seq  
 Data have been uploaded on the gee accession omnibus : <https://www.ncbi.nlm.nih.gov/geo/> with the following accession numbers :

GSE191308  
 GSE191309  
 GSE191310  
 GSE191311  
 GSE191312

## Field-specific reporting

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

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

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

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

Sample size	<p>No statistical methods were used to pre-determine sample sizes but our sample sizes are similar to those reported in our previous publications For all experiments, data distribution was assumed to be normal but this was not formally tested.</p> <p>In vitro experiments were carried out using at least 4 mice per group.</p> <p>In vivo experiments aimed to characterize the LCR1<sup>-/-</sup>, LCR2<sup>-/-</sup>, Gata3<sup>BS</sup><sup>-/-</sup>, Rora<sup>BS</sup><sup>-/-</sup> mice were carried out using at east 5 mice per group, with the exception of myeloid cells in the spleen that was carried with n=3 mice per group and the Rora<sup>BS</sup><sup>-/-</sup> Gata3dKO BM and lungs ILC2 screening that was performed with n= 2 Rora<sup>BS</sup><sup>-/-</sup> Gata3dKO<sup>-/-</sup> mice per experiment.</p> <p>In vivo papain and IL-33 acute challenges were carried out using at least 3 mice per group.</p> <p>In vivo HDM challenge experiments: Eight HDM challenge experiments containing a total of 58 male and 35 female HDM-challenged mice were performed as ILC2s are absent in both male and female. These eight experiments were used to measure the following parameters:</p> <ul style="list-style-type: none"> <li>- Three experiments were used to measure immune cell infiltration in the lung parenchyma</li> <li>- Three experiments were used to measure immune cell infiltration in the BALF and for histopathological analysis</li> <li>- Two experiments were used to measure IL-5 levels in serum.</li> </ul> <p>For the three experiments performed for measuring immune infiltration to the lung parenchyma, we challenged a total of 20 male and 16 female age and sex-matched mice (6.9-17.2-week-old), which were part of two independent cohorts of males and one independent cohort of females. Two representative experimental cohorts containing 10 male and 16 female mice (6.9-12.9-week-old) were pooled for statistical analysis and data representation.</p> <p>For the three experiments performed for measuring immune infiltration to the BALF, we challenged a total of 21 male and 15 female age-matched mice (6.9-19.5-week-old), which were part of three independent cohorts that contained both males and females. Two representative experimental cohorts containing 14 male and 15 female mice (6.9-8.0-week-old) were pooled for statistical analysis and data representation. Lung histology was performed on 4 animals.</p> <p>For the two experiments performed for measuring IL-5 levels in serum, we challenged a total of 13 male and 4 females age-matched mice (7.5-8.5-week-old mice), which were part of one independent cohort of males and one independent cohort that contained both males and females. These two experiments were pooled for statistical analysis and data representation.</p>
Data exclusions	<p>No data values were removed from in vitro experiments.</p> <p>No data values were removed from in vivo experiments aimed to characterize the LCR1<sup>-/-</sup>, LCR2<sup>-/-</sup>, Gata3<sup>BS</sup><sup>-/-</sup>, Rora<sup>BS</sup><sup>-/-</sup>, or Rora<sup>BS</sup><sup>-/-</sup> Gata3dKO mice.</p> <p>No data values were removed from the papain challenge experiment.</p> <p>No data values were removed from the rIL-33 experiments.</p> <p>No data values were removed from experiments in which immune cell infiltration to the BALF was measured during an HDM challenge.</p> <p>No data values were removed from experiments in which IL-5 levels in serum was measured during an HDM challenge.</p>

For the three experiments performed for measuring immune infiltration to the lung parenchyma during an HDM challenge, we excluded mice in which the HDM challenge likely failed due to technical reasons during our data analysis. For this purpose, we excluded mice that did not show an increase of CD45+ cells in the lung parenchyma of HDM-treated mice when compared to the PBS-treated animals. Using this criterion, we excluded 2 LCR1+/- mice from our analysis. We corroborated our exclusion criteria by performing an Iglewics and Hoagling's outlier test with modified z-scores, which confirmed that these two data points were outliers (modified z-score > 2,  $p < 0.05$ ).

## Replication

Experiments aimed at characterizing LCR1-/- mice were performed by WMK or MFM as follows:

- BM ILC2 characterization was performed ten times.
- Lung parenchyma ILC2 characterization was performed eight times.
- FAT ILC2 characterization was performed five times.
- siLP ILC2 characterization was performed four times.
- Skin ILC2 characterization was performed three times.
- Splenocyte characterization was performed nine times (out of the nine experiments : Tregs were analyzed three times, CD4 T cell activation analysis was performed six times, myeloid cell analysis was performed three times, NK cell maturation analysis was performed five times).
- BM ILC3p staining was performed one time.
- Eosinophil quantification in the blood, spleen and lung was performed one time with four mice only.
- Liver NK/ILC1 characterization was performed three times.
- Th2 differentiation was performed three times.

Experiments aimed at characterizing LCR2-/- mice were performed by WMK or MFM as follows:

- Lung ILC2 characterization was performed four times.
- siLP ILC2 characterization was performed three times.
- BM ID2 MFI was performed two times
- BM ILC2 characterization was performed two times.
- Liver ILC1 characterization was performed once.
- Splenocytes characterization was performed two times.
- Myeloid splenocyte characterization was performed once.

Experiments using the LCR1-/-; Arg1-YFP mice were repeated two times.

Experiments using the Gata3\_BS-/- mice were repeated four times.

Experiments using the Rora\_BS-/- mice were repeated two times.

Experiments using the Rora\_Gata3dKO mice were repeated two times with n=2 Rora\_Gata3dKO per experiment (total = 4).

E15.5 embryonic ILC2 experiment were performed two times.

HDM-induced AAI experiments:

Eight independent HDM experiments were performed:

- \*Three experiments were used for measuring immune cell infiltration in the lung parenchyma.
- \*Three experiments were used for measuring immune cell infiltration in the BALF.
- \*Two experiments were used for measuring IL-5 level in serum.

The Papain-induced AAI experiment was performed once (We only measured BM and Lung ILC2s and eosinophils).

The rIL-33-induced AAI experiments were performed two times (We only measured BM and Lung ILC2s and eosinophils).

Chimera experiments (CD45.2+ BM progenitors transduced with Id2\_RV or Empty RV transferred into CD45.1+ recipients) were performed 5 times. Among the five experiments lungs ILC2 rescue analysis was performed two times. All experiments listed were successful

## Randomization

All experiments were age-matched and performed over groups containing males and females as ILC2s are absent in both sexes in LCR1-deficient mice. For comparison of multiple genotypes, WT from each strain were used.

Data collection and analysis were not performed blind to the conditions of the experiments.

We color-coded males and females in the panels associated to in vivo HDM, papain, and IL-33-induced AAI

Eight HDM-induced AAI challenge experiments containing a total of 58 male and 35 female HDM-challenged mice were performed.

For the three experiments performed for measuring immune infiltration to the lung parenchyma, we challenged a total of 20 male and 16 female age and sex-matched mice (6.9-17.2-week-old), which were part of two independent cohorts of males and one independent cohort of females. Two representative experimental cohorts containing 10 male and 16 female mice (6.9-12.9-week-old) were pooled for statistical analysis and data representation.

For the three experiments performed for measuring immune infiltration to the BALF, we challenged a total of 21 male and 15 female age-matched mice (6.9-19.5-week-old), which were part of three independent cohorts that contained both males and females. Two representative experimental cohorts containing 14 male and 15 female mice (6.9-8.0-week-old) were pooled for statistical analysis and data representation. Lung histology was performed on 4 animals.

For the two experiments performed for measuring IL-5 levels in serum, we challenged a total of 13 male and 4 female age-matched mice (7.5-8.5-week-old mice), which were part of one independent cohort of males and one independent cohort that contained both males and females. These two experiments were pooled for statistical analysis and data representation.

## 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

- | n/a                                 | Involved in the study   |
|-------------------------------------|---|
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Antibodies                  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Eukaryotic cell lines       |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology          |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Animals and other organisms |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Human research participants            |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data                          |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern           |

### Methods

- | n/a                                 | Involved in the study                              |
|-------------------------------------|--|
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> ChIP-seq       |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Flow cytometry |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging    |

## Antibodies

### Antibodies used

Anti-mouse CD3 (clone: 145-2C11) Biolegend 100327 - Dilution 1/400  
 Anti-mouse CD3 molecular complex (17A2) BD 742175 - Dilution 1/400  
 Anti-mouse CD4 (RM4-5) BV711 Biolegend 100549 - Dilution 1/400  
 Anti-mouse CD4 (RM4-5) BV785 Biolegend 100552 - Dilution 1/400  
 Anti-mouse CD4 (GK1.5) FITC Biolegend 100405 - Dilution 1/400  
 Anti-mouse CD5 (57-7.3) FITC Biolegend 100606 - Dilution 1/400  
 Anti-mouse CD8 (53-6.7) FITC Biolegend 100706 - Dilution 1/400  
 Anti-mouse CD8 (53-6.7) PE-eFluor 610 eBioscience 61-0081-80 - Dilution 1/400  
 Anti-mouse CD8 (53-6.7) BV785 Biolegend 100749 - Dilution 1/400  
 Anti-mouse CD8 (53-6.7) APC-eFluor 780 eBioscience 47-0081-80 - Dilution 1/400  
 Anti-mouse CD8 (53-6.7) Pacific Blue Biolegend 100725 - Dilution 1/400  
 Anti-mouse/human CD11b (M1/70) APC Biolegend 101225 - Dilution 1/400  
 Anti-mouse CD11c (N418) APC Biolegend 117303 - Dilution 1/300  
 Anti-mouse CD11c (N418) FITC Biolegend 117306 - Dilution 1/300  
 Anti-mouse CD19 (6D5) BV785 Biolegend 115543 - Dilution 1/300  
 Anti-mouse KLRG1 (MAFA) - Biolegend 138425 1/400  
 Anti-mouse CD19 (1D3) APC-H7 BD 560245 - Dilution 1/300  
 Anti-mouse CD25 (PC61) PeCy7 Biolegend 102015 - Dilution 1/200  
 Anti-mouse CD25 (PC61) PerCP-Cy5.5 Biolegend 102030 - Dilution 1/200  
 Anti-mouse CD44 (IM7) BV785 Biolegend 103059 - Dilution 1/400  
 Anti-mouse CD44 (IM7) BB700 BD 566507 - Dilution 1/400  
 Anti-mouse CD44 (IM7) APC-eFluor 780 eBioscience 47-0441-80 - Dilution 1/400  
 Anti-mouse CD45 (30-F11) AF700 Biolegend 103128 - Dilution 1/200  
 Anti-mouse CD45 (30-F11) BV421 Biolegend 103133 - Dilution 1/200  
 Anti-mouse CD45R (RA3-6B2) FITC Biolegend 103205 - Dilution 1/200  
 Anti-mouse CD45.1 (A20) AF700 Biolegend 110724 - Dilution 1/200  
 Anti-mouse CD45.2 (104) APC Biolegend 109814 - Dilution 1/400  
 Anti-mouse CD49a (HM $\beta$ 1) PE Biolegend 142604 - Dilution 1/400  
 Anti-mouse CD49b (DX5) APC Biolegend 108910 - Dilution 1/400  
 Anti-mouse CD62L (MEL-14) Pacific Blue Biolegend 104435 - Dilution 1/400  
 Anti-mouse CD90.2 (30-H12) Pacific Blue ThermoFisher Scientific 140305 - Dilution 1/400  
 Anti-mouse CD90.2 (30-H12) BV421 eBioscience 105323 - Dilution 1/400  
 Anti-mouse CD90.2 (53-2.1) A700 eBioscience 140323 - Dilution 1/400  
 Anti-mouse/pig CD117 (2B8) ThermoFisher Scientific - Dilution 1/200  
 Anti-mouse CD103 (2E7) Pe Texas Red Biolegend 121429 - Dilution 1/400  
 Anti-mouse CD127 (A7R34) BV785 Biolegend 135037 - Dilution 1/300  
 Anti-mouse CD127 (A7R34) PE Biolegend 135009 - Dilution 1/300  
 Anti-mouse CD135 (A2F10) APC Biolegend 135309 - Dilution 1/400  
 Anti-mouse CCR3 (J0735) APC Biolegend 144511 - Dilution 1/300  
 Anti-mouse CD45R (RA3-6B2) PE Biolegend 103207 - Dilution 1/400  
 Anti-mouse CD45R (RA3-6B2) Alexa-fluor488 Biolegend 103228 - Dilution 1/400  
 Anti-mouse a4b7 Integrin (DATK32) PE Biolegend 120605 - Dilution 1/400  
 Anti-mouse Eomes (Dan11mag) APC ThermoFisher Scientific - Dilution 1/200  
 Anti-mouse F4/80 (BM8) PB Biolegend - Dilution 1/400  
 Anti-mouse GATA3 (L50-823) BV711 BD Biosciences 565449 - Dilution 1/300  
 Anti-mouse GATA3 (L50-823) PeCy7 BD Biosciences 560405 - Dilution 1/300  
 Anti-mouse Gr-1 FITC Biolegend 108405 - Dilution 1/400  
 Anti-mouse ID2 (ILCID2) APC eBioscience 17-9475-82 - Dilution 1/400

Anti-mouse IL-22 (Poly5164) APC Biolegend 516409 - Dilution 1/200  
 Anti-mouse GM-CSF (MP1-22E9) PE ThermoFischer 12-7331-82 - Dilution 1/200  
 Anti-mouse IL-4 (11B11) Biolegend 504103 - Dilution 1/200  
 Anti-mouse IL-5 (TRFK5) PE ThermoFisher Scientific 65-0866-14 - Dilution 1/200  
 Anti-mouse IL-5 (TRFK5) APC Biolegend 504305 - Dilution 1/200  
 Anti-mouse IL-13 (eBio13A) PE-eFluor 610 Invitrogen 61-7133-80 - Dilution 1/300  
 Anti-mouse IL-13 (eBio13A) Invitrogen efluor 450 48-7133-80 - Dilution 1/300  
 Anti-mouse IL-13 (eBio13A) Invitrogen Alexa488 (FITC) 53-7133-82 - Dilution 1/300  
 Anti-mouse IL-23R (12B2B64) BV421 Biolegend 150907 - Dilution 1/200  
 Anti-mouse NK1.1 (PK136) BV711 Biolegend 108745 - Dilution 1/400  
 Anti-mouse NK1.1 (PK136) FITC Biolegend 108705 - Dilution 1/400  
 Anti-mouse Nkp46 (29A1.4) PerCP-Cy5.5 Biolegend 137608 - Dilution 1/400  
 Anti-mouse Nkp46 (29A1.4) FITC Biolegend 137606 - Dilution 1/400  
 Anti-mouse RORgt (B2D) PE eBioscience 12-6981-82 - Dilution 1/300  
 Anti-mouse RORgt (B2D) PE-eFluor 610 eBioscience 61-6981-80 - Dilution 1/300  
 Anti-mouse Sca-1 (D7) PE Biolegend 108107 - Dilution 1/400  
 Anti-mouse Siglec-F (E50-2440) PE-CF594 BD Biosciences562757 - Dilution 1/400  
 Anti-mouse T1/ST2 (DIH9) PE Biolegend 146609 - Dilution 1/300  
 Anti-mouse T1/ST2 (U29-93) PE BD 566312 - Dilution 1/300  
 Anti-mouse T1/ST2 (DIH9) PeCy7 Biolegend 145304 - Dilution 1/300  
 Anti-mouse/Human T-bet (4B10) PE eBioscience 12-5825-80 - Dilution 1/200  
 Anti-mouse/Human T-bet (4B10) BV711 Biolegend 6448109 - Dilution 1/200  
 Anti-mouse/Human T-bet (4B10) BV421 Biolegend 644815 - Dilution 1/200  
 Anti-mouse TCR $\alpha$  (H57-597) PE Biolegend 109208 - Dilution 1/400  
 Anti-mouse TCR $\alpha$  (H57-597) Biolegend 109211 - Dilution 1/400  
 Anti-mouse TCR $\alpha$  (H57-597) APC Biolegend 109212 - Dilution 1/400  
 Anti-mouse TCR $\alpha$  (H57-597) PE Biolegend 109207 - Dilution 1/400  
 Anti-mouse TCR $\alpha$  (H57-597) BB700 BD 745846 - Dilution 1/400  
 Anti-mouse TCRgd (UC7-13D5) FITC Biolegend 107503 - Dilution 1/400  
 Anti-mouse Ter119 (TER-119) FITC Biolegend 116206 - Dilution 1/400  
 Viability Dye eFluor 506 ThermoFisher Scientific 65-0866-14 - Dilution 1/2000  
 Viability Dye eFluor 780 ThermoFisher Scientific 50-112-9035 - Dilution 1/2000  
 Purified NA/LE Hamster Anti-Mouse CD28 Biolegend (clone Clone 37.51) 102116 - 1ug/ml  
 Anti-CD3 $\epsilon$  Armenian Hamster Monoclonal Antibody [clone: 145-2C11], Size=1 mg 100340 - Dilution 10ug/ml

## Validation

All antibodies were tested before use either during this work or known to be working based on our previous works or based on multiple other studies in the field and following the manufacturer's recommendation, paper dilutions. Staining procedures were performed following the manufacturer's recommendation. Antibodies were validated based on combinations of known cell markers and using positive and negative controls

## Eukaryotic cell lines

### Policy information about cell lines

## Cell line source(s)

The ILC3-like MNK3 cell line was previously published by and characterized by (Allan et al., Mucosal Immunol 2015) and was obtained via Maria Ciofani.  
All other cell lines used for virus production and titration were obtained from ATCC (HEK293T / NIH-3T3 cells)

## Authentication

Cell lines were not authenticated

## Mycoplasma contamination

Cell lines were not tested for Mycoplasma contamination.

Commonly misidentified lines  
(See [ICLAC](#) register)

No commonly misidentified lines

## Animals and other organisms

### Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

## Laboratory animals

All mice used were C57BL/6J background and bred at the University of Pennsylvania. Mice. The LCR1 $^{-/-}$ , LCR2 $^{-/-}$ , Rora $_{BS}^{-/-}$  and Gata3 $_{BS}^{-/-}$ , Rora\_Gata3dKO mice were generated using the CRISPR/Cas9 system. The sgRNA sequences used for the generation of knockout mice are listed in the Supplementary Table 13. Genomic deletion of each locus was confirmed by PCR and Sanger sequencing. All strains were backcrossed onto the C57BL/6 background for at least 3 generations to control potential off-target effects. B6.SJL-Ptprca Pepcb/Boy (CD45.1+ - Strain #:002014), and B6.129S4-Arg1tm1.1Lky/J (Arg1-YFP - Strain #:015857) mice were purchased from The Jackson Laboratory.  
All experiments were performed using 6- to 19-week-old mice that were sex and age matched mice, and all experiments were performed in both males and females for this study. Rora\_Gata3dKO were backcrossed for 2 generations to control potential off-target effects and were used at an age of 4- to 5-week-old. Mice were housed at 21°C +/- 2°C, 55% humidity (+/- 10%) with 12h light dark/ cycle in 7-7 IVC caging with environmental enrichment of plastic houses plus paper bedding.

## Wild animals

This study does not use wild animals

## Field-collected samples

No field-collected samples were used for this study

## Ethics oversight

All mice were bred and maintained under pathogen-free conditions at an American Association for the Accreditation of Laboratory Animal Care accredited animal facility at the University of Pennsylvania. Mice were housed in accordance with the procedures outlined in the Guide for the Care and Use of Laboratory Animals under the animal study protocol 805188 approved by the institutional Animal Care and Use Committee at the University of Pennsylvania.

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

*May remain private before publication.*

All ChIP-Seq data comes from previously available datasets and can be found here :

ChIP-seq H3K27ac ILC2 (Takashi et al., 2019) GEO: GSE111871

ChIP-seq H3K27ac NK cell (Shih et al., 2016) GEO: GSE77695

Cut & Run MNK3 (Cella et al., 2019) GEO: GSE130775

ChIP-seq Gata3 (Takashi et al., 2019) GEO: GSE111871

ChIP-seq RORa\_T2A (Ferreira et al., 2021) GEO: GSE146745

## Files in database submission

N/A

## Genome browser session

(e.g. [UCSC](#))

N/A

## Methodology

Replicates N/A

Sequencing depth N/A

Antibodies N/A

Peak calling parameters N/A

Data quality N/A

Software N/A

## 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

Cells were plated on a 96-well U bottom plate and cell suspension was first resuspended in PBS and stained with Viability Dye for 10 mins on ice. After incubation, cells were washed with PBS and resuspended in FACS buffer (PBS + 2% FBS + 2mM EDTA) with surface antibodies for 30 minutes at 4°C. Cells were washed with FACS buffer and fixed either using the Fopx3 buffer (eBioscience, 30 mins) or with 2% PFA then subjected to flow cytometry. For intracellular staining cells were incubated in perm wash buffer for at least 30 minutes at 4°C, then washed 1 time with perm wash buffer and resuspended in FACS buffer then subjected to flow cytometry.

## Instrument

All samples were acquired on a LSR-II (BD), LSR-Fortessa (BD), FACS-Aria II (BD)

## Software

Flowjo, DIVA

## Cell population abundance

Sorted cells were checked for purity and was over 90% (BM ILC2 -- cells were enriched first using antibodies to deplete Lineage positive cells from the BM).

## Gating strategy

Cells were gated on SSC-A/FSC-A, FSC-A/FSC-H (Singlets) and viability negative.

## Gating strategy

Gatinn included as well listed marker for each populations :

Lungs ILC2 (CD45+, Lineage-, CD127+, CD90.2+, ST2+), VAT (CD45+, Lineage-, CD127+, CD90.2+, ST2+), skin (CD45+, CD90.2+, Lineage-, CD127+, GATA-3+, KLRG1+), siLP (Live, CD45+, Lineage-, CD127+, CD90.2+, ST2+), BM (CD45+, Lineage-, CD127+, CD90.2+, ST2+, GATA-3+, ICOS+), ILC1 in liver (CD45+, TCRb-, B220, NK1.1+, CD49a+, CD49b-), NK cells (CD45+, TCRb-, B220+, NK1.1+, CD49a-, CD49b+), siLP ILC3 numbers (CD45+, CD90.2+, Lineage-, CD127+, ST2-, RORyt+), ILC development (see Extended data figure 6c and e and legends), total T cell (TCRb+), CD4+ T cell (TCRb+,CD4+), activated CD4+ T cells (TCRb+,CD4+,CD44+), CD4+ Th2 cells (TCRb+, CD4+,GATA3+), CD8+ T cell (TCRb+,CD8+), activated CD8+ T cells (TCRb+,CD8+,CD44+), Tregs (TCRb+, CD4+,FoxP3+), NKT cells (TCRb+, CD4+,CD1d+), eosinophils (SSChigh,SiglecF+)

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