HHS Public Access

Author manuscript

Cell. Author manuscript; available in PMC 2017 May 19.

Published in final edited form as:

Cell. 2016 May 19; 165(5): 1120–1133. doi:10.1016/j.cell.2016.04.029.

Developmental Acquisition of Regulomes Underlies Innate Lymphoid Cell Functionality

Han-Yu Shih[1,5], **Giuseppe Sciumè**[1,5], **Yohei Mikami**[1,5], **Liying Guo**[2], **Hong-Wei Sun**[3], **Stephen R. Brooks**[3], **Joseph F. Urban Jr.**[4], **Fred P. Davis**[1], **Yuka Kanno**[1], and **John J. O'Shea**[1]

Han-Yu Shih: han-yu.shih@nih.gov; John J. O'Shea: John.Oshea@nih.gov

[1]Lymphocyte Cell Biology Section, Molecular Immunology and Inflammation Branch, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, MD, 20892, USA

[2]Laboratory of Immunology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, 20892, USA

[3]Biodata Mining and Discovery Section, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, MD, 20892, USA

[4]Diet, Genomics, and Immunology Laboratory, Beltsville Human Nutrition Research Center, Agricultural Research Service, United States Department of Agriculture, Beltsville, MD 20705, USA

Summary

Innate lymphoid cells (ILCs) play key roles in host defense, barrier integrity, and homeostasis, and they mirror adaptive CD4+ T helper (Th) cell subtypes in both usages of effector molecules and transcription factors. To better understand the relationship between ILC subsets and their Th cells counterparts, we measured genome-wide chromatin accessibility. We find that chromatin in proximity to effector genes is selectively accessible in ILCs prior to high-level transcription upon activation. Accessibility of these regions is acquired in a stepwise manner during development and changes little after in vitro or in vivo activation. Conversely, dramatic chromatin remodeling occurs in naïve CD4+ T cells during Th cell differentiation using a type 2-infection model. This

Correspondence to: Han-Yu Shih, han-yu.shih@nih.gov; John J. O'Shea, John.Oshea@nih.gov.

[5]Co-first authors.

Accession numbers

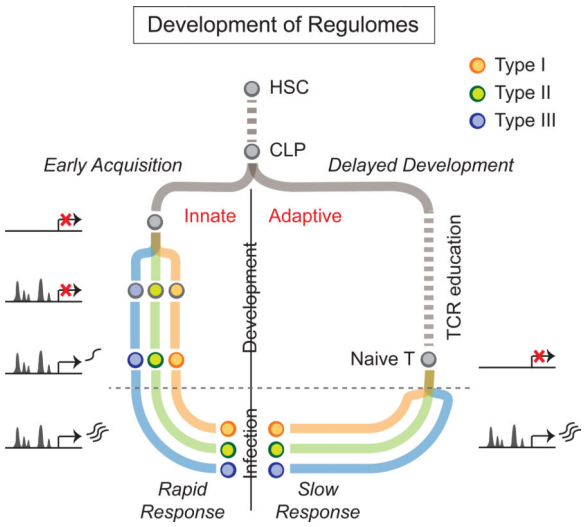
The ATAC-seq, RNA-seq and ChIP-seq data have been deposited in Gene Expression Omnibus (GEO) under the accession number GSE77695.

Author Contributions

H-Y.S. and G.S. designed the project and wrote the manuscript. H-Y.S., G.S. and Y.M. performed and interpreted all the experiments. H-Y.S. performed ATAC-seq experiments and computational analyses. Y.M. performed RNA-seq experiments. J.F.U. and L.G. helped *N. brasiliensis* infection experiments. H-W.S., S.R.B., and F.P.D. helped genomic data processing and analyses. Y.K. and F.P.D. contributed to data interpretation and manuscript editing. J.J.O'S. contributed to project design, data interpretation, and manuscript writing.

alteration results in a substantial convergence of Th2 cells toward ILC2 regulomes. Our data indicate extensive sharing of regulatory circuitry across the innate and adaptive compartments of the immune system, in spite of their divergent developing pathways.

Graphical abstract



Introduction

The immune system orchestrates host defense through complex effector networks mediated by an array of lymphocytes, including conventional T, B, and natural killer (NK) cells, along with an array of recently recognized innate lymphoid cells (ILCs) (Artis and Spits, 2015; Diefenbach et al., 2014; Eberl et al., 2015; Sonnenberg and Artis, 2015). Unlike T and B cells that mediate adaptive immunity against pathogenic microbes in an antigen-specific manner, ILCs respond to invaders promptly in the absence of somatically rearranged antigen receptors. Three classes of ILCs are presently recognized and categorized based on their selective cytokine-production profiles, mirroring previously identified $CD4^+$ Th cell subsets (Spits et al., 2013; Verykokakis et al., 2014). Group 1 ILCs includes conventional NK cells, the first identified ILC subset, along with ILC1s, which lack the cytotoxicity capability of NK cells. Both of these cells selectively produce IFN-γ, the key cytokine that defines T helper 1 (Th1) cells. Group 2 ILCs (encompassing ILC2) preferentially produce cytokines such as interleukin (IL)-5, IL-13 and IL-9, originally defined as Th2 cytokines. Finally, group 3 ILCs are a heterogeneous subset that comprise natural cytotoxicity receptor (NCR)-positive ILC3s and CD4-positive ILC3s (also known as lymphoid tissue inducer-like cells) that produce IL-17 and/or IL-22, the namesake cytokines of Th17 and/or Th22 cells.

Several important issues remain unresolved, including the regulatory mechanisms underlying ILC development, diversification and terminal differentiation, and how these mechanisms compare to those of T helper (Th) cell subsets. Like T and B lymphocytes, ILCs are derived from common lymphoid progenitors (CLPs), and further specified by an array of transcription factors (TFs) (De Obaldia and Bhandoola, 2015; Kang and Malhotra, 2015; Klose and Diefenbach, 2014). The transcriptional regulator inhibitor of DNA binding

2, Id2, for instance, counteracts the effects of E proteins to limit the development of T and B lymphocytes. Other TFs such as Nfil3, Plzf, Tox, Tcf7 and Runx3 are also involved in the lineage divergence during ILC development (Serafini et al., 2015). However, consistent with their selective cytokine production, ILCs also use the same lineage-determining transcription factors (LDTFs) that drive cognate T cell lineage specification (Shih et al., 2014; Spits et al., 2013). For instance, T-box transcription factors, including Eomesodermin and T-bet (encoded by *Eomes* and *Tbx21* genes, respectively), are involved in the specification of all IFN-γ-producers, whereas Th2 and Th17 master regulators, GATA-binding protein 3 (GATA-3) and retinoic acid receptor-related orphan receptor-γt (RORγt) are essential for the development of group 2 and 3 ILCs, respectively. However, the extent to which the ontogeny of ILCs truly parallels Th cell specification, especially at the genomic level, remains poorly understood.

Beyond the assessment of selective cytokine production and enumeration of LDTFs, the relationships between lineages can also be probed with genomic tools. Both microarray and RNA-seq have been extensively used to delineate cell type-specific transcriptomes (Kim and Lanier, 2013; Shay and Kang, 2013). Recently reported ILC transcriptomes suggest that the tissue microenvironment also has a substantial impact on gene expression profiles beyond lineage *per se* (Robinette et al., 2015). Thus, defining cell identity by transcriptome requires careful considerations of the local “environmental” factors and tissue residency.

Another strategy of determining cell fate and lineage relationships is to analyze global epigenetic information, which in contrast to gene expression, can be more stable and propagate information over time during development and differentiation (Lara-Astiaso et al., 2014). Epigenetic codes, including DNA methylation, histone modifications and chromatin accessibility, together construct unique chromatin landscapes at non-coding regulatory elements (REs), which contribute to gene expression by permitting or restricting access of transcriptional machinery to key loci.

It is now appreciated that distinct lineages exhibit thousands of highly distinctive genomic “switches”, which act in concert to govern tissue-specific and temporal control of gene expression. Among epigenomic elements, enhancers are intriguing due to their ability to control gene expression at a distance and contribute to lineage specificity (Heinz et al., 2015). Genome-wide enhancer distribution has been mapped in various lineages based on the characteristics of chromatin accessibility, histone modifications, and TF binding. Accumulating data reveals that the basal epigenomes (prior to cell activation) encode cell fate information and are progressively specified in response to developmental cues and environmental stimuli (Lara-Astiaso et al., 2014; Stergachis et al., 2015; Vahedi et al., 2012). Recent studies on macrophages highlight the environmental impact on tissue-specific chromatin states (Gosselin et al., 2014; Lavin et al., 2014). However, the contribution of development versus environment to the fate of ILC identity has not been systematically characterized at the genomic level.

In this study, we set out to answer a number of questions related to ILC ontogeny and regulation and their relationships with cognate adaptive immune cells. Using genomic tools, we comprehensively identified REs that comprise the lineage-specific regulome in

conjunction with measuring transcriptomes of prototypical ILCs and their progenitors. Our genome-wide analysis revealed that each ILC lineage possesses unique open chromatin landscapes and conform to the general view of ILC1, ILC2 and ILC3 as distinct lineages. These features were relatively static after ILC activation, despite dynamic changes in gene expression, revealing the poised status of ILCs prior to stimulation. The presence of lineage-specific REs in ILC precursors indicates that ILC functionality is pre-determined as the cells diverge into definite lineages. In contrast, naïve T cells exhibit markedly different chromatin landscapes that change dramatically after activation and final differentiation. Nonetheless, their regulomes converge with those of ILCs. The substantial overlap of REs between Th and ILCs is striking given their distinct routes of development. Together our data provide mechanistic underpinnings for ILC lineage commitment, acquisition of poised functionalities, as well as the relationships between innate and adaptive compartments.

Results

The chromatin landscapes of innate lymphocyte lineages reflect their distinct functionalities

ILCs have been categorized into 3 major groups based on selective cytokine production that parallels Th cell subsets (Spits et al., 2013). To better understand the regulatory logic of ILCs, we globally identified REs in the major types by an assay for transposase-accessible chromatin using sequencing (ATAC-seq). This method requires few cells, enabling us to characterize the regulomes of cells directly isolated *ex vivo* (Buenrostro et al., 2013). In parallel, we also assessed transcriptomes by RNA-seq. We first analyzed five prototypical ILC subsets, including conventional NK cells from spleen and ILC1 from liver (group 1); ILC2 from small intestine lamina propria (siLP, group 2); $CD4^+$ (LTi-like ILC3) and NCR^+ ILC3s both from siLP (group 3). These subsets were chosen based on the possibility to distinguish their identity using accepted surface markers (Figures 1A and S1A) and were confirmed by the expression of their LDTFs (Figure S1B).

We first examined chromatin landscapes of ILC signature cytokines that were differentially expressed in ILCs (Figures 1B **and** S1C–E). We found that the lineage-specific chromatin landscapes correlated with recognized functionality of each subset. The distinct patterns of accessibility encompassed not only promoters, but also intragenic and intergenic regions, extending as far as several kilobases away (**highlighted by blue and red, respectively**). These accessible regions included REs previously characterized in the *Ifng, Il4/5/13, Il17 and Il22* loci (Figures 1B and S1D, **marked by red triangles** (Balasubramani et al., 2010; Wilson et al., 2009)), as well as multiple lineage-specific REs not previously identified. Most of these REs appear to be putative enhancers as they co-localized with p300 and T-bet binding in NK cells, the latter being a critical TF known to regulate *Ifng* (Figure 1B). Notably, three *Ifng*-associated REs were specific for NK cells, and absent in ILC1 (Figure 1B, **red arrows**) suggesting that the same locus might be differentially regulated in distinct IFN-γ producing cells. On the other hand, an ILC2-specific RE 40 kb downstream of the *Ifng* gene showed GATA-3 binding in ILC2 (Zhong et al., 2015), suggesting the possibility of regulation by LDTFs, which antagonize alternative cell fates. We also observed that distinct REs at LDTF loci were accessible even in absence of their expression, which may

indicate opportunities for functional plasticity (Figures 1C **and** S1F–G). For example, a portion of *Rorc* and *Il17* REs was accessible in ILC2; this phenomenon may explain previous observations of IL-17 production in these cells (Huang et al., 2014). In addition, type 3 ILCs have been reported to express T-bet and produce low levels of IFN-γ (Bernink et al., 2013; Klose et al., 2014; Rankin et al., 2013; Sciumé et al., 2012); accordingly our data revealed that a portion of IFN-γ REs, including the promoter and distal T-bet binding sites, were accessible in NCR^+ILC3s but not in $CD4^+$ILC3s. Taken together, high-resolution profiling of chromatin landscapes by ATAC-seq identified lineage-specific REs near genes that specify ILC functions.

Global views of ILC chromatin landscapes reveal differential regulomes

Next, we sought to obtain a global picture of ILC regulomes by analyzing accessible REs genome-wide. Among a total of 82,305 accessible REs merged from five prototypical ILCs, about a quarter (25%) was common to all subsets, whereas the majority (75%) of the regions were either unique to a single cell type or shared by a subset of ILCs (here termed variable regions) (Figures 2A and 2B). As expected, differential RE accessibility was correlated with selective gene expression in ILC subsets (Figure S2A). Analyzing the genomic distribution of ILC REs revealed that promoters (REs within +/− 1kb of transcription start sites) were preferentially enriched among common REs (30%), but were depleted from variable REs (5%) (Figure 2C). This observation suggests that the divergence of ILC chromatin landscapes appear to be primarily shaped by distal REs. *For each lineage, we defined ATAC accessible regions that are not ubiquitously accessible in all lineages as their "signature REs", which differ from "specific REs" that are exclusively accessible in only one lineage*.

We next investigated potential TFs that could target ILC regulomes by searching for the enrichment of consensus TF motifs within signature REs. We found that target motifs for recognized LDTFs were differentially enriched among each ILC type, consistent with their known roles in driving ILC development (Figure 2D, **highlighted by red**). For instance, the T-box motif, which is shared by T-bet and Eomes, was prominent in group 1 ILC regulomes and depleted from those of ILC2s and ILC3s. Conversely, GATA-3 and RORγt motifs were enriched in group 2 and group 3 ILCs, respectively. In addition to known LDTFs, we also identified other TFs potentially important for ILC specification that mirrored the enrichment pattern of known LDTFs (Figure 2D). For instance, the enrichment of Runx motifs in group 1 and group 3 ILCs supports the very recent evidence of the essential role of Runx family members in their development (Ebihara et al., 2015). Since TFs belonging to a given family share the same motifs, we further filtered potential regulators based on their expression to identify both shared and unique regulatory networking modules for each ILC subset (Figure S2B). This analysis also recapitulated the fact that the TFs previously associated with NK cell development and/or function such as Ets and IRF family members (highlighted in the light red box) contribute to ILC regulation (Barton et al., 1998; Lohoff et al., 2000). In addition, the NF-κB family members were identified in group 2 and 3 ILC regulatory networks (highlighted in light blue box), correlated with the importance of IL1-β and IL-25/IL-33 functions in these subsets. Overall, these data point to distinctive regulomes that may reflect occupancy of key TFs contributing to ILC functionality and identity.

Having established the distinctive regulomes in ILCs, we next sought to determine how regulomes compared with transcriptomes in discerning ILC lineage identity. Our data indicate that comparisons of accessible chromatin landscapes of different ILC groups reveal greater differences (Pearson correlation r=0.62–0.78) than comparisons of expression profiles (Pearson correlation r=0.79–0.91) (Figure 2E **and** S2C). Specifically, ILC transcriptomes revealed the highest similarity among group 1 cells (NK and ILC1) and group 3 cells ($CD4^+$ and NCR^+ ILC3s) (Figure S2C, top panel). In addition, cells residing in the same tissue (intestinal ILC2 and group 3 ILCs) also displayed high similarity. However, the comparison of regulomes provided a rather different view (Figure S2C, bottom panel). The differences among subsets in the same ILC group were more pronounced, and similarities of ILC2s and ILC3s were less obvious. In particular, comparing the regulomes of ILC2 and ILC3 subsets suggests that these cell types are more distinct than expected based on the similarities of their transcriptomes (Figure 2E).

ILC regulomes are primed prior to activation

An important feature of ILCs is their ability to quickly respond to external cues and rapidly induce transcription and protein synthesis of effector genes to mediate host defense. Whether this process involves the rapid acquisition of new enhancers or utilization of pre-existing poised enhancers has not been determined. To understand the dynamics of enhancer landscapes and their impact on ILC activation, we stimulated NK cells, ILC2 and NCR^+ILC3 cells with relevant cytokines for 4–6 hours and compared the changes in transcriptomes and regulomes with their resting state (Figure 3A and 3B). *We observed more genes were down-regulated than those were up-regulated in NK cells and NCR^+ILC3 (> 2-fold change, p-value < 0.05) upon stimulation, whereas similar numbers were up- and down-regulated in ILC2 cells* (Figure 3B, **left panel)**. *We also observed concordant changes in the number of ATAC-accessible sites that were accessible before and after stimulation* (Figure 3B, **right panel**). *This observation suggests that the dynamic changes of transcriptomes and regulomes might be well correlated*. However, further global correlation analysis indicated that this was only partly true (Figures 3C **and** S3A). For instance, the transcriptomes of stimulated ILC2 and NCR^+ILC3 cells were more similar to each other than those of unstimulated cells (Figure 3C). In contrast, assessment of their respective regulomes revealed that the distinctive REs were relatively stable, congruent with subset identity.

We next explored the dynamics of REs near cytokine loci that are rapidly induced upon stimulation (Figures 3D **and** S3B–C). We found that even though transcript abundance increased dramatically, the chromatin landscapes of the *Ifng*, type 2 cytokine, and *Il22* loci changed little after activation. To further evaluate enhancer activity, we also investigated the dynamics of p300 binding and H3K27 acetylation levels in stimulated NK cells by ChIP-seq (Figure 3D). We found that most *Ifng* enhancers identified by ATAC-seq were pre-bound by p300 (12 of 13 sites) and were also H3K27-acetylated (7 of 13 sites) prior to activation. Upon stimulation, these sites exhibited enhanced p300 binding as well as increased H3K27-acetylation whereas ATAC signals remained unchanged. The observed dynamic enhancer activity was consistent with enhanced cytokine production from the locus. This suggests that

chromatin landscapes of stimulation-responsive elements are pre-determined before the cell is activated.

ILC Enhancer Landscapes Diverge Early in Development

Our finding that the enhancer landscapes of terminally differentiated ILCs are poised prior to cytokine stimulation raises the question of when these lineage-signature landmarks were initially established during development. To answer this, we profiled regulomes of developing ILCs, including immature NK (iNK), NK precursor (NKp), and ILC2 precursor (ILC2p) cells from bone marrow, and compared them with hematopoietic stem cells (HSCs), multipotent progenitor cells (MPPs), and common lymphoid progenitor cells (CLPs) (Figures 4A **and** S4A). The designation of NKp refers to lin^{-} $CD122^{hi}$ cells enriched for cells committed to the NK fate (Chiossone et al., 2009; Constantinides et al., 2014; Fathman et al., 2011; Hoyler et al., 2012; Rosmaraki et al., 2001). Comparison of these cells revealed stepwise loss of HSC signature REs and acquisition of ILC signature REs during ILC development (Figures 4B **and** S4B). The loss of ~50% of the HSC signature REs had already occurred in MPPs, with a further ~10% loss in CLPs and 20% loss in ILC precursors (NKp and ILC2p). Conversely, only 25–30% of ILC-signature features were present in early progenitors (HSCs, MPPs and CLPs), whereas another ~30% of these features was acquired as CLPs developed to NKp or ILC2p (Figure 4B). Of note, the later step involved minimum loss of HSC signature REs. In aggregate, these data indicate that the regulatory landscapes of ILC precursors are at states that adopt ILC-subset signature chromatin landscapes while retaining progenitor features.

Next, we compared the global views of ILC lineage identity defined by regulomes versus transcriptomes. Hierarchical clustering of either genome-wide gene expression or chromatin accessibility of ten cell types revealed relationships between bone marrow progenitors (HSCs, CLPs and MPPs) as well as peripheral mature ILCs (Figure 4C). However, assessment of transcriptomes revealed a different picture of NKp and ILC2p than regulomes. The transcriptomes of bone marrow ILC2p and NKp are highly correlated with each other and cluster with early progenitors (Figures 4C **and** S4C, **left panel**). On the other hand, comparing regulomes clustered bone marrow ILC2p with differentiated gut ILC2. (Figures 4C **and** S4C, **right panel**). Similarly, bone marrow NKp and iNK cells clustered with mature, splenic NK cells and ILC1s. One explanation for this difference is that ILC regulomes were established prior to terminal differentiation. To test this possibility, we examined the accessible chromatin landscapes of ILC signature genes like *Ifng* and Th2 cytokines. Indeed, the majority of lineage-signature REs was already accessible at precursor stages (Figures 4D and 4E), supporting our hypothesis.

To further understand how ILC regulomes were formed during development and how relevant they are to the ILC function, we compared REs among mature ILCs (NK and ILC2), ILC precursors (NKp and ILC2p) and CLPs (Figures 5A and 5B). We first identified REs gained and lost during ILC development, dividing them as early and late events (categories A and B for gained REs; categories C and D for lost REs). REs that were gained during ILC development were enriched for the motifs of specific LDTFs. In contrast, REs that were lost during ILC specification were enriched for motifs of early progenitor and

myeloid-associated TFs, such as Pu.1 and Spi-B. Remarkably, the transitioning REs, specifically accessible only in precursors, were enriched with motifs of both T-box and GATA families. This observation suggests the existence of a plastic stage prior to final differentiation in which titration of the level of these two TF families could determine differentiated cell identity.

Next, we analyzed the expression level of genes in proximity to differentially accessible REs (Figures 5C and 5D). As expected, we found that the genes associated with regions that only became accessible at the final developmental stage (category A) were also induced late. However, the genes with REs in ILC precursor cells (category B) showed expression trend similar to category A, suggesting REs of these genes were pre-deposited prior to terminal differentiation. Similarly, among the genes upregulated at the final differentiation stages, over half possessed pre-established REs (Figures 5E, 5F **and** S5). These results reinforce the notion that lineage-specific chromatin landscapes diverge early during ILC development.

Relationships between innate and adaptive lymphoid cells

ILC nomenclature was originally proposed on the basis of effector functions and LDTF expression reflecting cognate T cell subsets. However, unlike ILCs initiating both development and differentiation in bone marrow, T cells develop in the thymus and then further differentiate upon stimulation. To understand the relationships between ILC and T cell regulomes, we performed ATAC-seq on *ex vivo* isolated T cells in healthy mice, including naïve and memory $CD8^+$ T cells from bone marrow and naïve $CD4^+$ T cells from spleen. For memory/effector $CD4^+$ T cells, we isolated Th17 (GFP^+) cells from the intestine of Il17-GFP mice, using $CCR6^+CD25^-$, $CCR6^-CD25^+$ and $CCR6^-CD25^-$ $CD4^+$ T cells as controls.

To systematically compare the regulatory landscapes of different lineages, we performed hierarchical clustering on ATAC-seq similarity of in-house and published ATAC-seq (Figure 6). The latter dataset includes macrophages, microglia, and dendritic cells, as well as NK cells. This analysis revealed several features that suggest how regulomes can be useful for exploring cell identity. First, the analysis recapitulated our previous findings of the similarity between ILC precursors and their differentiated progeny (Figure 4C). Second, the analysis illustrated the relatively low impact of environment in determining characteristic regulomes of type 1 ILCs in different tissues (liver, bone marrow, spleen), which clustered together irrespective of the organ from which they were isolated. Third, the analysis showed ILCs (clade 1) having a closer relationship with T cells (clade 2) than with B cells and hematopoietic progenitors (clade 3), or with myeloid cells (clade 4). Finally, the results reinforced the current view of ILCs as distinct groups of cells with well-defined markers.

ILC and T Helper Chromatin Landscapes Converge Upon Infection

Given the clustering of T cells and ILCs in healthy hosts, we next compared their chromatin landscapes following infection. Naïve $CD4^+$ T cells require potent polarizing stimuli for effector differentiation; however, the similarity of effector ILC and Th cell transcriptomes and regulomes has not been assessed. To resolve this issue, we sought to obtain ILCs and T cells subjected to the same environment by infecting mice with *Nippostrongylus brasiliensis*,

a parasitic nematode provoking a Th2-dominant immune response (Finkelman et al., 2004). Both ILC2 and Th2 cells were enriched in lungs after 10 days of infection and then isolated for ATAC-seq and RNA-seq analysis (Figure 7A). By comparing REs of Th2-related genes (*Il4, Il5, Il9, Il10, and Il13*) in *N. brasiliensis* infected ILC2 and Th2 cells, we were struck by the similarities of the chromatin landscapes between these two cells, in spite of their differences prior to infection (Figure 7B). Notably, the majority of these REs were generated *de novo* in *N. brasiliensis*-infected Th2 cells compared to ILCs in which more modest changes were evident upon infection. Genome-wide analysis of accessible REs in *N. brasiliensis* induced Th2 cells revealed that these cells lost over half of naïve signature REs and two-third of Th2 REs were newly acquired (Figure 7C). Among these Th2-acquired REs, over 70% were also detectable in ILC2, demonstrating that Th2 cells gained a large portion of ILC2 regulomes upon infection. Consistently, pairwise comparison of gene expression amongst type 2 subsets, including ILC2p from bone marrow, ILC2 from small intestine, and both ILC2 and Th2 cells from lung of *N. brasiliensis*-infected mice, revealed minimal expression difference between *N. brasiliensis*-infected ILC2 and Th2 cells (Figure 7D). In contrast, the maximum difference in gene expression was observed between Th2 cells and naïve $CD4^+$ T cells. Hierarchical clustering emphasized the difference between the pre-established ILC and naïve $CD4^+$ T cell regulomes, although infection led to a convergence of circuitry (Figure 7E). To better define the basis of the convergence, we performed gene ontology analysis on GREAT (Ashburner et al., 2000; McLean et al., 2010) to identify genes that shared ATAC accessibility in infected Th2 and ILC2 cells. Comparison of accessible peaks present in both cell types relative to the background of the whole genome revealed significant enrichment for several immune response-related molecular functions. An additional explanation for the convergence Th2 and ILC2 regulomes might be similar changes in metabolic state; however, we found no significant enrichment in general metabolism or cell cycle terms (Figure S6). Therefore, the environmental impact of infection was able to synchronize gene regulation in ILC2 and Th2 cells, despite the greater impact on the latter. This implies substantial overlap of their regulatory networks even though they were established through distinct routes.

Discussion

In the present study, we sought to gain insight into the biology of ILCs by characterizing regulomes for the prototypic subsets and determining their roles in gene expression. Our data reveal how ILC regulomes “mature” progressively during development such that many key loci are primed prior to terminal differentiation. As a consequence, these loci are only moderately impacted by stimulation *in vitro* and *in vivo*. This contrasts with $CD4^+$ Th cells that undergo dramatic chromatin remodeling upon activation. However, during the course of infection, the regulomes of $CD4^+$ T cells and ILCs closely approximate one another, arguing for substantial sharing of mechanisms underlying regulation of lineage-specific functions.

Genomic views of ILC classification

With the expanding recognition of innate lymphocyte subsets, a classification of ILCs was proposed based on their analogy with $CD4^+$ Th cell subsets. However, the precise distinctions between conventional NK cells and ILC1 have been unclear and the

heterogeneity of ILC3s has challenged a simple classification. The identification of ILC3 expressing T-bet and the overlapping roles of LDTFs in ILC differentiation further complicate the picture (Tindemans et al., 2014). While the present classification posits that ILCs like helper T cells are specified to distinct “lineages”, the ability to comprehensively map the regulatory landscape of ILCs raises the question of how the current functional classification compares to a genomic perspective. The five recognized major subsets of ILCs, conventional NK, ILC1, ILC2, NCR^+ ILC3, and $CD4^+$ ILC3 cells can be discerned by chromatin landscapes and approximate in three main groups; although, the differences between NK and ILC1 cells and ILC3s are also apparent. The mechanisms establishing the similarities and differences in chromatin landscapes will be important to discern in the future, and the comprehensive regulomes defined in the current work will facilitate this endeavor.

Underpinnings of rapid effector responses in ILCs

A general characteristic of ILCs is their rapid and selective responses to infection. Locksley and colleagues first recognized the accessibility of the *Ifng* promoter in NK cells, as measured by histone acetylation (Stetson et al., 2003). Consistently, our global analysis revealed that the REs of effector gene loci were accessible in resting ILCs and changed little following activation, despite the increase of enhancer activity measured by p300 and histone acetylation. A striking finding of our work is that most of these REs are pre-formed in ILC precursors and become accessible in a stepwise manner during development. It should be noted that chromatin accessibility detected by ATAC-seq includes promoters, silencers and insulators as well as enhancers; however, we would argue that a significant proportion of ATAC accessible regions are enhancers, due to the fact that 40% of all ATAC peaks in NK cells are bound by p300, a useful proxy for enhancer activity in previous studies (Visel et al., Nature 2009) despite its functional redundancy with other co-activators (eg. CBP, Pcaf).

Environment versus ontogeny of ILC subsets

ILCs are important for barrier function where they are exposed to diverse exogenous and endogenous environmental stimuli; consequently, ILCs exhibit distinct functionalities in different tissues. This correlates well with the recently reported distinctive transcriptomes of intestinal ILCs (Robinette et al., 2015). In other innate cells, such as macrophages, environment also controls both gene expression and enhancers to define tissue-specific macrophage identities (Gosselin et al., 2014; Lavin et al., 2014). In view of these previous observations, we were struck by the contrast in viewing ILC identity by transcriptomes versus regulomes. ILC regulomes appear to be less sensitive to tissue localization and primarily reflect lineage relationships. In particular, regulomes define the lineage segregation between ILC2 and ILC3 better than transcriptome analyses. Also, comparison of transcriptomes revealed similarities between ILC2 and NK precursors, while inspection of regulomes revealed more similarities between precursors and their offspring. Understanding the factors that establish chromatin landscapes will be revealing; presumably, this is the consequence of the action of TFs that are either induced or activated during development. A better understanding of the regulatory logic of these developmental circuits will be important in elucidating the creation of these landscapes.

Mechanisms allowing ILC plasticity

While selective cytokine production is a major feature of the current classification of ILCs, substantial phenotypic plasticity has recently become evident (Bernink et al., 2013; Cella et al., 2009; Huang et al., 2014; Kearley et al., 2015; Klose et al., 2014; Serafini et al., 2015). Signature cytokine loci are most accessible in the expected subsets. However, it is clear that there are elements that are broadly accessible in all subsets. For instance, CNS-22, an element previously shown to be involved in activation-specific IFN-g induction in T-cells (Balasubramani et al., 2014), was accessible in all ILCs, including those not making this cytokine. Perhaps such elements represent “seed” enhancers (Factor et al., 2014), which have permissive actions on the general organization of the locus and precede selective, high level IFN-γ expression. Even more notable is that the broad accessibility of loci encoding LDTFs across ILCs provides a mechanism for phenotypic flexibility in the context of permitting rapid responses. Alternatively, the accessibility of loci encoding LDTF provide opportunities for cross-regulation. Taken together, the data make clear that despite mechanisms allowing selective cytokine production, multiple means exist to allow plasticity and flexible expression of effector genes.

T cells and ILCs: Regulomes Converge following Infection

The existing classification of ILCs and helper T cells implies functional relationships between them. However, the extent to which they truly share mechanisms to regulate common batteries of effector genes remains an important question that has not been previously examined. In principle, this question can be addressed both in terms of development (ontogeny) and evolution (phylogeny). From the view of development, our data indicate that ILCs are instructed in the bone marrow and undergo a stepwise process of specification. The unbiased view of the regulomes of naïve $CD4^+$ T cells would lead one to believe that they are distinct from ILCs, suggesting that their ability to acquire effector functions might involve mechanisms rather distinct from ILCs. This is consistent with a reasonably clear distinction between ILCs and T cell development (Shih et al., 2014). Nonetheless, following infection, the regulomes of ILCs and T cells approximate each other to a remarkable degree. This suggests that mechanisms underlying selective effector gene expression are shared between T cells and ILCs and may be evolutionarily ancient. While we know relatively little about the evolution of lymphocytes, the existence of two distinct modes of antigen-specific recognition in lymphocytes in vertebrate evolution (Hirano et al., 2011) could suggest that “innate” lymphoid cells may have preceded T cells evolutionarily (Serafini et al., 2015). Precisely why T cell development is associated with the absence of a pre-primed or poised regulome is not clear; however, this clearly allows for more room to control the system in which massive clonal expansion occurs along with acquisition of effector functions, rather than preexisting functionalities during infection.

Conclusion

The discovery of diverse, functionally specialized ILC subsets represents a major advance in our knowledge of how the immune system copes with infection, inflammation, tissue repair, and metabolic homeostasis. The apparent functional symmetry of the innate and adaptive systems makes the efforts to understand the molecular relationship between ILCs and T cells

of great interest. Our elucidation of the regulomes of ILCs and T cells provides insights into the genomic mechanisms that specify functions of different lineages. Deciphering the precise nature of circuits that shape the regulomes of ILC and T cells is an exciting area for future work.

Experimental procedures

Mice

Female wild-type C57BL/6J, Foxp3-GFP and Il17-GFP mice (6- to 12-week-old) were purchased from Jackson Laboratory. All animal studies were performed according to the National Institutes of Health guidelines for the use and care of live animals and were approved by the Institutional Animal Care and Use Committee of the National Institute of Arthritis, Musculoskeletal and Skin Diseases (NIAMS).

Cell isolation

NK cells were isolated from spleen and liver; ILC1s were isolated from liver; Treg, Il17-GFP$^+$ Th17, ILC2s, CD4$^+$ILC3 and NCR$^+$ILC3 were isolated from siLP; HSC, CLP, NKp, immature NK, and ILC2p were isolated from bone marrow; Nippo-infected ILC2 and Th2 cells were isolated from lung. Cells from bone marrow, liver, and spleen were obtained by mechanical disruption. Cells from lung were isolated after incubating lung fragments with 0.5 mg/ml Liberase TL for 1h followed by purification with 40% Percoll (GE Healthcare) (Meylan et al., 2013). Cells from siLP were isolated after incubating fine-cut intestine in HBSS solution with 0.5 mg/ml DNase I (Roche) and 0.25 mg/ml Liberase TL (Roche) followed by filtering with 100 μm cell strainer and purification with 40% Percoll (Sciumé et al., 2012). Isolated cells were further sorted as described previously. See Extended Experimental Procedures for antibodies used.

In vitro stimulation

All cells were stimulated in RPMI medium with 10% (vol/vol) FCS (Invitrogen), 2 mM glutamine (Invitrogen), 100 IU/ml penicillin (Invitrogen), 0.1 mg/ml streptomycin (Invitrogen), and 20 mM HEPES buffer, pH 7.2–7.5 (Invitrogen), and 2 mM β-mercaptoethanol (Sigma-Aldrich). NK cells were treated with 1000 U/ml IL-2 and 10 ng/ml IL-12 (R&D) for 6 hours; ILC2 cells were treated with 50 ng/ml IL-25 (BioLegend) and 50 ng/ml IL-33 (Biolegend) for 4 hours; NCR$^+$ ILC3 were treated with 50ug/ml of IL-23 (R&D Systems) for 6 hours.

In vivo ILC2 and Th2 induction

Mice of at least 8 weeks of age were infected with 500 infective third-stage *N. brasiliensis* by subcutaneous injection. Cells from lungs were isolated after 10 days of infection. Staining and sorting strategy of Th2 and ILC2 cells were as described in Figure 7.

RNA-seq

RNA-seq was performed as described previously with slight modification (Hirahara et al., 2015). Total RNA was prepared from approximately 30–50,000 cells by using TRIzol

following manufacture's protocol (Life Technologies). Total RNA was subsequently processed to mRNA-seq library using TruSeq SR mRNA sample prep kit (FC-122-1001, Illumina). The libraries were sequenced for 50 cycles (single read) with a HiSeq 2000 or HiSeq 2500 (Illumina). Raw sequencing data were processed with CASAVA 1.8.2 to generate FastQ files.

RNA-seq analysis

Sequence reads were mapped onto the mouse genome build mm9 using TopHat 2.1.0. Gene expression values (FPKM, fragments per kilobase exon per million mapped reads) were calculated with Cufflinks 2.2.1 (Trapnell et al., 2012). The BigWig tracks were generated from Bam files and converted into Bedgraph format using bedtools. These were further reformatted with UCSC tool bedGraphToBigWig. The differential gene expression was determined by DEseq using 2-fold change, p-value < 0.05 as threshold (Anders, S., and Huber, W., 2010). Downstream analyses and heatmaps were performed with R 3.0.1 (R Core Team, 2014) and custom R programs.

See also **Extended Experimental Procedures**.

ATAC-seq

ATAC-seq was performed according to published protocol (Buenrostro et al., 2013) with minor modification. 50,000 cells were pelleted and washed with 50ul 1xPBS, followed by treatment with 50ul lysis buffer (10mM Tris-HCl, pH7.4, 10mM NaCl, 3mM $MgCl_2$, 0.1% of IGEPAL CA-630). After pelleting the nuclei by centrifuging at 500x g for 10 min, the pellets were re-suspended in 40ul transposition reaction with 2ul Tn5 transposase (Illumina Cat# FC-121-1030) to tag and fragmentalize accessible chromatin. The reaction was incubated at 37°C with shaking at 300rpm for 30 min. The fragmentalized DNAs were then purified using Qiagen MinElute Kit and amplified with 10–11 cycles of PCR based on the amplification curve. Once the libraries are purified using Qiagen PCR Cleanup Kit, they were further sequenced for 50 cycles (paired-end reads) on HiSeq2500.

ATAC-seq analysis

ATAC-seq reads from two biological replicates for each sample were mapped to the mouse genome (mm9 assembly) using Bowtie 0.12.8. In all cases, redundant reads were removed using fastquniq (Xu et al., 2012) and customized python scripts were used to calculate fragment length of each pair of uniquely mapped PE reads. The fragment sizes distribute similarly to previously published data (data not shown). Only one mapped read to each unique region of the genome that was less than 175 base pairs was kept and used in peak calling. Regions of open chromatin were identified by MACS (version 1.4.2)(Zhang et al., 2008) using a *P*-value threshold of 1×10^{-5}. Only regions called in both replicates were used in downstream analysis. Peak intensities ('tags' column) were normalized as tags-per-10-million reads (RP10M) in the original library. Downstream analyses and heatmap generation were performed with the Hypergeometric Optimization of Motif EnRichment program (HOMER) v4.8 (Heinz et al., 2010) and R 3.0.1 (R Core Team, 2014).

See also **Extended Experimental Procedures**.

Chromatin immunoprecipitation sequencing (ChIP-seq)

ChIP-seq was performed using ex vivo purified NK cells without or with cytokine stimulation. At least 10 million cells were used for transcription factor ChIP, and 2 million cells for histone mark ChIP. After chemically cross-linking cells, chromatin was fragmented by sonication and immunoprecipitated by anti-H3K27Ac (ab4729, AbCam), anti-p300 (sc585, Santa Cruz) or anti-T-bet (sc21003, Santa Cruz). After recovering purified DNA, 10 ng or more of DNA was used to generate libraries according to the vendor's manual for illumina platform (NEB #E6240S/L, New England BioLabs). Illumina HiSeq2500 (H3K27Ac, T-bet) or Genome Analyzer II (p300) were used for 50 cycle single read sequencing. SICER (K27Ac) or MACS 1.4.2 (T-bet, p300) were used for peak calling using a reference genome mm9.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Drs. Y. Belkaid, K. Zhao, V. Sartorelli, J. Zhu, I. Fraser and A. Poholek for critically reading this manuscript. We also thank G. Gutierrez-Cruz (Biodata Mining Core Facility, NIAMS), J. Simone, J. Lay, and K. Tinsley (Flow Cytometry Section, NIAMS) for their technical support. This study utilized the high-performance computational capabilities of the Biowulf Linux cluster at the NIH. This work was supported by the Intramural Research Programs of NIAMS.

Reference

Artis D, Spits H. The biology of innate lymphoid cells. Nature. 2015; 517:293–301. [PubMed: 25592534]

Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biology. 2010; 11(10):R106. [PubMed: 20979621]

Ashburner, et al. Gene ontology: tool for the unification of biology. Nat Genet. 2000; 25(1):25–29. [PubMed: 10802651]

Balasubramani A, Winstead CJ, Turner H, Janowski KM, Harbour SN, et al. Deletion of a Conserved cis-Element in the Ifng Locus Highlights the Role of Acute Histone Acetylation in Modulating Inducible Gene Transcription. PLoS Genet. 2014; 10:e1003969. [PubMed: 24415943]

Balasubramani A, Mukasa R, Hatton RD, Weaver CT. Regulation of the Ifng locus in the context of T-lineage specification and plasticity. Immunol Rev. 2010; 238:216–232. [PubMed: 20969595]

Barton K, Muthusamy N, Fischer C, Ting CN, Walunas TL, Lanier LL, Leiden JM. The Ets-1 transcription factor is required for the development of natural killer cells in mice. Immunity. 1998; 9:555–563. [PubMed: 9806641]

Bernink JH, Peters CP, Munneke M, Velde te AA, Meijer SL, et al. Human type 1 innate lymphoid cells accumulate in inflamed mucosal tissues. Nature Immunology. 2013; 14:221–229. [PubMed: 23334791]

Bible PW, Kanno Y, Wei L, Brooks SR, O'Shea JJ, Morasso MI, et al. PAPST, a User Friendly and Powerful Java Platform for ChIP-Seq Peak Co-Localization Analysis and Beyond. PLoS ONE. 2015; 10(5):e0127285. [PubMed: 25970601]

Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. BuenrostroNatMeth. Nature Methods. 2013; 10:1213–1218. [PubMed: 24097267]

Cella M, Fuchs A, Vermi W, Facchetti F, Otero K, et al. A human natural killer cell subset provides an innate source of IL-22 for mucosal immunity. Nature. 2009; 457:722–725. [PubMed: 18978771]

Chiossone L, Chaix J, Fuseri N, Roth C, Vivier E, Walzer T. Maturation of mouse NK cells is a 4-stage developmental program. Blood. 2009; 113:5488–5496. [PubMed: 19234143]

Constantinides MG, McDonald BD, Verhoef PA, Bendelac A. A committed precursor to innate lymphoid cells. Nature. 2014; 508:397–401. [PubMed: 24509713]

De Obaldia ME, Bhandoola A. Transcriptional Regulation of Innate and Adaptive Lymphocyte Lineages. Annu. Rev. Immunol. 2015; 33:607–642.

Diefenbach A, Colonna M, Koyasu S. Development, Differentiation, and Diversity of Innate Lymphoid Cells. Immunity. 2014; 41:354–365. [PubMed: 25238093]

Eberl G, Colonna M, Di Santo JP, McKenzie ANJ. Innate lymphoid cells: A new paradigm in immunology. Science. 2015; 348:aaa6566–aaa6566. [PubMed: 25999512]

Ebihara T, Song C, Ryu SH, Plougastel-Douglas B, Yang L, Levanon D, Groner Y, Bern MD, Stappenbeck TS, Colonna M, et al. Runx3 specifies lineage commitment of innate lymphoid cells. Nature Immunology. 2015; 16:1124–1133. [PubMed: 26414766]

Factor DC, Corradin O, Zentner GE, Saiakhova A, Song L, Chenoweth JG, McKay RD, Crawford GE, Scacheri PC, Tesar PJ. Epigenomic comparison reveals activation of “seed” enhancers during transition from naive to primed pluripotency. Stem Cell. 2014; 14:854–863.

Fathman JW, Bhattacharya D, Inlay MA, Seita J, Karsunky H, Weissman IL. Identification of the earliest natural killer cell-committed progenitor in murine bone marrow. Blood. 2011; 118:5439–5447. [PubMed: 21931117]

Finkelman FD, Shea-Donohue T, Morris SC, Gildea L, Strait R, Madden KB, Schopf L, Urban JF. Interleukin-4- and interleukin-13-mediated host protection against intestinal nematode parasites. Immunol Rev. 2004; 201:139–155. [PubMed: 15361238]

Gosselin D, Link VM, Romanoski CE, Fonseca GJ, Eichenfield DZ, Spann NJ, Stender JD, Chun HB, Garner H, Geissmann F, et al. Environment Drives Selectionand Function of Enhancers Controlling Tissue-Specific Macrophage Identities. Cell. 2014; 159:1327–1340. [PubMed: 25480297]

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, et al. Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. Molecular Cell. 2010; 38:576–589. [PubMed: 20513432]

Heinz S, Romanoski CE, Benner C, Glass CK. The selection and function of cell type-specific enhancers. Nature Publishing Group. 2015; 16:144–154.

Hirahara K, Onodera A, Villarino AV, Bonelli M, Sciumé G, Laurence A, Sun H-W, Brooks SR, Vahedi G, Shih H-Y, et al. Asymmetric Action of STAT Transcription Factors Drives Transcriptional Outputs and Cytokine Specificity. Immunity. 2015; 42:877–889. [PubMed: 25992861]

Hirano M, Das S, Guo P, Cooper MD. The Evolution of Adaptive (Elsevier inc.). 2011

Hoyler T, Klose CSN, Souabni A, Turqueti-Neves A, Pfeifer D, et al. The transcription factor GATA-3 controls cell fate and maintenance of type 2 innate lymphoid cells. Immunity. 2012; 37:634–648. [PubMed: 23063333]

Huang Y, Guo L, Qiu J, Chen X, Hu-Li J, Siebenlist U, Williamson PR, Urban JF, Paul WE. IL-25-responsive, lineage-negative KLRG1hi cells are multipotential “inflammatory” type 2 innate lymphoid cells. Nature Immunology. 2014; 16:161–169. [PubMed: 25531830]

Kang J, Malhotra N. Transcription Factor Networks Directing the Development, Function, and Evolution of Innate Lymphoid Effectors. Annu. Rev. Immunol. 2015; 33:505–538. [PubMed: 25650177]

Kearley J, Silver JS, Sanden C, Liu Z, Berlin AA, White N, Mori M, Pham T-H, Ward CK, Criner GJ, et al. Cigarette Smoke Silences Innate Lymphoid Cell Function and Facilitates an Exacerbated Type I Interleukin-33-Dependent Response to Infection. Immunity. 2015; 42:566–579. [PubMed: 25786179]

Kim CC, Lanier LL. ScienceDirectBeyond the transcriptome: completion of act one of the Immunological Genome Project. Current Opinion in Immunology. 2013; 25:593–597. [PubMed: 24168965]

Klose, CSN.; Diefenbach, A. Current Topics in Microbiology and Immunology. Cham: Springer International Publishing; 2014. Transcription Factors Controlling Innate Lymphoid Cell Fate Decisions; p. 215-255.

Klose CSN, Kiss EA, Schwierzeck V, Ebert K, Hoyler T, d'Hargues Y, Göppert N, Croxford AL, Waisman A, Tanriver Y, et al. A T-bet gradient controls the fate and function of CCR6. Nature. 2014; 494:261–265. [PubMed: 23334414]

Lara-Astiaso D, Weiner A, Lorenzo-Vivas E, Zaretsky I, Jaitin DA, et al. Chromatin state dynamics during blood formation. Science. 2014

Lavin Y, Winter D, Blecher-Gonen R, David E, Keren-Shaul H, Merad M, Jung S, Amit I. Tissue-Resident Macrophage Enhancer Landscapes Are Shapedby the Local Microenvironment. Cell. 2014; 159:1312–1326. [PubMed: 25480296]

Lohoff M, Duncan GS, Ferrick D, Mittrücker HW, Bischof S, Prechtl S, Röllinghoff M, Schmitt E, Pahl A, Mak TW. Deficiency in the transcription factor interferon regulatory factor (IRF)-2 leads to severely compromised development of natural killer and T helper type 1 cells. J. Exp. Med. 2000; 192:325–336. [PubMed: 10934221]

McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, et al. GREAT improves functional interpretation of cis-regulatory regions. 2010; 28(5):495–501.

Meylan F, Hawley ET, Barron L, Barlow JL, Penumetcha P, Pelletier M, egrave GS, Richard AC, Hayes ET, Gomez-Rodriguez J, et al. The TNF-family cytokine TL1A promotes allergic immunopathology through group 2 innate lymphoid cells. 2013; 7:958–968.

R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2014.

Rankin LC, Groom JR, Chopin M, Herold MJ, Walker JA, et al. The transcription factor T-bet is essential for the development of NKp46+ innate lymphocytes via the Notch pathway. Nature Immunology. 2013; 14:389–395. [PubMed: 23455676]

Robinette ML, Fuchs A, Cortez VS, Lee JS, Wang Y, Durum SK, Gilfillan S, Colonna M, Shaw L, Yu B, et al. Transcriptional programs define molecular characteristics of innate lymphoid cell classes and subsets. Nature Immunology. 2015; 16:306–317. [PubMed: 25621825]

Rosmaraki EE, Douagi I, Roth C, Colucci F, Cumano A, Di Santo JP. Identification of committed NK cell progenitors in adult murine bone marrow. Eur. J. Immunol. 2001; 31:1900–1909.

Sciumé G, Hirahara K, Takahashi H, Laurence A, Villarino AV, et al. Distinct requirements for T-bet in gut innate lymphoid cells. Journal of Experimental Medicine. 2012; 209:2331–2338. [PubMed: 23209316]

Serafini N, Vosshenrich CAJ, Di Santo JP. Transcriptional regulation of innatelymphoid cell fate. Nat Rev Immunol. 2015; 15:415–428. [PubMed: 26065585]

Shay T, Kang J. Immunological Genome Project and systems immunology. Trends in Immunology. 2013; 34:602–609. [PubMed: 23631936]

Shih H-Y, Sciumé G, Poholek AC, Vahedi G, Hirahara K, Villarino AV, Bonelli M, Bosselut R, Kanno Y, Muljo SA, et al. Transcriptional and epigenetic networks of helper T and innate lymphoid cells. Immunol Rev. 2014; 261:23–49. [PubMed: 25123275]

Sonnenberg GF, Artis D. Innate lymphoid cells in the initiation, regulation and resolution of inflammation. Nat. Med. 2015; 21:698–708. [PubMed: 26121198]

Spits H, Artis D, Colonna M, Diefenbach A, Di Santo JP, et al. Innate lymphoid cells--a proposal for uniform nomenclature. Nat Rev Immunol. 2013; 13:145–149. [PubMed: 23348417]

Stergachis AB, Neph S, Sandstrom R, Haugen E, Reynolds AP, Zhang M, Byron R, Canfield T, Stelhing-Sun S, Lee K, et al. Conservation of trans-acting circuitry during mammalian regulatory evolution. Nature. 2015; 515:365–370. [PubMed: 25409825]

Stetson DB, Mohrs M, Reinhardt RL, Baron JL, Wang ZE, Gapin L, Kronenberg M, Locksley RM. Constitutive Cytokine mRNAs Mark Natural Killer (NK) and NK T Cells Poised for Rapid Effector Function. Journal of Experimental Medicine. 2003; 198:1069–1076. [PubMed: 14530376]

Tindemans I, Serafini N, Di Santo JP, Hendriks RW. GATA-3 Function in Innate and Adaptive Immunity. Immunity. 2014; 41:191–206. [PubMed: 25148023]

Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature Protocols. 2012; 7:562–578. [PubMed: 22383036]

Vahedi G, Takahashi H, Nakayamada S, Sun H-W, Sartorelli V, et al. STATs Shape the Active Enhancer Landscape of T Cell Populations. Cell. 2012; 151:981–993. [PubMed: 23178119]

Verykokakis M, Zook EC, Kee BL. ID'ing innate and innate-like lymphoid cells. Immunol Rev. 2014; 261:177–197. [PubMed: 25123285]

Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, Holt A, et al. ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature. 2009; 457(7231):854–858. [PubMed: 19212405]

Wilson CB, Rowell E, Sekimata M. Epigenetic control of T-helper-cell differentiation. Nat Rev Immunol. 2009; 9:91–105. [PubMed: 19151746]

Xu H, Luo X, Qian J, Pang X, Song J, Qian G, Chen J, Chen S. FastUniq: A Fast De Novo Duplicates Removal Tool for Paired Short Reads. PLoS ONE. 2012; 7:e52249. [PubMed: 23284954]

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nussbaum C, Myers RM, Brown M, Li W, et al. Model-based Analysis of ChIP-Seq (MACS). Genome Biol. 2008; 9:R137. [PubMed: 18798982]

Zhong C, Cui K, Wilhelm C, Hu G, Mao K, Belkaid Y, Zhao K, Zhu J. Group 3 innate lymphoid cells continuously require the transcription factor GATA-3 after commitment. Nature Immunology. 2015

Highlights

- Prototypical ILC subsets show distinctive regulomes
- Regulatory elements of ILC effector genes are poised prior to activation
- Regulomes of ILC subsets diverge early at precursor stages
- Regulomes of innate and adaptive cells converge upon infection

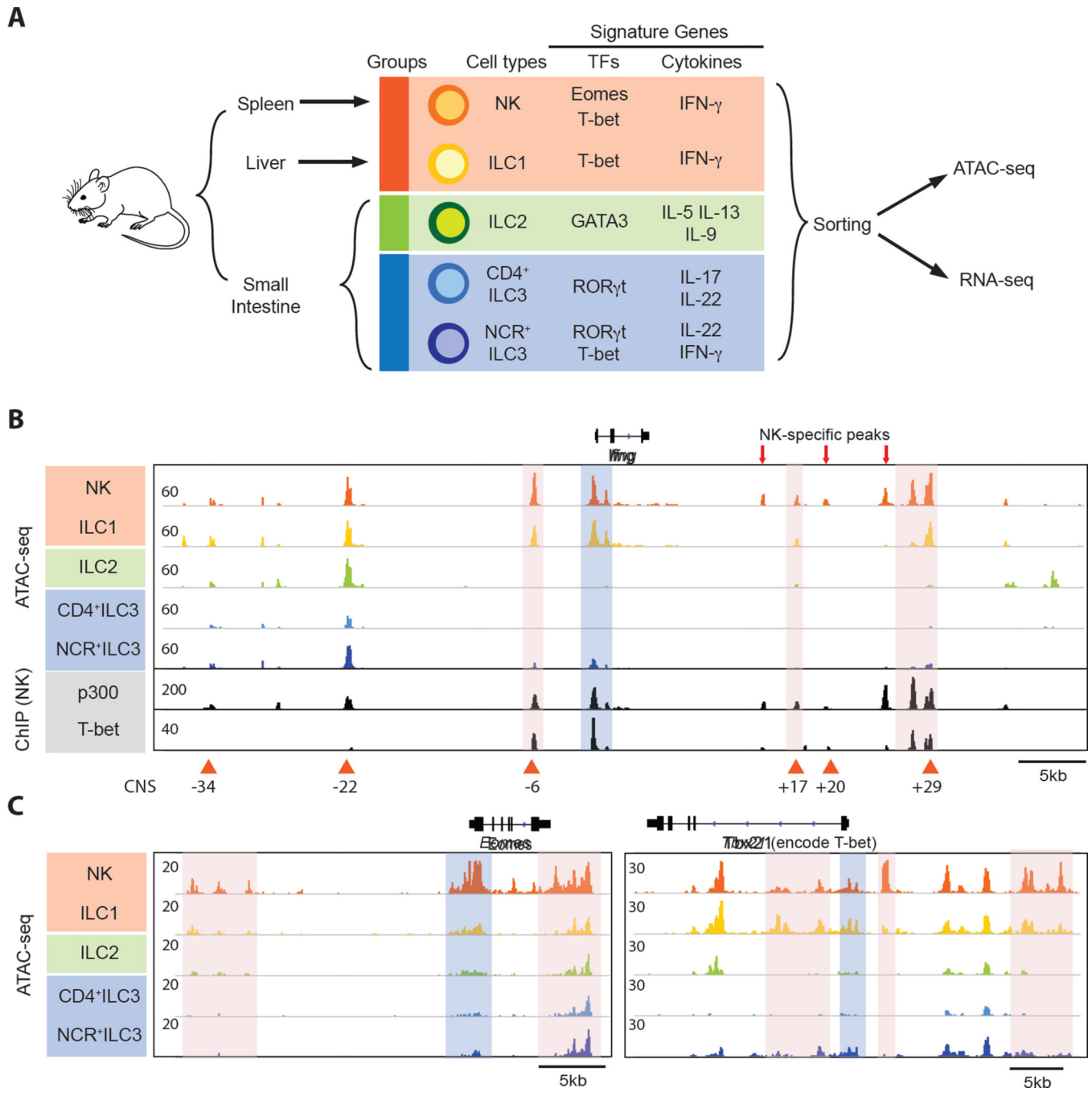


Figure 1. Identification of Regulatory Elements of Innate Lymphoid Cells

(A) Schematic illustration of experimental design. Five prototypical ILCs from various organs were isolated by flow cytometry for ATAC-seq and RNA-seq analysis. Listed in the table are ILC signature genes. All experiments were done in duplicates.

(B) – (C) Representative examples of normalized ATAC-seq signal profiles in ILCs across type I signature genes including **(B)** *Ifng*, **(C)** *Eomes* and *Tbx21*. **(B)** p300 and T-bet ChIP-seq were acquired from NK cells. Lineage-signature ATAC peaks at promoter and non-promoter regions are highlighted in blue and red, respectively. Red triangles denote known

regulatory elements (Balasubramani et al., 2010; Wilson et al., 2009). Red arrows denote NK cell-specific REs.

See also Figure S1 and **Experimental Procedures**.

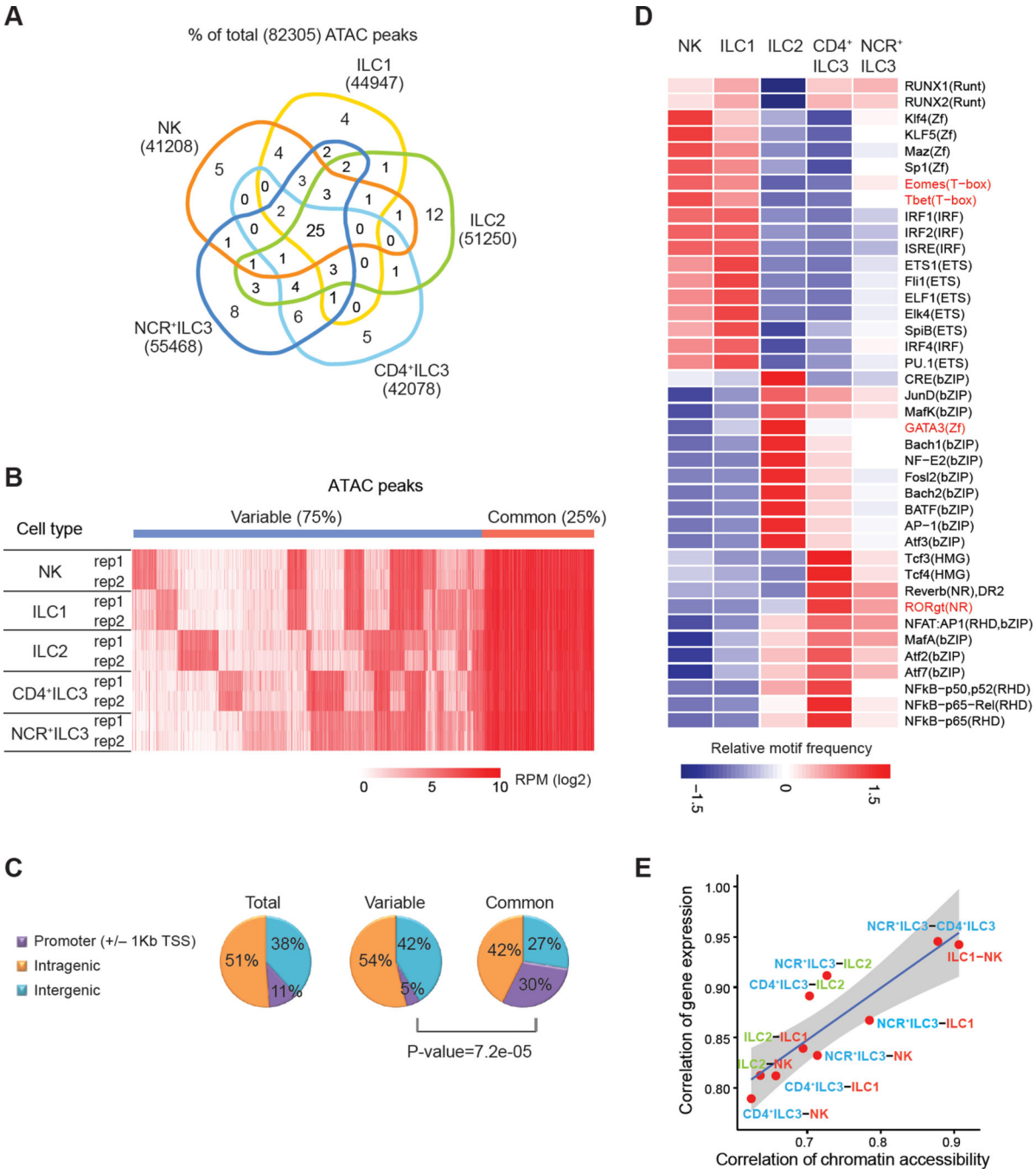


Figure 2. Genome-wide chromatin landscapes define distinct ILC subsets

(A) and (B) Comparison of global ATAC peaks in prototypical ILCs. **(A)** Venn diagram demonstrates percentages of ATAC peaks that are commonly or differentially present in ILCs. The total number of ATAC peaks in each ILC subset is shown next to the annotation. **(B)** Heatmap showing signal intensity (reads per million mapped reads by log2) of each ATAC peak.

(C) Pie charts illustrate the distribution of ATAC peaks across the genome (promoter, +/− 1kb of transcription start sites, intragenic or intergenic regions). P-value is determined by using Fisher's exact test.

(D) Heatmap showing relative enrichment of TF motifs among ILC signature REs. LDTFs were highlighted in red.

(E) Scatter plot showing the relationships for Pearson correlations of transcriptomes and regulomes between pairs of ILC subsets. Log2 transformed tag counts averaged from replicates were used for Pearson correlation analysis with threshold of 1 FPKM and 1 RPM for RNA-seq and ATAC-seq datasets, respectively. Blue line denotes linear regression line derived from all data points in the plot. Grey area denotes 95% confidence limits of linear regression.

See also Figure S2, **Experimental Procedures and** Supplemental Table S1.

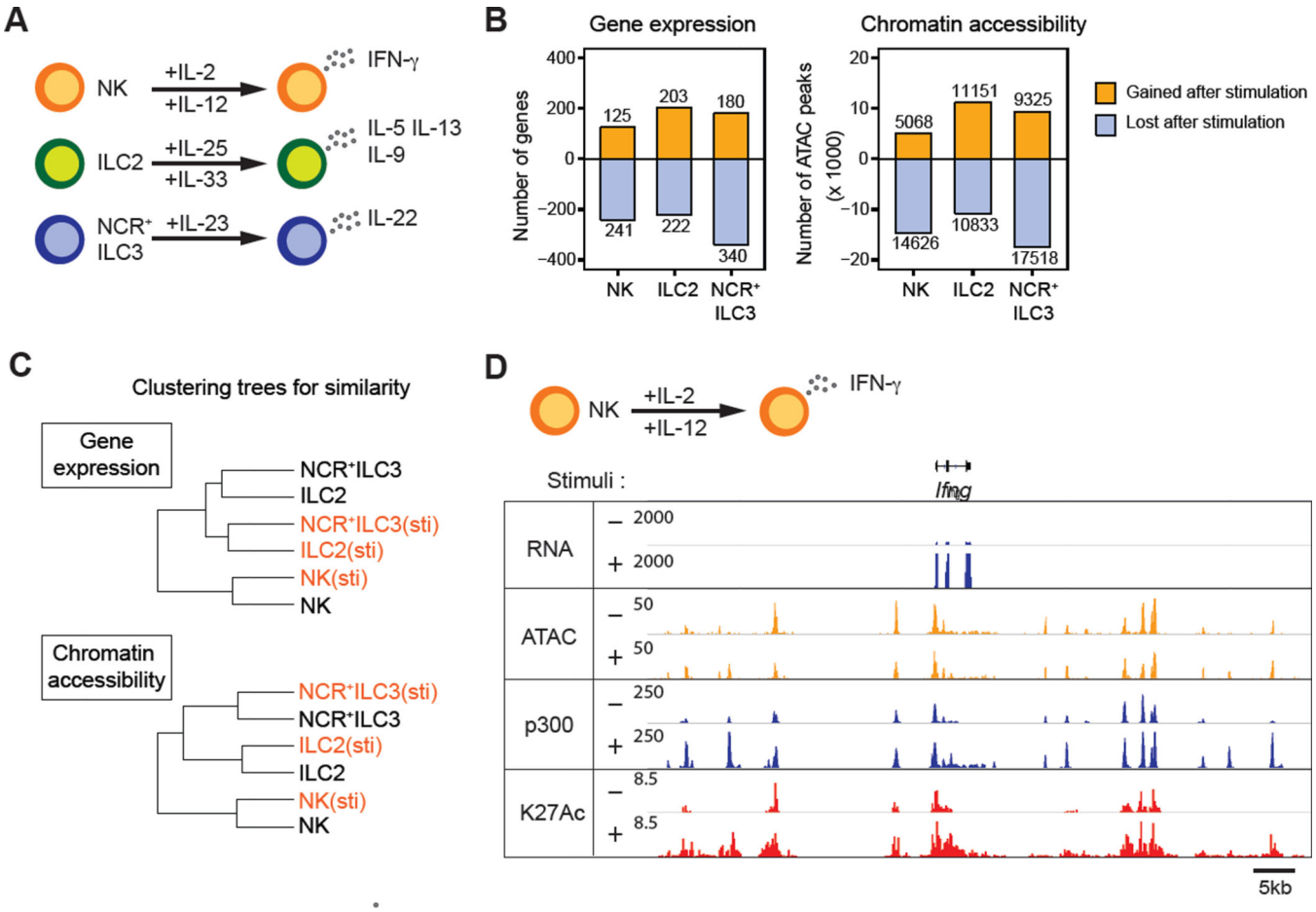


Figure 3. Cytokine Loci are Primed Prior to Activation

(A) Experimental designs for ILC stimulation. NK cells were treated with IL-2 (1000U/ml) and IL-12 (10ng/ml) for six hours, ILC2 cells were treated with IL-25 (50 ng/ml) and IL-33 (50 ng/ml) for four hours, and NCR^{+}ILC3 cells were treated with IL-23 (50ng/ml) for four hours.

(B) Changes in gene expression or chromatin accessibility upon ILC stimulation. Left panel, numbers of genes changed expression over two folds (p-value< 0.05) after stimulation. Right panel, number of peaks gained (orange) or lost (blue) after stimulation.

(C) Dendrogram showing hierarchical clustering analysis of ILC gene expression and chromatin accessibility before (marked in black) and after (marked in red) stimulation.

(D) Genome track view of the *Ifng* locus showing RNA-seq, ATAC-seq, H3K27 acetylation and p300 binding for stimulated and non-stimulated NK cells.

See also Figure S3, **Experimental Procedures and** Supplemental Table S1.

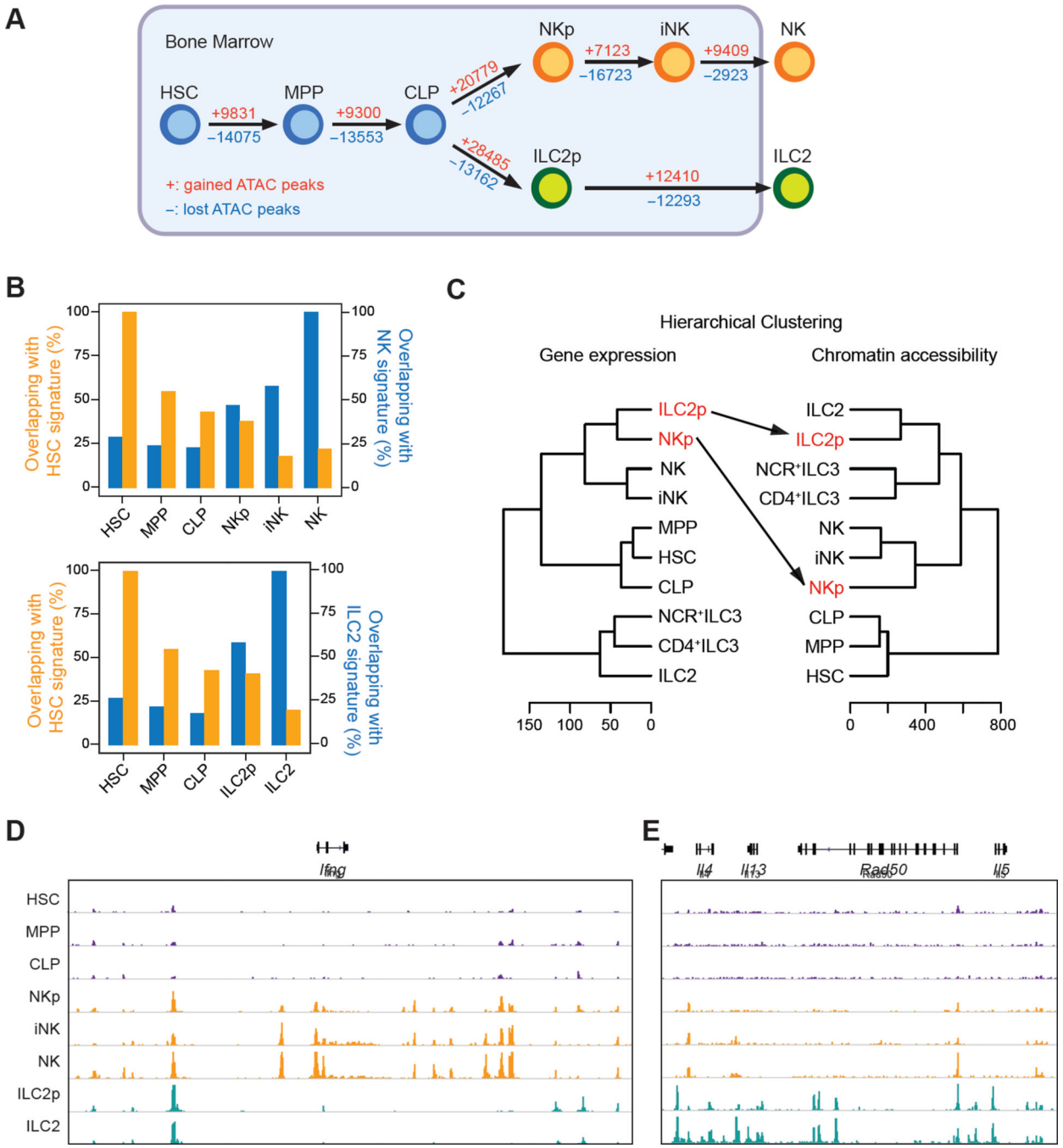


Figure 4. Distinctive ILC Enhancer Landscapes Diverge in Development

(A) Schematic diagram of the ILC developmental stages evaluated by ATAC-seq and RNA-seq. The numbers of ATAC peaks gained or lost during transition determined by PAPST program (Bible et al., 2015) are shown in red and blue, respectively, along the arrows.

(B) Bar plot illustrates progressive loss of HSC signature (in yellow) and reciprocal gain of mature lineage signature (in blue) for NK (left) or ILC2 (right) during ILC development.

(C) Dendrogram showing hierarchical clustering analysis of ILC gene expression (left) and chromatin accessibility (right) of developing ILCs. Log2 transformed tag counts averaged

from replicates were used for calculation of Euclidean distances between two cell types with threshold of 1 FPKM and 1 RPM for RNA-seq and ATAC-seq datasets, respectively, and were further clustered by hclust program in R using the ward method.

(D) – (E) Genome track view of the *Ifng* and *Th2* loci showing early establishment of divergent chromatin landscapes during ILC development.

See also Figure S4, **Experimental Procedures and** Supplemental Table S1.

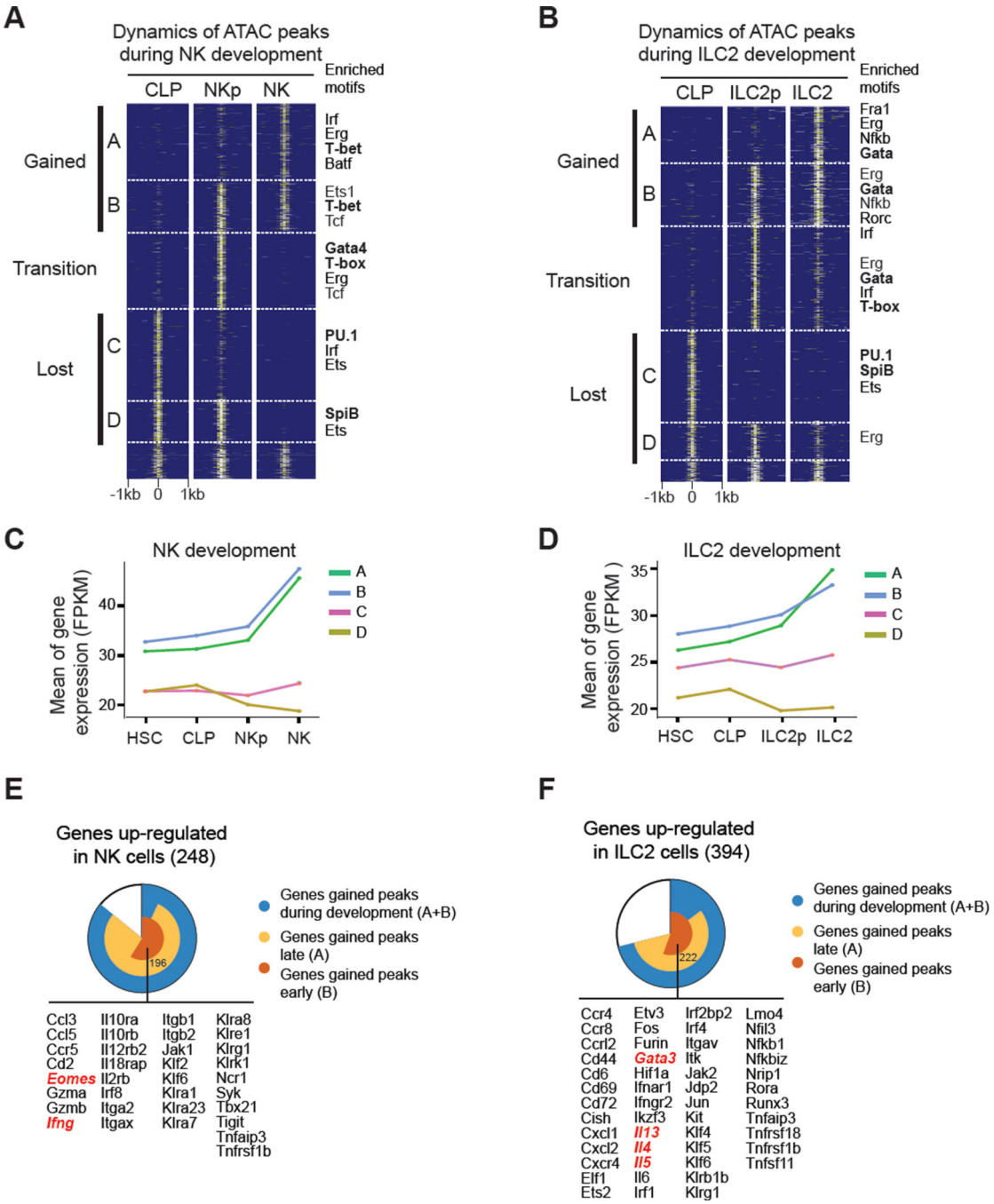


Figure 5. Regulatory Elements of ILC Signature Genes are Defined Prior to Maturation

(A) – (B) Dynamics of regulomes during NK **(A)** and ILC2 **(B)** development. ATAC-seq Peaks were classified into four categories based on their presence at different developmental stages. A: present in mature ILCs only; B: present in both ILC precursors and mature ILCs; C: present in CLP only; D: present in both CLP and ILC precursors. Representative motifs enriched in each group are listed on the right.

(C) – (D) Mean expression level (FPKM) of genes in proximity to REs categorized in **(A)** and **(B)** during ILC development was plotted.

(E) – (F) REs in proximity to genes upregulated during NK **(E)** and ILC2 (**F**) development were evaluated for their timing of acquisition (early vs late). Representative genes that acquire REs early (group 2) were listed. See also Figure S5, **Experimental Procedures and** Supplemental Table S1.

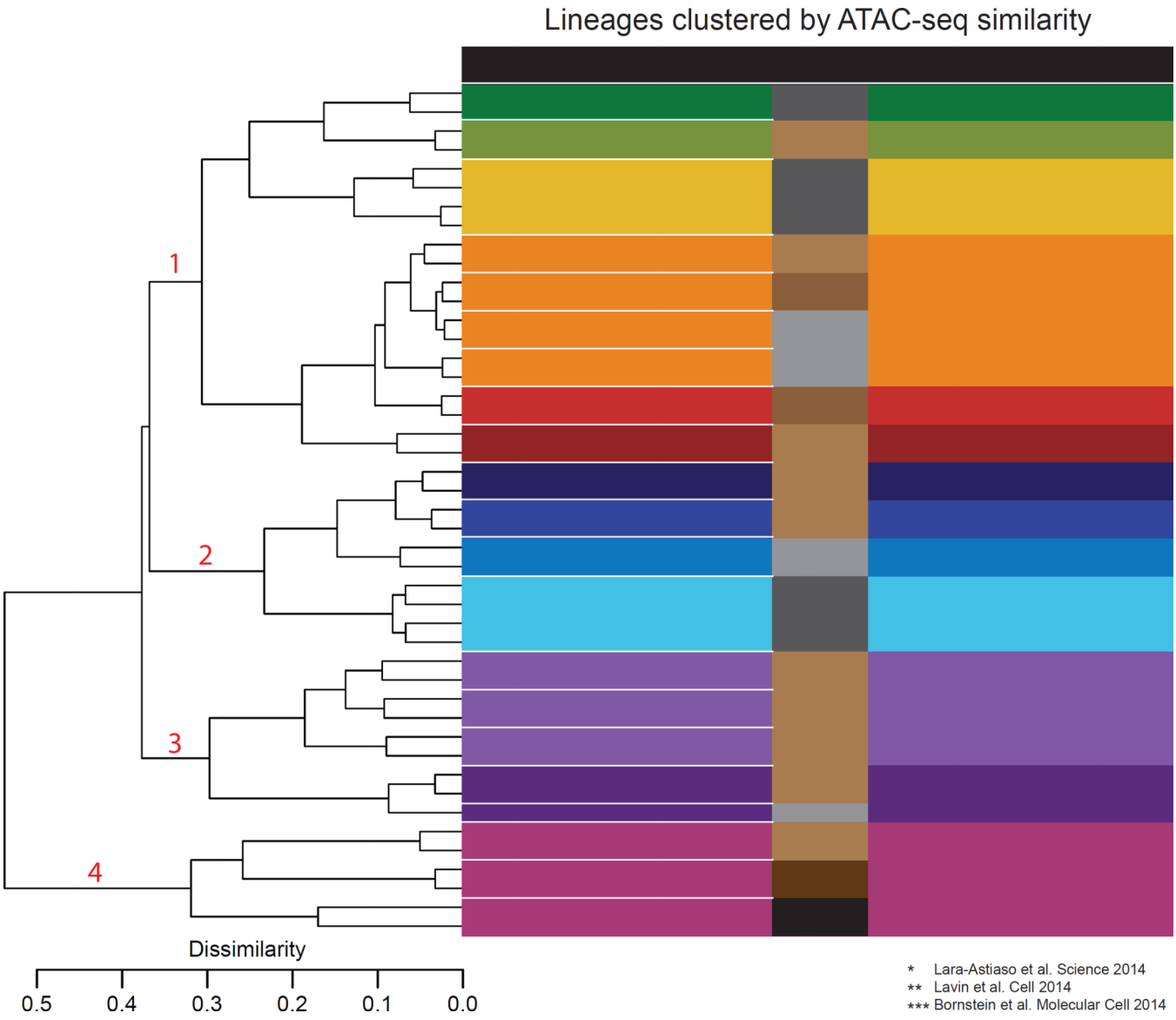


Figure 6. Relationships between innate and adaptive cell regulomes

Dendrogram showing unbiased hierarchical clustering analysis of lineage relationships based on 37 in-house and 8 previous published ATAC-seq data. Categories of cell types and tissue location from which the cells were harvested are indicated on the right and colored accordingly. The distance on the scale implies the dissimilarity (1-Pearson correlation coefficient) between two individual subjects. Log2 transformed tag counts were used for calculation of dissimilarity between two cell types with threshold of 1 RPM. See also **Extended Experimental Procedures**

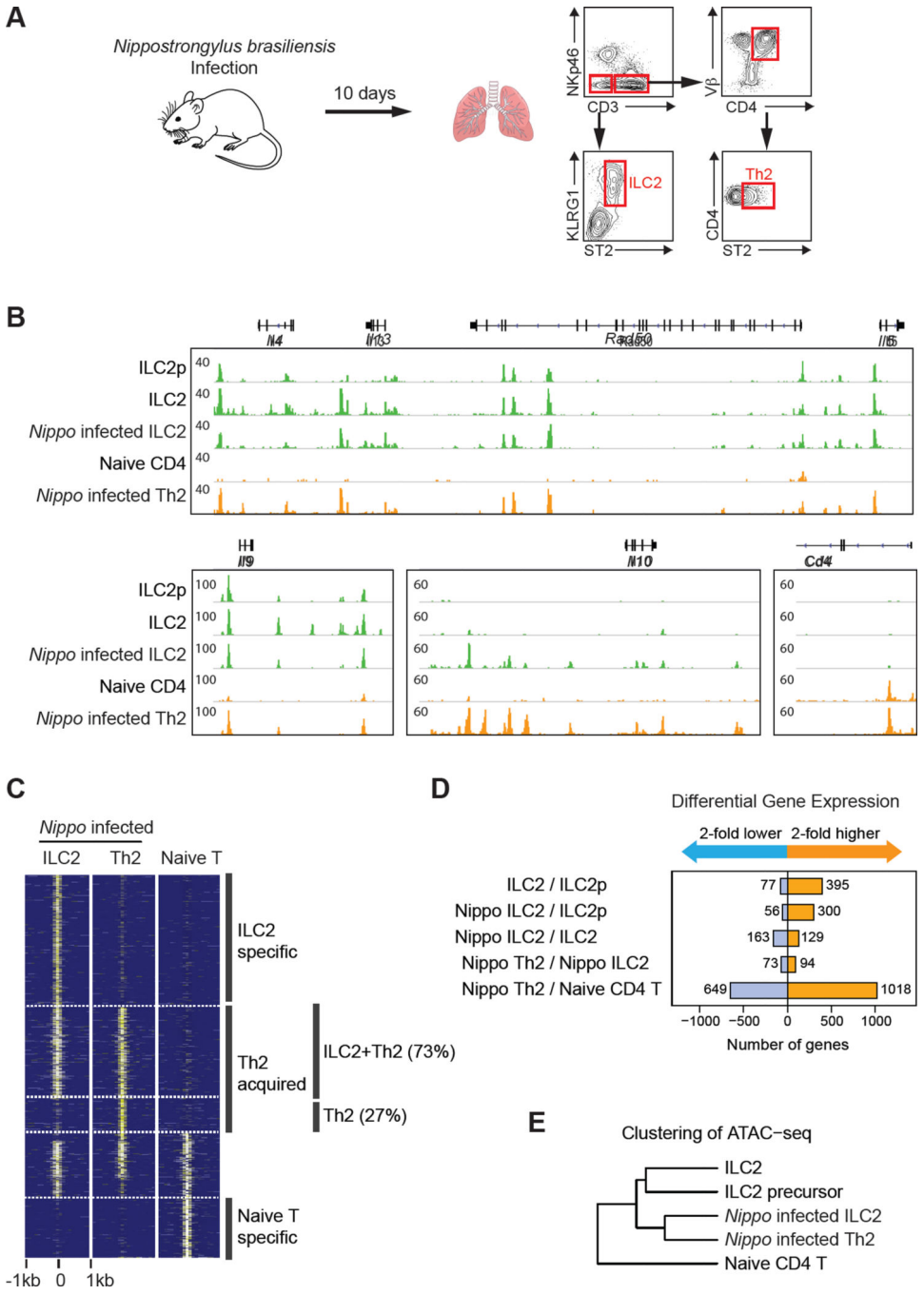


Figure 7. Similarity of ILC2 and Th2 regulomes upon infection

(A) Schematic illustration of experimental design. Lung cells from Foxp3-GFP mice infected with *N. brasiliensis* were sorted by flow cytometry. In the GFP-negative fraction, Th2 cells were sorted as $CD3\varepsilon^{+}V\beta^{+}CD4^{+}ST2^{+}$ cells, while ILC2 as $CD3\varepsilon^{-}NKp46^{-}KLRG1^{+}ST2^{+}$ cells. Purity of sorted cells ranges from 95 to 99% post sort.

(B) Representative examples of ATAC-seq signals in type 2 innate and adaptive cells at loci including Th2 cytokines, *Il9* and *Il10, Cd4* is a lineage marker that distinguishes ILC2 and Th2 cells.

(C) Comparison of ATAC-seq signals at signature REs in infected ILC2, Th2 and Naïve $CD4^+$ T cells. A substantial portion of Th2 ATAC-seq peaks acquired upon infection (73%) was shared with infected ILC2.
See also Supplemental Table S1 for accessible regions.
(D) Pairwise comparison of differential gene expression among type 2 innate and adaptive cells.
(E) Dendrogram showing hierarchical clustering of type 2 innate and adaptive cell regulomes to evaluate their similarities. Log2 transformed tag counts averaged from replicates were used for calculation of Euclidean distances between two cell types with threshold of 1 FPKM and 1 RPM for RNA-seq and ATAC-seq datasets, respectively, and were further clustered by hclust program in R using the ward method.
See also **Extended Experimental Procedures**