RUNX1 is required in granulocyte-monocyte progenitors to attenuate inflammatory cytokine production by neutrophils

Alexandra U. Zezulin[1], Daniel Yen[1], Darwin Ye[2], Elizabeth D. Howell[1], Erica Bresciani[3], Jamie Diemer[3], Jian-gang Ren[4], Mohd Hafiz Ahmad[5], Lucio H. Castilla[5], Ivo P. Touw[6], Andy J. Minn[2], Wei Tong[4], P. Paul Liu[3], Kai Tan[4,7], Wenbao Yu[4,7,8], and Nancy A. Speck[1,8]

[1]Department of Cell and Developmental Biology and Abramson Family Cancer Research Institute, Perelman School of Medicine, University of Pennsylvania, Philadelphia PA, 19104
[2]Department of Radiation Oncology and Abramson Family Cancer Research Institute, Perelman School of Medicine, University of Pennsylvania, Philadelphia PA, 19104
[3]Oncogenesis and Development Section, Division of Intramural Research, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892, USA
[4]Department of Pediatrics, Children's Hospital of Philadelphia and Perelman School of Medicine, University of Pennsylvania, Philadelphia PA, 19104
[5]Department of Molecular, Cell and Cancer Biology, University of Massachusetts Medical School, Worcester MA, 01605
[6] Department of Hematology, Erasmus MC, Rotterdam, NL
[7]Division of Oncology and Childhood Cancer Research, Children's Hospital of Philadelphia, Philadelphia PA, 19104
[8]Co-corresponding authors

Supplemental Materials

SUPPLEMENTAL METHODS

Bone Marrow Harvest: Bone marrow (BM) cells were harvested by spinning isolated bone at 12,000 x g for 1 min, filtering through a 70-μm filter, and resuspending in cold PBS with 2% heat-inactivated FBS. Red blood cells were lysed in ACK Lysing Buffer (Gibco) for 2 min on ice.

Transplant analyses: B6.SJL-*Ptprc*^{a}*Pepc*^{b}/BoyJ (CD45.1) mice were subjected to a split dose of 1000 cGy, 3 h apart. Each recipient received a 10:1 ratio of BM cells from control (CD45.1/2) and Runx1$^{\Delta HSC}$ mice (CD45.2) by retro-orbital injection. We assessed donor (CD45.2) and competitor (CD45.1/2) engraftment in peripheral blood at weeks 4, 16, and 20. Mice were sacrificed at 24 weeks post-transplant and bone marrow was harvested for final engraftment and cytometric bead array analyses.

***Ex vivo* culture and stimulation:** Following isolation of neutrophils via FACS, cells were rested for two hours at 37°C in Hank's media consisting of 1x Hank's Balanced Salt Solution (Gibco) with 25 mM HEPES and 10% heat-inactivated fetal bovine serum (FBS; Gemini Bio-Products) unless otherwise indicated. Cells were stimulated at 37°C with LPS (10 ng/mL or 100 ng/mL) (E. coli O111:B4, Imgen Technologies) in Hank's media. For the CD14 blocking experiments, neutrophils were incubated prior to LPS stimulation for 30 min at 37°C with 5, 15, or 50 μg/mL of function-blocking anti-CD14 or isotype control. For the ruxolitinib experiments, when indicated, neutrophils were incubated for 1 hour with 20 μM of ruxolitinib (Selleck Chemicals) prior to LPS stimulation. For intracellular flow assays, stimulation media also included Brefeldin A (BD GolgiPlug).

Cytokine quantification: 200,000 neutrophils were plated in a total volume of 100 μL into 96 well plates. After 8 hours of LPS stimulation, cells were pelleted, and supernatants were frozen until

analyzed. Absolute multiplex quantification of an 8-factor panel of cytokines, chemokines, and growth factors was performed using the Cytometric Bead Array (CBA) mouse soluble protein flex sets (BD) according to the manufacturer's instructions. A sigmoidal 4-parameter logistic regression was used to fit a standard curve and interpolate unknown concentrations.

Real-time quantitative reverse transcriptase-PCR (RT-qPCR): Isolated neutrophils (EasySep™ Mouse Neutrophil Enrichment Kit) were incubated in the presence of 20 μM of ruxolitinib (Selleck Chemicals), 1 μ/mL anti-mouse IFNAR-1 mAB, or vehicle, or isotype control for 30 min- 1 hr. Cells were then stimulated with 1,000 U/mL INF-α or vehicle for 1-2 hrs. Total RNA was isolated from neutrophils (RNeasy Mini Kit, Qiagen) and quantified by Nanodrop. Total RNA was reverse transcribed into cDNA (ProtoScript® II First Strand cDNA Synthesis Kit, NEB). Relative STAT1, CD14, IRGM2, and GBP2 mRNA levels were determined using Power SYBR Green Master Mix (Thermo Fisher Scientific). Cyclophilin was used as an endogenous control. qPCR primers are: CD14 F: GGCGCTCCGAGTTGTGACT; CD14 R: TACCTGCTTCAGCCCAGTGA; STAT1 F: GCCTCTCATTGTCACCGAAGAAC; STAT1 R: TGGCTGACGTTGGAGATCACCA; Cyclophilin F: ATGGCAAATGCTGGACCAA; Cyclophilin R: GCCATCCAGCCATTCAGTCT; IRGM2 F: CCCCTTCTTTCACGGCAGT; IRGM2 R: GGCAGTTGAGTCACCTGAGG; GBP2 F: CTGCACTATGTGACGGAGCTA; GBP2 R: CGGAATCGTCTACCCCACTC.

Western Blot: Neutrophils were starved in 0.05% BSA for 2-4 hrs, and then stimulated with vehicle or IFN-α for indicated time points and snap-frozen in dry ice. Cell pellets were lysed in LDS loading buffer and sonicated for homogenization. Samples were resolved by SDS-PAGE and transferred to a nitrocellulose membrane. For all primary phospho-antibody blots, membranes were blocked with 5% BSA (BP1600-100, Fisher Bioreagents) in TBS-T, while other primary antibody blots were blocked with 5% non-fat milk (sc2325, Santa Cruz). Membranes were

incubated with primary antibodies overnight in a cold room (complete list of antibodies listed in Table S1). Following primary antibody blots, membranes were washed with TBS-T, and then incubated with HRP-conjugated secondary antibody for 1 hr at room temperature. After washing, membranes were developed with ECL (#34095, Thermo Scientific). In certain cases, Western blots were stripped and reprobed with a second set of primary antibodies. Immunoblots were processed and developed by KwikQuant imager (Kindle Biosciences, LLC). Quantification of western blots was performed using ImageJ software.

H3K27ac ChIP-Seq: FACS purified neutrophils ($CD11b^{+}SiglecF^{-}F4/80^{-}Ly6G^{+}$) were fixed and cross-linked by 1% formaldehyde in 1× Fixing Buffer for 5 min at room temperature according to the vendor's protocol (Covaris). Cells were then stored at -80℃. Crosslinked cells were thawed on ice (100K cells) and resuspended in 1×Shearing Buffer and sonicated with Covaris E220 for 720s using the following settings: 5% duty factor, 105W Peak Incident Power, and 200 cycles per burst. 10% of sheared chromatin was used as the input and the remaining chromatin was divided into two equal aliquots for immunoprecipitation (IP) (50,000 cells per IP). IPs were performed using ChIP-IT high-sensitivity kit (Active Motif). IP and input samples were treated with RNase A followed by proteinase K. Crosslinking was reversed by incubation overnight at 65 ℃ and DNA was purified using a MinElute PCR purification kit (Qiagen). All IP DNA and 1-2 ng of input DNA were used for library preparation with the ThruPLEX DNA-Seq kit and Smarter DNA single index kit (Takara). 13 and 9 cycles were used for IP DNA and input DNA, respectively at step 5. Following library amplification, the libraries were bead purified. The concentrations were measured using both Qubit and KAPA qPCR. Agilent Bioanalyzer 2100 was used to check the quality of libraries. Libraries were sequenced on Illumina HiSeq 2500 sequencer in single-end mode with a read length of 75bp.

H3K27ac ChIP-Seq data processing: Sequencing reads were demultiplexed using Bcl2Fastq v2.20 then trimmed and filtered for quality using Trim Galore (Martin 2011) with the following settings: fastqc, and trim1. Reads were then aligned to the mouse genome (mm10) using bowtie 2.3.5.1 (Langmead et al. 2009; Langmead and Salzberg 2012). Only uniquely mapped reads with fewer than 2 mismatches were used for downstream analyses. Samtools v.1.1 (Danecek et al. 2021) was used to convert SAM files to BAM files, and Sambamba v0.6.6 (Tarasov et al. 2015) was used to filter out duplicates, multi-mappers, reads mapped to ChrM or blacklist regions, and unmapped reads. MACS2 2.1.4 (Zhang et al. 2008) was used for peak calling with the following parameters: narrow, q: 0.05. Control or Runx1$^{\Delta GMP}$ specific differential peaks were called if the RPKM fold change between Control and Runx1$^{\Delta GMP}$ peaks was greater than 2. Merged replicates were used to create bigwig files for visualization using Deeptools v3.3.0 (Ramirez et al. 2016)). The following parameters were used: normalized to reads per genomic content (RPGC), effective genome size: 2,308,125,349 bp, ignore for normalization: ChrX, min fragment length: 20, bin size: 10. IGV or UCSC genome browser was used for visualization. Deeptools was also used to plot regions of differential peaks. GREAT 4.0.4 (McLean et al. 2010; Hiller et al. 2013) was used for linking peak regions to genes and subsequent gene ontology annotation using the following parameters: species assembly: mm10; association rule: basal+extension: 5000 bp upstream, 1000 bp downstream, 1,000,000 bp max extension, curated regulatory domains included.

Bulk ATAC-seq: Neutrophils from human patient/control peripheral blood were isolated using the EasyStep Human Neutrophil Isolation Kit (Stem Cell). Patient and control neutrophils were isolated and processed side-by-side on the same day. 50,000 neutrophils were collected and washed with cold PBS. Cell pellets were resuspended in 50 µl of cold lysis buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$ and 0.1% Tween-20, 0.1% NP-40, 0.01% Digitonin, 1% BSA) and incubated on ice for 4 min. Lysis was halted with the addition of 50 µL of wash buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, and 0.1% Tween-20, 0.1%, 1% BSA).

Samples were centrifuged at 500 x g, 4 °C for 5 min. Nuclei pellets were resuspended in 50 µl of wash buffer and immediately centrifuged at 500 x g, 4 °C for 5 min. Nuclei pellets were resuspended in 50 µl of tran position reaction mix (1× Tagment DNA Buffer, 2.5 µl of Tagment DNA Enzyme 1) and incubated for 30 min at 37 °C. Subsequent steps of the protocol were performed as previously described (Buenrostro et al. 2013). Libraries were purified using a Qiagen MinElute Gel Purification kit for mouse samples and SPRI-Select bead purification for human samples. The concentrations were measured using both Qubit and KAPA qPCR. Agilent Bioanalyzer 2100 was used to determine the quality of libraries. Libraries were sequenced on the Illumina HiSeq 2500, with 75-bp paired-end reads. Each sample had two biological replicates.

Bulk ATAC-seq data processing for mouse samples: Sequencing reads were demultiplexed using Bcl2Fastq v2.20 then trimmed and filtered for quality using Trim Galore v0.6.4 (Martin 2011) with the following settings: fastqc, paired, and trim1. Reads were then aligned to the mouse genome (mm10) using bowtie 2.3.5.1 (Langmead et al. 2009; Langmead and Salzberg 2012). Only uniquely mapped reads with fewer than 2 mismatches were used for downstream analyses. Samtools v.1.1 (Danecek et al. 2021) was used to convert SAM files to BAM files, and Sambamba v0.6.6 (Tarasov et al. 2015) to filter out duplicates, multi-mappers, reads mapped to ChrM or blacklist regions, and unmapped reads. MACS2 2.2.7.1 (Zhang et al. 2008) was used for peak calling using the following parameters: BAMPE, q: 0.05. Peaks from Control and Runx1$^{\Delta GMP}$ were merged if there is at least 1bp overlap between two peaks. Reads Per Kilobase Million (RPKM) of a peak was then calculated using bedtools v2.25 (Quinlan and Hall 2010). Peaks with RPKM less than 0.5 in both Control and Runx1$^{\Delta GMP}$ cells were filtered from downstream analysis. Control or Runx1$^{\Delta GMP}$ specific peaks were called if the RPKM fold change between Control and Runx1$^{\Delta GMP}$ peaks was greater than 2. Merged replicates were used to create bigwig files for visualization using Deeptools v3.3.0 (Ramirez et al. 2016) and the following parameters: normalized to Bins Per Million mapped reads (BPM), bin size: 50. IGV or UCSC genome browser was used for

visualization. Deeptools was also used to plot regions of differential peaks. GREAT 4.0.4 (McLean et al. 2010; Hiller et al. 2013) was used for linking peak regions to genes and subsequent gene ontology annotation using the following parameters: species assembly: mm10; association rule: basal+extension: 5000 bp upstream, 1000 bp downstream, 1,000,000 bp max extension, curated regulatory domains included. Homer v4.11 (Heinz et al. 2010) was used for genomic annotation of peak regions.

Bulk ATAC-seq data processing for human samples: Human data were processed as described with the following changes. Reads were aligned to the human genome (hg38) using Bowtie 2.3.5.1 (Langmead et al. 2009; Langmead and Salzberg 2012). MACS2 2.2.7.1 (Zhang et al. 2008) was used for peak calling using the following parameters: BAMPE, q: 0.05. Peaks from the patient and control samples were called individually. For downstream analysis, we compared the patient and unaffected family members individually. The peaks from the patient sample and the corresponding control sample were merged for each comparison and RPKM per merged peak was calculated using BEDtools. Peaks with RPKM less than 0.5 in both control and patient were filtered from downstream analysis. Control or patient-specific peaks were called if the RPKM fold change between control and patient peaks was greater than 2. Individual samples were used to create bigwig files for visualization. GREAT 4.0.4 (McLean et al. 2010; Hiller et al. 2013) was used for linking peak regions to genes and subsequent gene ontology annotation using the following parameters: species assembly: hg38; association rule: basal+extension: 5000 bp upstream, 1000 bp downstream, 1,000,000 bp max extension, curated regulatory domains included.

Footprinting analysis and motif analysis: ATAC-seq footprinting analysis was performed using the Regulatory Genomics Toolbox (RGT) and HMM-based IdeNtification of Transcription factor footprints (HINT) software (Li et al. 2019). In brief, footprints were called from regions of chromatin accessibility (peaks on merged replicates) for each sample. Called footprints were then matched

to TF motifs using the JASPAR 2020 vertebrate motif database. Differential activity of transcription factors, as well as plots of each transcription factor footprint, were determined using HINT-differential. p-value and read counts were calculated using RGT HINT-differential for each transcription factor. Enriched transcription factor footprints at regions of chromatin with increased accessibility and corresponding p-values were determined using BiFET v.1.16.0 software (Youn et al. 2019). TF footprints were extracted for more in-depth analysis. TF motif scores were determined by the RGT toolbox (Gusmao et al. 2016). ATAC-seq motif analysis was done as follows: TF motifs were first scanned on all peaks by R package motifmatchr and a motif hit was called on a peak if the p-value was less than 10^{-5}. The TF motif was enriched on a list of Runx1$^{\Delta GMP}$ or Control specific peaks if the p-value of a binomial test, comparing the TF motif on the rest of the peaks, was less than 0.05.

Enrichment of transposable elements: An ATAC-seq peak is associated with a TE (annotated by RepeatMasker, Reference: Smit, AFA, Hubley, R & Green, P. *RepeatMasker Open-4.0*. 2013-2015 http://www.repeatmasker.org) if there is at least 1 base pair overlap between them. Enrichment of TEs was performed by the following permutation test for each TE class and family, respectively: For a given TE class (or family), we counted the number of gained peaks in Runx1$^{\Delta GMP}$ cells that overlapped with any TE belonging to the given TE class (or family). We randomly selected the same number of peaks from the rest of the peaks (lost peaks in Runx1$^{\Delta GMP}$ cells or stable peaks) and counted the number of peaks that overlapped with the given TE class (or family). This process was repeated 1000 times to generate a null distribution of the number of peaks overlapped with TE. The empirical p-value was then calculated as the number of times the permutation yielded values greater than the number of gained peaks overlapped with the TE class (or family), divided by 1000.

Cytosolic dsRNA immunoprecipitation and sequencing: dsRNA was immunoprecipitated as previously described (Tavora et al. 2020). Briefly, Protein G Dynabeads and Protein A Dynabeads (1:1) were washed and resuspended in NET-2 buffer. 5 µg of 9D5 dsRNA rabbit IgG mAb was bound to the beads for 1-2 hours in the cold room on a shaker. Six million FACS-purified neutrophils were lysed for 10 min by end-over-end rotation in the cold room in cytosolic lysis buffer. The cell lysate was centrifuged at 980 x g for 3 min at 4℃. The centrifugation steps were repeated for a total of three times. The cytosolic supernatant was then transferred to a fresh tube and spun at 17,000 x g for 10 min. For immunoprecipitation, the lysate was diluted 1:4 with NET-2-TurboDNase buffer. 95D-Dynabeads were added to the lysate and end-to-end rotated for 2 hr at 4 °C. Following magnetic separation, beads were washed with high salt washing buffer and then NET-2 buffer. 95D-bound dsRNA was extracted with Trizol reagent and purified using Direct-zol™ RNA Miniprep Kit (Zymo Research). RNA samples were ribo-depleted with Ribominus™ Eukaryote v2 kit (Thermo Fisher Scientific). Libraries were prepared using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina. Libraries were sequenced on Illumina HiSeq 2500 sequencer in paired-end mode with a read length of 75bp.

dsRNA-seq data processing: Raw dsRNA-seq sequencing data (.bcl files) was converted into Fastq files and de-multiplexed using Bcl2Fastq v2.20 software. The data in fastq file format was processed with the toolkit SQuIRE (Yang et al. 2019). The raw reads were aligned to the mm10 reference genome and TEs were annotated using RepeatMasker. The SQuIRE call module was used to perform differential expression analysis of genes, or TE by family, or TE by sub-family between Control and Runx1$^{\Delta GMP}$ cells.

Bulk RNA-seq: Cells were sorted, and total RNA samples were isolated (RNeasy Mini Kit, Qiagen). Sample QC, library preparations, and sequencing reactions were conducted at GENEWIZ, LLC./Azenta US, Inc (South Plainfield, NJ, USA) as follows: RNA samples were

quantified using Qubit 2.0 Fluorometer (ThermoFisher Scientific, Waltham, MA, USA) and RNA integrity was checked using TapeStation (Agilent Technologies). ERCC RNA Spike-In Mix kit (cat. 4456740) from ThermoFisher Scientific, was added to normalized cell number prior to library preparation following the manufacturer's protocol. The RNA sequencing libraries were prepared using the NEBNext Ultra II RNA Library Prep Kit for Illumina using the manufacturer's instructions (NEB). Briefly, mRNAs were initially enriched with Oligod(T) beads. Enriched mRNAs were fragmented for 15 min at 94 °C. First-strand and second-strand cDNA were subsequently synthesized. cDNA fragments were end-repaired and adenylated at 3'ends, and universal adapters were ligated to cDNA fragments, followed by index addition and library enrichment by PCR with limited cycles. The sequencing libraries were validated on the Agilent TapeStation (Agilent Technologies), and quantified by using Qubit 2.0 Fluorometer (ThermoFisher Scientific) as well as by quantitative PCR (KAPA Biosystems). The sequencing libraries were multiplexed and clustered onto a flow cell. After clustering, the flowcell was loaded onto the NovaSeq 6000 instrument and sequenced using a 2x150bp Paired End (PE) configuration according to the manufacturer's instructions.

Bulk RNA-seq data processing: Sequencing and base calling was performed by GeneWiz (https://www.genewiz.com/). Raw sequence data (.bcl files) were converted into fastq files and de-multiplexed using Illumina Bcl2Fastq v2.20 software. One mis- match was allowed for index sequence identification. After investigating the quality of the raw data, sequence reads were trimmed to remove possible adapter sequences and nucleotides with poor quality using Trimmomatic v.0.36. The trimmed reads were mapped to the *Mus musculus* reference genome available on ENSEMBL using the STAR aligner v.2.5.2b. BAM files were generated as a result of this step. Unique gene hit counts were calculated by using feature Counts from the Subread package v.1.5.2. Only unique reads that fell within exon regions were counted. After extraction of gene hit counts, the gene hit counts table was used for downstream differential expression

analysis. Using DESeq2, a comparison of gene expression between the groups of samples was performed. The Wald test was used to generate P values and Log2 fold changes. Genes with adjusted P values < 0.05 and absolute log2 fold changes >1 were called as differentially expressed genes for each comparison. Merged replicates were used to create bigwig files for visualization using Deeptools v3.5.1 (Ramirez et al. 2016) with the following parameters: normalized to Bins Per Million mapped reads (BPM), bin size: 50

CUT&RUN: RUNX1 CUT&RUN was performed as previously described (Shin et al. 2022). Briefly, 200K GMPs were sorted and bound to activated Concanavalin A beads(93569S) for 10 min at RT. Permeabilization and antibody binding was carried out for 2 hrs at 4°C with 90 ug/mL of Runx1 Ab (Ab23980) or IgG negative control (ABIN101961) in digitonin buffer (0.001% Digitonin, 150mM NaCl, 20mM HEPES, 0.5mM Spermidine, 0.1% BSA, 1X protease inhibitor) + 2mM EGTA. pAG-MNase (EpiCypher, 15-1016) binding was carried out for 1 hr at 4 °C in digitonin buffer. Cells were rinsed with a low-salt buffer (0.001% Digitonin, 20mM HEPES, 0.5mM Spermidine, 0.1% BSA, 1X protease inhibitor), and the cleavage reaction was carried out for 5 min at 0 °C with the addition of incubation buffer (0.001% Digitonin, 20mM HEPES, 0.5mM $CaCl_2$, 1x protease inhibitor). The incubation buffer was removed and the cleavage reaction was stopped with the addition of STOP buffer (0.001% Digitonin, 20mM EGTA, 170mM NaCl, 50 ug/mL RNase A, 25 ug/mL glycogen) + 5 ng/mL of Spike-in E.Coli DNA (EpiCypher, 18-1401). Cleaved DNA was released into the supernatant by incubation at 37 °C for 15 min. Beads and cells were pelleted by centrifugation for 5 min at 16000xg, 4°C. DNA from the supernatant was isolated by phenol-chloroform extraction. Sequencing libraries were prepared using the NEBNext Ultra II DNA Kit (E7645S). The purified DNA underwent standard QC (qPCR, KAPA, Bioanalyzer) and was sequenced using the NextSeq 500/550 High Output Kit v2.5 (150 cycles; 20024907).

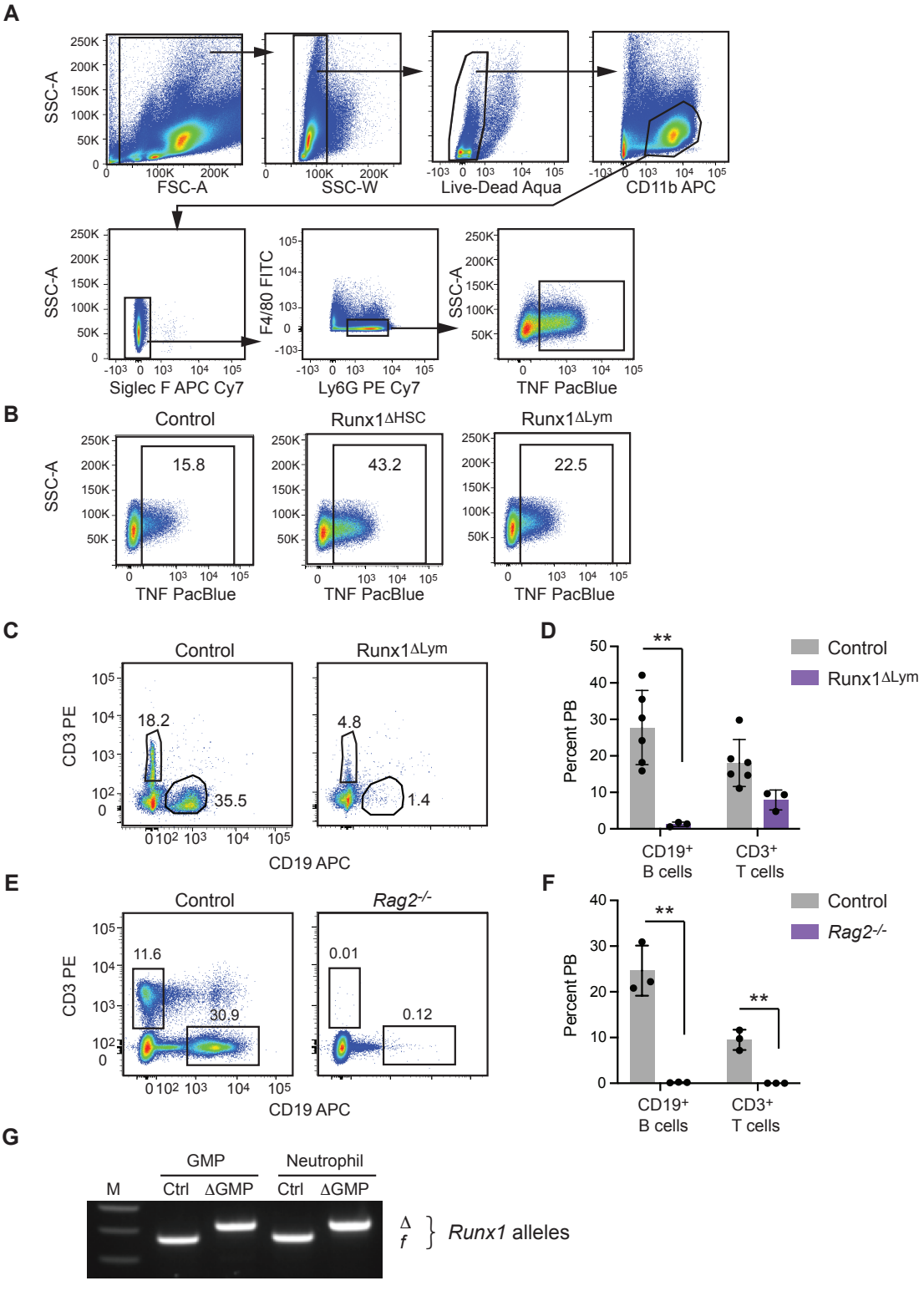
CUT&RUN data processing: Sequencing reads were demultiplexed using Bcl2Fastq v2.20 then trimmed and filtered for quality using Trim Galore (Martin 2011) with the following settings: fastqc, paired, and trim1. Reads were then aligned to the mouse genome (mm10) or the *E. coli* K12 genome using bowtie 2.3.5.1 (Langmead et al. 2009; Langmead and Salzberg 2012). Only uniquely mapped reads with fewer than 2 mismatches and MAPQ equal or greater than 30 were used for downstream analyses. Samtools v.1.1 (Danecek et al. 2021) was used to convert SAM files to BAM files, and Sambamba v0.6.6 (Tarasov et al. 2015) was used to filter out duplicates, multi-mappers, reads mapped to ChrM or blacklist regions, and unmapped reads. MACS2 2.1.4 (Zhang et al. 2008) was used for peak calling with the following parameters: narrow, q: 0.05. Merged replicates were used to create bigwig files for visualization using Deeptools v3.3.0 (Ramirez et al. 2016). The following parameters were used: normalized to Bins Per Million mapped reads (BPM bin size: 50. IGV or UCSC genome browser was used for visualization. GREAT 4.0.4 (McLean et al. 2010; Hiller et al. 2013) was used for linking peak regions to genes and subsequent gene ontology annotation using the following parameters: species assembly: mm10; association rule: basal+extension: 5000 bp upstream, 1000 bp downstream, 1,000,000 bp max extension, curated regulatory domains included.

Hi-C: Hi-C experiments were performed according to the Arima-Hi-C protocol (A510008). The proximally-ligated DNA was sonicated with Covaris E220 for 55s using the following settings: 10% duty factor, 140W Peak Incident Power, and 200 cycles per burst to obtain an average fragment size of 400 bp. Sequencing libraries were prepared using the NEBNext Ultra II DNA library Prep kit (E7645S). The purified DNA underwent standard QC (qPCR, TapeStation, Bioanalyzer) and was sequenced on the NovaSeq6000 following the manufacturer's protocols.

Hi-C data processing: Hi-C data was first demultiplexed by Bcl2Fastq v2.20 and then processed using software HiCPro v3.1.0

(https://genomebiology.biomedcentral.com/articles/10.1186/s13059-015-0831-x). Specifically, the ligation sites were obtained by running HiCPro utils script digest_genome.py with option '-r ^GATC,G^ANTC'. The reads were aligned to mm10 mouse genome. HiCPro was executed in a sequential mode with modules -s mapping -s proc_hic -s quality_checks for each sample. Valid pairs from different biological replicates were then put into the same folder followed up by running -s merge_persample. The contact matrices were constructed by running -s build_contact_maps and normalized by running -s ice_norm modules. The contact matrix was built in 50-kb resolution for downstream analysis. Hi-C compartments were identified at 50-kb resolution using a "sliding window" strategy as previously described (https://www.nature.com/articles/s41586-022-05365-x). First, the expected matrix was calculated by averaging Hi-C contacts at the same distance. Then the observed/expected matrix was obtained by summing the observed Hi-C contacts within a window of 500 kb centered at each bin divided by the sum of expected Hi-C contacts in the same window. The observed/expected matrix was then transformed into a Pearson's correlation matrix. The principal components were then obtained by calculating the covariance matrix of the Pearson's correlation matrix followed by eigenvector decomposition with the 'eigen' function in R. PC1 was used to assign the A and B compartment: regions with positive PC1 values corresponded to the A compartment and negative values corresponded to the B compartment based on their association with gene density.

SUPPLEMENTAL FIGURES


A
SSC-A
FSC-A
SSC-W
Live-Dead Aqua
CD11b APC
Siglec F APC Cy7
F4/80 FITC
Ly6G PE Cy7
TNF PacBlue
B
Control
Runx1ΔHSC
Runx1ΔLym
15.8
43.2
22.5
C
CD3 PE
CD19 APC
18.2
35.5
4.8
1.4
D
Percent PB
CD19+ B cells
CD3+ T cells
E
Rag2-/-
11.6
30.9
0.01
0.12
F
G
GMP
Neutrophil
M
Ctrl
ΔGMP
Δf
Runx1 alleles

Supplemental Figure 1. RUNX1 function in GMPs is necessary to restrict inflammatory cytokine production by neutrophils.

A) Scatter plots showing gating strategy for neutrophil isolation and analysis. Shown is an analysis of TNF^+ $CD11b^+Ly6G^+SiglecF^-F4/80^-$ neutrophils**.**

B) Scatter plots of intracellular TNF in Control, $Runx1^{\Delta HSC}$ and $Runx1^{\Delta Lym}$ neutrophils.

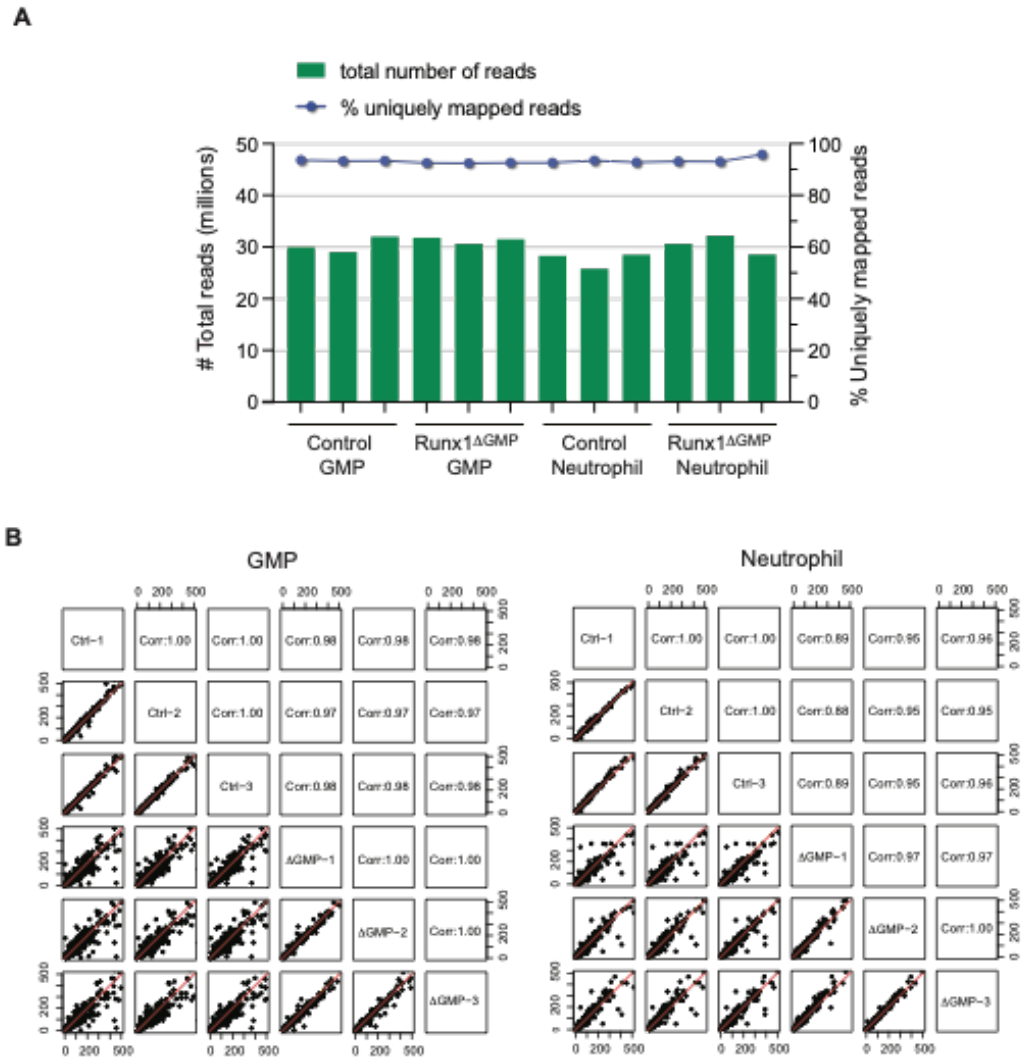
C) Analysis of $CD19^+$ B cells and $CD3^+$ T cells in the PB of Control (C57BL6/J) and $Runx1^{\Delta Lym}$ mice in which *Runx1* floxed alleles were deleted with Rag1-Cre.

D) Quantification of PB B and T cells in $Runx1^{\Delta Lym}$ mice. Mean ± SD, two-tailed, unpaired t-test, representative of 2 experiments, a total of 9 mice analyzed, **P≤0.01.

E) Analysis of B and T cells in the PB of Control and *Rag2*$^{-/-}$ mice.

F) Quantification of PB B and T cells in *Rag2*$^{-/-}$ mice, as in panel D. n= 6 mice analyzed.

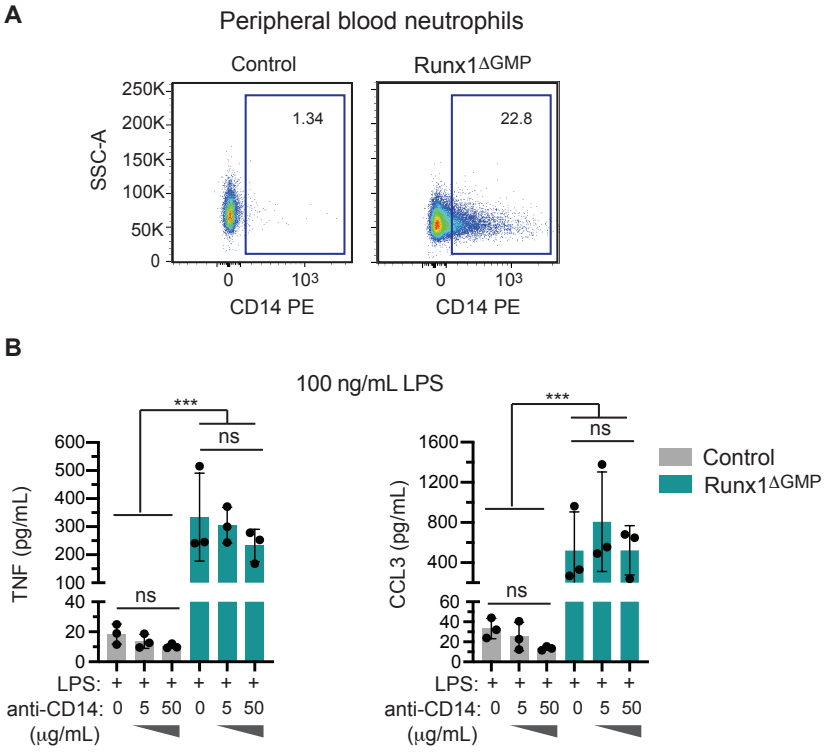
G) PCR showing *Runx1*$^{f/f}$ deletion with Cebpa-Cre in neutrophils and GMPs. Undeleted (f) and deleted (Δ) alleles are shown.



Supplemental Figure 2. Sample sequencing statistics.

A) Mapping statistics of RNA-seq data. RNA-Seq reads were mapped to the *Mus musculus* GRCm38 ERCC reference genome available on ENSEMBL using the STAR aligner v.2.5.2b. Green bars show the number of total reads (in millions). Blue dots show the percentages of uniquely mapped reads.

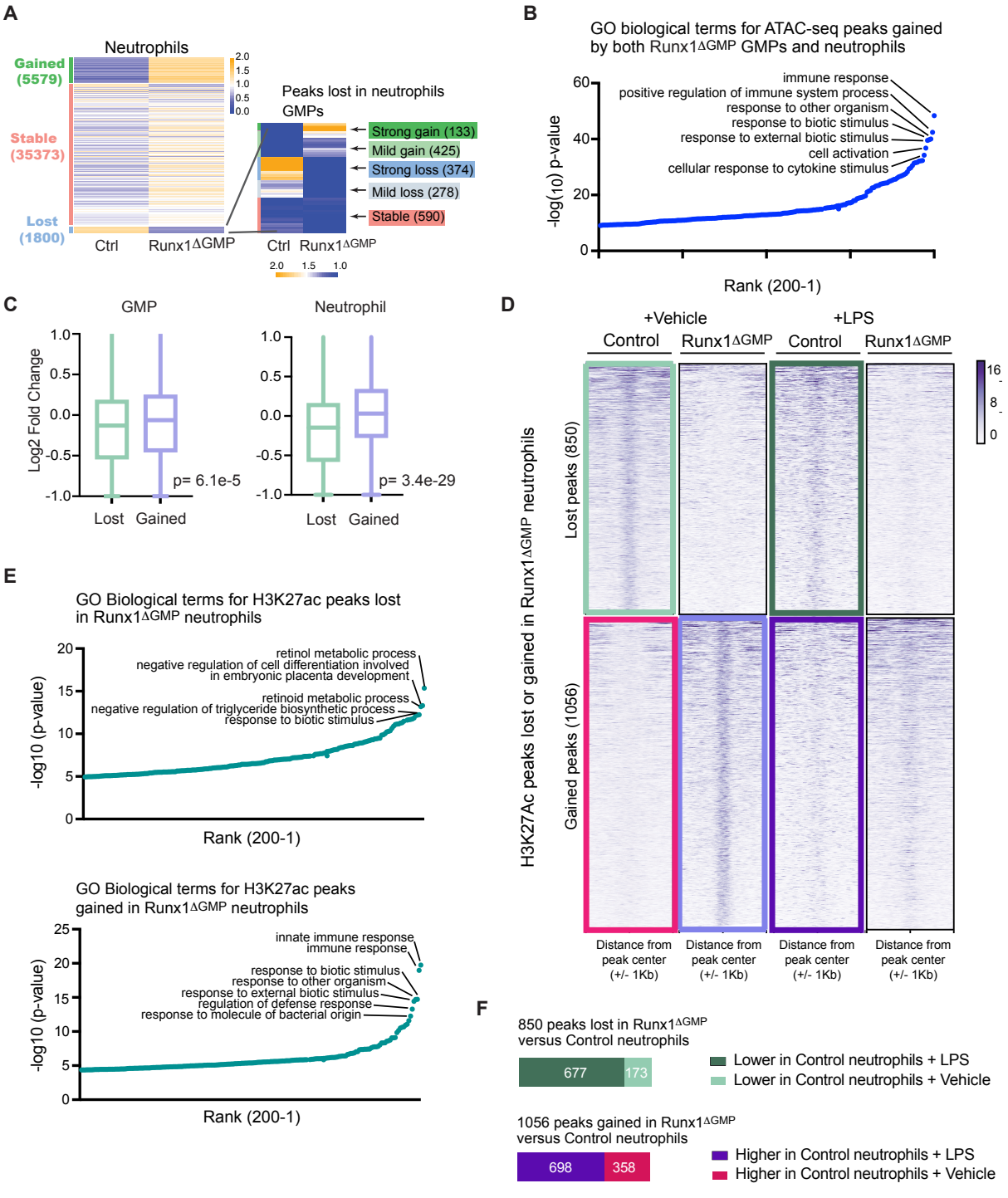
B) Pearson correlation of RNA-seq replicate samples. Scatter plot matrix of replicates. Pearson correlation coefficients were calculated based on transcripts per million (TPM) of genes for each pair of samples. The TPM of genes are plotted for each pair of samples, and the Pearson correlation coefficients are shown for each pair of samples.



Supplemental Figure 3. RUNX1 loss results in elevated levels of key TLR4 signaling molecules.

A) Representative scatter plots of CD14 expression on PB neutrophils from Control and Runx1$^{\Delta GMP}$ mice.

B) CBA analysis demonstrating the effect of CD14 blocking antibody on TNF and CCL3 production by purified BM-derived neutrophils stimulated for 8 hours with vehicle or a high dose (100 ng/mL) of LPS. Mean ± SD, one-way ANOVA plus Tukey multiple comparison test, representative of 3 experiments, a total of 16 mice were analyzed. ***P=0.001, ns= not significant.


A
Neutrophils
Gained (5579)
Stable (35373)
Lost (1800)
Ctrl
Runx1ΔGMP
Peaks lost in neutrophils GMPs
Strong gain (133)
Mild gain (425)
Strong loss (374)
Mild loss (278)
Stable (590)
B
GO biological terms for ATAC-seq peaks gained by both Runx1ΔGMP GMPs and neutrophils
immune response
positive regulation of immune system process
response to other organism
response to biotic stimulus
response to external biotic stimulus
cell activation
cellular response to cytokine stimulus
-log(10) p-value
Rank (200-1)
C
GMP
Neutrophil
Log2 Fold Change
p= 6.1e-5
p= 3.4e-29
Lost
Gained
D
+Vehicle
+LPS
Control
Runx1ΔGMP
H3K27Ac peaks lost or gained in Runx1ΔGMP neutrophils
Lost peaks (850)
Gained peaks (1056)
Distance from peak center (+/- 1Kb)
E
GO Biological terms for H3K27ac peaks lost in Runx1ΔGMP neutrophils
retinol metabolic process
negative regulation of cell differentiation involved in embryonic placenta development
retinoid metabolic process
negative regulation of triglyceride biosynthetic process
response to biotic stimulus
-log10 (p-value)
Rank (200-1)
GO Biological terms for H3K27ac peaks gained in Runx1ΔGMP neutrophils
innate immune response
immune response
response to biotic stimulus
response to other organism
response to external biotic stimulus
regulation of defense response
response to molecule of bacterial origin
F
850 peaks lost in Runx1ΔGMP versus Control neutrophils
677
173
Lower in Control neutrophils + LPS
Lower in Control neutrophils + Vehicle
1056 peaks gained in Runx1ΔGMP versus Control neutrophils
698
358
Higher in Control neutrophils + LPS
Higher in Control neutrophils + Vehicle

Supplemental Figure 4. Loss of RUNX1 activates genes that mediate innate immune responses.

A) Left, heatmap of ATAC-seq signals for Control (Ctrl) and Runx1$^{\Delta GMP}$ neutrophils. Peaks are categorized into Gained, Lost (in Runx1$^{\Delta GMP}$ GMPs), and Stable groups. Right, heatmap of ATAC-seq signal in Control and Runx1$^{\Delta GMP}$ GMPs for peaks lost in Runx1$^{\Delta GMP}$ neutrophils.

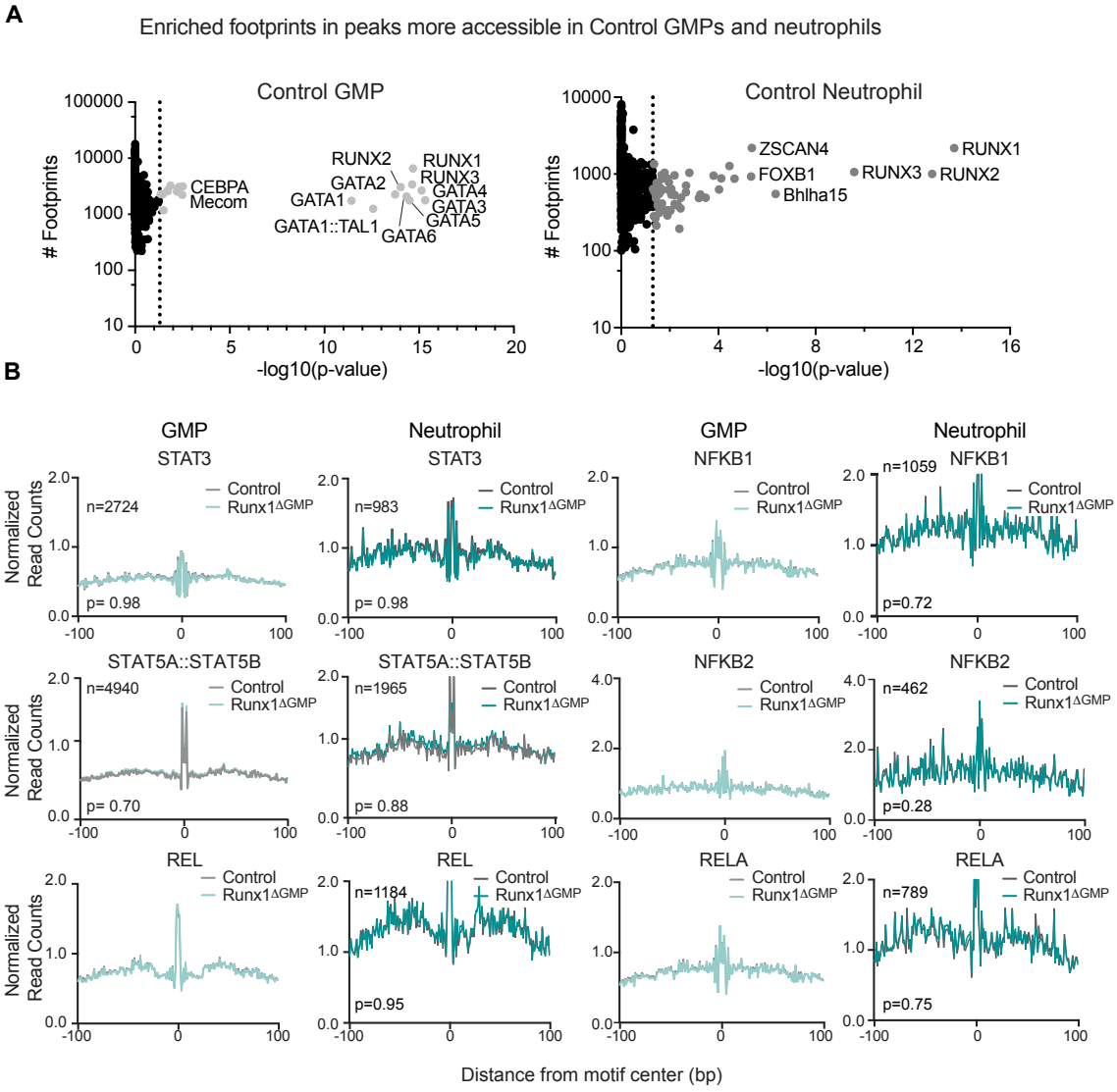
B) GO biological terms for gained ATAC-seq peaks shared by Runx1$^{\Delta GMP}$ GMPs and neutrophils.

C) Box and whisker plots of log2FC in RNA-seq signal relative to Control for genes in Runx1$^{\Delta GMP}$ GMPs and neutrophils that had differentially accessible regions. ATAC-seq peaks were assigned to genes by proximity.

D) Heat maps of H3K27ac signals in Control and Runx1$^{\Delta GMP}$ neutrophils. Light teal box; 850 peaks lost in Runx1$^{\Delta GMP}$ neutrophils (+Vehicle) compared to Control neutrophils (+Vehicle). Dark teal box; 677 peaks lower in Control neutrophils +LPS compared to Control neutrophils +Vehicle. Lavender box: 1056 peaks higher in Runx1$^{\Delta GMP}$ neutrophils +Vehicle compared to Control neutrophils +Vehicle (pink box). Dark purple box; 698 peaks higher in Control neutrophils +LPS compared to Control neutrophils +Vehicle. Scales are normalized to RPGC read counts.

E) Gene Ontology (GO) analysis of genes with H3K27ac lost or gained peaks in Control and Runx1$^{\Delta GMP}$ neutrophils (+Vehicle). The top 200 GO terms are plotted.

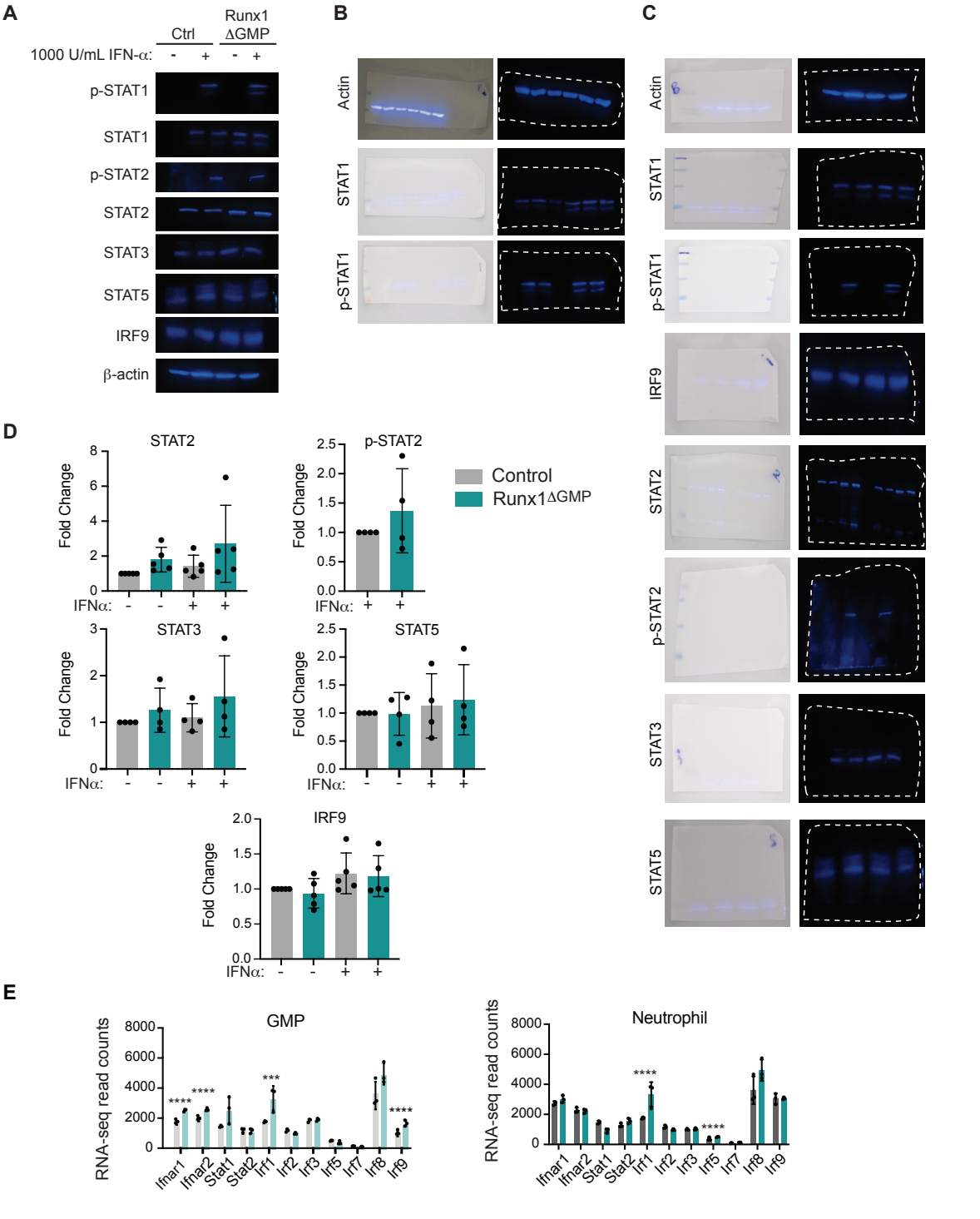
F) Bar graphs depicting the number of peaks lost in Runx1$^{\Delta GMP}$ neutrophils relative to Control neutrophils (top) or number of peaks gained in Runx1$^{\Delta GMP}$ neutrophils relative to Control neutrophils (bottom) from heat maps in panel D. Numbers inside bar graphs depict the number of peaks that were higher or lower in different conditions relative to the peaks lost (top) or peaks gained (bottom).


A
Enriched footprints in peaks more accessible in Control GMPs and neutrophils
Control GMP
Footprints
-log10(p-value)
CEBPA
Mecom
RUNX2
RUNX1
GATA2
RUNX3
GATA1
GATA4
GATA3
GATA1::TAL1
GATA5
GATA6
Control Neutrophil
ZSCAN4
RUNX1
FOXB1
RUNX3
RUNX2
Bhlha15
B
GMP
Neutrophil
STAT3
NFKB1
STAT5A::STAT5B
NFKB2
REL
RELA
Normalized Read Counts
Control
Runx1ΔGMP
n=2724
p= 0.98
n=983
p= 0.98
n=1059
p=0.72
n=4940
p= 0.70
n=1965
p= 0.88
n=462
p=0.28
n=1184
p=0.95
n=789
p=0.75
Distance from motif center (bp)

Supplemental Figure 5. Loss of RUNX1 affects type I IFN signaling.

A) Scatter plots showing enriched TF footprints in regions of chromatin with decreased accessibility in Runx1$^{\Delta GMP}$ neutrophils and GMPs relative to Controls. The number of footprints for each TF at regions of chromatin with decreased accessibility in Runx1$^{\Delta GMP}$ neutrophils and GMPs is displayed on the y-axis for the Control cells. Colored circles indicate $p<0.05$; p-value calculated using biFET.

B) Digital footprint profile plots for selected STAT and NF-κB TFs showing average normalized read counts and p-values calculated with HINT-differential using all peaks in Control and Runx1$^{\Delta GMP}$ cells.


A
Ctrl
Runx1 ΔGMP
1000 U/mL IFN-α: - + - +
p-STAT1
STAT1
p-STAT2
STAT2
STAT3
STAT5
IRF9
β-actin
B
Actin
STAT1
p-STAT1
C
Actin
STAT1
p-STAT1
IRF9
STAT2
p-STAT2
STAT3
STAT5
D
STAT2
p-STAT2
STAT3
STAT5
IRF9
Fold Change
IFNα: - - + +
Control
Runx1ΔGMP
E
GMP
Neutrophil
RNA-seq read counts
Ifnar1
Ifnar2
Stat1
Stat2
Irf1
Irf2
Irf3
Irf5
Irf7
Irf8
Irf9

Supplemental Figure 6. Loss of RUNX1 increases the levels of a subset of type I IFN signaling molecules.

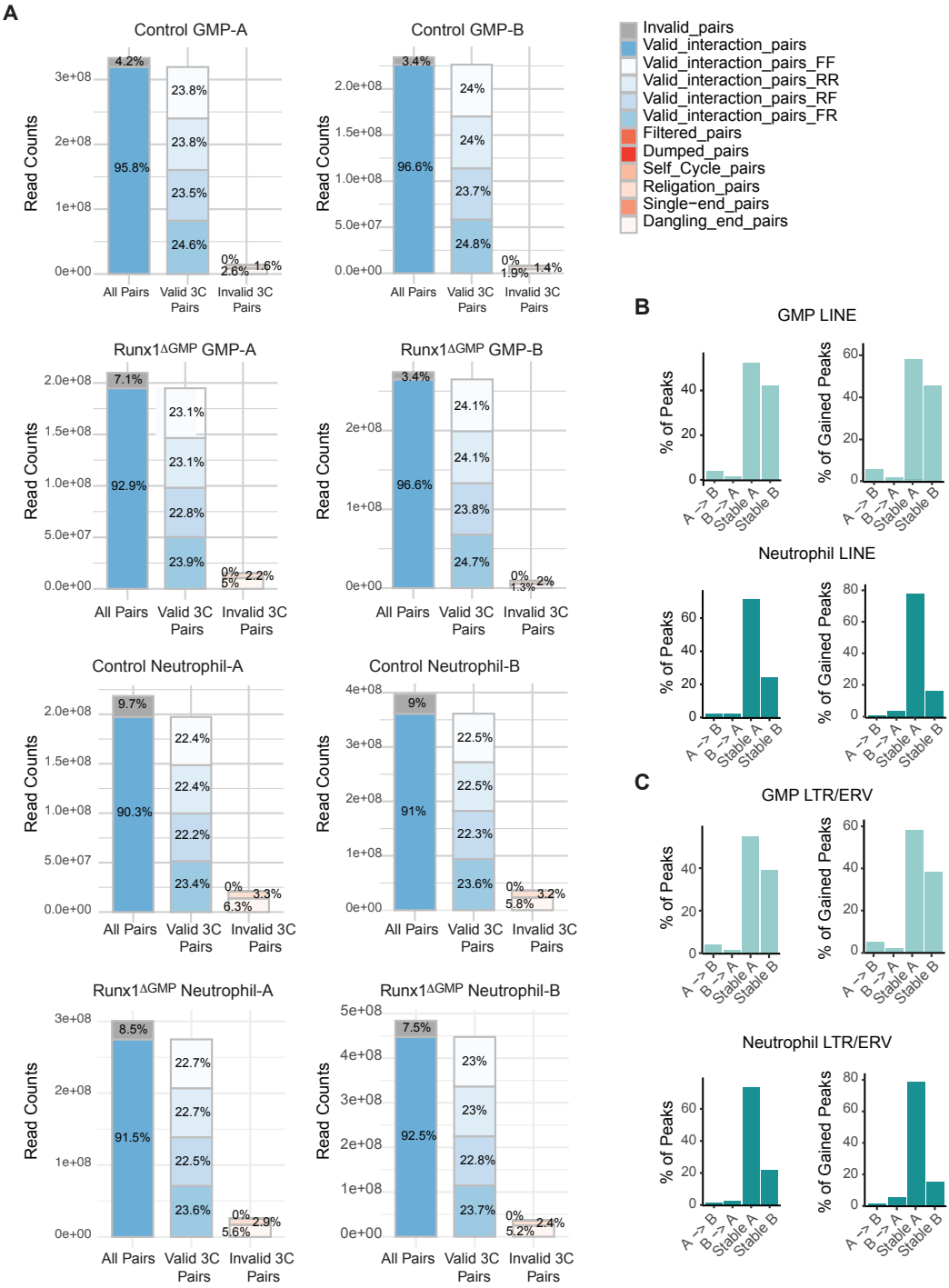
A) Western blot for total STAT, phosphorylated STAT (p-STAT), and IRF9 plus a β-actin control in Control and Runx1$^{\Delta GMP}$ neutrophils in the presence or absence of IFN-α.

B) Original Western blots for STAT1, p-STAT1, and β-actin in Control and Runx1$^{\Delta GMP}$ neutrophils for data in Figure 4F. The edges of each blot are depicted with a dotted line.

C) Original Western blots for data in panel A.

D) Quantification of Western blots for STAT2, p-STAT2, STAT3, STAT5, and IRF9. ANOVA plus Tukey multiple comparison tests; none of the differences except for STAT1 and p-STAT1 shown in Fig. 4G were significant.

E) RNA-seq read counts for genes encoding several components of the type I IFN signaling pathway including receptors (*Ifnar1, Ifnar2*), STATs, and IRFs in GMPs and neutrophils. Mean ± SD, unpaired two-tailed t-test. ***P=0.001, ****P<0.0001.

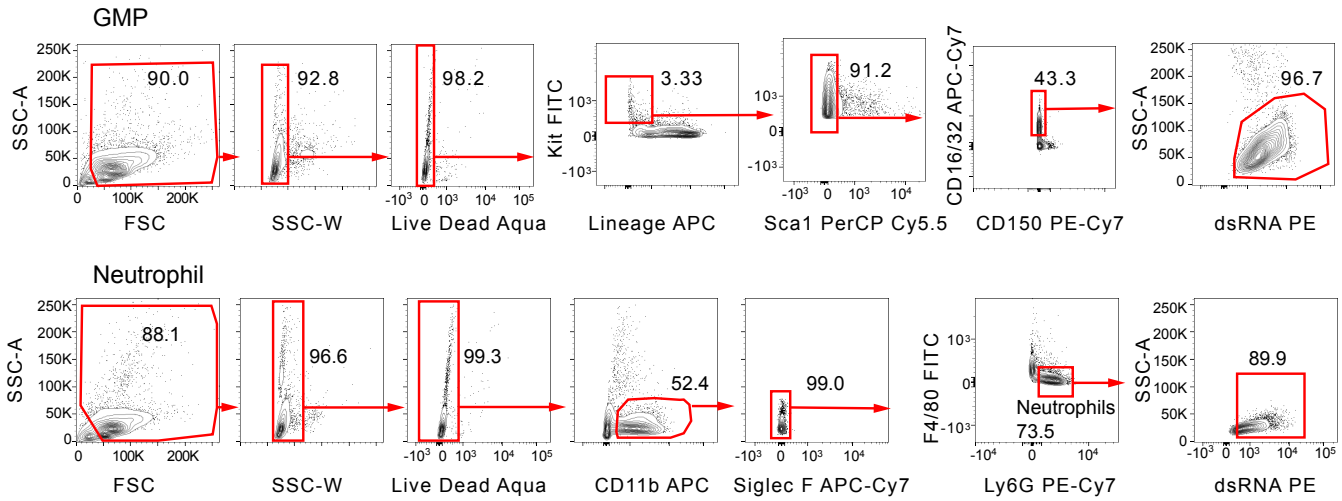

A
Control GMP-A
Control GMP-B
Runx1ΔGMP GMP-A
Runx1ΔGMP GMP-B
Control Neutrophil-A
Control Neutrophil-B
Runx1ΔGMP Neutrophil-A
Runx1ΔGMP Neutrophil-B
Read Counts
All Pairs
Valid 3C Pairs
Invalid 3C Pairs
Invalid_pairs
Valid_interaction_pairs
Valid_interaction_pairs_FF
Valid_interaction_pairs_RR
Valid_interaction_pairs_RF
Valid_interaction_pairs_FR
Filtered_pairs
Dumped_pairs
Self_Cycle_pairs
Religation_pairs
Single−end_pairs
Dangling_end_pairs
B
GMP LINE
Neutrophil LINE
C
GMP LTR/ERV
Neutrophil LTR/ERV
% of Peaks
% of Gained Peaks
A −> B
B −> A
Stable A
Stable B

Supplemental Figure 7: Hi-C quality and distribution of retroelements in compartments A/B.

A) Number and percentage of valid read pairs aligned to restriction fragments for Hi-C samples.

B) Percentage of all (left) and gained (right) ATAC-seq peaks in LINE retroelements distributed in chromatin compartments A/B in GMPs (top) and neutrophils (bottom).

C) Percentage of all (left panel) and gained (right panel) ATAC-seq peaks in LTR/ERV retroelements distributed in chromatin compartments A/B in GMPs (top) and neutrophils (bottom).



Supplemental Figure 8. Gating strategy for analyzing dsRNA in GMPs and neutrophils. Quantile contour FACS plots depicting gating strategies. Numbers on the x and y-axes are indicated on the first plot on the left, and unless changed are not depicted on plots to the right of the preceding plot.

SUPPLEMENTAL TABLES

Supplemental Table 1: List of all antibodies used for flow cytometry, western blot analysis, H3K27ac ChIP-seq, and RUNX1 CUT&RUN. For each antibody, the clone, fluorophore, dilution, manufacturer, and antibody registry number (or manufacturer's catalog number when no RRID number is available) are provided.

Antibody	Clone	Fluorophore	Supplier	RRID #	Dilution
c-Kit	2B8	FITC	Biolegend	AB_313215	1:200
CD3e	145-2C11	APC	Biolegend	AB_312677	1:200
CD3e	145-2C11	PE	Biolegend	AB_312672	1:200
CD4	RM4-5	BV 421	Biolegend	AB_2563052	1:200
CD8	53-6.7	PE-Cy7	Biolegend	AB_312761	1:200
CD14	Sa14-2	PE	Biolegend	AB_940584	1:200
CD14	Sa14-2	Unconjugated	Biolegend	AB_940588	1:200
CD14	4C1/CD14	Unconjugated	BD Biosciences	AB_396926	
CD16/32	93	APC-Cy7	Biolegend	AB_2104158	1:200
CD19	1D3	APC	Thermo Fisher	AB_1659676	1:200
CD19	1D3	PE	Biolegend	AB_2629817	1:200
CD45.1	A20	APC-Cy7	Biolegend	AB_313505	1:200
CD45.1	A20	PE-Cy7	Thermo Fisher	AB_469629	1:200
CD45.2	104	FITC	Thermo Fisher	AB_465062	1:200
B220	RA3-6B2	APC	Biolegend	AB_312997	1:200
CD150	TC15-12F12.2	PE-Cy7	Biolegend	AB_439797	1:200
F4/80	BM8	FITC	Biolegend	AB_893500	1:200
Gr-1	RB6-8C5	APC	Biolegend	AB_313377	1:200
Gr-1	RB6-8C5	PerCP-Cy5.5	Thermo Fisher	AB_906247	1:200
IFNAR-1	MAR1-5A3	Unconjugated (blocking Ab)	Bio X Cell	AB_2687723	
Ly-6G	1A8	APC	Biolegend	AB_2227348	1:200
Ly-6G	1A8	PE- Cy7	BD Biosciences	AB_1727562	1:200
Ly-6G	1A8	APC-Cy7	BD Biosciences	AB_1727561	1:200
Mac-1	M1/70	APC	Biolegend	AB_312795	1:200
Mac-1	M1/70	APC-Cy7	BD Biosciences	AB_396772	1:200
Nk1.1	PK136	APC	Biolegend	AB_313397	1:200
Rat IgG2a, κ Isotype Ctrl	RTK2758	PE	Biolegend	AB_326530	1:200
Mouse IgG1, κ Isotype Ctrl	monoclonal antibody	Unconjugated	BD Biosciences	AB_10050442	
Sca-1	D7	PerCP-Cy5.5	Thermo Fisher	AB_914372	1:200
Sca-1	E13-161.7	PE	Biolegend	AB_756193	1:200

Siglec F	E50-2440	PE	BD Biosciences	AB_394341	1:200
Siglec F	E50-2440	APC-Cy7	BD Biosciences	AB_2732831	1:200
Streptavidin		PerCP-Cy5.5	Biolegend	AB_2716577	1:200
Ter119	TER-119	APC	Biolegend	AB_313713	1:200
TNF-alpha	MP6-XT22	Pacific Blue	Biolegend	AB_893639	1:100
dsRNA	9D5	Unconjugated	Absolute Antibody	Cat. No: Ab00458-23.0	1:500
Donkey anti-rabbit IgG	Poly4064	PE	Biolegend	AB_2563484	1:400
Anti-Histone H3 (acetyl K27) antibody	polyclonal antibody		Abcam	AB_2118291 Cat. No: ab4729	
RUNX1 / AML1 antibody - ChIP Grade	polyclonal antibody		Abcam	AB_2184205 Cat. No: ab23980	
β-Actin	C4		Santa Cruz Biotechnology	AB_2714189 Cat. No: sc-47778 HRP	1:1000
GAPDH	D4C6R		Cell Signaling Technology	AB_2756824 Cat. No: 97166	1:1000
STAT1			Cell Signaling Technology	AB_2198300 Cat. No: 9172	1:1000
pSTAT1 (Tyr701)	58D6		Cell Signaling Technology	AB_561284 Cat. No: 9167	1:1000
STAT2	D9J7L		Cell Signaling Technology	AB_2799824 Cat. No: 72604	1:1000
pSTAT2 (Tyr689)			Millipore	AB_2198439 Cat. No: 07-224	1:1000
STAT3			Cell Signaling Technology	AB_2629499 Cat. No: 12640	1:1000
pSTAT3 (Tyr705)	D3A7		Cell Signaling Technology	AB_2491009 Cat. No: 9145	1:1000
STAT5	D2O6Y		Cell Signaling Technology	AB_2737403 Cat. No: 94205	1:1000
pSTAT5 (Y694)			Cell Signaling Technology	AB_2315225 Cat. No: 9351	1:1000
IRF9	D9I5H		Cell Signaling Technology	AB_2798964 Cat. No: 28845	1:1000
DAPI			Thermo Fisher	Cat. No: D1306	
LIVE/DEAD Fixable Aqua			Thermo Fisher	Cat. No: L34957	

Supplemental Table 2: Table showing the differentially expressed genes in Runx1$^{\Delta GMP}$ and control GMPs and neutrophils.

Supplemental Table 3: GO terms for peaks gained in Runx1$^{\Delta GMP}$ GMPs and neutrophils.

Supplemental Table 4: Table showing differential activity of transcription factors calculated using HINT-differential, enriched transcription factor footprints at regions of chromatin with increased accessibility calculated using BiFET v.1.16.0 software, and enriched motifs under ATAC-seq peaks.

Supplemental Table 5: RUNX1 CUT&RUN peaks in Control GMPs.

Supplemental Table 6: TE subfamilies that were enriched or depleted in the immunoprecipitant of Runx1$^{\Delta GMP}$ neutrophils.

SUPPLEMENTAL REFERENCES

Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. 2013. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* **10**: 1213-1218.

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM et al. 2021. Twelve years of SAMtools and BCFtools. *Gigascience* **10**.

Gusmao EG, Allhoff M, Zenke M, Costa IG. 2016. Analysis of computational footprinting methods for DNase sequencing experiments. *Nat Methods* **13**: 303-309.

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

Hiller M, Agarwal S, Notwell JH, Parikh R, Guturu H, Wenger AM, Bejerano G. 2013. Computational methods to detect conserved non-genic elements in phylogenetically isolated genomes: application to zebrafish. *Nucleic Acids Res* **41**: e151.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357-359.

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

Li Z, Schulz MH, Look T, Begemann M, Zenke M, Costa IG. 2019. Identification of transcription factor binding sites using ATAC-seq. *Genome Biol* **20**: 45.
Martin M. 2011. Cutadapt Removes Adapter Sequences From High-Throughput Sequencing Reads. *EMBnetjournal* **17**.
McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G. 2010. GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol* **28**: 495-501.
Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841-842.
Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dundar F, Manke T. 2016. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**: W160-165.
Shin B, Zhou W, Wang J, Gao F, Rothenberg EV. 2022. Runx factors launch T-cell and innate lymphoid programs via direct and gene network-based mechanisms. *bioRxiv*.
Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, Prins P. 2015. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* **31**: 2032-2034.
Tavora B, Mederer T, Wessel KJ, Ruffing S, Sadjadi M, Missmahl M, Ostendorf BN, Liu X, Kim JY, Olsen O et al. 2020. Tumoural activation of TLR3-SLIT2 axis in endothelium drives metastasis. *Nature* **586**: 299-304.
Yang WR, Ardeljan D, Pacyna CN, Payer LM, Burns KH. 2019. SQuIRE reveals locus-specific regulation of interspersed repeat expression. *Nucleic Acids Res* **47**: e27.
Youn A, Marquez EJ, Lawlor N, Stitzel ML, Ucar D. 2019. BiFET: sequencing Bias-free transcription factor Footprint Enrichment Test. *Nucleic Acids Res* **47**: e11.
Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W et al. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**: R137.