

Control of inducible gene expression links cohesin to hematopoietic progenitor self-renewal and differentiation

Sergi Cuartero¹, Felix D. Weiss^{1,15}, Gopuraja Dharmalingam^{3,15}, Ya Guo^{1,15}, Elizabeth Ing-Simmons^{1,2,12,15}, Silvia Masella⁴, Irene Robles-Rebollo¹, Xiaolin Xiao³, Yi-Fang Wang³, Iros Barozzi^{4,5}, Dounia Djeghloul¹, Mariane T. Amano^{1,13}, Henri Niskanen⁶, Enrico Petretto^{3,7}, Robin D. Dowell⁸, Kikuë Tachibana^{9,14}, Minna U. Kaikkonen^{10,6}, Kim A. Nasmyth⁹, Boris Lenhard², Gioacchino Natoli^{10,11}, Amanda G. Fisher¹ and Matthias Merkenschlager^{1*}

Cohesin is important for 3D genome organization. Nevertheless, even the complete removal of cohesin has surprisingly little impact on steady-state gene transcription and enhancer activity. Here we show that cohesin is required for the core transcriptional response of primary macrophages to microbial signals, and for inducible enhancer activity that underpins inflammatory gene expression. Consistent with a role for inflammatory signals in promoting myeloid differentiation of hematopoietic stem and progenitor cells (HSPCs), cohesin mutations in HSPCs led to reduced inflammatory gene expression and increased resistance to differentiation-inducing inflammatory stimuli. These findings uncover an unexpected dependence of inducible gene expression on cohesin, link cohesin with myeloid differentiation, and may help explain the prevalence of cohesin mutations in human acute myeloid leukemia.

Cohesin is a multiprotein complex that cooperates with the sequence-specific DNA binding protein CTCF in forming key features of 3D genome organization such as topologically associated domains (TADs), contact domains and chromatin loops. These features spatially compartmentalize genes and enhancers in interphase^{1–7} and are believed to facilitate preferential interactions between promoters and enhancers located in the same domain^{5,6,8–11}. Removal of architectural proteins, CTCF binding sites or domain boundaries weakens insulation between domains, thus exposing genes to regulatory elements in neighboring domains and potentially perturbing gene regulation^{5,7,12–17}. However, with a few notable exceptions of specific deregulated genes^{12–17} or DNA damage responses due to essential cohesin functions in the cell cycle^{18,19}, loss of cohesin or CTCF has shown limited impact on transcriptional control^{2,4}, chromatin marks or enhancer states^{3,7,20}. Even the complete removal of cohesin or CTCF, which abrogates the formation of CTCF–cohesin-based chromatin loops and substantially weakens TADs^{2,4}, does not result in clear gene regulatory phenotypes. This finding raised concerns as to whether current models overstate the significance of spatial genome compartmentalization for gene regulation. However, it is still unclear to what extent such limited impact of cohesin on gene regulation also applies to inducible responses,

such as the core myeloid inflammatory gene expression program^{21–25}. Here, hundreds of genes and thousands of gene regulatory elements are rapidly activated in a highly coordinated fashion, likely imposing an extraordinary level of regulatory requirements^{21–25}.

In addition to its role in genome compartmentalization in interphase, cohesin is essential for genome integrity in cycling cells²⁶. Because of this role, it may seem counterintuitive that cohesin mutations are frequently found in cancers, including acute myeloid leukemia^{27–29} (AML). However, partial loss of cohesin is compatible with cell proliferation and drives increased self-renewal of HSPCs^{30–34}. As increased self-renewal can facilitate leukemic transformation, it is important to elucidate the mechanisms that link cohesin to pathways that regulate the balance between self-renewal and differentiation. Defining these mechanisms in HSPCs with reduced cohesin function is complicated, as it is unclear whether changes in gene expression and chromatin state are cause or consequence of increased self-renewal and reduced differentiation^{30–34}.

To address these issues, we engineered mature, nonproliferating macrophages that can be depleted of cohesin in an inducible fashion after a normal history of differentiation. We found cohesin to be critically required for inflammatory gene expression in macrophages and HSPCs, and in primary human AML cells.

¹Lymphocyte Development Group, Epigenetics Section, MRC London Institute of Medical Sciences, Institute of Clinical Sciences, Faculty of Medicine, Imperial College London, London, UK. ²Computational Regulatory Genomics Group, Integrative Biology Section, MRC London Institute of Medical Sciences, Institute of Clinical Sciences, Faculty of Medicine, Imperial College London, London, UK. ³MRC London Institute of Medical Sciences, Institute of Clinical Sciences, Faculty of Medicine, Imperial College London, London, UK. ⁴Department of Experimental Oncology, European Institute of Oncology, Milan, Italy. ⁵Department of Surgery and Cancer, Department of Medicine, Imperial College London, London, UK. ⁶A. I. Virtanen Institute for Molecular Sciences, University of Eastern Finland, Kuopio, Finland. ⁷Cardiovascular and Metabolic Disorders, Duke-NUS Medical School, Singapore, Singapore. ⁸BioFrontiers Institute and Department of Molecular, Cellular and Developmental Biology, University of Colorado, Boulder, CO, USA. ⁹Department of Biochemistry, University of Oxford, Oxford, UK. ¹⁰Humanitas Clinical and Research Center, Milan, Italy. ¹¹Humanitas University, Milan, Italy. ¹²Present address: Max Planck Institute for Molecular Biomedicine, Muenster, Germany. ¹³Present address: Hospital Sírio-Libanês, Sao Paulo, Brazil. ¹⁴Present address: Institute of Molecular Biotechnology of the Austrian Academy of Sciences, Vienna Biocenter, Vienna, Austria. ¹⁵These authors contributed equally: Felix D. Weiss, Gopuraja Dharmalingam, Ya Guo, Elizabeth Ing-Simmons. *e-mail: matthias.merkenschlager@lms.mrc.ac.uk

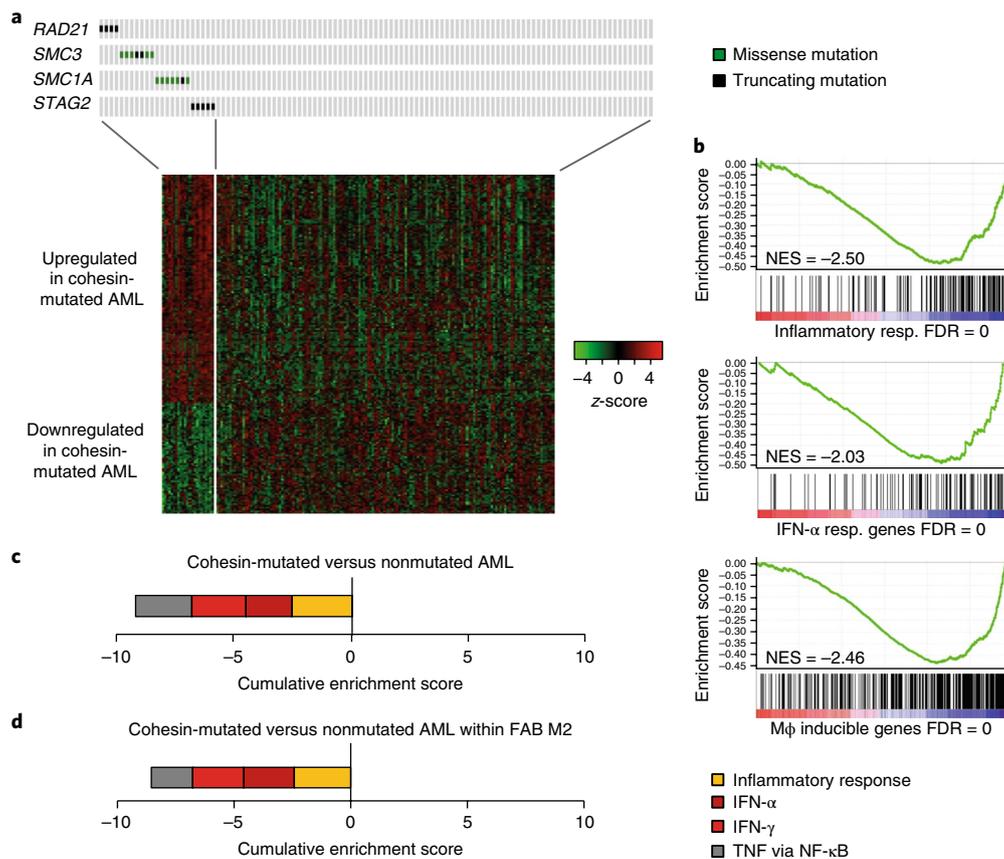


Fig. 1 | Cohesin links inflammation and cancer. **a**, Analysis of RNA-seq data from 173 primary human TCGA cases of AML, 23 of which had missense or truncating mutations in the cohesin genes *RAD21*, *SMC3*, *SMC1A* or *STAG2*, showed significant upregulation of 131 genes and downregulation of 63 genes in cohesin-mutated AML compared to 150 cases of AML without cohesin mutations (adjusted $P < 0.05$). Differential expression was analyzed and z-scores were calculated from normalized expression values as detailed in Methods. **b**, GSEA of 23 TCGA cases of AML with and 150 TCGA cases of AML without cohesin mutations. Top, inflammatory response (resp.) genes (normalized enrichment score (NES) = -2.5, FDR = 0; see Methods); middle, IFN- α response (NES = -2.03, FDR = 0); bottom, the human orthologs of inducible mouse macrophage (M ϕ) genes²³ in AML with cohesin mutations (NES = -2.46, FDR = 0). **c**, Cumulative NES of significantly enriched or depleted gene sets for the pathways inflammatory response, IFN- α , IFN- γ and TNF signaling via NF- κ B (FDR < 0.05) in 23 TCGA cases of AML with cohesin mutations compared to 150 TCGA cases of AML without cohesin mutations. **d**, Cumulative NES of significantly enriched or depleted gene sets for the pathways inflammatory response, IFN- α , IFN- γ and TNF signaling via NF- κ B (FDR < 0.05) in TCGA cases of AML of FAB subtype M2, comparing 10 cases of FAB subtype M2 AML with cohesin mutations and 27 cases of FAB subtype M2 AML without cohesin mutations.

As inflammatory signals regulate HSPC self-renewal and myeloid differentiation^{35–42}, our findings provide a mechanistic link between cohesin, inflammation and AML.

Results

Cohesin controls inflammatory gene expression in AML. To explore the role of cohesin in AML, we examined the correlation between cohesin mutations and gene expression by analyzing RNA sequencing (RNA-seq) data for 173 primary AML samples compiled by The Cancer Genome Atlas²⁸ (TCGA). Twenty-three had missense or truncating mutations in the genes encoding the cohesin subunits *RAD21*, *SMC3*, *SMC1A* or *STAG2* (Fig. 1a). Gene set enrichment analysis (GSEA) showed that inflammatory genes were the most strongly downregulated gene set in AML with cohesin mutations, closely followed by interferon (IFN)-responsive genes (Fig. 1b; false-discovery rate (FDR) values of <0.001 were rounded to 0). Genes involved in the IFN- α , IFN- γ and tumor necrosis factor (TNF) signaling pathways were similarly downregulated in cohesin-mutated AML (Fig. 1c).

AML samples of different French-American-British (FAB) subtypes⁴³ had characteristic patterns of inflammatory gene expression

(data not shown). Among the 37 AML samples classified by TCGA as FAB M2, 10 had cohesin mutations, providing sufficient power to compare gene expression within this subtype. GSEA identified inflammatory genes and genes involved in the IFN- α and IFN- γ pathways as the top three downregulated gene sets in FAB M2 AML with cohesin mutations (FDR = 0; Fig. 1d), linking reduced inflammatory gene expression to impaired cohesin function rather than AML subtype. This analysis of TCGA samples suggests a previously unrecognized role for cohesin in the regulation of inflammatory genes in human AML.

Inducible genes are sensitive to cohesin dosage. We next analyzed the impact of cohesin on inflammatory gene expression in primary mouse macrophages. These mature, quiescent myeloid cells are suitable for mechanistic studies of gene expression^{21–25}. To uncouple cohesin deletion from myeloid differentiation, we allowed myeloid progenitors to differentiate into mature macrophages and subsequently deleted the gene encoding the essential cohesin subunit *RAD21* (Supplementary Fig. 1a,b). *loxP*-flanked *Rad21* alleles were removed within 24 h of inducible Ert2Cre activation by 4-hydroxytamoxifen (4-OHT), and *RAD21* protein expression declined

gradually over 2–3 d (Supplementary Fig. 1b–e). This approach allowed homozygous cohesin deletion, as the cell cycle functions of cohesin are essential in cycling but not in quiescent cells¹⁷ (Supplementary Fig. 1d,f). The use of quiescent cells also precludes any selective expansion of immature cells (as seen in HSPCs with reduced cohesin function^{30–34}) and thereby enables like-for-like comparisons of gene expression and chromatin state between control and cohesin-deficient cells.

We used RNA-seq to profile gene expression in *Rad21*-deleted macrophages containing less than 15% residual RAD21 protein (Fig. 2a). As expected^{24,7,17}, the overall impact on gene expression was limited. Approximately 10% of constitutively expressed genes were up- or downregulated (adjusted $P < 0.05$; Fig. 2b). However, genes that are inducible by inflammatory signals²³ (Fig. 2b) were more severely affected by the loss of cohesin than constitutively expressed genes. Over 50% of inducible genes were deregulated at baseline (Fig. 2c).

In macrophages, activation of Toll-like receptor 4 (TLR4) by the bacterial cell wall component lipopolysaccharide (LPS) triggers a program of inducible gene expression²³, but not proliferation. Transcription factor and cytokine genes are activated early, and cytokines (notably IFN- β) trigger the auto- and paracrine induction of secondary response genes²³ (Fig. 2b). This program was curtailed in *Rad21*-deleted macrophages. The frequency of deregulated inducible genes progressively increased with time after LPS (Fig. 2c,d). Deregulated inducible genes²³ included genes classified as IFN-dependent²⁴ and, albeit to a lesser extent, IFN-independent genes²⁴. Specifically, 88% of IFN-dependent inducible genes and 68% of IFN-independent inducible genes were deregulated at adjusted $P < 0.05$ 8 h after LPS treatment.

Deregulation of inducible genes was profound not only in terms of the frequency of deregulated genes, but also in terms of the magnitude of the change (Fig. 2e). For example, 32% of inducible genes but only 5% of constitutive genes were deregulated fourfold or more 8 h after LPS stimulation (data not shown). The transcription factor NF- κ B is a key regulator of inducible genes in macrophages^{21–25}. After TLR4 activation by LPS, NF- κ B was predominantly nuclear in both *Rad21*-deleted and wild-type macrophages (Supplementary Fig. 1g). The deregulation of inducible genes was therefore not explained by unresponsiveness of *Rad21*-deficient macrophages to LPS.

Inducible macrophage genes are mostly proinflammatory^{23–25}, and genes related to the inflammatory response were predominantly downregulated in *Rad21*-deleted macrophages (FDR = 0.0; Fig. 2f). These changes in gene expression affected the secretion of inducible cytokines by *Rad21*-deleted macrophages. Of 26 LPS-responsive cytokines tested, 16 were deregulated, and 13 were decreased, including IFN- β , IL-6 and TNF (Supplementary Fig. 1h). Analysis of heterozygous *Rad21*^{+/-} macrophages, which retained $77 \pm 6\%$ of RAD21 protein expression compared to *Rad21*^{+/+} macrophages (data not shown), indicated that partially reduced cohesin function was sufficient to impair inducible gene expression (Fig. 2g).

Immediate impact of acute cohesin depletion. Because RAD21 protein abundance declined gradually after genetic deletion of *Rad21* (Supplementary Fig. 1c), cells were in a cohesin-depleted state for 24 to 48 h before RNA-seq analysis. It was therefore unclear to what extent inducible gene expression was under the direct control of cohesin. To address this question, we developed an experimental system for acute cohesin depletion based on *Rad21* alleles engineered by insertion of cleavage sites for the tobacco etch virus (TEV) protease into the endogenous locus⁴⁴. Fetal liver cells expressing TEV-cleavable RAD21 (RAD21-TEV) as their sole source of RAD21 protein were transduced with a cytoplasmic TEV-ERT2 fusion construct and differentiated into mature, quiescent macrophages. Addition of the ERT2 ligand 4-OHT released

TEV-ERT2 to the nucleus, and RAD21-TEV protein was rapidly degraded (Fig. 3a). RNA-seq 8 h after 4-OHT showed that acute depletion of cohesin resulted in the deregulation of 1,016 genes (adjusted $P < 0.05$ based on DEseq2 analysis of 3 RNA-seq replicates). The majority of these genes (557 of 1,016, or 55%) were also deregulated by *Rad21* deletion ($P < 2.22 \times 10^{-16}$, odds ratio = 2.65, Fisher's exact test). Acute depletion of cohesin deregulated a significantly higher fraction of inducible than constitutive genes (Fig. 3b, $P < 2.22 \times 10^{-16}$, odds ratio = 4.44). Deregulated genes were enriched for terms including signaling (adjusted $P = 1.33 \times 10^{-8}$), inflammatory response (adjusted $P = 7.93 \times 10^{-7}$), immune system (adjusted $P = 9.23 \times 10^{-6}$) and inflammation mediated by chemokine and cytokine signaling (adjusted $P = 2.44 \times 10^{-8}$). A 2-h pulse of LPS further deregulated inducible genes in 4-OHT-treated RAD21-TEV macrophages ($P < 2.22 \times 10^{-16}$, odds ratio = 3.54; Fig. 3b). Inducible genes²³ deregulated immediately after RAD21-TEV cleavage included both IFN-dependent²⁴ and IFN-independent ones (Fig. 3b). Inflammatory response genes were preferentially downregulated (Fig. 3c). Most inducible genes²³ that were downregulated by acute cohesin depletion were also downregulated by *Rad21* deletion (44 of 69, 64%, at baseline, $P < 0.0005$, odds ratio = 2.60; and 99 of 127, 78%, after 2 h LPS, $P < 9.17 \times 10^{-10}$, odds ratio = 3.94, Fisher's exact test). Transcripts downregulated immediately after cohesin cleavage in RAD21-TEV macrophages included regulators of inducible gene expression, such as transcription factors (*Fos*, *Jun*, *Irf2*, *Myc*, *Ets2*, *Prdm1/Blimp1*, *Egr2*, *Cebpa*, *Cebpb*), inflammatory cytokines (*Il1b*), chemokines (*Ccl3*, *Ccl7*, *Ccl9*), chemokine receptors (*Ccr1*, *Ccr3*, *Ccr5*) and receptors for inflammatory mediators (*Ifnar1*, *Ifnar2*, *Ifngr1*). Hence, inducible genes were under the immediate control of cohesin.

Restricted enhancer dynamics in cohesin-deficient macrophages.

The macrophage enhancer landscape is dynamically reconfigured in response to activation^{21,22}. Constitutive, activation-inducible and activation-repressed enhancers have been characterized on the basis of acetylation of Lys27 of histone H3 (H3K27ac), monomethylation of Lys4 of histone H3 (H3K4me1) and binding of the transcription factor PU.1 at promoter-distal sites²². We found that *Rad21* deletion did not affect H3K27ac at the great majority (97.2%) of constitutively active enhancers²² (DEseq2 adjusted $P < 0.05$; Fig. 4). In contrast, H3K27ac was broadly deregulated at LPS-inducible enhancers²² (2.6% of constitutive versus 24.8% of inducible enhancers; Fig. 4 and Supplementary Fig. 2a) and LPS-repressed enhancers²² (15.6%, $P < 0.05$, Supplementary Fig. 2a). In addition to H3K27ac, active enhancers are characterized by increased chromatin accessibility and enhancer transcription. We assessed enhancer accessibility by ATAC (assay for transposase-accessible chromatin)-seq (Supplementary Fig. 2b) and enhancer transcription by GRO (global run-on)-seq (Supplementary Fig. 2c,d) and confirmed that the activation of inducible enhancers was impaired in cohesin-deficient macrophages. We conclude that cohesin controls inducible gene expression and enhancer dynamics in macrophages.

Genomic organization of deregulated genes and enhancers.

Deregulated inducible genes²³ were enriched near deregulated inducible enhancers²² (adjusted $P < 0.005$ by nearest neighbor analysis, odds ratio = 2.11 in resting macrophages; adjusted $P = 4.05 \times 10^{-6}$, odds ratio = 1.70 after 6 h LPS for enhancers and 8 h LPS for transcripts). Deregulated genes and enhancers were significantly enriched within the same TADs (adjusted $P = 2.70 \times 10^{-11}$, odds ratio = 7.23 for H3K27ac deregulated enhancers; adjusted $P = 4.47 \times 10^{-43}$, odds ratio = 5.02 for GRO-seq deregulated enhancers, Supplementary Fig. 3a). Coherence between gene expression and enhancer states (Supplementary Fig. 3a) is illustrated by domains that harbor downregulated enhancers and clusters of downregulated chemokine genes⁴⁵ (Supplementary Fig. 3b). To assess LPS-induced

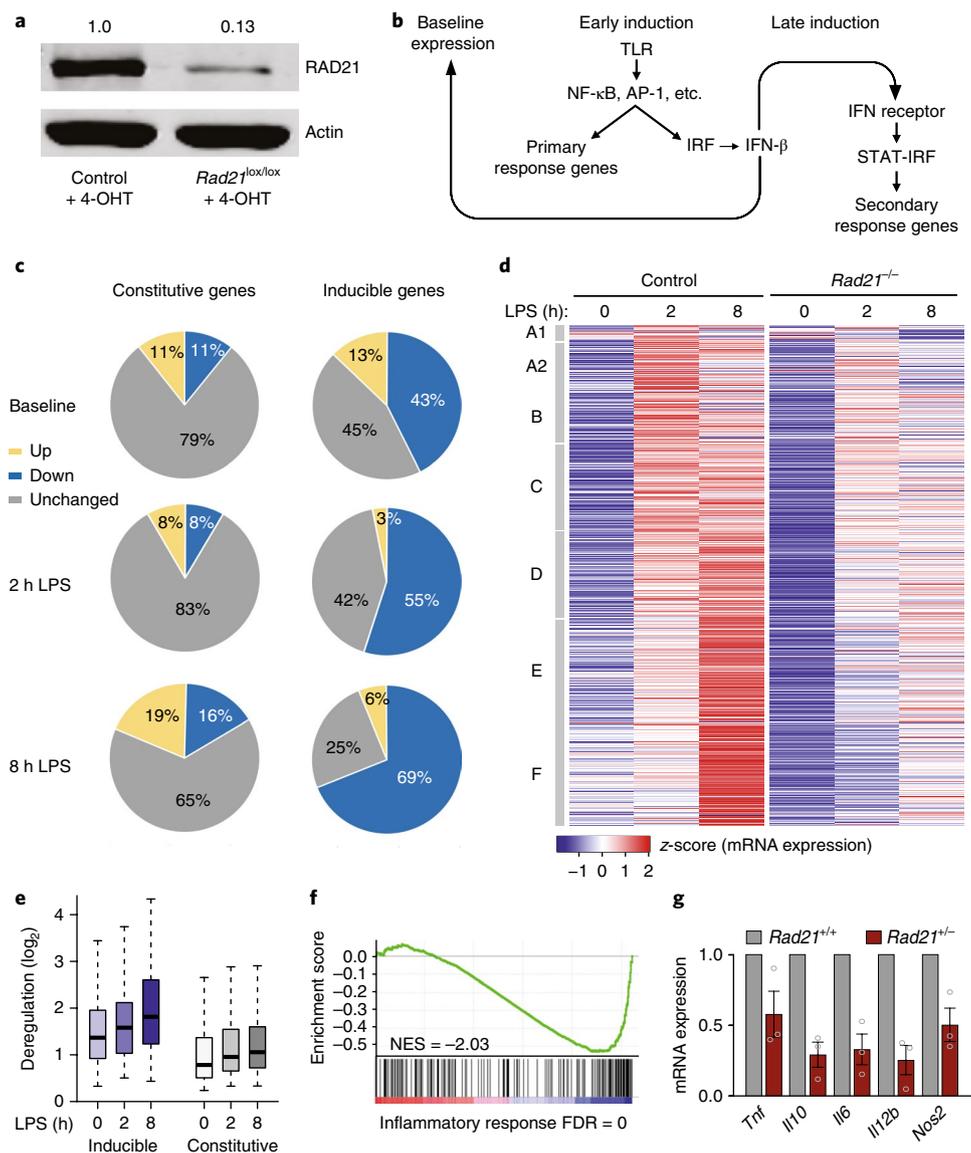


Fig. 2 | Cohesin promotes inducible gene expression. **a**, Immunoblot analysis of RAD21 protein expression in mature macrophages after *Rad21* deletion (mean 13.6% of control, 13 biological replicates). **b**, Regulation of inducible genes at baseline and in response to macrophage activation^{21–25}. **c**, DESeq2 analysis of RNA-seq data was used to determine the fraction of constitutively expressed ($n = 10,780$) and LPS-inducible²³ ($n = 560$) genes deregulated in *Rad21*^{-/-} macrophages at baseline and after 2 or 8 h of LPS stimulation ($P < 0.05$, 3 biological replicates per genotype and time point). **d**, Heat map of inducible gene expression by control (left) and *Rad21*-deleted macrophages (right) at 0, 2 and 8 h after LPS. Inducible gene classes²³ are indicated on the left. Average of 3 biological replicates. **e**, Extent of deregulation of inducible²³ versus constitutive genes in *Rad21*-deleted compared to wild-type macrophages at baseline and after LPS (\log_2 , irrespective of direction). Box plots are representative of 3 biological RNA-seq replicates per genotype and condition and show the median and lower and upper quartiles. Whiskers show the maximum and minimum data points up to 1.5 times the interquartile range. **f**, GSEA of inflammatory response genes in *Rad21*-deleted macrophages (normalized enrichment score (NES) = -2.03, FDR = 0). **g**, Quantitative RT-PCR of inflammatory gene expression 8 h after LPS stimulation of *Rad21*^{+/-} macrophages (mean \pm s.e.m. of 3 biological replicates).

changes in chromatin contacts, we applied serial 5C analyses of a ~5-Mb region rich in inducible genes and enhancers in wild-type macrophages. Most chromatin contacts remained unchanged in response to LPS (fold change < 2 ; Supplementary Fig. 3b). While average interactions between inducible promoters and enhancers did not increase significantly in response to LPS (Supplementary Fig. 3c), circular chromosome conformation capture (4C) analysis suggested that a subset of chromatin contacts at the inducible *Egr2* locus were reconfigured in response to LPS (Supplementary Fig. 4a). As is consistent with known cohesin functions^{1,3,4,8}, local chromatin contacts appeared reduced in *Rad21*-deleted macrophages at the *Egr2* locus (Supplementary Fig. 4b) and after acute

RAD21 depletion by TEV cleavage at the *Egr2*, *Ifnar1* and *Cebpb* loci (Supplementary Fig. 4c).

Chromatin accessibility of inducible enhancers. Global assessment of chromatin accessibility by ATAC-seq identified similar numbers of accessible sites in unstimulated wild-type and *Rad21*-deleted macrophages. In response to LPS, chromatin accessibility increased in wild-type but not in cohesin-deficient macrophages as judged by the number of ATAC-seq peaks and the percentage of reads in peaks (Supplementary Fig. 5a). This difference in accessibility was pronounced at the transcription start sites (TSS) of LPS-inducible enhancers⁴⁶ (Supplementary Fig. 5b). As cohesin can

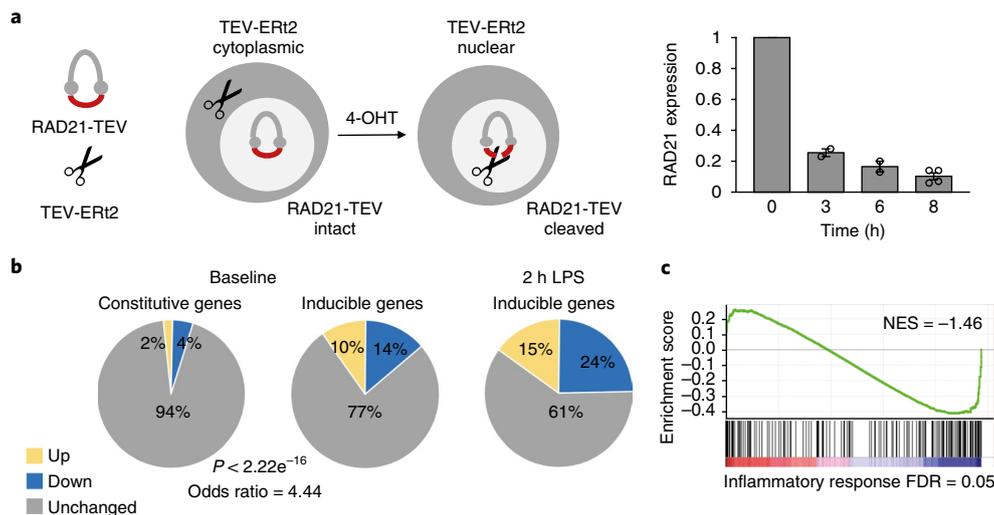


Fig. 3 | Identification of immediate cohesin target genes. **a**, Experimental system for TEV-induced cleavage of RAD21-TEV. Macrophages were generated from the livers of *Rad21*-TEV-Myc embryos⁴⁴. RAD21 protein was quantified by fluorescence immunoblotting for the Myc tag and normalized to actin (mean \pm s.e.m. of 2–4 biological replicates per time point). **b**, Changes in constitutive and inducible²³ gene expression in response to RAD21-TEV cleavage (4-OHT versus carrier) identified by DESeq2 analysis of RNA-seq data ($P < 0.05$, 3 biological replicates per genotype and time point). P -value and odds ratio were determined by two-sided Fisher's exact test between constitutive and inducible genes at baseline. Inducible genes²³ deregulated immediately after RAD21-TEV cleavage included 33 IFN-dependent²⁴ (14 up- and 19 downregulated) and 86 IFN-independent genes²⁴ (36 up- and 50 downregulated). **c**, GSEA of inflammatory response genes at baseline (normalized enrichment score (NES) = -1.46, FDR = 0.05).

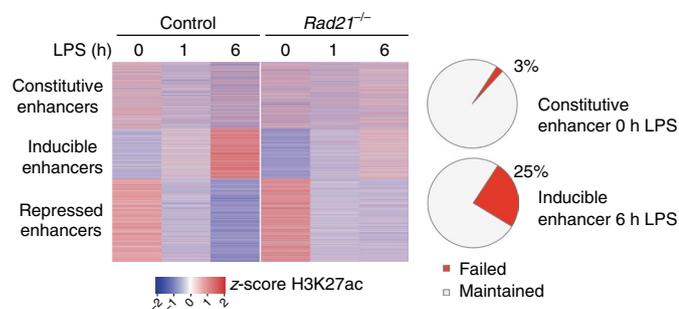


Fig. 4 | Restricted enhancer dynamics in cohesin-deficient macrophages. Left, heat map of H3K27ac ChIP-seq signals for constitutively active (8,991), inducible (6,708), and repressed (11,146) enhancers²². z-scores were calculated on the basis of fragments per kilobase of transcript per million mapped reads. Right, frequency of enhancers with deregulated H3K27ac in *Rad21*-deleted macrophages as determined by DESeq2 analysis of 2 H3K27ac ChIP-seq replicates (adjusted $P < 0.05$).

facilitate chromatin remodeling and accessibility^{47–50} we explored the relationship between cohesin binding and enhancer accessibility. Very few inducible enhancers acquired new RAD21 ChIP-seq peaks in response to LPS, but about a third of inducible enhancers showed increased RAD21 ChIP-seq reads (not peaks) in wild-type macrophages. Activation-induced cohesin binding in wild-type macrophages was not predictive of enhancer failure after *Rad21* deletion ($P = 0.67$, odds ratio = 0.96). These findings suggested that factors other than RAD21 binding contributed to enhancer failure in *Rad21*^{-/-} macrophages.

Failed enhancers have ISRE or IRF-PU.1 motifs. To understand enhancer deregulation in cohesin-deficient macrophages, we focused on inducible enhancers that showed LPS-induced upregulation of H3K27ac ($P < 0.05$ and FC ≥ 1.5) and active enhancer transcription in wild-type macrophages. We classified inducible

enhancers into those that remained intact versus those that failed to upregulate H3K27ac in *Rad21*^{-/-} macrophages (adjusted $P < 0.05$), and compared transcription factor motifs at their transcription start sites⁴⁶. Intact enhancers were enriched for NF- κ B ($P = 10^{-65}$ for H3K27ac, $P = 10^{-251}$ for GRO-seq) and NFAT ($P = 10^{-15}$ for H3K27ac, $P = 10^{-47}$ for GRO-seq) motifs. Failed enhancers were instead enriched in IFN-stimulated response elements (ISREs, targeted by the transcription factors STAT and IRF; $P = 10^{-14}$ for H3K27ac, $P = 10^{-49}$ for GRO-seq) and IRF-PU.1 composite motifs ($P = 10^{-19}$ for H3K27ac, $P = 10^{-9}$ for GRO-seq; Fig. 5a). ATAC-seq showed that chromatin accessibility of inducible enhancers with ISRE or IRF-PU.1 motifs was profoundly reduced in *Rad21*^{-/-} macrophages, more so than accessibility of inducible enhancers with NFAT or NF- κ B motifs (Fig. 5b). Inducible enhancers with ISRE or IRF-PU.1 motifs were more likely to fail ($P = 10^{-6}$, odds ratio = 8.94) while inducible enhancers with NF- κ B or NFAT motifs were less likely to fail ($P = 0.003$, odds ratio = 0.55; Table 1). Consistent with these findings, RNA-seq and quantitative reverse-transcription followed by PCR (RT-PCR) showed reduced expression of the LPS-inducible transcription factors *Stat1*, *Stat2* and *Irf7* in *Rad21*-deleted macrophages (Fig. 5c). These findings suggest that reduced expression of transcription factors contributed to enhancer failure in *Rad21*^{-/-} macrophages.

Partial rescue of inducible genes and enhancers by IFN. The organization of inducible gene expression is hierarchical (Supplementary Fig. 6a), as early events, including the induction of transcription factors and cytokines, are required for the appropriate regulation of downstream genes^{23–25}. In hierarchical networks, information propagates from a small number of upstream nodes—for example, TLRs or IFN receptors—to numerous downstream targets (Supplementary Fig. 6a). This strategy is vulnerable, as failure of early events can cause widespread defects⁵¹. We considered whether the organization of inducible gene expression in macrophages might compound the deregulation of inflammatory gene expression in cohesin-deficient cells. Our data implicate the IFN pathway as a key intermediate. First, inducible enhancers targeted by the IFN signaling pathway components STAT and IRF are prone to fail in cohesin-deficient

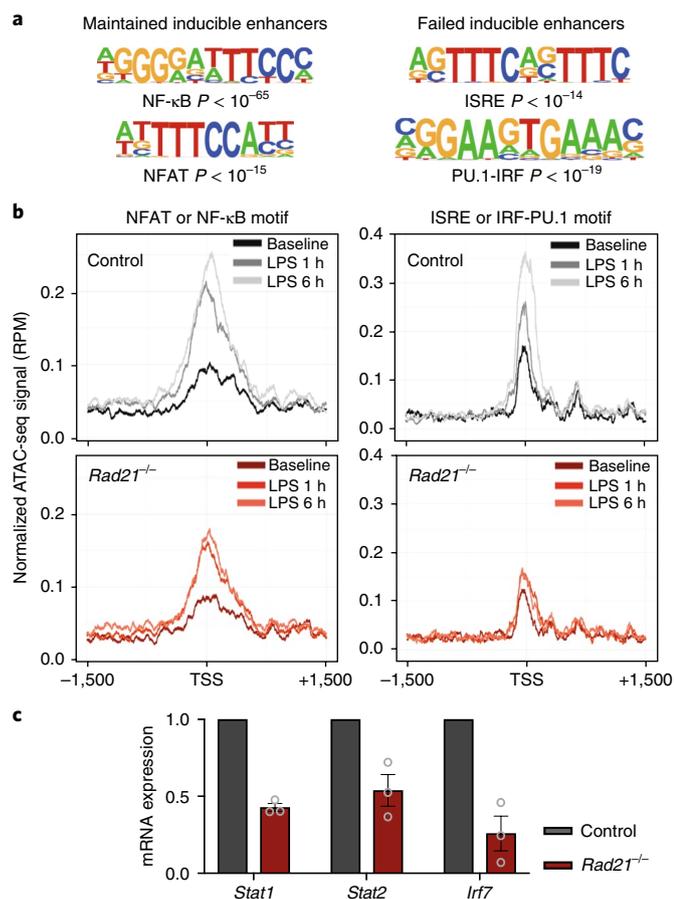


Fig. 5 | Inducible enhancers with ISRE or IRF-PU.1 motifs are more likely to fail. **a**, Inducible enhancers²² (H3K27ac $\log_2FC \geq 1.5$) with GRO-seq-mappable TSSs were classified as failed (reduced H3K27ac in *Rad21*^{-/-} macrophages, adjusted $P < 0.05$) or maintained, and compared for enrichment of TSS-proximal transcription factor motifs. The 10 most highly enriched motifs were NF- κ B (3 occurrences), NFAT (1), basic helix-loop-helix (2), nuclear receptors (2), JUN-CRE (1) and Fos12 (1) at maintained inducible enhancers, and ISRE (5 occurrences), PU.1-IRF (2), KLF (2) and STAT (1) at failed inducible enhancers. **b**, ATAC-seq accessibility of inducible enhancers with NFAT or NF- κ B motifs (top) versus ISRE or IRF-PU.1 motifs (bottom) in control (left) versus *Rad21*^{-/-} macrophages (right). RPM, reads per million. **c**, Differential expression of *Stat1*, *Stat2* and *Irf7* in *Rad21*^{-/-} macrophages relative to wild-type confirmed by quantitative RT-PCR. Mean \pm s.e.m. of 3 biological replicates, $P < 0.05$ by two-sided *t*-test.

macrophages. Second, IFN signaling genes are deregulated by acute cohesin depletion. On the basis of these observations, we tested the impact of exogenous IFN- β on inducible gene expression. IFN- β induced *Stat1*, *Stat2* and *Irf7* expression and significantly reduced the difference in the expression of these mediators between *Rad21*-deleted and wild-type macrophages (Fig. 6a).

We next assessed the impact of IFN on failed inducible enhancers with ISRE or IRF-PU.1 motifs and/or ChIP-seq evidence for STAT binding. Treatment of control and *Rad21*-deleted macrophages with IFN- β or IFN- γ followed by ChIP-PCR partially rescued H3K27ac (Fig. 6b). This result shows that cytokines and the transcription factors they regulate can promote enhancer activation in the absence of cohesin (Supplementary Fig. 5c).

Finally, we tested the global impact of IFN priming on LPS-inducible gene expression in *Rad21*-deleted macrophages by RNA-seq (Fig. 6c) and found partial rescue of early and late inducible gene classes²³. Rescue included a subset of domains with deregulated gene

Table 1 | Inducible enhancers with ISRE/IRF-PU.1 motifs are significantly more likely to fail than inducible enhancers with NF- κ B or NFAT motifs

| | Failed | Maintained |
|---|--------|------------|
| NF- κ B or NFAT motif | 57 | 90 |
| No NF- κ B or NFAT motif | 161 | 139 |
| $P = 0.003$, odds ratio = 0.55: less likely to fail | | |
| | Failed | Maintained |
| ISRE or IRF-PU.1 motif | 30 | 4 |
| No ISRE or IRF-PU.1 motif | 188 | 225 |
| $P = 1 \times 10^{-6}$, odds ratio = 8.94: more likely to fail | | |
| <i>P</i> -value and odds ratios were determined by Fisher's exact test. | | |

expression, as illustrated for clusters of *Sfn*, *Ccl*, *Gbp* and *Lrrc* genes (Fig. 6d). This rescue is most likely explained by the shared regulatory requirements of gene duplicates contained within these clusters⁴⁵.

At the genomic level, cohesin-dependent genes are enriched for cohesin binding¹⁷, as well as proximity to enhancers and super-enhancers^{4,7,20}. These features are evident for both constitutive and inducible genes deregulated by acute degradation of RAD21-TEV (Supplementary Fig. 6b), as illustrated by the IFN receptor genes *Ifnar1* and *Ifnar2* (Supplementary Fig. 6c). Acute cohesin depletion in RAD21-TEV macrophages preferentially deregulated cohesin-bound genes close to enhancers and superenhancers (Supplementary Fig. 6d). The deregulation of inducible genes became more extensive after prolonged cohesin depletion: in response to acute cohesin depletion in RAD21-TEV macrophages, 25% of inducible genes were deregulated at baseline. The fraction of deregulated inducible genes increased to 50% after 1 to 2 d of cohesin depletion in *Rad21*-deleted macrophages. Similarly, 39% of inducible genes were deregulated after 2 h LPS activation of acutely cohesin-depleted macrophages, which increased to 60% to 80% in LPS-stimulated macrophages 1 to 2 d after cohesin depletion. As deregulation spread to include most inducible genes in *Rad21*-deleted macrophages, it was no longer focused on enhancer-proximal and cohesin-bound genes (Supplementary Fig. 6d). These findings are consistent with the logic of the inducible gene expression network discussed above.

Overall, inducible genes were highly enriched for genomic proximity to enhancers, inducible enhancers and superenhancers ($P < 2.2 \times 10^{-16}$; Supplementary Fig. 6e) as a genomic correlate of cohesin dependence^{4,7,20}. Hence, inducible gene expression is vulnerable to cohesin depletion at two distinct levels: the cohesin dependence of its constituent components and the topology of the inducible gene expression network.

Cohesin controls inflammatory gene expression in HSPCs.

To address the relationship between cohesin, inflammatory gene expression and differentiation, we extended our analysis to HSPCs. Inflammatory signals cause gene expression changes in HSPCs^{35–42}. Conversely, HSPCs contribute to the production of inflammatory cytokines that regulate HSPC self-renewal and differentiation^{40,42,52}. Inflammatory signals promote myeloid differentiation at the expense of HSPC self-renewal^{35–42}, and HSPCs with reduced cohesin function display enhanced self-renewal in serial in vitro colony-forming assays^{30–32} and in vivo competitive reconstitution experiments³². We therefore examined gene expression in lineage-negative, Sca1⁺, c-Kit⁺ (LSK) progenitors with reduced cohesin function following *Stag2* RNA interference (RNAi) in vivo³¹ and found a notable down-regulation of inflammatory genes at baseline (Fig. 7). Equivalent results were seen in progenitors with reduced *Smc1a*³¹ and *Smc3*³²

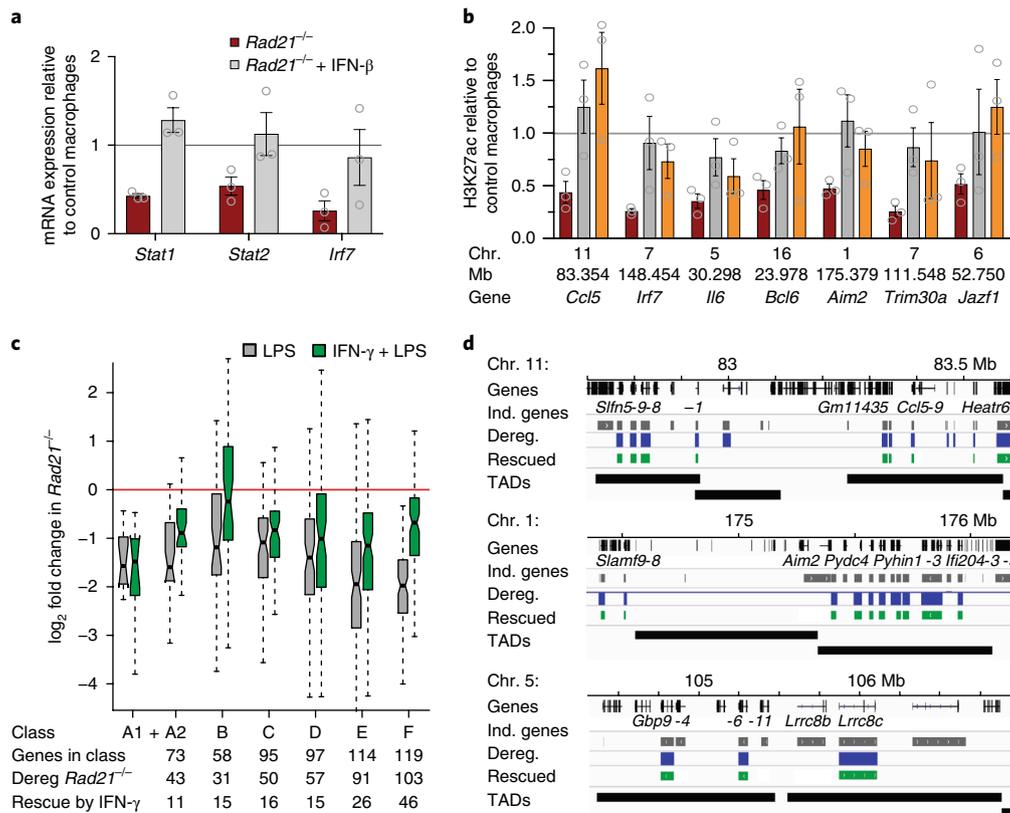


Fig. 6 | Rescue of inducible genes and enhancers in the absence of cohesin. **a**, Quantitative RT-PCR analysis of *Stat1*, *Stat2* and *Irf7* expression by control and *Rad21*^{-/-} macrophages. Expression in *Rad21*^{-/-} relative to control macrophages with (gray) and without (red) IFN-β treatment for 24 h. Mean ± s.e.m. of 3 biological replicates. **b**, ChIP quantitative PCR of histone H3-normalized H3K27ac in *Rad21*^{-/-} relative to control macrophages at candidate inducible enhancers. The genomic coordinates and the nearest inducible gene are shown for each enhancer. Cells were cultured in medium (red), 10 ng/ml IFN-γ (orange) or 100 U/ml IFN-β (gray) for 24 h before LPS stimulation for 6 h. Mean ± s.e.m. of 3 biological replicates. **c**, Fold change of early (2 h LPS, class A–D) and late (8 h LPS, class E, F) inducible genes²³ in *Rad21*^{-/-} over control macrophages with (green) or without (gray) 24 h IFN-γ pretreatment. The numbers of total, deregulated and rescued genes in each class are shown. Box plots show the median and lower and upper quartiles; whiskers show the maximum and minimum data points up to 1.5 times the interquartile range. **d**, Genomic view of inducible (Ind.) genes (gray), deregulated (Dereg.) genes (blue) and rescue of gene expression by IFN-γ pretreatment (green), assessed by fold change and/or DESeq-2 analysis of RNA-seq data from control and

expression (data not shown). Reanalysis of published gene expression data^{31,32,53} confirmed that proinflammatory pathways that are downregulated in cohesin-deficient HSPCs^{31,32} are reciprocally upregulated in HSPCs exposed to chronic *Mycobacterium avium* infection⁵³ ($P = 5.9 \times 10^{-28}$, odds ratio = 7.59, Table 2). These findings link cohesin to inflammatory gene expression in HSPCs.

HSPCs respond to cohesin-dependent inflammatory signals.

To evaluate the biological impact of cohesin-dependent cytokine secretion, we isolated LSKs by flow cytometry (Fig. 8a). Seven of ten stem cell genes tested showed reduced expression in LSKs exposed to medium conditioned by LPS-pulsed wild-type macrophages (Fig. 8b). Medium conditioned by LPS-pulsed *Rad21*-deleted macrophages had markedly less impact on stem cell gene expression (Fig. 8b). Common myeloid progenitors (CMPs) and granulocyte-macrophage progenitors (GMPs) upregulated the expression of the myeloid differentiation markers CD11b and CD16 in response to medium conditioned by wild-type macrophages, but medium conditioned by *Rad21*-deleted macrophages was less effective (Fig. 8c). These data show that HSPCs are sensitive to cohesin-dependent inflammatory signals.

Cohesin controls HSPC responses to inflammatory stimuli.

Finally, we asked how cohesin mutations affect the sensitivity of

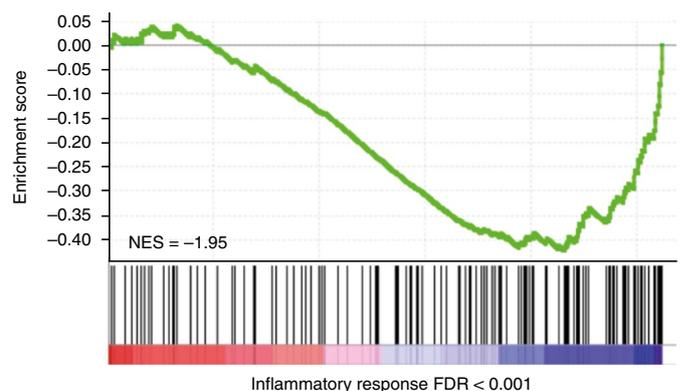


Fig. 7 | Cohesin controls inflammatory gene expression in HSPCs. GSEA analysis of inflammatory gene expression by HSPCs after *Stag2* versus control RNAi (FDR = 2.96×10^{-4} , data from ref. ³¹). NES, normalized enrichment score.

HSPCs to inflammatory signals. In response to the proinflammatory signal LPS, *Rad21*^{-/-} LSKs showed significantly lower inflammatory gene expression than wild-type LSKs (Fig. 8d). LPS

Table 2 | Genes downregulated in HSPCs with reduced cohesin function are upregulated in chronic inflammation

| | Odds ratio | P-value |
|---|------------|-------------------------|
| <i>Smc1</i> RNAi versus inflammation ^a | 5.97 | 2.4 × 10 ⁻⁸ |
| <i>Stag2</i> RNAi versus inflammation ^a | 6.09 | 1.9 × 10 ⁻¹⁵ |
| <i>Smc3</i> ^{+/-} versus inflammation ^b | 8.49 | 5.4 × 10 ⁻¹¹ |
| Combined | 7.59 | 5.9 × 10 ⁻²⁸ |

Downregulated genes in HSPCs with reduced cohesin function were intersected with genes upregulated in chronic inflammation⁵³. Gene ontology pathways enriched in the overlap included "Cytokine-mediated signaling pathway," "Cellular response to IFN-γ," "Cellular response to cytokine stimulus" and "Regulation of cytokine production." ^aDownregulated as determined by ref. ³¹ versus ref. ⁵³. ^bDownregulated as determined by ref. ³² versus ref. ⁵³.

reduced the expression of stem cell genes in wild-type LSKs, but stem cell gene expression was markedly more robust to LPS exposure in *Rad21*^{+/-} LSKs (Fig. 8e). We conclude that cohesin connects inflammation with self-renewal and differentiation by controlling the expression of inflammatory genes by HSPCs at baseline and in response to inflammatory stimuli, and the sensitivity of HSPCs to inflammatory signals.

Discussion

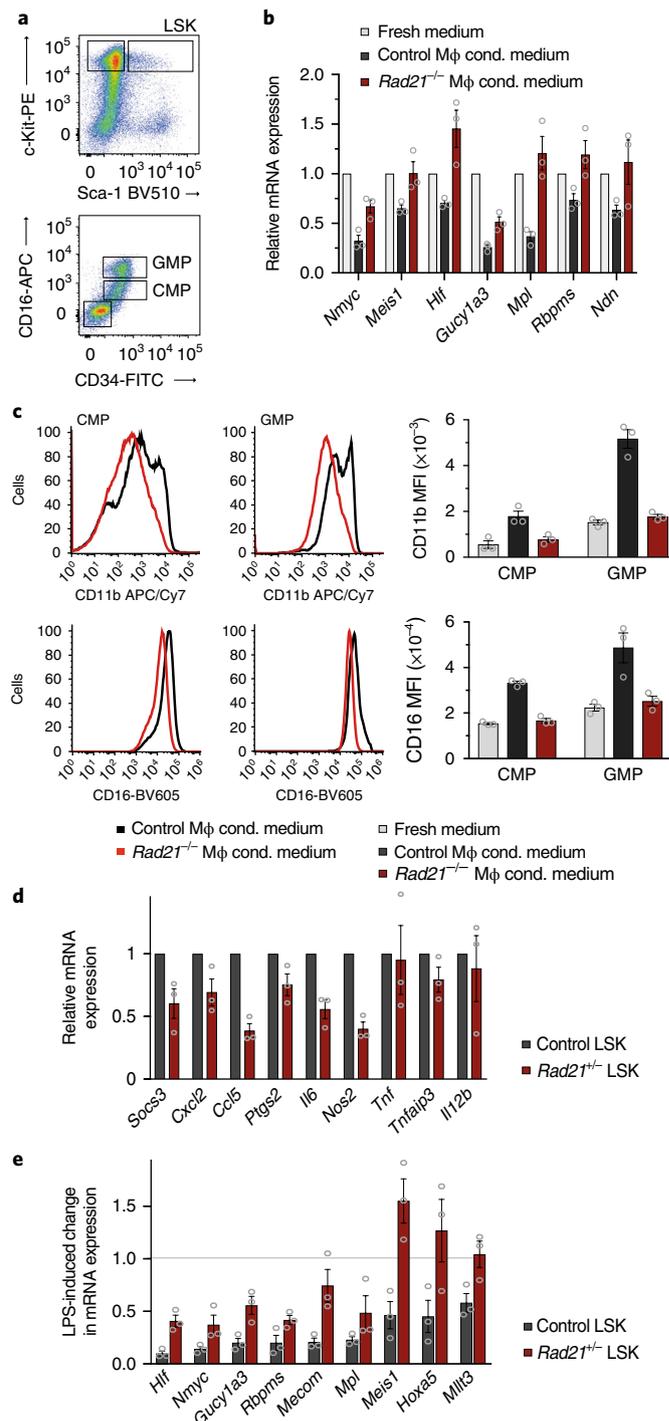
Depletion of cohesin or CTCF disrupts key features of 3D chromatin organization, but in previous studies had limited impact on the maintenance of gene expression and chromatin modifications^{3,4,7,20}. These findings called into question the significance of genome folding for the regulation of gene expression, chromatin state and enhancer activity. Here we show that cohesin is critically required for inducible gene expression and enhancer dynamics in primary myeloid cells. This indicates an important role for cohesin in the transition from a resting to an activated state and suggests that the impact of cohesin on gene expression may have been underestimated by studies confined to cell lines under steady-state conditions.

Inducible genes are subject to regulation by a complex network of enhancers^{21,22}, and our data show that inducible genes are significantly enriched in the vicinity of enhancers and superenhancers.

Fig. 8 | Cohesin controls the responsiveness of HSPCs to inflammatory stimuli.

a, Flow cytometric isolation of LSK, CMP and GMP populations. PE, phycoerythrin; BV, brilliant violet; APC, allophycocyanin; FITC, fluorescein isothiocyanate; Cy7, indotricarbocyanine. **b**, RT-PCR analysis of stem cell gene expression in LSKs exposed for 48 h to medium conditioned (cond.) by LPS-pulsed wild-type or *Rad21*-deleted macrophages. Normalized to gene expression in fresh medium. Mean ± s.e.m. of 3 biological replicates. Comparisons between wild-type and *Rad21*-deleted macrophage conditioned medium were *P* < 0.05 for all genes except *Ndn* (*t*-test). **c**, Flow cytometric analysis of the myeloid differentiation antigens CD11b and CD16 on CMPs and GMPs exposed to medium conditioned by LPS-pulsed wild-type or *Rad21*-deleted macrophages for 48 h. MFI: mean fluorescence intensity. Mean ± s.e.m. of 3 biological replicates. Comparisons between wild-type and *Rad21*-deleted macrophage conditioned medium were *P* < 0.05 for CD11b and CD16 (two-sided *t*-test). **d**, Quantitative RT-PCR analysis of inflammatory gene expression in *Rad21*^{+/-} relative to wild-type LSK cells exposed to LPS for 8 h (mean ± s.e.m. of 3 biological replicates, *P* < 0.05 for all genes except *Tnf*, *Tnfrsf3* and *Il12b*, two-sided *t*-test). RAD21 expression in *Rad21*^{+/-} bone marrow was 79% ± 8% of that in wild-type bone marrow (mean ± s.e.m. of 4 biological replicates). **e**, Quantitative RT-PCR analysis of stem cell gene expression in *Rad21*^{+/-} relative to wild-type LSKs exposed to LPS for 8 h. Mean ± s.e.m. of 3 biological replicates. *Hlf*, *Gucy1a3*, *Mecom*, *Meis1* and *Mllt3* were significantly higher in LPS-stimulated *Rad21*^{+/-} than wild-type LSKs (*P* < 0.05, two-sided *t*-test). *P* = 0.0014 by two-sided *t*-test over all transcripts.

Enhancer interactions are altered in the absence of cohesin^{4,20}, consistent with models whereby cohesin-dependent chromatin contacts facilitate enhancer–promoter contacts and counteract the segregation of chromatin regions according to chromatin state^{3,4,7,20}. These findings offer an explanation for the enrichment of inducible genes among immediate cohesin-regulated genes. The organization of the inducible gene expression network is hierarchical, and the expression of secondary response genes depends on inducible transcription factors and cytokines that act in an auto- and paracrine fashion^{23,25,54}. This regulatory logic and the cohesin-dependence of inducible genes, including IFN receptors and IFN-regulated transcription factors, render inducible gene expression particularly vulnerable to disruption by the loss of cohesin. In support of this



model, exogenous provision of inducible cytokines partially rescued inducible genes and enhancer dynamics in the absence of cohesin.

Cohesin is required for cell proliferation³⁶, yet many cancers accumulate cohesin mutations^{27,28}. These findings are reconciled by data that the amount of cohesin present in normal cells is in excess of what is required for sister chromatid cohesion⁵⁵. Modest reductions in cohesin function affect the development of multiple organ systems in humans⁵⁶, suggesting that the correct expression of developmental genes is highly sensitive to cohesin dosage. In HSPCs, reduced cohesin function tilts the balance between self-renewal and differentiation and allows increased proliferation of immature progenitors^{30–34}. Here we provide an explanation for this finding by demonstrating that HSPCs with reduced cohesin function show reduced inflammatory gene expression and increased resilience to the differentiation-inducing effect of inflammatory signals. Notably, we find that cohesin mutations impair the expression of inflammatory genes in human AML as well. The regulation of inflammatory gene expression and the sensitivity to inflammatory signals provide a mechanistic link between cohesin and myeloid differentiation^{35–42}. As inflammatory mediators control the self-renewal and differentiation of HSPCs in an auto- and paracrine fashion^{35–42}, this model suggests a mechanism for how cohesin mutations may favor self-renewal, delay differentiation and provide a selective advantage in AML. Our data establish precedent for cohesin dependence of gene regulatory networks. Similar mechanisms may operate in human development, where cohesin mutations disrupt multiple systems⁵⁶, and in cancers other than AML.

Methods

Methods, including statements of data availability and any associated accession codes and references, are available at <https://doi.org/10.1038/s41590-018-0184-1>.

Received: 5 January 2018; Accepted: 17 July 2018;
Published online: 20 August 2018

References

- Fudenberg, G. et al. Formation of chromosomal domains by loop extrusion. *Cell Rep.* **15**, 2038–2049 (2016).
- Nora, E. P. et al. Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell* **169**, 930–944.e22 (2017).
- Rao, S. S. P. et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665–1680 (2014).
- Rao, S. S. P. et al. Cohesin loss eliminates all loop domains. *Cell* **171**, 305–320.e24 (2017).
- Downen, J. M. et al. Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell* **159**, 374–387 (2014).
- Hnisz, D., Day, D. S. & Young, R. A. Insulated neighborhoods: structural and functional units of mammalian gene control. *Cell* **167**, 1188–1200 (2016).
- Schwarzer, W. et al. Two independent modes of chromatin organization revealed by cohesin removal. *Nature* **551**, 51–56 (2017).
- Merkenschlager, M. & Nora, E. P. CTCF and cohesin in genome folding and transcriptional gene regulation. *Annu. Rev. Genomics Hum. Genet.* **17**, 17–43 (2016).
- Dekker, J. & Heard, E. Structural and functional diversity of topologically associating domains. *FEBS Lett.* **589**, 2877–2884 (2015).
- Dekker, J. & Mirny, L. The 3D genome as moderator of chromosomal communication. *Cell* **164**, 1110–1121 (2016).
- Bickmore, W. A. & van Steensel, B. Genome architecture: domain organization of interphase chromosomes. *Cell* **152**, 1270–1284 (2013).
- Hnisz, D. et al. Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science* **351**, 1454–1458 (2016).
- Jiang, Y. et al. The methyltransferase SETDB1 regulates a large neuron-specific topological chromatin domain. *Nat. Genet.* **49**, 1239–1250 (2017).
- Mansour, M. R. et al. An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. *Science* **346**, 1373–1377 (2014).
- Lupiáñez, D. G. et al. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **161**, 1012–1025 (2015).
- Narendra, V. et al. CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *Science* **347**, 1017–1021 (2015).
- Seitan, V. C. et al. A role for cohesin in T-cell-receptor rearrangement and thymocyte differentiation. *Nature* **476**, 467–471 (2011).
- Kagey, M. H. et al. Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**, 430–435 (2010).
- Lavagnoli, T. et al. Initiation and maintenance of pluripotency gene expression in the absence of cohesin. *Genes Dev.* **29**, 23–38 (2015).
- Ing-Simmons, E. et al. Spatial enhancer clustering and regulation of enhancer-proximal genes by cohesin. *Genome Res.* **25**, 504–513 (2015).
- Kaikkonen, M. U. et al. Remodeling of the enhancer landscape during macrophage activation is coupled to enhancer transcription. *Mol. Cell* **51**, 310–325 (2013).
- Ostuni, R. et al. Latent enhancers activated by stimulation in differentiated cells. *Cell* **152**, 157–171 (2013).
- Bhatt, D. M. et al. Transcript dynamics of proinflammatory genes revealed by sequence analysis of subcellular RNA fractions. *Cell* **150**, 279–290 (2012).
- Chen, X. et al. Requirement for the histone deacetylase Hdac3 for the inflammatory gene expression program in macrophages. *Proc. Natl. Acad. Sci. USA* **109**, E2865–E2874 (2012).
- Glass, C. K. & Natoli, G. Molecular control of activation and priming in macrophages. *Nat. Immunol.* **17**, 26–33 (2016).
- Nasmyth, K. & Haering, C. H. Cohesin: its roles and mechanisms. *Annu. Rev. Genet.* **43**, 525–558 (2009).
- Kon, A. et al. Recurrent mutations in multiple components of the cohesin complex in myeloid neoplasms. *Nat. Genet.* **45**, 1232–1237 (2013).
- Ley, T. J. et al. Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N. Engl. J. Med.* **368**, 2059–2074 (2013).
- Thota, S. et al. Genetic alterations of the cohesin complex genes in myeloid malignancies. *Blood* **124**, 1790–1798 (2014).
- Mazumdar, C. et al. Leukemia-associated cohesin mutants dominantly enforce stem cell programs and impair human hematopoietic progenitor differentiation. *Cell Stem Cell* **17**, 675–688 (2015).
- Mullenders, J. et al. Cohesin loss alters adult hematopoietic stem cell homeostasis, leading to myeloproliferative neoplasms. *J. Exp. Med.* **212**, 1833–1850 (2015).
- Viny, A. D. et al. Dose-dependent role of the cohesin complex in normal and malignant hematopoiesis. *J. Exp. Med.* **212**, 1819–1832 (2015).
- Galeev, R. et al. Genome-wide RNAi screen identifies cohesin genes as modifiers of renewal and differentiation in human HSCs. *Cell Rep.* **14**, 2988–3000 (2016).
- Fisher, J. B. et al. The cohesin subunit Rad21 is a negative regulator of hematopoietic self-renewal through epigenetic repression of Hoxa7 and Hoxa9. *Leukemia* **31**, 712–719 (2017).
- Baldrige, M. T., King, K. Y., Boles, N. C., Weksberg, D. C. & Goodell, M. A. Quiescent haematopoietic stem cells are activated by IFN- γ in response to chronic infection. *Nature* **465**, 793–797 (2010).
- Binder, D. et al. Aplastic anemia rescued by exhaustion of cytokine-secreting CD8+ T cells in persistent infection with lymphocytic choriomeningitis virus. *J. Exp. Med.* **187**, 1903–1920 (1998).
- Dybedal, I., Bryder, D., Fossum, A., Rusten, L. S. & Jacobsen, S. E. Tumor necrosis factor (TNF)-mediated activation of the p55 TNF receptor negatively regulates maintenance of cycling reconstituting human hematopoietic stem cells. *Blood* **98**, 1782–1791 (2001).
- Essers, M. A. G. et al. IFN α activates dormant haematopoietic stem cells in vivo. *Nature* **458**, 904–908 (2009).
- Mossadegh-Keller, N. et al. M-CSF instructs myeloid lineage fate in single haematopoietic stem cells. *Nature* **497**, 239–243 (2013).
- Nagai, Y. et al. Toll-like receptors on hematopoietic progenitor cells stimulate innate immune system replenishment. *Immunity* **24**, 801–812 (2006).
- Pietras, E. M. et al. Chronic interleukin-1 exposure drives haematopoietic stem cells towards precocious myeloid differentiation at the expense of self-renewal. *Nat. Cell Biol.* **18**, 607–618 (2016).
- Zhao, J. L. et al. Conversion of danger signals into cytokine signals by hematopoietic stem and progenitor cells for regulation of stress-induced hematopoiesis. *Cell Stem Cell* **14**, 445–459 (2014).
- Bennett, J. M. et al. Proposals for the classification of the acute leukaemias. *Br. J. Haematol.* **33**, 451–458 (1976).
- Tachibana-Konwalski, K. et al. Rec8-containing cohesin maintains bivalents without turnover during the growing phase of mouse oocytes. *Genes Dev.* **24**, 2505–2516 (2010).
- Zlotnik, A., Yoshie, O. & Nomiya, H. The chemokine and chemokine receptor superfamilies and their molecular evolution. *Genome Biol.* **7**, 243 (2006).
- Azofeifa, J. G., Allen, M. A., Lladser, M. E. & Dowell, R. D. An annotation agnostic algorithm for detecting nascent RNA transcripts in GRO-Seq. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **14**, 1070–1081 (2017).
- Hakimi, M.-A. et al. A chromatin remodelling complex that loads cohesin onto human chromosomes. *Nature* **418**, 994–998 (2002).

48. Faure, A. J. et al. Cohesin regulates tissue-specific expression by stabilizing highly occupied cis-regulatory modules. *Genome Res.* **22**, 2163–2175 (2012).
49. Yan, J. et al. Transcription factor binding in human cells occurs in dense clusters formed around cohesin anchor sites. *Cell* **154**, 801–813 (2013).
50. Lopez-Serra, L., Kelly, G., Patel, H., Stewart, A. & Uhlmann, F. The Scc2–Scc4 complex acts in sister chromatid cohesion and transcriptional regulation by maintaining nucleosome-free regions. *Nat. Genet.* **46**, 1147–1151 (2014).
51. Schieber, T. A. et al. Information theory perspective on network robustness. *Phys. Lett. A* **380**, 359–364 (2016).
52. Painter, M. W., Davis, S., Hardy, R. R., Mathis, D. & Benoist, C. Transcriptomes of the B and T lineages compared by multiplatform microarray profiling. *J. Immunol.* **186**, 3047–3057 (2011).
53. Matatall, K. A. et al. Chronic infection depletes hematopoietic stem cells through stress-induced terminal differentiation. *Cell Rep.* **17**, 2584–2595 (2016).
54. Shalek, A. K. et al. Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. *Nature* **510**, 363–369 (2014).
55. Heidinger-Pauli, J. M., Mert, O., Davenport, C., Guacci, V. & Koshland, D. Systematic reduction of cohesin differentially affects chromosome segregation, condensation, and DNA repair. *Curr. Biol.* **20**, 957–963 (2010).
56. Deardorff, M. A., Noon, S. E. & Krantz, I. D. Cornelia de Lange syndrome. *GeneReviews* (2018).

Acknowledgements

We thank A. Innes, M. Spivakov, L. Rudolf (FLI, Jena) and D. Odom (CRUK Cambridge) for discussions, G. Sauvageau (University of Montreal) and the Leucegene consortium for early access to AML data, L. Game for sequencing, and J. Elliott and the LMS/NIHR Imperial Biomedical Research Centre Flow Cytometry Facility for cell sorting. This work was funded by Wellcome Investigator Award 099276/Z/12/Z (M.M.), Wellcome Project Grant P55504 (B.L.), Fundação de Amparo à Pesquisa do Estado de São Paulo 2014/20861-3 (M.T.A.), National Science Foundation grant ABI1262410 (R.D.D.), Academy of Finland grants 287478 and 294073 (M.U.K.), MRC Programme Grant ID 84637 and Wellcome Trust Programme Grant Ref 078241/Z/05/Z (K.A.M.), ERC grant

692789 (G.N.), and core support from the Medical Research Council UK to the London Institute of Medical Sciences.

Author contributions

S.C. conceived and designed the study; performed most experiments, including those based on *Rad21* deletion; analyzed data; designed figures; and contributed to writing the manuscript. F.D.W. performed experiments based on RAD21-TEV cleavage and contributed to writing the manuscript. G.D., X.X. and Y.-F.W. analyzed data. Y.G. performed and analyzed 4C experiments. E.I.-S. analyzed data, designed figures and contributed to writing the manuscript. S.M. performed 5C experiments. I.R.-R. performed and analyzed immunofluorescence experiments, designed figures and contributed to writing the manuscript. I.B. analyzed data and contributed to writing the manuscript. D.D. designed and performed flow cytometry experiments. M.T.A. performed experiments. H.N. designed and performed GRO-seq experiments. E.P. and B.L. designed and supervised data analysis. R.D.D. conceived and performed analysis of enhancer TSSs. K.T. and K.A.N. generated and provided essential reagents. M.U.K. designed and supervised GRO-seq experiments. G.N. designed and supervised 5C experiments and contributed to writing the manuscript. A.G.F. contributed to study design and writing the manuscript. M.M. conceived and designed the study, wrote the manuscript, made figures and supervised experiments. All authors discussed the results and commented on the manuscript.

Competing interests

The authors declare no competing financial interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41590-018-0184-1>.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to M.M.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

Mice and cell culture. Mouse work was performed according to the Animals (Scientific Procedures) Act under the authority of project license PPL70/7556 issued by the Home Office, UK following approval by the Imperial College London ethics review board. Bone marrow cells from *Rosa26-Ert2Cre⁺ Rad21^{WT/WT}* or *Rad21^{lox/lox}* mice¹⁷ were cultured in complete DMEM medium (10% FCS, 1% penicillin-streptomycin, 0.05 mM β -mercaptoethanol, 2 mM L-glutamine, 1 mM sodium pyruvate), 20% L929-conditioned media. Cre was induced on day 4 with 200 nM 4-hydroxytamoxifen (Sigma-Aldrich H7904). Macrophages were stimulated on day 7 with 10 ng/ml of LPS from *Salmonella typhosa* (Sigma-Aldrich L7895), where indicated after priming for 24 h with 10 ng/ml mouse IFN- γ (Invitrogen PMC4031) or 100 U/ml mouse IFN- β (Chemicon IF011).

For TEV cleavage of RAD21 protein, macrophages were isolated from E14 *Rad21^{tev/tev}* fetal livers⁴⁴ and plated in complete IMDM with 20% L929-conditioned media. Two days later, 4×10^6 cells were resuspended in 2 ml retroviral supernatant, 4 μ g/ml Polybrene, and centrifuged at 1,250g, 90 min, 37 °C. After 8 to 10 d, 500 nM 4-hydroxytamoxifen or carrier (ethanol) was added for 8 h. Where indicated, cells were treated with LPS (2 h, 10 ng/ml). The top 20–30% GFP-expressing cells were sorted for RNA and protein analysis.

LSKs, CMPs and GMPs were sorted from bone marrow depleted of lineage markers (CD4, CD8, B220, CD19, NK1.1, CD11b, Ter119, GR-1; Miltenyi 130-048-102 streptavidin beads). Cells were stained with Sca-1-BV510 (BD 565507), cKit-PE (eBioscience 12-1171-81), CD16-APC (eBioscience 17-0161-81) and CD34-FITC (eBioscience 11-0341-81), and the remaining lineage-positive cells were gated out using biotinylated streptavidin-eFluor 450 (eBioscience 48-4317-82). *Rad21^{+/-}* bone marrow was derived from mice with a germline deletion of one *Rad21* allele. Sorted populations were cultured in complete DMEM, 100 ng/ml recombinant mouse SCF (Peprotech 250-03). Where indicated, we added filtered medium conditioned for 24 h by macrophages that had been LPS-activated for 60 min and then washed.

FACS analysis. Macrophages were stained with CD11b-FITC (BD 561688), F4/80-PE (eBioscience 12-4801-80) with anti-mouse V α 11.1/11.2-TCR-PE (BD 553223) and V α 2-TCR-FITC (BD 553288) as isotype-matched control antibodies. CMPs and GMPs were analyzed by Sca-1-FITC (Biolegend 122505), cKit-Alexa Fluor 700 (eBioscience 56-1172-80), CD11b-APC-Cy7 (BD 557657) and CD16-BV605 (BD 563006).

TEV protein cloning and virus production. TEV cDNA was amplified from the pRNA vector and a v5 epitope tag (GKPIPNNLLGLDST) was inserted upstream of the TEV sequence. Ert2 and v5-TEV were fused by PCR using primers with XhoI and EcoRI sites and cloned into the XhoI-EcoRI site upstream of an internal ribosome entry site into pMSCV-IRES-GFP⁵⁸. Retrovirus was generated as described⁵⁸.

Immunoblots and antibody arrays. RAD21 (Abcam ab992), β -actin (Santa Cruz sc-69879), c-Myc (Santa Cruz sc-40 9E10) and GAPDH (Abcam ab8245) were used for immunoblots. Cytokine arrays (R&D ARY006) and IFN- β ELISA (RnD 42400-1) were performed following manufacturers' instructions using supernatant from macrophages collected 8 h after LPS stimulation (10 ng/ml). Immunoblots and antibody arrays were imaged using an Odyssey CLx instrument (LI-COR).

Immunofluorescence. Macrophages were seeded at 2×10^5 per coverslip, treated with LPS (10 ng/ml), fixed with formaldehyde (4%, 15 min), permeabilized with Triton X-100 (0.1%, 10 min), blocked with goat serum (10%, 30 min) and incubated with 1:100 p65 antibody (Abcam ab7970) in 10% serum for 1 h, followed by 1:750 goat anti-rabbit Alexa Fluor 488 (Invitrogen A11034) and mounting in ProLong Gold Antifade Mountant with DAPI (Invitrogen).

Image acquisition and analysis. Four 3D stacks were imaged per sample with a Leica SP8 microscope (between 113 to 340 cells imaged per sample, $1,024 \times 1,024$ pixels per image, with a pixel size of $0.2027 \times 0.2027 \mu\text{m}$, $\times 40$ oil objective). Maximum projections were analyzed in CellProfiler⁶⁹ using a pipeline that identifies nuclei (IdentifyPrimaryObjects) and the cell outline (IdentifySecondaryObjects) to determine the correlation between the DAPI signal and p65 fluorescence. Correlations 0.5 were considered indicative of nuclear translocation.

RT-qPCR. RNA was extracted with Trizol (Ambion) or RNA-bee (Amsbio) from macrophages and a PicoPure kit (Applied Biosystems KIT0204) from progenitors. cDNA synthesis used Superscript reverse transcriptase (Invitrogen) and qPCR with IQ SYBR Green Supermix (Bio-Rad) and a CFX Real-time PCR system (Bio-Rad). Primers are listed in Supplementary Table 1. Ct values were normalized to *Actb* and *Hprt*.

RNA-seq. RNA sequencing was performed from three biological replicates per condition. RNA from 2×10^6 cells was extracted with an RNeasy minikit and using Qiashredder (Qiagen). RNA was assessed for quality (Bioanalyzer, Agilent) and quantity (Qubit, Invitrogen). ERCC RNA Spike-Ins (Ambion) were added, and

strand-specific libraries prepared from 750 ng of total RNA using TruSeq Stranded total RNA Kit (Illumina RS-122-2201). RNA from liver-derived macrophages was purified with a PicoPure RNA Isolation kit (Applied Biosystems KIT0204), and 100 ng were used to prepare libraries using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina. Library quality and quantity were assessed on a Bioanalyzer and Qubit, respectively. Libraries were sequenced on an Illumina HiSeq2500 (v4 chemistry), generating 40 million paired-end 100-bp reads per sample.

GRO-seq. GRO-seq libraries²¹ were prepared from two biological replicates per condition from 5×10^6 cells. After nuclear run-on, RNA was extracted using Trizol (Ambion), treated with Turbo DNase (Ambion AM1907), fragmented (Ambion AM8740), purified on P-30 columns (Bio-Rad 732-6250), dephosphorylated with PNK (New England Biolabs Y904L) and purified using anti-BrdU beads (SantaCruz sc-214314). For reverse transcription, oligonucleotides with custom barcodes were used (Supplementary Table 1) and the cDNA was purified and PCR-amplified. The resulting product was gel purified (Novex 10% TBE gel) and cleaned using ChIP DNA Clean & Concentrator Kit (Zymo D5205).

ATAC-seq. ATAC-seq⁶⁰ was performed in two biological replicates per condition from 5×10^4 nuclei per replicate using Nextera Tn5 Transposase (Illumina FC-121-1030, 30 min, 37 °C). DNA was purified by Qiagen MinElute Kit. Transposed fragments were amplified with NEBNext High-Fidelity PCR Master Mix (NEB M0541). Libraries were cleaned and size-selected using AMPure beads (Agilent) and assessed by Bioanalyzer and Qubit.

ChIP-seq and ChIP-qPCR. For H3 and H3K27Ac ChIP, cells were cross-linked with 1% formaldehyde, lysed and sonicated (Bioruptor, Diagenode) for 40 cycles and power H in 1% Triton, 0.1% sodium deoxycholate, 0.5% SDS, 0.2 M NaCl, 10 mM Tris, pH 7.5, 10 mM EDTA. Lysates were incubated for 16 h with anti-H3 (Abcam ab1791) and anti-H3K27Ac (Abcam ab4729) prebound to protein G Dynabeads (Invitrogen 10004D) in RIPA buffer. Beads were washed and reverse cross-linked by incubation at 65 °C, 10% SDS. DNA was purified using ChIP DNA Clean & Concentrator Kit (Zymo D5205). For RAD21 ChIP, cells were sonicated for 25 cycles at power H in 1% Triton, 0.1% sodium deoxycholate, 0.1% SDS, 0.8 M NaCl, 10 mM Tris pH 7.5, 1 mM EDTA, and incubated for 16 h with anti-RAD21 (Abcam ab992). Libraries were prepared using a NEBNext Ultra DNA Library Prep kit (New England Biolabs E7370).

4C-seq. 4C template preparation was performed as described^{6,61} with modifications. Briefly, macrophages were cross-linked in PBS with 1% formaldehyde at 20–25 °C for 10 min and nuclei were isolated in lysis buffer (10 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.5% NP-40, 5 mM EDTA, protease inhibitors). The first digestion was performed by using MboI and digestion products were ligated by T4 DNA ligase. Then the 3C templates were digested by the second enzyme, NlaIII, and the digested DNA fragments were ligated again. 4C data analysis was performed using the 4Csepipe software suite⁶², and the setting values of nearcis were -stat_type mean -trend_resolution 5000. PCR primers used are listed in Supplementary Table 1.

5C. 3C templates were obtained crosslinking cells with 1% formaldehyde for 10 min at 20–25 °C. 1×10^7 cells were lysed in 500 μ l of lysis buffer (10 mM Tris-HCl pH8.0, 10 mM NaCl, 0.2% NP-40, 1 \times protease inhibitor) for 15 min on ice and disrupted with 15 strokes of a 1-ml pipette. After centrifugation, nuclei were resuspended in 500 μ l digestion buffer and lysed by adding SDS (0.1%, 10 min, 65 °C). SDS was quenched with Triton-X100 (1%). DNA was digested with HindIII (800 U, 37 °C, 16 h). After inactivation by SDS (1.6%, 65 °C, 20 min), samples were diluted in 7.5 ml ligation buffer with 3,000 NEB units T4 ligase and incubated at 16 °C for 4 h. Ligated chromatin was digested by proteinase K for 16 h. DNA was phenol-chloroform extracted and ethanol precipitated. Three femtomole 5C primers were annealed to the junctions of the 3C material for 16 h at 48 °C, joined with 10 U of NAD-dependent ligase for 1 h, and amplified by 25 PCR cycles using T3 and T7 universal primers. Libraries were sequenced to 30×10^6 100-bp paired-end reads on an Illumina Hi-seq 2000. Forward and reverse 5C primers were designed using my5C software (<http://3dg.umassmed.edu/my5Cprimers/5C.php>) to interrogate interactions between HindIII fragments containing transcription start sites (TSSs) and any other HindIII restriction fragments (distal fragments) in the ~5-Mb interval (80,141,160–85,160,410 on mouse chr. 11). Multiplex 5C libraries were produced by mixing 171 reverse primers annealing to the TSS of all genes in the region (~3 restriction fragments per TSS), 581 forward primers annealing to all other restriction fragments and 21 reverse primers with 20 forward primers corresponding to random restriction fragments on a gene desert region (chr. 14) to assess 99,351 possible contacts.

RNA-seq analysis. 100-bp paired-end RNA-seq reads were aligned to mouse genome mm9 using Tophat2⁶³ with arguments -library-type fr-firststrand -b2-very-sensitive -b2-L 25, with gene annotation from Ensembl version 67. Read counts on genes were summarized using HTSeq-count⁶⁴. Differentially expressed genes were identified using DESeq2⁶⁵. FPKM values were computed in R and heat

maps were drawn using rlog values for inducible genes²³ using R package heatmap3.

Gene set enrichment analysis. GSEA was carried out using ranked gene list based on Wald statistics from DESeq2 results using MSigDB gene sets⁶⁶. Genes with low read counts were excluded from the analysis by using the DESeq2 independent filtering approach. Gene ontology analysis was performed using the GOSeq R package⁶⁷ and pathway analysis using Panther⁶⁸.

ChIP-seq analysis. ChIP-seq and input libraries were sequenced and 50-bp single-end reads aligned to mouse genome mm9 using Bowtie version 0.12.8. Duplicate reads and reads aligning to more than one genomic position were discarded. Quality was assessed using ChIPQC⁶⁹. Genome-wide coverage tracks were generated using the 'coverage' function in the GenomicRanges R package, exported as bigwig files, and visualized using UCSC genome browser. ChIP-seq peaks were identified by MACS2⁷⁰ using input libraries. RAD21 consensus peaks were derived by taking the intersection of RAD21 peaks identified in each biological replicate. Genes were marked as RAD21-bound if there was a RAD21 peak overlapping or within 10 kb of the gene. Reads on enhancers²² were summarized using the summarizeOverlaps function of the GenomicAlignments R package. Enhancers with differential enrichment of H3K27Ac were identified by DESeq2.

GRO-seq analysis. GRO-seq libraries were sequenced as 50-bp single-end reads in two biological replicates. The 10 most 3' bases were discarded based on fastqc quality assessment. Reads were aligned to mouse genome mm9 using bowtie with arguments -l 30 -m 10 -n 2 --trim3 10. Read counts on enhancers were computed using summarizeOverlaps function from GeomicAlignments R Package. Differentially transcribed enhancers were identified using DESeq2.

Motif enrichment analysis. Enrichment of known transcription factor motifs in enhancer TSSs was performed using Homer's findMotifsGenome.pl program with default parameters⁷¹. The analysis was restricted to intergenic enhancer TSSs identified from GRO-seq signal⁴⁶ that were extended ± 100 bp. If an enhancer had multiple TSSs, all were included in the analysis. Strongly inducible enhancers were classified by DESeq2 analysis of H3K27ac in wild-type macrophages 1 or 6 h after LPS stimulation compared to unstimulated cells (\log_2 FC 1.5 and Benjamini and Hochberg-adjusted $P < 0.05$). Failed enhancers were identified by comparing H3K27ac in *Rad21*-deleted macrophages with wild-type macrophages at each time point (\log_2 FC = 0 and Benjamini and Hochberg-adjusted $P < 0.05$). Maintained enhancers were used as background for failed enhancer motif enrichment analysis and vice versa. Motif occurrences in enhancers were identified using Homer's findMotifsGenome.pl program.

Enhancer analysis. Unless otherwise indicated, the analysis of enhancers was based on ref. ²². Of 8,991 constitutive enhancers ("constitutive steady"²²), 3,775 were intergenic, and 7,082, 6,984 and 8,188 were included in DESeq2 at 0, 1 and 6 h. Of 6,708 inducible enhancers (union of "constitutive not steady," "poised activated" and "cryptic"²²), 2,893 were intergenic, and 3,713, 4,106 and 5,903 were included in DESeq2 at 0, 1 and 6 h. Of 11,146 LPS-repressed enhancers²², 4,914 were intergenic, and 8,787, 7,969 and 9,786 were included in DESeq2 at 0, 1 and 6 h. DESeq2 was used to identify enhancer deregulation within the three groups at each time point based on H3K27ac, GRO-seq or ATAC-seq (Benjamini and Hochberg-adjusted $P < 0.05$). FPKM values for H3K27ac, Rad21, H3 and GRO-seq datasets on enhancers were generated in R. Heat maps were generated using the heatmap3 R package. Superenhancers were defined using ROSE⁷². Peaks identified using H3K27ac ChIP-seq were used as input to ROSE. Promoters (TSS ± 2.5 kb) were excluded from the analysis.

ATAC-seq analysis. ATAC-seq libraries were sequenced as 100-bp paired-end reads in two biological replicates. FastQC and found that bases 35–100 were enriched for "Nextera transposase adapter" sequences. Therefore, reads were aligned to Mouse genome mm9 using bowtie v0.12.8 with arguments --chunkmbs 256 -S -n 2 -m 1 -p 8 -X 2000 by successively trimming 10 bases from the 3' end down to a read length of 40 bp. Uniquely aligned reads were retained. Duplicate reads were identified using Picard MarkDuplicates. Aligned reads in Watson strand were offset by +4 bp and reads aligned to Crick strand were offset by -5 bp as described⁴⁰. Reads from fragments <120 bp were considered unprotected. Accessibility peaks for each replicate were identified using MACS2⁷⁰ with arguments --nomodel --nolambda. We defined consensus accessibility peaks by taking the intersection of peaks from both biological replicates. Read counts on enhancers and accessibility peaks were computed using summarizeOverlaps and differentially accessible regions were identified by DESeq2. Accessibility plots were generated using SoGgi R Package with reads counts normalized to sequencing depth.

TCGA RNA-seq analysis. TCGA IlluminaHiSeq_RNASeqV2 dataset was obtained for 173 AML patients via the GDC Legacy Archive (<https://portal.gdc.cancer.gov/legacy-archive/search/f>). Raw gene counts for each patient were converted to counts per million (CPM) using the function cpm from the R/Bioconductor package edgeR (3.16.5)^{73,74}. Weakly expressed genes were removed if CPM was

< 1. Normalization was performed by trimmed mean of M -values (TMM)⁷⁵ using the calcNormFactors function in edgeR. R/Bioconductor package limma (3.30.12)⁷⁶ was applied for the differential expression analysis, and the function voom⁷⁷ was used to transfer raw counts to \log_2 (counts per million). Differential expression analysis between AMLs was performed by the lmFit and eBayes⁷⁸ functions in limma. Genes were ranked by moderated t -statistics and gene set enrichment analysis (GSEA)⁶⁶ was applied using hallmark gene sets from MSigDB⁷⁹. Oncoprint, mutations and clinical information were obtained from cBioportal⁸⁰.

5C analysis. After quality filtering, 101-nt paired-end reads were trimmed (4 bases at the 5' and 50 bases at the 3') using the fastx_trimmer tool (FASTX-Toolkit, http://hannonlab.cshl.edu/fastx_toolkit/). Trimmed reads were aligned to the primer pool using Novoalign (<http://novocraft.com>, version 3). Considering all possible forward-reverse pairs, interactions were summarized as a matrix. 5C data was analyzed at fragment level using HiTC (v1.18.1)⁸¹ and normalized using the square root of the coverage of each fragment³. Enhancer-promoter interactions were defined as interactions between a fragment overlapping an enhancer²² and a fragment overlapping an annotated promoter (Ensembl v67) using the linkOverlaps function from the InteractionSet package v1.2.1 with default parameters⁸². Data was visualized using GenomicInteractions v1.7.1 and Gviz (v1.18.2)^{83,84}. The normalized strength of the given set of interactions (for example, interactions involving inducible promoters) was compared at each time point to the normalized strength of all other enhancer-promoter interactions using a Wilcoxon rank sum test. P -values were corrected for multiple testing using the p.adjust function in R, and adjusted $P < 0.05$ was considered significant.

TAD analysis. CH12 TADs were defined using Tadttool⁸⁵ on preprocessed Hi-C matrices³ using the ninsulation algorithm with a window size of 400 kb and a cutoff of 0.15. To identify TADs enriched for classes of genes/enhancers, a binomial test approach was used. First, genes and enhancers were assigned to TADs using the findOverlaps function from the GenomicRanges package in R. Promoter regions were defined as 100-bp regions around annotated transcription start sites (Ensembl v67) and used to assign genes to TADs. Enhancers were assigned to TADs based on published enhancer coordinates²²; <1% of enhancers and promoters are assigned to 1 TAD.

For each class of enhancer and for each TAD, the number of enhancers of that class (for example, enhancers with downregulated H2K27ac), given the number of total enhancers in the TAD, was compared to the fraction of all enhancers in that class, using the binom.test function in R with alternative = "greater". The resulting P -values were corrected for multiple testing using the p.adjust function in R with method "BH" and domains were considered significantly enriched for a class of enhancers at adjusted $P < 0.05$. The same analysis was performed for genes.

Definition of IFN-dependent genes. IFN-dependent inducible genes included STAT target genes⁸⁶ and antiviral response genes⁸⁷ as curated previously²⁴.

Statistics and reproducibility. Statistical analysis were calculated with GraphPad Prism version 7 (two-tailed Student's t tests) or R version 3.2.3 (Fisher exact tests) as indicated in the figure legends. Statistical differences were considered significant when $P \leq 0.05$. Error bars are reported as s.e.m. Experiments were repeated independently at least three times.

Reporting Summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

Code availability. All code used in this study is available from the authors upon reasonable request.

Data availability. The data generated for this study have been deposited at the Gene Expression Omnibus (GEO) under accession code GSE108599. The data supporting the findings of this study are available from the corresponding author upon reasonable request.

References

- Seibler, J. et al. Rapid generation of inducible mouse mutants. *Nucleic Acids Res.* **31**, e12 (2003).
- Cobb, B. S. et al. Targeting of Ikaros to pericentromeric heterochromatin by direct DNA binding. *Genes Dev.* **14**, 2146–2160 (2000).
- Carpenter, A. E. et al. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* **7**, R100 (2006).
- Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013).
- Splinter, E., de Wit, E., van de Werken, H. J. G., Klous, P. & de Laat, W. Determining long-range chromatin interactions for selected genomic sites using 4C-seq technology: from fixation to computation. *Methods* **58**, 221–230 (2012).

62. van de Werken, H. J. G. et al. Robust 4C-seq data analysis to screen for regulatory DNA interactions. *Nat. Methods* **9**, 969–972 (2012).
63. Kim, D. et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).
64. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
65. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
66. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
67. Young, M. D., Wakefield, M. J., Smyth, G. K. & Oshlack, A. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* **11**, R14 (2010).
68. Mi, H. et al. PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res.* **45**, D183–D189 (2017). D1.
69. Carroll, T. S., Liang, Z., Salama, R., Stark, R. & de Santiago, I. Impact of artifact removal on ChIP quality metrics in ChIP-seq and ChIP-exo data. *Front. Genet.* **5**, 75 (2014).
70. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
71. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
72. Whyte, W. A. et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).
73. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
74. McCarthy, D. J., Chen, Y. & Smyth, G. K. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* **40**, 4288–4297 (2012).
75. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, R25 (2010).
76. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
77. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).
78. Phipson, B., Lee, S., Majewski, I. J., Alexander, W. S. & Smyth, G. K. Robust hyperparameter estimation protects against hypervariable genes and improves power to detect differential expression. *Ann. Appl. Stat.* **10**, 946–963 (2016).
79. Liberzon, A. et al. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).
80. Cerami, E. et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2**, 401–404 (2012).
81. Servant, N. et al. HiTC: exploration of high-throughput ‘C’ experiments. *Bioinformatics* **28**, 2843–2844 (2012).
82. Lun, A. T. L., Perry, M. & Ing-Simmons, E. Infrastructure for genomic interactions: Bioconductor classes for Hi-C, ChIA-PET and related experiments. *F1000Res.* **5**, 950 (2016).
83. Harmston, N., Ing-Simmons, E., Perry, M., Barešić, A. & Lenhard, B. GenomicInteractions: An R/Bioconductor package for manipulating and investigating chromatin interaction data. *BMC Genomics* **16**, 963 (2015).
84. Hahne, F. & Ivanek, R. Visualizing genomic data using Gviz and Bioconductor. *Methods Mol. Biol.* **1418**, 335–351 (2016).
85. Kruse, K., Hug, C. B., Hernández-Rodríguez, B. & Vaquerizas, J. M. TADtool: visual parameter identification for TAD-calling algorithms. *Bioinformatics* **32**, 3190–3192 (2016).
86. Amit, I. et al. Unbiased reconstruction of a mammalian transcriptional network mediating pathogen responses. *Science* **326**, 257–263 (2009).
87. Garber, M. et al. A high-throughput chromatin immunoprecipitation approach reveals principles of dynamic gene regulation in mammals. *Mol. Cell* **47**, 810–822 (2012).

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a | Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated
- Clearly defined error bars
State explicitly what error bars represent (e.g. SD, SE, CI)

Our web collection on [statistics for biologists](#) may be useful.

Software and code

Policy information about [availability of computer code](#)

Data collection

No software was used to collect data.

Data analysis

Prism 7
FlowJo v10
Microsoft Excel v14.5.5
GSEA (Java desktop version)
IGV v2.3.57
CellProfiler v2.2.0
R v3.2.3
R package ggplot v2.2.1
R package heatmap3

RNASeq:
Tophat version 2.0.10
Samtools version 0.1.18
htseq version 0.5.3p9
R package - DESeq2 version 1.10.1
R package - GenomicAlignments version 1.6.3
R package - GOseq v1.22.0

R package - edgeR (v3.16.5)
R package - limma (v3.30.12)

ChIPseq/GROSeq/ATAC-Seq/4C/5C:
R package ChIPQC v1.4.4
bowtie version 0.12.8
Samtools version 0.1.18
Picard version 1.90
MACS2 2.0.10
Homer v4.9.1
ROSE
R version 3.2.3
R package FASTQC v0.9.4
R package - rtracklayer version 1.30.4
R package - Rsamtools version 1.22.0
R package - GenomicRanges version 1.22.4
R package - DESeq2 version 1.10.1
R package - GenomicAlignments version 1.6.3
R package - Soggi version 1.8.0
R package - GenomicInteractions v1.7.1
R package - Gviz v.1.18.2
FASTX- toolkit
Novoalign v3
4CseqPipe software suite version 0.7
my5C software

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data generated for this study has been deposited at GEO under accession GSE108599.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/authors/policies/ReportingSummary-flat.pdf](https://www.nature.com/authors/policies/ReportingSummary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|-----------------|--|
| Sample size | No statistical test was used to determine sample size. Three or more biological replicates per group (as indicated in the figure legends) were used to allow statistical inference. In the case of genome-wide sequenced experiments (GRO-seq, ATAC-seq, ChIP-seq) two biological replicates per group were used according to current convention. |
| Data exclusions | No data were excluded except for standard quality control filtering during sequencing analysis. This includes filtering genes and enhancers with low read counts due to their low statistical power using a DESeq2 independent filtering approach. Standard filtering also involves removing reads aligning to multiple positions and duplicate reads, in order to exclude PCR bias. In the case of the human AML RNA-seq analysis, lowly expressed genes were removed if CPM was smaller than 1 in one of all group samples. In the GRO-seq analysis, analysis of quality of reads indicated that quality of the last 10bp at the 3' end was lower. Hence the last 10bp were excluded from the alignment. Finally, in order to define super-enhancers based on H3K27ac signal, promoter regions were excluded because they also could contain this histone mark and be misinterpreted as enhancers. |
| Replication | Replication attempts were successful and the results are represented as mean +/- SEM as indicated. |
| Randomization | No randomization was required. Mice used for bone marrow extraction were chosen according to genotype. |
| Blinding | Investigators were not blinded to group allocation. We considered that blinding was not relevant to this study because we used quantitative assays for all experiments and the computational framework was identical for all samples and replicates. |

Reporting for specific materials, systems and methods

Materials & experimental systems

- n/a Involved in the study
- Unique biological materials
- Antibodies
- Eukaryotic cell lines
- Palaeontology
- Animals and other organisms
- Human research participants

Methods

- n/a Involved in the study
- ChIP-seq
- Flow cytometry
- MRI-based neuroimaging

Antibodies

Antibodies used

Immunoblot: RAD21 (Abcam ab992) used at 1:1000 dilution, beta-Actin (Santa Cruz sc-69879) used at 1:10000 dilution, c-MYC (Santa Cruz sc-40 9E10) used at 1:1000 dilution. As secondary antibodies we used fluorescent (Alexa Fluor 680) goat anti-rabbit IgG and goat anti-mouse IgG at 1:10000 dilution (Life technologies A21109 and A21057). ChIP: anti-H3 (Abcam ab1791), anti-H3K27ac (Abcam ab4729) and Rad21 (Abcam ab992). For each IP, 1.5 ug of primary antibody was bound to Dynabeads protein G (Invitrogen 10004D) for incubation with sonicated extract. HSPC sorting: Sca-1-BV510 (BD 565507), cKit-PE (eBioscience 12-1171-81), CD16-APC (eBioscience 17-0161-81), CD34-FITC (eBioscience 11-0341-81), SAV-e450 (48-4317-82). HSPC staining: Sca-1-FITC (Biolegend 122505), cKit-Alexa Fluor 700 (eBioscience 56-1172-80), CD11b-APC-Cy7 (BD 557657) and CD16-BV605 (BD 563006). Antibodies for staining and sorting HSPC were used at 1:200. Macrophage staining: CD11b-FITC (BD 01714D) used at 1:100 dilution, F4/80-PE (eBioscience 12-4801-80) used at 1:200 dilution and anti-Mouse Vα 11.1, 11.2 TCR and Vα2 TCR (BD) were used as isotype controls.

Validation

All antibodies are commercially available and have been tested in mouse for the experiments they were used.

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals

Specific pathogen free male and female laboratory-bred mice of the indicated genotype on a mixed B6/129 background were used at 6-10 weeks of age.

Wild animals

No wild animals involved.

Field-collected samples

The study did not involve collecting samples from the field.

ChIP-seq

Data deposition

- Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).
- Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links

May remain private before publication.

GSE108599
reviewers token to access the data: spszqeogzlnhqn
<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE108599>

Files in database submission

GSM2905558 WT_unstimulated_Rad21_ChIP_rep1
GSM2905559 WT_LPS_1h_Rad21_ChIP_rep1
GSM2905560 WT_LPS_6h_Rad21_ChIP_rep1
GSM2905561 WT_unstimulated_Rad21_ChIP_rep2
GSM2905562 WT_LPS_1h_Rad21_ChIP_rep2
GSM2905563 WT_LPS_6h_Rad21_ChIP_rep2
GSM2905564 WT_unstimulated_Rad21_input_rep1
GSM2905565 WT_LPS_1h_Rad21_input_rep1
GSM2905566 WT_LPS_6h_Rad21_input_rep1
GSM2905567 WT_unstimulated_Rad21_input_rep2
GSM2905568 WT_LPS_1h_Rad21_input_rep2
GSM2905569 WT_LPS_6h_Rad21_input_rep2
GSM2913383 Rad21-/-_unstimulated_H3_ChIP_rep1
GSM2913384 Rad21-/-_LPS_1h_H3_ChIP_rep1

GSM2913385 Rad21-/-_LPS_6h_H3_ChIP_rep1
 GSM2913386 WT_unstimulated_H3_ChIP_rep1
 GSM2913387 WT_LPS_1h_H3_ChIP_rep1
 GSM2913388 WT_LPS_6h_H3_ChIP_rep1
 GSM2913389 Rad21-/-_unstimulated_H3_ChIP_rep2
 GSM2913390 Rad21-/-_LPS_1h_H3_ChIP_rep2
 GSM2913391 Rad21-/-_LPS_6h_H3_ChIP_rep2
 GSM2913392 WT_unstimulated_H3_ChIP_rep2
 GSM2913393 WT_LPS_1h_H3_ChIP_rep2
 GSM2913394 WT_LPS_6h_H3_ChIP_rep2
 GSM2913395 Rad21-/-_unstimulated_H3_K27Ac_input_rep1
 GSM2913396 Rad21-/-_LPS_1h_H3_K27Ac_input_rep1
 GSM2913397 Rad21-/-_LPS_6h_H3_K27Ac_input_rep1
 GSM2913398 WT_unstimulated_H3_K27Ac_input_rep1
 GSM2913399 WT_LPS_1h_H3_K27Ac_input_rep1
 GSM2913400 WT_LPS_6h_H3_K27Ac_input_rep1
 GSM2913401 Rad21-/-_unstimulated_H3_K27Ac_input_rep2
 GSM2913402 Rad21-/-_LPS_1h_H3_K27Ac_input_rep2
 GSM2913403 Rad21-/-_LPS_6h_H3_K27Ac_input_rep2
 GSM2913404 WT_unstimulated_H3_K27Ac_input_rep2
 GSM2913405 WT_LPS_1h_H3_K27Ac_input_rep2
 GSM2913406 WT_LPS_6h_H3_K27Ac_input_rep2
 GSM2913407 Rad21-/-_unstimulated_H3K27Ac_ChIP_rep1
 GSM2913408 Rad21-/-_LPS_1h_H3K27Ac_ChIP_rep1
 GSM2913409 Rad21-/-_LPS_6h_H3K27Ac_ChIP_rep1
 GSM2913410 WT_unstimulated_H3K27Ac_ChIP_rep1
 GSM2913411 WT_LPS_1h_H3K27Ac_ChIP_rep1
 GSM2913412 WT_LPS_6h_H3K27Ac_ChIP_rep1
 GSM2913413 Rad21-/-_unstimulated_H3K27Ac_ChIP_rep2
 GSM2913414 Rad21-/-_LPS_1h_H3K27Ac_ChIP_rep2
 GSM2913415 Rad21-/-_LPS_6h_H3K27Ac_ChIP_rep2
 GSM2913416 WT_unstimulated_H3K27Ac_ChIP_rep2
 GSM2913417 WT_LPS_1h_H3K27Ac_ChIP_rep2
 GSM2913418 WT_LPS_6h_H3K27Ac_ChIP_rep2

Genome browser session
 (e.g. [UCSC](#))

Not available.

Methodology

Replicates

Two biological replicates were performed per experiment.

Sequencing depth

ChIP-Seq libraries were sequenced along with input libraries as single end 50bp reads. Reads were aligned to mouse genome mm9 using Bowtie version 0.12.8 with default parameters. Reads aligning to multiple positions in the genome were discarded from the analysis.

| Sample | Total Reads | Aligned Reads | Reads% Duplicates | Duplicates% |
|----------------|-------------|---------------|-------------------|---------------------|
| KO0 H3 Rep1 | 22609822 | 15305559 | 67.69429233 | 1071078 6.997967209 |
| KO0 H3 Rep2 | 27119309 | 18032998 | 66.49504971 | 1148925 6.371236774 |
| KO0 Input Rep1 | 33288347 | 23591678 | 70.87068036 | 1285842 5.450405011 |
| KO0 Input Rep2 | 31358682 | 22124333 | 70.55249643 | 1385029 6.260206805 |
| KO0 K27 Rep1 | 25278150 | 21875600 | 86.53956085 | 2317774 10.59524767 |
| KO0 K27 Rep2 | 27781144 | 23593503 | 84.92631909 | 2787525 11.81479918 |
| KO1 H3 Rep1 | 25387465 | 16905434 | 66.58968905 | 1258427 7.443920103 |
| KO1 H3 Rep2 | 40703436 | 27296325 | 67.06147609 | 2138808 7.83551632 |
| KO1 Input Rep1 | 29805767 | 21093075 | 70.76843552 | 1113557 5.279253973 |
| KO1 Input Rep2 | 37127228 | 26389556 | 71.07871344 | 1985574 7.524090212 |
| KO1 K27 Rep1 | 29849706 | 25585312 | 85.71378224 | 2767563 10.81699922 |
| KO1 K27 Rep2 | 28677463 | 24273019 | 84.64144475 | 3002956 12.37158015 |
| KO6 H3 Rep1 | 31154307 | 21193449 | 68.02734851 | 1417149 6.686731357 |
| KO6 H3 Rep2 | 21022077 | 13748800 | 65.4017203 | 808317 5.879182183 |
| KO6 Input Rep1 | 22672509 | 16028248 | 70.69463728 | 808907 5.046758698 |
| KO6 Input Rep2 | 34350009 | 24004428 | 69.88186815 | 1513077 6.303324537 |
| KO6 K27 Rep1 | 27106962 | 23415327 | 86.38122929 | 2600942 11.1078611 |
| KO6 K27 Rep2 | 21982406 | 18628349 | 84.74208419 | 1918015 10.29621573 |
| WT0 H3 Rep1 | 28664543 | 19819582 | 69.14319897 | 1441557 7.273397592 |
| WT0 H3 Rep2 | 29712633 | 20463358 | 68.87090081 | 1429656 6.986419335 |
| WT0 Input Rep1 | 24275431 | 17342299 | 71.43971615 | 865354 4.989845925 |
| WT0 Input Rep2 | 39174296 | 28740889 | 73.36670198 | 1745459 6.073086327 |
| WT0 K27 Rep1 | 21350236 | 18711541 | 87.64090945 | 2041258 10.90908547 |
| WT0 K27 Rep2 | 30454900 | 26496451 | 87.00225908 | 3333386 12.58049993 |
| WT1 H3 Rep1 | 31286852 | 21357477 | 68.2634258 | 1566038 7.332504677 |
| WT1 H3 Rep2 | 45213347 | 30459749 | 67.36893201 | 2797620 9.184645612 |
| WT1 Input Rep1 | 41142016 | 30028759 | 72.98805921 | 1730887 5.764097677 |
| WT1 Input Rep2 | 22939630 | 16515463 | 71.99533297 | 906780 5.490490942 |

WT1 K27 Rep1 27223264 23657327 86.90114088 2515333 10.6323635
 WT1 K27 Rep2 36933781 31620977 85.61532598 3841506 12.14859996
 WT6 H3 Rep1 25821798 17739240 68.69870177 1239530 6.987503411
 WT6 H3 Rep2 32259224 21773607 67.49575563 1475326 6.775753783
 WT6 Input Rep1 28525927 20282270 71.10117754 1141329 5.627225158
 WT6 Input Rep2 29495520 21047517 71.35835205 1404818 6.674507021
 WT6 K27 Rep1 34666003 29973863 86.46472165 3722803 12.42016419
 WT6 K27 Rep2 32982649 28158509 85.37370361 3513445 12.47738295
 Sample Name Total Reads Aligned Reads Aligned Reads % Duplicates Reads Duplicate %
 WT0 Rad21 Rep1 48273783 35238650 72.99749017 4012543 11.3868
 WT1 Rad21 Rep1 56406004 39204817 69.50468783 4217228 10.7569
 WT6 Rad21 Rep1 109991223 76631453 69.67051635 6167779 8.0486
 WT0 Input Rep1 59364171 41486953 69.88550889 2405021 5.7971
 WT1 Input Rep1 80533810 55761614 69.24000491 3658310 6.5606
 WT6 Input Rep1 67528746 45837468 67.87845283 3005451 6.5568
 WT0 Rad21 Rep2 80939929 57562029 71.11697491 21695972 37.6915
 WT1 Rad21 Rep2 33049137 25007481 75.66757643 1918994 7.6737
 WT6 Rad21 Rep2 86963760 60705101 69.80505558 19573540 32.2436
 WT0 Input Rep2 38490800 27276533 70.86507165 897809 3.2915
 WT1 Input Rep2 18655009 13465319 72.18071565 373078 2.7707
 WT6 Input Rep2 37297789 26057953 69.86460511 835909 3.2079

Antibodies

H3: ab1791 Abcam
 H3K27ac: ab4729 Abcam
 Rad21: ab992 Abcam

Peak calling parameters

ChIP-Seq Peaks were identified by MACS2 using input libraries using default parameters. RAD21 consensus peaks were derived by taking the intersection of RAD21 peaks identified in each biological replicate. The 3 time points contain 53762, 43710, 37037 peaks respectively (5%FDR).

Data quality

Reads were aligned to the mouse genome mm9 using Bowtie version 0.12.8 with default parameters. Reads aligning to multiple positions in the genome were discarded from the analysis. Quality of the ChIP-Seq libraries were assessed using ChIPQC. Duplicate reads were identified using Picard MarkDuplicates and excluded from the downstream analysis.

Software

R package ChIPQC v1.4.4
 bowtie version 0.12.8
 Samtools version 0.1.18
 Picard version 1.90
 MACS2 2.0.10
 R version 3.2.3
 R package - rtracklayer version 1.30.4
 R package - Rsamtools version 1.22.0
 R package - GenomicRanges version 1.22.4
 R package - DESeq2 version 1.10.1

Flow Cytometry

Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

Bone marrow was extracted from femurs and tibiae using 25G needles and complete DMEM media. Total bone marrow cells, bone marrow derived macrophages and sorted HSPCs were stained with the corresponding antibodies diluted in PBS-2%FCS for 20 minutes in the dark at 4 degrees.

Instrument

FACSAria Fusion

Software

BD FACSDiva 8.0
 FlowJo v10
 Prism 7

Cell population abundance

HSPC populations were sorted from Lineage negative bone marrow cells. LSK population was usually 1-3% of the lineage

Cell population abundance

negative population. cKit+ Sca1- were ~30% of the lineage negative, and within this population CMP and GMP were usually 20-30% each.

Gating strategy

Cells were first gated for live cells (FSC/SSC) and gated to exclude doublets. To exclude lineage-positive cells, total bone marrow cells were first stained with biotin-lineage antibodies (CD4, CD8, B220, CD19, NK1.1, CD11b, Ter119, GR-1) and subsequently depleted with streptavidin beads (Miltenyi 130-048-102), further confirmed by streptavidin-negative gating (SAV-e450). LSK were Sca-1 and cKit positive. The Kit-positive Sca-1 negative population was gated, and from this gate CD16-intermediate CD34-positive were sorted as CMP, and CD16-high CD34-positive were sorted as GMP.

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.