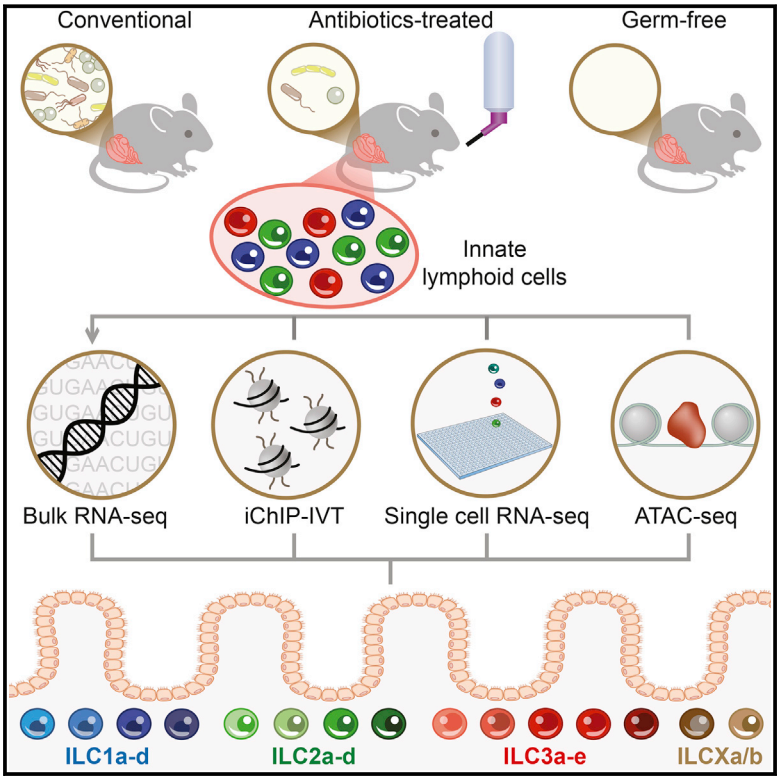
Article

Cell

The Spectrum and Regulatory Landscape of Intestinal Innate Lymphoid Cells Are Shaped by the Microbiome

Graphical Abstract



Authors

Meital Gury-BenAri, Christoph A. Thaiss, Nicolas Serafini, ..., James P. Di Santo, Eran Elinav, Ido Amit

Correspondence

eran.elinav@weizmann.ac.il (E.E.),
ido.amit@weizmann.ac.il (I.A.)

In Brief

The gut ILC population is a lot more diverse than we thought, and this diversity requires constant signaling from the local microbiota.

Highlights

- Transcriptional and chromatin landscapes of innate lymphoid cells
- Functional compartmentalization within ILC subsets revealed by single-cell RNA-seq
- Impact of commensal microbiota on epigenetic regulation and gene expression of ILCs
- Single-cell analysis of whole-transcriptome responses to microbial colonization

Data Resources

GSE85157
GSE85156
GSE85154
GSE85153
GSE85152

Gury-BenAri et al., 2016, Cell *166*, 1–16
August 25, 2016
http://dx.doi.org/10.1016/j.cell.2016.07.043

CellPress

Please cite this article in press as: Gury-BenAri et al., The Spectrum and Regulatory Landscape of Intestinal Innate Lymphoid Cells Are Shaped by the Microbiome, Cell (2016), http://dx.doi.org/10.1016/j.cell.2016.07.043

Article

Cell

The Spectrum and Regulatory Landscape of Intestinal Innate Lymphoid Cells Are Shaped by the Microbiome

Meital Gury-BenAri,[1,10] Christoph A. Thaiss,[1,10] Nicolas Serafini,[2,3] Deborah R. Winter,[1] Amir Giladi,[1] David Lara-Astiaso,[1] Maayan Levy,[1] Tomer Meir Salame,[4] Assaf Weiner,[1] Eyal David,[1] Hagit Shapiro,[1] Mally Dori-Bachash,[1] Meirav Pevsner-Fischer,[1] Erika Lorenzo-Vivas,[1] Hadas Keren-Shaul,[1] Franziska Paul,[1] Alon Harmelin,[5] Gérard Eberl,[6,7] Shalev Itzkovitz,[8] Amos Tanay,[9] James P. Di Santo,[2,3] Eran Elinav,[1,*] and Ido Amit[1,11,*]

[1]Department of Immunology, Weizmann Institute of Science, Rehovot 76100, Israel
[2]Innate Immunity Unit, Institut Pasteur, 75015 Paris, France
[3]Institut National de la Santé et de la Recherche Médicale (INSERM), U1223 Paris, France
[4]Biological Services Unit, Weizmann Institute of Science, Rehovot 76100, Israel
[5]Department of Veterinary Resources, Weizmann Institute of Science, Rehovot 76100, Israel
[6]Department of Immunology, Institut Pasteur, 75015 Paris, France
[7]CNRS, URA 1961, 75015 Paris, France
[8]Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot 76100, Israel
[9]Department of Computer Science and Applied Mathematics and Department of Biological Regulation, Weizmann Institute of Science, Rehovot 76100, Israel
[10]Co-first author
[11]Lead Contact
*Correspondence: eran.elinav@weizmann.ac.il (E.E.), ido.amit@weizmann.ac.il (I.A.)
http://dx.doi.org/10.1016/j.cell.2016.07.043

SUMMARY

Innate lymphoid cells (ILCs) are critical modulators of mucosal immunity, inflammation, and tissue homeostasis, but their full spectrum of cellular states and regulatory landscapes remains elusive. Here, we combine genome-wide RNA-seq, ChIP-seq, and ATAC-seq to compare the transcriptional and epigenetic identity of small intestinal ILCs, identifying thousands of distinct gene profiles and regulatory elements. Single-cell RNA-seq and flow and mass cytometry analyses reveal compartmentalization of cytokine expression and metabolic activity within the three classical ILC subtypes and highlight transcriptional states beyond the current canonical classification. In addition, using antibiotic intervention and germ-free mice, we characterize the effect of the microbiome on the ILC regulatory landscape and determine the response of ILCs to microbial colonization at the single-cell level. Together, our work characterizes the spectrum of transcriptional identities of small intestinal ILCs and describes how ILCs differentially integrate signals from the microbial microenvironment to generate phenotypic and functional plasticity.

INTRODUCTION

Innate lymphoid cells (ILCs) are the most recently discovered arm of the innate immune system, consisting of cytotoxic cells (NK cells) and "helper-like" ILCs (Eberl et al., 2015a; Serafini et al., 2015). The latter are characterized by expression of the interleukin-7 receptor (IL-7Ra/CD127) and are enriched at mucosal sites, where they contribute to the maintenance of tissue homeostasis and host defense against infection (Artis and Spits, 2015; Eberl et al., 2015b; Sonnenberg and Artis, 2015; Xu and Di Santo, 2013). Helper-like ILCs have been categorized into three distinct subsets based on their functional similarity to T helper cell subsets, with which they share transcription factors as well as cytokine signatures (Diefenbach et al., 2014; Serafini et al., 2015; Walker et al., 2013). Type I ILCs (ILC1s) express the transcription factor T-bet and produce interferon-γ (IFN-γ) (Bernink et al., 2013; Fuchs et al., 2013; Klose et al., 2014); type II ILCs (ILC2s) are regulated by Gata3 and secrete IL-5 and IL-13 (Hoyler et al., 2012; Mjösberg et al., 2012; Roediger et al., 2013); and type III ILCs (ILC3s) are characterized by RORγt and IL-22 expression (Cording et al., 2014). Additionally, recent evidence suggests the existence of an intermediate ILC subgroup between type I and type III ILCs (Bernink et al., 2015; Klose et al., 2013; Verrier et al., 2016; Vonarbourg et al., 2010), and between type I and type II ILCs (Bal et al., 2016; Lim et al., 2016; Ohne et al., 2016; Silver et al., 2016), but whether these findings represent general plasticity between ILC subsets and whether additional subsets exist that are lacking equivalent T cell counterparts remains unknown.

Single-cell genomics enables unbiased characterization of cell types (Grün et al., 2015; Jaitin et al., 2014; Paul et al., 2015; Treutlein et al., 2014), implicating comprehensive genome-wide sampling by single-cell RNA sequencing (RNA-seq) as an effective tool to systematically resolve ILCs heterogeneity. Transcriptional profiling studies have assigned gene expression signatures to each of the known ILC subsets and suggested key molecules involved in the control of ILC activity (Robinette

et al., 2015; Serafini et al., 2015). At the same time, heterogeneous transcript abundances within each subset point toward a higher complexity and plasticity that is not captured by the current classification scheme (Bjorklund et al., 2016).

The activity of ILCs is shaped by multiple tissue-specific signals, including nutrients and environmental xenobiotics (Kiss et al., 2011; Spencer et al., 2014; van de Pavert et al., 2014). It has also been reported that the activity of type III ILCs is influenced by the intestinal microbiota (Sanos et al., 2009; Satoh-Takayama et al., 2008; Sawa et al., 2011). However, the full scope by which the microbiota impacts the function of all three ILC subsets remains elusive. Similarly, it remains to be investigated how ILCs integrate changes in the local tissue environment, including alterations in microbial colonization, into the gene regulatory landscapes. Recent investigations into the chromatin landscape of human and mouse ILCs suggested a potential role for subset-specific regulomes in this process (Koues et al., 2016; Shih et al., 2016).

In this study, we used genome-wide chromatin and transcriptional profiling coupled to unbiased single-cell transcriptomic analysis to generate the first comprehensive map of the entirety of ILCs residing in the small intestinal lamina propria. We discover a strong degree of functional compartmentalization within all ILC subsets, manifesting as specialized expression of defined cytokines and metabolic pathways by distinct subpopulations. Beyond the known ILC subtypes, we also identify two additional transcriptional states that fail to fit within the current canonical classification. Using this high-resolution subset mapping, we profile the impact of commensal bacterial colonization on gene regulatory programs in ILCs on the single-cell level, identify epigenetic mechanisms of ILC specification, and determine influences of the microbiota on the chromatin landscapes of ILCs.

RESULTS

The Transcriptional Signature of Helper-like ILC Subsets

To study the dynamics of gene regulation in ILCs, we first profiled the transcriptomes of the three canonical subsets of helper-like ILCs (Figure 1A). To this end, we purified CD127$^+$ ILCs from the small intestine of RORγt-GFP reporter mice, and sorted RORγt$^-$NKp46$^+$ ILC1s, RORγt$^-$ KLRG-1$^+$ ILC2s, and RORγt$^+$ ILC3s for transcriptional profiling by RNA-seq (Figures 1A and S1A). We carefully ruled out contaminations from other cell types by comparing the obtained transcriptional profiles with those from purified intestinal myeloid populations (Figures S1B and S1C) (Lavin et al., 2014). We then focused on the set of 1,821 genes that were most highly expressed and differential across the three types (see STAR★Methods and Table S1). In order to identify global patterns of gene expression, we performed k-means clustering (k = 6) that divided the data into genes that were uniquely expressed by each of the types (Figure 1B, clusters I–III) and those that are shared between two types of ILCs (Figure 1B, clusters IV–VI). Successful purification of ILC subsets was confirmed by the specific expression pattern of the lineage transcription factors *Tbx21* encoding T-bet (ILC1), *Gata3* (ILC2), and *Rorc* encoding RORγ (ILC3), as well as the signature cytokines *Ifng* (ILC1), *Il4*, *Il5*, *Il13* (ILC2), and *Il22* (ILC3) (Figures 1B and 1C). Interestingly, KEGG analysis of the 1,161 subset-specific genes identified pathways involved in intracellular metabolism, such as mTOR and Notch signaling in ILC1s, sphingolipid and amino acid metabolism in ILC2s, and carbohydrate metabolism as well as glycolysis in ILC3s (Figure S1D), suggesting a differential metabolic profile in each type of ILC that might be coupled to their unique activities. In line with this finding, several additional transcription factors, soluble mediators, and metabolic enzymes featured subset-specific expression patterns, including the hexokinases *Hk2* and *Hk3* as well as enolase-1 and phosphofructokinase that likely represent active glycolysis in ILC3s (Figures S1E–S1G).

We found 660 genes shared between two groups of ILCs. Of note, ILC1s and ILC3s shared the largest group of common transcripts (306), including the transcription factor *Hopx* and the cytokine *Il16*, supporting the notion these subsets share a functional relationship and flexibility in classification (Bernink et al., 2015; Klose et al., 2013). Among the genes expressed by more than one subset were the urea cycle enzymes *Arg1* (ILC2/3) and *Odc1* (ILC1/3), the chemokines *Cxcl2* and *Cxcl3* (ILC2/3), as well as the growth factor *Vegfa* (ILC2/3) (Figures 1C and S1E–S1G). Together, these data define rich transcriptional, immunologic, and metabolic programs unique to each subset of ILCs but also identify elements spanning across the defined subsets (Figures S1D–S1G).

Transcriptional Regulation in ILC Subsets Is Linked to the Chromatin Landscape

We next sought to characterize the chromatin landscape underlying the observed transcriptional programs in the three ILC subsets. Because the small intestine of a mouse typically harbors less than 50,000 ILCs in each subset, obtaining accurate epigenetic profiling with conventional technology presents a major challenge. To enable high-sensitivity in vivo chromatin analysis from such a low number of cells obtained from individual mice, we developed a modification of our previously described indexing-first chromatin immunoprecipitation approach (iChIP) (Lara-Astiaso et al., 2014), which involves the integration of a linear amplification step using in vitro transcription (IVT) preceding DNA library preparation (Figures 2A and S2A). Since IVT amplification requires only a single ligation event to label and amplify the chromatin molecule, this iChIP-IVT protocol is 8-fold more sensitive and results in an increased ChIP library complexity from small cell numbers (Figure S2B). Indeed, highly reproducible H3K4me2 signals were achieved from as little as 100 cells (Figure S2C).

We then applied iChIP-IVT on the three subsets of small intestinal ILCs. We assayed several histone modifications including trimethylation of histone H3 at lysine 4 (H3K4me3), H3K4me2, and H3K27 acetylation (H3K27ac) in order to profile the global landscape for a variety of regulatory elements. The obtained chromatin signatures were free of contamination by other cell types, as evidenced by the absence of signal from myeloid or T lymphocyte-specific regulatory regions (Figures S3A and S3B) (Lara-Astiaso et al., 2014). Promoters, characterized by the enrichment of H3K4me3, tended to be shared among all three ILC subsets (6,206 out of 8,770 H3K4me3 regions)

Please cite this article in press as: Gury-BenAri et al., The Spectrum and Regulatory Landscape of Intestinal Innate Lymphoid Cells Are Shaped by the Microbiome, Cell (2016), http://dx.doi.org/10.1016/j.cell.2016.07.043

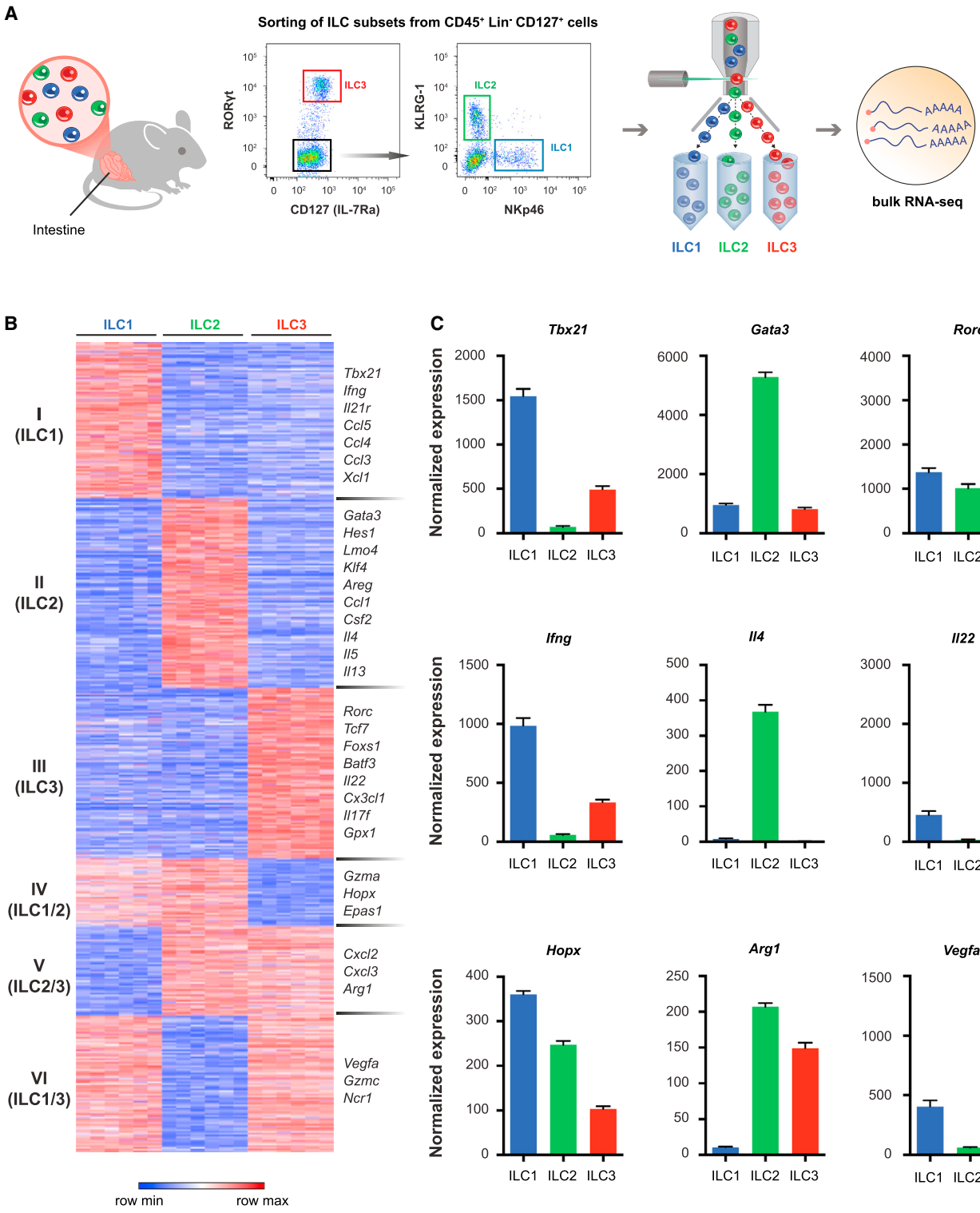


Figure 1. The Transcriptional Signature of ILC Subtypes

(A) Schematic and gating strategy for defining the transcriptomes of the three canonical ILC subtypes, isolated from the small intestinal lamina propria of RORc-GFP mice, and subjected to RNA-seq.

(B) Heatmap showing k-means clustering (k = 6) of 1,821 differentially expressed genes in the canonical ILC subsets. Clusters names and gene examples are given.

(C) Bar graphs of individual gene expression (a.u.). Error bars, SEM.

See also Figure S1 and Table S1.

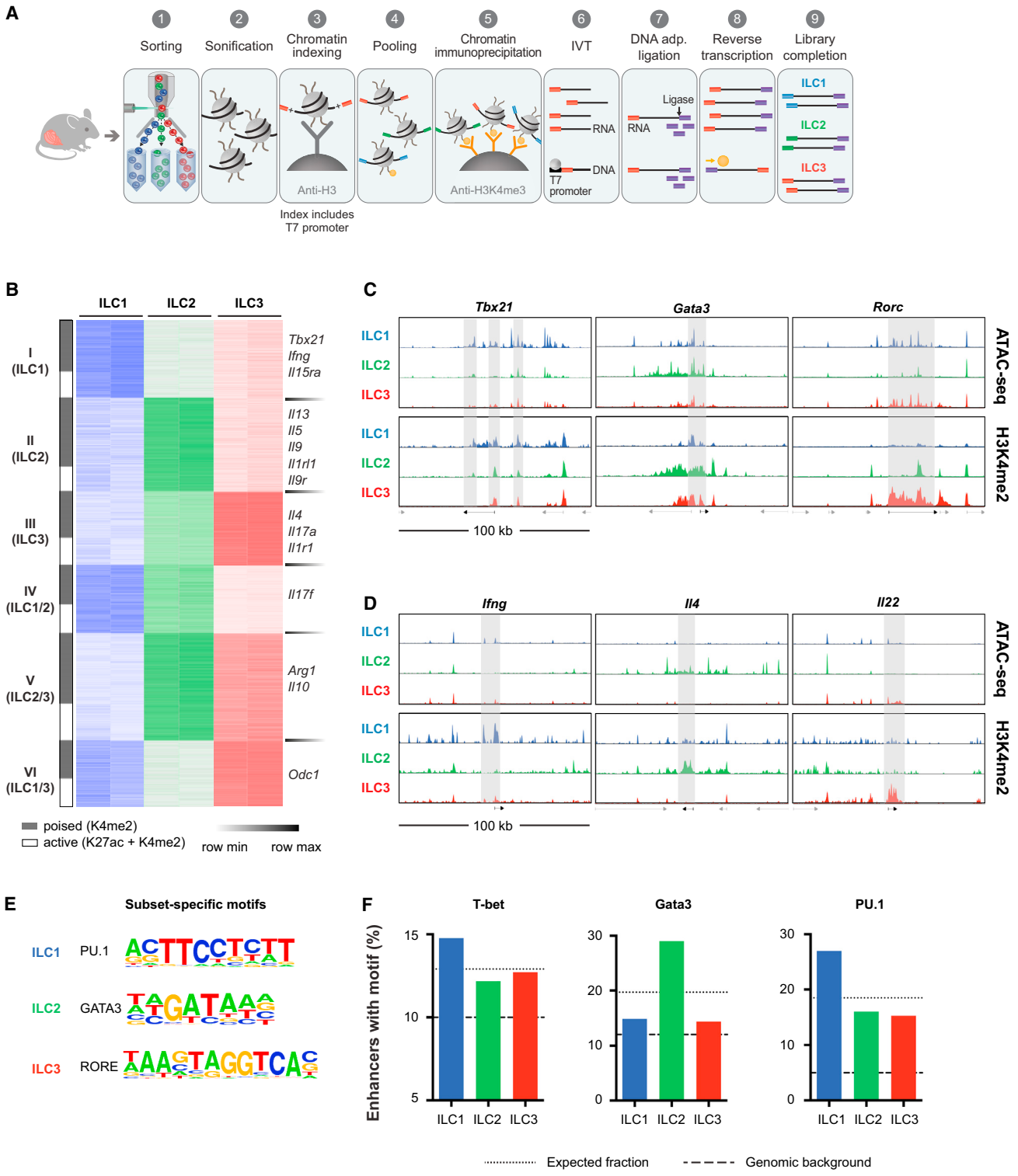


Figure 2. The Chromatin Landscape of ILCs Reveals Subtype-Specific Enhancers and Transcription Factor Binding

(A) Schematic diagram of the iChIP-IVT approach of indexing-first chromatin immunoprecipitation involving an initial T7 promoter-coupled chromatin barcoding step, ChIP with the desired antibody, linear amplification, and ssDNA adaptor ligation.

(B) K-means clustering (k = 6) of H3K4me2 intensity in 26,015 high-confidence enhancer regions. Examples of gene loci are provided. The proportion of enhancers that are active (H3K27ac high) in at least one sample in each cluster is indicated with a white bar on the right.

(legend continued on next page)

(Figure S3C; see STAR★Methods) consistent with the early developmental acquisition of chromatin organization in ILCs (Shih et al., 2016). In contrast, a large set of enhancers, distal regions marked by H3K4me2, were unique to each subset (1,268 in ILC1s, 5,934 in ILC2s, and 4,998 in ILC3s), implicating the enhancer landscape in the regulation of subset-specific gene expression (Figures 2B–2D and S3C; Table S2). Indeed, the six gene expression clusters were differentially linked to enhancers from the three ILC subsets, particularly when the enhancers are active (H3K27ac enrichment) (Figures 2B and S3D).

In order to more precisely analyze specific transcription factor binding sites, we performed an assay for transposase-accessible chromatin (ATAC-seq) (Buenrostro et al., 2013) on all ILC subsets and analyzed peaks of open chromatin within enhancer regions. At each of the loci encoding the three key transcription factors (*Tbx21, Gata3*, and *Rorc*), we observed differential chromatin landscape between the ILC subsets (Figure 2C). The same was true for the signature cytokines. For example, the locus of *Ifng* depicts at least one unique enhancer candidate upstream of the gene (Figure 2D), suggesting subset-specific transcription factor binding. In the case of *Arg1*, which is expressed in ILC2 and ILC3 (Figure 1C), both subsets appeared to be enriched for H3K4me2 at the transcription start site, while ILC2s featured an additional peak 25 kb upstream (Figure S3E), possibly associated with more complex regulation leading to higher expression in this subset (Figure 1C). In addition, we used a de novo motif-finding algorithm to search for enriched sequences in subset-specific regions compared to the full set of enhancers. We identified a unique motif signature for each ILC type, suggesting widespread binding of the relevant transcription factor: GATA3 in ILC2s ($p = 10^{-36}$) and ROR response element (RORE) in ILC3s ($p = 10^{-102}$) (Figure 2E; Table S3). Surprisingly, the top enriched motif signature for ILC1-specific enhancers was PU.1 ($p = 10^{-50}$), which is not known to play a role in ILC1 regulation. We also analyzed each enhancer cluster for enrichment of known motifs ($p < 10^{-5}$) against the genomic background (Figures 2F and S3F). This approach confirmed the presence of T-bet binding in ILC1-specific enhancers and revealed new candidate transcriptional regulators (Figures 2F and S3F). Collectively, these data describe the chromatin landscape and regulatory control regions of the three ILC subsets and highlight a network of transcription factors that uniquely control the transcriptional identity of each cell type.

Single-Cell RNA-Seq Identifies Transcriptional States of Small Intestinal ILCs

As the transcriptional and epigenetic profiles described above were obtained from populations sorted according to a small number of canonical markers of ILC subtypes, the observed gene expression signatures likely represent a composite picture of related and plastic subsets, such as NKp46$^+$ and NKp46$^-$ ILC3s (Bernink et al., 2015; Klose et al., 2013; Vonarbourg et al., 2010), and may be obscuring additional intermediate states of innate lymphocyte types. To obtain an unbiased description of all transcriptional states assumed by ILCs under homeostatic conditions, we performed massively parallel single-cell RNA-seq (MARS-seq) (Jaitin et al., 2014) of the entire CD127$^+$ cell population from the small intestinal lamina propria (Figure 3A), without enrichment for the known ILC subsets (Figure S4A). We used an index sorting strategy that allowed for retrospective analysis of surface marker combinations of each individual cell (Paul et al., 2015). We then clustered all cells based on 6,637 differential genes using an unbiased expectation-maximization algorithm previously described (Paul et al., 2015). This approach created a map of single cells that clustered into 23 transcriptionally homogeneous subgroups. Eight clusters containing small amounts (<2%) of cells were identified as non-ILCs based on marker gene expression and removed from further analyses (Figures S4B and S4C). T cell contamination was ruled out based on the absence of *Cd5* in the dataset (Figure S4D). The remaining 15 ILC clusters featured distinct transcriptional profiles, marker gene expression, and indexed protein markers (Figures 3B; see STAR★Methods). To classify the groups according to known ILC subsets, we generated gene modules characteristic for ILC1, ILC2, and ILC3 and evaluated each cluster based on similarity to these gene modules (Figure 3C). This approach identified four distinct transcriptional states within ILC1 (referred to as ILC1a-d), four clusters within ILC2 (ILC2a-d), and five states of ILC3 (ILC3a-e) (Figures 3B–3D; Tables S4 and S5). The observed within-subset heterogeneity was most pronounced within ILC1s and ILC3s (Figures 3B and S4E). Among ILC3s, cluster ILC3b showed strong similarity to transcriptional profiles reported for bulk-sorted NKp46$^+$ ILC3s (Robinette et al., 2015), while cluster ILC3e contained mostly cells with an lymphoid tissue inducer (LTi)-like signature (Figure S4F). We also found two transcriptional profiles (referred to as ILCXa and ILCXb) that did not feature a clear membership to one of the three canonical ILC groups (Figures 3B and 3C).

Functional Compartmentalization within Small Intestinal ILC Subsets

We next explored the characteristics of each individual transcriptional state identified by the transcriptome-based clustering approach. ILC1a-d featured a gradient of *Tbx21* expression, with the levels being highest in ILC1d, which also expressed the highest levels of *Ccl4*, *Itga1*, *Gzmb*, and *Gzmc* (Figure 4A). ILC1b and ILC1c expressed the transcription factor *Spfi1*, in line with our finding of PU.1-binding sites in ILC1-enriched regions (Figure 2E; Table S4). Other transcriptional states of ILC1s were characterized by expression of *Itga4* (ILC1b) and *Ccl5* (ILC1c). Indeed, flow cytometry analysis confirmed the presence of ILC1 subsets within the small intestinal RORγt$^-$ NKp46$^+$ population that can be

(C and D) Normalized profiles of H3K4me2 and ATAC-seq signal in 100-kb regions of the indicated loci of subset-specific genes encoding transcription factors (C) and cytokines (D).

(E) Top hits for motifs enriched in subset-specific enhancer regions. $p = 10^{-50}$ (PU.1); $p = 10^{-36}$ (GATA3); $p = 10^{-102}$ (RORE).

(F) Fraction of enhancers in subset-specific gene clusters containing the sequence motif for T-bet, Gata3, and PU.1. Dotted line indicates the expected distribution of enhancers across clusters. Dashed line indicates the genomic background level.

See also Figures S2 and S3 and Tables S2 and S3.

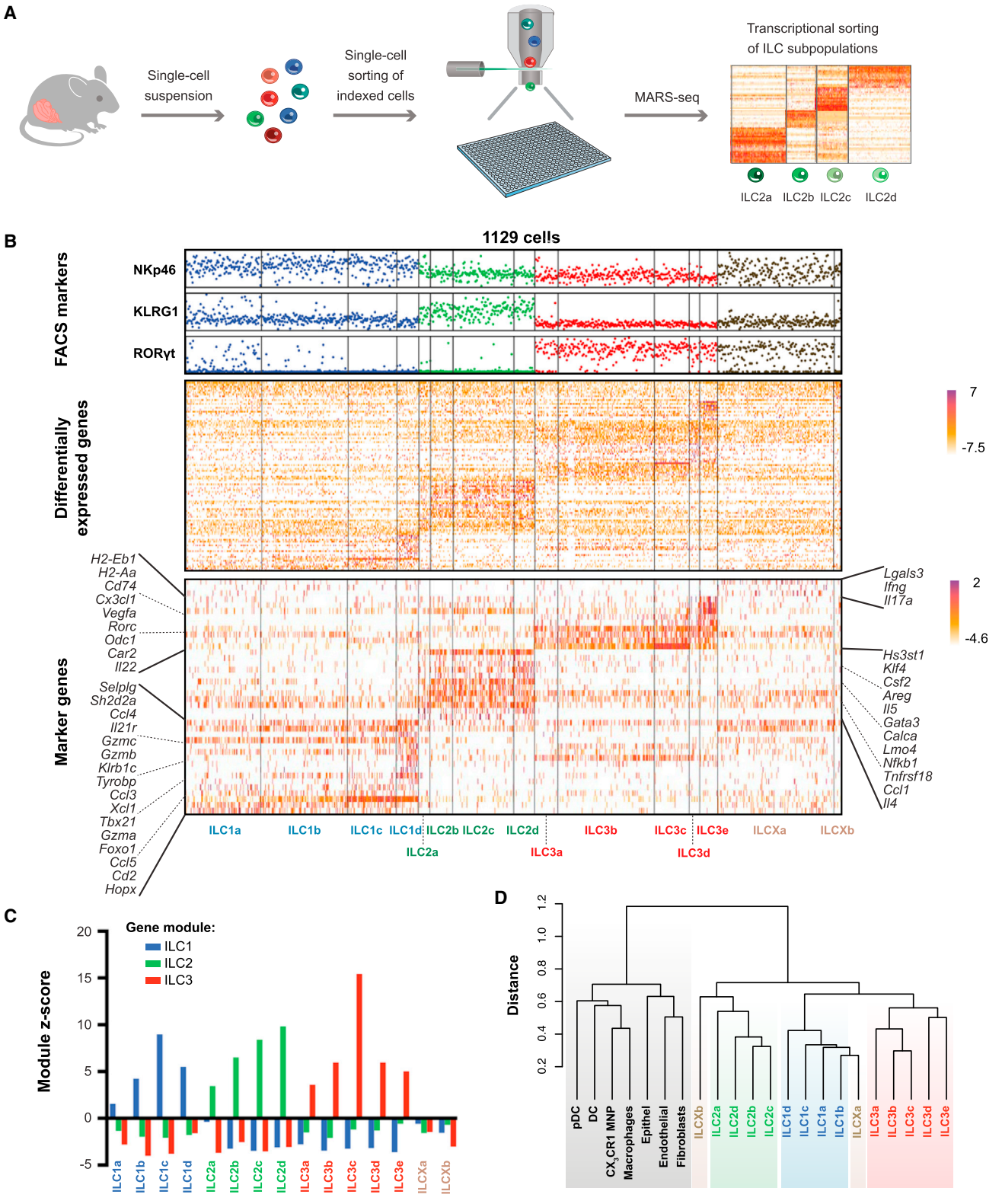


Figure 3. Single-Cell Transcriptome Analysis of Small Intestinal ILCs

(A) Schematic of the single cell assay. Single-cell suspension was prepared from mouse small intestinal lamina propria. Lin$^-$ CD127$^+$ single cells were then index sorted and processed for MARS-seq.

(legend continued on next page)

distinguished based on their expression of CD49a (encoded by *Itga1*) and CD49d (encoded by *Itga4*), as well as a subset of CCL5-producing cells (Figure 4B). ILC1a expressed the lowest levels of *Tbx21*, *Ifng*, and *Ccl5* (Figure 4A) among all ILC1s, but show detectable expression of *Gata3*, *Areg*, and *Klrg1* (Figure 4C), indicating that this subset might represent the recently described plasticity between ILC2 and ILC1 (Bal et al., 2016; Lim et al., 2016; Ohne et al., 2016; Silver et al., 2016). In silico fluorescence-activated cell sorting (FACS) analysis of each individual cell according to index sorting (Paul et al., 2015) confirmed protein expression of KLRG-1 in ILC1a (Figure S5A). The receptor for IL-21 (IL-21R) was highly expressed in all clusters of ILC1, confirming earlier reports of bulk ILC transcriptomes (Robinette et al., 2015). We further validated subset-specific IL-21R expression in ILC1s on the protein level by flow cytometry (Figure 4B).

Compartmentalization was also observed within ILC2s, albeit with less sharp boundaries between the different transcriptional states (Figures S4E and S5A). Similar to what we had observed for *Tbx21*, graded expression of *Gata3* characterized four different groups (ILC2a-d) that showed similarly graded levels of *Klf4*, *Klrg1*, *Ly6a*, and *Il2ra* (Figure 4C; Table S4). ILC2d, which was highest in the expression of all of these marker genes, uniquely expressed high levels of *Il5* and *Csf2* (Figure 4C). Similarly, the expression of amphiregulin (*Areg*), a factor critical for intestinal tissue protection (Monticelli et al., 2015), was confined to ILC2d (Figure 4C). In contrast, all ILC2s expressed high levels of *Arg1*, in line with the recent discovery of arginase-1 as an important metabolic mediator of ILC2 function (Monticelli et al., 2016). Interestingly, *Il4* expression was observed in an opposite gradient, being highest in ILC2a and lowest in ILC2d. Flow cytometry analysis confirmed the existence of subsets producing IL-4, which was distinct from the subset producing IL-5 (Figure 4D). IL-5 production correlated with high expression of KLRG-1 and Sca-1 (encoded by *Ly6a*), while IL-4 production was highest in KLRG-1$^-$ cells.

Within ILC3s, five diverse transcriptional states (ILC3a-e) were distinguished. One distinctive feature of these subsets was the abundance of the hallmark cytokine *Il22*, whose expression was almost exclusively restricted to group ILC3c, despite the equal distribution of the transcription factor *Rorc* across all subgroups (Figure 5A). Group ILC3a displayed high levels of *Ltb*, *Il17a*, and *Ccr6* (Figure 5A). Another subgroup, ILC3e, featured high levels of *Il17f*, *Ccr6*, and of genes involved in antigen presentation via major histocompatibility complex (MHC) class II (Figure 5A), an ILC3-associated function previously reported to aid in the deletion of microbiota-reactive T cells (Hepworth et al., 2013, 2015; von Burg et al., 2014). ILC3e thus appeared to correspond to the previously reported CCR6$^+$ NKp46$^-$ subset expressing MHC II molecules. Given the binary distinction between the IL-22-producing and MHC-II-expressing ILC3s, we further explored the characteristics of these two groups, using flow and mass cytometry. As predicted by the RNA-seq results, these groups could be distinguished by the surface expression of CCR6, CD49a, NKp46, and CD4 (Figure 5B). Expression profile analysis suggested the MHC molecules H2-Aa, H2-Ab1, and H2-Eb1, as well as the chaperone that regulates antigen presentation (CD74) as further distinctive markers for ILC3e, while ILC3c expressed *Gzmc*, *Cxcl10*, *Ramp3*, and *Lmo4*, apart from high levels of *Il22* (Figure 5C). We further validated these results using mass cytometry (Spitzer and Nolan, 2016), which allowed integrative analysis of more than 20 surface and intracellular markers (Figure S5B). Clusters ILC3c and ILC3e were identified by t-SNE clustering as distinct non-overlapping groups of cells with dichotomic expression of IL-22 and MHC-II (Figure S5B), further corroborating the results obtained from single-cell RNA-seq.

Thus, NKp46, CCR6, CD49a, and CD4 can be used to distinguish five clusters of ILC3s, only one of which highly expresses MHC-II (Figure 5D). As predicted by the RNA-seq data, CD49a$^+$ and CD49a$^-$ ILC3s also differed with respect to the production of IL-17A and IL-17F. While the majority of IL-17A$^+$ and IL-17F$^+$ double producers were found among the CD49a$^-$ group, distinct cells within the CD49a$^+$ group produced the IL-17A and IL-17F cytokines (Figure S5C).

Interestingly, the different clusters of ILC3 appeared to feature distinct metabolic functions, characterized by strongly enriched levels of the hexokinase *Hk2* in group ILC3b, a gene catalyzing the first step of glycolysis, and *Odc1* in group ILC3d, the rate-limiting enzyme in polyamine biosynthesis (Figure 3B; Table S4). This suggests that differential metabolic regulation might underlie the observed functional compartmentalization within ILC subsets.

Among the ILCX populations that did not cluster with any of the three canonical ILC subsets, subgroup ILCXa expressed high levels of the FACS markers NKp46 and RORγt as well as high levels of *Ifng* (Figures 3B, 3C, 4A, and S5C), reminiscent of earlier reports on "ex-ILC3" cells that acquire an ILC1-like expression profile (Bernink et al., 2015; Klose et al., 2013; Vonarbourg et al., 2010). Consistently, this cluster contained transcriptional features of both ILC1 and NKp46$^+$ ILC3 profiles (Figure S4E). The small cluster ILCXb most highly expressed *Il2* and *Ccl22* (Figure 3B) and did not show a strong similarity to any ILC gene module (Figure 3C), which might represent a transient or plastic ILC state.

Together, these results identify the heterogeneity of transcriptional states of small intestinal ILCs, revealing a high degree of compartmentalization within each of the subsets, manifested in clearly distinct expression patterns of transcription factors and functional mediators by different subgroups, which collectively explain the composite expression picture obtained from the analysis of bulk sorted ILC subsets.

(B) Heatmap of 1,129 single ILCs, sorted by clusters. Top: index sorting measurements of the three FACS markers used to discriminate between ILC subtypes. Middle panel: expression of the 100 most differential genes. Bottom: expression of marker genes.
(C) Average co-expression score of ILC gene modules for each identified transcription cluster. Gene modules were defined by all genes whose expression is 2-fold higher in cells of the respective subsets. Score was calculated as the deviation from the expected co-expression based on a shuffled expression matrix.
(D) Correlation-based hierarchical clustering analysis for the identified clusters of ILCs compared to other cell types.
See also Figures S4 and S5 and Tables S4 and S5.

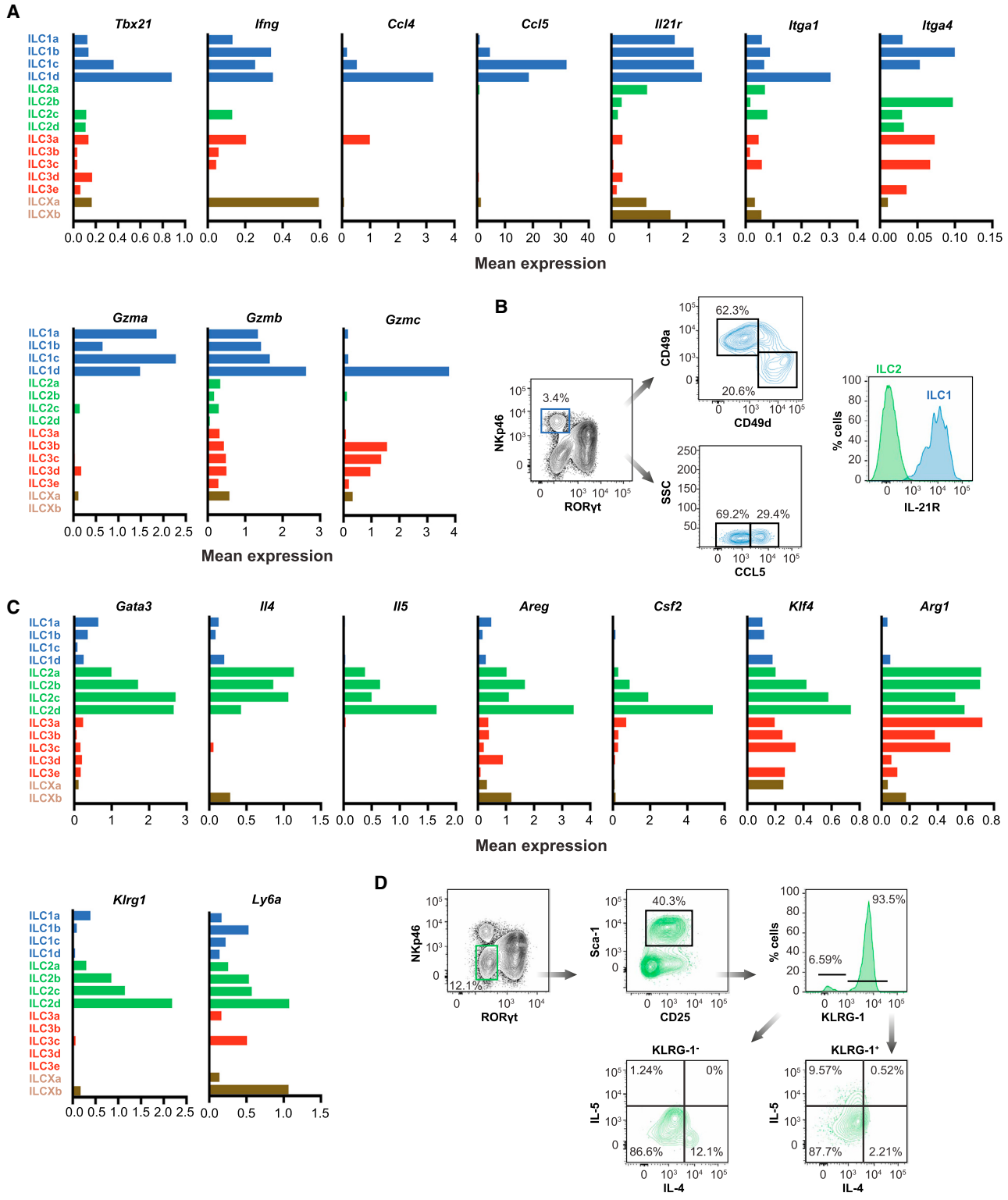


Figure 4. Functional Compartmentalization within Subsets ILC1 and ILC2

(A) Mean expression (per 1,000 molecules) of key marker genes across 15 transcriptional clusters of ILCs.

(B) Flow cytometry analysis of ILC1 subgroups expressing CD49a, CD49d, CCL5, and IL-21R.

(C) Mean expression (per 1,000 molecules) of key marker genes across 15 transcriptional clusters of ILCs.

(D) Flow cytometry analysis of ILC2 subgroups expressing IL-4 and IL-5.

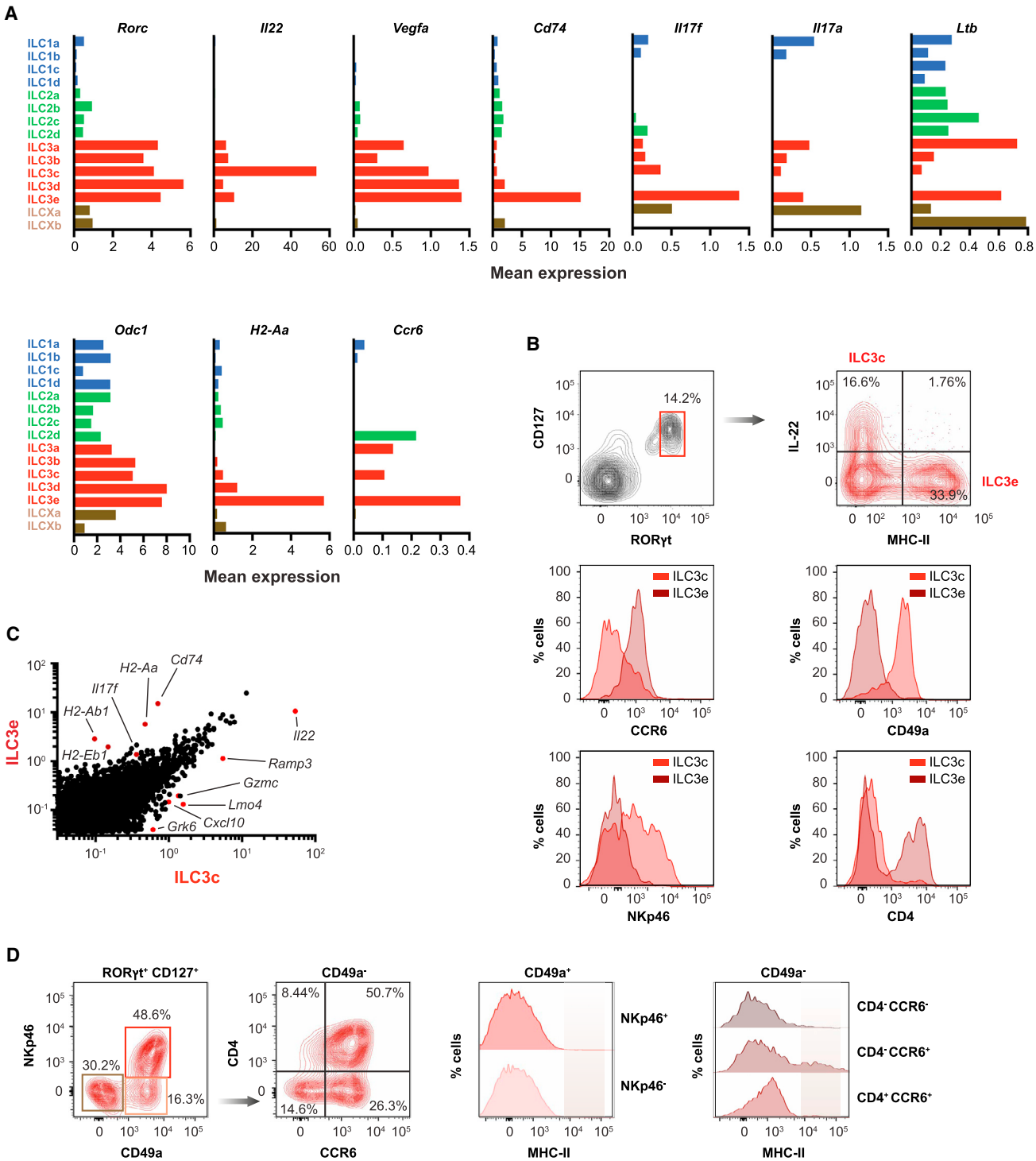


Figure 5. Functional Compartmentalization within Subset ILC3

(A) Mean expression (per 1,000 molecules) of key marker genes across 15 transcriptional clusters of ILCs.

(B) Flow cytometry analysis of the indicated surface markers on IL-22$^+$ and MHC-II$^+$ RORγt$^+$ ILCs.

(C) Comparison of gene expression between clusters ILC3c and ILC3e.

(D) Flow cytometry analysis of ILC3 subgroup-specific expression of MHC-II.

See also Figure S5.

The Transcriptional and Epigenetic Response of ILCs to Microbial Colonization

ILCs residing in the intestinal lamina propria are located in close proximity to a large community of microorganisms colonizing the lumen and mucosal surface of the gastrointestinal tract. To determine the impact of this microbial colonization on the homeostatic gene regulatory program of ILCs, we perturbed the community of mucosal-associated bacteria either by administration of broad-spectrum antibiotics (vancomycin, neomycin, ampicillin, and metronidazole) administered in the drinking water of adult RORγt-GFP reporter mice for 3 weeks, or by using adult germ-free (GF) mice born and raised in sterile isolators (Figure 6A). Expectedly, antibiotic treatment greatly reduced the total number of commensal bacteria colonizing the intestine, while germ-free mice featured undetectable levels of microorganisms, as confirmed by qPCR and electron microscopy (Figures 6B and S6A).

To determine the microbial effects on the ILC transcriptional landscape, we first determined the global effect of microbiota depletion by antibiotics (Abx) on the activity of ILCs within the three canonical subsets by RNA-seq and compared their gene expression profiles to those obtained from mice with normal (pathogen-free) microbial colonization (SPF). While the general transcriptional identity of all ILC subsets was preserved upon antibiotic treatment, the levels of several hundreds of transcripts were significantly altered (Figures S6B–S6D). Of note, the gene expression profiles of ILC1s and ILC2s were more strongly affected by Abx-induced microbiota depletion than ILC3s (Figures 6C and 6D). Remarkably, antibiotic treatment appeared to render the transcriptional profiles of ILC1s and ILC2s more similar to that of ILC3s (Figure 6C), and a significant overlap was noted between genes that were upregulated in either ILC1s or ILC2s and those that were classified as ILC3-specific in Figure 1B ($p < 10^{-10}$) (Figure 6D), including *Atf5* (Figure 6E), *Gpx1* (Figure S6E), and *Cxcl9* (Figure S6F). In addition, KEGG assignment of the genes most significantly changed upon antibiotic treatment revealed pathways associated with cellular adhesion and interaction with the extracellular matrix (ECM), chemokine signaling, as well as MAPK signaling (Figures S6B–S6D), all characteristic of the ILC3 transcriptome under homeostatic conditions (Figure S1D). These results further highlight the notion that ILCs acquire additional elements of the ILC3 program upon administration of antibiotics.

We next sought to characterize the epigenetic changes underlying these transcriptional alterations upon microbiota perturbation. To this end, we performed iChIP-IVT on ILCs from antibiotic-treated mice and compared the epigenetic profiles to those from SPF mice. Similar to the transcriptional landscape, the global chromatin identity of each subset was maintained upon microbiota depletion, but shifts were seen in several thousands of H3K4me2 regions—including the loss of subset specificity in previously clustered enhancers (Figures S6G and S6H), most prominently within ILC1s. For instance, in the case of *Tcf7*, *Cd93*, and *Gjb2*, administration of antibiotics diminished lineage specificity in H3K4me2 marks (Figure 6F). To link the transcriptional alterations associated with antibiotic treatment with changes in the chromatin landscape, we next focused on the set of differential genes most significantly changed by antibiotics and examined their associated chromatin status in normal conditions. We found that the loci of genes upregulated in ILC1s and ILC2s were enriched for enhancers that were unique to ILC3s in homeostasis (Figure 6G), further substantiating the notion that upon treatment with antibiotics, all subsets acquired transcriptional elements associated with ILC3s. Furthermore, motif search of transcription factors binding sites preferentially found in the gained regions compared with lost identified the ROR elements in both ILC1s and ILC2s (Figure 6G). Together, these results characterize the global changes in the ILC transcriptome and epigenome induced by microbiota depletion and suggest that all three subsets at least partially assume a RORγt-driven ILC3-like expression and chromatin profile upon loss of normal microbial colonization.

The Microbiome Differentially Influences Gene Regulation of ILC Subgroups

Given the transcriptional heterogeneity of intestinal ILCs on the single-cell level, we next sought to identify the particular subgroups of ILCs that account for these bulk transcriptional responses. We therefore performed MARS-seq on 706 and 944 cells obtained from the total population of $CD127^+$ helper-like ILCs isolated from the small intestine of antibiotic-treated and GF mice, respectively, and clustered them together with the single-cell transcriptomes obtained from naive mice. Interestingly, when comparing the relative abundances of each of the 15 ILC subgroups between all three conditions of microbial colonization, we noted an expansion in the proportion of cells categorized as ILC3s and to a lesser extent ILC2s (Figure 7A). Overall, the single-cell level ILC compositional profiles of Abx-treated and GF mice looked similar to each other, but distinct from SPF mice, indicating that short-term antibiotic-induced microbiota depletion phenocopies the effect of life-long absence of a microbiota on the population structure of ILCs.

To assess the degree of microbiota responsiveness of each individual cluster, we performed principal component analysis on the mean expression profiles of all 15 clusters and determined the relative distances between each cluster in the Abx and GF groups and their corresponding SPF counterparts (Figures S7A and S7B). Interestingly, the degree of alterations in gene expression was highly distinct between subgroups, but the relative changes in gene expression within each subgroup were similar between Abx and GF mice (Figures S7A and S7B). We identified seven "ILC states" (ILC1a, ILC1c, ILC1d, ILC2a, ILC2c, ILC2d, and ILC3a) to be most responsive to alterations in microbial colonization, further corroborating the high degree of microbiota responsiveness of subsets ILC1 and ILC2. Comparison of gene expression profiles among all 15 subgroups revealed several interesting features. In line with the observations from bulk population RNA-seq, we noted an upregulation of ILC3-associated genes in ILC1s and ILC2s in the scenario of microbiota depletion, as exemplified by the appearance of *Hk2* expressing ILC1 and ILC2 clusters (Figure 7B). Such loss of sharp expression differences between the subsets upon Abx treatment was also observed for ILC2-specific genes, as exemplified by *Klrg1* (Figure S7C). The expression of several genes was lost across all subgroups in both Abx and GF groups, including *Il17a* (Figure 7C), consistent with previous reports on the microbiota

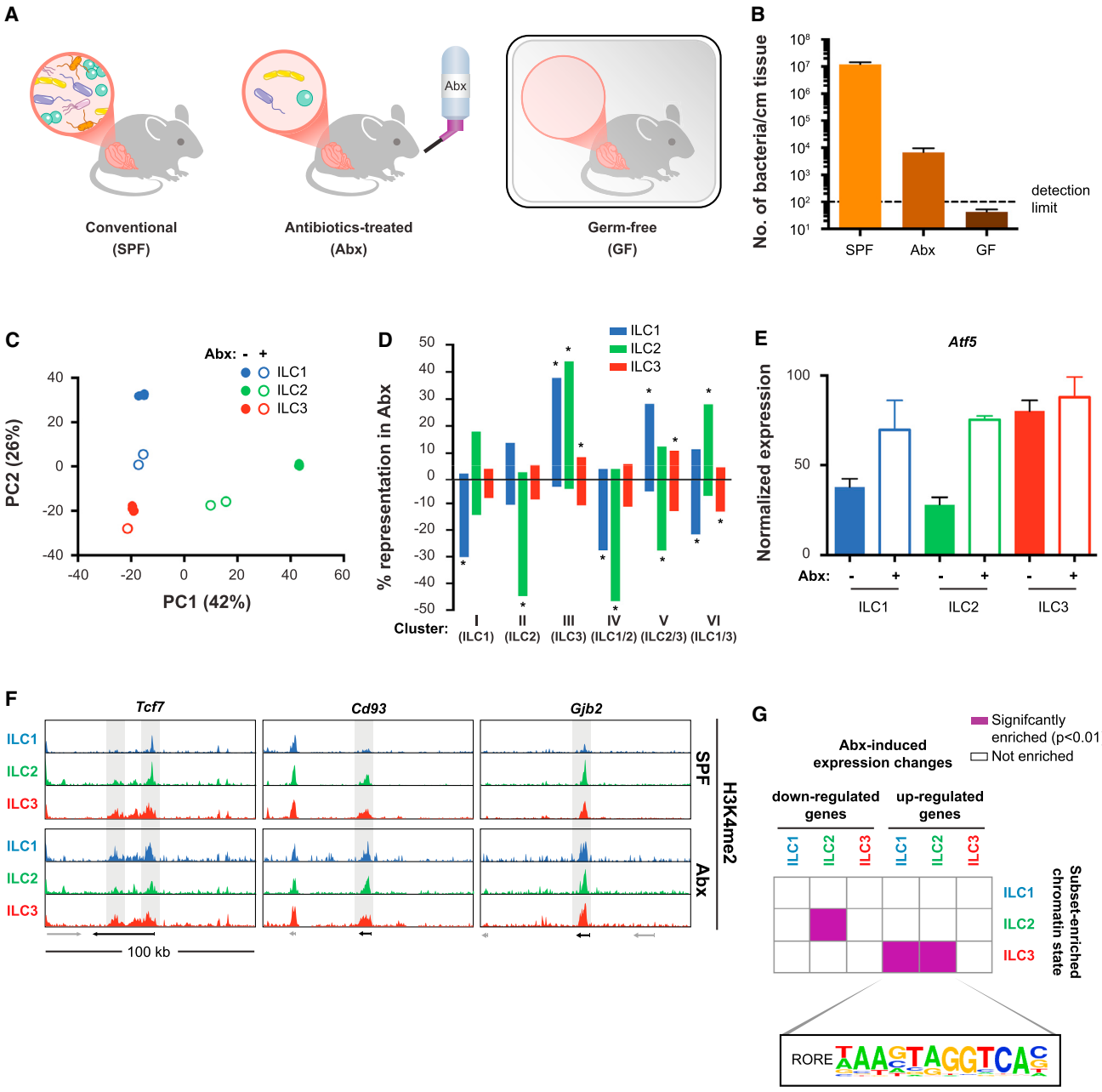


Figure 6. The Microbiota Impacts the Transcriptional and Epigenetic Landscape of ILCs

(A) Schematic of the three states of microbial colonization analyzed. Left, specific pathogen-free (SPF); middle, antibiotics treatment (3 weeks in drinking water); right, germ-free housing in sterile isolators.

(B) Quantification of mucosal bacterial colonization per cm of intestinal tissue.

(C) Principal components analysis of differentially expressed genes from Figure 1B across ILCs obtained from SPF mice and mice treated with Abx.

(D) Fraction of genes in each expression cluster from Figure 1B with differential expression (at least 2-fold change) upon antibiotic treatment. *Significance ($p < 0.05$; hypergeometric distribution)

(E) Example showing expression change in an ILC3-associated gene that is upregulated in ILC1 and ILC2 upon antibiotic treatment. Error bars, SEM.

(F) Examples showing H3K4me2 intensity in ILC3-associated loci modified in ILC1 upon antibiotic treatment. 100-kb regions are shown.

(G) Enrichment of candidate enhancers (distal H3K4me2 regions) from the ILC subset-specific clusters in SPF mice from Figure 2B. A hypergeometric distribution determined p values. Analysis of gained compared with lost enhancer regions in ILC1s and ILC2s identified RORE as the most significantly enriched motif. See also Figure S6.

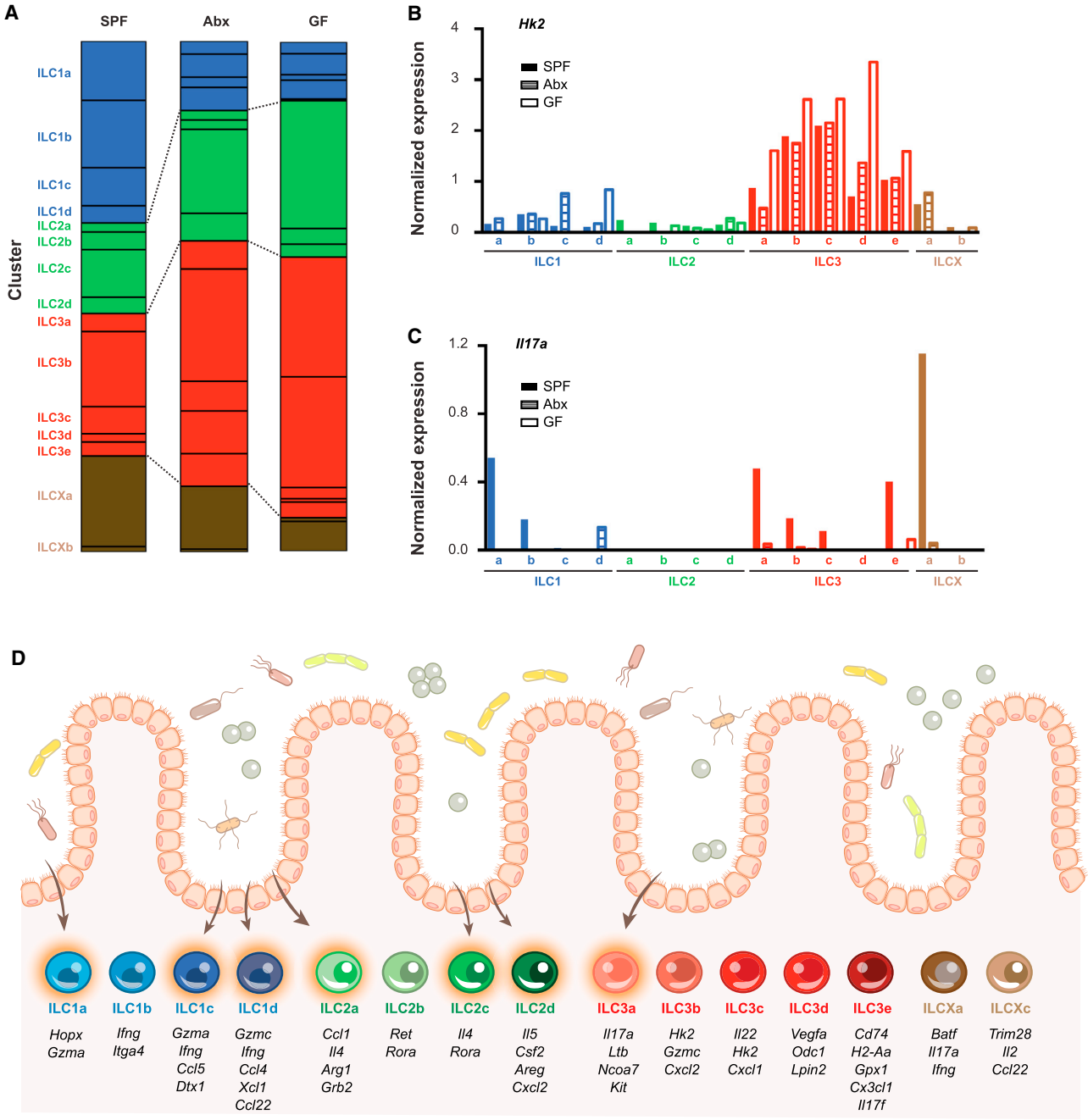


Figure 7. Single-Cell Analysis of ILC Transcriptional Responses to Microbial Colonization

(A) Changes in distribution of single cells mapping to clusters 1–15 in ILCs from microbiota-perturbed mice (Abx, GF) as compared to SPF mice.

(B and C) Bar graphs of gene expression across 15 clusters of ILC single-cell subpopulations in SPF, Abx, and GF mice.

(D) Schematic summarizing the heterogeneous spectrum of ILCs in the small intestinal lamina propria. Arrows to highlighted subsets indicate ILCs whose transcriptomes are most strongly affected by the microbiota.

See also Figure S7.

dependency of this cytokine (Ivanov et al., 2009). Nevertheless, the functional compartmentalization within ILC3s remained intact, as exemplified by *Cd74* expression in the MHC-II-expressing cluster (Figure S7D). In line with previous reports, we found a strong microbiome-dependency of clusters featuring plastic characteristics between ILC3 and ILC1 (Atarashi et al., 2015; Klose et al., 2013; Sanos et al., 2009; Satoh-Takayama et al., 2008; Sawa et al., 2011; Vonarbourg et al., 2010), manifesting in reduced numbers, altered transcriptomes, and reduced *Tbx21* and *Il22* expression of clusters ILC1a and ILCXa in the

absence of the microbiome (Figures 7A, S7B, and S7E). In contrast, *Il22* levels in *Tbx21*-negative ILC3s remained unaltered (Figure S7E). Together, our analysis identifies unique cluster-specific ILC responses to intestinal microbial colonization on the single-cell level (Figure 7D).

DISCUSSION

In this study, we systematically characterize the genomic landscape of intestinal helper-like ILCs by integrating genome-wide RNA-seq, ChIP-seq, ATAC-seq, and single-cell transcriptome analysis, to obtain the first comprehensive description of the transcriptional identity and regulatory landscape of the diverse ILC subsets present in the small intestine. We identify thousands of genomic control elements and transcription factor circuits that regulate gene expression in the three canonical ILC subsets. In addition, our study assesses the unclassified total population of helper-like ILCs in the small intestine by single-cell transcriptomics, revealing 15 transcriptional states assumed by ILCs under homeostatic conditions and a high degree of functional specialization within the known subsets. Alteration of commensal microbial colonization by antibiotic treatment or in germ-free mice results in subgroup-specific responses toward a global acquisition of genomic elements characteristic of the ILC3 profile.

The discovery of ILCs less than a decade ago has provided numerous critical insights into our understanding of the innate immune response involved in the protection of tissue homeostasis and the immune response to infection at mucosal surfaces. Nonetheless, several fundamental questions of ILC physiology have remained unsolved. First, ILCs have a unique role at the interface between the immune system of the host and the mucosal environment, including microorganisms and dietary components. The mechanisms by which ILCs perform this integration task are poorly understood. Our systematic charting of the in vivo chromatin landscape of ILCs in the presence or absence of commensal communities provides insight into the subset-specific regulation of gene expression and suggests regulatory mechanisms driving the versatile responses of ILCs to changes in the microbial environment. With the improved sensitivity of the iChIP-IVT process, we were able to assemble a comprehensive in vivo map of the ILC control elements, consisting of accessible chromatin, promoters, as well as active and poised enhancers. Using this methodology, we demonstrate that the chromatin landscape can adjust to changes in commensal microbial colonization, even at the timescale of weeks, while preserving cellular epigenetic identity. Our data therefore suggest that ILCs evaluate the state of microbial colonization by adjusting the enhancer landscape and the accessibility of transcription factor-binding sites within the chromatin architecture (Lavin et al., 2014).

Second, while previous studies assessing the transcriptional profiles of ILCs have focused on their activity in immune responses, our data highlight an additional functional specialization with respect to the intracellular metabolic pathways employed by the three subsets of ILCs. Genes expressed by ILC3s were highly enriched in elements of glycolysis, a metabolic phenotype typically associated with effector lymphocytes (Pearce and Pearce, 2013) and furthermore featured a signature of fructose, mannose, and galactose metabolism. In contrast, ILC2s exhibited a profile characterized by amino acid and polyamine metabolism, including genes involved in arginine, proline, alanine, and ornithine metabolism. This metabolic profile resembles the one found in M2-polarized and tumor-associated macrophages (Colegio et al., 2014). Given the close connection and mechanistic coupling of intracellular metabolism and immune cell function (Kelly and O'Neill, 2015), the elucidation of functional crosstalk between these pathways and ILC function will present an exciting area of future study.

Third, our single-cell functional map of CD127^{+} ILCs reassembles the landscape of small intestinal ILCs, confirms the existence and functional division of the three classical subsets from an unbiased perspective, identifies two transcriptional states that do not feature a high congruency with any subset-specific gene program, and demonstrates a high degree of intra-subset specialization. Thereby, albeit restricted to the limitations of current single-cell technology, our analysis highlights the existence of discrete functional compartments within the known ILC subsets, pointing towards transcriptional diversity within defined ILC populations, as exemplified by antigen-presenting ILC3s that inhibit microbiota-directed T cell responses (Hepworth et al., 2015) versus IL-22-producing ILC3s that promote microbiota-directed T cell responses (Sano et al., 2015). Elucidating the role of within-tissue distribution, cell-cell interactions, exposure to soluble mediators, and response to luminal metabolites on these transcriptional clusters awaits further study. Furthermore, temporal and spatial dynamics of marker gene expression not captured by our analysis might constitute another layer of regulation that can be resolved by longitudinal application of the technologies presented here (Verrier et al., 2016; Vonarbourg et al., 2010).

Finally, our study provides more insights into the impact of the microbiota on gene regulation in ILCs, by demonstrating that the responsiveness of ILCs to the microbiota is highly heterogeneous, even within the defined subsets. Integration of all three levels of genomic assessment—population transcriptomics, population epigenetics, and single-cell transcriptomics—point toward an unexpected phenomenon, namely, the acquisition of ILC3-like expression profiles across multiple subsets upon depletion of the microbiota. This finding suggests that homeostatic commensal colonization may suppress the regulatory elements involved in ILC3 fate determination and the execution of the associated transcriptional program (Sawa et al., 2011). Thus, in the case of antibiotic treatment or under germ-free conditions, the absence of signals from the microbiota may release the default ILC3-associated regulatory inhibition. At the same time, ILC3 plasticity toward the ILC1 fate is suppressed in the absence of the microbiota, in line with previous reports (Klose et al., 2013; Sanos et al., 2009; Satoh-Takayama et al., 2008; Vonarbourg et al., 2010). This finding provides an example of the global and unexpected effects potentially mediated by antibiotic use on homeostatic immune function. It further suggests that proper transcriptional regulation of the ILC-microbiota crosstalk may play critical roles in preservation of a healthy intestinal microenvironment, while preventing auto-inflammatory disorders, some of which involve aberrant ILC3 activity (Buonocore et al., 2010; Geremia et al., 2011).

Collectively, the insights gained in this study present a first step toward a more comprehensive understanding of ILC regulation and functional plasticity, as well as their roles in a multitude of microbiota-associated disorders. Similar studies of ILC subsets in other tissues from both mice and humans, and in response to various environmental perturbations, have the potential to define new markers, regulatory regions, targets, and pathways perturbed across a wide range of diseases, with prospective for therapeutic intervention.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Mice
- METHOD DETAILS
  - Isolation of ILCs from Small Intestinal Lamina Propria for Flow Cytometry
  - Mass Cytometry
  - Electron Microscopy
  - Bulk RNA Isolation, Library Construction, and Analysis
  - RNA-Seq Processing and Analysis
  - iChIP-IVT
  - Primers Used during iChIP-IVT Library Construction
  - ATAC-Seq
  - Sequencing
  - Processing of ChIP-Seq and ATAC-Seq
  - Chromatin and Motif Analysis
  - Gene Tracks and Normalization
  - Single-Cell Sorting
  - Index Sorting
  - MARS-Seq Library Preparation
  - Single-Cell Analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY
  - Data Resources

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and five tables and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2016.07.043.

AUTHOR CONTRIBUTIONS

M.G.-B. and C.A.T. conceived the project; designed, performed, and analyzed experiments; and wrote the manuscript. N.S., D.L.-A., M.L., T.M.S., H.S., M.D.-B., M.P.-F., E.L.-V., H.K.-S., and F.P. developed experimental protocols and performed experiments. D.R.W., A.G., A.W., and E.D. developed computational methods and performed bioinformatic analysis. A.L., G.E., S.I., A.T., and J.P.D. contributed essential tools. E.E. and I.A. conceived and directed the project, designed experiments, interpreted the results, and wrote the manuscript.

ACKNOWLEDGMENTS

We thank the members of I.A.'s and E.E.'s labs for fruitful discussions, Carmit Bar-Nathan for GF mouse care taking, Elena Kartvelishvily for help with electron microscopy, Melanie Flach for technical advice, and Maayan Wigelmann for art work. C.A.T. received a Boehringer Ingelheim Fonds PhD Fellowship. E.E. is supported by Y. and R. Ungar; the Abisch Frenkel Foundation for the Promotion of Life Sciences; the Gurwin Family Fund for Scientific Research; the Leona M. and Harry B. Helmsley Charitable Trust; the Crown Endowment Fund for Immunological Research; the estate of J. Gitlitz; the estate of L. Hershkovich; the Benoziyo Endowment Fund for the Advancement of Science; the Adelis Foundation; J.L. and V. Schwartz; A. and G. Markovitz; A. and C. Adelson; the French National Center for Scientific Research (CNRS); D.L. Schwarz; the V.R. Schwartz Research Fellow Chair; L. Steinberg; J.N. Halpern; A. Edelheit, and by grants funded by the European Research Council; a Marie Curie Integration grant; the German-Israeli Foundation for Scientific Research and Development; the Israel Science Foundation; the Minerva Foundation; the Rising Tide Foundation; the Helmholtz Foundation; and the European Foundation for the Study of Diabetes. E.E. is the incumbent of the Rina Gudinski Career Development Chair and a senior fellow of the Canadian Institute For Advanced Research (CIFAR). I.A. is supported by the European Research Council (309788), the I-CORE for chromatin and RNA regulation, and personal grants from the Israel Science foundation (782/11) and the BLUEPRINT FP7 consortium, the Ernest and Bonnie Beutler Research Program of Excellence in Genomic Medicine, a Minerva Stiftung research grant, the National Human Genome Research Institute Center for Excellence in Genome Science (1P50HG006193), the Israeli Ministry of Science, Technology and Space, the David and Fela Shapell Family Foundation, and the Abramson Family Center for Young Scientists. I.A. is the incumbent of the Alan and Laraine Fischer Career Development Chair.

Received: March 14, 2016
Revised: May 18, 2016
Accepted: July 27, 2016
Published: August 18, 2016

REFERENCES

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

Atarashi, K., Tanoue, T., Ando, M., Kamada, N., Nagano, Y., Narushima, S., Suda, W., Imaoka, A., Setoyama, H., Nagamori, T., et al. (2015). Th17 cell induction by adhesion of microbes to intestinal epithelial cells. Cell *163*, 367–380.

Bal, S.M., Bernink, J.H., Nagasawa, M., Groot, J., Shikhagaie, M.M., Golebski, K., van Drunen, C.M., Lutter, R., Jonkers, R.E., Hombrink, P., et al. (2016). IL-1β, IL-4 and IL-12 control the fate of group 2 innate lymphoid cells in human airway inflammation in the lungs. Nat. Immunol. *17*, 636–645.

Bernink, J.H., Peters, C.P., Munneke, M., te Velde, A.A., Meijer, S.L., Weijer, K., Hreggvidsdottir, H.S., Heinsbroek, S.E., Legrand, N., Buskens, C.J., et al. (2013). Human type 1 innate lymphoid cells accumulate in inflamed mucosal tissues. Nat. Immunol. *14*, 221–229.

Bernink, J.H., Krabbendam, L., Germar, K., de Jong, E., Gronke, K., Kofoed-Nielsen, M., Munneke, J.M., Hazenberg, M.D., Villaudy, J., Buskens, C.J., et al. (2015). Interleukin-12 and -23 control plasticity of CD127(+) group 1 and group 3 innate lymphoid cells in the intestinal lamina propria. Immunity *43*, 146–160.

Bjorklund, A.K., Forkel, M., Picelli, S., Konya, V., Theorell, J., Friberg, D., Sandberg, R., and Mjosberg, J. (2016). The heterogeneity of human CD127 innate lymphoid cells revealed by single-cell RNA sequencing. Nat. Immunol. *17*, 451–460.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat. Methods *10*, 1213–1218.

Buonocore, S., Ahern, P.P., Uhlig, H.H., Ivanov, I.I., Littman, D.R., Maloy, K.J., and Powrie, F. (2010). Innate lymphoid cells drive interleukin-23-dependent innate intestinal pathology. Nature *464*, 1371–1375.

Colegio, O.R., Chu, N.Q., Szabo, A.L., Chu, T., Rhebergen, A.M., Jairam, V., Cyrus, N., Brokowski, C.E., Eisenbarth, S.C., Phillips, G.M., et al. (2014).

Functional polarization of tumour-associated macrophages by tumour-derived lactic acid. Nature *513*, 559–563.

Cording, S., Medvedovic, J., Cherrier, M., and Eberl, G. (2014). Development and regulation of RORγt(+) innate lymphoid cells. FEBS Lett. *588*, 4176–4181.

Dennis, G., Jr., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C., and Lempicki, R.A. (2003). DAVID: Database for annotation, visualization, and integrated discovery. Genome Biol. *4*, 3.

Diefenbach, A., Colonna, M., and Koyasu, S. (2014). Development, differentiation, and diversity of innate lymphoid cells. Immunity *41*, 354–365.

Eberl, G., Colonna, M., Di Santo, J.P., and McKenzie, A.N. (2015a). Innate lymphoid cells. Innate lymphoid cells: A new paradigm in immunology. Science *348*, aaa6566.

Eberl, G., Di Santo, J.P., and Vivier, E. (2015b). The brave new world of innate lymphoid cells. Nat. Immunol. *16*, 1–5.

Fuchs, A., Vermi, W., Lee, J.S., Lonardi, S., Gilfillan, S., Newberry, R.D., Cella, M., and Colonna, M. (2013). Intraepithelial type 1 innate lymphoid cells are a unique subset of IL-12- and IL-15-responsive IFN-γ-producing cells. Immunity *38*, 769–781.

Geremia, A., Arancibia-Cárcamo, C.V., Fleming, M.P., Rust, N., Singh, B., Mortensen, N.J., Travis, S.P., and Powrie, F. (2011). IL-23-responsive innate lymphoid cells are increased in inflammatory bowel disease. J. Exp. Med. *208*, 1127–1133.

Grün, D., Lyubimova, A., Kester, L., Wiebrands, K., Basak, O., Sasaki, N., Clevers, H., and van Oudenaarden, A. (2015). Single-cell messenger RNA sequencing reveals rare intestinal cell types. Nature *525*, 251–255.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

Hepworth, M.R., Monticelli, L.A., Fung, T.C., Ziegler, C.G., Grunberg, S., Sinha, R., Mantegazza, A.R., Ma, H.L., Crawford, A., Angelosanto, J.M., et al. (2013). Innate lymphoid cells regulate CD4+ T-cell responses to intestinal commensal bacteria. Nature *498*, 113–117.

Hepworth, M.R., Fung, T.C., Masur, S.H., Kelsen, J.R., McConnell, F.M., Dubrot, J., Withers, D.R., Hugues, S., Farrar, M.A., Reith, W., et al. (2015). Immune tolerance. Group 3 innate lymphoid cells mediate intestinal selection of commensal bacteria-specific $CD4^+$ T cells. Science *348*, 1031–1035.

Hoyler, T., Klose, C.S., Souabni, A., Turqueti-Neves, A., Pfeifer, D., Rawlins, E.L., Voehringer, D., Busslinger, M., and Diefenbach, A. (2012). The transcription factor GATA-3 controls cell fate and maintenance of type 2 innate lymphoid cells. Immunity *37*, 634–648.

Ivanov, I.I., Atarashi, K., Manel, N., Brodie, E.L., Shima, T., Karaoz, U., Wei, D., Goldfarb, K.C., Santee, C.A., Lynch, S.V., et al. (2009). Induction of intestinal Th17 cells by segmented filamentous bacteria. Cell *139*, 485–498.

Jaitin, D.A., Kenigsberg, E., Keren-Shaul, H., Elefant, N., Paul, F., Zaretsky, I., Mildner, A., Cohen, N., Jung, S., Tanay, A., and Amit, I. (2014). Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. Science *343*, 776–779.

Kelly, B., and O'Neill, L.A. (2015). Metabolic reprogramming in macrophages and dendritic cells in innate immunity. Cell Res. *25*, 771–784.

Kiss, E.A., Vonarbourg, C., Kopfmann, S., Hobeika, E., Finke, D., Esser, C., and Diefenbach, A. (2011). Natural aryl hydrocarbon receptor ligands control organogenesis of intestinal lymphoid follicles. Science *334*, 1561–1565.

Klose, C.S., Kiss, E.A., Schwierzeck, V., Ebert, K., Hoyler, T., d'Hargues, Y., Göppert, N., Croxford, A.L., Waisman, A., Tanriver, Y., and Diefenbach, A. (2013). A T-bet gradient controls the fate and function of CCR6-RORγt+ innate lymphoid cells. Nature *494*, 261–265.

Klose, C.S., Flach, M., Möhle, L., Rogell, L., Hoyler, T., Ebert, K., Fabiunke, C., Pfeifer, D., Sexl, V., Fonseca-Pereira, D., et al. (2014). Differentiation of type 1 ILCs from a common progenitor to all helper-like innate lymphoid cell lineages. Cell *157*, 340–356.

Koues, O.I., Collins, P.L., Cella, M., Robinette, M.L., Porter, S.I., Pyfrom, S.C., Payton, J.E., Colonna, M., and Oltz, E.M. (2016). Distinct gene regulatory pathways for human innate versus adaptive lymphoid cells. Cell *165*, 1134–1146.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

Lara-Astiaso, D., Weiner, A., Lorenzo-Vivas, E., Zaretsky, I., Jaitin, D.A., David, E., Keren-Shaul, H., Mildner, A., Winter, D., Jung, S., et al. (2014). Immunogenetics. Chromatin state dynamics during blood formation. Science *345*, 943–949.

Lavin, Y., Winter, D., Blecher-Gonen, R., David, E., Keren-Shaul, H., Merad, M., Jung, S., and Amit, I. (2014). Tissue-resident macrophage enhancer landscapes are shaped by the local microenvironment. Cell *159*, 1312–1326.

Levy, M., Thaiss, C.A., Zeevi, D., Dohnalová, L., Zilberman-Schapira, G., Mahdi, J.A., David, E., Savidor, A., Korem, T., Herzig, Y., et al. (2015). Microbiota-modulated metabolites shape the intestinal microenvironment by regulating NLRP6 inflammasome signaling. Cell *163*, 1428–1443.

Lim, A.I., Menegatti, S., Bustamante, J., Le Bourhis, L., Allez, M., Rogge, L., Casanova, J.L., Yssel, H., and Di Santo, J.P. (2016). IL-12 drives functional plasticity of human group 2 innate lymphoid cells. J. Exp. Med. *213*, 569–583.

Lochner, M., Peduto, L., Cherrier, M., Sawa, S., Langa, F., Varona, R., Riethmacher, D., Si-Tahar, M., Di Santo, J.P., and Eberl, G. (2008). In vivo equilibrium of proinflammatory IL-17+ and regulatory IL-10+ Foxp3+ RORgamma t+ T cells. J. Exp. Med. *205*, 1381–1393.

Mjösberg, J., Bernink, J., Golebski, K., Karrich, J.J., Peters, C.P., Blom, B., te Velde, A.A., Fokkens, W.J., van Drunen, C.M., and Spits, H. (2012). The transcription factor GATA3 is essential for the function of human type 2 innate lymphoid cells. Immunity *37*, 649–659.

Monticelli, L.A., Osborne, L.C., Noti, M., Tran, S.V., Zaiss, D.M., and Artis, D. (2015). IL-33 promotes an innate immune pathway of intestinal tissue protection dependent on amphiregulin-EGFR interactions. Proc. Natl. Acad. Sci. USA *112*, 10762–10767.

Monticelli, L.A., Buck, M.D., Flamar, A.L., Saenz, S.A., Tait Wojno, E.D., Yudanin, N.A., Osborne, L.C., Hepworth, M.R., Tran, S.V., Rodewald, H.R., et al. (2016). Arginase 1 is an innate lymphoid-cell-intrinsic metabolic checkpoint controlling type 2 inflammation. Nat. Immunol. *17*, 656–665.

Ohne, Y., Silver, J.S., Thompson-Snipes, L., Collet, M.A., Blanck, J.P., Cantarel, B.L., Copenhaver, A.M., Humbles, A.A., and Liu, Y.J. (2016). IL-1 is a critical regulator of group 2 innate lymphoid cell function and plasticity. Nat. Immunol. *17*, 646–655.

Paul, F., Arkin, Y., Giladi, A., Jaitin, D.A., Kenigsberg, E., Keren-Shaul, H., Winter, D., Lara-Astiaso, D., Gury, M., Weiner, A., et al. (2015). Transcriptional heterogeneity and lineage commitment in myeloid progenitors. Cell *163*, 1663–1677.

Pearce, E.L., and Pearce, E.J. (2013). Metabolic pathways in immune cell activation and quiescence. Immunity *38*, 633–643.

Rakoff-Nahoum, S., Paglino, J., Eslami-Varzaneh, F., Edberg, S., and Medzhitov, R. (2004). Recognition of commensal microflora by toll-like receptors is required for intestinal homeostasis. Cell *118*, 229–241.

Robinette, M.L., Fuchs, A., Cortez, V.S., Lee, J.S., Wang, Y., Durum, S.K., Gilfillan, S., and Colonna, M.; Immunological Genome Consortium (2015). Transcriptional programs define molecular characteristics of innate lymphoid cell classes and subsets. Nat. Immunol. *16*, 306–317.

Roediger, B., Kyle, R., Yip, K.H., Sumaria, N., Guy, T.V., Kim, B.S., Mitchell, A.J., Tay, S.S., Jain, R., Forbes-Blom, E., et al. (2013). Cutaneous immunosurveillance and regulation of inflammation by group 2 innate lymphoid cells. Nat. Immunol. *14*, 564–573.

Sano, T., Huang, W., Hall, J.A., Yang, Y., Chen, A., Gavzy, S.J., Lee, J.Y., Ziel, J.W., Miraldi, E.R., Domingos, A.I., et al. (2015). An IL-23R/IL-22 circuit regulates epithelial serum amyloid A to promote local effector Th17 responses. Cell *163*, 381–393.

Sanos, S.L., Bui, V.L., Mortha, A., Oberle, K., Heners, C., Johner, C., and Diefenbach, A. (2009). RORgammat and commensal microflora are required for

the differentiation of mucosal interleukin 22-producing NKp46+ cells. Nat. Immunol. *10*, 83–91.

Satoh-Takayama, N., Vosshenrich, C.A., Lesjean-Pottier, S., Sawa, S., Lochner, M., Rattis, F., Mention, J.J., Thiam, K., Cerf-Bensussan, N., Mandelboim, O., et al. (2008). Microbial flora drives interleukin 22 production in intestinal NKp46+ cells that provide innate mucosal immune defense. Immunity *29*, 958–970.

Sawa, S., Lochner, M., Satoh-Takayama, N., Dulauroy, S., Bérard, M., Kleinschek, M., Cua, D., Di Santo, J.P., and Eberl, G. (2011). RORγt+ innate lymphoid cells regulate intestinal homeostasis by integrating negative signals from the symbiotic microbiota. Nat. Immunol. *12*, 320–326.

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

Setty, M., Tadmor, M.D., Reich-Zeliger, S., Angel, O., Salame, T.M., Kathail, P., Choi, K., Bendall, S., Friedman, N., and Pe'er, D. (2016). Wishbone identifies bifurcating developmental trajectories from single-cell data. Nat. Biotechnol. *34*, 637–645.

Shekhar, K., Brodin, P., Davis, M.M., and Chakraborty, A.K. (2014). Automatic classification of cellular expression by nonlinear stochastic embedding (ACCENSE). Proc. Natl. Acad. Sci. USA *111*, 202–207.

Shih, H.Y., Sciumè, G., Mikami, Y., Guo, L., Sun, H.W., Brooks, S.R., Urban, J.F., Jr., Davis, F.P., Kanno, Y., and O'Shea, J.J. (2016). Developmental acquisition of regulomes underlies innate lymphoid cell functionality. Cell *165*, 1120–1133.

Silver, J.S., Kearley, J., Copenhaver, A.M., Sanden, C., Mori, M., Yu, L., Pritchard, G.H., Berlin, A.A., Hunter, C.A., Bowler, R., et al. (2016). Inflammatory triggers associated with exacerbations of COPD orchestrate plasticity of group 2 innate lymphoid cells in the lungs. Nat. Immunol. *17*, 626–635.

Sonnenberg, G.F., and Artis, D. (2015). Innate lymphoid cells in the initiation, regulation and resolution of inflammation. Nat. Med. *21*, 698–708.

Spencer, S.P., Wilhelm, C., Yang, Q., Hall, J.A., Bouladoux, N., Boyd, A., Nutman, T.B., Urban, J.F., Jr., Wang, J., Ramalingam, T.R., et al. (2014). Adaptation of innate lymphoid cells to a micronutrient deficiency promotes type 2 barrier immunity. Science *343*, 432–437.

Spitzer, M.H., and Nolan, G.P. (2016). Mass cytometry: Single cells, many features. Cell *165*, 780–791.

Thaiss, C.A., Zeevi, D., Levy, M., Zilberman-Schapira, G., Suez, J., Tengeler, A.C., Abramson, L., Katz, M.N., Korem, T., Zmora, N., et al. (2014). Transkingdom control of microbiota diurnal oscillations promotes metabolic homeostasis. Cell *159*, 514–529.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: Discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105–1111.

Treutlein, B., Brownfield, D.G., Wu, A.R., Neff, N.F., Mantalas, G.L., Espinoza, F.H., Desai, T.J., Krasnow, M.A., and Quake, S.R. (2014). Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. Nature *509*, 371–375.

van de Pavert, S.A., Ferreira, M., Domingues, R.G., Ribeiro, H., Molenaar, R., Moreira-Santos, L., Almeida, F.F., Ibiza, S., Barbosa, I., Goverse, G., et al. (2014). Maternal retinoids control type 3 innate lymphoid cells and set the offspring immunity. Nature *508*, 123–127.

Verrier, T., Satoh-Takayama, N., Serafini, N., Marie, S., Di Santo, J.P., and Vosshenrich, C.A. (2016). Phenotypic and functional plasticity of murine intestinal NKp46+ group 3 innate lymphoid cells. J. Immunol. *196*, 4731–4738.

von Burg, N., Chappaz, S., Baerenwaldt, A., Horvath, E., Bose Dasgupta, S., Ashok, D., Pieters, J., Tacchini-Cottier, F., Rolink, A., Acha-Orbea, H., and Finke, D. (2014). Activated group 3 innate lymphoid cells promote T-cell-mediated immune responses. Proc. Natl. Acad. Sci. USA *111*, 12835–12840.

Vonarbourg, C., Mortha, A., Bui, V.L., Hernandez, P.P., Kiss, E.A., Hoyler, T., Flach, M., Bengsch, B., Thimme, R., Hölscher, C., et al. (2010). Regulated expression of nuclear receptor RORγt confers distinct functional fates to NK cell receptor-expressing RORγt(+) innate lymphocytes. Immunity *33*, 736–751.

Walker, J.A., Barlow, J.L., and McKenzie, A.N. (2013). Innate lymphoid cells—how did we miss them? Nat. Rev. Immunol. *13*, 75–87.

Xu, W., and Di Santo, J.P. (2013). Taming the beast within: Regulation of innate lymphoid cell homeostasis and function. J. Immunol. *191*, 4489–4496.

Please cite this article in press as: Gury-BenAri et al., The Spectrum and Regulatory Landscape of Intestinal Innate Lymphoid Cells Are Shaped by the Microbiome, Cell (2016), http://dx.doi.org/10.1016/j.cell.2016.07.043

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-mouse B220 eFluor 450 (clone RA3-2B6)	eBioscience	Cat#48-0452-82
Anti-mouse CD19 eFluor 450 (clone 1D3)	eBioscience	Cat#48-0193-82
Anti-mouse CD3 eFluor 450 (clone 17A2)	eBioscience	Cat#48-0032-82
Anti-mouse GR-1 eFluor 450 (clone RB6-8C5)	eBioscience	Cat#48-5931-82
Anti-mouse CD45 Biotin (clone 30-F11)	Biolegend	Cat#103112
Anti-mouse IL-7Ra APC (clone A7R34)	Biolegend	Cat#135012
Anti-mouse NKp46 PE (clone 29A1.4)	Biolegend	Cat#137604
Anti-mouse KLRG-1 PerCP-Cy5.5 (clone 2F1)	Biolegend	Cat#368612
Streptavidin BV605	Biolegend	Cat#405229
Anti-Mouse Ly-6G/C 141Pr	Fluidigm	Cat#3141005
Anti-Mouse CD11c-142Nd	Fluidigm	Cat#3142003
Anti-Human/ Mouse IL-5 143Nd	Fluidigm	Cat#3143003
Anti-Mouse CD5 146Nd	This study	BLG-1006
Anti-Mouse CD45 147Sm	Fluidigm	Cat#3147003
Anti-Mouse CD11b 148Nd	Fluidigm	Cat#3148003
Anti-Human/ Mouse IL-22 151Eu	This study	BLG-5175
Anti-Mouse CD3e 152Sm	Fluidigm	Cat#3152004
Anti-Mouse CD335 153Eu	Fluidigm	Cat#3153006
Anti-Mouse CD19 154Sm	This study	BLG-1155
Anti-Mouse IL-17F 156Gd	This study	210-401-B31
Anti-Mouse RORgt 159Tb	Fluidigm	Cat#3159019
Anti-Human/ Mouse T-bet Gd160	This study	BLG-6448
Anti-Mouse Ly-6C Gd162	This study	BLG-1280
Anti-Mouse IFNg 165Ho	Fluidigm	Cat#3165003
Anti-Mouse IL-4 166Er	Fluidigm	Cat#3166003
Anti-Human/Mouse Gata-3-167Er	Fluidigm	Cat#3167007
Anti-Mouse IL-17A-169Tm	Fluidigm	Cat#3169005
Anti-Mouse NK1.1 170Er	Fluidigm	Cat#3170002
Anti-Mouse I-A/I-E 174Yb	Fluidigm	Cat#3174003
Anti-Mouse CD127 (IL7Ra) 175Lu	Fluidigm	Cat#3175006
Anti-Human/ Mouse CD45R (B220) 176Yb	Fluidigm	Cat#3176002
anti-H3K4me2	Abcam	Cat#ab32356
anti-H3K4me3	Millipore	Cat#07-473
anti-H3K27ac	Abcam	Cat#ab4729
Chemicals, Peptides, and Recombinant Proteins		
Vacomycin	Sigma-Aldrich	Cat#V1130
Ampicillin	Sigma-Aldrich	Cat#A1593
Kanamycin	Sigma-Aldrich	Cat#60615
Metronidazole	Sigma-Aldrich	Cat#M3761
Critical Commercial Assays		
Lamina propria dissociation kit	Miltenyi Biotech	Cat#130-097-410
Foxp3 / Transcription Factor Staining Buffer Set	eBioscience	Cat#00-5523-00

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited Data		
Raw data files for bulk RNA sequencing	NCBI Gene Expression Omnibus	GEO: GSE85154
Raw data files for ChIP sequencing	NCBI Gene Expression Omnibus	GEO: GSE85156
Raw data files for ATAC sequencing	NCBI Gene Expression Omnibus	GEO: GSE85153
Raw data files for single-cell RNA sequencing	NCBI Gene Expression Omnibus	GEO: GSE85152
Experimental Models: Organisms/Strains		
C57Bl/6JOlaHsd	Harlan	Cat#705
Rorc(γt)-Gfp^{TG}	(Lochner et al., 2008)	N/A
Sequence-Based Reagents		
Chromatin first indexing adaptor: CGATTGAGGCCGGT AATACGACTCACTATAGGGGCGACGTGTGCTCTTCC GATCTXXXXXXXXT modified with a C3 spacer (blocker) at the 5′. XXXXXXXX is the barcode for sample multiplexing.	This paper	N/A
ligation adaptor: XXXXNNNNNAGATCGGAAGAGCGTCG TGTAG modified with a phosphate group at 5′ and a C3 spacer (blocker) at the 3′. XXXX is the barcode for sample multiplexing.	This paper	N/A
Second: TCTAGCCTTCTCGCAGCACATC	This paper	N/A
P5_Rd1: AATGATACGGCGACCACCGAGATCTACACTC TTTCCCTACAC	This paper	N/A
P7_Rd2: CAAGCAGAAGACGGCATACGAGATGTGACTG GAGTTCAGAC	This paper	N/A
Software and Algorithms		
ACCENSE	(Shekhar et al., 2014)	http://www.cellaccense.com
NIH's Database for Annotation, Visualization and Integrated Discovery	(Dennis et al., 2003)	https://david.ncifcrf.gov/
TopHat (v2.0.10)	(Trapnell et al., 2009)	https://ccb.jhu.edu/software/tophat/index.shtml
HOMER software	(Heinz et al., 2010)	http://homer.salk.edu

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for reagents may be directed to, and will be fulfilled by the corresponding authors Eran Elinav (eran.elinav@weizmann.ac.il) and Ido Amit (ido.amit@weizmann.ac.il).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mice

C57Bl/6 RORc-GFP mice were previously described (Lochner et al., 2008). Germ-free C57Bl/6 mice were born in the Weizmann Institute germ-free facility and routinely monitored for sterility as previously described (Levy et al., 2015). In all experiments, adult female mice were used. Mice were provided with food and water ad libitum and housed under a strict 12 hr light-dark cycle. For antibiotic treatment, mice were given a combination of vancomycin (0.5 g/l), ampicillin (1 g/l), kanamycin (1 g/l), and metronidazole (1 g/l) in their drinking water (Rakoff-Nahoum et al., 2004; Thaiss et al., 2014). All antibiotics were obtained from Sigma Aldrich. Mice were 8-9 weeks of age at the beginning of antibiotic treatment experiments. All experimental procedures were approved by the local IACUC.

METHOD DETAILS

Isolation of ILCs from Small Intestinal Lamina Propria for Flow Cytometry

Mice were sacrificed by cervical dislocation, and small intestines were excised and transferred to ice-cold PBS. Adipose tissue and Peyer's patches were surgically removed. For sorting experiments, small intestines then were cleaned from fecal content and processed using a lamina propria dissociation kit (Miltenyi Biotech) according to the manufacturer's instructions. The isolated cells were washed with cold PBS and resuspended in PBS containing 1% BSA for direct cell surface staining. Single-cell suspensions for cell

sorting were stained on ice for 30 min with antibodies against CD45, CD3, CD19, Gr-1, CD127, NKp46, and KLRG-1. Cell sorting was performed using a BD-Fusion cell sorter.

For flow cytometry analysis experiments, tissues were opened and the lumen contents were removed by shaking in cold PBS (GIBCO; 3 times). Epithelial cells were removed by shaking tissue (200 rpm) in EDTA buffer (5mM EDTA, 2% FCS in DMEM; GIBCO) for 20 min at 37°C. Pieces of tissue were washed with PBS and digested in Liberase TL (25μg/ml; Roche), DNaseI (50μg/ml, Sigma Aldrich) and 2% FCS DMEM (GIBCO) solution at 37°C for 30 min with shaking (80 g). Remaining tissue was mechanically dissociated and the lymphocytes were separated by Percoll gradient. For the FACS analysis, cells were stained with antibodies to the following markers: CD45.2 (104); T-bet (eBio4B10), CD49d (R1-2), CD49a (Ha31/8), CD49b (DX5), CCL5 (2G9), IL-21R (eBio4A9), GATA-3 (TWAJ), CD127 (A7R34), KLRG1 (2F1), Sca-1 (D7), CD25 (PC61.5), NKp46, IL-4 (11B11), IL-5 (TRFK5), IA/IE (5M5/114.15.2), CD117 (ACK2), CCR6 (29-2L17), IL-17A (TC11-18H10), IL17F (9D3.1C8) and CD4 (RM4-5). All antibodies were from BD, eBioscience and Biolegend. Dead cells were excluded from analysis using eF506 viability stain (eBioscience). For experiment involving intracellular staining, cells were isolated from RORc-GFP *Rag2*$^{-/-}$ mice and stimulated for 3 hr with IL-2/12/18 (Figure 4B), IL-2/33 (Figure 4D), IL-1β/23 (Figure S5C), and PMA (10 ng/ml) and ionomycin (100 ng/ml) (Figures 4B, 4D, and S5C). For intranuclear staining, cells were fixed and permeabilized using Foxp3 staining kit (ebioscience). For cytokine production, cells were stimulated ex vivo by incubation for 3h with cytokines, PMA, ionomycin and Golgiplug (BD). The cells were fixed in 4% PFA in PBS and permeabilized using 0.3% saponin HBSS medium. FACS fortessa (BD) were used for cell acquisition and the flow cytometry data were analyzed with FlowJo 10 (TreeStar).

Mass Cytometry

Upon isolation of cells from the small intestinal lamina propria, cells were stained with metal-conjugated antibodies as previously described (Setty et al., 2016). Briefly, cells were stained with cell-ID TM Cisplatin (Fluidigm) (5 min RT). Next cells were stained with surface antibodies (30 min RT), and fixed with 1.6% PFA (10 min RT). After permeabilization with 100% ice-cold Methanol (15 min, 4°C), the cells were stained with intracellular antibodies (30 min, RT). Finally, the cells were labeled with Iridium DNA intercalator for DNA content and analyzed by CyTOF mass cytometry using CyTOF2. Data were normalized using bead normalized with bead standards. For tSNE analysis, ACCENSE was used (Shekhar et al., 2014).

Electron Microscopy

Mice were perfused with fixative containing 2% glutaraldehyde and 3% PFA in 0.1M sodium cacodylate. Intestinal samples were extensively washed from fecal matter and fixed for 24hrs. Samples were rinsed three times in sodium cacodylate buffer and postfixed in 1% osmiumtetroxide for 1hr, stained in 1% uranyl acetate for a further hour, then rinsed, dehydrated, and dried using critical point drying. Samples were then gold-coated and viewed in an ULTRA 55 FEG (ZEISS).

Bulk RNA Isolation, Library Construction, and Analysis

10^4-10^5 cells from each population were sorted into 100–200 mL of lysis/binding buffer (Life Technologies). mRNA was captured with 12 μl of Dynabeads oligo(dT) (Life Technologies), washed, and eluted at 70°C with 10 μl of 10 mM Tris-Cl (pH 7.5). We used a derivation of MARS-seq as described (Jaitin et al., 2014), developed for single-cell RNA-seq to produce expression libraries with a minimum of two replicates per population. We sequenced an average of 4 million reads per library and aligned them to the mouse reference genome (NCBI 37, mm9) using TopHat v2.0.10 (Trapnell et al., 2009) with default parameters. Expression levels were calculated and normalized using ESAT software (http://garberlab.umassmed.edu/software/esat).

RNA-Seq Processing and Analysis

We aligned the RNA-seq reads to the mouse reference genome (NCBI 37, mm9) using TopHat v2.0.13 with default parameters (Trapnell et al., 2009). Duplicate reads were filtered if they aligned to the same base and had identical UMIs. Expression levels were calculated and normalized for each sample to the total number of reads using HOMER software (http://homer.salk.edu) with the command "analyzeRepeats.pl rna mm9 -d [sample files] -count 3utr -condenseGenes" (Heinz et al., 2010). For the RNA-seq analysis in Figure 1, we focused on highly expressed genes with 2-fold differential over the noise (8 reads) between the means of any two subtypes (1821 genes). KEGG analysis was done using DAVID (Dennis et al., 2003).

iChIP-IVT

Cells were cross-linked for 8 min in 1% formaldehyde and quenched for 5 min in 0.125 M glycine prior to sorting. Sorted and frozen cell pellets were lysed in 0.5% SDS and sheared with the NGS Bioruptor Sonicator (Diagenode). Sheared chromatin was immobilized on 12 μl Dynabeads Protein G (Invitrogen) with 1.3 μg of anti-H3 antibody (ab1791). Magnetized chromatin was then washed with 10 mM Tris-HCl supplemented with 1X protease inhibitors. Chromatin was end repaired, dA-tailed and ligated with 5 ul of 0.75 uM partial Illumina Read2 sequencing adapters containing T7 polymerase promotor. Indexed chromatin was pooled, split to 3 IP pools and incubated with 2.5 μg anti-H3K4me2 antibody (ab32356)/anti-H3K4me3 (Millipore, 07-473)/anti-H3K27Ac (ab4729) at 4°C for 3h and for an additional hour with Protein G magnetic beads (Invitrogen). Magnetized chromatin was washed and reverse cross-linked. DNA was subsequently purified with 1.65X SPRI. In vitro transcription step of linear amplification were introduced to produce RNA transcripts out of the DNA fragments using the T7 High Yield RNA polymerase IVT kit (NEB). After IVT, DNase treatment

was performed to eliminate the DNA fragments. Next, a partial Illumina Read1 sequencing adaptor that includes a pool barcode was single strand ligated to the fragmented RNA using a T4 RNA ligase I (New England Biolabs). The ligated product was reverse transcribed using Affinity Script RT enzyme (Agilent) and a primer complementary to the ligated adaptor. The library was completed and amplified through a PCR reaction with 0.5 μM of each primer and PCR ready mix (Kapa Biosystems). The forward primer contains the Illumina P5-Read1 sequences and the reverse primer contains the P7-Read2 sequences. DNA concentration was measured with a Qubit fluorimeter (Invitrogen) and mean molecule size was determined with 2200 TapeStation analyzer (Agilent) and library quality was further determined by qPCR.

Primers Used during iChIP-IVT Library Construction

Primer name	Sequence and modifications
Chromatin first indexing adaptor	CGATTGAGGCCGGTAATACGACTCACTATAGGGGCGACGTGTGCTCTTCCGATCTXXXXXXXXT modified with a C3 spacer (blocker) at the 5′. XXXXXXXX is the barcode for sample multiplexing.
ligation adaptor	XXXXNNNNNAGATCGGAAGAGCGTCGTGTAG modified with a phosphate group at 5′ and a C3 spacer (blocker) at the 3′. XXXX is the barcode for sample multiplexing.
Second RT primer	TCTAGCCTTCTCGCAGCACATC
P5_Rd1 PCR forward	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT
P7_Rd2 PCR reverse	CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT

ATAC-Seq

To profile open chromatin, we used the ATAC-seq protocol developed by (Buenrostro et al., 2013) with modifications described by (Lara-Astiaso et al., 2014). In brief, cell populations were sorted in 400ul of MACS buffer (1x PBS, 0.5% BSA, 2mM EDTA) and pelleted by centrifugation for 15min at 500 g and 4°C with low acceleration and brake settings. Cell pellets were washed once with 1x PBS and cells were pelleted by centrifugation using the previous settings. Cell pellets were re-suspended in 25 μl of lysis buffer (10mM Tris-HCl [pH 7.4], 10mM NaCl, 3mM MgCl2, 0.1% Igepal CA-630) and nuclei were pelleted by centrifugation for 30min at 500 g, 4°C with low acceleration and brake settings. Supernatant was discarded and nuclei were re-suspended in 25 μl reaction buffer containing 2 μl of Tn5 transposase and 12.5 μl of TD buffer (Nextera Sample preparation kit from Illumina). The reaction was incubated at 37°C for one hour. Then 5 μl of clean up buffer (900mM NaCl, 300mM EDTA), 2 μl of 5% SDS and 2 μl of Proteinase K (NEB) were added and incubated for 30 min at 4°C. Tagmentated DNA was isolated using 2x SPRI beads cleanup. For library amplification, two sequential 9-cycle PCR were performed in order to enrich small tagmentated DNA fragments. We used 2 μl of indexing primers included in the Nextera Index kit and KAPA HiFi HotStart ready mix. After the first PCR, the libraries were selected for small fragments (less than 600 bp) using SPRI cleanup. Then a second PCR was performed with the same conditions in order to obtain the final library. DNA concentration was measured with a Qubit fluorometer (Life Technologies) and library sizes were determined using TapeStation (Agilent Technologies).

Sequencing

ATAC-Seq and ChIP-seq libraries (pooled at equimolar concentration) were sequenced using Illumina sequencers, NextSeq 500, at average sequencing depth per sample of 30 million reads for ATAC-Seq, and 15 million reads per sample for ChIP-seq.

Processing of ChIP-Seq and ATAC-Seq

Reads were aligned to the mouse reference genome (mm9, NCBI 37) using Bowtie2 aligner version 2.2.5 (Langmead et al., 2009) with default parameters. The Picard tool MarkDuplicates from the Broad Institute (http://broadinstitute.github.io/picard/) was used to remove PCR duplicates. To identify regions of enrichment (peaks) from ChIP-seq (H3K4me2, H3K4me3, H3K27ac) and ATAC-seq, we used the HOMER package makeTagDirectory followed by findPeaks command with the histone parameter or 500bp centered regions, respectively (Heinz et al., 2010). Union peaks file were generated for each of H3K4me2 and H3K4me3 by combining and merging overlapping peaks in all samples.

Chromatin and Motif Analysis

The read density (number of reads in 10 million total reads per 1000 bp) was calculated in each region from the union peaks files. We consider promoters to be H3K4me3 regions within ± 2000bp of a TSS (n = 8770). We defined 26,015 high confidence enhancers based on their presence in at least two replicates of H3K4me2 from the same population and their distal location (i.e., excluding promoters). The region intensity was given in log-base2 of the normalized density (log2(x+1)). Enhancer activity was determined by the intensity of H3K27ac in the region (active if greater than 4) from the relevant sample(s) resulting in approximately 2:1 ratio of poised to active enhancers. Kmeans clustering wae performed using MATLAB function *kmeans* with the distance metric set to "correlation." Enhancers were assigned to the nearest gene within 50kb. The relative contribution of enhancers to the gene expression (Figure S3B) was determined by counting the number of (active or poised) enhancers associated with genes in each cluster and dividing by the

total number in any cluster. In order to narrow down the region of interest within the enhancer to the most likely transcription factor binding locale, we overlapped the high confidence enhancers in each cluster (Figure 2B) or previously annotated enhancers that were 2-fold differential in SPF versus Abx (Figure 4G) with ATAC-seq peaks. Then, these narrowed regions were used as input for the HOMER package motif finder algorithm findMotifGenome.pl (Heinz et al., 2010).

Gene Tracks and Normalization

All gene tracks were visualized as bigWig files of the combined replicates normalized to 10,000,000 reads and created by the HOMER algorithm makeUCSCfile (Heinz et al., 2010). For visualization, the tracks were smoothed by averaging over a sliding window of 500 bases.

Single-Cell Sorting

Isolated cells were single-cell sorted into 384-well cell capture plates containing 2 μl of lysis solution and barcoded poly(T) reverse-transcription (RT) primers for single- cell RNA-seq. Barcoded single cell capture plates were prepared with a Bravo automated liquid handling platform (Agilent) as described previously (Jaitin et al., 2014). Four empty wells were kept in each 384-well plate as a no-cell control during data analysis. Immediately after sorting, each plate was spun down to ensure cell immersion into the lysis solution, snap frozen on dry ice and stored at −80°C until processed.

Index Sorting

To record marker levels of each single cell, the FACS Diva 7 "index sorting" function was activated during single-cell sorting. During index sorting, single cells were sorted from the entire $Lin^{-}CD45^{+}CD127^{+}$ population instead of gating for a specific ILC subtype. The intensities of RORγt, KLRG-1, and NKp46 were recorded and linked to each cell's position within the 384-well plate.

MARS-Seq Library Preparation

Single-cell libraries were prepared as previously described (Jaitin et al., 2014). Briefly, mRNA from cells sorted into MARS-seq capture plates were barcoded and converted into cDNA and pooled using an automated pipeline. The pooled sample was then linearly amplified by T7 in-vitro transcription and the resulting RNA was fragmented and converted into sequencing ready library by tagging the samples with pool barcodes and Illumina adaptor sequences during ligation followed by reverse transcription and PCR. Each pool of cells was tested for library quality and concentration was assessed as described earlier (Jaitin et al., 2014).

Single-Cell Analysis

MARS-seq reads were processed as previously described (Paul et al., 2015), with a few changes. Mapping of reads was done using hisat (version 0.1.6) and reads with multiple mapping positions were excluded. Reads were associated with genes if they were mapped to an exon defined by a reference set obtained from the UCSC genome browser. Exons of different genes that share genomic position on the same strand are considered as a single gene with concatenated gene symbol. Cells with less than 200 UMIs were discarded from the analysis. Genes with mean expression smaller than 0.005 UMIs/cell or with above average expression and low coefficient of variance (< 1.2) were also discarded.

In order to assess the heterogeneity of innate lymphocyte subtypes, we used a recently published batch-aware multinomial mixture-model clustering algorithm (Paul et al., 2015). A brief summary of the algorithm is described below.

Low-level processing of MARS-Seq reads results in a matrix U with n rows and m columns, where rows represent genes and columns represent cells. Entry U_{ij} contains the number of unique molecular identifiers (UMIs) from gene i that were found in cell j. The model assumes that each cell belongs to one of K cell types, and that each cell type defines a different distribution of transcripts within cells.

Our model assumes that cells are sampled uniformly from the population, and that each cell type dictates a multinomial distribution over the sample of sequenced RNA molecules. The model consists of a three types of parameters:

$\boldsymbol{map_j}$ - The assignment of cell j to one of K cell types.

$\alpha_{i,\boldsymbol{map}^j}$ - The probability of observing gene i in cell j, assuming that j belongs to cell type $\boldsymbol{map}_j$.

β_{ib^j} - A positive inflation factor accounting for batch effect on the expression of gene i (b^j is the batch of cell j).

A pseudo EM algorithm is used to infer the assignment of cells to types, gene probability within cell type, and magnitude of batch effect. The algorithms outline is as follows:

1. Initialize the model:
   a. Estimate β (batch effect vector) from the gene expression in each batch.
   b. Sample a first seed for the cell types at random by drawing uniformly from the list of cells.
2. Repeat a-d to generate the desired number of clusters:
   a. Initialize a pre-seed model using the regularized transcripts of the seed cell.
   b. Find the set of D cells with the highest likelihood to the pre-seed model. Using these cells, initialize a new set of parameters α and optimize the likelihood of the selected cells given these parameters and given constant batch parameters. Optimization

is done using non-linear optimization procedure (L-BFGS-B). Add the optimized parameters as a new component to the model.

c. Compute the log-likelihood of each cell to each of the current initialized types. Assign each cells to its corresponding maximum-likelihood cell type.

d. Sample a new seed cell at random by drawing from quantiles 0-0.05 of the maximum likelihoods obtained in c.

3. Given the current set of multinomial models, calculate the assignment for each cell ($\boldsymbol{map}_j$) by calculating the maximum assignment probability.
4. Given the current β and *map* parameters, use L-BFGS-B to find α (for each cell type) that maximizes the likelihood of the U matrix.
5. Given the current α and the map parameters, use L-BFGS-B to find β that maximizes the likelihood of the U matrix.
6. Return to step 2 and repeat until the likelihood function converges, or the maximum number of iterations is reached.

QUANTIFICATION AND STATISTICAL ANALYSIS

In each experiment, multiple mice were analyzed as biological replicates: n = 6 mice for data reported in Figure 1; n = 2 mice for data reported in Figure 2; n = 1129 cells from 3 individual mice for data reported in Figure 3; four independent experiments in flow cytometry data shown in Figures 4 and 5; n = 6 and 2 for SPF and Abx mice, respectively, in Figure 6; and n = 706 and 944 cells from 3 individual Abx and germ-free mice, respectively, for data reported in Figure 7. Graphs show mean ± SEM. Hypergeometric testing was used to assess statistically significant enrichments. * indicates $p < 0.05$, unless denoted otherwise.

DATA AND SOFTWARE AVAILABILITY

Data Resources

The accession number for the sequencing data reported in this paper is GEO: GSE85157. This parent directory includes the following data sets: GSE85156 (ChIP-seq), GSE85154 (RNA-seq), GSE85153 (ATAC-seq), and GSE85152 (single cell RNA-seq).

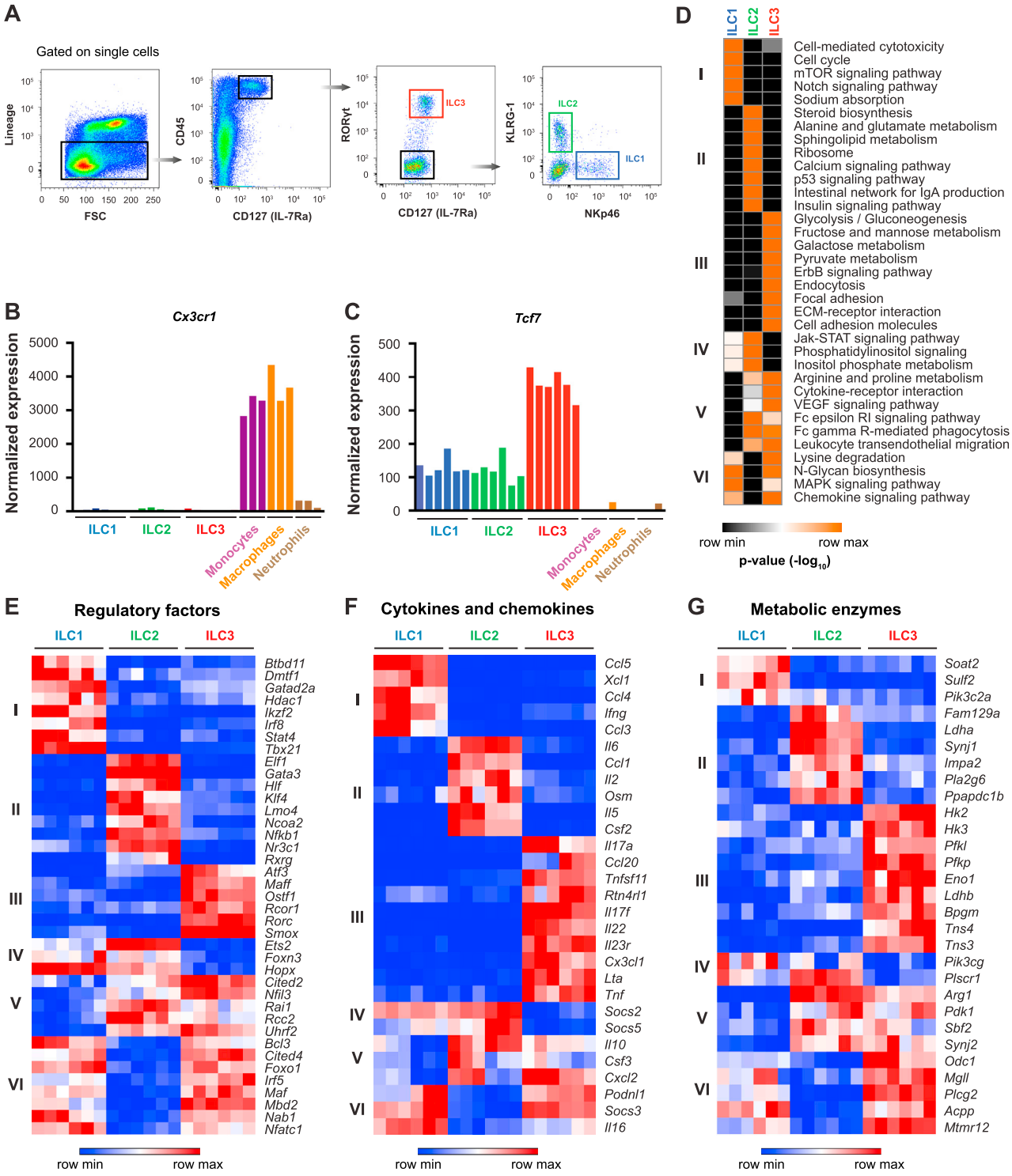


Figure S1. Transcriptional Characteristics of ILC Subsets, Related for Figure 1

(A) FACS plots showing gating strategy for sorting of canonical ILCs subtypes.

(B and C) Comparative analysis of gene expression in ILCs and myeloid cells.

(D) KEGG pathway annotation for each of the 6 clusters in Figure 1B.

(E–G) Heatmap showing gene expression in selected examples of regulatory factors (E), cytokines and chemokines (F), and metabolic enzymes (G) from each of the gene clusters in Figure 1B.

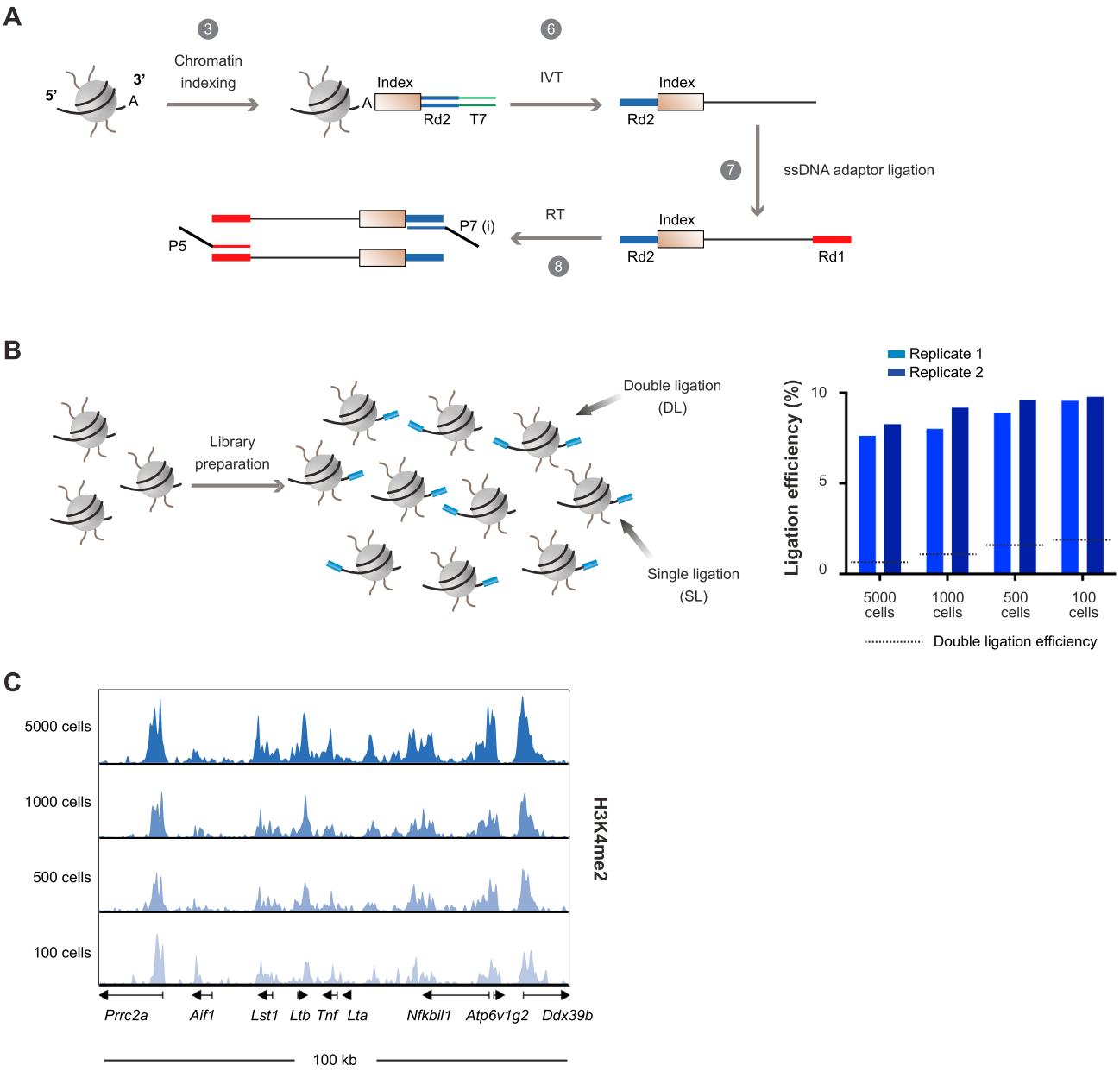


Figure S2. A Modified Protocol for High-Sensitivity In Vivo Chromatin Analysis, Related to Figure 2

(A) Schematic view of important steps in the iChIP-IVT protocol. Step 3: indexing adaptor including T7 promotor. Step 6: linear amplification by T7 *in-vitro* tanscriptase (IVT). Step 7: ssDNA second adaptor ligation followed by reverse transcription and PCR library amplification.

(B) Quantification of ligation efficiency. Schematic indicating adaptor ligation efficiency. Approximately 10% of molecules undergo adaptor ligation events. Bar graph shows ligation efficiency quantification of iChIP-IVT libraries, 2 replicates of 4 different cell numbers. Dashed line indicates double ligation efficiency.

(C) Normalized H3K4me3 profiles of peaks found in the 100 kb region of the tumor necrosis factor (TNF) locus obtained with iChIP-IVT of titrated amounts of bone marrow-derived dendritic cells.

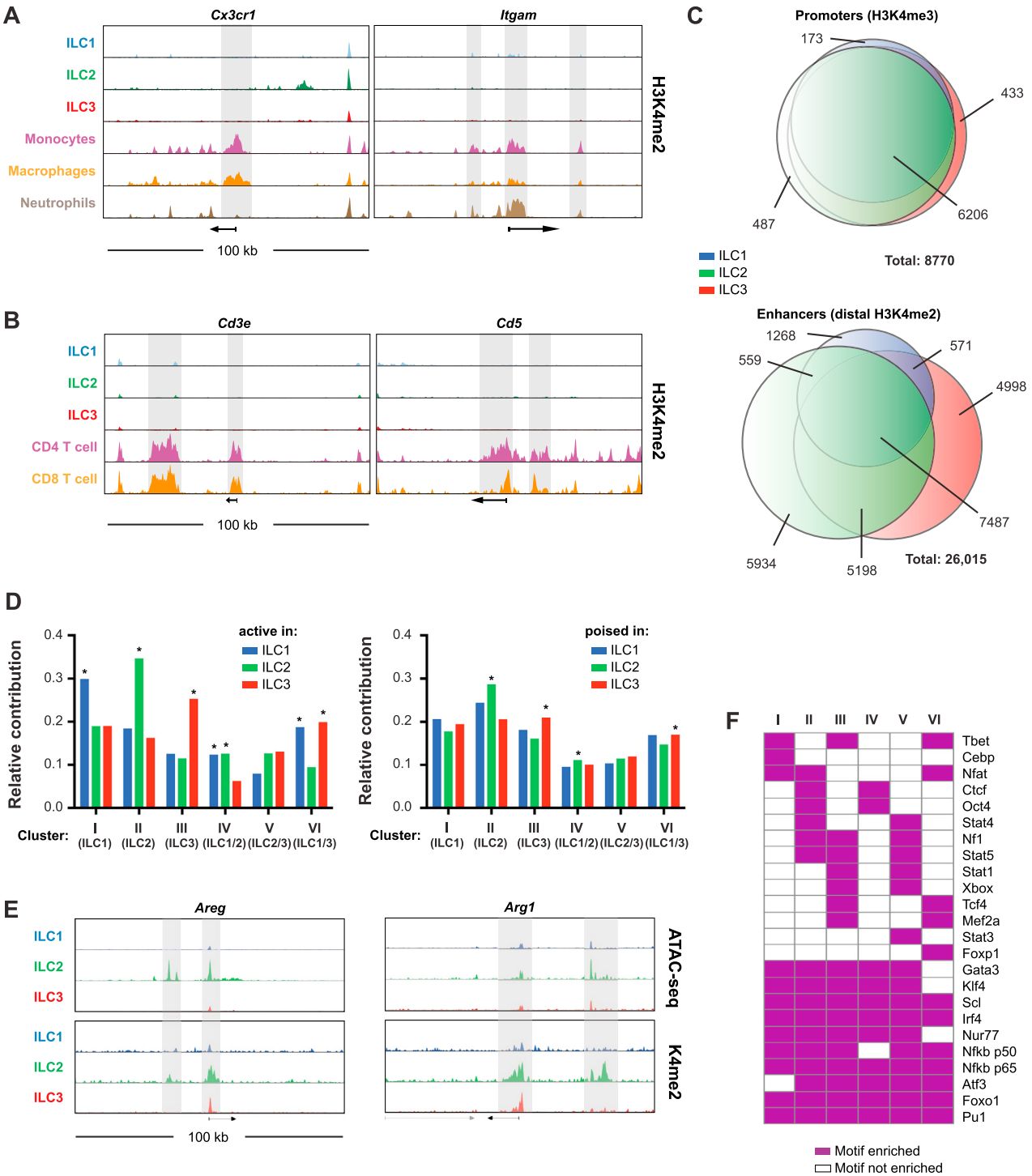


Figure S3. Chromatin Regulation of ILC-Specific Gene Expression, Related to Figure 2

(A and B) Individual examples of normalized profiles of H3K4me2 signals in 100 kilobase (kb) regions around the indicated genes in ILCs compared to other immune cells.

(C) Venn diagram of the overlap of promoters (left) and enhancers (right) among the three ILCs subtypes.

(D) Relative contribution of active (left) and poised (right) enhancer regions to the expression clusters in Figure 1B. * indicates significance ($p < 0.05$, hyper-geometric distribution)

(E) Individual examples of normalized profiles of H3K4me2 and ATAC-seq signal in 100 kilobase (kb) regions around the *Areg* (left) and *Arg1* (right) loci.

(F) Heatmap of significantly enriched motifs ($p < 10^{-5}$) in ATAC-seq peaks of H3K4me2-marked enhancer regions from each cluster in Figure 2B.

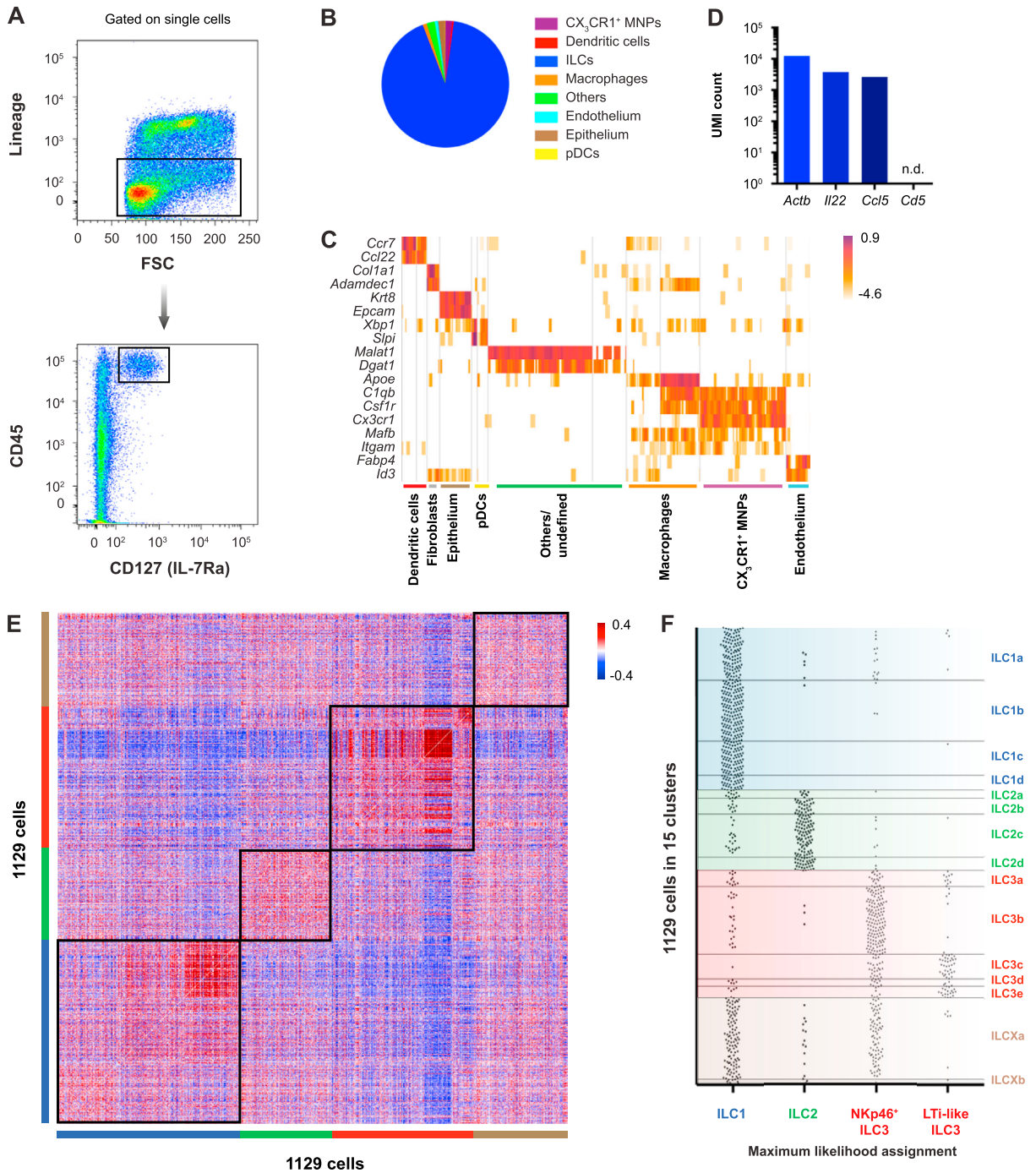


Figure S4. The Single-Cell Landscape of Small Intestinal ILCs, Related to Figure 3

(A) FACS plots showing gating strategy for indexed single-cell sorting of ILCs.

(B) Quantification of cell types obtained by the gating strategy in (A).

(C) Heatmap showing marker genes for all non-ILCs shown in (B).

(D) UMI count of the indicated genes across all single cells.

(E) Pairwise correlation between the transcriptomes of 1129 single cells. The color-coding indicates post-clustering inference of subset membership (blue = ILC1; green = ILC2; red = ILC3; brown = ILCx).

(F) Maximum likelihood-based assignment of each single-cell transcriptome of the 15 ILC clusters to bulk transcriptional clusters of ILC1, ILC2, NKp46^{+} ILC3, and LTi-like ILC3.

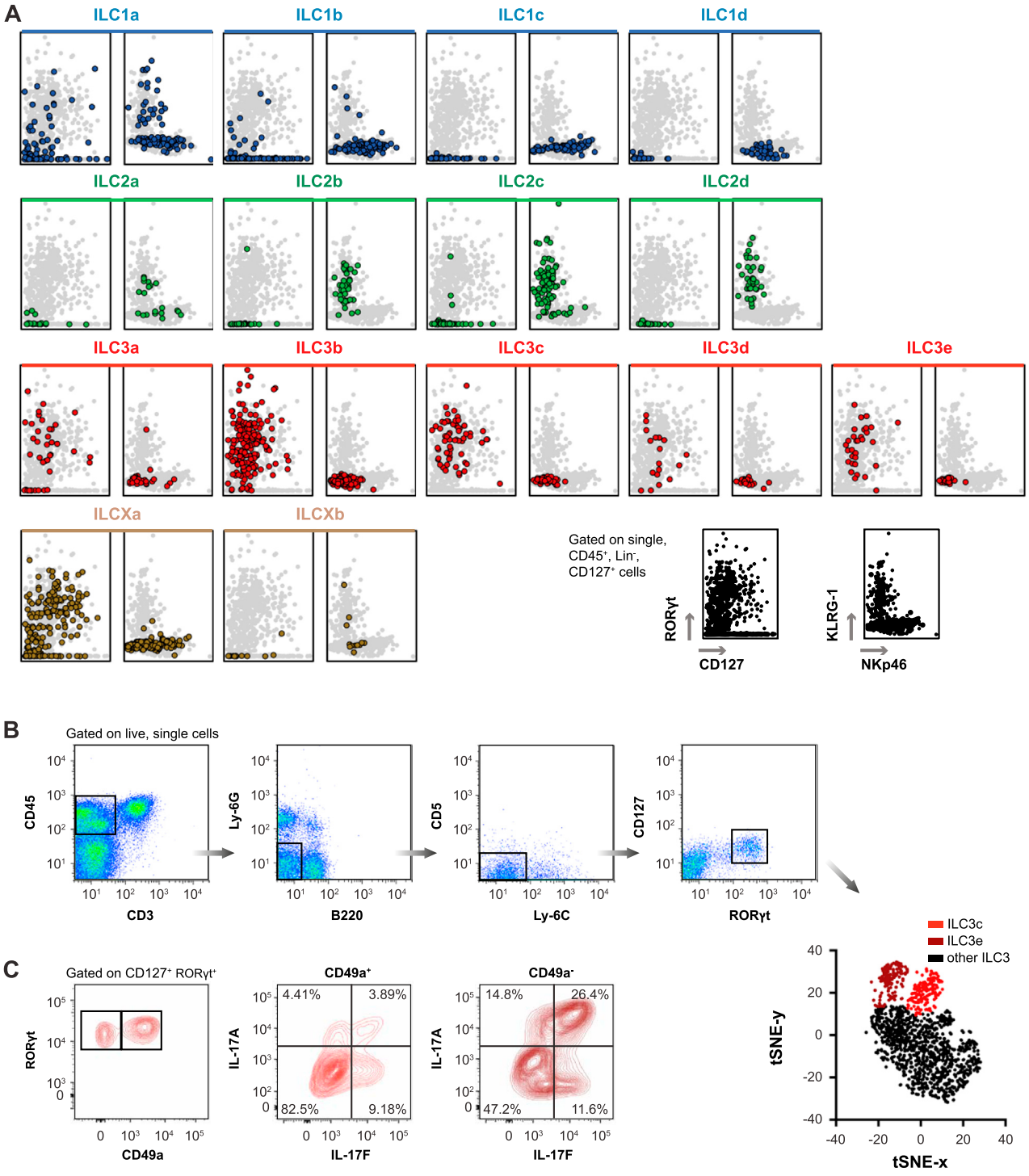


Figure S5. Gating Analysis of Index-Sorted and Mass Cytometry, Related to Figures 3 and 5

(A) In silico gating of index-sorted populations after clustering by gene expression. Indexing data of single cell clusters is projected on FACS plots, reflecting the heterogeneity of surface marker expression among all 15 clusters. Cells are gated on CD45$^+$ Lin$^-$ CD127$^+$ cells.

(B) Gating strategy and tSNE analysis of mass cytometry data of RORγt$^+$ ILCs. ILC3c and ILC3e were identified based on expression of IL-22 and MHC-II, respectively.

(C) Flow cytometry analysis of CD49a$^+$ and CD49a$^-$ ILC3s expressing IL-17A and IL-17F.

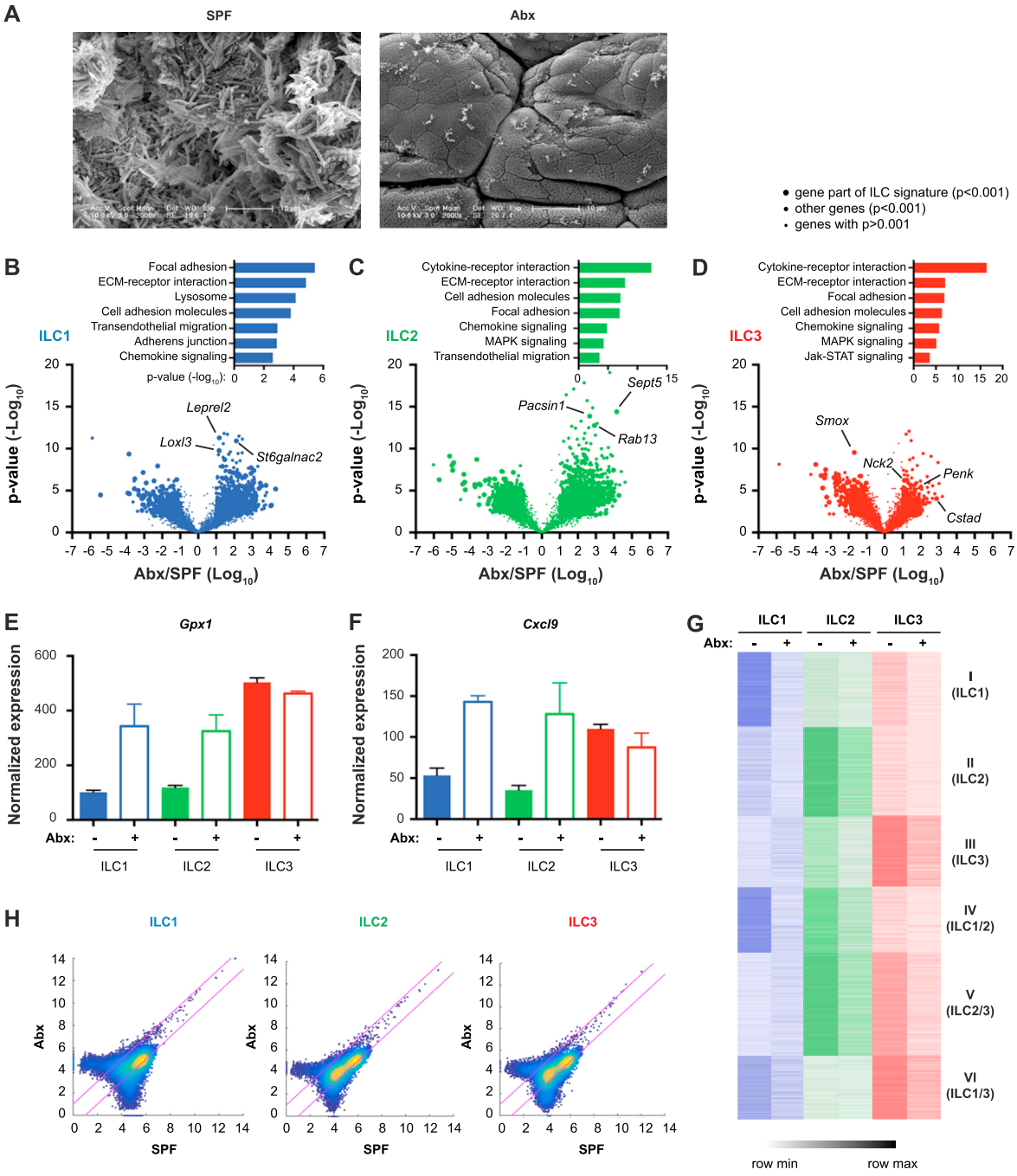


Figure S6. Transcriptome and Epigenome Regulation by the Microbiota, Related to Figure 6

(A) Scanning electron micrographs of mouse intestines in SPF condition (left) and microbiota-depleted condition after antibiotic treatment (Abx; right).

(B–D) Volcano plots comparing gene expression of the three ILC populations in SPF and Abx mice. Dot size indicates significance of expression changes. Significantly changed genes belonging to ILC signature genes in Figure 1B are marked with the largest dot size. Examples indicate genes from ILC3-associated signature. Bar graphs indicate KEGG pathway annotation of significantly altered genes.

(E and F) Examples of ILC3-associated genes upregulated in ILC1 and ILC2 upon antibiotic treatment. Error bars indicate SEM.

(G) Heatmap of H3K4me2 intensity in clustered enhancers (distal regions from Figure 2B) in ILC subsets from SPF and Abx-treated mice. Columns represent merged replicates.

(H) Scatterplot comparing H3Kme2 intensity in all enriched regions from each ILC subset in SPF and Abx-treated mice. Pink lines indicate 2-fold change.

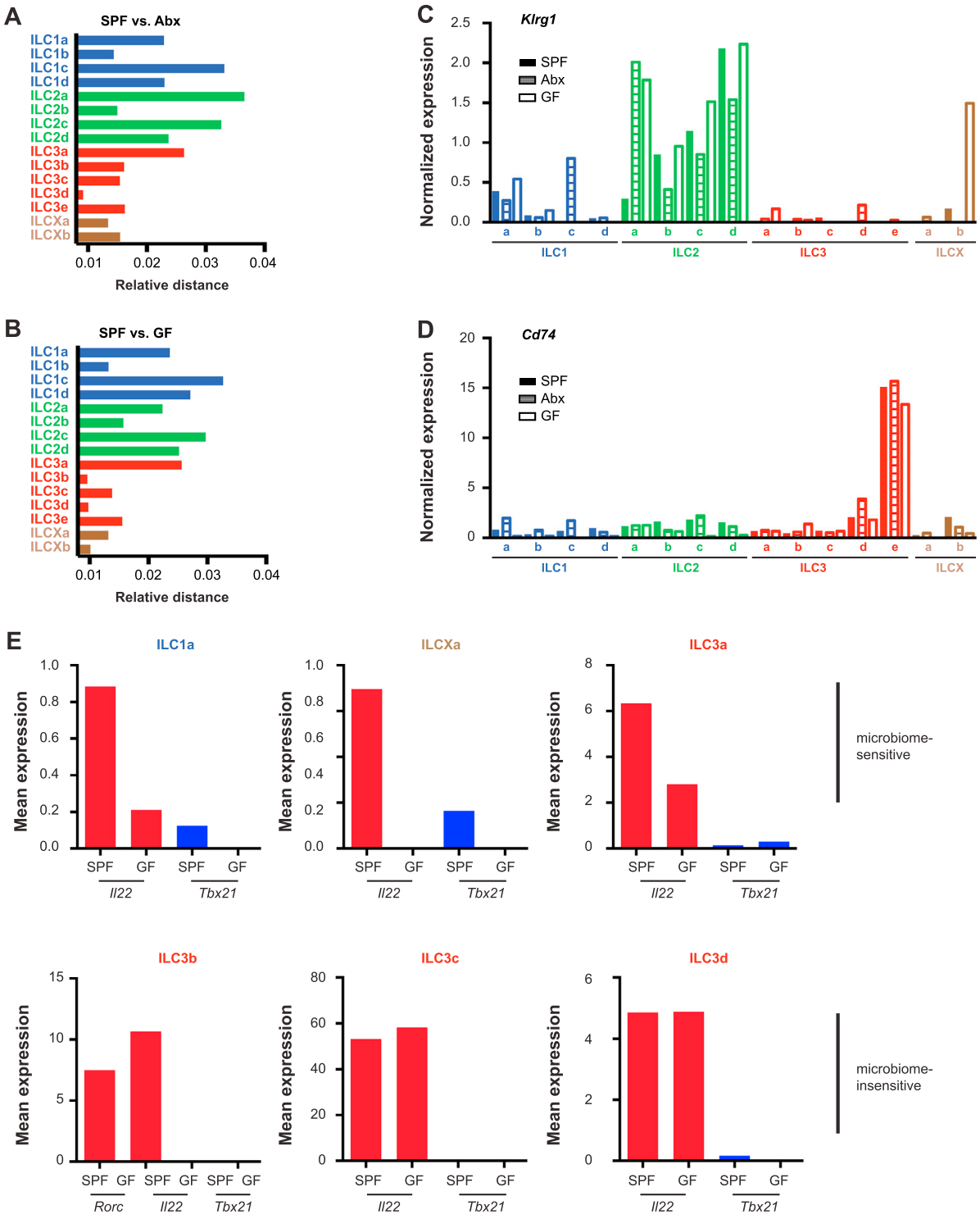


Figure S7. Single-Cell Cluster Responses to the Intestinal Microbiota, Related to Figure 7

(A and B) Bar graph of Euclidian distances between the gene expression across 15 ILC subpopulations in SPF, Abx, and GF mice.

(C and D) Bar graphs showing gene expression of individual genes across 15 ILC subpopulations in WT, Abx, and GF mice.

(E) Cluster-specific effects of the microbiota on transcript levels of *Il22* and *Tbx21*.