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

Grégoire Stik^{1*}, Enrique Vidal¹, Mercedes Barrero¹, Sergi Cuartero^{1,2}, Maria Vilas-Casadesús¹, Julen Mendieta-Esteban³, Tian V. Tian^{1,4}, Jinmi Choi⁵, Clara Berenguer^{1,2}, Amaya Abad¹, Beatrice Borsari¹, François le Dily¹, Patrick Cramer⁵, Marc A. Marti-Renom^{1,3,6}, Ralph Stadhouders^{7,8*} and Thomas Graf^{1,9*}

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.

*Corresponding authors: gregoire.stik@crg.eu; r.stadhouders@erasmusmc.nl; thomas.graf@crg.eu

SUMMARY

Three-dimensional organization of the genome has emerged as an important player in transcriptional regulation¹⁻⁷. In mammals, CTCF and the cohesin complex create sub-megabase structures with elevated internal chromatin contact frequencies, called topologically associating domains (TADs)⁸⁻¹¹. Although TADs are generally considered important for transcriptional regulation, ablation of TAD organization by disrupting CTCF or the cohesin complex only caused modest gene expression changes¹²⁻¹⁵. In contrast, CTCF is required for cell cycle regulation¹⁶, early embryo development and for the formation of various adult cell types¹⁷. 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 causes an almost complete loss of TAD organization yet, surprisingly, the cells remain capable of efficiently transdifferentiating into functional macrophages. 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.

MAIN TEXT

Lineage instructive transcription factors establish new cell identities by activating a novel gene expression program while silencing the old one. Whereas they largely achieve this 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 level into active ('A') and inactive ('B') compartments¹⁸, themselves subdivided into topologically associating domains (TADs). TADs are separated by boundaries enriched for cohesin and the CCCTC-binding factor (CTCF)^{8,19,20}. Large deletions overlapping boundaries can cause a fusion of adjacent domains that can lead to developmental abnormalities²¹. In addition, inversion or deletion of individual CTCF binding sites can induce a loss of specific contacts or insulation from active chromatin²²⁻²⁴. Mechanistically, genomic insulation by TADs is thought to facilitate enhancer-promoter interactions while inhibiting cross-boundary communication between regulatory elements to prevent aberrant gene activation²⁵. Surprisingly, while acute depletion of CTCF or cohesin showed a dramatic loss of TADs, only modest changes in gene expression were detected^{12,13}. In contrast, CTCF gene deletion broadly perturbs mammalian development¹⁷, although it also impairs cell cycle progression¹⁶. Hence, the importance of CTCF and TAD organization in facilitating transcriptional rewiring during cell state transitions - often accompanied by extensive cell division - remains controversial.

We have recently developed a system uniquely suitable to study the role of CTCF in cell state transitions, consisting 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**; Supplementary Note 1)²⁶. Using this system, we analyzed a time-series of transdifferentiating cells for genome-wide changes in 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 compartment regions were dynamic during transdifferentiation, showing transcriptional changes correlating with the altered compartmentalization state (**Fig.1b-f**, **Extended Data Fig.1a-d**; Supplementary Note 2). Next, we used chromosome-wide insulation potential²⁷ to identify between 3,100-3,300 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 analysed 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 significantly more enriched at stable than at dynamic boundaries (**Fig.1h**), as observed earlier²⁸. 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 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 indeed showed 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 approach²⁹ (**Fig.2a**;

Supplementary Note 3). Addition of auxin to these cells triggered proteasome-dependent CTCF degradation, resulting in a loss of mCherry⁺ cells and a rapid CTCF depletion to levels 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 cells¹³. We next performed Hi-C on cells cultured in the presence of auxin or DMSO (as a control) at 24 and 168 hours post-induction (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 wildtype 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 24hpi and 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 24hpi and iMac stages (**Extended Data Fig.2g**). Consequently, the ratio of contact enrichment inside TADs over outside was also strongly decreased (**Extended Data Fig.2h**). Whereas stable borders exhibited a dramatic loss of insulation after CTCF depletion, dynamic borders showed essentially no change (**Fig.2e**), in agreement with their low CTCF occupancy (**Fig.1h**).

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**; see 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 0h, 24hpi and in iMacs. E-P interaction frequencies significantly decreased during inactivation, which was somewhat accelerated in auxin treated cells at 24hpi and iMac stages (**Fig.2f**). Similarly, E-P interaction frequencies significantly 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 are 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 Mac-1 (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 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 24hpi 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 after 24h, 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 knock down of CTCF accelerates myeloid commitment of common myeloid precursor cells³⁰. A heatmap 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 *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 like the cell cycle, GTPase signalling 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 interactions (5 -10 Mb) between enhancers and promoters revealed that their activation and 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 during transdifferentiation from one compartment to another 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 a previous study¹³, we observed an approximately 10% reduction in compartment strength of the auxin-treated samples (**Fig.3e**), which could explain the slight acceleration of compartment transitions observed in auxin treated cells at 24hpi (**Fig.3f**). An example is provided by the *MAFB* locus, a myeloid-specific 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.3g**), including at an enhancer region that becomes decorated with H3K27ac at 24hpi (**Extended Fig.3j**). In short, our Hi-C data revealed that although CTCF appears dispensable for genome compartmentalization, its depletion slightly decreased compartmentalization strength, which could facilitate compartmental rearrangements.

Previous reports indicated that CTCF plays a role in controlling macrophage gene expression³¹ and that cohesin is required for an optimal inflammatory response of macrophages³². This raised the possibility of an involvement of genome topology in mounting an acute inflammatory response, as CTCF is known to stabilize the interaction between cohesin and chromatin^{33,34}. Accordingly, aggregates of our Hi-C signals at previously described cohesin-bound loops³⁵ showed that these interaction punctuates disappeared after auxin treatment (**Extended Data Fig.4a**). Using a public dataset of lipopolysaccharide (LPS) responsive genes³⁶ 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 iMac3 exposed for 2h to LPS (**Fig.4a**), revealing a reduced induction of critical LPS-responsive genes such as *IL6*, *TNFA*, *CCL2* (**Fig.4b**). Even more pronounced changes were observed at the level of secreted cytokines 8h 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 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 *IL-6*, *TNFA* and *CCL2*) were strongly associated to pathways related to the inflammatory response to bacterial stimuli (**Extended Data Fig.4e-f**). Although a sizeable fraction of differentially expressed genes were already altered in auxin-treated iMacs prior to LPS stimulation, the total number of affected genes doubled after LPS exposure (**Extended Data Fig.4g**) and the observed upregulation upon LPS treatment was significantly blunted after CTCF depletion (**Extended Data Fig.4h**). Of note, the expression of these genes was not significantly changed in CTCF-depleted iMacs prior to LPS treatment (**Extended Data Fig.4i**) and most of the key transcription factors and receptors involved in the LPS response were unaffected after 24h 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 repressor³⁷ in this context. A phagocytosis assay with DMSO- and 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 also more strongly insulated than random gene sets (**Extended Data Fig.5a-b**), suggesting they could be extra susceptible to deregulation by a 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 neighbouring TADs but also decreased the frequency of *IL6* E-P interactions (**Fig.4d** and **Extended Data Fig.5c**). Interestingly, the neighbouring 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 restraints³⁸. 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, are not essential for developmental gene regulation but instead provide 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 significantly 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^{12,13,39-41}. In addition, our findings indicate a critical role of 3D chromatin organization in providing an optimal response to external signals, in agreement with studies of developmentally regulated loci and nuclear hormone receptor

signaling^{39,42}. Future studies are required to assess whether this can be further generalized to signalling 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 raise the possibility that lineage instructive transcription factors themselves shape multi-level topological genome dynamics relevant for major transcriptional rewiring.

Acknowledgments

We thank M. T. Kanemaki for the degron plasmids; Roderic Guigo laboratory, and Silvia Perez-Lluch 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 the laboratory of T.G. 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., M.A.M.-R.), the Ministerio de Educacion y Ciencia (SAF.2012-37167 to T.G and BFU2017-85926-P to M.A.M.-R.), the AGAUR to T.G, 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 antibody arrays. G.S realized 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 Statement

The authors declare no competing interests

Additional information

Extended data is available for this paper

Supplementary information is available for this paper

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

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 (80-.)*. **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. P. *et al.* Cohesin Loss Eliminates All Loop Domains. *Cell* **171**, 305–320 e24 (2017).
13. Nora, E. P. *et al.* Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell* **169**, 930–944 e22 (2017).
14. Schwarzer, W. *et al.* Two independent modes of chromatin organization revealed by cohesin removal. *Nature* **551**, 51–56 (2017).
15. Haarhuis, J. H. I. *et al.* The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. *Cell* **169**, 693–707.e14 (2017).
16. Heath, H. *et al.* CTCF regulates cell cycle progression of $\alpha\beta$ T cells in the thymus. *EMBO J.* **27**, 2839–2850 (2008).
17. Arzate-Mejía, R. G., Recillas-Targa, F. & Corces, V. G. Developing in 3D: the role of CTCF in cell differentiation. *Development (Cambridge, England)* vol. 145 (2018).

18. Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* (80-.). **326**, 289–293 (2009).
19. Phillips-Cremins, J. E. & Corces, V. G. Chromatin insulators: linking genome organization to cellular function. *Mol Cell* **50**, 461–474 (2013).
20. 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).
21. Lupiáñez, D. G. *et al.* Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* (2015) doi:10.1016/j.cell.2015.04.004.
22. Guo, Y. *et al.* CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. *Cell* **162**, 900–10 (2015).
23. 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.* (2015) doi:10.1073/pnas.1518552112.
24. Narendra, V. *et al.* CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *Science* (80-.). (2015) doi:10.1126/science.1262088.
25. de Laat, W. & Duboule, D. Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature* **502**, 499–506 (2013).
26. Rapino, F. *et al.* C/EBPalpha induces highly efficient macrophage transdifferentiation of B lymphoma and leukemia cell lines and impairs their tumorigenicity. *Cell Rep* **3**, 1153–1163 (2013).
27. Crane, E. *et al.* Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature* **523**, 240–244 (2015).
28. Stadhouders, R. *et al.* Transcription factors orchestrate dynamic interplay between genome topology and gene regulation during cell reprogramming. *Nat Genet* **50**, 238–249 (2018).
29. 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).
30. Ouboussad, L., Kreuz, S. & Lefevre, P. F. CTCF depletion alters chromatin structure and transcription of myeloid-specific factors. *J. Mol. Cell Biol.* **5**, 308–22 (2013).

31. Nikolic, T. *et al.* The DNA-binding factor Ctfc critically controls gene expression in macrophages. *Cell. Mol. Immunol.* **11**, 58–70 (2014).
32. Cuartero, S. *et al.* Control of inducible gene expression links cohesin to hematopoietic progenitor self-renewal and differentiation. *Nat Immunol* **19**, 932–941 (2018).
33. Parelho, V. *et al.* Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* **132**, 422–433 (2008).
34. Wendt, K. S. *et al.* Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* **451**, 796–801 (2008).
35. Mumbach, M. R. *et al.* HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat Methods* **13**, 919–922 (2016).
36. Faridi, M. H. *et al.* CD11b activation suppresses TLR-dependent inflammation and autoimmunity in systemic lupus erythematosus. *J. Clin. Invest.* **127**, 1271–1283 (2017).
37. 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).
38. 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**, (2017).
39. 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).
40. Ghavi-Helm, Y. *et al.* Highly rearranged chromosomes reveal uncoupling between genome topology and gene expression. *Nat. Genet.* (2019) doi:10.1038/s41588-019-0462-3.
41. Williamson, I. *et al.* Developmentally regulated Shh expression is robust to TAD perturbations. *Dev.* **146**, (2019).
42. 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).
43. 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).

- 440 44. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for
441 Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
- 442 45. Ay, F. *et al.* Identifying multi-locus chromatin contacts in human cells using
443 tethered multiple 3C. *BMC Genomics* **16**, (2015).
- 444 46. Vidal, E. *et al.* OneD: increasing reproducibility of Hi-C samples with abnormal
445 karyotypes. *Nucleic Acids Res* **46**, e49 (2018).
- 446 47. Ramírez, F. *et al.* High-resolution TADs reveal DNA sequences underlying
447 genome organization in flies. *Nat. Commun.* **9**, (2018).
- 448 48. Flyamer, I. M., Illingworth, R. S. & Bickmore, W. A. Coolpup.py: versatile pile-
449 up analysis of Hi-C data. *Bioinformatics* (2020)
450 doi:10.1093/bioinformatics/btaa073.
- 451 49. Dobin, A. *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**,
452 15–21 (2013).
- 453 50. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and
454 dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, (2014).
- 455 51. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2.
456 *Nat. Methods* **9**, 357–359 (2012).
- 457 52. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics*
458 **25**, 2078–2079 (2009).
- 459 53. Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A. & Manke, T. DeepTools: A
460 flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42**,
461 (2014).
- 462 54. Feng, J., Liu, T., Qin, B., Zhang, Y. & Liu, X. S. Identifying ChIP-seq
463 enrichment using MACS. *Nat. Protoc.* **7**, 1728–1740 (2012).
- 464 55. Quinlan, A. R. & Hall, I. M. BEDTools: A flexible suite of utilities for comparing
465 genomic features. *Bioinformatics* **26**, 841–842 (2010).
- 466 56. Heinz, S. *et al.* Simple Combinations of Lineage-Determining Transcription
467 Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell
468 Identities. *Mol. Cell* **38**, 576–589 (2010).
- 469 57. Ross-Innes, C. S. *et al.* Differential oestrogen receptor binding is associated
470 with clinical outcome in breast cancer. *Nature* **481**, 389–393 (2012).
- 471 58. Miguel-Escalada, I. *et al.* Human pancreatic islet three-dimensional chromatin
472 architecture provides insights into the genetics of type 2 diabetes. *Nat. Genet.*
473 (2019) doi:10.1038/s41588-019-0457-0.

59. Stefano, M. Di *et al.* Dynamic simulations of transcriptional control during cell reprogramming reveal spatial chromatin caging. *bioRxiv* 1–29 (2019) doi:10.1101/642009.
60. Baù, D. & Marti-Renom, M. A. Genome structure determination via 3C-based data integration by the Integrative Modeling Platform. *Methods* vol. 58 300–306 (2012).
61. Pettersen, E. F. *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

Figure Legends

Fig.1 | Transcription factor-driven transdifferentiation re-wires 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 50Mb DNA region of B cells and iMacs. Color scale represents the normalized number of contacts per read. **c**, Transformation of the Hi-C map based on the PC1 values of a PCA on the Hi-C correlation matrix. PC1 values for A and B compartment 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 grey arrow indicating transdifferentiation trajectory. **e**, Proportion of dynamic compartment bins (dyn.) including its distribution of different sub-categories. **f**, Integration of gene expression associated with dynamic compartment using RNA-seq (n represents the number of genes and p 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**, Top: Differential contact map (iMac minus B cell signal) at the *DDX54* locus. Color scale represents differential contacts per 100,000 reads. Bottom: Snapshot of genome browser showing CTCF ChIP-seq signals at the locus. CTCF peaks at the

507 newly created border are highlighted. All box plots depict the first and third quartiles
508 as the lower and upper bounds of the box, with a thicker band inside the box showing
509 the median value and whiskers representing 1.5x the interquartile range.

Fig.2 | Auxin-mediated depletion of CTCF impairs chromatin insulation at stable but not dynamic TAD boundaries. **a**, Schematic representation of auxin-mediated 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 3 times with similar results. **c**, Western blot showing loss of CTCF in CTCF-mAID B cells treated with auxin. Detection of TUBA4A was used as a loading control. The blots have been cropped from original blots available in Source Data. The experiment was repeated 3 times with similar results. **d**, Top: Representative Hi-C contact maps (20 kb resolution) of a 10 Mb region in chromosome 7 from transdifferentiating cells (24hpi) treated with DMSO or auxin. Color scale represents the normalized number of contacts. Bottom: insulation score line graphs across the locus. **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 log2 fold changes. Estimations are computed using all 9 samples in a single linear mixed model.

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. 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 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). Grey points connected by an arrow represent non-tagged B cell transdifferentiation. Ellipses group 24hpi 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 biological 2 replicates (p, two-sided Wilcoxon rank-sum test). **d**, Heatmap 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**, Top: saddle plot showing pairwise enrichment of the 20% top and bottom PC1 values from Hi-C contacts at 100 kb bins (see Methods). Lower 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 anti-diagonal 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. 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.

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 24h followed by 2-8h of LPS treatment and assayed for cytokine expression. **b**, qRT-PCR of selected cytokine genes 2h after LPS stimulation of iMacs pre-treated with DMSO or auxin. Error bars represent standard error; sample sizes (n) are indicated and represent biologically independent samples and p unpaired two-tailed t test. **c**, Secreted cytokine levels by iMacs treated with DMSO or auxin and stimulated for 8h with LPS. Error bars represent standard error (n=3 biologically independent samples) **d**, Top: 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. Color scale represents differential contacts per 100,000 reads. The location of the *IL6* and *STEAP1B* genes is indicated; Middle part: Virtual 4C extracted from Hi-C data at the *IL6* locus in iMacs treated with DMSO or auxin, using the *IL6* transcription start site (TSS) as a viewpoint; Bottom part: browser snapshot showing CTCF and H3K27ac ChIP-seq signals. *IL6* enhancers (e1 and e2) are highlighted and the green, red and blue spheres represent the *STEAP1B* promoter, the *IL6* enhancers and the *IL6* TSS, respectively. **e**, Distance distribution between TSS and enhancer regions (n=1,000 3D models based on Hi-C data of iMacs treated with DMSO or auxin). Median (solid

line), first and third quartile (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.

Methods

Cell Culture

BLaER cell line²⁶ is derived from RCH-ACV lymphoblastic leukaemia cell line in which CEBPA fused with the estrogen receptor hormone binding domain and the GFP marker are expressed. BLaER cells and subclones were cultured in RPMI medium (Gibco, 22400089) supplemented with 10% fetal bovine serum (GIBCO, 10100147), 1% glutamine (GIBCO, 25030081), 1% penicillin/streptomycin antibiotic (Thermo Fisher Scientific, 15140122), 550 μ M β -mercaptoethanol (GIBCO, 31350010). Cells were maintained at a density of $0.1\text{--}6 \times 10^6$ cells/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 completed with 100 nM β -estradiol, IL3 and CSF1 (100 ng/mL). iMacs were collected after 7 days of incubation. For auxin-inducible degradation, indole-3-acetic acid (IAA, a chemical analogue of auxin) was added to the medium at 500 μ M from a 1000X stock diluted in dimethyl sulfoxide. 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 μ g/mL) for 2h to collect RNA or 8h to collect supernatant.

Plasmid construction

The CTCF-mAID-mCherry targeting vectors were cloned by serial modification of the base vectors pMK292 (Addgene #72830) and pMK293 (Addgene #72831). Homology arms (HA) of the last exon of CTCF were synthesised (IDT®). The TIR1-AAVS1 donor vector pMK232 (Addgene #105924) and the pX330 vector expressing the sgRNA to target the AAVS1 locus in human cells (Addgene #72833) were kindly provided by Masato Kanemaki²⁹. CTCF-targeting sgRNAs were cloned in pX330 by annealing oligonucleotides caccgTGATCCTCAGCATGATGGAC and aaacGTCCATCATGCTGAGGATCAc .

Gene Targeting

For transfection, plasmids were prepared using 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 (Amaza Nucleofector, Lonza®), using Kit C and program X-001, according to manufacturer's instructions. One microgram of each plasmid (pMK232 and pX330-AAVS-sgRNA) was added per 100 μ L of solution mix and 1 million cells. Eight millions of 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 μ g/mL, Gibco A1113803) was added to the medium in order to select edited cells. Selection medium was changed each 2-3 days and the selection was performed for 10 days. 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 qPCR allowing the selection of the clone with the most robust expression (cell line #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 μ g of each plasmid (px330-mCherry-sgRNA; pHA-mAID-mCherry-Neo R or pHA-mAID-mCherry-Hygro R) was added per 100 μ L of solution mix and 1 million cells. Eight millions of 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 in order to select edited cells (500 μ g/mL G418, Life Technologies #11811031 and/or 100 μ g/mL of Hygromycin B, GIBCO #10687010). Selection medium was changed each 2-3 days and the selection was performed for 17-20 days. 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 (DAPI-negative) were sorted by live flow cytometry on a BD InfluxTM 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 (RT) using Human FcR Binding Inhibitor (1:20 dilution, eBiosciences, 16-9161-73) and cells were then stained with antibodies against CD19 (APC-Cy7 Mouse Anti-Human CD19, BD Pharmingen, 557791) and Mac-1 (APC Mouse Anti-Human CD11b/Mac-1, BD Pharmingen, 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/mL in medium and fluoresbrite carboxy bright blue beads (1µm, Polysciences 17458) were added (300 beads/cell) and incubated 24h before 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 the 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 1X and lysis was performed in 30 µL of Laemmli Buffer 1X (50 mM Tris-HCl pH6.8, 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 $5 \cdot 10^5$ cells) were separated by electrophoresis in a 7.5% polyacrylamide gel (Bio-Rad #4561023) before transfer to nitrocellulose membrane. Membranes were blocked with 5% non-fat milk TBS-Tween medium (50 mM Tris, 150 mM NaCl, 0.1% Tween 20) for 1h at RT. Incubation with primary antibody was performed at 4°C shaking overnight (anti-CTCF, 07-729 Millipore; anti- α -tubulin, ab7291 Abcam; anti-mAID-tag MBL Life Science M214-3; 1:1000 in 5% Milk in TBS-Tween). Membranes were washed with TBS-Tween (3 x 10 min) before secondary incubation with antibodies fused to HRP (goat anti-mouse IgG, Sigma Aldrich #A3682, dilution 1:5000) for 1h at RT. After 3 final washes, membranes were incubated in ECL™ Start Western Blotting Detection Reagent mix (Sigma Aldrich, GERPN3243) for 2 min at RT before development on X-ray film. Cytokine arrays (R&D ARY006) were performed following manufacturers' instructions using supernatant from iMacs collected 8 h after LPS stimulation (1µg/mL). Antibody arrays were imaged using an Odyssey CLx instrument (LI-COR).

Immunofluorescence

iMacs were grown on glass-coverslips, fixed with 3% formaldehyde in PBS1X for 10 min at RT. After washing with PBS, imaging was performed using a Leica TCS SPE inverted microscope. Images were post-processed 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 glycine. Cell pellets were lysed by incubating 10 min on ice with 5mM Pipes pH 8, 85 mM KCL, 0.5% IGEPAL, 1X protease inhibitor (Roche®). After centrifugation, pellets were incubated in 1% SDS, 10 mM EDTA pH 8, 50 mM Tris-HCL pH 8.1 and 1X PIC for 10 min on ice. Chromatin was sheared on a Bioruptor pico sonicator (Diagenode) at 4°C for 14 cycles of 30 sec ON and 30 sec OFF. After sonication, the solution was left on ice for 1h to allow SDS precipitation and clarified by centrifugation at 16,000g at 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 1X cold IP buffer (Diagenode). 10 µg of anti-CTCF (Milipore, 07-729) was added followed by overnight incubation at 4°C on a rotator. 42 µL of beads (Unblocked Protein A beads, kch-503-008, Diagenode) were used per IP after blocking them using 1% bovine serum albumin cold IP buffer for 15 min at 4°C under rotation. Blocked beads were added to the chromatin solution and incubated 3h at 4°C with rotation. Beads were then collected by centrifugation for 2 min at 3000 rpm at 4°C and washed 3 times with cold IP buffer and 2 times with cold TE buffer (10mM Tris pH 8, 1mM EDTA). Beads were then eluted with freshly prepared elution buffer (1% SDS, 0.1M NaHCO₃) and incubated 25 min at RT. The supernatant was transferred into a new tube and cross-linking was reversed by adding NaCl (final concentration 200mM) and incubating overnight at 65°C. Protein digestion was achieved by adding Tris pH 6.5 (40mM), EDTA pH 8 (10mM) and proteinase K (4 µg/µL) and incubating 1h at 45°C. DNA was then purified by phenol chloroform isoamyl (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 adaptor 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 bp). Libraries were sequenced on Illumina HiSeq2000 or NextSeq 500 instrument using 50 or 75 nucleotides paired-end mode, respectively.

Quantitative RT-PCR and RNA-seq

RNA was extracted with the miRNeasy mini kit (Qiagen) and quantified with a NanoDrop spectrophotometer. cDNA was produced with a High Capacity RNA-to-cDNA kit (Applied Biosystems) and was used for qRT-PCR analysis in triplicate reactions with SYBR Green QPCR Master Mix (Applied Biosystems). Primers are available upon request. Libraries were prepared with an Illumina TrueSeq Stranded total RNA Library Preparation Kit after Ribo-zero® depletion, and single-end sequencing (75 nt) was performed on an Illumina HiSeq2500 instrument.

ATAC-seq

ATAC-seq was performed as previously described⁴³. Briefly, 5 million cells were harvested and treated with Nextera Tn5 Transposase (Illumina, FC-121-1030) for 45 min at 37°C. Library fragments were amplified using 1x NEBNext High-Fidelity 2x PCR Master Mix (NEB, M0541S) and 1.25 µM of custom Nextera PCR primers. PCR amplification was done with 11 cycles, determined by KAPA Real-Time Library Amplification Kit (Peqlab, KK2701) to stop prior to saturation. Then, the samples were purified using MinElute PCR Purification Kit (Qiagen, 28004) and with Agencourt AMPure XP beads (Beckman Coulter, A63881) in 3:1 ratio. The libraries were sequenced paired-end (50bp) on a HiSeq2000 instrument.

***In situ* Hi-C library preparation**

In situ Hi-C was performed as previously described²⁰ with the following modifications: (i) two million cells were used as starting material; (ii) chromatin was initially digested with 100 U Mbol (New England BioLabs) for 2h, and then another 100 U (2h incubation) and a final 100 U were added before overnight incubation; (iii) before fill-in with bio-dATP, nuclei were pelleted and resuspended in fresh 1× NEB2 buffer; (iv) ligation was performed overnight at 24°C with 10,000 cohesive end units per reaction; (v) 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 sec on and 60 sec off); (vi) DNA fragment-size selection was performed only after final library amplification; (vii) library preparation was performed with an NEBNext DNA Library Prep Kit (New England BioLabs) with 3 µL NEBNext adaptor in the ligation step; (viii) 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 based on TADbit³⁸. First, quality of the reads was checked using FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) to discard problematic samples and detect systematic artefacts. Trimmomatic⁴⁴ 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 protocols⁴⁵. Briefly, each side of the sequenced read was mapped in full length to the reference genome (hg38, Dec 2017 GRCh38). 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 non-informative 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 artefact. This filter only removed extra copies of the original pair.

• random breaks: start position of one of the reads was too far from RE cutting site, possibly due to non-canonical enzymatic activity or random physical breaks. Threshold was set to 750 bp (default), > 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 OneD⁴⁶. 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 topologically associated domains (TADs)

To segment the genome into A/B compartments, normalized Hi-C matrices at 100kb resolution were corrected for decay as previously published, grouping diagonals when signal-to-noise was below 0.05²⁰. 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. Significant differences of PC1 values between conditions were calculated using two-sided Wilcoxon rank-sum tests. Normalized contact matrices at 50kb resolution were used to define TADs, using a previously described method²⁷ with default parameters. First, for each bin, an insulation score was obtained based on 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 score²⁷. 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. Significant differences of insulation scores between conditions were calculated using two-sided Wilcoxon rank-sum tests.

Inter and intra-compartment strength measurements

We followed a previously reported strategy to measure overall interaction strengths within and between A and B compartments¹⁴. Briefly, we based our analysis on the 100kb 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 log2 enrichment over the expected counts by distance decay. Finally, we summarize each type of interaction (A-A, B-B and A-B/B-A) by taking the median values of the log2 contact enrichment.

Meta-analysis of borders

To study the behaviour 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 log2 fold change of contact enrichment between bins inside and outside 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 labelled as part of an enhancer, promoter or “others” if they did not belong to previous types. Then the observed contacts were gathered between the different types of bins within their TAD and expected contact frequencies were computed based on the genomic distance that separate each pair (the expected distance decay was calculated excluding entries outside TADs). Then a linear mixed models including TAD ID as random effect was used to estimate the quantities of interest. Results are expressed as Log2 of the ratio observed on expected frequencies of contacts.

849

850 **Long-range interactions between enhancers and promoters**

851 HiC matrices were generated at 10-kb resolution using HiCExplorer⁴⁷ and long range
852 interactions (5-10 Mb) between promoters and enhancers activated or inactivated
853 were computed using the HiCExplorer tool hicAggregateContacts.

854

855 **Meta-analysis of cohesin loops**

856 Hi-C matrices in cool format were used to generate genome-wide aggregate plots at
857 SMC1 loops detected by HiChIP³⁵. We used *coolpup.py*⁴⁸ to pile-up normalized Hi-C
858 signals at a 10-kb resolution at SMC1 loop previously identified³⁵, and plotted 100 kb
859 upstream and downstream of the SMC1 anchor coordinates.

860

861 **Virtual 4C analysis**

862 Hi-C matrices for virtual 4C profiles were further smoothed using a focal (moving
863 window) average of one bin. The profiles were generated from these normalized
864 matrices and correspond to histogram representation of the lines of the matrices
865 containing the baits (therefore expressed as counts per thousand of normalized
866 reads within the region depicted).

867

868 **Gene expression analysis using RNA-seq data**

869 Reads were mapped using STAR⁴⁹ (standard options) and the Ensembl human
870 genome annotation (GRCh38v27). Gene expression was quantified using STAR (--
871 quantMode GeneCounts). Batch effects were removed using ComBat function from
872 sva R package (v3.22). Sample scaling and statistical analysis were performed using
873 the R package DESeq2⁵⁰ (R 3.3.2 and Bioconductor 3.0). Genes changing
874 significantly at any time point were identified using the nbinomLRT test (FDR<0.01)
875 and fold-change>2 between at least two time points. Log2-vsds (variance stabilized
876 DESeq2) counts were used for further analysis unless stated otherwise. To compare
877 expression of various set of genes the data were mean-centered log-transformed and
878 significant differences were calculated using two-sided Wilcoxon rank-sum tests.

879

880 **Chromatin accessibility analysis using ATAC-seq data**

881 Reads were mapped to the UCSC human genome build (hg38) using Bowtie2⁵¹ with
882 standard settings. Reads mapping to multiple locations in the genome were removed

using SAMtools⁵²; PCR duplicates were filtered using Picard (http://broadinstitute.github.io/picard). Bam files were parsed to deepTools⁵³ for downstream analyses and browser visualization. Peaks in ATAC-seq signal were identified using MACS2⁵⁴ (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 MACS2⁵⁴ 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 BEDTools⁵⁵. H3K27ac coverage and CTCF binding heatmaps were performed using deepTools⁵³.

DNA motif analysis

ATAC peaks specific to the TAD borders were identified using bedtools⁵⁵. DNA motif analysis of the ATAC peaks were analysed using HOMER⁵⁶ (*findMotifs.pl*) and the Homer motif results were shown. It uses ZOOPS scoring (zero or one occurrence per sequence) coupled with 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 H3K27ac signals at these regions was quantified using the Diffbind R package to define activated and inactivated regions from 0h to 168h (differences were computed with DiffBind R package (v2.2.12) using DBA_DESEQ2 method -filter for significance was set at foldchange 1 and FDR<0.05) as previously described⁵⁷. This analysis allowed to profile 29,711 dynamic enhancers and 8,439 dynamic promoters of which about half became activated and the other half inactivated (**Extended Data Fig.2h**), also reflected by the expression of the associated genes (**Extended Data Fig.2i**).

3D modelling and analysis

The processed Hi-C datasets were binned at 10 kb resolution and then normalized using OneD⁴⁶. Then, we defined the regions to be modelled given the genomic

context around the enhancer and promoters of interest by following the steps: (i) select key elements contained in the region (i.e., enhancers and promoters); (ii) retrieve the top 5% interactors of each of these elements; (iii) 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. Added the edge twice if it was in the top 5% of both members; (iv) group the networks allowing a genomic distance gap of 50 kb and filtered out the groups in which the ratio between (number of edges) / (Number of nodes) was smaller than 5; and (v) for each of the regions, ensure that the modelled 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 modelled as previously described⁵⁸ using TADdyn⁵⁹, a molecular dynamic-based protocol implemented in the TADbit software³⁸. Similarly to TADbit, TADdyn generates models using a restraint-based approach, in which experimental frequencies of interaction are transformed into a set of spatial restraints⁶⁰. A total of 1,000 models were generated for each genomic region and cell type. Contact maps generated from the ensembled models highly correlated with the input Hi-C normalized interaction matrices. Each ensemble of models was next clustered based on structural similarity as implemented in TADbit³⁸. The absence of major structural differences between clusters prompted us to use all them in further analysis. Next, TADbit was used to measure the following features of the models: (i) distance between particles containing genomic regions of interest in the model ensemble; distances distribution between selected pairs or particles, and (iii) significant differential distance distributions assessed by two sample Kolmogorov-Smirnov statistic. Finally, model images were generated with Chimera⁶¹.

Statistics and reproducibility

RNA-seq and in situ Hi-C data throughout the paper were generated by analysis of 2 biologically independent samples from 2 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 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-side manner using the module `scipy.stats.ks_2sam` in the SciPy software.

Reporting Summary

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

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.

Methods-only-References

43. 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).
44. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
45. Ay, F. *et al.* Identifying multi-locus chromatin contacts in human cells using tethered multiple 3C. *BMC Genomics* **16**, (2015).
46. Vidal, E. *et al.* OneD: increasing reproducibility of Hi-C samples with abnormal karyotypes. *Nucleic Acids Res* **46**, e49 (2018).
47. Ramírez, F. *et al.* High-resolution TADs reveal DNA sequences underlying genome organization in flies. *Nat. Commun.* **9**, (2018).
48. Flyamer, I. M., Illingworth, R. S. & Bickmore, W. A. Coolpup.py: versatile pile-up analysis of Hi-C data. *Bioinformatics* (2020) doi:10.1093/bioinformatics/btaa073.

49. Dobin, A. *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
50. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, (2014).
51. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
52. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
53. Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A. & Manke, T. DeepTools: A flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42**, (2014).
54. Feng, J., Liu, T., Qin, B., Zhang, Y. & Liu, X. S. Identifying ChIP-seq enrichment using MACS. *Nat. Protoc.* **7**, 1728–1740 (2012).
55. Quinlan, A. R. & Hall, I. M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
56. 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).
57. Ross-Innes, C. S. *et al.* Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481**, 389–393 (2012).
58. Miguel-Escalada, I. *et al.* Human pancreatic islet three-dimensional chromatin architecture provides insights into the genetics of type 2 diabetes. *Nat. Genet.* (2019) doi:10.1038/s41588-019-0457-0.
59. Stefano, M. Di *et al.* Dynamic simulations of transcriptional control during cell reprogramming reveal spatial chromatin caging. *bioRxiv* 1–29 (2019) doi:10.1101/642009.
60. Baù, D. & Marti-Renom, M. A. Genome structure determination via 3C-based data integration by the Integrative Modeling Platform. *Methods* vol. 58 300–306 (2012).
61. Pettersen, E. F. *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

a **b** **c** **d** **e** **f** **g** **h** **i**

a **b** **c** **d** **e** **f**

a **b** **c** **d** **e** **f** **g**

a **b** **c** **d** **e** **f**

SUPPLEMENTARY INFORMATION

Supplementary Notes

Supplementary Note 1. Transdifferentiation of human B cells into macrophages

In this study, we used a system consisting of a B leukemia cell line (BLaER) that can be efficiently converted by exogenous C/EBP α expression into functional induced macrophages (iMacs) with only one cell division on average (**Fig.1a**)¹. Given that C/EBP α exerts its lineage-determining capacity by predominantly binding to distal enhancers^{2,3}, we reasoned that our transdifferentiation system would heavily rely on long-range transcriptional control and therefore on CTCF-mediated 3D genome organization. BLaER cells expressing C/EBP α fused to the estrogen receptor were treated with β -estradiol (β -est), converting cells to phagocytic iMacs 168 hours post-induction (hpi) (**Fig.1a**).

Supplementary Note 2. Chromatin compartmentalization dynamics during transdifferentiation of B cells into macrophages

First, we studied the A/B chromatin compartmentalization dynamic on the basis of the first eigenvector values of a principal component analysis on the Hi-C correlation matrix (PC1 values), showing local transitions between compartments (**Fig.1b-c**). Overall, A-B compartmentalization was highly reproducible between biological duplicates and revealed changes across the genome over time (**Extended Data Fig.1a-b**). PCA highlighted a progressive rewiring of genome compartmentalization during transdifferentiation (**Fig.1d**). Although most regions assigned to either the A (40%) or B (60%) compartment remained essentially constant throughout transdifferentiation (**Extended Data Fig.1c**), around 14% were dynamic (**Fig.1e**). Of these, ~25% switched from B-to-A, ~40% from A-to-B and the remaining regions showed transient transitions (A-B-A and B-A-B) (**Fig.1e**). Genes associated to regions exhibiting a B-to-A switch were upregulated on average and found to be enriched for 'immune system'-related functions whereas genes in A-to-B switching domains were downregulated and enriched for 'transcription'-related processes (**Fig.1f, Extended Data Fig.1d**). In contrast, expression of genes in A-B-A regions remained unchanged, while genes in B-A-B domains were slightly upregulated

(Fig.1f), raising the possibility of a ‘memory effect’ of regions that have transiently moved into the A compartment.

Supplementary Note 3. Auxin-inducible degron approach to deplete CTCF during transdifferentiation of B cells into macrophages

We used CRISPR/Cas9 technology to introduce into BLaER cells, a transgene encoding the Tir1 F-box protein from *Oryza sativa* and targeted the stop codon of both CTCF alleles to introduce a 68-amino acid version of the auxin-inducible degron (mAID) tag together with a mCherry cassette. Addition of auxin to culture medium then triggers proteasome-dependent CTCF degradation also resulting in a loss of mCherry⁺ fluorescence (Fig.2b-c).

Supplementary Table 1. Sequences of oligonucleotides used for qRT-PCR

Oligonucleotide	Sequence
IL6_Foward	AGTGAGGAACAAGCCAGAGC
IL6_Reverse	ATTTGTGGTTGGGTCAGGGG
IL1B_Foward	CGCCAGTGAAATGATGGCTT
IL1B_Reverse	ATCCAGAGGGCAGAGGTCC
TNF_Foward	ACTTTGGAGTGATCGGCCC
TNF_Reverse	CATTGGCCAGGAGGGCATT
CCL2_Foward	ATCAATGCCCCAGTCACCTG
CCL2_Reverse	TCTCCTTGGCCACAATGGTC
CCL1_Foward	CGGAGCAAGAGATTCCCCTG
CCL1_Reverse	TGCCTCTGAACCCATCCAAC
CCL5_Foward	TCATTGCTACTGCCCTCTGC
CCL5_Reverse	CACACTTGGCGGTTCTTTTCG
CSF3_Foward	GAGTGTGCCACCTACAAGCT
CSF3_Reverse	CCGCTATGGAGTTGGCTCAA
GAPDH_Foward	CAGCCTCAAGATCATCAGCA
GAPDH_Reverse	TGTGGTCATGAGTCCTTCCA

Supplementary References

1. Rapino, F. *et al.* C/EBPalpha induces highly efficient macrophage transdifferentiation of B lymphoma and leukemia cell lines and impairs their tumorigenicity. *Cell Rep* **3**, 1153–1163 (2013).
2. Di Stefano, B. *et al.* C/EBPalpha creates elite cells for iPSC reprogramming by

upregulating Klf4 and increasing the levels of Lsd1 and Brd4. *Nat Cell Biol* **18**, 371–381 (2016).

3. van Oevelen, C. *et al.* C/EBPalpha Activates Pre-existing and De Novo Macrophage Enhancers during Induced Pre-B Cell Transdifferentiation and Myelopoiesis. *Stem Cell Reports* **5**, 232–247 (2015).

a **b** **c** **d** **e** **f** **g** **h**

	Stable borders	Gained borders	Lost borders
Rank1	CTCF CCACIAGATGGC $p=1.e-101$	SPIB AAAGAGGAAAGT $p=1.e-92$	EBF CTCCTIAGGG $p=1.e-12$
Rank2	ETS1 ACAGGAAGTG $p=1.e-72$	PU.1 GAAATGAAC $p=1.e-27$	RUNX CTGTGGTTTC $p=1.e-11$
Rank3	RUNX1 AAACCCACA $p=1.e-40$	BORIS CGCCCTAGTG $p=1.e-17$	ZBTB18 TGGCCATCGT $p=1.e-11$

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 are calculated using hypergeometric statistical tests).

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 24h 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 log₂ 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 | 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**, qRT-PCR of selected cytokine genes 2h after LPS stimulation of 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 $10\mu\text{m}$). 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 $1.5\times$ the interquartile range.

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 4C 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 4C 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.