nature genetics

LETTERS

https://doi.org/10.1038/s41588-020-0643-0

CTCF is dispensable for immune cell transdifferentiation but facilitates an acute inflammatory response

Grégoire Stik[1]✉, Enrique Vidal[1], Mercedes Barrero[1], Sergi Cuartero[1,2], Maria Vila-Casadesús[1], Julen Mendieta-Esteban[3], Tian V. Tian[1,4], Jinmi Choi[5], Clara Berenguer[1,2], Amaya Abad[1], Beatrice Borsari[1], François le Dily[1], Patrick Cramer[5], Marc A. Marti-Renom[1,3,6], Ralph Stadhouders[7,8]✉ and Thomas Graf[1,9]✉

Three-dimensional organization of the genome is important for transcriptional regulation[1–7]. In mammals, CTCF and the cohesin complex create submegabase structures with elevated internal chromatin contact frequencies, called topologically associating domains (TADs)[8–12]. Although TADs can contribute to transcriptional regulation, ablation of TAD organization by disrupting CTCF or the cohesin complex causes modest gene expression changes[13–16]. In contrast, CTCF is required for cell cycle regulation[17], embryonic development and formation of various adult cell types[18]. To uncouple the role of CTCF in cell-state transitions and cell proliferation, we studied the effect of CTCF depletion during the conversion of human leukemic B cells into macrophages with minimal cell division. CTCF depletion disrupts TAD organization but not cell transdifferentiation. In contrast, CTCF depletion in induced macrophages impairs the full-blown upregulation of inflammatory genes after exposure to endotoxin. Our results demonstrate that CTCF-dependent genome topology is not strictly required for a functional cell-fate conversion but facilitates a rapid and efficient response to an external stimulus.

Lineage-instructive transcription factors establish new cell identities by activating a new gene expression program while silencing the old one. Whereas this is largely achieved through binding to promoters and enhancers, genome topology has recently emerged as a new player in gene regulation. Chromatin contact maps obtained by chromosome conformation capture techniques, such as Hi-C, revealed that chromatin can be separated at the megabase (Mb) level into active (A) and inactive (B) compartments[19], which are themselves subdivided into TADs. Large deletions overlapping boundaries can cause fusion of adjacent domains, which can lead to developmental abnormalities[20]. In addition, inversion or deletion of individual CTCF-binding sites can induce a loss of specific contacts or insulation from active chromatin[21–23]. Mechanistically, genomic insulation by TADs is thought to facilitate enhancer–promoter interactions while inhibiting cross-boundary communication between regulatory elements to prevent aberrant gene activation[1]. Hence, the importance of CTCF and TAD organization in facilitating transcriptional rewiring during cell-state transitions—often accompanied by extensive cell division—remains controversial[24].

We have recently developed a system that is particularly suitable for studying the role of CTCF in cell-state transitions, which consists of a B leukemia cell line (BLaER) that can be efficiently converted by exogenous CEBPA expression into functional induced macrophages (iMacs) with only one cell division on average (Fig. 1a and Supplementary Note 1)[25]. Using this system, we analyzed a time series of transdifferentiating cells for genome-wide changes in three-dimensional (3D) genome organization (in situ Hi-C), enhancer activity (ChIP–seq of histone modifications), chromatin accessibility (ATAC-seq) and gene expression (RNA-seq).

We first determined genome segmentation into A and B compartments on the basis of the first eigenvector values of a principal component analysis (PCA) on the Hi-C correlation matrix ('PC1 values'). Overall, although most of the genome remained stable, around 14% of A or B compartments were dynamic during transdifferentiation, showing transcriptional changes correlating with the altered compartmentalization (Fig. 1b–f, Extended Data Fig. 1a–d and Supplementary Note 2). Next, we used chromosome-wide insulation potential[26] to identify between 3,100 and 3,300 TAD borders per time point (Fig. 1g). Boundaries were highly reproducible between biological replicates (Jaccard index > 0.99) and enriched in binding sites for CTCF (Extended Data Fig. 1e). Genome-wide insulation scores analyzed by PCA over time revealed progressive changes, reflecting a transdifferentiation trajectory (Extended Data Fig. 1f). While 70% of TAD borders were stable across all stages, 18% were lost or gained and 12% were transiently altered (Fig. 1g). CTCF binding was substantially more enriched at stable than at dynamic boundaries (Fig. 1h), as observed earlier[27]. Furthermore, while lost borders showed some CTCF occupancy in B cells that decreased in iMacs, gained borders were depleted for CTCF in both cell states (Fig. 1h), indicating CTCF-independent mechanisms driving local insulation. The dynamic rearrangement of TAD borders during transdifferentiation is illustrated by the *DDX54* locus (Fig. 1i), in which a new boundary appears in iMacs without apparent changes in CTCF binding. Furthermore, border gain or loss did not correlate with changes in local gene expression (Extended

---

[1]Centre for Genomic Regulation (CRG) and Institute of Science and Technology (BIST), Barcelona, Spain. [2]Josep Carreras Leukaemia Research Institute (IJC), Barcelona, Spain. [3]CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Barcelona, Spain. [4]Vall d'Hebron Institute of Oncology (VHIO), Barcelona, Spain. [5]Max Planck Institute for Biophysical Chemistry, Göttingen, Germany. [6]ICREA, Barcelona, Spain. [7]Department of Pulmonary Medicine, Erasmus MC, Rotterdam, the Netherlands. [8]Department of Cell Biology, Erasmus MC, Rotterdam, the Netherlands. [9]Universitat Pompeu Fabra (UPF), Barcelona, Spain. ✉e-mail: gregoire.stik@crg.eu; r.stadhouders@erasmusmc.nl; thomas.graf@crg.eu

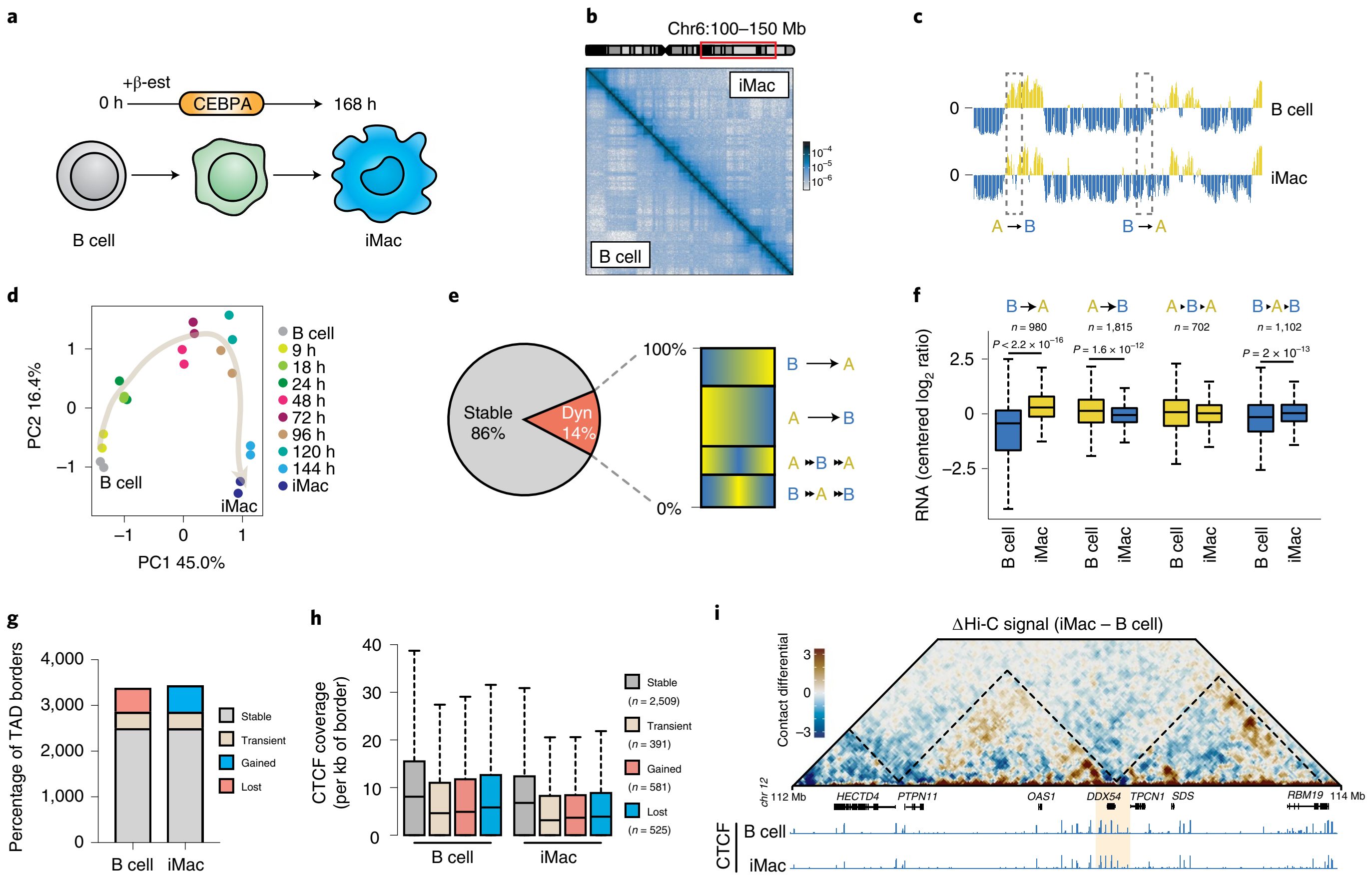


Fig. 1 | Transcription-factor-driven transdifferentiation rewires nuclear compartments and modulates TAD borders independently of CTCF binding. **a**, Schematic overview of the transdifferentiation system. CEBPA-ER in B cells (BLaER cell line) translocates to the nucleus after β-estradiol (β-est) treatment, activating the factor. A week after treatment the cells convert into induced macrophages ('iMac' stage). **b**, Representative in situ Hi-C contact maps (100-kb resolution) of a 50-Mb DNA region of B cells and iMacs. The color scale represents the normalized number of contacts per read. **c**, Transformation of the Hi-C map on the basis of the PC1 values of a PCA on the Hi-C correlation matrix. PC1 values for A and B compartments are shown in yellow and blue, respectively; dotted rectangles highlight local compartment changes during transdifferentiation. **d**, PCA of PC1 compartment values ($n = 28{,}749$ bins), with the beige arrow indicating the transdifferentiation trajectory. **e**, Proportion of dynamic compartment bins (Dyn), including the distribution of its different subcategories. **f**, Integration of gene expression associated with the dynamic compartment using RNA-seq (n represents the number of genes and P values are calculated using a two-sided Wilcoxon rank-sum test). **g**, Number of stable, transiently changed, gained or lost TAD borders in B cells and iMacs. **h**, CTCF-peak coverage at the different types of borders (n represents the number of borders in each category) in B cells and iMacs. **i**, Differential contact map (iMac minus B cell signal) at the *DDX54* locus (top). Color scale represents differential contacts per 100,000 reads. Snapshot of genome browser showing CTCF ChIP–seq signals at the locus (bottom). CTCF peaks at the newly created border are highlighted. All box plots depict the first and third quartiles as the lower and upper bounds of the box, with a thicker band inside the box showing the median value and whiskers representing 1.5× the interquartile range.

Data Fig. 1g), indicating that transcription is not a driver of the observed changes. However, whereas motif analysis at ATAC-seq peaks within stable borders did indeed show a strong enrichment for the CTCF motif, dynamic borders were enriched for PU.1 and EBF1 motifs (Extended Data Fig. 1h), raising the possibility that lineage-restricted transcription factors are involved in disrupting and/or establishing these borders.

To directly assess the importance of CTCF during CEBPA-induced transdifferentiation we devised an auxin-inducible degron approach[28] (Fig. 2a and Supplementary Note 3). Addition of auxin to these cells triggered proteasome-dependent CTCF degradation, resulting in a loss of mCherry$^+$ cells and rapid CTCF depletion to levels that were undetectable by western blot (Fig. 2b,c). Likewise, 80% of CTCF peaks were no longer detected after auxin treatment, and the enrichment level of persistent peaks was substantially reduced (Extended Data Fig. 2a,b), as previously described for mouse embryonic stem cells[14]. We next performed Hi-C on cells cultured in the presence of auxin or dimethyl sulfoxide (DMSO) (as a control) at 24 and 168 h postinduction (hpi) of transdifferentiation. Scaling of contact probabilities as a function of genomic separation did not change after CTCF depletion (Extended Data Fig. 2c). Analysis of chromosome-wide insulation potential in wild-type and CTCF-AID B cells showed that fusing the mAID tag to CTCF only had a negligible impact on TAD organization (Extended Data Fig. 2d,e). However, ~70% of TAD borders became undetectable and not visible in Hi-C contact maps after auxin treatment, both at 24 hpi and at iMac stages (Fig. 2d and Extended Data Fig. 2f). Overall, insulation scores at borders detected in the control cells (DMSO) were dramatically reduced upon auxin treatment, both at 24 hpi and at iMac stages (Extended Data Fig. 2g). Consequently, the ratio of contact enrichment inside TADs and outside TADs was also strongly decreased (Extended Data Fig. 2h). Whereas stable borders exhibited a dramatic loss of insulation after CTCF depletion, dynamic borders showed no change essentially (Fig. 2e), in agreement with their low CTCF occupancy (Fig. 1h).

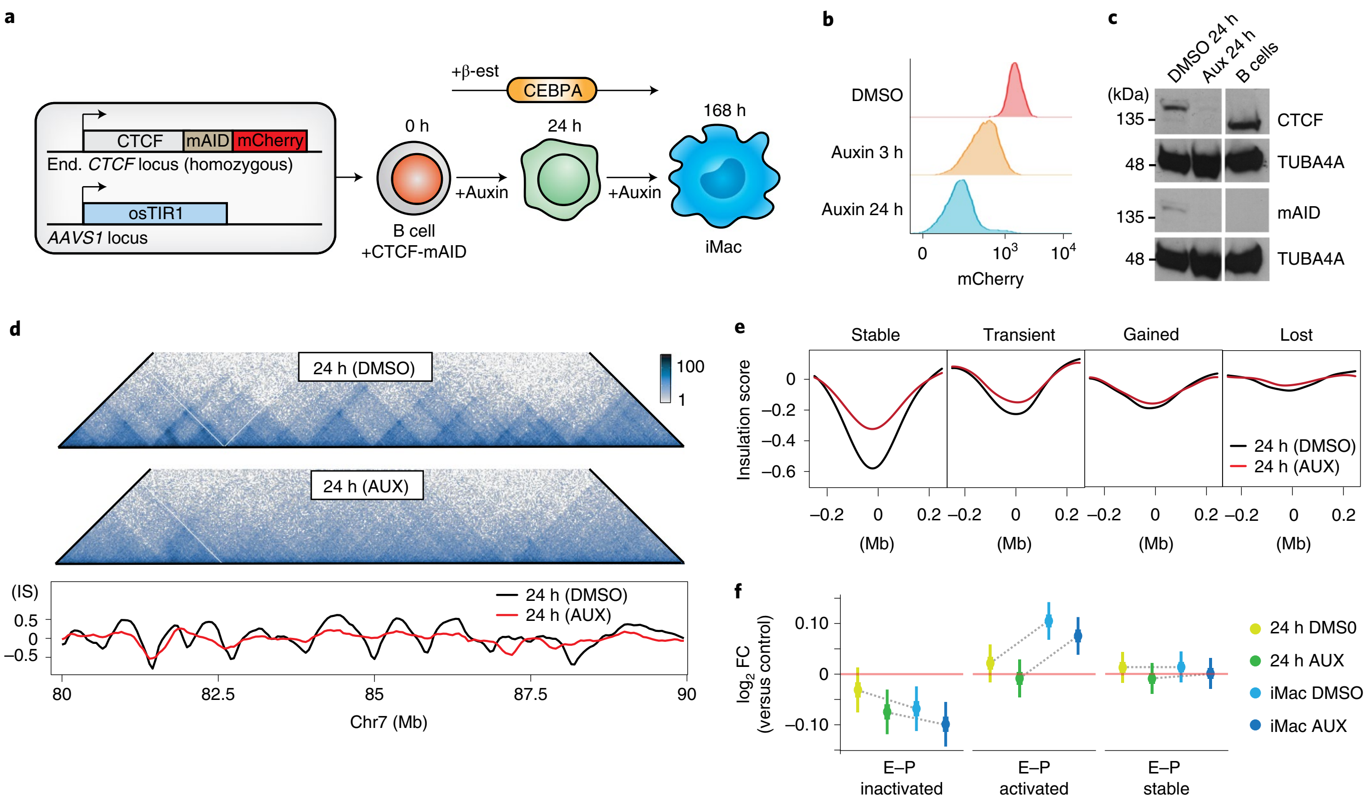


Fig. 2 | Auxin-mediated depletion of CTCF impairs chromatin insulation at stable but not at dynamic TAD boundaries. a, Schematic representation of auxin-mediated endogenous (End.) CTCF degradation, showing the constructs used and the design of the experimental setup. **b**, Flow cytometry analysis showing decreased mCherry fluorescence intensity after auxin treatment, as a proxy for CTCF levels. The experiment was repeated three times with similar results. **c**, Western blot showing loss of CTCF in CTCF-mAID B cells treated with auxin. TUBA4A was used as a loading control. The blots have been cropped from the original blots, which are available in the Source Data. The experiment was repeated three times with similar results. **d**, Representative Hi-C contact maps (20-kb resolution) of a 10-Mb region in chromosome 7 from transdifferentiating cells (24 hpi) treated with DMSO or auxin (top). The color scale represents the normalized number of contacts. Insulation score (IS) line graphs across the locus (bottom). **e**, Insulation scores at stable, transient, gained and lost borders of cells treated with DMSO or auxin. Areas shown are centered on boundary regions ±250 kb. **f**, Changes of enhancer–promoter (E–P) intra-TAD contacts during transdifferentiation with DMSO ($n=2$ biologically independent samples) or auxin ($n=2$ biologically independent samples) in comparison to B cells ($n=1$). Dots represent point estimates and bars (wide and narrow) indicate confidence intervals (50% and 95%, respectively) for the $\log_2$ FC, where FC is fold change. Estimations are computed using all nine samples in a single linear mixed model.

We next used ATAC-seq and H3K4me1/H3K27ac ChIP–seq to identify promoters and enhancers that are either activated, inactivated or remain stable during transdifferentiation (Extended Data Fig. 2i–l; Methods). Transdifferentiation was accompanied by extensive chromatin-state dynamics focused at enhancers, which were preferential targets of CEBPA binding (Extended Data Fig. 2j,k). We then interrogated how CTCF depletion affects intra-TAD enhancer–promoter (E–P) contacts at 0 h, 24 hpi and in iMacs. E–P interaction frequencies noticeably decreased during inactivation, which was somewhat accelerated in auxin-treated cells at 24 hpi and at iMac stages (Fig. 2f). Similarly, E–P interaction frequencies noticeably increased during activation, while E–P interactions at stable regulatory elements were not affected by CTCF depletion (Fig. 2f). These data demonstrate that, although auxin-treated samples show a minor overall reduction of intra-TAD E–P contacts, E–P interaction dynamics accompanying transdifferentiation seem to be independent of CTCF.

To assess whether CTCF depletion and a loss of TAD organization impacts CEBPA-induced transdifferentiation, we monitored the expression of the B cell marker CD19 and the macrophage marker Mac1 (CD11b) by flow cytometry at 0, 24, 48, 96 and 168 hpi. Surprisingly, CTCF-depleted cells converted into macrophage-like cells with even slightly accelerated kinetics at intermediate time points (Fig. 3a and Extended Data Fig. 3a), which was confirmed using a different clone of CTCF-mAID B cells (Extended Data Fig. 3b). The iMacs obtained under conditions of CTCF depletion were phagocytic and they activated inflammatory cytokine genes in response to endotoxin treatment (Extended Data Fig. 3c,d). Our findings show that CTCF depletion and widespread loss of TAD organization neither blocks nor delays transdifferentiation of B cells into functional macrophages.

We next analyzed how gene expression is affected upon CTCF depletion in cells at 24 hpi and at the iMac stage. A PCA of the entire transcriptome showed that CTCF depletion does not impair the overall rewiring of gene expression induced during cell-fate conversion (Fig. 3b). Instead, auxin-treated cells were more advanced towards transdifferentiation at 24 hpi, which was further confirmed by analyzing gene expression dynamics of B cell and myeloid cell signature genes (Fig. 3b,c). These observations agree with previous findings suggesting that a partial knockdown of CTCF accelerates myeloid commitment of common myeloid precursor cells[29]. A heat map of 8,595 annotated transcripts that changed significantly (fold change > 2, $P<0.05$) during transdifferentiation highlighted the overall similarity between DMSO- and auxin-treated samples (Fig. 3d). In fact, 76% of differentially expressed genes in control iMacs were similarly regulated under conditions of CTCF depletion (Extended Data Fig. 3e). This is illustrated for cell-type-restricted transcription factors by activation of *CEBPB*, *JUN*, *CEBPA* and

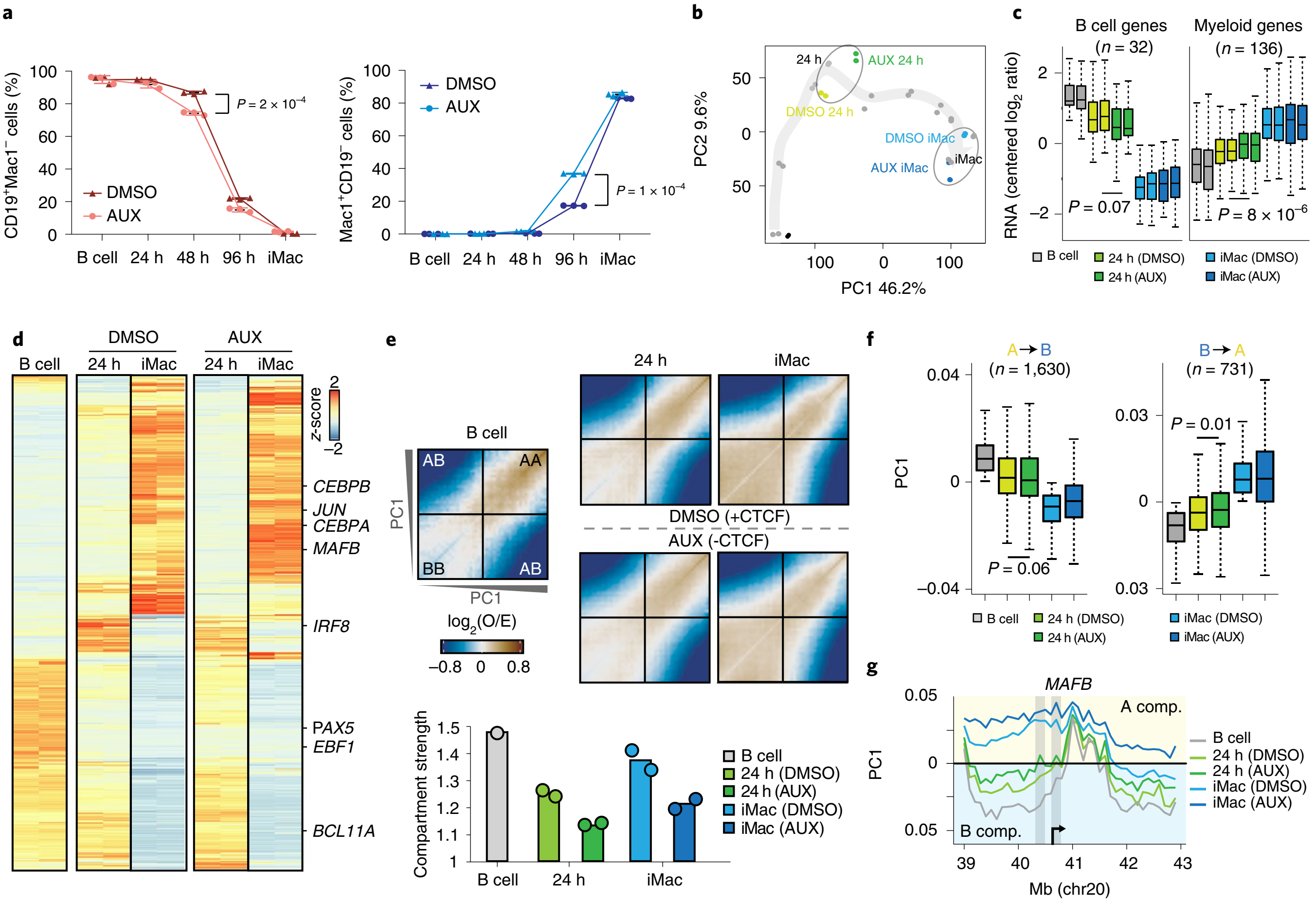


Fig. 3 | CTCF is dispensable for transcription-factor-induced cell-fate conversion. **a**, Flow cytometry analysis during transdifferentiation of cells treated with DMSO or auxin (AUX). Graphs show percentages of CD19$^+$Mac1$^-$ cells (left) or CD19$^-$Mac1$^+$ (right) cells ($n = 3$ biologically independent samples, error bars show standard deviation and *P* from unpaired two-tailed *t*-test). **b**, PCA analysis of transcriptome changes during transdifferentiation of CTCF-mAID B cells treated with DMSO or auxin ($n = 23,680$ genes). Gray points connected by an arrow represent nontagged B cell transdifferentiation. The ellipses group 24 hpi and iMac stage samples. **c**, RNA expression of selected B cell ($n = 32$) and myeloid cell genes ($n = 136$) during transdifferentiation with DMSO and auxin for the two biological replicates (*P*, two-sided Wilcoxon rank-sum test). **d**, Heat map of differentially expressed annotated transcripts ($n = 2$ biologically independent samples, FC > 2 and $P < 0.01$, two-tailed likelihood ratio test followed by FDR correction) in cells treated with DMSO or auxin. Myeloid and B cell regulator genes are indicated on the right. **e**, Saddle plot showing pairwise enrichment of the 20% top and bottom PC1 values from Hi-C contacts at 100-kb bins (Methods) (top part). Compartmentalization strength scores derived from B cell ($n = 1$) and DMSO ($n = 2$) or auxin-treated cell ($n = 2$) biologically independent samples. The score corresponds to the ratio between same-compartment and different-compartment contacts (diagonal corners over antidiagonal corners in the saddle plots). **f**, Average PC1 values of dynamic compartment bins (A to B, $n = 1,630$, and B to A, $n = 731$) in B cell ($n = 1$) and DMSO ($n = 2$) or auxin-treated cell ($n = 2$) biologically independent samples (*P*, two-sided Wilcoxon rank-sum test). **g**, Plot of PC1 values (100-kb bins) at the *MAFB* locus during transdifferentiation in the presence of DMSO or auxin. comp., compartment. All box plots depict the first and third quartiles as the lower and upper bounds of the box, with a thicker band inside the box showing the median value and whiskers representing 1.5× the interquartile range.

MAFB; by transient upregulation of *IRF8*; and by silencing of *PAX5*, *EBF1* and *BCL11A* in a similar fashion under both conditions (Fig. 3d). Although iMacs produced in the presence of auxin functionally resemble macrophages, they still show substantial differences in gene expression (~13% of expressed genes) compared to DMSO controls, mostly involving ubiquitous cellular processes, such as the cell cycle, GTPase signaling or ribosome biogenesis (Extended Data Fig. 3f).

The finding that CTCF depletion impacts TAD organization without substantially altering transdifferentiation capacity and kinetics prompted us to explore other features of 3D genome organization. Analyzing our Hi-C data for inter-TAD, long-range E–P interactions (5–10 Mb) revealed that their activation or inactivation is associated with the formation or dissolution of interacting clusters, respectively (Extended Data Fig. 3g). Remarkably, this occurred independently of CTCF (Extended Data Fig. 3g), suggesting that the observed 3D clusters are linked to compartmentalization changes involving transcription factors bound to these regulatory regions. Further analyses revealed that 78% of the regions that switched from one compartment to another during transdifferentiation do so in both DMSO- and auxin-treated cells (Extended Data Fig. 3h), showing that CTCF is largely dispensable for these large-scale genome rearrangements. In line with findings from a previous study[14], we observed ~10% reduction in compartment strength in auxin-treated samples (Fig. 3e), which could explain the slight acceleration of compartment transitions observed in auxin-treated cells at 24 hpi (Fig. 3f). An example is provided by the *MAFB* locus, a myeloid-expressed gene that is upregulated during transdifferentiation (Extended Data Fig. 3i), whereby the B-to-A switch was faster and more pronounced in auxin-treated cells than in DMSO controls

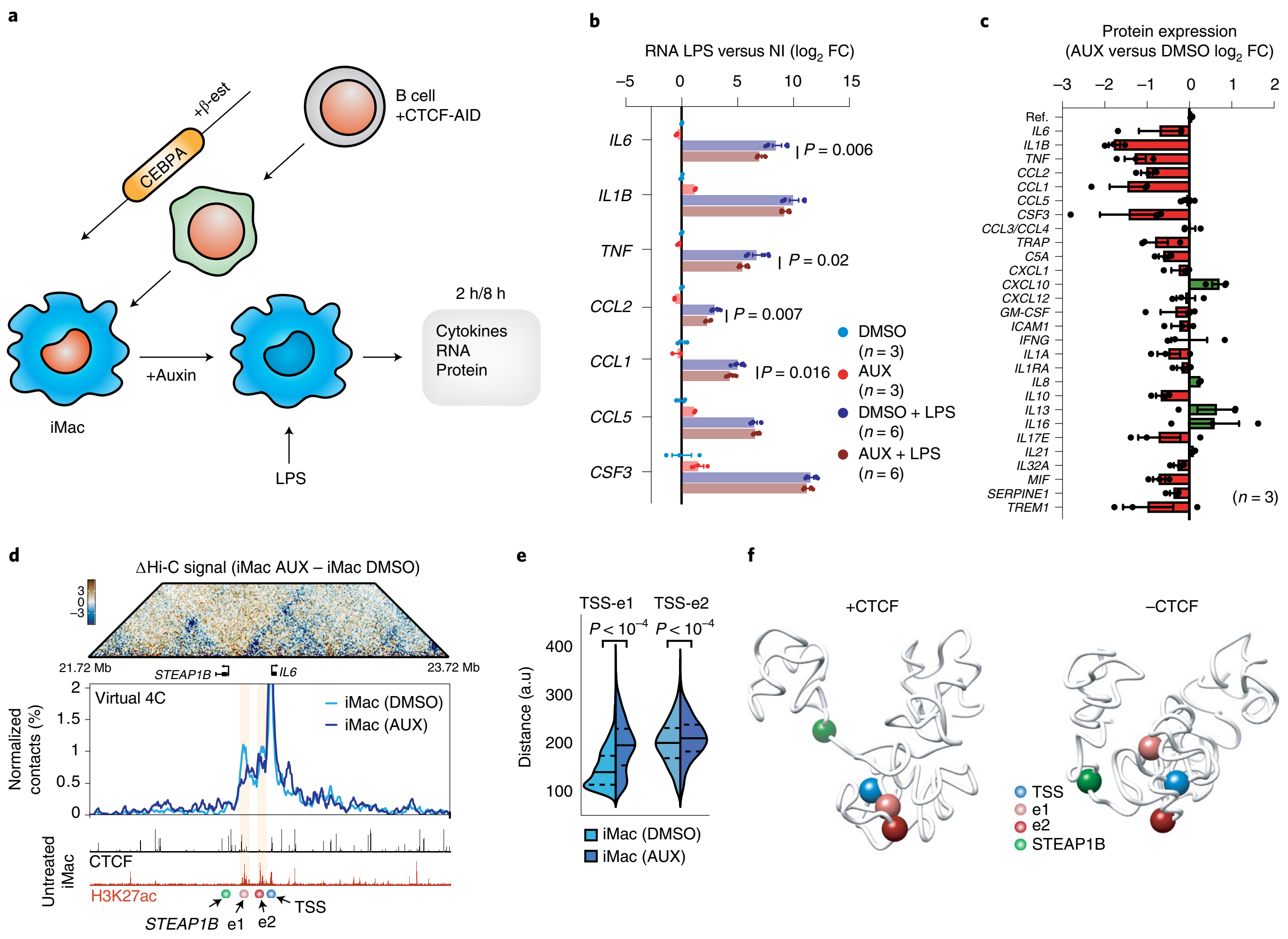


Fig. 4 | CTCF depletion attenuates the acute inflammatory response of iMacs to endotoxin. a, Schematic overview of the experiment. iMacs generated in the presence of CTCF were treated with either DMSO or auxin for 24 h followed by 2–8 h of LPS treatment and were assayed for cytokine expression. **b**, RNA expression, measured by qRT–PCR of selected cytokines in noninduced (NI) or 2-h LPS-induced iMacs pretreated with DMSO or auxin. Error bars represent standard error, sample sizes (*n*) are indicated and represent biologically independent samples and *P* values derive from unpaired two-tailed *t*-test. **c**, Secreted cytokine levels by iMacs treated with DMSO or auxin and stimulated for 8 h with LPS. Error bars represent standard error ($n=3$ biologically independent samples) **d**, Differential in situ Hi-C contact maps (10-kb resolution) at the *IL6* locus (chr7: 21.72–23.72 Mb) in iMacs generated in the presence of DMSO or auxin. The color scale represents differential contacts per 100,000 reads. The location of the *IL6* and *STEAP1B* genes is indicated (top). Virtual 4C extracted from Hi-C data at the *IL6* locus in iMacs treated with DMSO or auxin, using the *IL6* TSS as a viewpoint (middle). Browser snapshot showing CTCF and H3K27ac ChIP-seq signals. *IL6* enhancers (e1 and e2) are highlighted and green, red and blue spheres represent the *STEAP1B* promoter, the *IL6* enhancers and the *IL6* TSS, respectively (bottom). **e**, Distance distribution between TSS and enhancer regions ($n=1{,}000$ 3D models on the basis of Hi-C data of iMacs treated with DMSO or auxin). Median (solid line), first and third quartiles (dashed line) are indicated (*P*, two-sided Komogorov–Smirnov test). **f**, 3D chromatin conformation model of the *IL6* locus in DMSO or auxin-treated cells.

(Fig. 3g), including at an enhancer region that becomes decorated with H3K27ac at 24 hpi (Extended Data Fig. 3j). In short, our Hi-C data revealed that, although CTCF appears dispensable for genome compartmentalization, its depletion slightly decreased the strength of compartmentalization, which could facilitate compartmental rearrangements.

Previous reports indicated that CTCF plays a role in controlling macrophage gene expression[30] and that cohesin is required for an optimal inflammatory response of macrophages[31]. This raised the possibility of the involvement of genome topology in mounting an acute inflammatory response, as CTCF is known to stabilize the interaction between cohesin and chromatin[32,33]. Accordingly, aggregates of our Hi-C signals at previously described cohesin-bound loops[34] showed that these interactions disappeared after auxin treatment (Extended Data Fig. 4a). Using a public dataset of lipopolysaccharide (LPS)-responsive genes[35] we found that both the enhancers and promoters of such genes are enriched for CTCF and that their promoters are closer to enhancers as compared to unresponsive genes (Extended Data Fig. 4b–d). Therefore, we tested the effect of CTCF depletion in iMacs exposed to LPS for 2 h (Fig. 4a), revealing a reduced induction of critical LPS-responsive genes, such as *IL6*, *TNFA* and *CCL2* (Fig. 4b). Even more pronounced changes were observed at the level of secreted cytokines 8 h after LPS treatment (Fig. 4c). We next used RNA-seq to investigate the genome-wide effect of CTCF depletion on LPS-treated iMacs. Out of 39,963 detected genes, 746 were found to be significantly upregulated ($P<0.01$), although pathway enrichment analysis could not detect any significant associations (Extended Data Fig. 4e,f). Conversely, the 694 downregulated genes (among them *IL6*, *TNFA* and *CCL2*) were strongly associated with pathways related to the inflammatory response to bacterial stimuli (Extended Data Fig. 4e,f). Although a sizable fraction of differentially expressed genes was already altered in auxin-treated iMacs before

LPS stimulation, the total number of affected genes doubled after LPS exposure (Extended Data Fig. 4g) and the observed upregulation upon LPS treatment was noticeably blunted after CTCF depletion (Extended Data Fig. 4h). Of note, the expression of these genes was not noticeably changed in CTCF-depleted iMacs before LPS treatment (Extended Data Fig. 4i) and most of the key transcription factors and receptors involved in the LPS response were unaffected after 24 h of auxin treatment (Extended Data Fig. 4j), suggesting a direct role for CTCF in fine-tuning the expression of inflammatory response genes. A similar proportion of promoters of upregulated or downregulated genes after CTCF depletion and LPS stimulation were bound by CTCF in iMacs (Extended Data Fig. 4k), indicating no dominant role for CTCF as a promoter–proximal repressor[36] in this context. A phagocytosis assay with DMSO- or auxin-treated iMacs showed that, although CTCF-depleted cells were still functional, the number of engulfed beads per cell was reduced (Extended Data Fig. 4l,m), in line with the observed attenuation of the acute inflammatory response.

We next investigated whether CTCF-mediated 3D genome organization could underlie the apparent sensitivity of inflammatory response genes to CTCF depletion. Genes activated by LPS that were downregulated in CTCF-depleted iMacs were located closer to TAD borders and were also more strongly insulated than random gene sets (Extended Data Fig. 5a,b), suggesting that they could be extra susceptible to deregulation by loss of CTCF. To validate this and assess the impact of CTCF depletion on E–P interactions at key inflammatory response genes, we used our Hi-C data to conduct a 'virtual' 4C analysis of the *IL6* and *CCL2* loci, centered on their promoters. Active enhancers within these loci were identified by H3K27ac enrichment. At the *IL6* locus, CTCF depletion not only disrupted insulation from neighboring TADs, but also decreased the frequency of *IL6* E–P interactions (Fig. 4d and Extended Data Fig. 5c). Interestingly, the neighboring gene, *STEAPB1*, located just upstream of the *IL6* TAD, was found to be ectopically expressed upon CTCF depletion, likely resulting from aberrant contacts with *IL6* enhancers that are normally suppressed by the *IL6* TAD border (Extended Data Fig. 5d,e). To gain further insight into local chromatin conformation changes we generated 3D models of the *IL6* locus using Hi-C interaction data, transforming the interaction frequencies between genomic segments into spatial restraints[37]. This revealed that, initially, the *IL6* locus resides in a constrained space isolated from adjacent regions and that, upon CTCF depletion, the regions collapsed into less well-defined domains, separating the enhancers from their cognate target promoter (Fig. 4e,f). These models also confirm the decreased distance between *STEAPB1* and the *IL6* enhancers in the absence of CTCF (Fig. 4f and Extended Data Fig. 5f). Similar observations were made at the *CCL2* locus, where CTCF depletion also induced a loss of chromatin insulation and a decrease in E–P contacts (Extended Data Fig. 5g–j). These findings indicate that in macrophages, CTCF-mediated chromatin insulation and E–P interactions maintain acute inflammatory response genes in a primed configuration, permitting their rapid and robust activation in response to bacterial stimuli.

Our study has shown that the architectural protein CTCF is dispensable for the transdifferentiation of B cells into macrophages, while it is required for a full-blown inflammatory response. These findings indicate that CTCF-mediated genome topology, including TADs formed by cohesin-mediated loop extrusion, is not essential for developmental gene regulation, but instead provides robustness and precision to an acute transcriptional response to bacterial endotoxins. Nevertheless, we cannot exclude that in other biological contexts, gene regulatory circuits especially dependent on CTCF-mediated genome topology might be more critically relevant. Importantly, our study uncouples the critical role of CTCF in cell proliferation from its role as genome organizer and transcriptional regulator, and provides nuanced insights into the role of 3D chromatin organization for gene regulation. The observation that genome-wide CTCF is dispensable for a mammalian cell-state transition notably extends recent findings showing that CTCF or cohesin depletion during steady-state conditions, TAD rearrangements in flies or deletion of CTCF-mediated TAD boundaries in mice only caused minor changes in gene expression[13,14,38–40]. In addition, our findings indicate a critical role for 3D chromatin organization in providing an optimal response to external signals, in agreement with studies of developmentally regulated loci and nuclear hormone receptor signaling[38,41]. Future studies are required to assess whether this can be further generalized to other signaling responses during differentiation or development. In summary, we propose that cell-fate transitions can occur in the absence of CTCF, while the effects of CTCF on genome topology are highly relevant for an acute transcriptional response. The observation that CTCF and global TAD organization are not strictly required for cell-fate changes raises the possibility that lineage-instructive transcription factors themselves shape multilevel topological genome dynamics relevant for major transcriptional rewiring.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41588-020-0643-0.

Received: 12 November 2019; Accepted: 8 May 2020;

References

1. de Laat, W. & Duboule, D. Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature* **502**, 499–506 (2013).
2. Gorkin, D. U., Leung, D. & Ren, B. The 3D genome in transcriptional regulation and pluripotency. *Cell Stem Cell* **14**, 762–775 (2014).
3. Dekker, J. & Mirny, L. The 3D genome as moderator of chromosomal communication. *Cell* **164**, 1110–1121 (2016).
4. Spielmann, M., Lupiáñez, D. G. & Mundlos, S. Structural variation in the 3D genome. *Nat. Rev. Genet.* **19**, 453–467 (2018).
5. Furlong, E. E. M. & Levine, M. Developmental enhancers and chromosome topology. *Science* **361**, 1341–1345 (2018).
6. Stadhouders, R., Filion, G. J. & Graf, T. Transcription factors and 3D genome conformation in cell-fate decisions. *Nature* **569**, 345–354 (2019).
7. Kim, S. & Shendure, J. Mechanisms of interplay between transcription factors and the 3D genome. *Mol. Cell* **76**, 306–319 (2019).
8. Dixon, J. R. et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
9. Hou, C., Li, L., Qin, Z. S. & Corces, V. G. Gene density, transcription, and insulators contribute to the partition of the *Drosophila* genome into physical domains. *Mol. Cell* **48**, 471–484 (2012).
10. Nora, E. P. et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381–385 (2012).
11. Sexton, T. et al. Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* **148**, 458–472 (2012).
12. Rao, S. S. et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665–1680 (2014).
13. Rao, S. S. P. et al. Cohesin loss eliminates all loop domains. *Cell* **171**, 305–320.e24 (2017).
14. Nora, E. P. et al. Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell* **169**, 930–944.e22 (2017).
15. Schwarzer, W. et al. Two independent modes of chromatin organization revealed by cohesin removal. *Nature* **551**, 51–56 (2017).
16. Haarhuis, J. H. I. et al. The cohesin release factor WAPL restricts chromatin loop extension. *Cell* **169**, 693–707.e14 (2017).
17. Heath, H. et al. CTCF regulates cell cycle progression of αβ T cells in the thymus. *EMBO J.* **27**, 2839–2850 (2008).
18. Arzate-Mejía, R. G., Recillas-Targa, F. & Corces, V. G. Developing in 3D: the role of CTCF in cell differentiation. *Development* **145**, dev137729 (2018).
19. Lieberman-Aiden, E. et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293 (2009).

20. Lupiáñez, D. G. et al. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **161**, 1012–1025 (2015).
21. Guo, Y. et al. CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function. *Cell* **162**, 900–910 (2015).
22. Sanborn, A. L. et al. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl Acad. Sci. USA* **112**, E6456–E6465 (2015).
23. Narendra, V. et al. CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *Science* **347**, 1017–1021 (2015).
24. Beagan, J. A. & Phillips-Cremins, J. E. On the existence and functionality of topologically associating domains. *Nat. Genet.* **52**, 8–16 (2020).
25. Rapino, F. et al. C/EBPα induces highly efficient macrophage transdifferentiation of B lymphoma and leukemia cell lines and impairs their tumorigenicity. *Cell Rep.* **3**, 1153–1163 (2013).
26. Crane, E. et al. Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature* **523**, 240–244 (2015).
27. Stadhouders, R. et al. Transcription factors orchestrate dynamic interplay between genome topology and gene regulation during cell reprogramming. *Nat. Genet.* **50**, 238–249 (2018).
28. Natsume, T., Kiyomitsu, T., Saga, Y. & Kanemaki, M. T. Rapid protein depletion in human cells by auxin-inducible degron tagging with short homology donors. *Cell Rep.* **15**, 210–218 (2016).
29. Ouboussad, L., Kreuz, S. & Lefevre, P. F. CTCF depletion alters chromatin structure and transcription of myeloid-specific factors. *J. Mol. Cell Biol.* **5**, 308–322 (2013).
30. Nikolic, T. et al. The DNA-binding factor Ctcf critically controls gene expression in macrophages. *Cell. Mol. Immunol.* **11**, 58–70 (2014).
31. Cuartero, S. et al. Control of inducible gene expression links cohesin to hematopoietic progenitor self-renewal and differentiation. *Nat. Immunol.* **19**, 932–941 (2018).
32. Parelho, V. et al. Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* **132**, 422–433 (2008).
33. Wendt, K. S. et al. Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* **451**, 796–801 (2008).
34. Mumbach, M. R. et al. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat. Methods* **13**, 919–922 (2016).
35. Faridi, M. H. et al. CD11b activation suppresses TLR-dependent inflammation and autoimmunity in systemic lupus erythematosus. *J. Clin. Invest.* **127**, 1271–1283 (2017).
36. Bell, A. C., West, A. G. & Felsenfeld, G. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* **98**, 387–396 (1999).
37. Serra, F. et al. Automatic analysis and 3D-modelling of Hi-C data using TADbit reveals structural features of the fly chromatin colors. *PLoS Comput. Biol.* **13**, e1005665 (2017).
38. Despang, A. et al. Functional dissection of the *Sox9*–*Kcnj2* locus identifies nonessential and instructive roles of TAD architecture. *Nat. Genet.* **51**, 1263–1271 (2019).
39. Ghavi-Helm, Y. et al. Highly rearranged chromosomes reveal uncoupling between genome topology and gene expression. *Nat. Genet.* **51**, 1272–1282 (2019).
40. Williamson, I. et al. Developmentally regulated *Shh* expression is robust to TAD perturbations. *Development* **146**, dev179523 (2019).
41. Le Dily, F. L. et al. Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. *Genes Dev.* **28**, 2151–2162 (2014).

Methods

Cell culture. The BLaER cell line[25] is derived from the RCH-ACV lymphoblastic leukemia cell line in which CEBPA fused with the estrogen receptor (ER) hormone-binding domain and the GFP marker are expressed. BLaER cells and subclones were cultured in RPMI medium (Gibco, catalog no. 22400089) supplemented with 10% fetal bovine serum (Gibco, catalog no. 10100147), 1% glutamine (Gibco, catalog no. 25030081), 1% penicillin/streptomycin antibiotic (Thermo Fisher Scientific, catalog no. 15140122), 550 μM β-mercaptoethanol (Gibco, catalog no. 31350010). Cells were maintained at a density of $0.1–6\times10^6$ cells per ml. Cells were checked for mycoplasma infection every month and tested negative. To induce transdifferentiation, BLaER cells were seeded at 0.3 million cells per ml in a culture medium supplemented with 100 nM β-estradiol, IL-3 and CSF-1 ($100\,\text{ng ml}^{-1}$). iMacs were collected after 7 d of incubation. For auxin-inducible degradation, indole-3-acetic acid (IAA, a chemical analog of auxin) was added to the medium at 500 μM from a 1,000× stock diluted in DMSO. Stocks were kept at 4 °C up to 4 weeks or −20 °C for long-term storage. For endotoxin stimulation, cells were treated with LPS ($1\,\mu\text{g ml}^{-1}$) for 2 h to collect RNA or 8 h to collect supernatant.

Plasmid construction. The CTCF-mAID-mCherry targeting vectors were cloned by serial modification of the base vectors pMK292 (Addgene, catalog no. 72830) and pMK293 (Addgene, catalog no. 72831). Homology arms (HA) of the last exon of *CTCF* were synthesized (IDT). The TIR1-AAVS1 donor vector pMK232 (Addgene, catalog no. 105924) and the pX330 vector expressing the single guide RNA (sgRNA) to target the *AAVS1* locus in human cells (Addgene, catalog no. 72833) were kindly provided by M. Kanemaki[28]. *CTCF*-targeting sgRNAs were cloned in pX330 by annealing oligonucleotides caccgTGATCCTCAGCATGATGGAC and aaacGTCCATCATGCT GAGGATCAc.

Gene targeting. For transfection, plasmids were prepared using a Plasmid Midi Kit (Qiagen), followed by ethanol precipitation. Constructs were not linearized. The BLaER cell line was used to generate the parental line expressing the OsTIR1 enzyme. Transfection was carried out by electroporation (Amaxa Nucleofector, Lonza), using Kit C and program X-001, according to the manufacturer's instructions. One microgram of each plasmid (pMK232 and pX330-AAVS-sgRNA) was added per 100 μl of solution mix and 1 million cells. Eight sets of 1 million cells were transfected using the same conditions, and the day after the transfection dead cells were eliminated by centrifugation and alive cells were pooled together. Three days after transfection, puromycin ($1\,\mu\text{g ml}^{-1}$, Gibco, catalog no. A1113803) was added to the medium to select edited cells. Selection medium was changed every 2–3 d and the selection was performed for 10 d. Single-cell sorting of resistant cells was performed and AAVS PCR genotyping allowed the selection of homozygous insertion of the TIR1 expression cassette at the *AAVS* locus. Several clones were selected and tested for TIR1 expression by quantitative PCR (qPCR) allowing the selection of the clone with the most robust expression (cell line no. 2B10). This clone was used for the targeting of *CTCF*. Two runs of gene targeting were performed (the first using a neomycin-targeting plasmid and the second with a hygromycin-targeting plasmid) to obtain homozygous recombined alleles. One microgram of each plasmid (px330-mCherry-sgRNA; pHA-mAID-mCherry-Neo R or pHA-mAID-mCherry-Hygro R) was added per 100 μl of solution mix and 1 million cells. Eight sets of 1 million cells were transfected using the same conditions, and the day after the transfection, dead cells were eliminated by centrifugation and alive cells were pooled together. Three days after transfection, antibiotic was added to the medium to select edited cells ($500\,\mu\text{g ml}^{-1}$ G418, Life Technologies, catalog no. 11811031 and/or $100\,\mu\text{g ml}^{-1}$ of hygromycin B, Gibco, catalog no. 10687010). Selection medium was changed each 2–3 d and the selection was performed for 17–20 d. Single-cell sorting of resistant cells expressing mCherry was performed and a genotyping PCR allowed the selection of homozygous mAID-CTCF-targeted cells.

Flow cytometry. BLaER cells and derived clones were resuspended in culture medium, spun down and resuspended in 4% FBS–PBS, and live cells (DAPI$^-$) were sorted by live flow cytometry on a BD Influx instrument (BD Bioscience). For monitoring transdifferentiation, cells were subjected to a specific cell surface marker staining. Briefly, blocking was carried out for 10 min at room temperature using human FcR binding inhibitor (1:20 dilution, eBiosciences, catalog no. 16–9161–73) and cells were then stained with antibodies against CD19 (APC-Cy7 mouse anti-human CD19, BD Pharmingen, catalog no. 557791) and Mac1 (APC mouse anti-human CD11b/Mac1, BD Pharmingen, catalog no. 550019) at 4 °C for 20 min in the dark. After washing, DAPI staining was performed just before analysis. For monitoring phagocytosis, cells were seeded at a density of 0.5 million cells per ml in medium and Fluoresbrite carboxy bright blue beads (1 μm, Polysciences, catalog no. 17458) were added (300 beads per cell) and incubated for 24 h before fluorescence-activated cell sorting (FACS) analysis. Dissociation, wash and flow buffers were supplemented with auxin, when appropriate, to avoid re-expression of the CTCF–mAID–mCherry fusion. All the analyses were performed using an LSR Fortessa instrument (BD Biosciences). Data analysis was performed using FlowJo software.

Western blots and antibody arrays. One million cells were centrifuged, washed with PBS 1× and lysis was performed in 30 μl of Laemmli buffer 1× (50 nM Tris–HCl pH 6.8, 2% glycerol, 2% SDS, 0.01% 2-mercaptoethanol, 0.05% bromophenol blue). After heating at 95 °C for 10 min, the protein extracts (corresponding to 5×10^5 cells) were separated by electrophoresis in a 7.5% polyacrylamide gel (Bio-Rad, catalog no. 4561023) before transfer to nitrocellulose membrane. Membranes were blocked with 5% nonfat milk TBS–Tween medium (50 mM Tris, 150 mM NaCl, 0.1% Tween 20) for 1 h at room temperature. Incubation with primary antibody was performed at 4 °C with shaking overnight (anti-CTCF, Millipore, catalog no. 07–729; anti-α-tubulin, Abcam, catalog no. ab7291; anti-mAID-tag, MBL Life Science, catalog no. M214–3; 1:1,000 in 5% milk in TBS–Tween). Membranes were washed with TBS–Tween (3 × 10 min) before secondary incubation with antibodies fused to horse radish peroxidase (HRP) (goat anti-mouse IgG, Sigma Aldrich, catalog no. A3682, dilution 1:5,000) for 1 h at room temperature. After three final washes, membranes were incubated in ECL Start Western Blotting Detection Reagent mix (Sigma Aldrich, catalog no. GERPN3243) for 2 min at room temperature before development on X-ray film. Cytokine arrays (R&D, catalog no. ARY006) were performed following the manufacturer's instructions using supernatant from iMacs collected 8 h after LPS stimulation ($1\,\mu\text{g ml}^{-1}$). Antibody arrays were imaged using an Odyssey CLx instrument (LI-COR).

Immunofluorescence. iMacs were grown on glass coverslips and fixed with 3% formaldehyde in PBS 1× for 10 min at room temperature. After washing with PBS, imaging was performed using a Leica TCS SPE inverted microscope. Images were postprocessed using Fiji Is Just ImageJ (FIJI).

ChIP–seq. Cells were cross-linked for 10 min using 1% formaldehyde and quenched using a final concentration of 0.125 M glycin. Cell pellets were lysed by incubating for 10 min on ice with 5 mM PIPES pH 8, 85 mM KCl, 0.5% IGEPAL, 1× protease inhibitor (Roche). After centrifugation, pellets were incubated in 1% SDS, 10 mM EDTA pH 8, 50 mM Tris–HCl pH 8.1 and 1× PIC for 10 min on ice. Chromatin was sheared on a Bioruptor pico sonicator (Diagenode) at 4 °C for 14 cycles of 30 s on and 30 s off. After sonication, the solution was left on ice for 1 h to allow SDS precipitation and clarified by centrifugation at 16,000*g* for 10 min at 4 °C. Supernatant was transferred in a new tube, 10% was saved as input and the rest was diluted to 1.2 ml with 1× cold immunoprecipitation (IP) buffer (Diagenode). A 10-μg portion of anti-CTCF (Milipore, catalog no. 07–729) was added followed by overnight incubation at 4 °C on a rotator. Beads (42 μl) (unblocked protein A beads, Diagenode, catalog no. kch-503–008) were used per IP after blocking them using 1% BSA cold IP buffer for 15 min at 4 °C under rotation. Blocked beads were added to the chromatin solution and incubated 3 h at 4 °C with rotation. Beads were then collected by centrifugation for 2 min at 3,000 r.p.m. at 4 °C and washed three times with cold IP buffer and two times with cold TE buffer (10 mM Tris pH 8, 1 mM EDTA). Beads were then eluted with freshly prepared elution buffer (1% SDS, 0.1 M $NaHCO_3$) and incubated for 25 min at room temperature. The supernatant was transferred into a new tube and cross-linking was reversed by adding NaCl (final concentration 200 mM) and incubating overnight at 65 °C. Protein digestion was achieved by adding Tris pH 6.5 (40 mM), EDTA pH 8 (10 mM) and proteinase K ($4\,\mu\text{g}\,\mu\text{l}^{-1}$) and incubating for 1 h at 45 °C. DNA was then purified by phenol:chloroform:isoamyl alcohol (25:24:1) extraction. The entire DNA sample was used to construct Illumina sequencing libraries. Library preparation was performed using the NEBNext DNA Library Prep Kit (New England BioLabs) with 2 μl NEBNext adapter in the ligation step. Libraries were amplified for 14 cycles with Herculase II Fusion DNA Polymerase (Agilent) and were purified/size-selected with Agencourt AMPure XP beads (>200 base pairs (bp)). Libraries were sequenced on Illumina HiSeq2000 or NextSeq500 instruments using the 50- or 75-nucleotide paired-end mode, respectively.

Quantitative PCR with reverse transcription and RNA-seq. RNA was extracted with the miRNeasy mini kit (Qiagen) and quantified with a NanoDrop spectrophotometer. Complementary DNA was produced with a High Capacity RNA-to-cDNA kit (Applied Biosystems) and was used for quantitative PCR with reverse transcription (qRT–PCR) analysis in triplicate reactions with SYBR Green QPCR Master Mix (Applied Biosystems). Oligonucleotide sequences are indicated in Supplementary Table 1. Libraries were prepared with an Illumina TrueSeq Stranded total RNA Library Preparation Kit after Ribo-zero depletion, and single-end sequencing (75 nucleotides) was performed on an Illumina HiSeq2500 instrument.

ATAC-seq. ATAC-seq was performed as previously described[42]. Briefly, 5 million cells were harvested and treated with Nextera Tn5 Transposase (Illumina, catalog no. FC-121–1030) for 45 min at 37 °C. Library fragments were amplified using 1× NEBNext High-Fidelity 2× PCR Master Mix (NEB, catalog no. M0541S) and 1.25 μM of custom Nextera PCR primers. PCR amplification was done with 11 cycles, determined by KAPA Real-Time Library Amplication Kit (Peqlab, catalog no. KK2701) so as to stop before saturation. Then, the samples were purified using MinElute PCR Purification Kit (Qiagen, catalog no. 28004) and with Agencourt

AMPure XP beads (Beckman Coulter, catalog no. A63881) in a 3:1 ratio. The libraries were sequenced as paired-end (50 bp) on a HiSeq2000 instrument.

In situ Hi-C library preparation. In situ Hi-C was performed as previously described[12] with the following modifications: (1) two million cells were used as starting material; (2) chromatin was initially digested with 100 U MboI (New England BioLabs) for 2 h, and then another 100 U (2 h incubation) and a final 100 U were added before overnight incubation; (3) before fill-in with bio-dATP, nuclei were pelleted and resuspended in fresh 1× NEB2 buffer; (4) ligation was performed overnight at 24 °C with 10,000 cohesive end units per reaction; (5) de-cross-linked and purified DNA was sonicated to an average size of 300–400 bp with a Bioruptor Pico (Diagenode; 7 cycles of 20 s on and 60 s off); (6) DNA fragment-size selection was performed only after final library amplification; (7) library preparation was performed with an NEBNext DNA Library Prep Kit (New England BioLabs) with 3 μl NEBNext adapter in the ligation step; (8) libraries were amplified for 8–12 cycles with Herculase II Fusion DNA Polymerase (Agilent) and were purified/size-selected with Agencourt AMPure XP beads (>200 bp). Hi-C library quality was assessed by low-coverage sequencing on an Illumina NextSeq500 instrument, after which every biological replicate ($n = 2$) was sequenced at high coverage on an Illumina HiSeq2500 instrument to obtain ~0.5 billion reads in total per time point per biological replicate.

In situ Hi-C data processing and normalization. Hi-C data were processed using an in-house pipeline on the basis of TADbit[37]. First, quality of the reads was checked using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) to discard problematic samples and detect systematic artifacts. Trimmomatic[43] with the recommended parameters for paired-end reads was used to remove adapter sequences and poor-quality reads (ILLUMINACLIP:TruSeq3-PE.fa:2:30:12:1:true; LEADING:3; TRAILING:3; MAXINFO:targetLength:0.999; and MINLEN:36). For mapping, a fragment-based strategy, as implemented in TADbit, was used, which is similar to previously published protocols[44]. Briefly, each side of the sequenced read was mapped in full length to the reference genome (hg38, GRCh38, December 2017). After this step, if a read was not uniquely mapped, we assumed the read was chimeric due to ligation of several DNA fragments. We next searched for ligation sites and discarded reads in which no ligation site was found. Remaining reads were split as often as ligation sites were found. Individual split read fragments were then mapped independently. These steps were repeated for each read in the input FASTQ files. Multiple fragments from a single uniquely mapped read will result in as many contacts as possible pairs can be made between the fragments. For example, if a single read was mapped through three fragments, a total of three contacts (all-versus-all) was represented in the final contact matrix. We used the TADbit filtering module to remove noninformative contacts and to create contact matrices. The different categories of filtered reads applied are:

- Self-circle: reads coming from a single restriction enzyme (RE) fragment and point to the outside.
- Dangling-end: reads coming from a single RE fragment and point to the inside.
- Error: reads coming from a single RE fragment and point in the same direction
- Extra dangling-end: reads coming from different RE fragments but are close enough and point to the inside. The distance threshold used was left to 500 bp (default), which is between percentile 95 and 99 of average fragment lengths.
- Duplicated: the combination of the start positions and directions of the reads was repeated, pointing at a PCR artifact. This filter only removed extra copies of the original pair.
- Random breaks: start position of one of the reads was too far from the RE cutting site, possibly due to noncanonical enzymatic activity or random physical breaks. Threshold was set to 750 bp (default), greater than percentile 99.9. From the resulting contact matrices, low-quality bins (those presenting low contacts numbers) were removed as implemented in TADbit's 'filter_columns' routine. The matrices obtained were normalized for sequencing depth and genomic biases using OneD[45]. Then, they were further normalized for local coverage within the region (expressed as normalized counts per thousand within the region) without any correction for the diagonal decay. For differential analysis, the resulting normalized matrices were directly subtracted from each other.

Identification of subnuclear compartments and TADs. To segment the genome into A/B compartments, normalized Hi-C matrices at 100-kilobase (kb) resolution were corrected for decay as previously published, grouping diagonals when signal-to-noise was below 0.05 (ref. [12]). Corrected matrices were then split into chromosomal matrices and transformed into correlation matrices using the Pearson product–moment correlation. The first component of a PCA (PC1) on each of these matrices was used as a quantitative measure of compartmentalization, and AT content was used to assign negative and positive PC1 categories to the correct compartments. If necessary, the sign of the PC1 (which is randomly assigned) was inverted so that positive PC1 values corresponded to A-compartment regions, and vice versa for the B compartment. Noticeable differences in PC1 values between conditions were calculated using two-sided Wilcoxon rank-sum tests. Normalized contact matrices at 50-kb resolution were used to define TADs, using a previously described method[26] with default parameters. First, for each bin, an insulation score was obtained on the basis of the number of contacts between bins on each side of a given bin. Differences in insulation score between both sides of the bin were computed and borders were called searching for minima within the insulation score[26]. This procedure resulted in a set of borders for each time point and replicate. Between replicates, overlapping or borders distant of less than 1 bin were merged to obtain a list of conserved borders for each time point. Conserved borders overlapping or distant of less than 1 bin among each time point were considered as stable while the others were considered as dynamic. Noticeable differences in insulation scores between conditions were calculated using two-sided Wilcoxon rank-sum tests.

Inter- and intracompartment strength measurements. We followed a previously reported strategy to measure overall interaction strengths within and between A and B compartments[15]. Briefly, we based our analysis on the 100-kb bins showing the most extreme PC1 values, discretizing them by percentiles and taking the bottom 20% as B compartment and the top 20% as A compartment. We classified each bin in the genome according to PC1 percentiles and gathered contacts between each category, computing the $\log_2$ enrichment over the expected counts by distance decay. Finally, we summarized each type of interaction (A–A, B–B and A–B/B–A) by taking the median values of the $\log_2$ contact enrichment.

Meta-analysis of borders. To study the behavior of TAD borders, all TADs of sizes ranging from 0.5 to 1.5 Mb were selected. Then we defined a flanking region of 1 Mb around the border and gathered the observed and expected (by distance decay) matrix counts. Setting up their relative position to the corresponding border, the matrices were stacked to obtain a meta-contact matrix around TAD borders for each condition. This information was summarized by comparing the average $\log_2$(fold change) of contact enrichment between bins inside and outside the TAD.

Enhancer–promoter intra-TAD contacts analysis. By using Hi-C matrices at 5-kb resolution, we focused on TADs containing enhancers and promoters. Each bin was labeled as part of an enhancer or promoter, or 'others' if it did not belong to those previous types. Then, the observed contacts were gathered between the different types of bins within their TAD and expected contact frequencies were computed on the basis of the genomic distance separating each pair (the expected distance decay was calculated excluding entries outside TADs). Then a linear mixed model including TAD identity as a random effect was used to estimate the quantities of interest. Results are expressed as $\log_2$ of the ratio observed on expected frequencies of contacts.

Long-range interactions between enhancers and promoters. Hi-C matrices were generated at 10-kb resolution using HiCExplorer[46] and long-range interactions (5–10 Mb) between promoters and enhancers that were activated or inactivated were computed using the HiCExplorer tool hicAggregateContacts.

Meta-analysis of cohesin loops. Hi-C matrices in cool format were used to generate genome-wide aggregate plots at SMC1-bound loops detected by HiChIP[34]. We used coolpup.py[47] to pile-up normalized Hi-C signals at a 10-kb resolution at SMC1-bound loops previously identified[34], and plotted 100 kb upstream and downstream of the SMC1 anchor coordinates.

Virtual 4C analysis. Hi-C matrices for virtual 4C profiles were further smoothed using a focal (moving window) average of one bin. The profiles were generated from these normalized matrices and correspond to histogram representation of the lines of the matrices containing the baits (therefore, they are expressed as counts per hundred normalized reads within the region depicted).

Gene expression analysis using RNA-seq data. Reads were mapped using STAR[48] (standard options) and the Ensembl human genome annotation (GRCh38 v.27). Gene expression was quantified using STAR (--quantMode GeneCounts). Batch effects were removed using the ComBat function from the sva R package (v.3.22). Sample scaling and statistical analysis were performed using the R package DESeq2 (ref. [49]) (R v.3.3.2 and Bioconductor v.3.0). Genes with expression changing markedly at any time point were identified using the nbinomLRT test (false discovery rate (FDR) < 0.01) and fold change (FC) > 2 between at least two time points. The $\log_2$(vsd) (variance stabilized DESeq2) counts were used for further analysis unless stated otherwise. To compare expression of various set of genes, the data were mean-centered log-transformed and significant differences were calculated using two-sided Wilcoxon rank-sum tests.

Chromatin accessibility analysis using ATAC-seq data. Reads were mapped to the UCSC human genome build (hg38) using Bowtie2 (ref. [50]) with standard settings. Reads mapping to multiple locations in the genome were removed using SAMtools[51] and PCR duplicates were filtered using Picard (http://broadinstitute.github.io/picard). Bam files were parsed to deepTools[52] for downstream analyses and browser visualization. Peaks in ATAC-seq signal were identified using MACS2 (ref. [53]) (callpeak --nolambda --nomodel -g hs -q 0.01).

ChIP–seq data analysis. Reads were mapped and filtered as described for ATAC-seq. CTCF peaks were identified using MACS2 (ref. [53]) with the 'narrowpeaks' option. CTCF peaks not called in both independent biological replicates were excluded in all subsequent analyses. Coverage of CTCF peaks per TAD border was computed using BEDTools[54]. H3K27ac coverage and CTCF-binding heat maps were performed using deepTools[52].

DNA motif analysis. ATAC-seq peaks specific to the TAD borders were identified using BEDtools[54]. DNA motif analysis of the ATAC-seq peaks were analyzed using HOMER[55] (findMotifs.pl) and the Homer motif results were shown. This uses ZOOPS scoring (zero or one occurrence per sequence) coupled with a hypergeometric test to determine motif enrichment and statistical significance.

Identification of dynamic regulatory regions. Intersecting ATAC-seq peaks with H3K4me1 peaks allowed the identification of 63,665 enhancers, while the overlap with transcription start sites (TSS) revealed 24,932 promoters (Extended Data Fig. 2g). The intensity of the H3K27ac signals at these regions was quantified using the Diffbind R package (v.2.2.12) to define activated and inactivated regions from 0 h to 168 h. Differences were computed with using the DBA_DESEQ2 method and -filter for significance was set at fold change > 2 and FDR < 0.05, as previously described[56]. This analysis allowed profiling of 29,711 dynamic enhancers and 8,439 dynamic promoters, of which about half became activated and the other half inactivated (Extended Data Fig. 2h), which is also reflected by the expression of the associated genes (Extended Data Fig. 2i).

3D modeling and analysis. The processed Hi-C datasets were binned at 10-kb resolution and then normalized using OneD[45]. Then, we defined the regions to be modeled given the genomic context around the enhancer and promoters of interest using following the steps: (1) select key elements contained in the region (that is, enhancers and promoters); (2) retrieve the top 5% interactors of each of these elements; (3) build a network joining the key elements with their retrieved top 5% interactors, and the top 5% interactors among them in the cases where this interaction (interactor with interactor) was present in the top 5% of at least one of them. We added the edge twice if it was in the top 5% of both members; (4) group the networks allowing a genomic distance gap of 50 kb and filtered out the groups in which the ratio (number of edges)/(number of nodes) was smaller than 5; and (5) for each of the regions, ensure that the modeled region contained most of the nodes (genomic coordinate from one bin start until end) appearing in the groups that passed the previous filter. Once regions were selected, normalized interaction matrices were modeled as previously described[57] using TADdyn[58], a molecular dynamic-based protocol implemented in the TADbit software[37]. Similarly to TADbit, TADdyn generates models using a restraint-based approach, in which experimental frequencies of interaction are transformed into a set of spatial restraints[59]. A total of 1,000 models were generated for each genomic region and cell type. Contact maps generated from the ensemble models highly correlated with the input Hi-C normalized interaction matrices. Each ensemble of models was next clustered on the basis of structural similarity as implemented in TADbit[37]. The absence of major structural differences between clusters prompted us to use all of them in further analysis. Next, TADbit was used to measure the following features of the models: (1) distance between particles containing genomic regions of interest in the model ensemble; (2) distances distribution between selected pairs or particles; and (3) notable differential distance distributions assessed by the two-sample Kolmogorov–Smirnov statistic. Finally, model images were generated with Chimera[60].

Statistics and reproducibility. RNA-seq and in situ Hi-C data throughout the paper were generated by analysis of two biologically independent samples from two transdifferentiation experiments. Representative data are shown only if results were similar for both biologically independent replicates. All box plots depict the first and third quartiles as the lower and upper bounds of the box, with a band inside the box showing the median value and whiskers representing 1.5× the interquartile range. Wilcoxon rank-sum tests were performed with the wilcox.test() function in R in a two-sided manner. Student's *t*-tests were performed with the t.test() function in R in an unpaired and two-sided fashion with $(n-2)$ degrees of freedom. Kolmogorov–Smirnov tests were performed in a two-sided manner using the module scipy.stats.ks_2sam in the SciPy software.

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

Data availability

The Hi-C, RNA-seq, CTCF ChIP–seq, ATAC-seq datasets generated and analyzed for the current study are available in the Gene Expression Omnibus (GEO) database under accession number GSE140528. ATAC-seq and CEBPA ChIP–seq datasets used in the current study are available in the GEO database under accession number GSE131620. The H3K27ac and H3K4me1 ChIP–seq datasets used in this study are available in the ArrayExpress database under accession number E-MTAB-9010. Source data for Fig. 2c are provided with the paper.

References

42. 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).
43. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
44. Ay, F. et al. Identifying multi-locus chromatin contacts in human cells using tethered multiple 3C. *BMC Genomics* **16**, 121 (2015).
45. Vidal, E. et al. OneD: increasing reproducibility of Hi-C samples with abnormal karyotypes. *Nucleic Acids Res.* **46**, e49 (2018).
46. Ramírez, F. High-resolution TADs reveal DNA sequences underlying genome organization in flies. *Nat. Commun.* **9**, 189 (2018).
47. Flyamer, I. M., Illingworth, R. S. & Bickmore, W. A. Coolpup.py: versatile pile-up analysis of Hi-C data. *Bioinformatics* **36**, 2980–2985 (2020).
48. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
49. 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).
50. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
51. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
52. Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A. & Manke, T. DeepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42**, W187–W191 (2014).
53. Feng, J., Liu, T., Qin, B., Zhang, Y. & Liu, X. S. Identifying ChIP-seq enrichment using MACS. *Nat. Protoc.* **7**, 1728–1740 (2012).
54. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
55. 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).
56. Ross-Innes, C. S. et al. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481**, 389–393 (2012).
57. Miguel-Escalada, I. et al. Human pancreatic islet three-dimensional chromatin architecture provides insights into the genetics of type 2 diabetes. *Nat. Genet.* **51**, 1137–1148 (2019).
58. Stefano, M. Di et al. Dynamic simulations of transcriptional control during cell reprogramming reveal spatial chromatin caging. Preprint at https://doi.org/10.1101/642009 (2019).
59. Baù, D. & Marti-Renom, M. A. Genome structure determination via 3C-based data integration by the Integrative Modeling Platform. *Methods* **58**, 300–306 (2012).
60. Pettersen, E. F. et al. Chimera - a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

Acknowledgements

We thank M. T. Kanemaki for the degron plasmids; R. Guigó's laboratory, and S. Pérez-Luch in particular, for the H3K27ac and H3K4me1 ChIP–seq, produced in the framework of the RNA-MAPS project (ERC-2011-AdG-294653-RNA-MAPS); Y. Cuartero for help with sequencing and CTCF ChIP–seq; C. Segura for help with immunofluorescence microscopy; the CRG Genomics and flow cytometry core facilities and the CRG-CNAG Sequencing Unit for sequencing; and members of T.G.'s laboratory for discussions. This work was supported by the European Research Council under the 7th Framework Programme FP7/2007-2013 (ERC Synergy Grant 4D-Genome, grant agreement 609989, to T.G. and M.A.M.-R.), the Ministerio de Educación y Ciencia (SAF.2012-37167, to T.G., and BFU2017-85926-P, to M.A.M.-R.), the AGAUR (to T.G.) and the Marató TV3 (201611) (to M.A.M.-R.). P.C. was supported by the Deutsche Forschungsgemeinschaft (SFB860, SPP1935, EXC 2067/1-390729940), the European Research Council (advanced investigator grant TRANSREGULON, grant agreement no. 693023) and the Volkswagen Foundation. G.S. was supported by a Marie Skłodowska-Curie fellowship (H2020-MSCA-IF-2016, miRStem) and by the 'Fundación Científica de la Asociación Española Contra el Cáncer'. T.V.T. was supported by Juan de la Cierva postdoctoral fellowship (MINECO; FJCI-2014-22946). B.B. was supported by the fellowship 2017FI_B00722 from the Secretaria d'Universitats i Recerca del Departament d'Empresa i Coneixement (Generalitat de Catalunya) and the European Social Fund (ESF). R.S. was supported by a Netherlands Organisation for Scientific Research Veni fellowship (91617114) and an Erasmus MC Fellowship. We also acknowledge support from 'Centro de Excelencia Severo Ochoa 2013-2017' (SEV-2012-0208), the Spanish Ministry of Science and Innovation to the EMBL partnership and the CERCA Program Generalitat de Catalunya to the CRG, as well as the support of the Spanish Ministry of Science and Innovation through the Instituto de Salud Carlos III, the Generalitat de Catalunya through Departament de Salut and Departament d'Empresa i Coneixement, and co-financing by the Spanish Ministry of Science and Innovation with funds from the European Regional Development Fund (ERDF) corresponding to the 2014-2020 Smart Growth Operating Program to CNAG.

Author contributions

G.S., R.S. and T.G. conceived the study and wrote the manuscript with input from all coauthors. G.S. performed molecular biology, RNA-seq, ChIP–seq and in situ Hi-C experiments. T.V.T., J.C. and A.A. performed ChIP–seq and ATAC-seq. G.S., E.V., M.V.-C., J.M.-E. and B.B. performed bioinformatic analyses. G.S., E.V., M.V.-C., J.M.-E. and R.S. integrated and visualized data. G.S. and M.B. performed the CTCF-degron CRISPR targeting and S.C. performed cytokine arrays. G.S. performed the transdifferentiation experiments with help from M.B. and C.B. F.L.D., P.C., M.A.M.-R. and R.S. provided valuable advice and T.G. supervised the research.

Competing interests

The authors declare no competing interests.

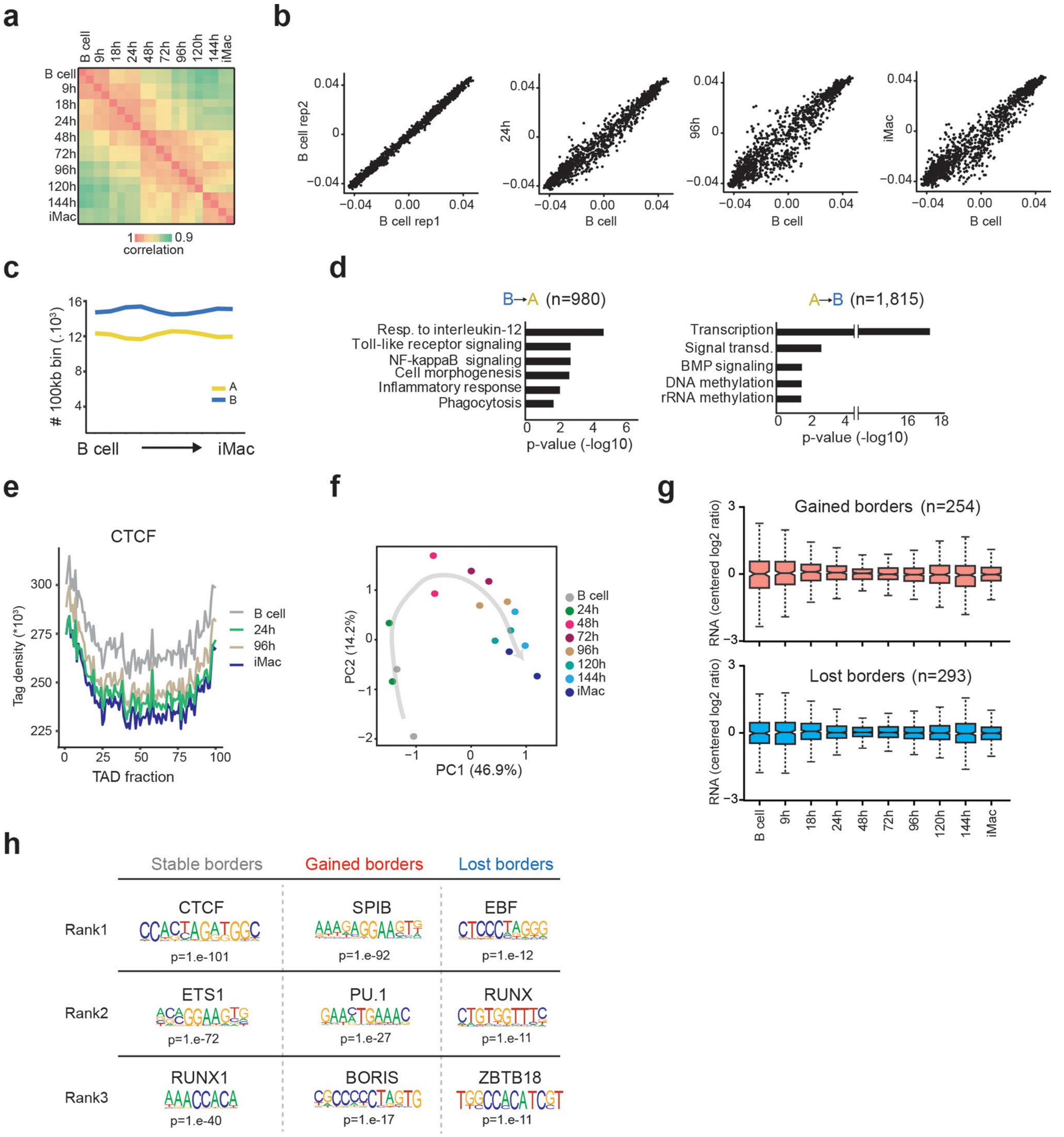
Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41588-020-0643-0.

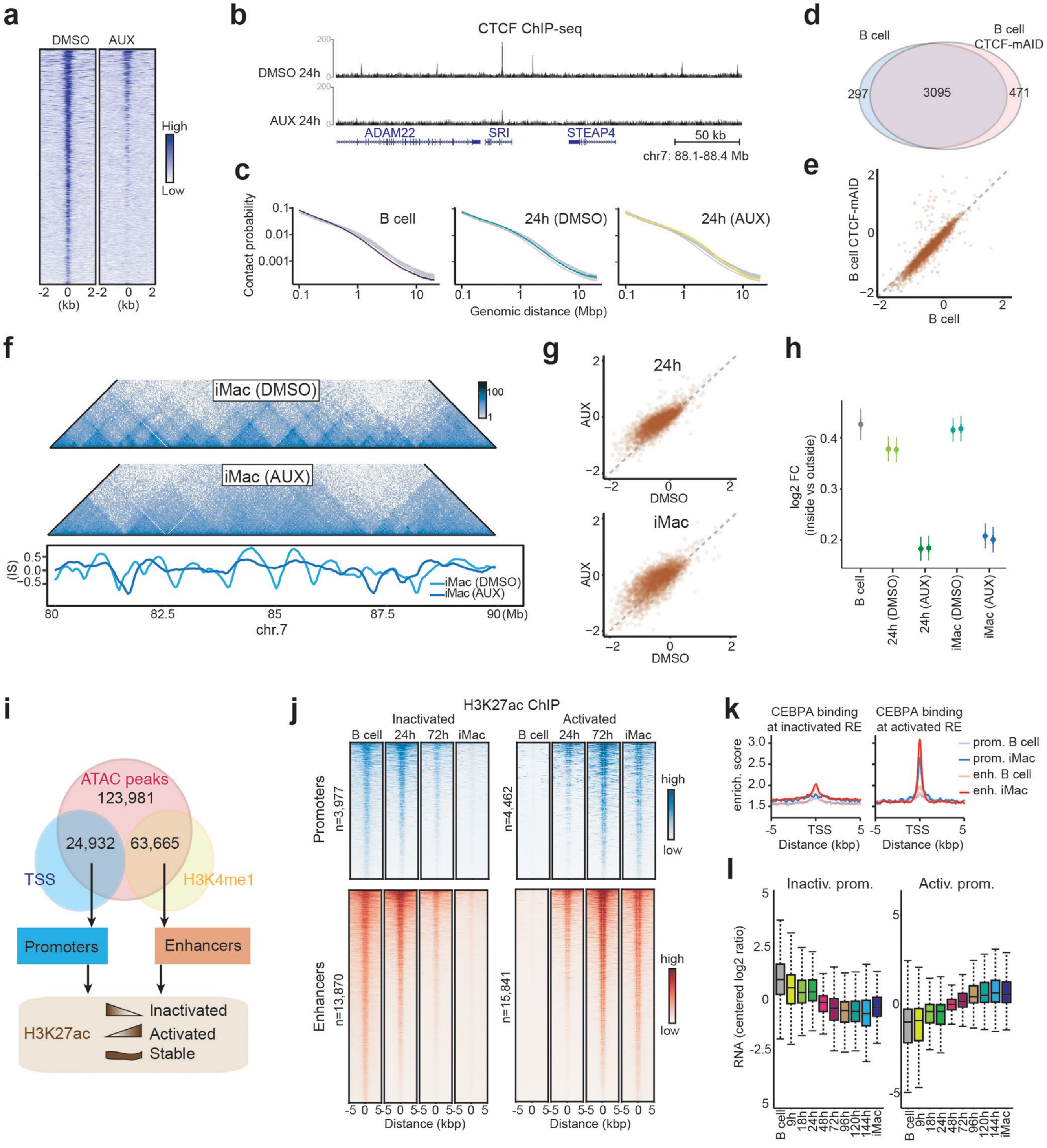
Supplementary information is available for this paper at https://doi.org/10.1038/s41588-020-0643-0.

Correspondence and requests for materials should be addressed to G.S., R.S. or T.G.

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

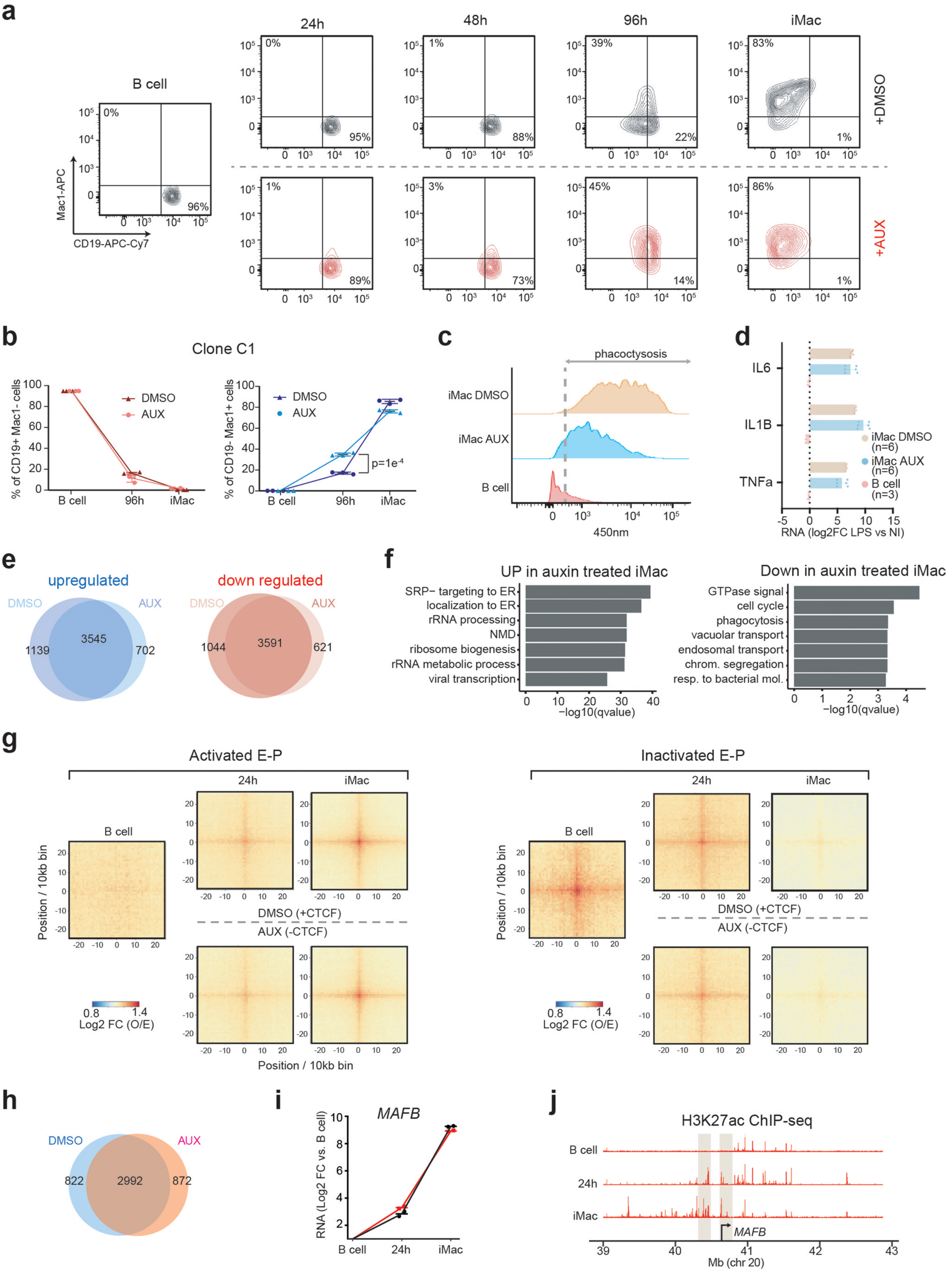


Extended Data Fig. 1 | Characterization of chromatin compartmentalization and TAD dynamics during transdifferentiation. **a**, Genome-wide Pearson correlation matrix between PC1 values of Hi-C samples at different time points. **b**, Scatter plots of PC1 values (n = 1,332 100-kb bins) showing changes relative to initial B cell genome compartmentalization for chromosome 12. **c**, Line chart depicting fractions of the genome assigned to A or B compartments at 10 time points during transdifferentiation. Y-axis represents the number of 100-kb bins. **d**, Gene ontology analysis of genes in regions switching from B to A (n = 980 genes) or A to B (n = 1,815 genes) compartments (*P* values, FDR corrected Fisher test). **e**, CTCF binding signal at TADs, normalized for TAD size in samples at various transdifferentiation time points. **f**, PCA of insulation score values at TAD borders during transdifferentiation (n = 4,006 TAD borders). Grey arrow depicts an averaged trajectory. **g**, RNA expression of genes at TAD borders gained (n = 254 genes) or lost (n = 293 genes) during transdifferentiation. All box plots depict the first and third quartiles as the lower and upper bounds of the box, with a thicker band inside the box showing the median value and whiskers representing 1.5x the interquartile range. **h**, Homer DNA motif analysis at ATAC-seq peaks detected at stable (n = 2,044), gained (n = 591) or lost (n = 135) TAD borders (*P* values are calculated using hyper-geometric statistical tests).



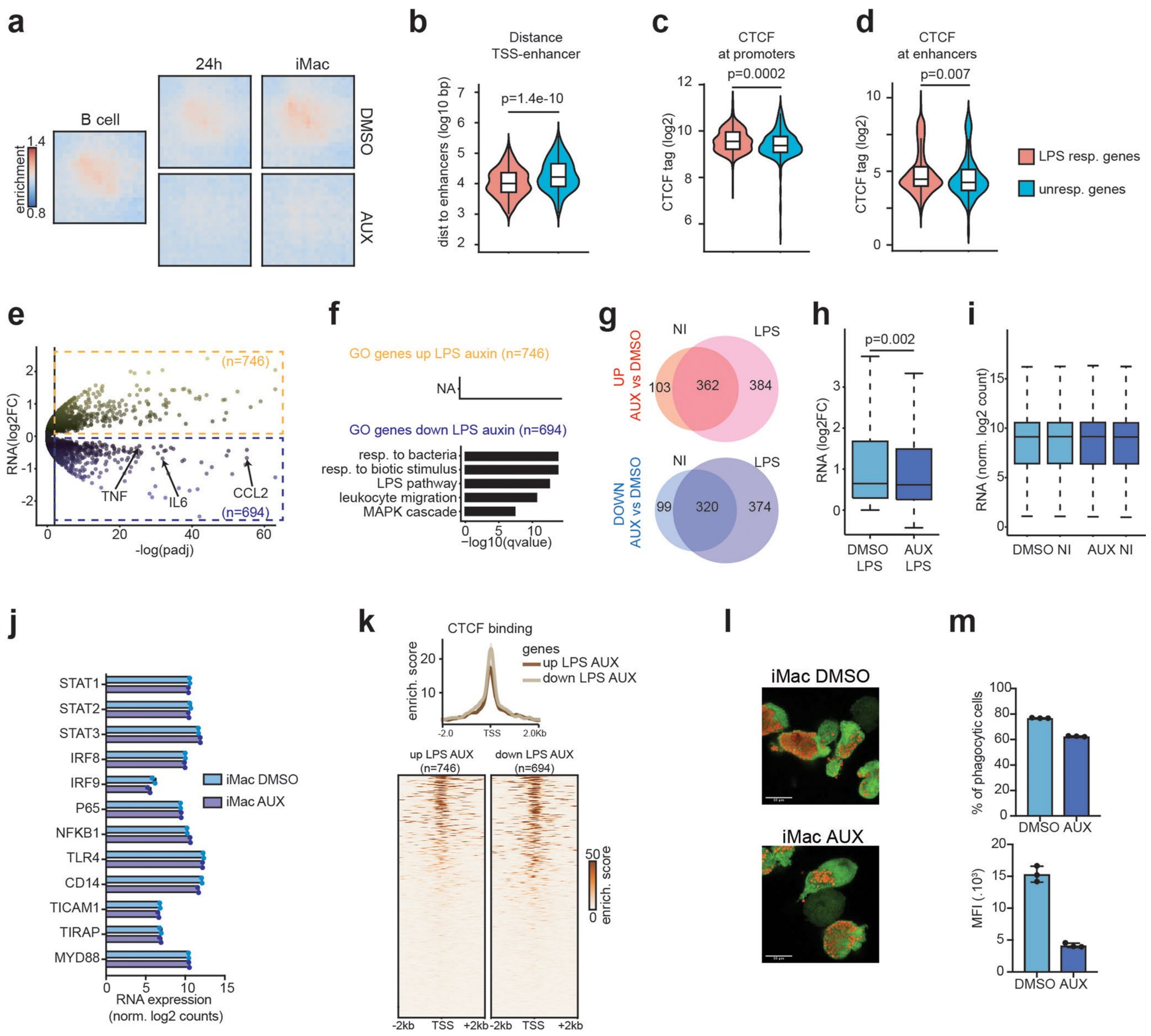
Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Molecular characterization of CTCF-mAID BLaER cells during transdifferentiation. **a**, Heatmap of CTCF binding signal at CTCF ChIP-seq peaks detected in untreated CTCF-mAID cells. **b**, Browser snapshot showing CTCF binding loss upon 24 h of auxin treatment. **c**, Overall scaling of Hi-C contact frequency as a function of genomic distance in cells treated with DMSO or auxin. **d**, Venn diagram showing the overlap of TAD borders detected in B cells and in CTCF-AID B cells. **e**, Scatter plots comparing insulation scores at TAD borders at B cell and at CTCF-AID B cells. Lower values indicate stronger insulation. **f**, Top: Representative in situ Hi-C contact maps (20-kb resolution) of iMacs obtained after treatment with DMSO or auxin. Color scale represents the normalized number of contacts. Bottom: plots of the corresponding insulation scores for each bin within the 10-Mb region shown. **g**, Scatter plots comparing insulation scores at TAD borders at 24 hpi and at the iMac stage after DMSO or auxin treatment. Lower values indicate stronger insulation. **h**, Contact enrichment of interactions inside TADs versus outside TADs at the indicated time points for B cell (n = 1), DMSO- (n = 2) or auxin-treated cell (n = 2) biologically independent samples. Dots represent point estimates and bars (wide and narrow) indicate confidence intervals (50% and 95 %, respectively) for the $\log_2$ fold changes. All estimations are computed using all 9 samples in a single linear mixed model. **i**, Outline of strategy used to identify dynamic promoters and enhancers during transdifferentiation. Numbers of ATAC-seq peaks intersecting with TSS (promoters) and H3K4me1 peaks (enhancers) are indicated. **j**, H3K27ac decoration at dynamic promoters and enhancers that become either inactivated or activated. **k**, CEBPA binding at activated and inactivated regulatory elements (RE) in B cell and iMacs. **l**, RNA expression of genes associated with inactivated (n = 1,259) and activated (n = 1,421) promoters during transdifferentiation. All box plots depict the first and third quartiles as the lower and upper bounds of the box, with a thicker band inside the box showing the median value and whiskers representing 1.5x the interquartile range.

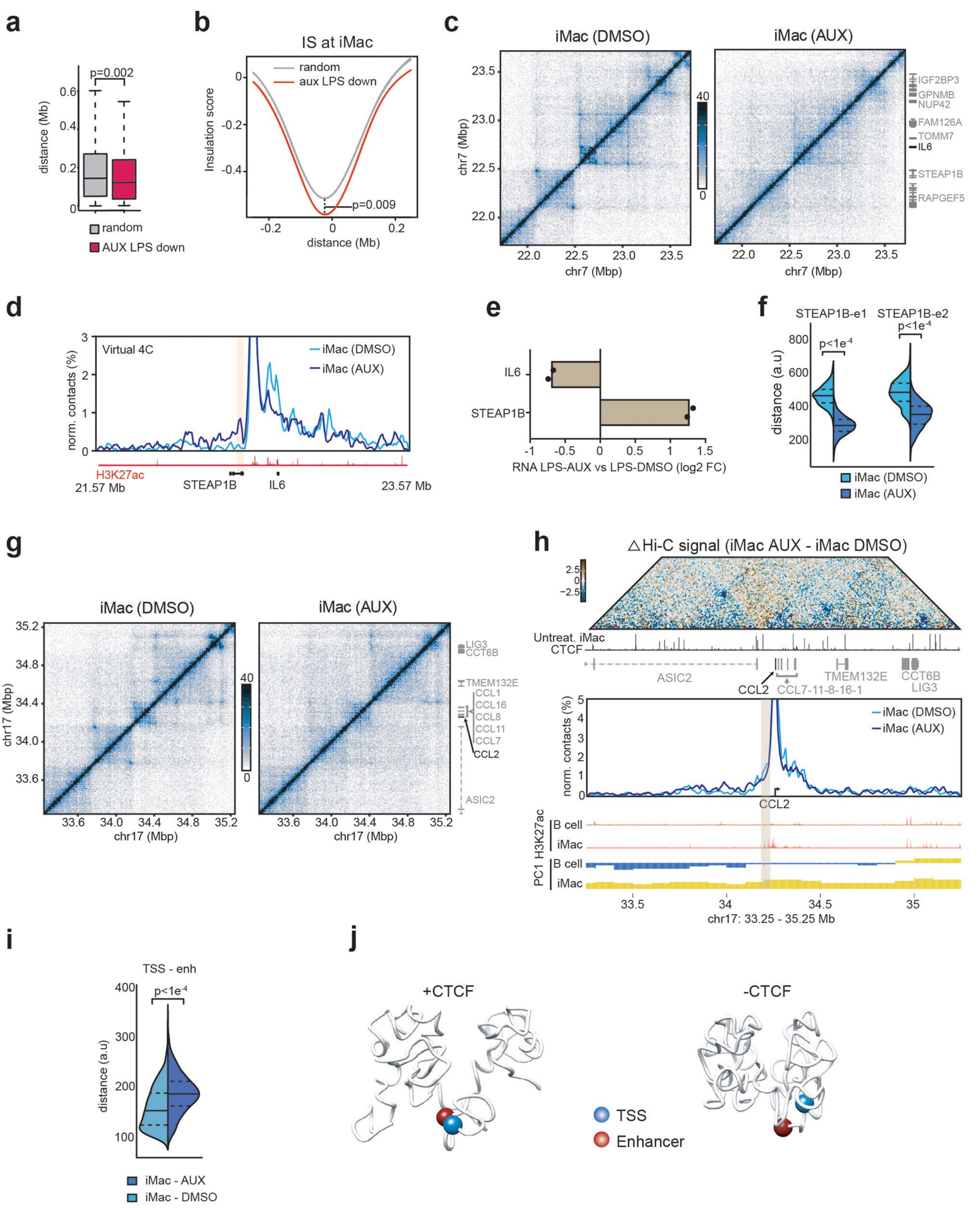


Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | CTCF depletion does not impair transdifferentiation or long-range enhancer-promoter contact dynamics. a, Representative flow cytometry analysis of CD19 and Mac-1 marker expression during transdifferentiation of CTCF-mAID B cells treated with DMSO or auxin. The experiment was repeated 3 times with similar results. **b**, Transdifferentiation kinetics of CTCF-mAID B cells (clone C1) in the presence of DMSO or auxin analysed at 0, 96 and 168 hpi by flow cytometry for CD19 and Mac-1 expression (n = 3 biologically independent samples). Centre indicates mean, error bars show standard deviation and *P* unpaired two-tailed t-test. **c**, Phagocytosis assay of iMacs analyzed by flow cytometry showing uptake of blue fluorescent beads. The experiment was repeated 3 times with similar results. **d**, RNA expression measured by qRT-PCR of cytokines in noninduced (NI) or 2h LPS-induced iMacs DMSO (n = 6), iMacs AUX (n = 6) or B cells (n = 3). Mean values are shown, error bars represent standard error and *n* represents biologically independent samples. **e**, Venn diagram showing the overlap of genes upregulated (left) and downregulated (right) in iMacs after transdifferentiation in the presence of DMSO or auxin based on RNA-Seq (n = 2 biologically independent samples). **f**, Gene ontology analysis of genes specifically upregulated (n = 419) and downregulated (n = 744) specifically in CTCF-depleted iMacs (q-value, FDR corrected Fisher exact test). **g**, Aggregate metaplots (10-kb resolution) depicting long range (5–10-Mb) interaction frequencies between enhancers and promoter (E-P) during transdifferentiation. Area shown is centered on enhancers or promoters ± 250-kb). **h**, Venn diagram showing the number of switching compartment regions (100-kb bins) during transdifferentiation in presence of DMSO or auxin. **i**, Expression of *MAFB* during transdifferentiation with or without CTCF, as measured by RNA-seq (n = 2 biologically independent samples, lines connect mean values). **j**, Enhancer activity at the *MAFB* locus during transdifferentiation. Browser snapshot showing H3K27ac ChIP-seq profiles of a 4-Mb domain surrounding the *MAFB* locus. The enhancer and the promoter shown in Fig. 3g are highlighted in light brown.



Extended Data Fig. 4 | CTCF depletion in iMacs attenuates the acute inflammatory response to endotoxins. a, Genome-wide aggregation of normalized Hi-C signal anchored at cohesin loops during transdifferentiation with DMSO or auxin. **b**, Distance distribution between enhancers and TSS of genes responsive (n = 378) or unresponsive (n = 380) to LPS (*P*, two-sided Wilcoxon rank-sum test). **c**, CTCF enrichment at promoters and enhancers **d**, of genes responsive (n = 378) or unresponsive (n = 378) to LPS (*P*, two-sided Wilcoxon rank-sum test). **e**, Differential gene expression between LPS-induced iMac treated with auxin or DMSO (n = 2 biologically independent samples, *P*-adj two-tailed likelihood ratio test followed by FDR correction). **f**, Gene ontology analysis of the significantly ($p < 0.01$) upregulated (n = 746) and downregulated (n = 694) genes in LPS-induced iMacs treated with auxin compared to DMSO (q-value, FDR-corrected Fisher exact test). **g**, Overlap of upregulated (top) and downregulated (bottom) genes (AUX vs DMSO) between non-induced iMacs (NI) and iMacs treated with LPS (LPS). **h**, LPS-upregulated genes in iMacs exposed to DMSO compared to auxin (n = 2,470 genes, *P* two-sided Wilcoxon rank-sum test). **i**, RNA expression in non-induced (NI) iMacs of genes upregulated after LPS stimulation of DMSO treated iMacs (n = 2,470). **j**, RNA expression of key transcription factors and receptors of the LPS signalling pathway (n = 2 biologically independent samples). **k**, CTCF binding at promoters of genes deregulated in LPS-induced iMacs treated with auxin as compared to DMSO. **l**, Micrographs show uptake of fluorescent beads (shown in red) by iMacs treated with DMSO or auxin (Scale bar represents 10 µm). The experiment was repeated 3 times with similar results. **m**, Quantification of phagocytosis assay. Upper panel shows percentage of cells with bead uptake; lower panel shows mean fluorescent intensity (MFI). Bars represent mean values of n = 3 biologically independent samples and error bars denote standard deviation. All box plots depict the first and third quartiles as the lower and upper bounds of the box, with a thicker band inside the box showing the median value and whiskers representing 1.5x the interquartile range.



Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | CTCF depletion in iMacs impairs 3D chromatin organization at inflammatory response gene loci. a, Distance distribution between promoters and their closest TAD borders of genes downregulated in auxin treated iMacs after LPS induction (as compared to iMacs exposed to DMSO) or for a random set of genes with a similar size (n = 687) (*P*, two-sided Wilcoxon rank-sum test). Box plots depict the first and third quartiles as the lower and upper bounds of the box, with a thicker band inside the box showing the median value and whiskers representing 1.5x the interquartile range. **b**, Average insulation scores of TAD borders closest to genes downregulated in auxin treated iMacs after LPS induction or closest to a random set of genes with a similar size (n = 687). Area shown is centered on boundary regions ± 250-kb (*P*, two-sided Wilcoxon rank-sum test). **c**, Hi-C maps (10-kb resolution) at the *IL6* locus. Color scale represents the normalized number of contacts and the genes within the locus are indicated on the right. **d**, Virtual 4 C of iMacs treated with DMSO (dark blue) or auxin (light blue) using *IL6* enhancer 1 (e1) as viewpoint; Browser snapshot of H3K27ac ChIP-seq signal is shown and the *STEAP1B* promoter is highlighted. **e**, Differential expression of *IL6* and *STEAP1B* in LPS-induced iMacs treated with auxin as compared to DMSO (bars represent the mean values of n = 2 biologically independent samples). **f**, Distance distribution between STEAP1B promoter and IL6 enhancer regions (n = 1,000 models). Median (solid line), first and third quartile (dashed line) are indicated (*P*, two-sided Komogorov-Smirnov test). **g**, Hi-C maps (10-kb resolution) at the *CCL2* locus in iMacs generated in the presence of DMSO or auxin. Color scale represents the normalized number of contacts. Genes within the locus are indicated on the right. **h**, Top: Differential Hi-C maps of the *CCL2* locus (10-kb resolution) in iMacs generated in the presence of DMSO or auxin; CTCF ChIP-seq signal and gene positions are shown below the Hi-C map. Middle: Virtual 4 C of the *CCL2* locus of iMacs treated with DMSO (dark blue) or auxin (light blue), using *CCL2* promoter as viewpoint. Bottom: browser snapshots showing H3K27ac ChIP-seq and PC1 A/B compartment tracks. The *CCL2* enhancer is highlighted. **i**, Distance distribution between CCL2 TSS and enhancer regions (n = 1,000 models). Median (solid line), first and third quartiles (dashed line) are indicated (*P*, two-sided Komogorov-Smirnov test). **j**, 3D model of the *CCL2* locus in DMSO or auxin treated iMacs.

nature research

Corresponding author(s): Gregoire Stik
Ralph Stadhouders
Thomas Graf

Last updated by author(s): May 5, 2020

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.

Statistics

For all statistical analyses, confirm that the following items are present in the 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
☐	☒	A statement on 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 statistical parameters 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

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No softwares were used to collect the data
Data analysis	STAR (Dobin et al., 2013), DESeq2 (Love et al., 2014), Bowtie2 (Langmead and Salzberg, 2012), SAMtools (Li et al., 2009), HOMER (Heinz et al., 2010), TADbit (Serra et al., 2017), BEDTools (Quinlan and Hall, 2010), FastQC (http://www.bioinformatics.babraham.ac.uk / projects/fastqc/), Trimmomatic (Bolger et al., 2014), Mfuzz R package (2.26.0), sva R package (3.22), MACS2 (Zhang et al. 2008), MSigDB (Liberzon et al. 2011) OneD (Vidal et al. 2018), HiCExplorer (https://hicexplorer.readthedocs.io/en/latest/), DiffBind R package (https:// bioconductor.org/packages/release/bioc/html/DiffBind.html), coolpup.py (Flyamer et al., 2020), wilcox.test() R function, t.test() R function, SciPy (https://www.scipy.org), Prism (v8 for macOS), FlowJo (version 10.5.3 for macOS), Fiji (Schindelin et al., 2012)

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. 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 Hi-C, RNA-seq, CTCF-ChIP-seq, ATAC-seq datasets generated and analysed for the current study are available in the Gene Expression Omnibus (GEO) database under accession number GSE140528. ATAC-seq and CEBPA ChIP-seq datasets used in the current study are available in the GEO database under accession number GSE131620. The H3K27ac and H3K4me1 ChIP-seq datasets used in this study are available in ArrayExpress database under accession number E-MTAB-9010.

Field-specific reporting

Please select the one below that is 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/documents/nr-reporting-summary-flat.pdf

Life sciences study design

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

Sample size	Statistical methods were not used to predetermine sample sizes. All the sample sizes of seq data were based on standard sequencing analysis practices and were determined to be suitable for statistical analyses. Samples sizes for other experiments were based on standard practices and were determined to be suitable based on statistical power calculated following the experiments
Data exclusions	TAD border detection criteria were pre-established in order to increase the accuracy of the TAD borders detection. TAD borders not called in both independent biological replicates were excluded in all subsequent analyses (see Method Section). All read and bin filtering strategies used for Hi-C data analysis are described in detail in the Method section. Peak detection criteria were pre-established in order to increase the accuracy of the peak calling. CTCF Chip-seq peak not called in both independent biological replicates were excluded in all subsequent analyses (see Method Section).
Replication	Results were highly reproducible between biologically independent replicates (e.g. Extended Data Fig. 1). All attempts to replicate experiments and data analysis were successful.
Randomization	Randomization is not relevant to this study because no comparisons between experimental groups were made.
Blinding	Blinding was not relevant to this study because all metrics were derived from absolute quantitative methods without human subjectivity.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study	
☐	☒	Antibodies
☐	☒	Eukaryotic cell lines
☒	☐	Palaeontology
☒	☐	Animals and other organisms
☒	☐	Human research participants
☒	☐	Clinical data

Methods

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

Antibodies

Antibodies used	Human FcR Binding Inhibitor (eBiosciences, cat#16-9161-73, lot#4350612, dilution#1/20) APC-Cy7 Mouse Anti-Human CD19 (BD Pharmingen, cat#557791, clone#SJ25-C1, lot#9178618, dilution#1/100) APC Mouse Anti-Human CD11b/Mac-1 (BD Pharmingen, cat#550019, clone#ICRF44, lot#8009510, dilution#1/20) anti-α-tubulin (Abcam, cat#ab7291,clone#DM1A, dilution#1/1000) anti-mAID-tag (MBL Life Science,cat#M214-3, clone#1E4, lot#004, dilution#1/1000) anti-CTCF (Milipore, cat#07-729, lot#3059608, dilution#WP#1/1000 dilution#ChIP#1/200)
Validation	Human FcR Binding Inhibitor (eBiosciences, 16-9161-73) validated for flow cytometry (https://www.thermofisher.com/antibody/product/Fc-Receptor-Binding-Inhibitor-Antibody-Polyclonal/16-9161-73) and Turunen et al., 2016 APC-Cy7 Mouse Anti-Human CD19 (BD Pharmingen, 557791) validated for flow cytometry analysis of human cells (https://www.bdbiosciences.com/us/applications/research/clinical-research/oncology-research/blood-cell-disorders/surface-markers/human/apc-cy7-mouse-anti-human-cd19-sj25c1-also-known-as-sj25-c1/p/557791) APC Mouse Anti-Human CD11b/Mac-1validated for flow cytometry analysis of human cells (https://www.bdbiosciences.com/eu/applications/research/stem-cellresearch/mesenchymal-stem-cell-markers-bone-marrow/human/negative-markers/apc-mouse-anti-human-cd11bmac-1-icrf44-also-known-as-44/p/550019) anti-α-tubulin (Abcam ab7291) validated by Abcam (https://www.abcam.com/alpha-tubulin-antibody-dm1a-loading-controlab7291.html) and Kline et al., 2019 anti-mAID-tag (MBL Life Science M214-3) was validated by Nishimura, K. and Kanemaki, M. T., Curr. Protoc. Cell Biol. 64, 20.9.1-20.9.16 (2014)

anti-CTCF: Millipore, 07-729 (validated by ENCODE, see https://www.encodeproject.org/antibodycharacterizations/890eca82-62f8-406c-868a-c94a5a3d748e/@@download/attachment/human_CTCF_07-729_validation_Snyder.pdf)

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	We used the BLaER cell line previously established in our laboratory (Rapino et al., 2013)
Authentication	The cell line was validated by FACS analysis and by comparing the transcriptome with previous published data.
Mycoplasma contamination	The cell line was tested negative for Mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

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.

https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE140528

Files in database submission

CTCF0-1.bw
CTCF0-2.bw
CTCF24-1.bw
CTCF24-2.bw
CTCF96-1.bw
CTCF96-2.bw
CTCF168-1.bw
CTCF168-2.bw
C3_DMSO_24h.bw
C3_AUX_24h.bw
CTCF0-1_peaks.narrowPeak
CTCF0-2_peaks.narrowPeak
CTCF24-1_peaks.narrowPeak
CTCF24-2_peaks.narrowPeak
CTCF96-1_peaks.narrowPeak
CTCF96-2_peaks.narrowPeak
CTCF168-1_peaks.narrowPeak
CTCF168-2_peaks.narrowPeak
C3_DMSO_24h_peaks.narrowPeak
C3_AUX_24h_peaks.narrowPeak
CTCF0-1_R1.fastq.gz
CTCF0-2_R1.fastq.gz
CTCF24-1_R1.fastq.gz
CTCF24-2_R1.fastq.gz
CTCF96-1_R1.fastq.gz
CTCF96-2_R1.fastq.gz
CTCF168-1_R1.fastq.gz
CTCF168-2_R1.fastq.gz
C3_DMSO_24h_R1.fastq.gz
C3_AUX_24h_R1.fastq.gz
CTCF0-1_R2.fastq.gz
CTCF0-2_R2.fastq.gz
CTCF24-1_R2.fastq.gz
CTCF24-2_R2.fastq.gz
CTCF96-1_R2.fastq.gz
CTCF96-2_R2.fastq.gz
CTCF168-1_R2.fastq.gz
CTCF168-2_R2.fastq.gz
C3_DMSO_24h_R2.fastq.gz
C3_AUX_24h_R2.fastq.gz

Genome browser session
(e.g. UCSC)

No longer applicable

Methodology

Replicates

sample_name nb_of_replicates jaccard_index

Replicates	CTCF_0h 2 0.652404 CTCF_24h 2 0.681697 CTCF_96h 2 0.733294 CTCF_168h 2 0.699885 C3_DMSO_24h 1 C3_AUX_24h 1
Sequencing depth	sample name Total number of reads Uniquely mapped reads length of reads paired end CTCF0-1 26429075 21030038 50 nt PE CTCF0-2 25593098 19953487 50 nt PE CTCF24-1 20289338 16069419 50 nt PE CTCF24-2 26143838 20767865 50 nt PE CTCF96-1 32180358 25390504 50 nt PE CTCF96-2 25364829 19840847 50 nt PE CTCF168-1 17782525 13913789 50 nt PE CTCF168-2 21835140 17130058 50 nt PE C3_DMSO_24h 81068034 63835692 75 nt PE C3_AUX_24h 89550180 69430321 75 nt PE
Antibodies	anti-CTCF - Millipore - 07-729 (lot:3059608)
Peak calling parameters	/software/bin/macs2 callpeak -n ${describer} -t ${folder}${describer}.rmdup.bam -f BAMPE -g hs -q 0.05
Data quality	CTCF Chip-seq peak not called in both independent biological replicates were excluded in all subsequent analyses. sample_name nb_of_peaks peaks_above_5-fold_enrichment CTCF0-1_peaks.narrowPeak 46556 46556 CTCF0-2_peaks.narrowPeak 43253 43253 CTCF24-1_peaks.narrowPeak 45246 45246 CTCF24-2_peaks.narrowPeak 48699 48699 CTCF96-1_peaks.narrowPeak 46895 46895 CTCF96-2_peaks.narrowPeak 45546 45546 CTCF168-1_peaks.narrowPeak 33261 33261 CTCF168-2_peaks.narrowPeak 36803 36803 C3_DMSO_24h_peaks.narrowPeak 23639 23639 C3_AUX_24h_peaks.narrowPeak 6512 6512
Software	MACS2

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	Cells were collected and washed twice with PBS before staining. More detailed protocol is described in the Method section.
Instrument	All the analyses were performed using the LSR Fortessa instrument (BD Biosciences, San Jose, CA).
Software	Data analysis was performed using Flowjo software.
Cell population abundance	Cells were analyzed for the expression of the cell surface markers CD19 and ITGAM (Mac1). All the BLaER cells were positive for CD19 and after 7 days of transdifferentiation around 85 % of the cells lose CD19 marker and acquired ITGAM.
Gating strategy	Gates were defined using non stained cells as negative controls. An example of the gating strategy used is shown in Extended Data Figure 3a

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