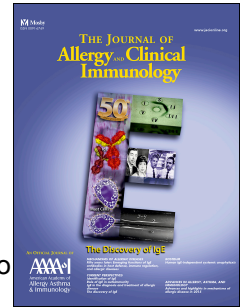


Accepted Manuscript

Epigenome analysis links gene regulatory elements in group 2 innate lymphocytes to asthma susceptibility

Ralph Stadhouders, PhD, Bobby W.S. Li, MSc, Marjolein J.W. de Bruijn, BSc, Antonio Gomez, PhD, Tata Nageswara Rao, PhD, Hans Jörg Fehling, PhD, Wilfred F.J. van IJcken, PhD, Ai Ing Lim, PhD, James P. Di Santo, MD, PhD, Thomas Graf, PhD, Rudi W. Hendriks, PhD



PII: S0091-6749(18)30295-1

DOI: [10.1016/j.jaci.2017.12.1006](https://doi.org/10.1016/j.jaci.2017.12.1006)

Reference: YMAI 13311

To appear in: *Journal of Allergy and Clinical Immunology*

Received Date: 8 June 2017

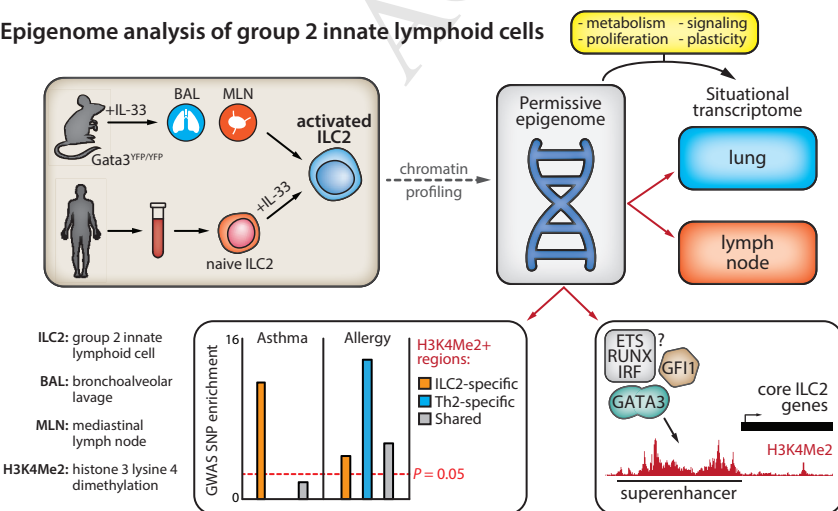
Revised Date: 12 December 2017

Accepted Date: 15 December 2017

Please cite this article as: Stadhouders R, Li BWS, de Bruijn MJW, Gomez A, Rao TN, Fehling HJ, van IJcken WFJ, Lim AI, Di Santo JP, Graf T, Hendriks RW, Epigenome analysis links gene regulatory elements in group 2 innate lymphocytes to asthma susceptibility, *Journal of Allergy and Clinical Immunology* (2018), doi: 10.1016/j.jaci.2017.12.1006.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Epigenome analysis of group 2 innate lymphoid cells



Epigenome analysis links gene regulatory elements in group 2 innate lymphocytes to asthma susceptibility

Authors

Ralph Stadhouders, PhD,^{1,2,3,*} Bobby W.S. Li, MSc,¹ Marjolein J.W. de Bruijn, BSc,¹ Antonio Gomez, PhD,^{2,3} Tata Nageswara Rao, PhD,^{4,†} Hans Jörg Fehling, PhD,⁴ Wilfred F.J. van IJcken, PhD,⁵ Ai Ing Lim, PhD,^{6,7} James P. Di Santo, PhD,^{6,7} Thomas Graf, PhD,^{2,3} and Rudi W. Hendriks, PhD^{1,*}

Affiliations

¹Department of Pulmonary Medicine, Erasmus MC Rotterdam, the Netherlands; ²Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology (BIST) and ³Universitat Pompeu Fabra, Barcelona, Spain; ⁴Institute of Immunology, University Clinics, Ulm, Germany; ⁵Center for Biomimics, Erasmus MC Rotterdam, the Netherlands; ⁶Innate Immunity Unit, Institut Pasteur, Paris, France and ⁷Institut National de la Santé et de la Recherche Médicale (INSERM) U1223, Paris, France

†Current address: Department of Biomedicine, Experimental Hematology, University Hospital Basel, Switzerland

*Corresponding authors

Contact information

Ralph Stadhouders, PhD (r.stadhouders@erasmusmc.nl)

Rudi W. Hendriks, PhD (r.hendriks@erasmusmc.nl)

Erasmus MC | Department of Pulmonary Medicine | Room Ee2251a
3000 CA Rotterdam | the Netherlands | Phone: +31-10-7043700

Supplementary Material:

- Methods
- Supplementary Figures 1-10
- Supplementary Tables 1-5

Methods

Mice

Wildtype C57BL/6 mice were purchased from Harlan. *Gata3*^{YFP/YFP} 'GATIR' mice, carrying a Venus YFP sequence inserted between the last *Gata3* exon and the 3' UTR (T.N.R., H.J.F., et al., manuscript in preparation), were bred on a C57BL/6 background. We have extensively validated GATA3 levels and immunological parameters in untreated GATIR mice or animals treated with IL-33 or house dust mite: under all conditions GATIR mice were identical to wildtype mice (**Supplementary Fig.1** and ref.1). Mice were housed at the Erasmus MC Animal Centre under specific pathogen-free conditions and were 8-10 weeks old at time of analysis. All experiments were approved by the Erasmus MC Animal Ethics Committee.

Isolation of in vivo activated murine ILC2s

GATIR mice were exposed to 0.5 µg recombinant IL-33 (BioLegend) three times intranasally with one day of rest in between each challenge. Mice were sacrificed one day after the final administration. BAL fluid was obtained by flushing the lungs with 3 x 1 mL PBS containing 0.5 mM EDTA. Single-cell suspensions were prepared from mediastinal lymph nodes (MLN) in PBS containing 0.5% bovine serum albumin (BSA) and 5 mM EDTA, using standard procedures².

For flowcytometric analysis and FACS sorting cells were stained with fluorescently labelled monoclonal antibodies (listed in **Supplementary Table 4**). Lineage-negative cells were defined as cells not expressing CD3, CD4, CD5, CD8, CD11b, CD11c, CD19, CD20, B220, NK1.1, FcεRIα, TER-119 and Gr-1. ILC2s were sorted using a FACSAria flow cytometer equipped with three lasers and FACSDiva software (Beckton Dickinson). Data analysis was performed with FlowJo software (Tree Star Inc.).

Isolation and culture of human ILC2s and Th2 cells

Blood samples (buffy coats from 500 ml peripheral blood) from healthy donors for ILC2 isolation were obtained from Etablissement Français du Sang (Paris, France) in an agreement signed with the Institut Pasteur (Paris, France). FACS strategies for isolation of naïve ILC2s (viable CD45⁺Lin⁻CD7⁺CD127⁺CRTh2⁺ cells) from peripheral blood and subsequent short-term culture (7 days) of ILC2s in the presence of IL-2, IL-7, IL-25 and IL-33 have been described^{3, 4}. Peripheral blood samples from healthy donors for T helper subset isolation were obtained after informed consent

in procedures approved by the Medical Ethical committee of the Erasmus MC Rotterdam. Memory Th2 cells were FACS-sorted from peripheral blood mononuclear cell fractions, using antibodies specific for CD3, CD4, CCR6, CXCR3 and CCR4 and CD45RA⁵. Antibodies used are listed in **Supplementary Table 4**.

RNA isolation and RNA sequencing (RNA-Seq)

RNA was extracted from two independent ILC2 pools (obtained from 5-6 GATIR mice) using the RNeasy Micro kit (Qiagen) according to the manufacturer's instructions. Library preparation was performed using the Smart-seq2 methodology⁶ and sequenced according to the Illumina TruSeq Rapid v2 protocol on an Illumina HiSeq2500 (single read mode, 51 bp read length). Reads were aligned to the mouse genome (mm10 build) using HISAT2⁷. Sample scaling and statistical analysis were performed using the R package DESeq2⁸; genes with >1 log2 fold change and $P < 0.05$ (Wald test) were considered differentially expressed. Standard RPKM values were used as an absolute measure of gene expression. Genes with average RPKM < 2 were considered not expressed.

Confocal microscopy

Lungs from GATIR mice were inflated with OCT embedding medium containing 2% PFA (Thermo Fisher Scientific, USA) and snap frozen in liquid nitrogen to preserve morphology and YFP fluorescence. Lymph nodes were fixed in 2% PFA, placed in PBS containing 30% sucrose overnight, embedded in OCT and stored at -80°C . 7 μm thick cryosections were cut at -20°C using a cryostat (Thermo Fisher Scientific) and stained at room temperature with primary antibodies for 1 hour and secondary antibodies (**Supplementary Table 4**) for 30 minutes⁹. Slides were counterstained with DAPI (Invitrogen, USA) for 5 minutes, sealed with Vectashield (Vector Laboratories, USA) and examined with an LSM 510-META confocal microscope equipped with a 405 nm, 488 nm, 543 nm and 633 nm laser (Zeiss, Germany). Images were processed and analyzed in Fiji, an open source scientific image processing application based on ImageJ (<http://fiji.sc/>).

Chromatin immunoprecipitation and sequencing using ChIPmentation

Epigenome analysis of murine ILC2s was performed on two independent ILC2 pools (obtained from 5-6 GATIR mice). For human peripheral blood ILC2s and Th2 cells, samples from 3-4 healthy donors were analyzed. After FACS isolation (or short-term culture for human ILC2s) cells were

immediately crosslinked in PBS containing 1% formaldehyde (Sigma) for 10 min. at room temperature for ChIPmentation analysis¹⁰. Crosslinking was quenched by adding glycine (0.125M final concentration) followed by 5 min. incubation at room temperature. Cells were placed on ice, washed with PBS and snap-frozen for storage at -80°C. Pellets (20-50k cells) were processed in parallel to minimize technical variation. Cells were resuspended in 100µl sonication buffer (1% SDS, 10mM EDTA, 50mM Tris-HCl pH8 and 1x EDTA-free complete protease inhibitors; Roche) and transferred to a 0.65ml Bioruptor sonication tube (Diagenode). After a 15 min. incubation on ice, cells were sonicated for 30 cycles (30 sec ON - 30 sec OFF) using a Bioruptor Pico sonicator (Diagenode) to shear chromatin down to ± 250 bp fragments. Chromatin was equilibrated by adding 900µl 10x ChIP dilution buffer (0.01% SDS, 1.1% Triton X-100, 1.2mM EDTA, 16.7mM Tris-HCl pH8, 167mM NaCl) and incubated overnight at 4°C with 1µl of H3K4Me2-specific antibody (ab32356, Abcam; previously validated as highly sensitive and specific for H3K4Me2¹¹) or normal rabbit IgG as a negative control (sc-2027, Santa Cruz). In addition, 20µl of protein A Dynabeads (Thermo Fisher Scientific) per IP were blocked in PBS containing 0.1% BSA (Sigma) by incubation overnight at 4°C. The next day, beads were resuspended in the original volume with ChIP dilution buffer and added to the chromatin extracts. After 2 hours of incubation at 4°C, beads were collected and washed with Low Salt buffer (0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-HCl pH8, 150mM NaCl), High Salt buffer (0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-HCl pH8, 500mM NaCl) and LiCl buffer (10mM Tris-HCl pH8, 1mM EDTA, 250mM LiCl, 0.5% NP-40, 0.5% deoxycholic acid). Chromatin-antibody immobilized on magnetic beads were then subjected to tagmentation as recently described¹². Eluted DNA was purified using MinElute spin columns (Qiagen) and amplified for 8-12 cycles using Nextera PCR primers. Libraries were purified using dual (0.5x-2.0x) SPRI Ampure XP beads (Beckman Coulter), pooled (up to 10 per sequencing run) and sequenced on a NextSeq500 (Illumina) running a single-read 75bp protocol.

Initial ChIPmentation data processing, quality assessment and visualization

Reads were demultiplexed using BaseSpace (Illumina) and aligned to the mouse genome (mm10 build) using Bowtie2¹³ with standard settings, removing reads that could not be uniquely mapped. Indexed and sorted bam files were parsed to HOMER¹⁴ for further analysis. Data quality and signal-to-noise ratios for all samples were in line with published guidelines¹⁵: fraction of reads in peaks (FRiP) was on average 49.4% with <7% duplicate reads. Over 95% of all expressed genes (>2 RPKM) showed above-background H3K4Me2 promoter enrichment levels, demonstrating

robust sensitivity. GATA3 and GFI1 ILC2 ChIP-Seq data were downloaded from the Gene Expression Omnibus (GEO, series GSE71198¹⁶ for GATA3 and series GSE50806¹⁷ for GFI1). Tag directories were generated for each sample with removal of duplicate reads (*-tbp 1* option). BedGraph files displaying normalized counts (reads per million) were generated for direct visualization in the UCSC Genome Browser (<https://genome.ucsc.edu/>) using the *makeUCSCfile* HOMER script.

Peak calling, annotation and downstream analyses

H3K4Me2 enriched regions were identified using HOMER *findPeaks* with *-region -size 1000 -minDist 2500* options, utilizing signal obtained from a negative control IgG experiment as background. GATA3 or GFI1 peaks were called using the *-style factor* option. Annotation of peaks as defined by their overlap with known genomic features was performed using the *annotatePeaks* script. Overlapping and non-overlapping regions/peaks between two samples were identified using the *intersect* function of BEDTools¹⁸ or the HOMER *mergePeaks* script (*-d given* option) requiring a minimal overlap of 1bp. Regions/peaks of similar and differential enrichment (fold change >2) were determined using HOMER *getDifferentialPeaks* (*-F 2* option). Histograms of ChIP signal were generated using the *annotatePeaks* script (*size 5000 -hist 10 -d* options) and visualized as heatmaps with Java TreeView¹⁹. Venn diagrams were drawn using the VennDiagram package²⁰ in RStudio (v0.99.484).

Peak-gene assignment, gene ontology and molecular signature analysis

Regions/peaks were assigned to putative target genes using the 'single nearest gene' method of GREAT²¹. GREAT was subsequently used to calculate enrichments of these genes for known Gene Ontology (GO) terms and molecular signatures (from the MSigDB²²) using the whole genome as background.

Transcription factor binding motif discovery

The HOMER *findMotifsGenome* script (*-size 1000 -mask -len 6,8,10,12 -S 20* options) was used to search for known transcription factor binding motifs (n=255) in H3K4Me2 enriched regions. Motifs with $P < 0.001$ were considered enriched and visualized as a heatmap.

Superenhancer identification

The HOMER *findPeaks* script (-style super -L 1 options) was used to identify superenhancers in ILC2s using the same criteria as previously defined by Whyte et al.²³ We utilized signals obtained from a negative control mock (IgG) experiment as background for superenhancer identification. To verify that our superenhancers localized to regions of open chromatin we downloaded previously published²⁴ ILC2 ATAC-Seq data (GSE77695). BigWig files were used to generate meta-profiles of ATAC-Seq signal across the ILC2 superenhancers as described above using HOMER software. Overlapping and non-overlapping superenhancers were identified using the *intersect* function of BEDTools as described above.

Integration of epigenomic data with asthma-associated human genetic variation

SNPs associated with human asthma phenotypes in genome-wide association studies ($P < 1.0 \times 10^{-5}$) were downloaded from the NHGRI-EBI GWAS Catalog²⁵. SNAP²⁶ was used to retrieve SNPs in high pairwise linkage disequilibrium (CEU population panel, $r^2 > 0.8$, distance limit = 500) with the GWAS Asthma SNPs, yielding a total of 3405 unique SNPs. Putative causal SNPs for various human disorders were previously published²⁷. Differentially methylated regions associated with asthma were obtained from earlier epi-GWAS studies^{28, 29}. We used GREAT to generate a list of genes closest to these SNPs ('two nearest gene' association rule, $n=664$ genes) and added ALOX5³⁰. Murine orthologues of human genes were retrieved using Ensembl BioMart. GREs containing Asthma-associated SNPs were identified using the *intersect* function of BEDTools. Statistical significance of overlap between SNP and GRE positions was assessed using the *fisher* function of BEDTools (Fisher's exact test). Random shuffling of the 3405 SNPs across the genome ($n=10$) consistently yielded non-significant enrichments using this test ($P > 0.45$). SNP enrichment ratios were also obtained using BEDTools's *fisher* function ('ratio' parameter).

Statistical analysis

In the allergic airway inflammation experiments Mann-Whitney U tests were applied for comparison of the numbers of various immune cell populations between groups of mice. Fisher's exact test was performed using the BEDTools *fisher* function. All other tests for statistical significance were performed by the HOMER analysis suite or using RStudio. P values < 0.05 were considered statistically significant.

Accession numbers

All mouse and human ChIPmentation and RNA-Seq datasets were deposited in the Gene Expression Omnibus (GEO), accession number GSE98843.

ACCEPTED MANUSCRIPT

Supplementary Figure Legends

Supplementary Figure 1. GATIR mouse validation experiments. (A) Histograms showing GATA3 mean fluorescent intensity levels (antibody staining) in ILC2s isolated from BAL or MLN of GATIR reporter or wildtype (WT) mice. Lineage(Lin)⁺ CD4⁻ cells are shown (dashed line) as a GATA3⁻ control. (B) Flowcytometric quantification of eosinophils in BAL fluid of wildtype (WT) and GATIR mice after *in vivo* treatment with IL-33. (C) Flowcytometric quantification of CD4⁺ T cells in BAL fluid or MLN of wildtype (WT) and GATIR mice after *in vivo* treatment with IL-33. (D) Proportions of cytokine-producing ILC2s (after *in vitro* stimulation with PMA and ionomycin for 4 hours²) in BAL fluid or MLN of wildtype (WT) and GATIR mice after *in vivo* treatment with IL-33. All panels depict concatenated data from 3-5 mice; a representative of three independent experiments is shown.

Supplementary Figure 2. Transcriptome, microscopy and epigenome analyses of ILC2s in AI. (A) Scatterplot of log₂ RPKM values for both RNA-Seq biological replicate experiments per tissue. Pearson correlation values (R^2) are indicated. (B) Gene expression levels (n=2 pools of 5-6 mice, mean values + SEM are shown) in purified ILC2s of archetypical marker genes specifically expressed in other hematopoietic cell types. Dashed line indicates the 2 RPKM cut-off used throughout this study to define actively expressed genes. (C) Lung or MLN cryosections from IL-33 treated GATIR mice counterstained with CD3 (T cell marker) and B220 (B cell marker). ILC2s (YFP⁺CD3⁻) are indicated by white arrowheads. Dashed line denotes airway epithelium.

Supplementary Figure 3. Quality control metrics of murine ILC2 H3K4Me2⁺ epigenome maps. (A) H3K4Me2 (black) and non-specific IgG (gray) signals at example loci that depict highly expressed genes (e.g. *Gata3*), moderately expressed genes (e.g. *Kit*) and lowly or non-expressed genes (e.g. the neuronal *Olig1* gene) in BAL ILC2s. (B) Meta-profiles depicting average enrichment levels for all H3K4Me2⁺ GREs (black line) in BAL (left) or MLN (right) ILC2s for both replicate pools. Signal measured in a mock experiment (using a non-specific IgG) is shown in grey. (C) Orange bars denote average gene expression levels (errors bars showing 95% confidence interval) of genes ranked by promoter H3K4Me2 levels, which were grouped in four quartiles of increasing H3K4Me2 enrichment (average log₂ normalized H3K4Me2 levels are indicated by black lines within the orange bars). Note that gene promoter H3K4Me2 enrichment levels positively

correlate with mRNA expression levels. A Wilcoxon rank-sum test was used to compute *P* values for quartile comparisons.

Supplementary Figure 4. Highly similar H3K4Me2⁺ epigenomes in transcriptionally divergent ILC2 populations. (A) Number of H3K4Me2⁺ regions detected in both ChIPmentation biological replicates per tissue. (B) Box plots depicting expression levels of genes differentially expressed between ILC2s from BAL or MLN (either upregulated in BAL ILC2s or upregulated in MLN ILC2s) are shown at the top, lower boxplots show normalized promoter H3K4Me2 levels of the same gene sets. Differences between BAL and MLN expression or H3K4Me2 levels were tested for statistical significance using a Wilcoxon rank-sum test (N.S. = not significant, *P* > 0.05). (C) Venn diagram showing the overlap between H3K4Me2⁺ regions detected in paired ChIPmentation biological replicates per tissue. (D) Average RPKM expression values for all listed genes in the Interferon Signaling Gene Ontology term associated with BAL-specific H3K4Me2⁺ regions (*P* = 0.00443, paired two-tailed Student's t-test).

Supplementary Figure 5. Epigenetic priming in activated ILC2s. (A) Molecular pathway enrichment (MSigDB) amongst genes associated with active (yellow bars) or primed (grey bars) GREs in ILC2s. (B) Genome browser views of the H3K4Me2 landscape around selected loci implicated in airway remodeling in BAL ILC2s. (C) Gene expression levels (n=2 pools of 5-6 mice, mean values + SEM are shown) in purified ILC2s of genes shown in panel B. Dashed line indicates the 2 RPKM cut-off used throughout this study to define actively expressed genes.

Supplementary Figure 6. The GATA3/GFI1 binding landscape in ILC2s. (A) Heatmap showing GATA3 (left, in red) and GFI1 (right, in blue) normalized ChIP-sequencing signal^{16, 17} (reads per million, RPM) for all genomic binding sites. Input genomic DNA signals are shown as negative control experiments. (B,C) Bar graph showing the distance distribution of GATA3 (panel B) or GFI1 (panel C) binding sites to the nearest transcription start site (TSS). Pie charts depict the proportion of ILC2 superenhancers bound by GATA3 (panel B) or GFI1 (panel C). (D,E) Number of binding sites (top) and percentage of overlap with H3K4Me2 enriched regions (bottom) for GATA3 (panel D) or GFI1 (panel E). (F) Venn diagrams showing overlap between GATA3 and GFI1 target genes (nearest gene within 1 Mb, left) or binding sites (at least 1 bp overlap, right). (G) Molecular pathway enrichment (MSigDB) amongst genes co-bound by GATA3 and GFI1 or bound specifically by one transcription factor.

Supplementary Figure 7. ILC2 superenhancers reside in open, transcription-permissive chromatin. (A) Meta-profiles of ILC2 ATAC-Seq signals²⁴ across all H3K4Me2-based ILC2 superenhancers (centered on superenhancer midpoints). (B) Genome browser view of BAL ILC2 H3K4Me2 (blue) and RNA-Seq (grey) profiles for chromosome 15. Superenhancer locations are shown below. (C) Genome browser view of the H3K4Me2 landscape around the *Gata3* gene in BAL ILC2s. Superenhancer (SE) elements are indicated by yellow shading, *tce1* is a previously identified³¹ *Gata3* enhancer in T cells.

Supplementary Figure 8. Similar expression levels of asthma-associated genes in ILC2s generated by IL-33 or chronic allergen exposure. Heatmap shows average gene expression levels in BAL ILC2s isolated from GATIR mice in which airway inflammation was induced by IL-33 or chronic allergen (house dust mite, HDM) exposure (mean of n=2 pools of 5-6 mice; log₂ RPKM counts generated using RNA-Seq). Expression levels are shown for the 665 asthma-associated genes identified by GWAS, which were further divided into active, primed or silent genes based on their combined epigenome/transcriptome profile (see Fig.4) in IL-33 activated ILC2s.

Supplementary Figure 9. H3K4Me2 dynamics during human ILC2 activation. (A) Number of H3K4Me2⁺ regions detected in all ChIPmentation biological replicates for naïve ILC2s (nILC2), activated ILC2s (aILC2) and circulating memory Th2 cells. The number and average percentage of shared enriched regions is indicated on the right. (B) Heatmap depicting normalized H3K4Me2 signal (in reads per million, RPM) for nILC2-specific (n=863), shared (n=16541) and aILC2-specific (n=14957) H3K4Me2⁺ gene regulatory elements (GREs). (C) Venn diagram indicating the overlap between H3K4Me2⁺ regions detected in nILC2s and aILC2s. (D) Pie charts showing the overlap between regions of differential DNA methylation associated with serum IgE or asthma from the indicated epigenome-wide association studies (epiGWAS) and H3K4Me2⁺ GREs in human ILC2s.

Supplementary Figure 10. Inter-species comparison of epigenomic features in activated ILC2s. (A) Human and murine IL-33 activated ILC2s share a conserved core of 117 superenhancer (SE)-associated genes. Selected genes are shown boxed; genome browser views of the *Gata3/GATA3* and *Rora/RORA* loci are shown to indicate conserved ILC2 SE structure (denoted by a grey rectangle) between humans and mice. (B) Genome browser view of key ILC1 (*TBX21*) and

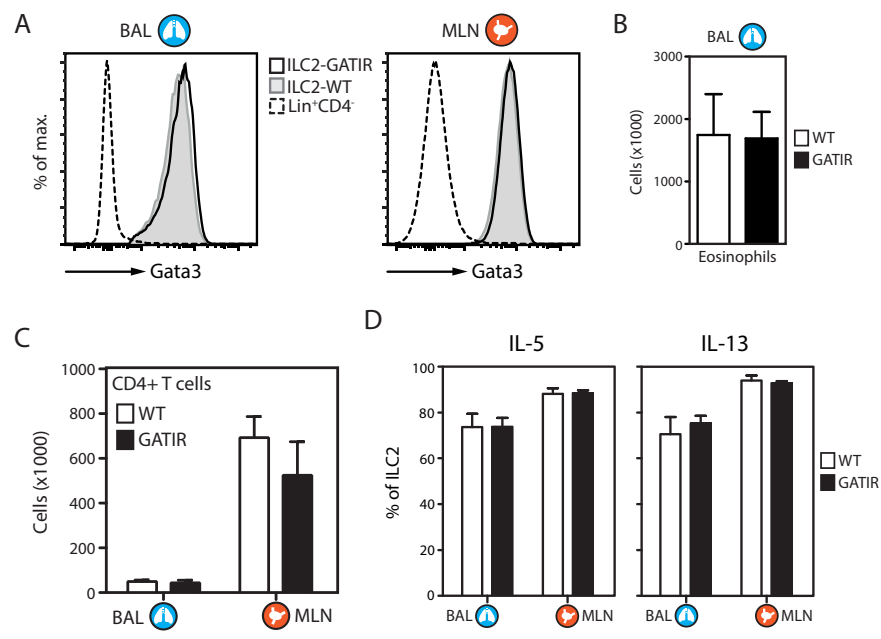
ILC₃ (*IL17*, *RORC*) loci in ILC2s. Note the extensive decoration of these loci with H₃K₄Me₂, indicating a primed state in ILC2s. (C) Venn diagrams showing a comparison of GWAS asthma genes (n=665) associated with (top) or not associated with (bottom) H₃K₄Me₂⁺ gene regulatory elements in murine (red) or human (blue) ILC2s.

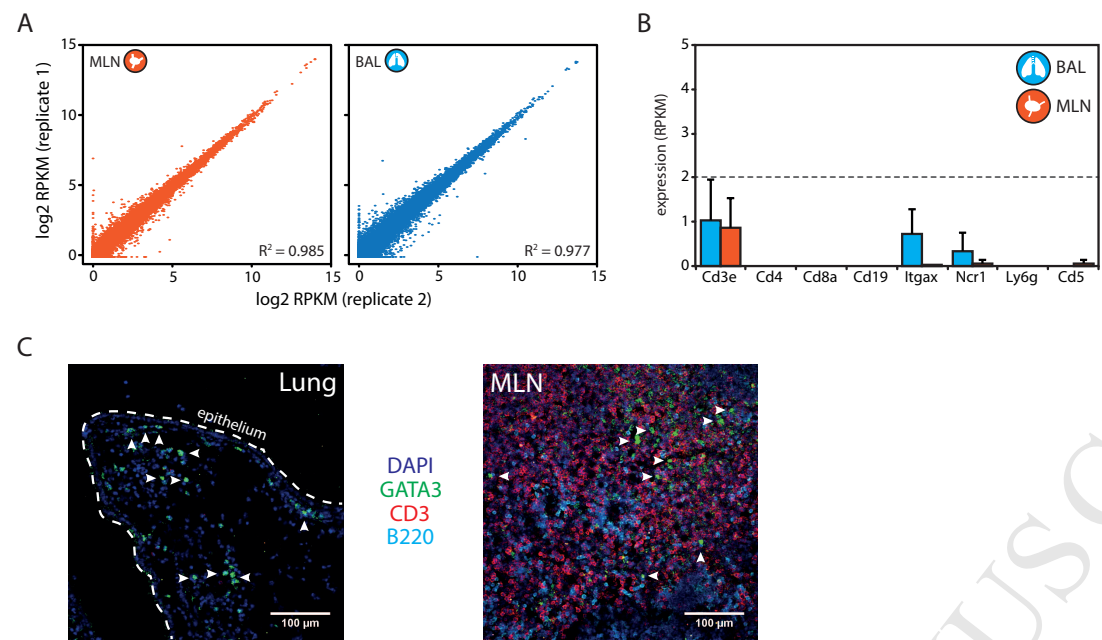
ACCEPTED MANUSCRIPT

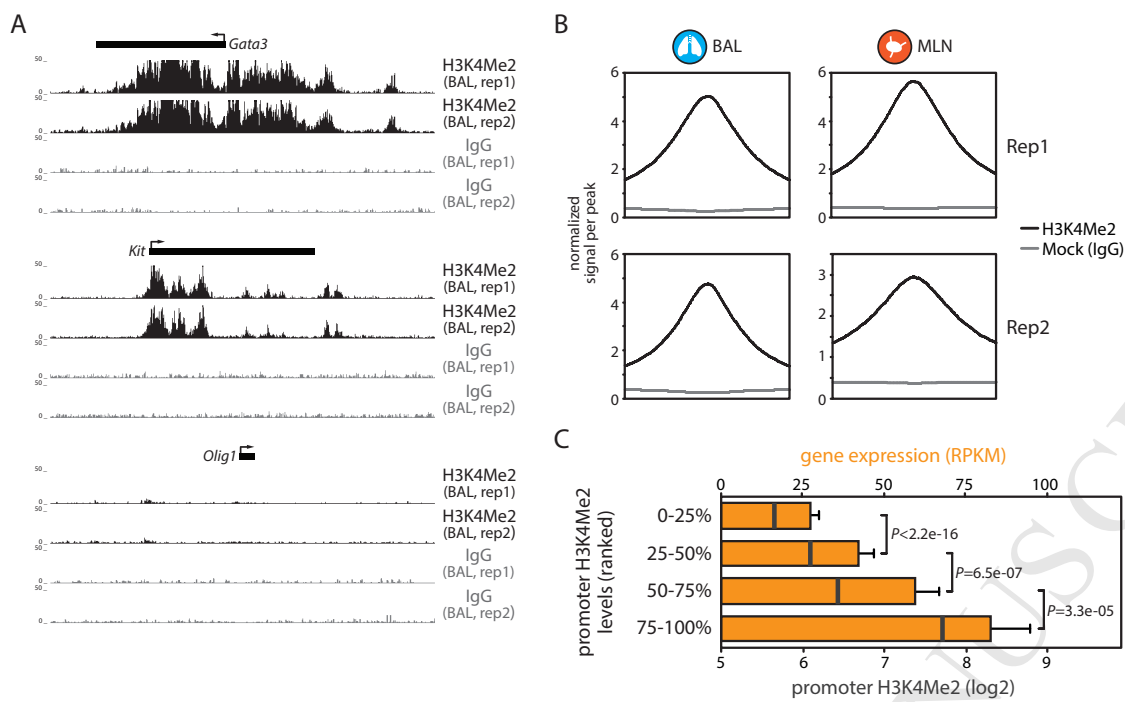
References

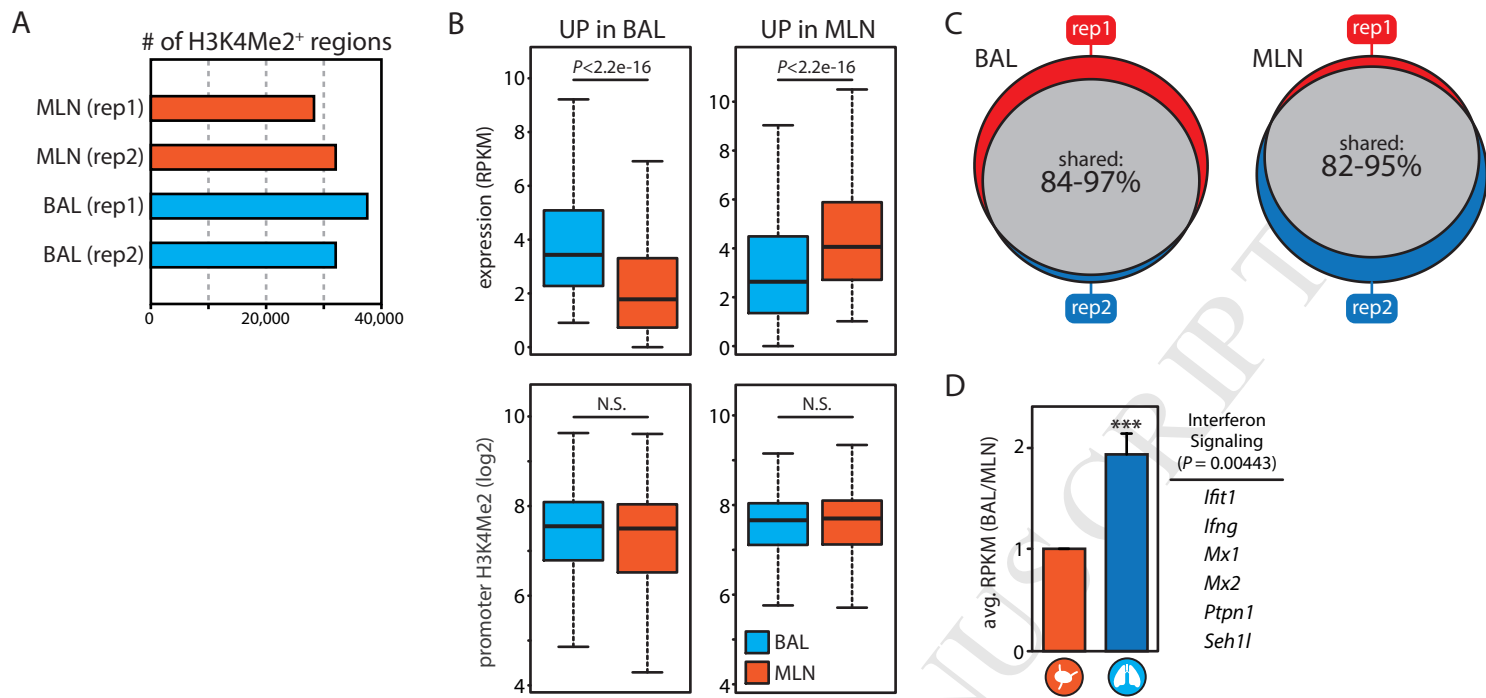
1. Li BWS, Stadhouders R, de Bruijn MJW, Lukkes M, Beerens DMJM, Brem MD, et al. Group 2 Innate Lymphoid Cells Exhibit a Dynamic Phenotype in Allergic Airway Inflammation. *Front Immunol* 2017; doi:10.3389/fmmu.2017.01684.
2. Klein Wolterink RG, Kleinjan A, van Nimwegen M, Bergen I, de Bruijn M, Levani Y, et al. Pulmonary innate lymphoid cells are major producers of IL-5 and IL-13 in murine models of allergic asthma. *Eur J Immunol* 2012; 42:1106-16.
3. Lim AI, Menegatti S, Bustamante J, Le Bourhis L, Allez M, Rogge L, et al. IL-12 drives functional plasticity of human group 2 innate lymphoid cells. *J Exp Med* 2016; 213:569-83.
4. Lim AI, Li Y, Lopez-Lastra S, Stadhouders R, Paul F, Casrouge A, et al. Systemic Human ILC Precursors Provide a Substrate for Tissue ILC Differentiation. *Cell* 2017; 168:1086-100 e10.
5. Broos CE, van Nimwegen M, In 't Veen JC, Hoogsteden HC, Hendriks RW, van den Blink B, et al. Decreased Cytotoxic T-Lymphocyte Antigen 4 Expression on Regulatory T Cells and Th17 Cells in Sarcoidosis: Double Trouble? *Am J Respir Crit Care Med* 2015; 192:763-5.
6. Picelli S, Bjorklund AK, Faridani OR, Sagasser S, Winberg G, Sandberg R. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat Methods* 2013; 10:1096-8.
7. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* 2015; 12:357-60.
8. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014; 15:550.
9. Vroman H, Bergen IM, Li BW, van Hulst JA, Lukkes M, van Uden D, et al. Development of eosinophilic inflammation is independent of B-T cell interaction in a chronic house dust mite-driven asthma model. *Clin Exp Allergy* 2017; 47:551-64.
10. Stadhouders R, Thongjuea S, Andrieu-Soler C, Palstra RJ, Bryne JC, van den Heuvel A, et al. Dynamic long-range chromatin interactions control Myb proto-oncogene transcription during erythroid development. *EMBO J* 2012; 31:986-99.
11. Egelhofer TA, Minoda A, Klugman S, Lee K, Kolasinska-Zwierz P, Alekseyenko AA, et al. An assessment of histone-modification antibody quality. *Nat Struct Mol Biol* 2011; 18:91-3.
12. Schmidl C, Rendeiro AF, Sheffield NC, Bock C. ChIPmentation: fast, robust, low-input ChIP-seq for histones and transcription factors. *Nat Methods* 2015; 12:963-5.
13. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012; 9:357-9.
14. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* 2010; 38:576-89.
15. Landt SG, Marinov GK, Kundaje A, Kheradpour P, Pauli F, Batzoglou S, et al. ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res* 2012; 22:1813-31.
16. Zhong C, Cui K, Wilhelm C, Hu G, Mao K, Belkaid Y, et al. Group 3 innate lymphoid cells continuously require the transcription factor GATA-3 after commitment. *Nat Immunol* 2016; 17:169-78.
17. Spooner CJ, Lesch J, Yan D, Khan AA, Abbas A, Ramirez-Carrozzi V, et al. Specification of type 2 innate lymphocytes by the transcriptional determinant Gfi1. *Nat Immunol* 2013; 14:1229-36.
18. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010; 26:841-2.

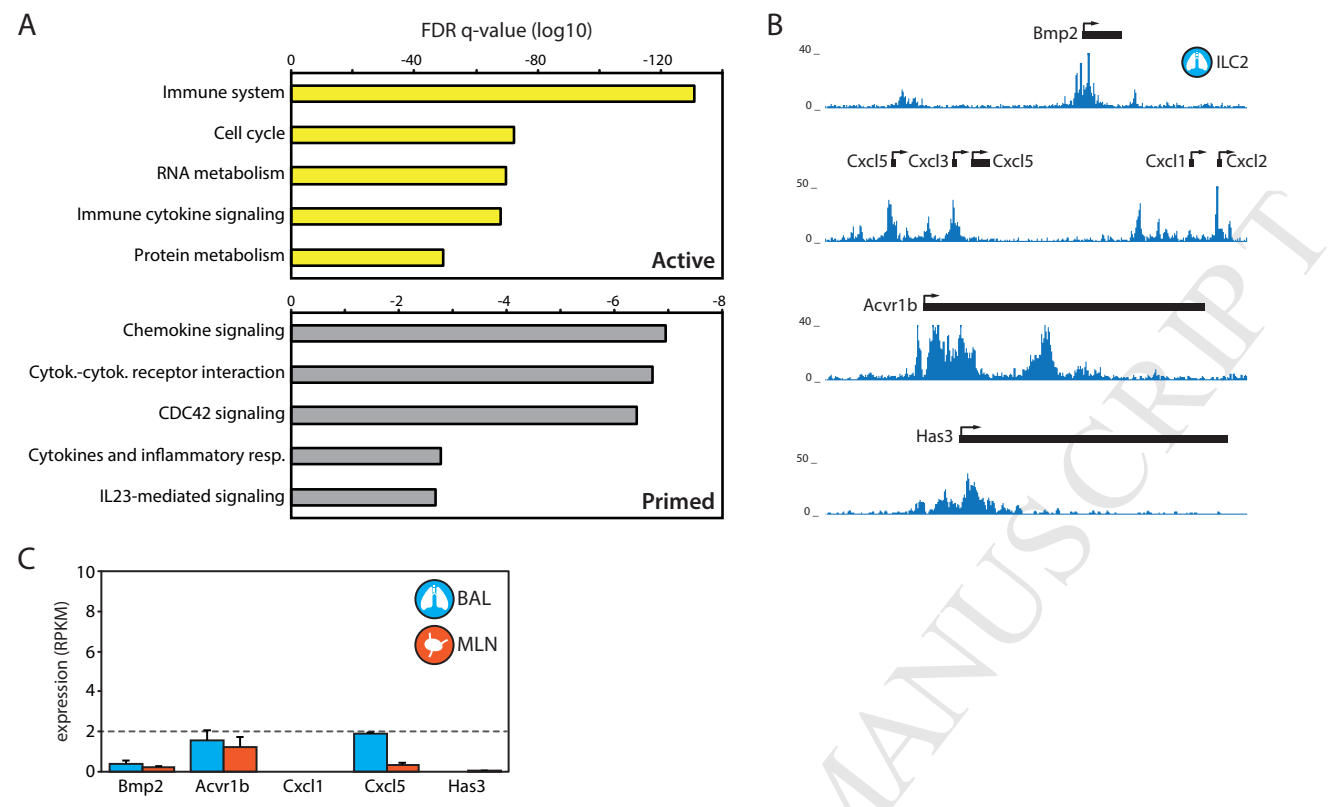
19. Saldanha AJ. Java Treeview--extensible visualization of microarray data. *Bioinformatics* 2004; 20:3246-8.
20. Chen H, Boutros PC. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics* 2011; 12:35.
21. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, et al. GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol* 2010; 28:495-501.
22. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005; 102:15545-50.
23. Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* 2013; 153:307-19.
24. Shih HY, Sciume G, Mikami Y, Guo L, Sun HW, Brooks SR, et al. Developmental Acquisition of Regulomes Underlies Innate Lymphoid Cell Functionality. *Cell* 2016; 165:1120-33.
25. Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res* 2014; 42:D1001-6.
26. Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* 2008; 24:2938-9.
27. Farh KK, Marson A, Zhu J, Kleinewietfeld M, Housley WJ, Beik S, et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* 2015; 518:337-43.
28. Liang L, Willis-Owen SA, Laprise C, Wong KC, Davies GA, Hudson TJ, et al. An epigenome-wide association study of total serum immunoglobulin E concentration. *Nature* 2015; 520:670-4.
29. Yang IV, Pedersen BS, Liu A, O'Connor GT, Teach SJ, Kattan M, et al. DNA methylation and childhood asthma in the inner city. *J Allergy Clin Immunol* 2015; 136:69-80.
30. Drazen JM, Yandava CN, Dube L, Szczerback N, Hippensteel R, Pillari A, et al. Pharmacogenetic association between ALOX5 promoter genotype and the response to anti-asthma treatment. *Nat Genet* 1999; 22:168-70.
31. Hosoya-Ohmura S, Lin YH, Herrmann M, Kuroha T, Rao A, Moriguchi T, et al. An NK and T cell enhancer lies 280 kilobase pairs 3' to the gata3 structural gene. *Mol Cell Biol* 2011; 31:1894-904.

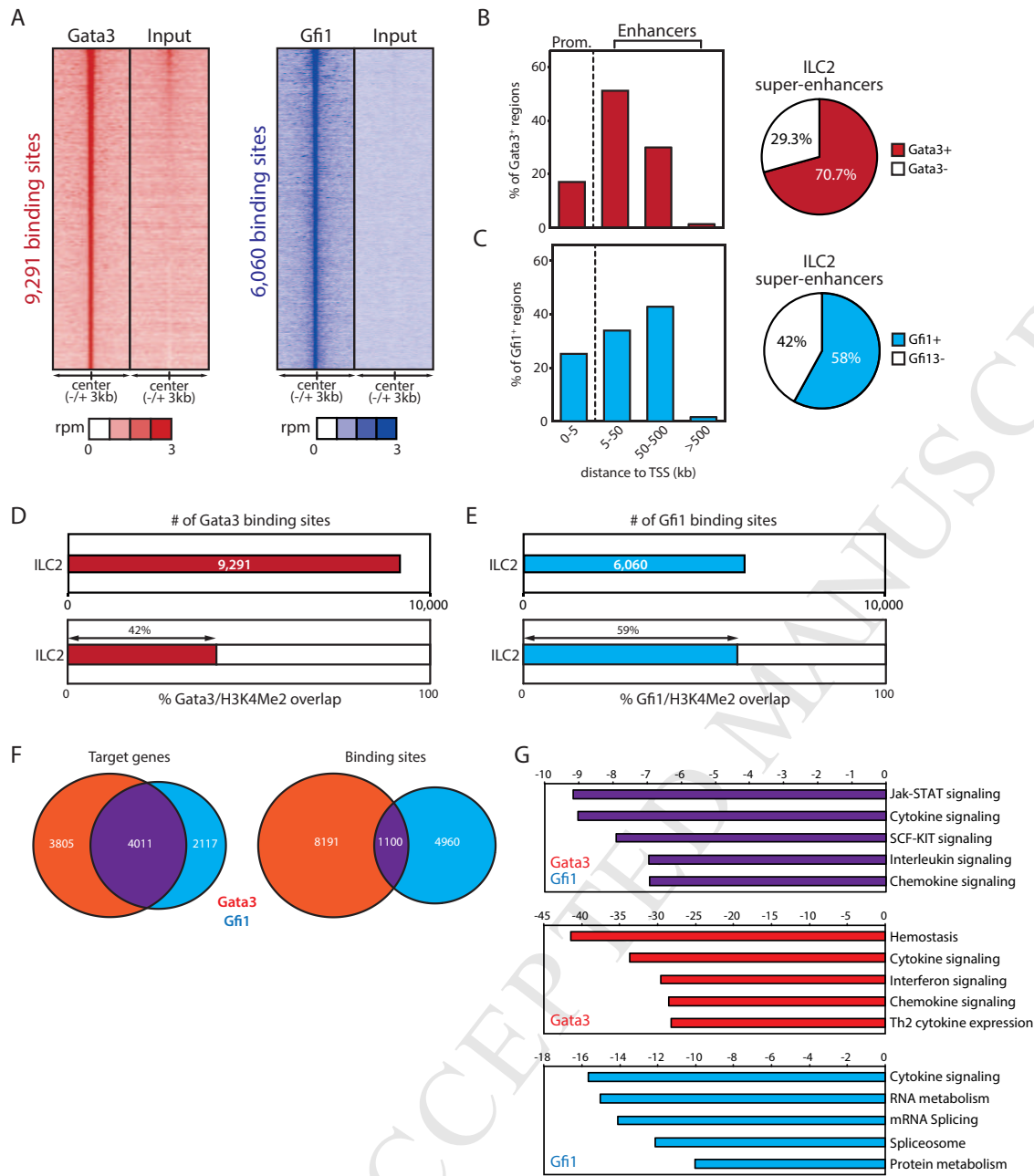


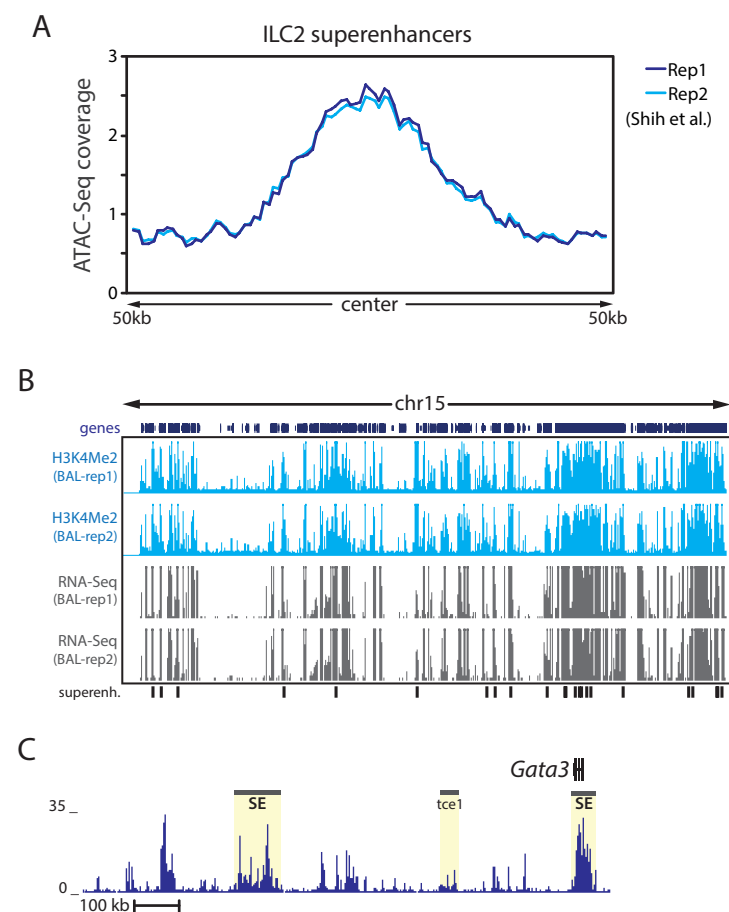


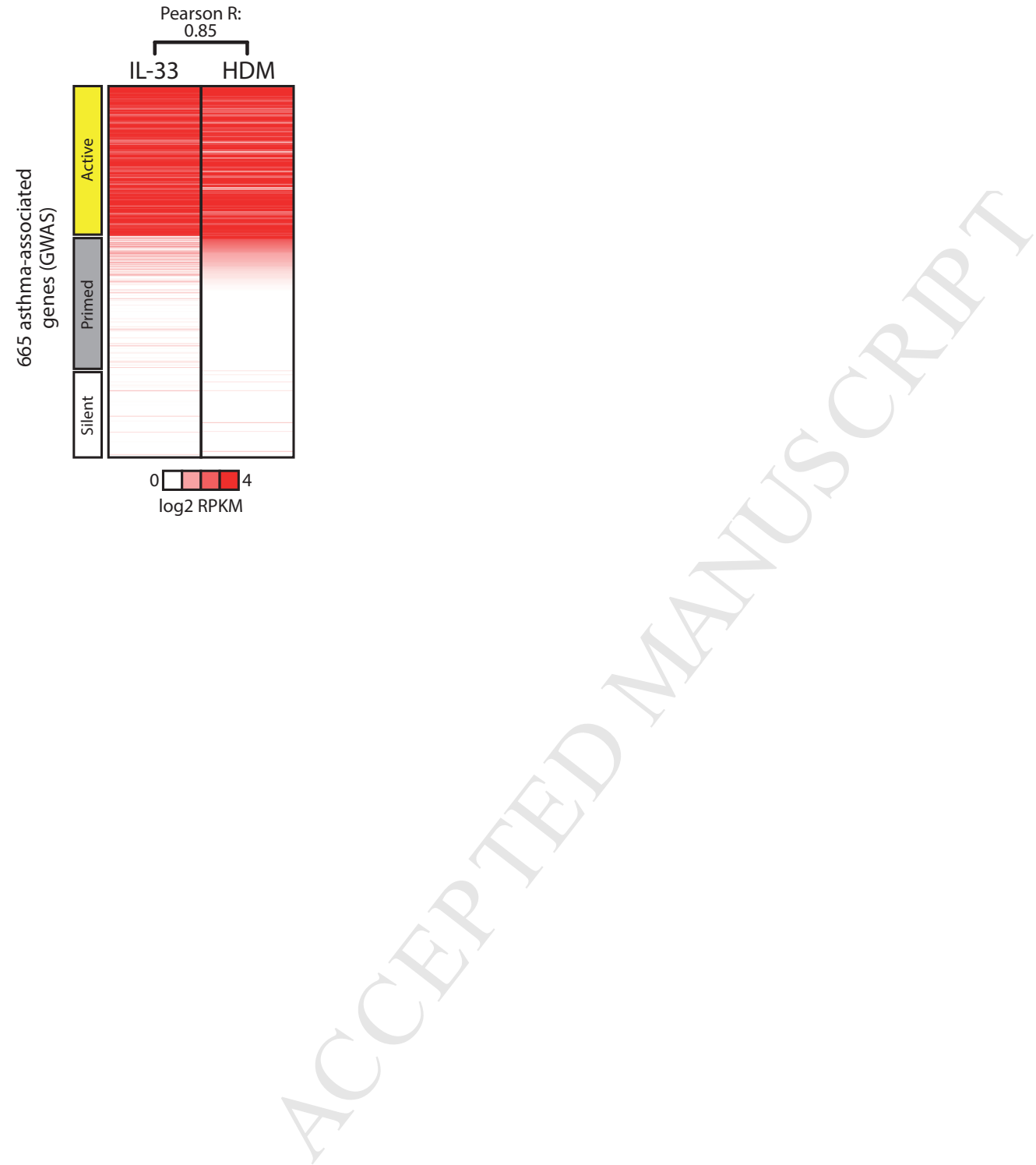


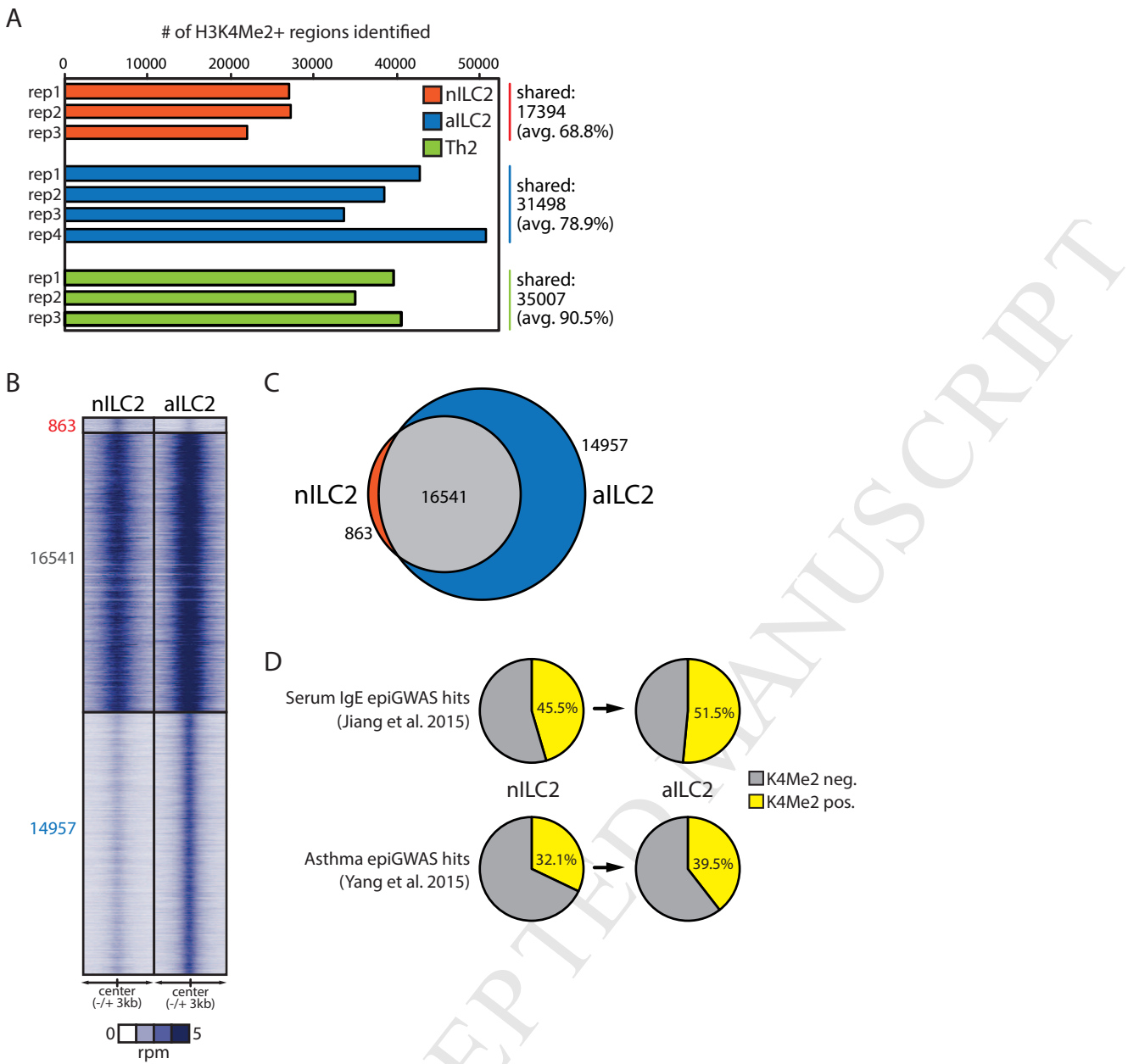


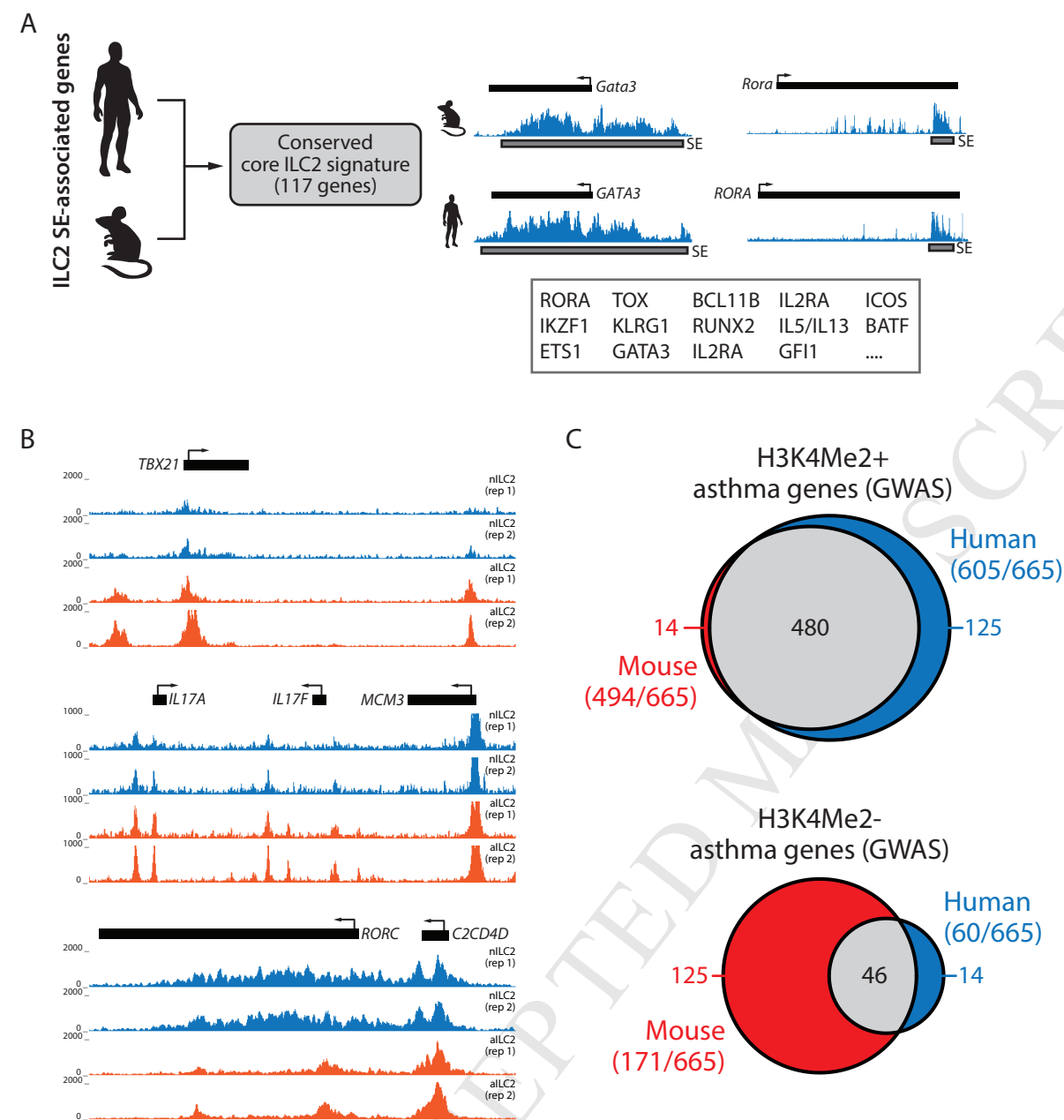












Supplementary Table 1. Genes differentially expressed between BAL and MLN

gene	baseMean	log2FC_BAL/MLN	lfcSE	stat
Nupr1	1124.765018	6.566712637	0.285836286	-22.97368446
Akr1c18	178.7247329	5.083376028	0.426862847	-11.90868698
E230016K23R	180.0855449	4.877855397	0.49430343	-9.868139888
Ccl6	177.0206257	4.859988152	0.494589836	-9.826300091
Ciart	350.2079944	4.794742478	0.354994664	-13.50651987
Slc7a3	401.0748525	4.794320496	0.314741001	-15.23258958
Tph1	2302.745111	4.438899672	0.19347387	-22.94314817
C920021L13R	362.5088269	4.394478603	0.339940389	-12.92720355
Aqp9	1238.587296	4.248168771	0.503399029	-8.438968935
LOC10263175	845.2402185	4.232723967	0.217869324	-19.42781064
Tff1	1145.621966	4.213589386	0.196659294	-21.42583395
Ccl7	66.8162463	4.211923699	0.516866039	-8.148965855
Asic3	70.27663739	4.208226696	0.518252664	-8.120029072
Cd63	434.7150615	4.159530333	0.288007436	-14.44244077
Il9	65.6915651	4.143052508	0.519473105	-7.975489915
Amica1	113.1843796	4.115360021	0.455272445	-9.039334717
Clec7a	76.14699732	4.072313951	0.525256914	-7.752994474
Atp6v0d2	98.16295111	4.064407687	0.522599494	-7.777289754
Spsb1	109.3978617	3.966787984	0.456473442	-8.690073984
Aldh1l2	63.16931183	3.966610385	0.51483268	-7.704659281
Mmp12	83.44783437	3.878521116	0.526289613	-7.369556651
Mgll	252.0416958	3.872139095	0.33175371	-11.67172807
Ctla4	1365.937796	3.837311952	0.178335602	-21.51736335
Hmox1	219.0031544	3.834873702	0.388820148	-9.862847197
Tnfrsf4	833.2381917	3.803883714	0.224484442	-16.94497705
Lad1	50.50004455	3.734563239	0.530246238	-7.043073521
Ermn	78.97296482	3.653739741	0.501362137	-7.287625986
Tnfsf8	1053.192122	3.635239443	0.209193615	-17.37739194
Il21	47.49811708	3.626298386	0.532997945	-6.803587937
Fcgr3	63.0073817	3.542278342	0.525483087	-6.74099401
Myo16	290.8570139	3.529941159	0.306963992	-11.49952846
Atf3	72.83888524	3.462567495	0.500976188	-6.911640873
Slc6a9	118.6278738	3.461957598	0.4414586	-7.842088919
Gcnt2	60.30784748	3.445648983	0.510997106	-6.74299119
Ccl9	51.20713885	3.382016513	0.540795254	-6.253783642
Tnc	128.3048925	3.350894087	0.416625176	-8.042946697
Egfr	47.19977458	3.35042628	0.53428049	-6.270912637
Havcr2	161.8433009	3.317680826	0.37311291	-8.891895031
Asns	3896.24922	3.292457354	0.170232703	-19.34092151
Dusp18	49.4732649	3.287799136	0.530077292	-6.202490059
Cth	42.55679802	3.282332756	0.534932645	-6.135973919

Supplementary Table 2. Genes associated with superenhancers in murine and I

Mouse ILC2 (BAL/MLN)	Human nILC2-aILC2 (shared)	Human nILC2-specific
0610039K10Rik	HIST1H4E	GPR183
0610040F04Rik	VAMP2	PIP4K2A
1110028F11Rik	DAD1	MPZL3
1110059E24Rik	SLAMF1	TARP
1700007L15Rik	RORA	IL7R
1700016G22Rik	SOS1	DDIT4
1700061F12Rik	IKZF2	TRG-AS1
1700061F12Rik	SMAD3	ARHGAP15
1810013L24Rik	CYTH1	GRAP2
1810026B05Rik	SENCR	LOC389641
2010111I01Rik	TESPA1	VPS45
2010320M18Rik	ATN1	ITGB2
2310036O22Rik	MBNL1-AS1	ATP8B2
2500004C02Rik	MIR3175	ZBTB16
4833419F23Rik	IL2RA	SELL
4921511C10Rik	CELF2	AXIN1
4930412O13Rik	GATA3	CD7
4930429F24Rik	STK17B	PRSS3P2
4930478L05Rik	HOXB5	ZFP36L1
4930523C07Rik	CCDC85B	SESN1
4932438H23Rik	RARA	MGAT5
4933411E08Rik	ITK	RORC
5031425E22Rik	CYTIP	PRR3
5830428M24Rik	CDC42	HELB
6030468B19Rik	ETS1	TLE4
6330415G19Rik	CDC42SE2	MIR6855
9330175E14Rik	TNFAIP8	AKT1S1
9530052E02Rik	CARNS1	ATPIF1
A630072M18Rik	TSPOAP1-AS1	ITPKB-IT1
A930006K02Rik	CHST11	DENND2D
Abhd15	LOC101928120	TNFRSF9
Ablim1	RGL2	LINC01814
Abt1	RASSF1	FASLG
Acot7	EVL	TRIB2
Actr3	GTPBP8	MALAT1
Adamtsl4	MIDN	NR4A2
Adat3	NDUFV2	TP53INP1
Adcy7	DVL2	CD28
Adipor2	ITGB7	LINC01307
Aebp2	HIST1H2BB	IDI1
Aff4	ZEB2-AS1	UCP2

Supplementary Table 3. Genes located adjacent asthma-associated SNPs identified by GWAS

AAGAB	rs17294280 (+78789), rs2278546 (+80472), rs72743482 (+80475), rs72743477 (+82
AARD	rs921067 (+6863), rs2921722 (+9709), rs3020107 (+9776), rs3020108 (+9948), rs30
ABI3BP	rs9823506 (+235646)
ACMSD	rs10496731 (+1511), rs3739029 (+3290), rs1530556 (+3617), rs1530555 (+3622), rs
ACO1	rs10970975 (+47313), rs1556138 (+47947), rs1467713 (+48029), rs1467712 (+4806
ACPL2	rs7624084 (+142615), rs6440003 (+143539), rs4683605 (+144099), rs6764769 (+14
ACSL6	rs115057521 (+108988), rs10063686 (+147674), rs28689629 (+212802), rs2850316
ACTN4	rs74449234 (+56260), rs76878864 (+91926)
ADAD1	rs45605540 (-159114), rs72687036 (-142027), rs45515895 (-115341), rs7693745 (-1
ADAMTSL1	rs114646018 (+416169)
ADCY8	rs7000447 (+531389)
ADCY9	rs35714728 (+92717), rs66801518 (+92998), rs67335251 (+94464), rs71388556 (+9
ADSS	rs4658627 (+104260)
AGER	rs204993 (-3480)
AIM2	rs114923675 (-45346), rs3122355 (-37506)
AIPL1	rs73976923 (+56034), rs79392787 (+88451)
AJAP1	rs12568266 (-521554)
AKAP11	rs347412 (-59609), rs347414 (-58889), rs347416 (-58427), rs369829 (-43159), rs23
AKR1E2	rs11252394 (-627576)
ALDH1A2	rs566767 (+276249), rs487836 (+324757)
ALDH7A1	rs114776225 (+78998), rs28468587 (+79645), rs75561589 (+81628), rs77885874 (+
ALOX5	promoter variants (Drazen et al. Nature Genetics 1999)
ALPK2	rs9963330 (+275680)
ANAPC1	rs7594583 (+148658), rs11677031 (+170628), rs11683921 (+170634), rs11122896 (
ANXA1	rs3000643 (+893271), rs1032365 (+893619), rs1463699 (+894611), rs1491440 (+89
AP3B1	rs73132886 (+241939)
AP5S1	rs4815617 (+26106)
ARHGAP15	rs2890729 (+704986), rs2178734 (+705223), rs2593719 (+706278), rs2593718 (+70
ARHGAP21	rs10828735 (-227487)
ARHGAP24	rs342958 (-115629), rs16995986 (+297050)
ARL4A	rs17167307 (+956802)
ARPP21	rs17033506 (-42052)
ARRDC4	rs77062951 (-71211), rs1442808 (+260919)
ASAP1	rs7000447 (-109066)
ASAP2	rs73912949 (-55660)
ASXL3	rs117022740 (-100659), rs116850624 (-73322), rs78173567 (-70708), rs76650652 (
ATG3	rs2705520 (+11523)
ATP10A	rs79252495 (-613661)
ATP1A1	rs61786305 (-178305), rs74346392 (-154795), rs61786312 (-138148), rs76298144 (
ATP1B1	rs1556793 (-374523), rs1556792 (-374499), rs1556791 (-374450), rs1556790 (-374
ATRN	rs171165 (+5929)
B4GALNT3	rs80251893 (-19527), rs77331693 (-1521), rs79836015 (+9426), rs215230 (+24707)

Supplementary Table 4. Antibodies used for flow cytometry and confocal microscopy

Antibody	Conjugate	Clone	Company
<i>mouse studies</i>			
B220	PE	RA3-6B2	eBioscience
CD11b	PE	M1/70	eBioscience
CD11c	EF450	N418	eBioscience
CD11c	PE	N418	eBioscience
CD19	PE	1D3	BD Biosciences
CD19	PerCP-Cy5.5	eBio1D3	eBioscience
CD3	PE	145-2c11	eBioscience
CD4	FITC	RM4-5	eBioscience
CD4	PE	GK1.5	eBioscience
CD5	PE	53-7.3	eBioscience
CD8	APC	53-7.3	eBioscience
CD8	PE	53-6.7	eBioscience
FcεRIα	PE	MAR-1	eBioscience
Gr-1	APC-eFluor 780	RB6-8C5	eBioscience
Gr-1	PE	RB6-8C5	BD Biosciences
NK1.1	PE	PK136	eBiosciences
Sca-1	Pacific Blue	D7	BioLegend
Siglec-F	PE-CF594	E50-2440	BD Biosciences
Streptavidin	APC		eBioscience
T1/ST2	Biotin	DJ8	MD Bioproducts
TER-119	PE	TER-119	eBioscience
KLRG1	APC	2F1	BD Biosciences
Ki-67	FITC	SolA15	eBioscience
B220 (confocal)	Biotin	TY25	eBioscience
CD3 (confocal)	Biotin	145-2C11	BD Biosciences
α-rat (confocal)	Cy5	Polyclonal	Jackson ImmunoResearch
α-hamster (confocal)	Cy3	Polyclonal	Jackson ImmunoResearch
<i>human studies</i>			
CD3	PE-Cy7	SK7	BD Biosciences
CD4	PE	OKT4	eBioscience
CCR6	Biotin	11A9	BD Biosciences
CXCR3	BV711	1C6	BD Biosciences
CCR4	FITC	205410	R&D systems
CD45RA	PE-TxR	MEM-56	Invitrogen

Supplementary Table 5. Human ILC2 GREs carrying GWAS (combined) or epi-GWAS (Yang et al.

chr	start	end	Size (bp)	Study
chr1	10554204	10557956	3752	GWAS
chr10	6069659	6117790	48131	GWAS
chr10	6121330	6137242	15912	GWAS
chr10	8938304	8941852	3548	GWAS
chr10	28590056	28594370	4314	GWAS
chr11	76298901	76300387	1486	GWAS
chr11	102979142	102982690	3548	GWAS
chr11	113928553	113958743	30190	GWAS
chr11	128185865	128189560	3695	GWAS
chr12	567579	571360	3781	GWAS
chr12	56365961	56370639	4678	GWAS
chr12	56389473	56392822	3349	GWAS
chr12	56433858	56438238	4380	GWAS
chr12	56471933	56475332	3399	GWAS
chr12	117255266	117258526	3260	GWAS
chr15	67436190	67444021	7831	GWAS
chr16	27412818	27414697	1879	GWAS
chr17	1925524	1935948	10424	GWAS
chr17	1951065	1964468	13403	GWAS
chr17	37909163	37915328	6165	GWAS
chr17	37929525	37935291	5766	GWAS
chr17	37937167	37939601	2434	GWAS
chr17	37968353	37977145	8792	GWAS
chr17	38016057	38023267	7210	GWAS
chr17	38076378	38086266	9888	GWAS
chr17	65358485	65364202	5717	GWAS
chr17	65371501	65376938	5437	GWAS
chr17	65386644	65388099	1455	GWAS
chr18	48493313	48496917	3604	GWAS
chr19	52529037	52533531	4494	GWAS
chr19	52549994	52554202	4208	GWAS
chr2	8440414	8445302	4888	GWAS
chr2	8449157	8454966	5809	GWAS
chr2	53160359	53163976	3617	GWAS
chr2	102926378	102933460	7082	GWAS
chr2	102935245	102937589	2344	GWAS
chr2	102948822	102958570	9748	GWAS
chr2	102966976	102968858	1882	GWAS
chr2	102970484	102986900	16416	GWAS
chr2	112463003	112464004	1001	GWAS
chr2	135674101	135681292	7191	GWAS