HHS Public Access

Author manuscript

J Immunol. Author manuscript; available in PMC 2022 August 15.

Published in final edited form as:

J Immunol. 2021 August 15; 207(4): 1044–1054. doi:10.4049/jimmunol.2000207.

Epigenetic analysis of the murine chromatin landscape identifies a repertoire of eosinophil-specific PU.1-bound enhancers

Jennifer M. Felton[*], **Sushmitha Vallabh**[*], **Sreeja Parameswaran**[#], **Lee E. Edsall**[#], **Kevin Ernst**[#], **Benjamin Wronowski**[*], **Astha Malik**[*], **Michael Kotliar**[*], **Matthew T. Weirauch**[‡,#,¶], **Artem Barski**[*,§,‡], **Patricia C. Fulkerson**[*,‡,1], **Marc E. Rothenberg**[*,‡,1,2]

[*]Division of Allergy and Immunology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA

[§]Division of Human Genetics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA

[‡]Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH, USA

[#]Center for Autoimmune Genomics and Etiology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA

[¶]Divisions of Biomedical Informatics and of Developmental Biology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA

Abstract

Eosinophils develop in the bone marrow from hematopoietic progenitors into mature cells capable of a plethora of immunomodulatory roles via the choreographed process of eosinophilopoiesis. However, the gene regulatory elements and transcription factors (TFs) orchestrating this process remain largely unknown. The potency and resulting diversity fundamental to an eosinophil's complex immunomodulatory functions and tissue specialization likely result from dynamic epigenetic regulation of the eosinophil genome—a dynamic eosinophil regulome. Herein, we applied a global approach using broad-range, next-generation sequencing to identify a repertoire of eosinophil-specific enhancers. We identified over 8,200 active enhancers located within 1 - 20 kB of expressed eosinophil genes. TF binding motif analysis revealed PU.1 (*Spi1*) motif enrichment in eosinophil enhancers, and ChIP-seq confirmed PU.1 binding in likely enhancers of genes highly expressed in eosinophils. A substantial proportion (>25%) of these PU.1-bound enhancers were unique to murine, culture-derived eosinophils when compared among enhancers of highly expressed genes of three closely related myeloid cell subsets (macrophages, neutrophils, immature granulocytes). Gene ontology analysis of eosinophil-specific, PU.1-bound enhancers revealed enrichment for genes involved in migration, proliferation, degranulation and survival. Furthermore, eosinophil-specific super enhancers were enriched in genes whose homologues are associated with risk loci for eosinophilia and allergic diseases. Our collective data identify

[2]**Corresponding Author**: Marc E. Rothenberg, MD, PhD– Marc.Rothenberg@cchmc.org +1 (513) 636-7177.

[1]Joint senior author

Author contributions

J.M.F, S.V., B.W. and A.M. performed experiments. J.M.F., S.V., S.P., L.E.E., K.E., M.K., M.T.W. and A.B. performed and assisted in bioinformatic analysis. J.M.F and L.E.E. generated the UCSC Genome Browser session. J.M.F., L.E.E., M.K., M.E.R., M.T.W., A.B. and P.C.F. contributed to the interpretation of the results. J.M.F., P.C.F., M.T.W. and M.E.R designed experiments and wrote the manuscript with intellectual contributions from all authors.

eosinophil-specific enhancers regulating key eosinophil genes through epigenetic mechanisms (H3K27 acetylation) and TF binding (PU.1).

Keywords

Eosinophil; enhancer; promoter; epigenetic; gene regulation; transcription factor; PU.1; Spi1

Introduction

The long-held view of eosinophils as end-stage cells involved primarily in allergy and host protection against parasites has been transformed by demonstrating that eosinophils are pleiotropic, multi-functional leukocytes involved in a plethora of biological processes, including tissue homeostasis, mammary gland development, estrus cycling, antigen presentation, cytokine production, epithelial remodeling, fat metabolism and cancer (1–8). Moreover, two subsets of regulatory eosinophils have recently been identified (2, 9), with transcriptomes and cellular phenotypes distinct from those of the conventional, inflammation-recruited eosinophil. The potency and resulting diversity fundamental to an eosinophil's complex immunomodulatory functions and tissue specialization likely result from dynamic epigenetic regulation of the genome—a dynamic eosinophil regulome. However, the gene regulatory elements (REs) that orchestrate the eosinophil regulome remain largely unknown.

The importance of key transcription factor (TF) families in eosinophil lineage commitment and maturation has been well documented. Most notably, there are critical roles for GATA, CEBPε, and PU.1 in eosinophil development, granule production and function (10–12). However, the role of these TFs in regulating enhancers has not been deeply studied. The development of new technologies that enable the global identification of genomic elements and the TFs that regulate expression of associated genes has provided opportunities to discern how *cis* REs, such as enhancers, control gene expression programs under normal and disease states (13). Epigenetic features enable identification of enhancer-like regions within cell populations, facilitating systematic investigation of their importance in gene transcription and cellular phenotype. Promoters are often identified by the presence of open regions marked by histone H3 lysine 4 trimethylation (H3K4me3), surrounded by regions marked by histone H3 lysine 4 monomethylation (H3K4me1) and located proximal (within 1 kB) to the transcription start site (TSS). In contrast, enhancers are often identified by open regions marked by H3K4me1 and are distal (>1 kB) to the TSS (14–19). Subsequently, enhancers can be classified as poised (lacking enriched acetylation of histone H3 lysine 27 (H3K27ac)) or active (enriched for H3K27ac). Despite having a clear practical utility, global analysis and chromatin mapping of mature eosinophils has been limited, predominantly due to difficulties in isolating intact RNA and chromatin from these sensitive, lysis-prone, enzyme rich granule containing cells. At present, there is only one published study investigating H3K4me3 marks within promoters of eosinophils and eosinophil progenitors (20). Thus, the use of bioinformatic and epigenetic tools to investigate eosinophil cell populations is under-represented in the literature.

Taking a global approach to identify the REs present in mature, bone marrow culture–derived murine eosinophils (bmEos), we performed chromatin immunoprecipitation coupled with massively parallel sequencing (ChIP-seq) for H3K27ac, H3K4me1 and H3K4me3, assay for transposase-accessible chromatin (ATAC-seq) and RNA sequencing (RNA-seq). Using these approaches, we defined a repertoire of eosinophil-specific enhancers that are located within 1-20 kB of genes expressed by mature eosinophils and then identified associated TFs. Throughout this manuscript, we use distance to define classify promoter and enhancers, with promoters located proximally (within 1 kB) and enhancers located distally (1 – 20 kB) to the TSS of expressed genes. We focused our study on enhancers, and excluded from analysis epigenetic marks detected within 1 kB of the TSS. Our analysis revealed the presence of numerous distal enhancers located predominantly within intronic regions. Gene ontology (GO) analysis revealed significant enrichment for genes associated with leukocyte migration, activation, chemotaxis, proliferation, degranulation and cell death, highlighting that important functional genes may be dynamically regulated by distal enhancers. DNA binding-site motifs for PU.1 were significantly enriched at enhancers that were predominantly near genes highly expressed in eosinophils, and PU.1 binding at these loci was further confirmed by ChIP-seq. Comparing our data with publicly available macrophage (21), neutrophil (22) and immature granulocyte (23) datasets revealed that >25% of the PU.1-bound regions were unique to eosinophils. Furthermore, eosinophil-specific super enhancers (regions of the genome containing multiple strong enhancers that have been shown to correlate with cell identity (24)) were enriched for genes associated with genome-wide association study (GWAS)–derived risk loci for eosinophilia and allergic diseases.

Collectively, our data identify eosinophil-specific enhancers likely regulating key eosinophil genes through epigenetic mechanisms (open chromatin, H3K4 mono- and tri-methylation and H3K27 acetylation) and TF binding (PU.1). These findings are supportive of a dynamic, epigenetic eosinophil regulome underlying eosinophil diversity. Future comprehensive analyses of eosinophil epigenetic and transcriptional profiling from homeostatic, developmental or disease states will further advance our understanding of allergic and eosinophil-related disorders and the potential development of therapeutic strategies.

Methods

Mice

Six- to eight-week-old BALB/c mice were used for all experiments. Animals were housed under standard specific pathogen-free conditions and handled under approved protocols of the Institutional Animal Care and Use Committee of Cincinnati Children's Hospital Medical Center (IACUC2018-0028, IACUC2016-0083).

Bone marrow–derived eosinophil culture and flow cytometry analysis

Murine bone marrow–derived eosinophils (bmEos) were generated from unselected bone marrow progenitor cells as first described by Dyer *et al.* (25, 26), and were used for all experiments. After 14 days, eosinophil maturity was assessed on the basis of cellular

morphology and Siglec-F (Sig-F)+CCR3+ positivity as determined by flow cytometry. Eosinophils were stained with SiglecF-AF647 (1:100, BD Biosciences) and CCR3-PE (1:100, R&D) for 25 min and Zombie NIR fixable viability (1:1000, Biolegend) for an additional 5 min. Cells were washed prior to analysis on FACSCanto (BD Biosciences). Cell cultures that were >95% SigF+CCR3+ were harvested and used for downstream experiments.

mRNA sequencing

Eosinophil RNA was isolated from 2-5x10^6 bmEos by double chloroform extraction prior to purification with the Direct-zol™ RNA MiniPrep (Zymo Research). RNA quality was assessed via Agilent RNA NanoCHIP, and only samples with an RNA quality number >8 were sent for sequencing analysis. Libraries were constructed with the TruSeq Stranded mRNA kit (Illumina) and sequenced on the Illumina NovaSeq 6000 (S1 Flow Cell) or Illumina HiSeq 2500, paired end, 100 bases for each end by the Cincinnati Children's Hospital Medical Center DNA Sequencing and Genotyping Core, Cincinnati, Ohio, targeting ~30 million reads per sample.

ATAC sequencing

A total of 6.5x10^4 viable bmEos were pelleted at 500 *g* for 5 minutes at 4°C prior to resuspension in 50 uL ATAC-Resuspension Buffer (ATAC-RSB; 10 mM Tris-HCl, 10 mM NaCl, 3 mM $MgCl_2$ made up to 50 mL in ddH_2O) with added 0.1% NP-40, 0.1% Tween-20 and 0.01% Digitonin. Samples were incubated on ice for 3 minutes prior to washing in 1 mL cold ATAC-RSB containing 0.1% Tween-20 only. Nuclei were pelleted at 500 *g* for 10 minutes at 4°C prior to addition of 50 uL transposition mixture (2X TD buffer, 100 nM transposase enzyme, 5 uL ddH2O, 1% Digitonin, 10% Tween-20 made up to 50 mL in 1X PBS) and incubated at 37°C for 30 minutes. After tagmentation, the transposased DNA was purified using the MiniElute Reaction Cleanup Kit (28204, Qiagen). Libraries were constructed with the OMNI-ATAC protocol (27) and sequenced on the Illumina HiSeq 4000, paired end, 150 bases each end, Novogene.

ChIP sequencing

A total of 1-3x10^7 viable bmEos were pelleted, and eosinophil chromatin was cross-linked by the addition of 1:10 of 10X crosslinking solution (8.8% formaldehyde, 0.1M NaCl, 1 mM EDTA, 1 mM EGTA, 50 mM HEPES made up to a total volume of 1 mL in ddH_2O) and incubated on ice for 5 minutes. Addition of 1:20 2.5M glycine was used to stop the cross-linking reaction. The cells were centrifuged at 2000 *g* for 5 minutes at 4°C, after which the pellets were washed twice with $PBS^{-/-}$ and frozen at −80°C until use. Following storage, the cell pellets were thawed on ice and re-suspended in 1 mL 'L1 lysis buffer' (50 mM HEPES, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP-40, 0.25% Triton X100 made up to a volume of 50 mL in ddH_2O) with added protease inhibitor (1:100) and NaButyrate (1:1000, stock 1M) and rotated at 4°C for 10 minutes. Cells were centrifuged at 1900 *g* for 3 minutes at 4°C prior to addition of 1 mL 'L2 buffer' (200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 10 mM Tris made up to a volume of 50 mL in ddH_2O) with added protease inhibitor (1:100) and NaButyrate (1:1000, stock 1M) and rotated at room temperature for 10 minutes. Cells were centrifuged at 2000 *g* for 5 minutes at 4°C to remove supernatant, followed

by two additional washes with Tris-EDTA (TE) + 0.1% SDS to remove all excess salts, which inhibit the sonication process. The pellet was re-suspended in 1 mL TE + 0.1% SDS and cooled on ice for 5 minutes prior to sonication using the Covaris S-series sonicator at 170.0 peak power, 10.0 duty factor and 200 burst. A small amount (10 uL) of the chromatin suspension was saved for agarose gel analysis to confirm adequate sonication of the sample, and salts were re-added (1% Triton X100, 5% glycerol, 0.5M NaCl, 0.1% NaDOC, 1:100 protease inhibitors, 1:1000 NaButyrate) to the remaining suspension. The chromatin was then pre-cleared following incubation with Protein-A beads prior to performing the ChIP with antibodies for H3K4me1 (C15410194, Diagenode), H3K4me3 (17-614, Millpore), H3K27ac (C15410196, Diagenode) and PU.1 (sc-352, SantaCruz) overnight in the SX-8G IP-Star® CHIP Robot (Diagenode). Libraries were constructed by ChIPmentation (28) and sequenced on the Illumina HiSeq 4000, single-end or paired-end, 150 bases per end, Novogene.

Functional genomics analysis

RNA-seq data were processed using the pipeline NextGenAligner (Mario Pujato (2019), https://github.com/MarioPujato/NextGenAligner). Briefly, the raw sequence read quality was analyzed using FastQC/0.11.7 (Simon Andrews, FastQC—a quality control application for FastQ files (2018)), https://github.com/s-andrews/FastQC). Adapter sequences were trimmed using CutAdapt/1.8.1 (29) and Trim Galore/0.4.2 (Felix Kreuger, TrimGalore: a wrapper for Cutadapt and FastQC (2015), https://github/FelixKreuger/TrimGalore), prior to alignment of sequence reads to the mm9 genome and RefSeq-based transcriptome using STAR/2.5 (30). Duplicates were removed using Picard/1.89 (Broad Institute, Picard Toolkit (2019), http://broadinstitute.github.io/picard). FeatureCounts (31) and the calculateFPKM module were used to quantify gene transcript fragments per kilobase per million mapped reads (FPKM). To characterize gene expression by eosinophils, thresholds were determined using expression levels of known expressed eosinophil genes, with ≥5 FPKM in both replicates indicating genes that are expressed by eosinophils and <5 FPKM indicating genes that are silent, repressed or minimally expressed. Gene ontology (GO) and genome-wide association study (GWAS) analysis was performed using Enrichr (32, 33). ATAC and ChIP sequencing analysis was performed as above except for read alignment, which was performed with BOWTIE/2.3.4.1 (34). Peaks were called using MACS/2.1.0 (35) with a q value of 0.01. Read and peak information for ATAC-seq and ChIP-seq samples are in Supplementary Table 1. Presence of ATAC-seq or ChIP-seq peaks within 20 kB of expressed eosinophil genes was determined using BEDtools/2.27.0 (36). ATAC-seq and ChIP-seq tracks were visualized using the Integrative Genomics Viewer (IGV) browser (37). Differential binding analysis of ChIP-seq and ATAC-seq peaks was performed with DiffBind/2.16.0 (38).

TF binding site and super enhancer analysis

TF binding motif enrichment analyses of identified REs were performed using the HOMER software package (39). Briefly, HOMER uses a library of >7,000 TF binding models (in the form of position weight matrices, taken from the CisBP database (40)) to scan a set of input sequences for statistical enrichment of each position weight matrix. Calculations are performed using ZOOPS (zero or one occurrence per sequence) scoring coupled

with hypergeometric enrichment analysis to determine motif enrichment. Input eosinophil enhancer sequences were also assessed for statistical enrichment of motifs for PU.1 binding sites using the findPeaks module and factor mode within HOMER. Significantly enriched TF binding site motifs are expressed as log P values. Super enhancer data were generated using the ROSE algorithm (24, 41) within Scientific Data Analysis Platform (https://scidap.com, Datirium) (42, 43).

Additional analyses

Data were analyzed with GraphPad Prism software (v8; GraphPad Software), with flow cytometric data analyzed using FlowJo software (TreeStar). A p value < 0.05 was considered as statistically significant.

Data availability

The next-generation sequencing data are accessible through NCBI's Gene Expression Omnibus (44) GSE144951 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE144951. The USCS Genome browser has been generated and is accessible at http://genome.ucsc.edu/s/feljj4/FELTON_mm9_bmEos.

Results

Chromatin accessibility and histone modifications associated with expressed eosinophil genes

To assess the chromatin landscape profile of mature, culture-derived, murine eosinophils (bmEos), open areas of chromatin and histone modifications for H3K4me1, H3K4me3 and H3K27ac were assessed across the genome. Prior to downstream analysis, eosinophils were evaluated for maturity, and >95% of cells were found to express CCR3 and Siglec-F (data not shown). Biological replicates for each of the four epigenetic marks showed a high correlation in peak location and strength as assessed by differential binding analysis of ATAC- and ChIP-seq peak data (DiffBind) output for correlation and replication of murine NGS data sets (Supp. Figure 1A–B) and per-peak read count (Supp. Figure 1C). Biological replicates were subsequently pooled for downstream analysis. Next we paired the chromatin accessibility and histone findings (ATAC-seq and ChIP-seq) with eosinophil transcriptomic data (RNA-seq).

Gene expression analysis revealed that 6,288 genes were expressed (≥5 FPKM) by mature culture-derived bmEos. These genes included well-described marker genes for eosinophils, such as major basic protein (MBP, *Prg2*), eosinophil cationic protein (ECP, *Ear1*), eosinophil peroxidase (EPX, *Epx*), CCL11 (eotaxin) receptor (CCR3, *Ccr3*), IL-5Rα (*Il5ra*), Siglec-F (*Siglec5*) and the TFs PU.1 (*Spi1*) and GATA-1 (*Gata1*) and the Ikaros TF family members Aiolos (*Ikzf3*) and Helios (*Ikzf2*). In contrast, >18,000 genes were considered to be silent, repressed or minimally expressed (<5 FKPM, data not shown). Expressed eosinophil genes had higher chromatin accessibility and H3K4me1, H3K4me3 and H3K27ac levels in comparison to the silent, repressed or minimally expressed genes (Supp. Figure 1D). Expressed genes were associated with the classical 'see-saw' promoter profile (open chromatin, H3K27ac marks and H3K4me3 marks surrounded by H3K4me1) within 1 kB

of the TSS (45), whereas the silent, repressed or minimally expressed genes did not present with this profile (Supp. Figure 1E).

Diverse promoter structure across eosinophil genes

Active promoters are traditionally identified by open regions of chromatin (ATAC) co-marked with H3K4me3 and H3K27ac (Figure 1A, **top**). Interestingly, when assessing the proportion of eosinophil genes found to possess active promoters within 1 kB of the TSS, we observed that there was an almost even split of active promoters lacking or possessing H3K4me1 at the H3K4me3 peak location (Figure 1A, **bottom**). Of the 6,288 expressed eosinophil genes, 2,542 (40.43%) were found lacking H3K4me1 at the peak location (H3K4me1-), whereas 2,776 (44.15%) were found to possess H3K4me1 (H3K4me1+) (Figure 1B). When we compared the expression level of eosinophil genes associated with an H3K4me1+ or H3K4me1− active promoter, it was observed that genes with an H3K4me1− active promoter were expressed to a higher level than those with an H3K4me1+ active promoter ($p < 0.0001$) (Figure 1B). GO analysis of the 2,542 genes with an H3K4me1− active promoter revealed significant enrichment of genes associated with leukocyte-mediated immunity, migration, activation and degranulation (Figure 1C). In contrast, GO analysis of the 2,776 genes with an H3K4me1+ active promoter revealed significant enrichment for genes associated with mRNA processing, gene expression and translation (Figure 1D). Interestingly, 970 (15.4%) of expressed eosinophil genes did not possess an H3K27ac-marked promoter within 1 kB of the TSS.

Active chromatin regions in highly expressed eosinophil genes show enrichment for PU.1 motifs

Next we interrogated the regulation of key eosinophil genes by comparing the prevalence of active REs, identified by the presence of H3K27ac in open chromatin (REs lacking H3K27ac in open chromatin are classified as ‘poised’ and were not included in this analysis) and TF binding motifs between the top 10% (653 genes, n = 2, average FPKM range 86.5–4870.8) and bottom 10% (653 genes, n = 2, average FPKM range 5–6.9) of expressed eosinophil genes. We hypothesized that highly expressed eosinophil genes would be enriched for H3K27ac relative to lowly expressed eosinophil genes. As expected, the top 10% of expressed eosinophil genes (Supp. Figure 2A**; blue line**) had an increase in H3K27ac signal compared to that of the bottom 10% of expressed eosinophil genes (Supp. Figure 2A**; green line**). Analysis of TF binding site motifs revealed preferential enrichment for PU.1 binding motifs in the top 10% of expressed eosinophil genes compared to the bottom 10% of expressed genes (Supp. Figure 2B **and** Supp. Figure 2C). Highly expressed eosinophil genes were also preferentially enriched for YY1, FOS, JUN and CEBP family members, albeit to a much lower extent than PU.1 (Supp. Figure 2B **and** Supp. Figure 2C). Gene expression analysis revealed PU.1 to be the second most highly expressed TF in bmEos (Supp. Figure 2D), further supporting its importance in the eosinophil regulome.

To determine whether the proximal and/or distal REs were driving the PU.1 epigenetic signature, we examined DNA sequences within active (H3K37ac+) chromatin regions on the basis of genomic location; promoter regions (+/−1 kB from TSS) and enhancers (1 – 20 kB from TSS) were identified by H3K27ac density to assess TF binding site motif enrichment.

Interestingly, HOMER analysis revealed that the DNA binding site motifs for PU.1 were the most enriched within the distal (1 – 20 kB from TSS) active chromatin regions identified as enhancers (Supp. Figure 2E) compared to the promoters (+/−1 kB of the TSS) (Supp. Figure 2F). Together, these data suggest that regulation of key eosinophil genes occurs by a select repertoire of TFs, including PU.1, at epigenetically regulated distal enhancers.

Active eosinophil enhancers are associated with highly expressed genes

Enhancers are described as discrete regions of DNA that increase transcriptional activity of promoters from a distance, often many kB away, in part via binding of lineage-specific TFs (46, 47). Due to 1) the key role of PU.1 in granulocyte maturation and function and 2) the observed enrichment of PU.1 in active enhancers distal (1 – 20 kB) to the TSS, we aimed to further characterize the chromatin modification profiles at these locations. Although the classical description of enhancers has an inverted profile of H3K4me1 and H3K4me3 (45), there are some non-canonical examples of H3K4me3 depositing at a subset of enhancers acting in lineage-specific functions (47). Considering this, we speculated that highly expressed genes associated with key eosinophil functional responses would have active enhancers with greater H3K4me3 deposition.

Notably, over 83% (5,243/6,288) of expressed eosinophil genes had at least one type of enhancer (active, poised, inactive) within 1 – 20 kB of its TSS; genes associated with active enhancers were expressed at a higher level than those associated with poised enhancers (Figure 2A). Furthermore, we observed that a greater proportion of active enhancers than poised enhancers were marked by H3K4me3 deposition (Figure 2B). Poised enhancers were identified by regions of open chromatin co-marked with H3K4me1 (with or without H3K4me3) and lacking H3K27ac (Figure 2C, **top**). Active enhancers were identified by regions of open chromatin co-marked with H3K4me1 (with or without H3K4me3) and possessing H3K27ac (Figure 2C, **bottom**). Active enhancers were enriched for gene pathways associated with granulocyte-mediated immunity, migration and phagocytosis (Figure 2D) to a greater degree than were poised enhancers (Figure 2E). Key functional genes with at least one active distal eosinophil enhancer located within 1 - 20 kB included *Il5ra*, *Gata2, Epx*, *Ccr3, Pgr2, Mcl1*, *Ikzf1, Ikzf2, Ikzf3, Irf2* and *Stat5b*. These data identify epigenetically marked REs consistent with active distal enhancers in the eosinophil regulome and are suggestive of their involvement in key lineage-specific eosinophil functions.

Active distal enhancers show enrichment for PU.1 DNA binding site motifs

To elucidate which TFs are involved in the regulation of the eosinophil enhancer epigenome, we examined DNA sequences within active enhancers for TF binding site motif enrichment. HOMER analysis revealed that the DNA binding site motifs for PU.1 were the most enriched (Figure 2F **and** Supp. Figure 3A), with both the motif score (Figure 2G, **left**; $p = 0.0014$) and proportion of predicted binding sites (Figure 2G, **right**; $p < 0.001$, chi-squared test) being greater in active than poised enhancers. Other highly represented DNA binding site motifs included those for ELK:ELF family members, CEBP isoforms, IRFs and GATA. *De novo* analysis revealed enrichment for DNA binding site motifs for KLF family members, AP1, ELK:ELF family members and SMAD3 (Supp. Figure 3B).

PU.1-bound eosinophil active enhancers

We paired PU.1 binding sites (identified by ChIP-seq for PU.1) with enhancer data to determine the locations at which PU.1 is bound to enhancers across the eosinophil epigenome. The PU.1 motif was enriched in PU.1 ChIP-seq peaks compared to H3K27ac peaks alone or background (Figure 3A). Over 95% (96.7%, 7970/8228) of active enhancers were bound by PU.1, in contrast to only 61.3% (3586/5848) of poised enhancers (Figure 3B; $p < 0.001$, chi-squared test). To determine which TF may be facilitating PU.1 binding at the 7.2% of PU.1 peaks that did not possess a PU.1 or similar binding motifs, we performed *de novo* TF binding motif enrichment. First, all peaks with a GGAA within their genomic sequence were removed to ensure that ETS/ELK/ELF family TFs were not included due to the similarities in their core binding sequences. Interestingly, *de novo* motif enrichment revealed that FOXK1, members of the KLF family, STAT3, STAT5, SMAD3 and GATA1 were the most enriched *de novo* motifs (Figure 3C), with a strong enrichment in the target sequences compared to background (Supplementary Figure 3C). Of the total active enhancers with predicted PU.1 binding motifs identified, 83.2% (5374/6461) were confirmed by ChIP-seq (data not shown), including the active enhancer identified within the *Il5ra* gene (Figure 3D).

Eosinophil-specific enhancers are regulated by PU.1 and are enriched for pathways involved in degranulation, activation and regulation of granulocyte-mediated immunity

To determine which of the identified PU.1-bound active enhancers were unique to murine culture-derived eosinophils, we compared our datasets with publicly available data (H3K27ac, PU.1 ChIP-seq and RNA-seq) from other myeloid cell types: bone marrow–derived macrophages (mBMDM: SRR502912, SRR502934 and RNA-seq SRR502948 (21)), neutrophils (Neuts: SRR6751044, SRR6751056, and RNA-seq SRR6751062 (22)) and immature granulocytes (ImmGrans: SRR611737, SRR611738 (23)). No paired RNA-seq data were available for ImmGrans. Due to differences in mouse strain/background, we generated bmEos ATAC and H3K27ac datasets from mice on a BL6/C57 background to assess the correlation and overlap of called peaks between background strains. A strong correlation and overlap of peak locations was observed, with the most influential variable being epigenetic mark/assay type (PC1 = 61%; ATAC or H3K27ac) over mouse strain (PC2 = 14%; BL6/C57 or BALB/c) (Supp. Figure 3D–E), with 71.3-82.3% of ATAC and H3K27ac ChIP-seq peaks shared between BALB/c and BL6/C57 datasets (data not shown). Furthermore, the strongest correlation and clustering between PU.1 ChIP-seq datasets were within PU.1 myeloid datasets (Supp. Figure 3F).

Differential binding and occupancy analysis (DiffBind) of active PU.1-bound bmEos enhancers compared to active PU.1 co-bound regions in mBMDMs, Neuts and ImmGrans revealed 1,691 PU.1-bound active enhancers that were unique to eosinophils (26%) (Figure 4A). Eosinophils had ~70% of their PU.1-bound enhancers in common with at least one of the other myeloid cell types. A total of 1,445 PU.1-bound enhancers were common between all cells of the myeloid lineage assessed, describing a shared myeloid PU.1 regulome. The 1,691 unique, active PU.1-bound, eosinophil-specific enhancers were associated with a total of 1,300 genes with increased expression (FPKM) in bmEos (compared to mBMDMs ($p = 0.0257$) and Neuts ($p = 0.0219$) (Figure 4B), and GO analysis revealed enrichment for

gene pathways associated with myeloid differentiation, granulocyte activation, degranulation and migration (Figure 4C), including but not limited to *Prg2*, *Epx* (Figure 4D), *Ikzf1*, *Ikzf2*, *Ikf3* (Supplementary Figure 4A), *Cebpa*, *Cebpb*, *Cebpe* (Supplementary Figure 4B), *Trib1* (Figure 5D), *Smad3*, *Cd44*, *Cd9*, *Ccl9*, *Ccl17*, *Il5ra* (Supplementary Figure 4C), *Ear6*, *Grb7*, *Klf13*, *Nfkb1*, *Elmo1* and *Gata1* (Supplementary Figure 4D).

Together, these data demonstrate that relevant, functionally expressed genes are associated with PU.1-bound, lineage-specific and cell-specific enhancers.

Eosinophil-specific super enhancers direct cell identity and are enriched for GWAS-associated genes

Lastly, we used the H3K27ac datasets to identify super enhancers, defined as regions of the genome containing multiple strong enhancers that have been shown to correlate with cell identity (24). To this end, we identified eosinophil super enhancers using the ROSE algorithm (24). Eosinophil super enhancers were associated with key genes involved in maturation and function, including *Trib1*, *Ikzf2, Spi1* (PU.1), *Il4, Ikzf1, Prg2* (proMBP) and *Cebpe* (Figure 5A; **top left**). Super enhancers for mBMDMs, Neuts and ImmGrans were also identified using the ROSE algorithm (Figure 5A). DiffBind analysis of super enhancers from each cell type—bmEos, Neuts, mBMDMs, and ImmGrans—revealed a total of 643 eosinophil-specific super enhancers (Figure 5B). Interestingly, eosinophils had the greatest proportion of cell type–specific super enhancers out of the four populations (Figure 5C; $p < 0.001$, chi-squared test). Genes associated with eosinophil-specific super enhancers included but were not limited to *Ccr3*, *Ikzf1, Ikzf2, Trib1* (Figure 4E), *Cebpe* (Supplementary Figure 4B), *Smad3* and *Prg2.* Furthermore, murine, eosinophil-specific super enhancers were enriched in gene whose homologues are known to be associated with GWAS hits for eosinophil levels (eosinophilia, percentage of white blood cells or granulocytes) and allergic diseases in humans (Figure 5D).

Together, these data demonstrate that relevant, functionally expressed genes are associated with lineage-specific and cell type–specific super enhancers, which are enriched in genes associated with allergic disease risk loci.

Discussion

It is becoming increasingly appreciated that eosinophils are pleiotropic, multi-functional cells involved in a plethora of biological processes, suggesting the presence of complex cell autonomous pathways that determine eosinophil gene expression and function. As the genomic regulatory landscape of eosinophils was previously unknown, we took a global approach to analyze genome-wide transcriptome and epigenome of mature, murine, culture-derived bmEos in order to describe the cell's enhancer profiles. Our aim was to identify and define the characteristics of unique, eosinophil-specific enhancers controlling expression of key genes, in addition to identifying TFs crucial for regulating eosinophil gene expression at the identified enhancers.

Active enhancers likely have roles in key eosinophil pathways and lineage determination. Highly expressed eosinophil genes were enriched for active (H3K27ac+) distal (1-20 kB

of TSS) enhancers, particularly in genes associated with migration, activation, proliferation and degranulation, suggesting that these key pathways may be regulated by enhancers. Furthermore, deposition of H3K4me3 was observed at a greater portion in active enhancers, suggesting that these could be involved in regulating lineage-defining functions (47). Russ *et al.* (47) observed a greater proportion of H3K4me3 deposition at predominantly poised enhancers (H3K27ac-) in naïve T cells, yet we observed a greater proportion of H3K4me3 deposition in active (H3K27ac+) enhancers. When considering the terminally differentiated state of eosinophils in contrast to naïve T cells, we speculate that the lineage-defining developmental processes regulated by these enhancer regions were activated at an earlier point in eosinophil development. Future studies will focus on delineating the epigenetic changes that occur during eosinophilopoiesis in addition to post-maturation activation using a range of cytokines and inflammatory mediators.

Correspondingly, the observed increase in expression of genes possessing an H3K4me1- active promoter, compared to those possessing an H3K4me1+ active promoter, further highlights the functional specificity of RE subtypes. These findings suggest that traditional (H3K4me1-) active promoters regulate highly expressed genes enriched for cell-specific functions (e.g., granulocyte activation and degranulation), whereas non-traditional (H3K4me1+) active promoters regulate lowly expressed genes involved in general cellular functions (e.g., regulation of translation and mRNA processing). Furthermore, the differential enrichment of TF binding motifs demonstrated between proximal promoters (+/−1 kB of the TSS) and distal enhancers (1 – 20 kB from TSS) demonstrates that select TF are involved in these discrete processes.

TFs have a major role in determining cell maturation and development in both cell type– and lineage-specific contexts. The most prominent of those involved in eosinophil maturation and lineage commitment are PU.1, GATA-1 and members of the CCAAT/enhancer binding protein (CEBP) family (10, 11, 20, 48, 49). Our data are consistent with these findings, showing high expression levels for all of these TFs, in addition to enrichment for their binding motifs within identified eosinophil enhancers. Interestingly, previous data published by our group showed enrichment for binding site motifs of members of the IKZF TF family (IKZF2, IKZF3) in promoters of eosinophil progenitors and FACS-sorted, bone marrow–resident eosinophils (20). Despite elevated transcript levels of these TFs and their association with both PU.1-bound eosinophil-specific enhancers and super enhancers, enrichment for their binding motifs in distal enhancers of cultured eosinophils was relatively low. This potentially suggests that 1) IKZF family members regulate eosinophil gene expression at an earlier time point in eosinophil development and maturation and may not have as prominent a role in mature, cultured cells under the direction of IL-5 or 2) IKZF family members have a more prominent role in the regulation of gene transcription in eosinophils at promoters (proximal REs). Similarly, the prevalence of PU.1 binding at enhancers in mature eosinophils supports the hypothesis that PU.1 is not essential for progenitor specification but instead has an important role in the terminal differentiation of mature eosinophils (50–54).

The roles of less prominent TFs in determining cell maturation and development in both cell type– and lineage-specific contexts are not to be discounted. Interestingly, *de novo*

motif analysis of active enhancers revealed enrichment for TFs including AP-1, KLF family members, STATs, SMAD3 and GATA. These same elements were also enriched in the subpopulation of PU.1-bound active enhancers without a canonical DNA binding motif for PU.1 within their sequence. Notably, AP-1–dependent elements have substantial overlap with risk loci for multiple immune diseases, including allergic disease (55). In addition, the activity of the *IL5Ra* P1 promoter has been shown to be mediated by AP-1 binding (56), with the consensus AP-1 binding motif present adjacent to the EOS1 enhancer-like element in the human *IL5Ra* promoter (57). KLF family members have been shown to interact with PU.1, CEBP and GATA to facilitate myeloid and monocyte differentiation (58–60), in addition to regulating eosinophil function in adipose tissue (61). Furthermore, SMAD3 has known functional roles in regulation of TGF-β signaling (62), and SMAD3-deficient mice exhibit attenuated pathology in a model of experimental eosinophilic esophagitis (EoE) (63). The attenuated pathology comprises reduced levels of esophageal remodeling, fibrosis and angiogenesis, despite no reduction in eosinophil number (63), suggesting that SMAD3 may be more involved in eosinophil functional responses than eosinophilopoiesis. Our discovery of an eosinophil-specific super enhancer associated with SMAD3 substantiates a key potential role of SMAD3 in regulating gene expression in eosinophils; thus, modifying interactions between SMAD3 and active enhancers may be a novel strategy to reduce eosinophil-associated tissue pathology, including esophageal fibrosis and its associated complications, in EoE (63). Collectively, these data highlight the regulation of cell type– and maturation-specific changes in the TF repertoire at eosinophil-specific REs.

In an effort to define the eosinophil-specific epigenome, we compared our findings with publicly available data describing PU.1-bound enhancers in macrophages (21), neutrophils (22) and immature granulocytes (23). Differential binding analysis revealed that >25% of the identified PU.1-bound eosinophil active enhancers were unique to eosinophils and thus were described as PU.1-bound, eosinophil-specific enhancers; these enhancers were enriched for pathways involved in activation, degranulation and granulocyte-mediated immunity. PU.1-bound, common myeloid–lineage enhancers were also identified. Our discovery of both a sizeable common myeloid and a unique, eosinophil-specific PU.1-bound enhancer repertoire reinforces the importance of PU.1 in myeloid gene expression and development (10, 11, 20, 48, 51–54) and supports previous studies using PU.1-deficient mice; these studies showed an absence of *Epx*, *Il5ra* and *Prg2* (MBP) mRNA expression in fetal liver cells from embryonic lethal PU.1$^{-/-}$ mice and a total loss of the eosinophil population within the bone marrow and spleen of viable PU.1$^{-/-}$ mice (50). Further studies utilizing gene-editing techniques, including CRISPR/Cas9, will allow investigation of the consequence of loss of PU.1-bound, eosinophil-specific enhancers in eosinophil development and functional responses. We are currently in the process of generating an *Il5ra* enhancer deletion mouse to interrogate the impact of the loss of this PU.1-bound, eosinophil-specific enhancer on eosinophil function and development.

One limitation encountered routinely when utilizing gene set enrichment analysis is the limited set of genes associated with GO terms outlining eosinophil functional responses; currently, there are only 38 genes associated with the eosinophil-related gene lists available in the GO database (64, 65), which may be due to the challenges in isolating RNA from these sensitive, lysis-prone cells. In contrast, there are 199 genes associated with

neutrophil-related gene lists. Enrichment for gene pathways involved in ‘neutrophil’ degranulation, activation and immunity for genes associated with PU.1-bound eosinophil-specific enhancers emphasizes the gap in the literature describing and outlining the genes involved in eosinophil functional responses. Although these PU.1-bound, eosinophil-specific enhancers were absent in neutrophils, neutrophil-related terms were the most enriched due to limited GO terms outlining relevant degranulation, activation and functional responses in eosinophils. Although we included these neutrophil-related terms in the analysis, we aim to utilize the current dataset to contribute to the GO consortium to develop terms relating specifically to eosinophil maturation and functional responses for more accurate distinction of these granulocyte populations within the literature and for the progression of the field of eosinophil-related research.

Finally, the description of the eosinophil-specific super enhancer landscape provides exciting opportunities for investigation into the role of regulator regions associated with driving cell identity (24). Parallel recent findings have identified the importance of *Trib1* in the terminal differentiation of myeloid populations; *Trib1* regulates both granulocyte precursor lineage commitment and mature eosinophil identity (66). From these findings, we speculate that ongoing studies investigating the tissue- and inflammatory-specific changes in the chromatin landscape will uncover novel key regulators of eosinophil tissue specificity and function, as well as elucidate potential approaches for treating allergic disorders.

Collectively, these novel data describe the epigenetic landscape of eosinophil-specific enhancers in mature, murine, cultured eosinophils, providing an initial step towards an understanding of the eosinophil chromatin landscape and its role in the eosinophil regulome. Future studies investigating the tissue- and inflammatory-specific changes in the chromatin landscape that drive eosinophil maturation, recruitment, tissue specificity and function will elucidate potential opportunities and approaches for treating allergic disorders, such as manipulation of enhancer-regulated genes for therapeutic benefit.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Shawna Hottinger for her editorial assistance, Dr. Mark Rochman for thoughtful contributions and Dr. Mario Pujato for continued assistance and development of bioinformatic pipelines. RNA sequencing was performed in collaboration with the Cincinnati Children’s DNA Sequencing and Genotyping Core, Cincinnati, Ohio.

This work was supported by the National Institutes of Health through the National Institute of Allergy and Infectious Diseases (NIAID) grant R01 AI130133.

Disclosure/Conflict of Interests

J.M.F., S.V., L.E.E., S.P., B.W., K.E., M.K., A.M. and M.T.W. have no conflicts of interest to disclose. M.E.R. is a consultant for Pulm One, Spoon Guru, ClostraBio, Celgene, Astra Zeneca and Arena Pharmaceuticals and has an equity interest in the first three listed and royalties from reslizumab (Teva Pharmaceuticals), PEESSv2 (Mapi Research Trust) and UpToDate. M.E.R. is an inventor of patents owned by Cincinnati Children’s. A.B. is a co-founder of Datirium, LLC. P.C.F. has received research funding from Knopp Biosciences.

Abbreviations

ATAC-seq	Assay for transposase-accessible chromatin (ATAC) with massively parallel sequencing
ddH_2O	Double-distilled water
bmEos	Bone marrow–derived eosinophils
ChIP-seq	Chromatin immunoprecipitation (ChIP) with massively parallel sequencing
DMEM	Dulbecco's modified Eagle medium
EDTA	Ethylenediaminetetraacetic acid
EGTA	Ethylene glycol-bis(β-aminoethyl ether)-N,N,N′,N′-tetraacetic acid
FPKM	Fragments per kilobase of transcript per million mapped reads
FSC	Forward scatter
FLT3	FMS-like receptor tyrosine kinase 3
FLT3-L	FLT3-ligand
GO	Gene ontology
GWAS	Genome-wide association studies
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
H3K4me1	Histone H3 lysine 4 monomethylation
H3K4me3	Histone H3 lysine 4 trimethylation
H3K27ac	Histone H3 lysine 27 acetylation
IGV	Integrative Genomics Viewer
IMDM	Iscove's modified Dulbecco medium
kB	Kilobase
PE	Paired end
RE	Regulatory element
RBC	Red blood cell
SCF	Stem cell factor
SigF / Siglec-F	Sialic acid–binding Ig-like lectin F

SSC	Side scatter
TE	Tris-EDTA
TF	Transcription factor
TES	Transcription end site
TSS	Transcriptional start site
ZOOPS	Zero or one occurrence per sequence

References

1. Felton JM, Lucas CD, Rossi AG, and Dransfield I. 2014. Eosinophils in the Lung – Modulating Apoptosis and Efferocytosis in Airway Inflammation. Front. Immunol. 5: 302. [PubMed: 25071763]
2. Mesnil C, Raulier S, Paulissen G, Xiao X, Birrell MA, Pirottin D, Janss T, Starkl P, Ramery E, Henket M, Schleich FN, Radermecker M, Thielemans K, Gillet L, Thiry M, Belvisi MG, Louis R, Desmet C, Marichal T, and Bureau F. 2016. Lung-resident eosinophils represent a distinct regulatory eosinophil subset. J. Clin. Invest. 126: 3279–3295. [PubMed: 27548519]
3. Marichal T, Mesnil C, and Bureau F. 2017. Homeostatic Eosinophils: Characteristics and Functions. Front. Med. 4: 101.
4. Shi H-Z2004. Eosinophils function as antigen-presenting cells. J. Leukoc. Biol. 76: 520–527. [PubMed: 15218055]
5. Kita H1996. The eosinophil: A cytokine-producing cell?J. Allergy Clin. Immunol. 97: 889–892. [PubMed: 8655882]
6. Rothenberg ME, and Hogan SP. 2006. The Eosinophil. Annu. Rev. Immunol. 24: 147–174. [PubMed: 16551246]
7. Lee SD, and Tontonoz P. 2014. Eosinophils in Fat: Pink Is the New Brown. Cell157: 1249–1250. [PubMed: 24906141]
8. Wu D, Molofsky AB, Liang H-E, Ricardo-Gonzalez RR, Jouihan HA, Bando JK, Chawla A, and Locksley RM. 2011. Eosinophils sustain adipose alternatively activated macrophages associated with glucose homeostasis. Science332: 243–247. [PubMed: 21436399]
9. Becher B, Schlitzer A, Chen J, Mair F, Sumatoh HR, Teng KWW, Low D, Ruedl C, Riccardi-Castagnoli P, Poidinger M, Greter M, Ginhoux F, and Newell EW. 2014. High-dimensional analysis of the murine myeloid cell system. Nat. Immunol. 15: 1181–1189. [PubMed: 25306126]
10. Du J, Stankiewicz MJ, Liu Y, Xi Q, Schmitz JE, Lekstrom-Himes JA, and Ackerman SJ. 2002. Novel combinatorial interactions of GATA-1, PU.1, and C/EBPepsilon isoforms regulate transcription of the gene encoding eosinophil granule major basic protein. J. Biol. Chem. 277: 43481–43494. [PubMed: 12202480]
11. Gombart AF, Kwok SH, Anderson KL, Yamaguchi Y, Torbett BE, and Koeffler HP. 2003. Regulation of neutrophil and eosinophil secondary granule gene expression by transcription factors C/EBP epsilon and PU.1. Blood101: 3265–3273. [PubMed: 12515729]
12. Lekstrom-Himes JA2001. The role of C/EBP(epsilon) in the terminal stages of granulocyte differentiation. Stem Cells19: 125–133. [PubMed: 11239167]
13. Wamstad JA, Wang X, Demuren OO, and Boyer LA. 2014. Distal enhancers: new insights into heart development and disease. Trends Cell Biol. 24: 294–302. [PubMed: 24321408]
14. Link VM, Gosselin D, and Glass CK. 2015. Mechanisms Underlying the Selection and Function of Macrophage-Specific Enhancers. Cold Spring Harb. Symp. Quant. Biol. 80: 213–221. [PubMed: 26582787]
15. Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, Boyer LA, Young RA, and Jaenisch R. 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc. Natl. Acad. Sci. 107: 21931–21936. [PubMed: 21106759]

16. Heinz S, Romanoski CE, Benner C, and Glass CK. 2015. The selection and function of cell type-specific enhancers. Nat. Rev. Mol. Cell Biol. 16: 144–154. [PubMed: 25650801]

17. Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, Van Calcar S, Qu C, Ching KA, Wang W, Weng Z, Green RD, Crawford GE, and Ren B. 2007. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat. Genet. 39: 311–318. [PubMed: 17277777]

18. Wang Z, Zang C, Rosenfeld JA, Schones DE, Barski A, Cuddapah S, Cui K, Roh T-Y, Peng W, Zhang MQ, and Zhao K. 2008. Combinatorial patterns of histone acetylations and methylations in the human genome. Nat. Genet. 40: 897–903. [PubMed: 18552846]

19. Ernst J, Kheradpour P, Mikkelsen TS, Shoresh N, Ward LD, Epstein CB, Zhang X, Wang L, Issner R, Coyne M, Ku M, Durham T, Kellis M, and Bernstein BE. 2011. Systematic analysis of chromatin state dynamics in nine human cell types. Nature473: 43–49. [PubMed: 21441907]

20. Bouffi C, Kartashov AV, Schollaert KL, Chen X, Bacon WC, Weirauch MT, Barski A, and Fulkerson PC. 2015. Transcription Factor Repertoire of Homeostatic Eosinophilopoiesis. J. Immunol. 195: 2683–2695. [PubMed: 26268651]

21. Ostuni R, Piccolo V, Barozzi I, Polletti S, Termanini A, Bonifacio S, Curina A, Prosperini E, Ghisletti S, and Natoli G. 2013. Latent Enhancers Activated by Stimulation in Differentiated Cells. Cell152: 157–171. [PubMed: 23332752]

22. Fischer J, Walter C, Tönges A, Aleth H, Jordão MJC, Leddin M, Gröning V, Erdmann T, Lenz G, Roth J, Vogl T, Prinz M, Dugas M, Jacobsen ID, and Rosenbauer F. 2019. Safeguard function of PU.1 shapes the inflammatory epigenome of neutrophils. Nat. Immunol. 20: 546–558. [PubMed: 30911105]

23. Luyten A, Zang C, Liu XS, and Shivdasani RA. 2014. Active enhancers are delineated de novo during hematopoiesis, with limited lineage fidelity among specified primary blood cells. Genes Dev. 28: 1827–1839. [PubMed: 25128499]

24. Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, Rahl PB, Lee TI, and Young RA. 2013. Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. Cell153: 307–319. [PubMed: 23582322]

25. Dyer KD, Moser JM, Czapiga M, Siegel SJ, Percopo CM, and Rosenberg HF. 2008. Functionally Competent Eosinophils Differentiated Ex Vivo in High Purity from Normal Mouse Bone Marrow. J. Immunol. 181: 4004–4009. [PubMed: 18768855]

26. Dyer KD, Czapiga M, Foster B, Foster PS, Kang EM, Lappas CM, Moser JM, Naumann N, Percopo CM, Siegel SJ, Swartz JM, Ting-De Ravin S, and Rosenberg HF. 2007. Eosinophils from Lineage-Ablated dblGATA Bone Marrow Progenitors: The dblGATA Enhancer in the Promoter of GATA-1 Is Not Essential for Differentiation Ex Vivo. J. Immunol. 179: 1693–1699. [PubMed: 17641035]

27. Corces MR, Trevino AE, Hamilton EG, Greenside PG, Sinnott-Armstrong NA, Vesuna S, Satpathy AT, Rubin AJ, Montine KS, Wu B, Kathiria A, Cho SW, Mumbach MR, Carter AC, Kasowski M, Orloff LA, Risca VI, Kundaje A, Khavari PA, Montine TJ, Greenleaf WJ, and Chang HY. 2017. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nat. Methods14: 959–962. [PubMed: 28846090]

28. Schmidl C, Rendeiro AF, Sheffield NC, and Bock C. 2015. ChIPmentation: fast, robust, low-input ChIP-seq for histones and transcription factors. Nat. Methods12: 963–965. [PubMed: 26280331]

29. Martin M2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal17: 10–12.

30. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. Bioinforma. 29: 15–21.

31. Liao Y, Smyth GK, and Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics30: 923–930. [PubMed: 24227677]

32. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, Clark NR, and Ma'ayan A. 2013. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics14: 128. [PubMed: 23586463]

33. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A, McDermott MG, Monteiro CD, Gundersen GW, and Ma'ayan A.

2016. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. 44: W90–97. [PubMed: 27141961]

34. Langmead B, and Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods9: 357–359. [PubMed: 22388286]
35. Gaspar JM2018. Improved peak-calling with MACS2. bioRxiv496521.
36. Quinlan AR, and Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics26: 841–842. [PubMed: 20110278]
37. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, and Mesirov JP. 2011. Integrative genomics viewer. Nat. Biotechnol. 29: 24–26. [PubMed: 21221095]
38. Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ, Brown GD, Gojis O, Ellis IO, Green AR, Ali S, Chin S-F, Palmieri C, Caldas C, and Carroll JS. 2012. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature481: 389–393. [PubMed: 22217937]
39. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, and Glass CK. 2010. Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. Mol. Cell38: 576–589. [PubMed: 20513432]
40. Lambert SA, Yang AWH, Sasse A, Cowley G, Albu M, Caddick MX, Morris QD, Weirauch MT, and Hughes TR. 2019. Similarity regression predicts evolution of transcription factor sequence specificity. Nat. Genet. 51: 981–989. [PubMed: 31133749]
41. Lovén J, Hoke HA, Lin CY, Lau A, Orlando DA, Vakoc CR, Bradner JE, Lee TI, and Young RA. 2013. Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell153: 320–334. [PubMed: 23582323]
42. Kartashov AV, and Barski A. 2015. BioWardrobe: an integrated platform for analysis of epigenomics and transcriptomics data. Genome Biol. 16: 158. [PubMed: 26248465]
43. Kotliar M, Kartashov AV, and Barski A. 2019. CWL-Airflow: a lightweight pipeline manager supporting Common Workflow Language. GigaScience8(7): giz084. [PubMed: 31321430]
44. Edgar R, Domrachev M, and Lash AE. 2002. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 30: 207–210. [PubMed: 11752295]
45. Sharifi-Zarchi A, Gerovska D, Adachi K, Totonchi M, Pezeshk H, Taft RJ, Schöler HR, Chitsaz H, Sadeghi M, Baharvand H, and Araúzo-Bravo MJ. 2017. DNA methylation regulates discrimination of enhancers from promoters through a H3K4me1-H3K4me3 seesaw mechanism. BMC Genomics18: 964. [PubMed: 29233090]
46. Banerji J, Rusconi S, and Schaffner W. 1981. Expression of a β-globin gene is enhanced by remote SV40 DNA sequences. Cell27: 299–308. [PubMed: 6277502]
47. Russ BE, Olshansky M, Li J, Nguyen MLT, Gearing LJ, Nguyen THO, Olson MR, McQuilton HA, Nüssing S, Khoury G, Purcell DFJ, Hertzog PJ, Rao S, and Turner SJ. 2017. Regulation of H3K4me3 at Transcriptional Enhancers Characterizes Acquisition of Virus-Specific CD8+ T Cell-Lineage-Specific Function. Cell Rep. 21: 3624–3636. [PubMed: 29262339]
48. Fulkerson PC2017. Transcription Factors in Eosinophil Development and As Therapeutic Targets. Front. Med. 4.
49. Uhm TG, Kim BS, and Chung IY. 2012. Eosinophil Development, Regulation of Eosinophil-Specific Genes, and Role of Eosinophils in the Pathogenesis of Asthma. Allergy Asthma Immunol. Res. 4: 68–79. [PubMed: 22379601]
50. Ackerman SJ, Du J, Xin F, Dekoter R, Mckercher S, Maki R, Singh H, and Yamaguchi Y. 2000. Eosinophilopoiesis: To be or not to be (an eosinophil)? That is the question: transcriptional mechanisms regulating eosinophil genes and development. Respir. Med. 94: 1135–1138.
51. Anderson KL, Smith KA, Conners K, McKercher SR, Maki RA, and Torbett BE. 1998. Myeloid development is selectively disrupted in PU.1 null mice. Blood91: 3702–3710. [PubMed: 9573007]
52. Scott EW, Fisher RC, Olson MC, Kehrli EW, Simon MC, and Singh H. 1997. PU.1 Functions in a Cell-Autonomous Manner to Control the Differentiation of Multipotential Lymphoid–Myeloid Progenitors. Immunity6: 437–447. [PubMed: 9133423]
53. Fisher RC, and Scott EW. 1998. Role of PU.1 in hematopoiesis. Stem Cells16: 25–37. [PubMed: 9474745]

54. Simon MC, Olson M, Scott E, Hack A, Su G, and Singh H. 1996. Terminal myeloid gene expression and differentiation requires the transcription factor PU.1. Curr. Top. Microbiol. Immunol. 211: 113–119. [PubMed: 8585941]
55. Yukawa M, Jagannathan S, Vallabh S, Kartashov AV, Chen X, Weirauch MT, and Barski A. 2020. AP-1 activity induced by co-stimulation is required for chromatin opening during T cell activation. J. Exp. Med. 217(1): e20182009. [PubMed: 31653690]
56. Baltus B, van Dijk TB, Caldenhoven E, Zanders E, Raaijmakers JAM, Lammers J-WJ, Koenderman L, and de Groot RP. 1998. An AP-1 site in the promoter of the human IL-5Rα gene is necessary for promoter activity in eosinophilic HL60 cells. FEBS Lett. 434: 251–254. [PubMed: 9742933]
57. Sun Z, Yergeau DA, Wong IC, Tuypens T, Tavernier J, Paul CC, Baumann MA, Auron PE, Tenen DG, and Ackerman SJ. 1996. Interleukin-5 receptor alpha subunit gene regulation in human eosinophil development: identification of a unique cis-element that acts lie an enhancer in regulating activity of the IL-5R alpha promoter. Curr. Top. Microbiol. Immunol. 211: 173–187. [PubMed: 8585949]
58. Shahrin NH, Diakiw S, Dent LA, Brown AL, and D'Andrea RJ. 2016. Conditional knockout mice demonstrate function of Klf5 as a myeloid transcription factor. Blood128: 55–59. [PubMed: 27207790]
59. Feinberg MW, Wara AK, Cao Z, Lebedeva MA, Rosenbauer F, Iwasaki H, Hirai H, Katz JP, Haspel RL, Gray S, Akashi K, Segre J, Kaestner KH, Tenen DG, and Jain MK. 2007. The Kruppel-like factor KLF4 is a critical regulator of monocyte differentiation. EMBO J. 26: 4138–4148. [PubMed: 17762869]
60. Barbarani G, Fugazza C, Strouboulis J, and Ronchi AE. 2019. The Pleiotropic Effects of GATA1 and KLF1 in Physiological Erythropoiesis and in Dyserythropoietic Disorders. Front. Physiol. 10. [PubMed: 30740057]
61. Knights AJ, Vohralik EJ, Houweling PJ, Stout ES, Norton LJ, Alexopoulos SJ, Yik JJ, Mat Jusoh H, Olzomer EM, Bell-Anderson KS, North KN, Hoehn KL, Crossley M, and Quinlan KGR. 2020. Eosinophil function in adipose tissue is regulated by Krüppel-like factor 3 (KLF3). Nat. Commun. 11: 2922. [PubMed: 32523103]
62. Kanzaki M, Shibagaki N, Hatsushika K, Mitsui H, Inozume T, Okamoto A, Dobashi Y, Ogawa H, Shimada S, and Nakao A. 2007. Human eosinophils have an intact Smad signaling pathway leading to a major transforming growth factor-beta target gene expression. Int. Arch. Allergy Immunol. 142: 309–317. [PubMed: 17135762]
63. Cho J, Doshi A, Rosenthal P, Beppu A, Miller M, Aceves S, and Broide D. 2014. Smad3 deficient mice have reduced esophageal fibrosis and angiogenesis in a mouse model of egg induced eosinophilic esophagitis. J. Pediatr. Gastroenterol. Nutr. 59: 10–16. [PubMed: 24590208]
64. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, and Sherlock G. 2000. Gene Ontology: tool for the unification of biology. Nat. Genet. 25: 25–29. [PubMed: 10802651]
65. Gene Ontology Consortium T2019. The Gene Ontology Resource: 20 years and still GOing strong. Nucleic Acids Res. 47: D330–D338. [PubMed: 30395331]
66. Mack EA, Stein SJ, Rome KS, Xu L, Wertheim GB, Melo RC, and Pear WS. 2019. Trib1 regulates eosinophil lineage commitment and identity by restraining the neutrophil program. Blood133(22): 2413–2426. [PubMed: 30917956]

Key points:

- Active (H3K27ac+) eosinophil-specific enhancers are enriched for PU.1
- Active eosinophil-specific enhancers are enriched in key functional gene pathways
- Eosinophil super enhancers associate with genes with risk loci for allergic disease

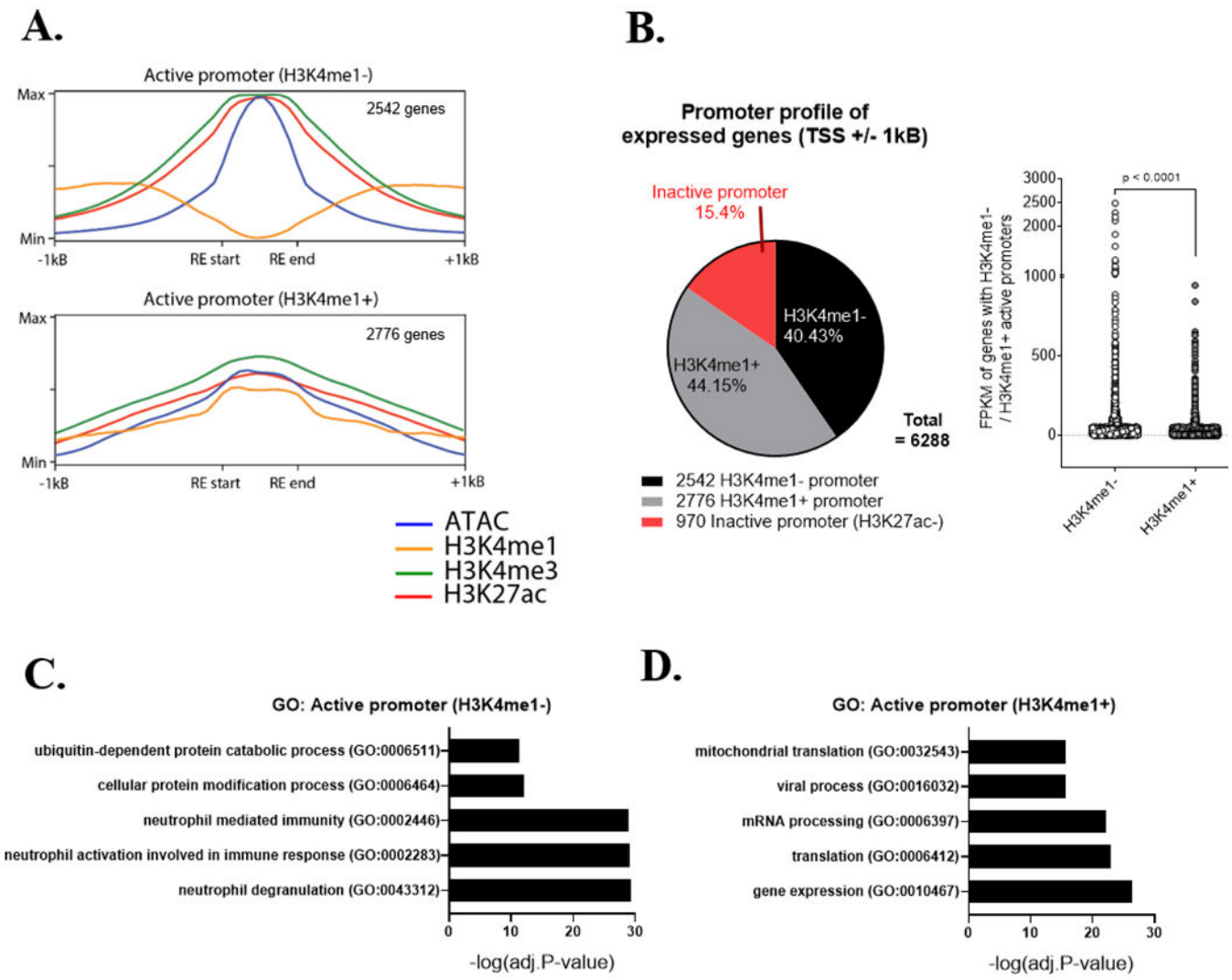

Figure 1. Eosinophil promoter profile.

(A) Average tag density profile plots for open areas of chromatin (ATAC; blue), H3K4me1 (orange), H3K4me3 (green) and H3K27ac (red) showing the binding locations relative to the identified regulatory elements (RE) within 1 kB of the TSS of expressed genes in the eosinophil epigenome. Active promoters were identified by ATAC, H3K4me3 and H3K27ac co-marked peaks within 1 kB of the TSS and classified depending on their H3K4me1 status (Top; H3K4me1-, Bottom; H3K4me1+). (B) Pie chart showing proportion of expressed genes associated with either promoter type, with scatter plot showing corresponding gene expression (FPKM) levels (mean ± SEM). (C) Gene ontology (GO) analysis of biological processes for the 2,542 expressed genes possessing an H3K4me1− active promoter. (D) Gene ontology (GO) analysis of biological processes for the 2,776 expressed genes possessing an H3K4me1+ active promoter.

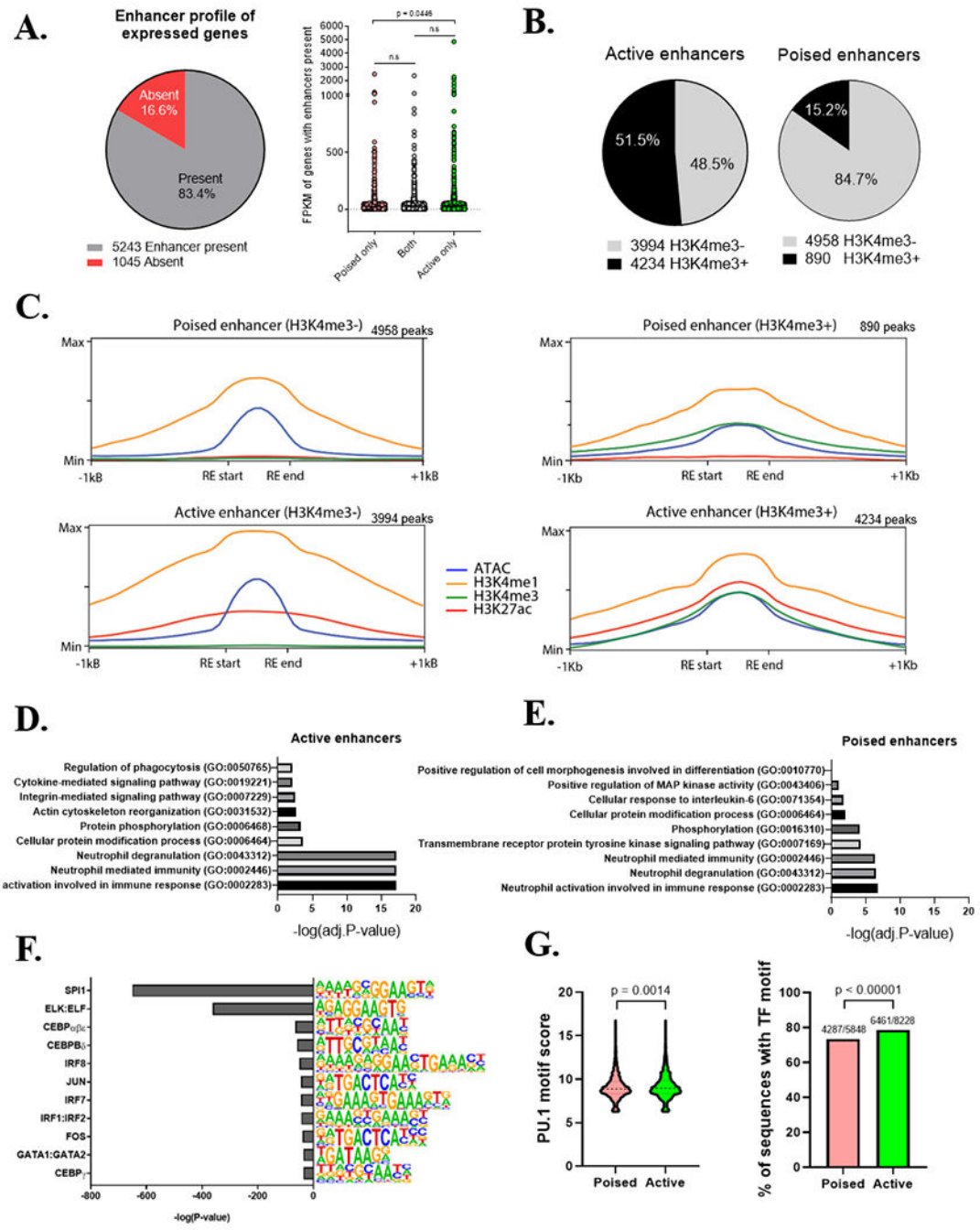

Figure 2. Active eosinophil enhancers are associated with gene pathways involved in granulocyte-mediated immunity.

(A) Pie chart showing proportion of expressed genes associated with any enhancer type (located 1 – 20 kB from the TSS of expressed genes), with scatter plot showing corresponding gene expression (FPKM) levels (mean ± SEM). (B) Pie charts showing proportion of enhancers with/without H3K4me3. (C) Average tag density profile plots for open areas of chromatin (ATAC; blue), H3K4me1 (orange), H3K4me3 (green) and H3K27ac (red) showing the binding locations relative to the identified regulatory elements (RE) within the eosinophil epigenome. Enhancers lacking H3K27ac are classified as poised (top),

whereas those with H3K27ac are classified as active (bottom). (Left) H3K4me3− enhancers were identified by ATAC, H3K4me1 and H3K27ac co-marked peaks within 1 - 20 kB of an expressed gene, lacking H3K4me3. (Right) H3K4me3+ enhancers were identified by ATAC, H3K4me1 and H3K27ac co-marked peaks within 1 - 20 kB of an expressed gene, and by the presence of H3K4me3. (D) Gene ontology (GO) analysis of biological processes for the 4,326 expressed genes possessing an active enhancer. (E) Gene ontology (GO) analysis of biological processes for the 3,483 expressed genes possessing a poised enhancer. (F) HOMER motif enrichment of enhancer genomic regions indicating highly represented transcription factor binding site motifs. The DNA sequence of each transcription factor binding motif is shown to the right, with the nucleotide position and probability shown visually. Significance is expressed as −log(P-value). (G) Left; Binding motif score for PU.1 between poised and active enhancers, showing a significant increase in the motif score in active enhancers. Significance was accepted with $p < 0.05$ by student's t test. Right; Proportion of poised and active enhancers predicted to have a PU.1 binding motif present by HOMER. Significance was accepted with $p < 0.05$ by chi-squared test.

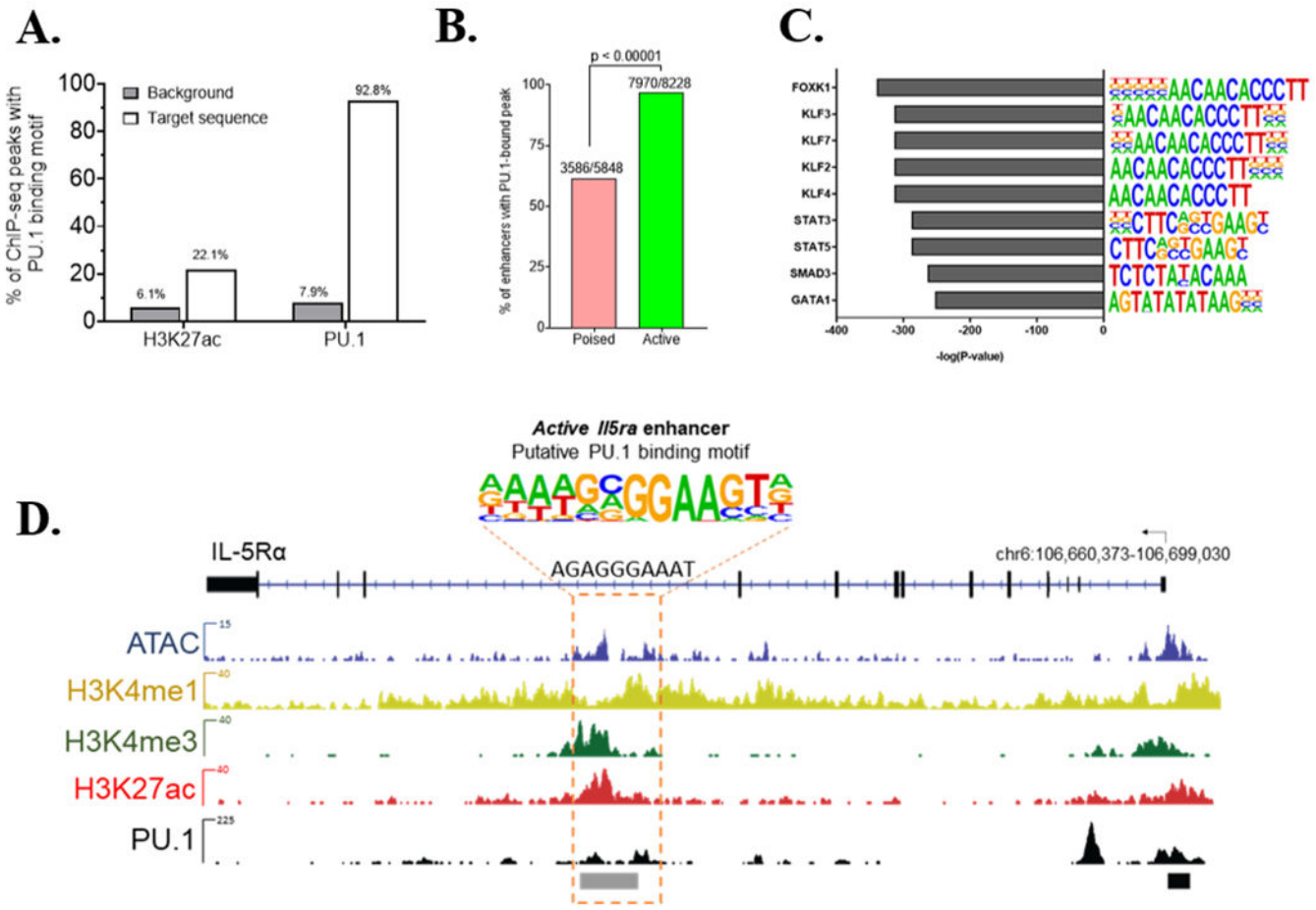

Figure 3. Active eosinophil enhancers are enriched for PU.1-bound regions.

(A) Percentage of ChIP-seq peaks predicted to contain PU.1 binding motifs. (B) Proportion of poised and active enhancer with a PU.1-bound peak as determined by ChIP-seq. Significance was accepted with $p < 0.05$ by chi-squared test. (C) HOMER *de novo* motif enrichment analysis of PU.1-bound ChIP-seq peaks that did not express a PU.1 binding motif (GGAA). (D) Representative image showing epigenetic signature of identified enhancers and PU.1 binding in the *Il5ra* gene locus. The image was captured from the IGV genome browser and shows the tag density profile of aligned reads for open chromatin (ATAC), H3K4me1, H3K4me3, H3K27ac and PU.1. Black arrow starting at the transcription start site (TSS) indicates direction of gene transcription. Black box underneath aligned tracks shows location of identified promoter region. Grey box underneath aligned tracks shows location of identified enhancer region. Orange box overlay highlights an eosinophil enhancer region with a predicted PU.1 binding site and confirmed binding by PU.1 ChIP-seq.

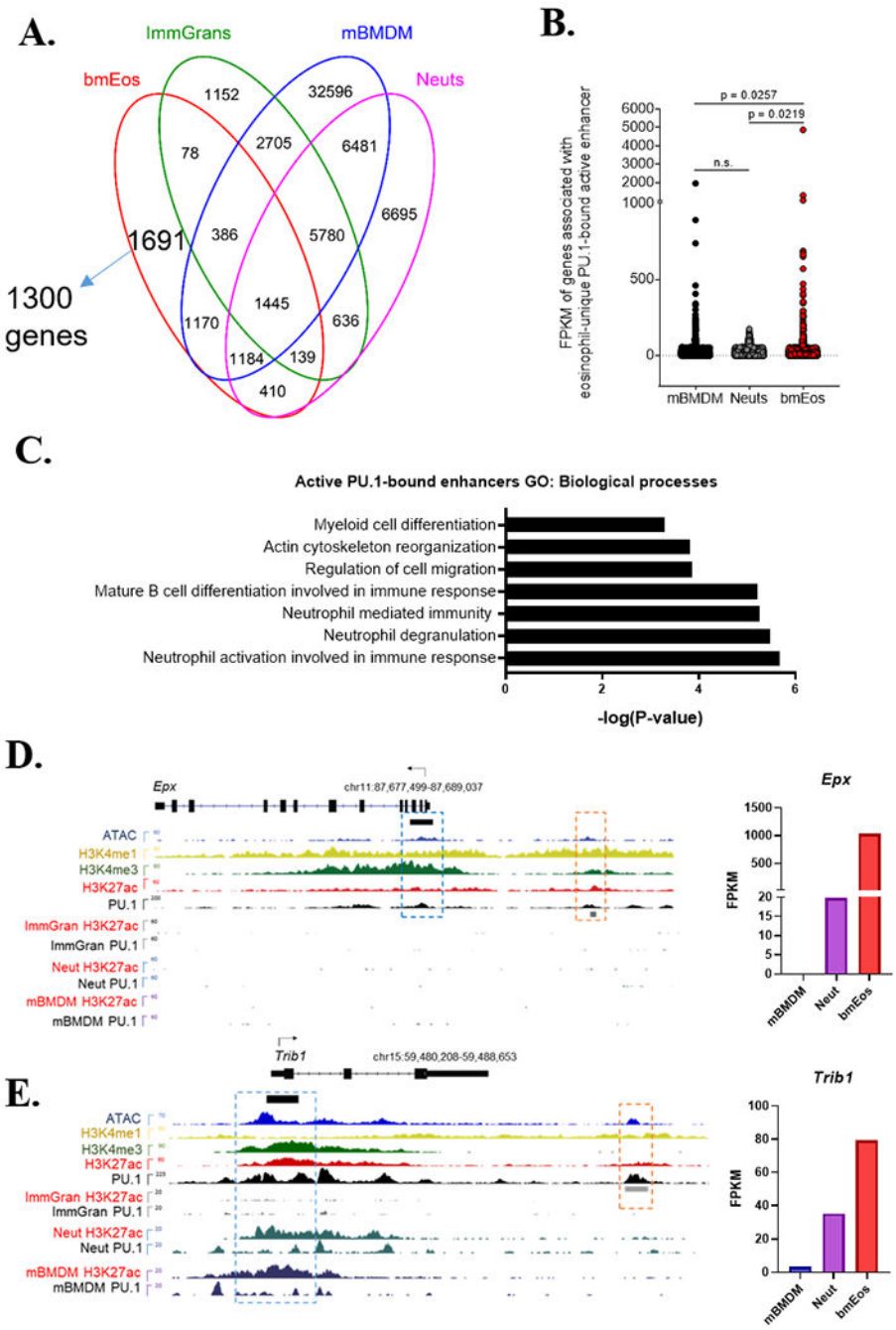

Figure 4. Identification of PU.1-bound, eosinophil-specific active enhancers.

Differential binding and occupancy analysis of identified bmEos PU.1-bound active enhancers (1 – 20 kB from TSS) against PU.1-bound active enhancers identified from three publically available cell populations, bone marrow–derived macrophages (mBMDMs; (21)), neutrophils (Neuts; (22)) and immature granulocytes (ImmGrans; (20)) to define unique and shared, active PU.1-buond enhancers between eosinophils and the described myeloid populations. (A) Multi-way comparison of PU.1-bound active enhancers identified between each myeloid cell type. Numbers enumerate individual super enhancers that are

unique or common between eosinophils and defined myeloid populations. (B) Scatter plot showing gene expression (FPKM) levels (mean ± SEM) of the 1300 genes found to contain unique, eosinophil-specific, PU.1-bound active enhancers between myeloid cell populations. (C) Gene Ontology (GO) analysis of biological processes for the set of 1300 expressed genes containing at least one unique, PU.1-bound, eosinophil-specific enhancer. Epigenetic signatures of representative key eosinophil genes (D) *Epx* and (E) *Trib1* with bar chart showing gene expression level (FPKM) between bmEos and the two myeloid populations (mBMDM and Neuts) with paired ChIP- and RNA-seq data publically available. The images were captured from the IGV genome browser and show the tag density profile of aligned ATAC- and ChIP-seq reads. Black arrows starting at the transcription start site (TSS) indicate the direction of gene transcription. PU.1-bound, eosinophil-specific enhancers (orange boxes) were identified by their presence (grey boxes underneath alignment tracks) in eosinophils and their absence in mBMDMs, Neuts and ImmGrans. Gene promoter regions are also shown (blue boxes) with black bars underneath the gene structure.

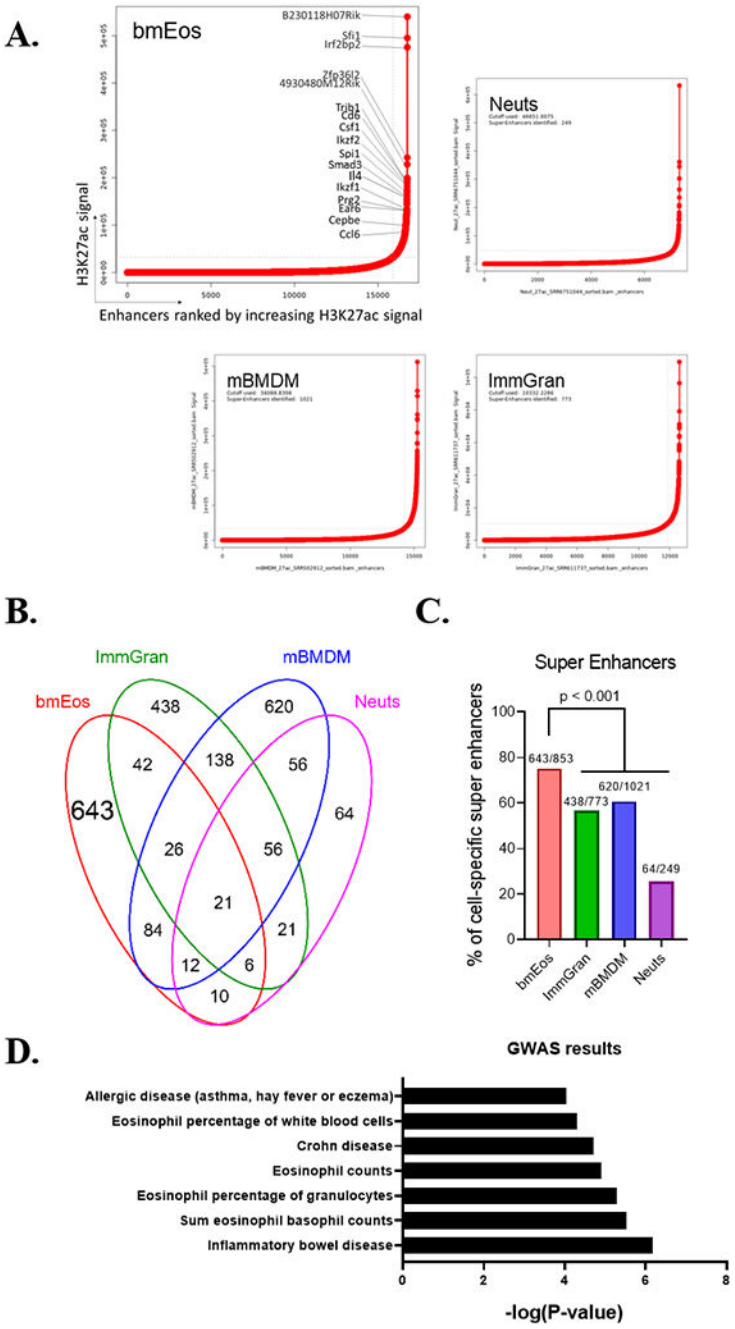

Figure 5. Eosinophil-specific super enhancers.

Super enhancers identified by ranked H3K27ac intensity using the ROSE algorithm (A) Top left; eosinophil super enhancers in murine bone marrow-derived eosinophils (bmEos). Bottom left; murine bone marrow–derived macrophage (mBMDM) super enhancers. Top right; murine neutrophil (Neuts) super enhancers. Bottom right; murine immature granulocyte (ImmGran) super enhancers. (B) Venn diagram showing proportion of cell type–specific super enhancers identified as unique or shared between murine myeloid populations. (C) Proportion of cell type–specific super enhancers ($p < 0.001$ by chi-squared test). (D)

GWAS analysis for the set of 643 eosinophil-specific super enhancers shows significant enrichment of murine genes whose human homologues are associated with eosinophil levels (eosinophilia, percentage of white blood cells or granulocytes) and allergic disease in humans.