Title: Replication timing maintains the global epigenetic state in human cells.

Authors: Kyle N. Klein[1]†, Peiyao A. Zhao[1]†, Xiaowen Lyu[2]†, Takayo Sasaki[1,3], Daniel A. Bartlett[1], Amar M. Singh[4], Ipek Tasan[5], Meng Zhang[5], Lotte P. Watts[6], Shin-ichiro Hiraga[6], Toyoaki Natsume[7], Xuemeng Zhou[8], Timour Baslan[9], Danny Leung[8], Masato T. Kanemaki[7], Anne D. Donaldson[6], Huimin Zhao[5], Stephen Dalton[4], Victor G. Corces[2], David M. Gilbert[1,3*].

Affiliations:

[1]Department of Biological Science, 319 Stadium Drive, Florida State University, Tallahassee, FL 32306, USA.

[2]Department of Human Genetics, Emory University School of Medicine, Atlanta, GA 30322, USA.

[3]San Diego Biomedical Research Institute, La Jolla, CA 92121, USA.

[4]Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602, USA.

[5]Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.

[6]Institute of Medical Sciences, University of Aberdeen, Aberdeen, AB25 2ZD, UK.

[7]Department of Chromosome Science, National Institute of Genetics, Research Organization of Information and Systems (ROIS), Yata 1111, Mishima, Shizuoka 411-8540, Japan. Department of Genetics, The Graduate University for Advanced Studies (SOKENDAI), Yata 1111, Mishima, Shizuoka 411-8540, Japan.

[8]Divison of Life Science, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China.

[9]Cancer Biology and Genetics Program, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA.

[*]Correspondence to: gilbert@bio.fsu.edu

[†]These authors contributed equally.

Abstract: The temporal order of DNA replication (replication timing; RT) is correlated with chromatin modifications and 3D genome architecture, however causal links have not been established, largely due to an inability to manipulate the global RT program. We show that loss of RIF1 causes near complete elimination of the RT program by increasing heterogeneity between individual cells. RT changes are coupled with widespread alterations in chromatin modifications and genome compartmentalization. Conditional depletion of RIF1 causes replication-dependent disruption of histone modifications and alterations in genome architecture. These effects were magnified with successive cycles of altered RT. These results support models in which the timing of chromatin replication and thus assembly plays a key role in maintaining the global epigenetic state.

One Sentence Summary: Replication timing controlled by RIF1 helps to dictate chromatin state and 3D genome organization.

Main Text: DNA is replicated during S phase of the cell cycle in a temporal order known as the replication timing (RT) program. RT is conserved among eukaryotes, developmentally controlled, and correlates with many important epigenomic features (*1*). Early replicating chromatin generally contains active histone modifications, is located in the nuclear interior and correlates with the A compartment defined by Hi-C (*2*) while late replicating chromatin is associated with transcriptionally repressive histone modifications, localization at the nuclear periphery, and the B-compartment. Histone modifications are both recycled from parental chromatin and added de novo after passage of the replication fork with different chromatin states showing differing dynamics of reassembly (*2*, *3*). Despite these close correlations, the mechanistic link between RT and the accurate maintenance of chromatin through cell cycles remains elusive. It has long been hypothesized that RT influences chromatin maintenance. Indeed, microinjection of plasmids into mammalian nuclei revealed that plasmids replicated in early S phase were decorated with acetylated histones, while those replicated later in S phase were devoid of acetylated histones (*4*, *5*). However, there is no direct evidence implicating RT in epigenetic state maintenance, largely due to the inability to manipulate genome wide RT. The conserved protein RIF1 has been shown to affect RT in many eukaryotes, however, because the effects have been partial or localized, RIF1 disruption has not been exploited to study the effects of RT abrogation (*6*, *7*).

RIF1 KO causes heterogenous RT

To gain insight into the role of genome-wide RT in shaping the epigenome we knocked out (KO) RIF1 in H9 hESC, HCT116, and HAP1 cell lines (Fig. S1, A to C). As previously reported (*8*), all three RIF1 KO cell lines proceeded through the cell cycle with nearly WT kinetics (Fig. S1, D and E) and exhibited genome wide aberrations in RT albeit with varying degrees of severity (Fig. 1A). Similar to prior reports in mammalian cells (*6*), domains changed RT either from early to late (EtL) or late to early (LtE) in log2(E/L) RT profiles of HCT116 and HAP1 cells. RIF1 KO caused 43% of the genome to change RT in HCT116 (23% EtL and 20% LtE) and 39% to change in HAP1 cells (20% EtL and 19% LtE) (Fig S1F). However, in H9 hESCs (Fig. 1A), nearly the entire genome acquired a log2(E/L) close to zero (Fig S1F). RIF1 control of RT in hESCs was dosage dependent as partial knockdown (KD) of RIF1 in H9 hESCs (Fig S2A) resulted in a partial effect on the RT program (Fig S2B and C). Replication foci (*9*) in RIF1 KO cells showed a ‘blending’ of early and middle spatial patterns (Fig S3). These results demonstrate a considerably more extensive role for RIF1 in RT control in hESCs versus other cell types.

Previous reports have interpreted RT changes in RIF1 KO cells as distinct RT switches (*6*) but the severity of the RT phenotype in RIF1 KO hESCs suggested an alternative mechanism. To address this, we performed high resolution Repli-seq (*10*) (Fig S4A) in RIF1 KO cells, which identifies peaks of replication initiation termed initiation zones (IZs) and valleys of late replication containing broadly distributed, low efficiency initiation events (*10*). We detected dramatic diffusion of RT patterns and loss of defined IZs in both HCT116 and H9 hESCs (Fig 1B, S4B) indicating major RT variation within the cell population (Fig S4C and D). Most remarkably, there was a nearly complete genome-wide abrogation of the RT program in RIF1 KO H9 hESCs (Fig 1B and S4B). In RIF1 KO HCT116 cells, EtL and LtE regions called using E/L Repli-seq showed substantial loss of temporal control (Fig S4E and F). Even early regions that were not called as EtL switches in E/L Repli-seq also lost defined patterns (Compare Figs. 1A and B, S4E and F) indicating that the entirety of the early replicating genome lost RT

specificity upon RIF1 KO in both cell lines. By contrast, many late replicating regions in HCT116 retained late replication (Fig S4E and F) indicating a RIF1-independent mechanism controlling RT for these regions. We next divided each IZ called in WT high resolution Repli-seq into four S phase timing classes (*10*) and plotted the cumulative percentage of DNA replicated through S phase (Fig 1C). WT cells showed typical segregation of IZs according to the temporal order (*10*) while RIF1 KO cells showed major overlap of IZ classes and flatter sigmoidal-like curves (Fig 1G). We subsequently calculated the genome wide heterogeneity parameter T_{width} which is positively correlated with heterogeneity (Methods). We found that genome wide T_{width} was greatly increased in both cell types upon RIF1 KO (Fig S4G) indicating a significant increase in RT heterogeneity.

To directly validate increased cell to cell RT heterogeneity in RIF1 KO we performed single-cell Repli-seq (*11*) on HAP1 WT and RIF1 KO cells. RIF1 KO cells showed co-replication of early and late domains within the same cell (Fig S5) and significantly larger T_{width} compared to WT (Fig 1D) directly confirming that RIF1 KO disrupts RT by substantially increasing RT heterogeneity rather than causing discrete RT shifts in all cells.

RIF1 KO alters distribution of heterochromatic histone modifications

To map RIF1 binding, we performed Cut&Run against GFP on GFP tagged RIF1 in HCT116 and H9 hESCs (Fig S6A). RIF1 was enriched in the late replicating portion of the genome in both cell lines (Fig S6B) and bound chromatin in broad domains (Fig 2A). In both cell lines, RIF1 binding was enriched at regions that lost RT control upon RIF1 KO (affected regions) (Fig 2, A, B, and E, S6C to E) while those regions that maintained their late RT (unaffected regions) had low RIF1 enrichment (Fig 2, C, D, and E, S6C and D). However, in RIF1 KO H9 hESCs, the regions defined as ‘unaffected regions’ are rare and still display RT dysregulation, but significantly less than ‘affected regions’ (compare Fig 2D to Fig 2B, Methods).

We next performed ChIP-seq (spike-in) on the late replication associated histone mark H3K9me3 in RIF1 KO cells (Fig S7). In WT cells, affected regions were enriched for smaller H3K9me3 peaks while unaffected regions contained large H3K9me3 domains (Fig 2, A, C, and E, S8A and B). In HCT116, small H3K9me3 peaks at affected regions (n=390) were lost (Fig 2, A and E), while the large H3K9me3 domains at unaffected regions (n=209) were strengthened (Fig 2, C and E). H3K9me3 was globally depleted in H9 hESCs (Fig S8C, D). The rare unaffected domains in RIF1 KO H9 hESCs (n=49) were large H3K9me3 domains that were only slightly diminished for H3K9me3 compared to affected domains (Fig 2E). H3K9me3 depletion in HCT116 RIF1 KO cells (Fig S8F, Methods) caused partial earlier replication of unaffected regions (Fig 2F, Fig S8G, H) indicating a role for H3K9me3 in their late replication but implicating additional unknown mechanisms (possibly association with the nuclear lamina) in maintaining late RT at unaffected regions. By contrast, RIF1 KO caused cell type specific changes to H3K27me3 with H9 hESCs primarily exhibiting downregulation and overall upregulation in HCT116 (Fig S9). These results reveal that multiple mechanisms, including RIF1 and large H3K9me3 domains, orchestrate late RT to different extents in different cell types.

Affected and unaffected late replicating regions form separate interaction hubs

The strong association between late replication and B compartmentalization (*12*) compelled us to investigate the genomic compartmentalization of RIF1 KO cells by Hi-C. Surprisingly, we found that affected and unaffected regions had distinct PC1 eigenvector associations despite being similarly late replicating in WT cells and this distinction was furthered upon RIF1 KO (Fig

S10A). WT interactions between unaffected regions were significantly stronger than interactions between affected regions or those between unaffected regions and affected regions (Fig S10B). Upon RIF1 KO, unaffected regions formed new interactions and strengthened existing interactions (Fig 2G, S10C). Strengthened interactions were correlated with upregulated H3K9me3 domains (Fig S10D). We then sorted affected and unaffected regions according to the extent of H3K9me3 changes (negative: downregulated, positive: upregulated) and calculated their interaction frequencies. In WT cells upregulated and downregulated H3K9me3 peaks form separate interaction hubs (Fig 2H left column) that were strengthened and weakened respectively upon RIF1 KO (Fig 2H right column), accentuating the separation between the two. In individual cases, the juxtaposition of affected/unaffected domains created new compartment and TAD boundaries (Fig S10 E and F). These results demonstrate that affected and unaffected H3K9me3 regions are two intrinsically different classes of late replicating chromatin with distinct interaction preferences; unaffected domains form strong interactions with one another to form a compartment that maintains late RT without RIF1 while affected domains form a separate hub of interactions and require RIF1 to enforce late replication.

RIF1 KO causes depleted active histone modifications and attenuated A compartment interactions

ChIP-seq of H3K27ac (spike-in) and H3K4me3 revealed that both marks became depleted in both RIF1 KO cells (Fig S7, S11A to C). Loss of active marks from the A compartment was concurrent with changes in compartmentalization and Hi-C interactions in both cell lines. In H9 hESCs, the majority of compartment changes involved the disappearance of A compartment chromatin into neighboring B compartments (Fig 3A and B; Fig S11D), while HCT116 exhibited discrete shifts in compartmentalization in both directions (Fig 3A, Fig S11D and E). Loss of A compartment interactions in Hi-C contact maps corresponded to reduced H3K27ac peaks in both cell lines [Fig 3C (inset arrows), Fig S11F (arrows)]. Within A to B compartment switches, the levels of H3K27ac and H3K4me3 were significantly depleted (Fig 3D) and the interactions between H3K27ac peaks were reduced (Fig 3E). These data show that depletion of active histone modifications, particularly H3K27ac, strongly correlates with loss of A compartment interactions in RIF1 KO cells.

Both cell lines exhibited genome wide weakening of A/A compartment interactions accompanied by strengthening of B/B compartment interactions (Fig S12A). This was further confirmed when we called statistically significant differential interactions using diffHiC. Strengthened interactions were seen predominantly within the B compartment and weakened interactions were concentrated in the A compartment in both cell lines (Fig S12B and C). This is likely the combined effect of strengthened interactions between upregulated H3K9me3 domains within the B compartment (Fig S12D) and weakened interactions between downregulated H3K27ac peaks in the A compartment (Fig S12E) and suggests a redefining of compartment identities in RIF1 KO cells. EtL and LtE regions showed occasional loss and gain of TAD boundaries respectively (Fig S13A) but globally, the positioning and number of TAD boundaries was not significantly affected by RIF1 KO in either cell line (Fig S13B), while the strength of TAD boundaries was increased (Fig S13 C and D) independent of RAD21 binding (Fig S14). Together these data indicate a substantial reorganization of genomic compartments correlating with epigenome changes in RIF1 KO cells.

DNA replication is required for epigenomic changes

To directly address RIF1's primary role, we used an auxin inducible degron system (*13*) in HCT116 cells to rapidly degrade RIF1 (Fig S15A). We first confirmed that RT defects observed in RIF1 KO occur during the first S phase after RIF1 loss (Fig S15B and C) by degrading RIF1 in G1-synchronized cells and performing E/L Repli-seq (Methods) upon release into S phase. We next synchronized cells in G1, degraded RIF1, and either held them in G1 or released them into early, middle, or late S phase/G2 (Fig S15D and E), and performed Hi-C and ChIP-seq. We found no significant difference in genome organization between RIF1 degraded and control cells held in G1 (Fig 4A, S16A to C). We observed extensive architectural changes at both A and B compartment chromatin as control cells transition from G1 into and through S phase (Fig 4A, S16A to C), controlled for copy number (Fig S16D), characterized by initial increase and subsequent decrease in interaction frequencies within A compartment and the converse within B compartment regions (Fig 4A, control lines). RIF1 KD cells showed diminished dynamics with decreased intra-A interactions and increased intra-B interactions during early S phase and the reverse during mid and late S phase compared to control cells (Fig 4A KD lines, S16C). These changes are coincidental with RT disruptions at corresponding loci (Fig 4B, S16E). These results demonstrate that RIF1 is required for extensive architectural changes during S phase, suggesting that these changes are linked to the normal temporal order of replication.

H3K27ac and H3K9me3 ChIP-seq (spike-in) also revealed no significant difference between RIF1 degraded and control cells held in G1 (Fig 4C to H, S16 F and G, G1 data), indicating that histone mark changes are also not a direct result of RIF1 loss. H3K27ac peaks were not depleted until cells reached middle S phase with the extent of depletion greatest at late S phase [Fig 4C, D, S16F (blue dots)] suggesting that disruption of early-enriched histone marks takes place only when they are replicated late. H3K9me3 peaks at affected regions became depleted in late S/G2 [Fig 4E, F, S16G (blue dots)] while H3K9me3 domains at unaffected regions became more enriched for H3K9me3 at late S phase/G2 [Fig 4G, H, S16G (red dots)]. Together these results indicate that disruption of genomic structure and chromatin identity in RIF1 degraded cells is a result of a dysregulated DNA replication program.

Epigenomic changes are quantitatively correlated with lengths of RT dysregulation

We further hypothesized that extended degradation of RIF1 and RT disruption through multiple rounds of DNA replication would exacerbate the observed epigenomic phenotypes. To test this, we degraded RIF1 for 24, 48, and 96 hours (Fig S17A) constituting ~1, 2, and 4 cell cycles respectively. After 24 hours of RIF1 degradation the RIF1 KO RT phenotype was fully recapitulated and remained so throughout the time course (Fig S17B and C). H3K27ac peaks that were depleted in RIF1 KO cells, and the interactions between them, progressively diminished toward RIF1 KO levels with time (Fig 5A and B, S17D and E). H3K9me3 peaks at affected late replicating regions became significantly depleted after 24 hours of RIF1 degradation and remained depleted throughout the time course (Fig 5C and D, S17F). Unaffected regions progressively accumulated H3K9me3 throughout the time course (Fig 5E and F), however the strengthening of interactions between H3K9me3 domains seen in RIF1 KO was not observed even after 96 hours of RIF1 degradation (Fig S17G). Sorting chromatin interactions by ΔRT revealed that regions that changed RT (both EtL and LtE) gradually lost interactions with regions of normally similar RT (Fig 5G, upper left and lower right corners) and gained interactions with regions of normally different RT (Fig 5G, upper right and lower left corners), indicating a strong correlation between RT loss and disruption of chromatin contacts. Together these data show that as cells progress through multiple rounds of DNA replication with dysregulated RT, they

gradually accumulate aberrant epigenomic and structural signatures, supporting the model that stochastic RT gradually affects newly synthesized chromatin with each round of mis-timed chromatin assembly.

RIF1 KO causes limited gene expression changes

RIF1 KO caused expression changes of 2284 genes in H9 hESCs and 1737 genes in HCT116 (Fig S18A and B), which did not include pluripotency factors or histone modification writers (Fig S18C). Genes affected in HCT116 RIF1 KO gradually changed their expression during the RIF1-AID degradation time course toward RIF1 KO levels (Fig S18D), however only 6 significantly changed genes were shared between all RNA-seq samples (Fig S18E and Table S1). Single cell RNA-seq revealed an increase in global cell-to-cell heterogeneity in gene expression in RIF1 KO H9 hESCs and HAP1 cells, but not HCT116 (Fig S18 F to I).

RT and chromatin compartmentalization are correlated with gene expression (*1*), however neither changes in RT nor compartment switches were able to predict gene expression changes in RIF1 KO cells (Fig S19 A to D). Differentially expressed genes showed cell type specific changes in the distribution of specific histone modifications around their transcription start sites that correlated with expression changes (Fig S19 E and F). Altogether, these results suggest that gene expression changes are indirect effects of RIF1 depletion.

Discussion

We show that deletion of RIF1 results in a dramatic increase in cell-to-cell RT heterogeneity rather than discrete population level RT switches. Although RIF1 is bound to late replicating chromatin, the entire early replicating genome as well as most of the late replicating genome becomes disrupted for RT control without RIF1 (Fig 1). This is reminiscent of observations in budding yeast where de-repression of late replicating origins increased competition for limiting replication initiation factors and delayed the RT of early origins (*13*) and suggests a similar mechanism of RT disruption in RIF1 KO cells. The late replicating genome is composed of two types of domains whose delayed replication is enforced by different mechanisms and that form separate chromatin hubs. In both hESCs and HCT116 cells, RIF1 KO leads to widespread aberrant histone modification patterns that correlate with distinct genome-wide changes in 3D genome architecture. We further show that disruptions to epigenome and chromatin structure require DNA replication and that continuous rounds of replication with heterogeneous RT quantitatively exacerbate these changes. We propose that RT changes due to RIF1 KO result in aberrant re-establishment of epigenetic marks that cause profound changes in the epigenetic landscape that then alter genome architecture (Fig 5H). This work provides the first mechanistic evidence linking the RT program with maintenance of the global epigenetic state and genome compartmentalization and establishes RIF1 as a key regulator of epigenome maintenance through its role in RT control.

References:

1. N. Rhind, D. M. Gilbert, DNA Replication Timing. *Cold Spring Harb. Perspect. Biol.* **5**, a010132 (2013).
2. N. Reverón-Gómez, C. González-Aguilera, K. R. Stewart-Morgan, N. Petryk, V. Flury, S. Graziano, J. V. Johansen, J. S. Jakobsen, C. Alabert, A. Groth, Accurate Recycling of Parental Histones Reproduces the Histone Modification Landscape during DNA Replication. *Mol. Cell* (2018), doi:10.1016/j.molcel.2018.08.010.
3. T. M. Escobar, O. Oksuz, R. Saldaña-Meyer, N. Descostes, R. Bonasio, D. Reinberg,

Active and Repressed Chromatin Domains Exhibit Distinct Nucleosome Segregation during DNA Replication. *Cell* (2019), doi:10.1016/j.cell.2019.10.009.

4. J. Zhang, F. Xu, T. Hashimshony, I. Keshet, H. Cedar, Establishment of transcriptional competence in early and late S phase. *Nature* (2002), doi:10.1038/nature01150.
5. L. Lande-Diner, J. Zhang, H. Cedar, Shifts in Replication Timing Actively Affect Histone Acetylation during Nucleosome Reassembly. *Mol. Cell*. **34**, 767–74 (2009).
6. S. B. C. Buonomo, in *Advances in experimental medicine and biology* (2017; http://www.ncbi.nlm.nih.gov/pubmed/29357062), vol. 1042, pp. 259–272.
7. C. A. Seller, P. H. O'Farrell, Rif1 prolongs the embryonic S phase at the Drosophila mid-blastula transition. *PLoS Biol.* **16**, e2005687 (2018).
8. R. Foti, S. Gnan, D. Cornacchia, V. Dileep, A. Bulut-Karslioglu, S. Diehl, A. Buness, F. A. A. Klein, W. Huber, E. Johnstone, R. Loos, P. Bertone, D. M. M. Gilbert, T. Manke, T. Jenuwein, S. C. B. C. B. B. Buonomo, Nuclear Architecture Organized by Rif1 Underpins the Replication-Timing Program. *Mol. Cell*. **61**, 260–273 (2016).
9. R. T. O'Keefe, S. C. Henderson, D. L. Spector, Dynamic organization of DNA replication in mammalian cell nuclei: spatially and temporally defined replication of chromosome-specific alpha-satellite DNA sequences. *J. Cell Biol.* **116**, 1095–1110 (1992).
10. P. A. Zhao, T. Sasaki, D. M. Gilbert, High-resolution Repli-Seq defines the temporal choreography of initiation, elongation and termination of replication in mammalian cells. *Genome Biol.* (2020), doi:10.1186/s13059-020-01983-8.
11. V. Dileep, D. M. Gilbert, Single-cell replication profiling to measure stochastic variation in mammalian replication timing. *Nat. Commun.* **9**, 427 (2018).
12. T. Ryba, I. Hiratani, J. Lu, M. Itoh, M. Kulik, J. Zhang, T. C. Schulz, A. J. Robins, S. Dalton, D. M. Gilbert, Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. *Genome Res.* **20**, 761–770 (2010).
13. K. Nishimura, T. Fukagawa, H. Takisawa, T. Kakimoto, M. Kanemaki, An auxin-based degron system for the rapid depletion of proteins in nonplant cells. *Nat. Methods*. **6**, 917–922 (2009).
14. K. Yoshida, J. Bacal, D. Desmarais, I. Padioleau, O. Tsaponina, A. Chabes, V. Pantesco, E. Dubois, H. Parrinello, M. Skrzypczak, K. Ginalski, A. Lengronne, P. Pasero, The Histone Deacetylases Sir2 and Rpd3 Act on Ribosomal DNA to Control the Replication Program in Budding Yeast. *Mol. Cell*. **54**, 691–697 (2014).
15. C. Marchal, T. Sasaki, D. Vera, K. Wilson, J. Sima, J. C. Rivera-Mulia, C. Trevilla-García, C. Nogues, E. Nafie, D. M. Gilbert, Genome-wide analysis of replication timing by next-generation sequencing with E/L Repli-seq. *Nat. Protoc.* **13**, 819–839 (2018).
16. T. Baslan, J. Kendall, L. Rodgers, H. Cox, M. Riggs, A. Stepansky, J. Troge, K. Ravi, D. Esposito, B. Lakshmi, M. Wigler, N. Navin, J. Hicks, Genome-wide copy number analysis of single cells. *Nat. Protoc.* **7**, 1024–1041 (2012).
17. X. Lyu, M. J. Rowley, V. G. Corces, Architectural Proteins and Pluripotency Factors Cooperate to Orchestrate the Transcriptional Response of hESCs to Temperature Stress. *Mol. Cell*. **71**, 940-955.e7 (2018).
18. D. A. Orlando, M. W. Chen, V. E. Brown, S. Solanki, Y. J. Choi, E. R. Olson, C. C. Fritz, J. E. Bradner, M. G. Guenther, Quantitative ChIP-Seq normalization reveals global modulation of the epigenome. *Cell Rep.* (2014), doi:10.1016/j.celrep.2014.10.018.
19. P. J. Skene, S. Henikoff, An efficient targeted nuclease strategy for high-resolution

mapping of DNA binding sites. *Elife* (2017), doi:10.7554/elife.21856.
20. Y. Zhang, T. Liu, C. A. Meyer, J. Eeckhoute, D. S. Johnson, B. E. Bernstein, C. Nussbaum, R. M. Myers, M. Brown, W. Li, X. S. Shirley, Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* (2008), doi:10.1186/gb-2008-9-9-r137.
21. E. B. Stovner, P. Sætrom, epic2 efficiently finds diffuse domains in ChIP-seq data. *Bioinformatics* (2019), doi:10.1093/bioinformatics/btz232.
22. A. Soufi, G. Donahue, K. S. Zaret, Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell* (2012), doi:10.1016/j.cell.2012.09.045.
23. E. Lieberman-Aiden, N. L. van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B. R. Lajoie, P. J. Sabo, M. O. Dorschner, R. Sandstrom, B. Bernstein, M. A. Bender, M. Groudine, A. Gnirke, J. Stamatoyannopoulos, L. A. Mirny, E. S. Lander, J. Dekker, Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science (80-.).* **326**, 289–293 (2009).
24. M. Imakaev, G. Fudenberg, R. P. McCord, N. Naumova, A. Goloborodko, B. R. Lajoie, J. Dekker, L. A. Mirny, Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat. Methods* (2012), doi:10.1038/nmeth.2148.
25. J. R. Dixon, S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, M. Hu, J. S. Liu, B. Ren, Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*. **485**, 376–380 (2012).
26. B. R. Lajoie, J. Dekker, N. Kaplan, The Hitchhiker's guide to Hi-C analysis: Practical guidelines. *Methods* (2015), doi:10.1016/j.ymeth.2014.10.031.
27. A. T. L. Lun, G. K. Smyth, diffHic: A Bioconductor package to detect differential genomic interactions in Hi-C data. *BMC Bioinformatics* (2015), doi:10.1186/s12859-015-0683-0.
28. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* (2009), doi:10.1093/bioinformatics/btp616.
29. Q. Jiang, J. Y. J. Ang, A. Y. Lee, Q. Cao, K. Y. Li, K. Y. Yip, D. C. Y. Leung, G9a Plays Distinct Roles in Maintaining DNA Methylation, Retrotransposon Silencing, and Chromatin Looping. *Cell Rep.* (2020), doi:10.1016/j.celrep.2020.108315.
30. A. Dobin, C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T. R. Gingeras, STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* (2013), doi:10.1093/bioinformatics/bts635.
31. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* (2009), doi:10.1093/bioinformatics/btp352.
32. A. R. Quinlan, I. M. Hall, BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* (2010), doi:10.1093/bioinformatics/btq033.
33. L. Breiman, Random forests. *Mach. Learn.* **45**, 5–32 (2001).

Acknowledgments: We would like to thank Ruth Didier and Beth Alexander of the FSU Flow Cytometry and Confocal Microscopy Facilities for their excellent help with flow cytometry and FACS for this project. Thanks to Amber Brown of the FSU Biological Science Core Labs and to Yangming Yang and Cindy Vied of the FSU Translational Labs. Thanks to Sandy R. Westermann of SCIGRAPHIX for generating the model figure. Thanks to Bas van Steensel, Jennifer Phillips-Cremins, and Peter Fraser for critical reading of the manuscript. **Funding:** This

work was supported by NIH grant GM083337 to D.M.G., GM035463 to V.G.C., and GM085354 to D.M.G. and V.G.C.; T.B. is supported by the William C. and Joyce C. O'Neil Charitable Trust, Memorial Sloan Kettering Single Cell Sequencing Initiative. **Author contributions:** Conceptualization, K.N.K., P.A.Z., V.G.C., and D.M.G; Methodology, K.N.K., A.S., I.T., M.Z., L.P.W., S.H., T.N., and T.B.; Formal Analysis, P.A.Z.; Investigation, K.N.K., X.L., T.S., D.A.B., and X.Z.; Writing – Original Draft, K.N.K., P.A.Z., and D.M.G.; Writing – Review and Editing, K.N.K., P.A.Z., X.L., V.G.C., and D.M.G.; Visualization, K.N.K. and P.A.Z.; Supervision, D.L., M.T.K., A.D.D., H.Z., S.D., V.G.C., and D.M.G.; Funding Acquisition, D.M.G. and V.G.C. **Competing interests:** Authors declare no competing interests. **Data and materials availability:** All data is available on the Gene Expression Omnibus (GEO) #GSE160563. Data analysis scripts are available upon request.

Supplementary Materials:

Materials and Methods

Figures S1-S19

Tables S1-S5

References (*14-33*)

Figure 1

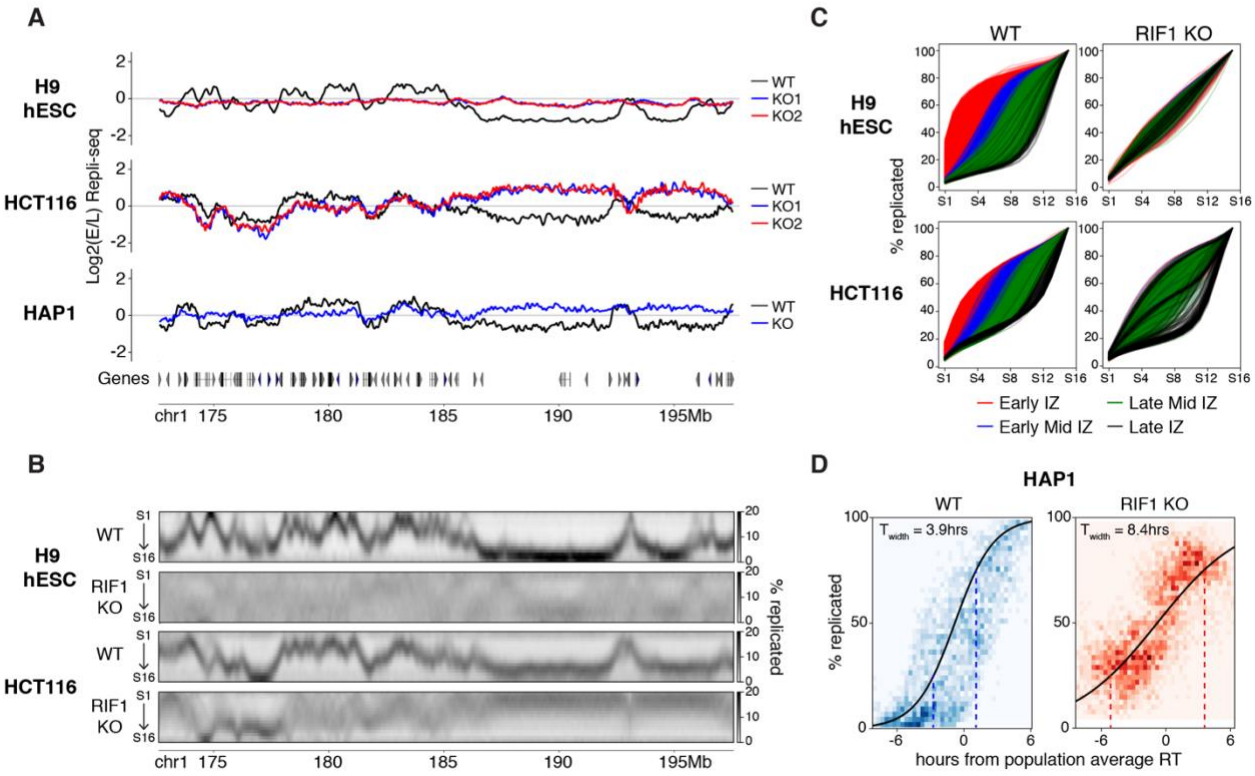


Fig. 1. RIF1 controls RT by reducing cell to cell variation in replication timing.

(A) Log2(E/L) Repli-seq plots of Chr1 172.6-197.6 Mb in WT (black) and two RIF1 KO clones (blue and red) in H9 hESCs (top) and HCT116 (middle) and HAP1 (bottom) cell lines. (B) High resolution Repli-seq plots of Chr1 172.6-197.6 Mb in WT and RIF1 KO in H9 hESCs (top two) and HCT116 (bottom two); same locus as (A). (C) Cumulative percent-replicated plots for each IZ called in WT cells versus S phase fraction of 16 fraction Repli-seq color coded by their timing (red: early, blue: early mid, green: late mid, black: late). (D) Sigmoidal fitting of the percentage replicated (y-axis) against time in hours from population average RT (x-axis) for HAP1 WT (left) and RIF1 KO (right). The heatmaps (blue: WT and red: KO) represent the data spread for all 50kb bins genome wide in all cells. Dotted lines at 25% of cells replicated and 75% of cells replicated indicate the span of T_{width}.

Figure 2

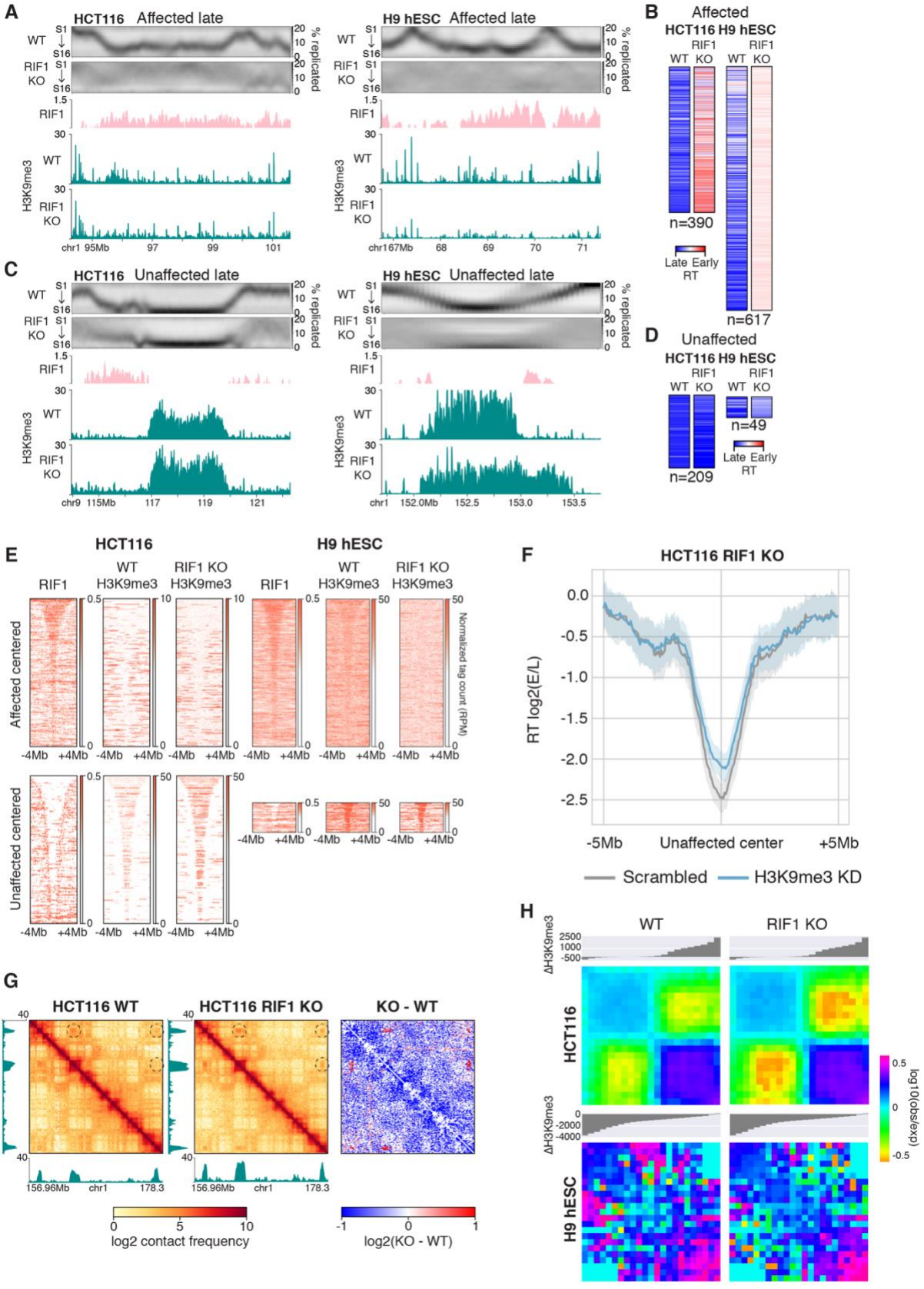

A
HCT116 Affected late
H9 hESC Affected late
WT
RIF1 KO
% replicated
RIF1
H3K9me3
B
Affected
HCT116 H9 hESC
RIF1
WT KO
n=390
n=617
Late Early RT
C
HCT116 Unaffected late
H9 hESC Unaffected late
D
Unaffected
n=49
n=209
E
HCT116
H9 hESC
RIF1
WT H3K9me3
RIF1 KO H3K9me3
Affected centered
Unaffected centered
Normalized tag count (RPM)
F
HCT116 RIF1 KO
RT log2(E/L)
-5Mb
Unaffected center
+5Mb
Scrambled
H3K9me3 KD
G
HCT116 WT
HCT116 RIF1 KO
KO - WT
log2 contact frequency
log2(KO - WT)
H
WT
RIF1 KO
ΔH3K9me3
HCT116
H9 hESC
log10(obs/exp)

Fig. 2. RT affected and unaffected late regions are distinct classes of chromatin.

(A) Affected late region at Chr1 94.3-101.6 Mb in HCT116 (left) and Chr1 66.75-71.4 Mb in H9 hESCs (right) showing from top to bottom: high resolution Repli-seq in WT and RIF1 KO cells, RIF1 fold enrichment in WT cells, and H3K9me3 ChIP-seq tag counts in WT and RIF1 KO cells. (B) Heat maps of RT indices for affected late regions in WT and RIF1 KO of HCT116 and H9 hESCs. (C) Unaffected late region at Chr9 113.95-122.3 Mb in HCT116 (left) and Chr1 151.7-153.75 Mb in H9 hESCs (right) showing the same panels as (A). (D) Heat maps of RT indices for unaffected late regions in WT and RIF1 KO of HCT116 and H9 hESCs. (E) RIF1 binding and tag of H3K9me3 in HCT116 (left) and H9 hESC (right) WT or RIF1 KO cells centered on affected late regions (top) or unaffected late regions (bottom) ±4 Mb and sorted by size. (F) Mean log2(E/L) Repli-seq scores centered at all unaffected late regions ±5 Mb in HCT116 RIF1 KO scrambled control (grey) or H3K9me3 KD (blue) cells. Shadow represents 95% confidence interval. (G) H3K9me3 tag count tracks beside ICE normalized Hi-C and subtraction contact maps of HCT116 WT and RIF1 KO at Chr1 156.95-178.3Mb. Dotted circles denote regions of increased interaction. (H) Log10(obs/exp) aggregate interactions between late regions in WT and RIF1 KO HCT116 (top) and H9 hESCs (bottom). The interactions were binned into 11 equal segments, which were ranked by increasing ΔH3K9me3 where negative and positive values respectively indicate decrease and increase in H3K9me3 in RIF1 KO compared to WT.

Figure 3

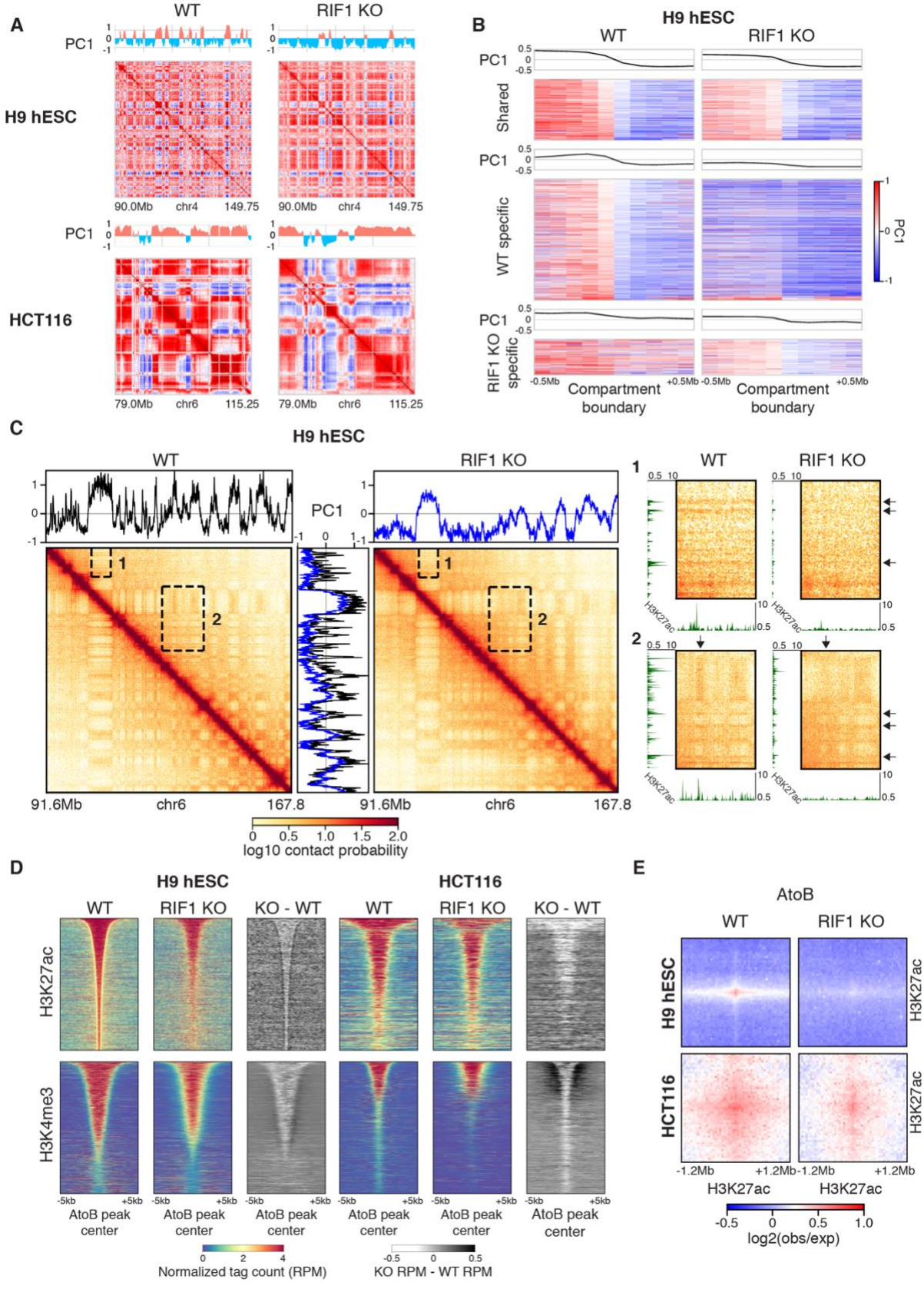

A
WT
RIF1 KO
PC1
H9 hESC
90.0Mb chr4 149.75
HCT116
79.0Mb chr6 115.25
B
H9 hESC
Shared
WT specific
RIF1 KO specific
Compartment boundary
C
H9 hESC
91.6Mb chr6 167.8
log10 contact probability
D
H9 hESC
HCT116
H3K27ac
H3K4me3
AtoB peak center
Normalized tag count (RPM)
KO RPM - WT RPM
E
AtoB
log2(obs/exp)

Fig. 3. RIF1 KO causes global alterations of compartments and epigenetic state.

(A) Correlation matrices and PC1 Eigenvector of Chr4 90.0-149.75Mb in WT and RIF1 KO H9 hESCs (top) and Chr6 79.0-115.25Mb in WT and RIF1 KO HCT116 (bottom). (B) Heatmaps of PC1 values centered on shared (top), WT specific (middle), and RIF1 KO specific (bottom) compartment boundaries ±0.5 Mb in WT and RIF1 KO H9 hESCs. (C) ICE normalized Hi-C contact map of Chr6 91.6-167.8 Mb in H9 hESC WT and RIF1 KO cells with accompanying PC1 Eigenvector plots (black: WT, blue: KO). To the right are expanded views of insets 1 and 2 with accompanying H3K27ac ChIP-seq plots. Arrows indicate compartments and ChIP-seq peaks that are lost upon RIF1 KO. (D) Normalized tag counts (RPM) signal and subtraction plots for H3K27ac and H3K4me3 centered on peaks within AtoB compartment switching regions ± 5kb in WT and RIF1 KO cells sorted by peak size. (E) Aggregate Hi-C log2(obs/exp) interactions between H3K27ac peaks within AtoB compartment switching regions ±1.2 Mb in WT and RIF1 KO cells.

Figure 4

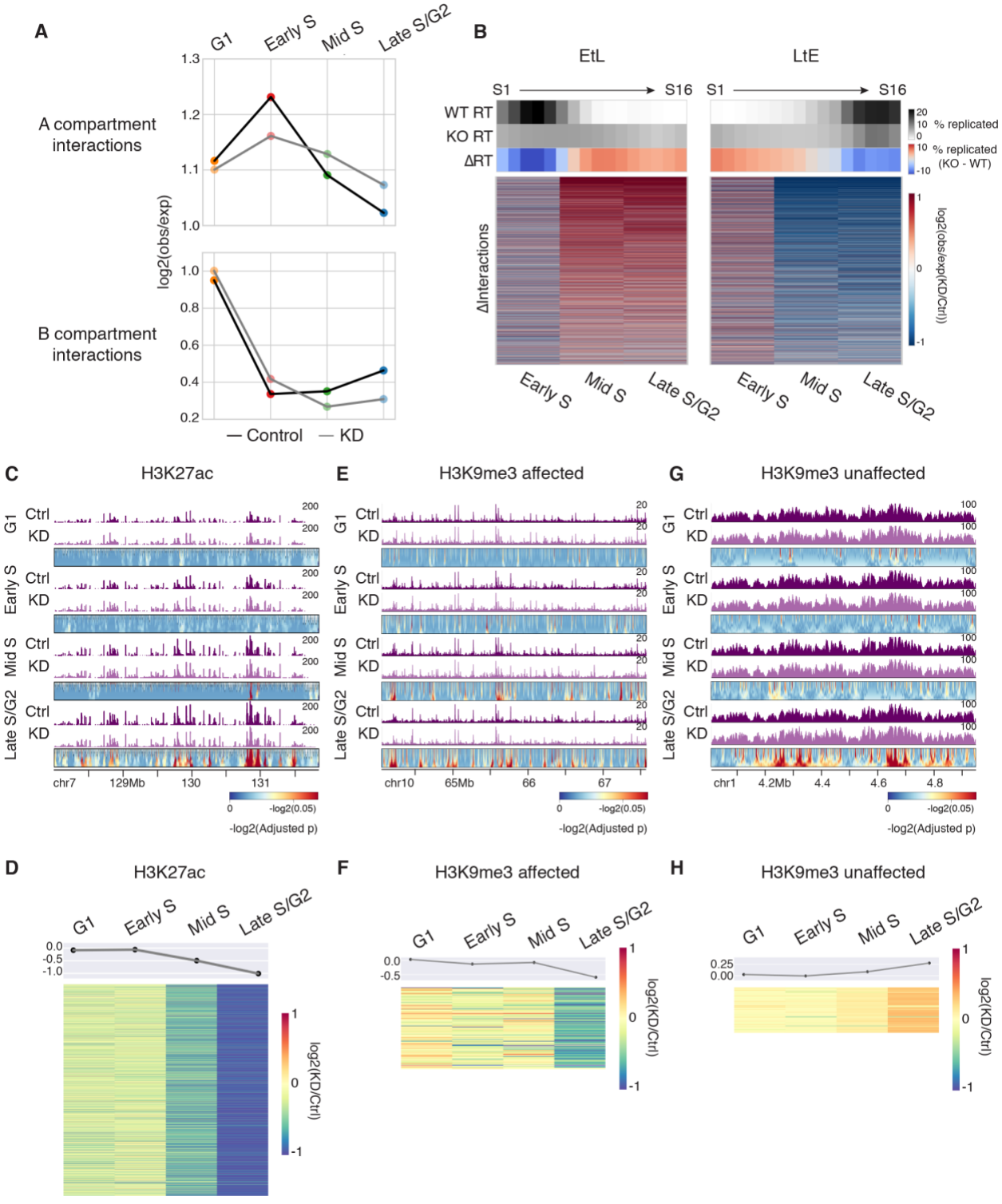

A
G1
Early S
Mid S
Late S/G2
A compartment interactions
B compartment interactions
log2(obs/exp)
— Control — KD
B
EtL
LtE
S1
S16
WT RT
KO RT
ΔRT
% replicated
% replicated (KO - WT)
ΔInteractions
log2(obs/exp(KD/Ctrl))
C
H3K27ac
E
H3K9me3 affected
G
H3K9me3 unaffected
Ctrl
KD
chr7
129Mb
130
131
chr10
65Mb
66
67
chr1
4.2Mb
4.4
4.6
4.8
-log2(0.05)
-log2(Adjusted p)
D
H3K27ac
F
H3K9me3 affected
H
H3K9me3 unaffected
log2(KD/Ctrl)

Fig. 4. Epigenome affects require DNA replication and occur during the first S phase after RIF1 degradation.

(A) Average log2(observed/expected) within top 10% A- (top panel) and B- (bottom panel) compartment in control (black line) and RIF1 KD (gray line) cells at G1, early S, middle S, and late S/G2 time points. (B) Interaction frequency differences between control and RIF1 KD at top 10% EtL and LtE regions at G1, early S, middle S and late S/G2 timepoints whose RT is indicated by average high resolution Repli-Seq heatmaps. (C) H3K27ac ChIP-seq of Chr7 128-131.9 Mb in control (top), RIF1 KD (middle), and domainogram (bottom) indicating -log2 (Benjamini-Hochberg adjusted p-values) calculated for the differences of control subtracted from KD tracks (Methods) at G1, early S, middle S, and late S/G2 time points. (D) Heatmap showing log2(KD/control) of H3K27ac peaks at G1, early S, middle S, and late S/G2 time points. (E) H3K9me3 ChIP-seq of Chr10 64.05-67.6Mb in control (top), RIF1 KD (middle), and domainogram (bottom) at G1, early S, middle S, and late S/G2 time points. (F) Heatmap showing log2(KD/control) of H3K9me3 peaks at affected regions at G1, early S, middle S, and late S/G2 time points. (G) H3K9me3 ChIP-seq of Chr1 4.05-5.0 Mb in control (top), RIF1 KD (middle), and domainogram (bottom) at G1, early S, middle S, and late S/G2 time points. (H) Heatmap showing log2(KD/control) of H3K9me3 domains at unaffected regions at G1, early S, middle S, and late S/G2 time points.

Figure 5

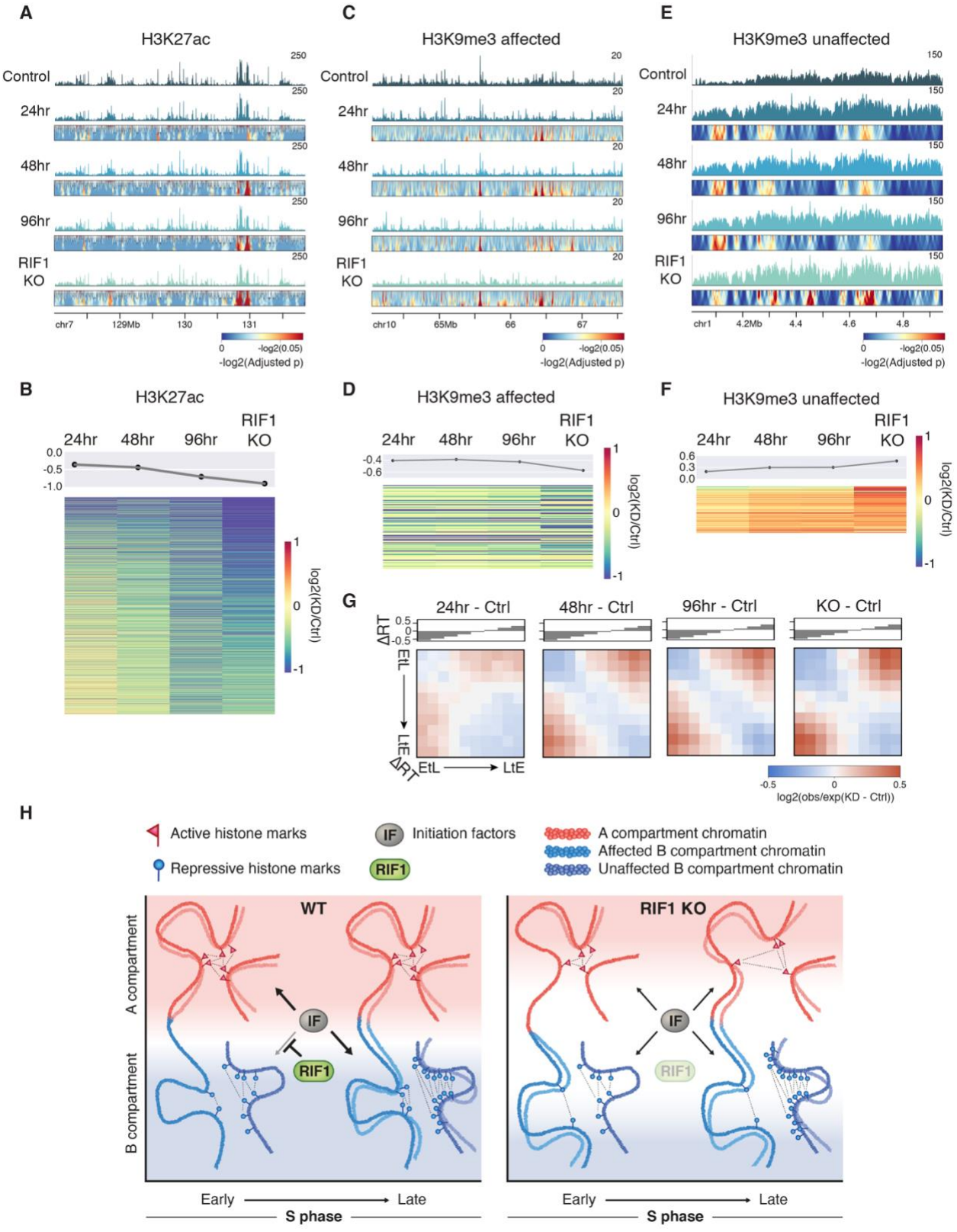

A
H3K27ac
Control
24hr
48hr
96hr
RIF1 KO
chr7
129Mb
130
131
-log2(Adjusted p)
C
H3K9me3 affected
chr10
65Mb
66
67
E
H3K9me3 unaffected
chr1
4.2Mb
4.4
4.6
4.8
B
H3K27ac
24hr 48hr 96hr RIF1 KO
log2(KD/Ctrl)
D
H3K9me3 affected
F
H3K9me3 unaffected
G
24hr - Ctrl
48hr - Ctrl
96hr - Ctrl
KO - Ctrl
ΔRT
EtL
LtE
log2(obs/exp(KD - Ctrl))
H
Active histone marks
Repressive histone marks
IF
Initiation factors
RIF1
A compartment chromatin
Affected B compartment chromatin
Unaffected B compartment chromatin
WT
RIF1 KO
A compartment
B compartment
Early
Late
S phase

Fig. 5. Epigenome affects are exacerbated as cell cycle without RT control.

(A) H3K27ac ChIP-seq of Chr7 128-131.9 Mb in control (top), RIF1 KD (middle), and domainograms (bottom) in control, 24hr, 48hr, 96hr degradation, and RIF1 KO. (B) Heatmap showing log2(KD/control) of H3K27ac tags at 24hr, 48hr, 96hr degradation, and RIF1 KO. (C) H3K9me3 ChIP-seq of Chr10 64.05-67.6Mb in control (top), RIF1 KD (middle), and domainograms (bottom) in control, 24hr, 48hr, 96hr degradation, and RIF1 KO. (D) Heatmap showing log2(KD/control) of H3K9me3 tags at affected regions in 24hr, 48hr, 96hr degradation, and RIF1 KO. (E) H3K9me3 ChIP-seq of Chr1 4.05-5.0 Mb in control (top), RIF1 KD (middle), and domainograms (bottom) in control, 24hr, 48hr, 96hr degradation, and RIF1 KO. (F) Heatmap showing log2(KD/control) of H3K9me3 tags at unaffected regions in 24hr, 48hr, 96hr degradation, and RIF1 KO. (G) Subtraction saddle plots of *cis* log2(observed/expected) contacts of control sample from 24hr, 48hr, 96hr degradation, and RIF1 KO sorted by ΔRT (KD RT - Control RT). (H) Model figure illustrating the role of RT in epigenome maintenance. In WT cells RIF1 prevents replication initiation factors from activating replication within repressive chromatin, allowing early replication of active chromatin which is then assembled with active histone marks (red flags) which self-interact to form the A compartment. Repressive chromatin is replicated in late S phase and is assembled with repressive histone marks (blue balls) which self-interact to form the B compartment. The B compartment is divided into regions that depend on RIF1 for late replication (affected) and those that do not (unaffected). RIF1 KO allows limiting replication initiation factors to associate heterogeneously with both active and repressive chromatin, delaying replication of the former and advancing replication of the latter. Delayed replication of active chromatin causes depletion of active histone marks and weakens interactions within the A compartment. Advanced replication of affected repressive chromatin causes depletion of repressive histone marks and weakens interactions within the affected B compartment hub. Unaffected B compartment hubs associated with repressive domains maintain their late replication in RIF1 KO and become enriched for repressive histone marks.

Science
AAAS

Supplementary Materials for

Replication timing maintains the global epigenetic state in human cells.

Kyle N. Klein[1]†, Peiyao A. Zhao[1]†, Xiaowen Lyu[2]†, Takayo Sasaki[1,3], Daniel A. Bartlett[1], Amar M. Singh[4], Ipek Tasan[5], Meng Zhang[5], Lotte P. Watts[6], Shin-ichiro Hiraga[6], Toyoaki Natsume[7], Xuemeng Zhou[8], Timour Baslan[9], Danny Leung[8], Masato T. Kanemaki[7], Anne D. Donaldson[6], Huimin Zhao[5], Stephen Dalton[4], Victor G. Corces[2], David M. Gilbert[1,3*].

Correspondence to: gilbert@bio.fsu.edu

This PDF file includes:

Other Supplementary Materials for this manuscript include the following:

aba5545_Suppl. Excel_seq1_v1

aba5545_Suppl. Excel_seq2_v1

aba5545_Suppl. Excel_seq3_v1

Materials and Methods

Cell lines: H9 hESCs were grown in feeder free conditions on Geltrex matrix (Thermo Fisher A14133) coated dishes in StemPro (Thermo Fisher A100701) media according to manufacturer's specifications. hESCs were passaged with ReLeSR (StemCell Technologies 05872) according to manufacturer's specifications. RIF1 KO H9 lines were established by transfecting two px330 CRISPR/Cas9 plasmids (addgene.com Plasmid #42230, Table S2) containing separate sgRNAs (Table S3) that would cut upstream and downstream of RIF1 exon 2. Cells were subcloned as in (35) and screened by PCR for loss of amplification from exon 2. HCT116 cells were grown in McCoy's 5A media plus 10% FBS and 1% Pen/Strep and passaged by trypsinization. HCT116 lines were established by CRISPR/Cas9 knock in of eGFP-RIF1-FLOX construct (Table S4) and blasticidin selection. Selected cells were dissociated to single cells and diluted to 100 cells per 10cm plate and allowed to grow into colonies for 1-2 weeks. Clones with homozygous integration of the eGFP-RIF1-FLOX construct were screened by PCR. RIF1 was deleted from HCT116 eGFP-RIF1-FLOX cell lines by transient transfection with pAAV-Ef1a-mCherry-IRES-Cre (addgene.com Plasmid #55632) and FACS sorting of mCherry positive cells. Positive cells were dissociated to single cells and diluted to 100 cells per 10cm plate and allowed to grow into colonies for 1-2 weeks. Clones with homozygous deletion of RIF1 exon3 were screened by PCR. RIF1-mAID-mClover HCT116 cells were established as in (*12*). HAP1 WT and RIF1 KO cells were purchased from Horizon Discovery (HZGHC000663c010) and grown in IMDM plus 10% FBS and 1% Pen/Strep and passaged by trypsinization according to manufacturer's instructions.

HCT116 RIF1-AID synchronization and degradation: For the first cell cycle time course ChIP-seq and Hi-C experiments, RIF1-mAID-mClover HCT116 cells were synchronized in G1, early, middle, and late S phase. To synchronize in G1 cells were incubated with 20μM lovastatin (Mevinolin; LKT Labs M1687) for 24 hours. Cells were synchronized at early S phase by washing away lovastatin with 2x PBS washes and releasing cells into fresh media plus 2mM Mevalonic acid (Mevalonolactone; Sigma-Aldrich M4667) and 20μM thymidine (Sigma-Aldrich T9250) and incubating for 16 hours. Cells were synchronized at mid and late S/G2 by washing away thymidine with 2x PBS washes and releasing into fresh media for 3.5 hours and 5.75 hours respectively. For all first cell cycle time course samples, RIF1 was degraded by adding 2ug/mL doxycycline and 500 uM IAA (Sigma-Aldrich I2886) to G1 synchronization media 18 hours after initial synchronization allowing for 6 hours of RIF1 degradation before release. G1 synchronized cells remained in lovastatin for 12 hours following RIF1 degradation before collection. Doxycycline and IAA were maintained in the media during further synchronization and release steps for all RIF1 degradation samples. For the first cell cycle Repli-seq, cells were synchronized in G1 and RIF1 was degraded as above. Cells were then released from the G1 block for 10, 14, 18, or 22 hours and labeled and fixed for E/L Repli-seq (*14*). Equal numbers of cells from each release time point were pooled together for standard E/L Repli-seq sorting and processing (*14*). For all first cell cycle experiments, control cells were synchronized with the same protocol, but doxycycline and IAA were omitted from the media. For extended time course, RIF1 was degraded by adding 2ug/mL doxycycline and 500 uM IAA to media and maintained in the media for 24, 48, or 96 hours before cells were collected. RIF1 degradation was confirmed by microscopy and western blotting.

E/L Repli-seq and analysis: Genome-wide RT profiles were generated as previously described (*15*) with the following modifications: RT datasets were not normalized by quantile

normalization since this processing step will result in the enforcement of WT distribution on KO, thereby preventing the detection of any genome-wide RT changes. The Loess smoothed coverages are scaled using robust scaler from the python package scikit-learn (http://scikit-learn.org) instead. The scaler removes median and scales the datasets to their collective inter-quantile range according to:

$$xscaled=(x-median(x))/x.quantile(0.75)-x.quantile(0.25)$$

(Equation I)

As a result, different datasets with varying dynamic ranges directly comparable to each other. Sex chromosomes were excluded from Repli-seq analysis.

Replication foci staining: Asynchronous cells were grown on coverslips until 70% confluence and pulsed with BrdU for 30 minutes. Cells were then washed with PBS and fixed in cold 70% EtOH. Chromatin was denatured with 1.5 N HCl 30min RT and washed away. Cells were permeabilized with PBS plus 0.5% Tween20 for 5min RT. Mouse anti-BrdU (BD 55567) (1:50) in PBS plus 10% goat serum was added for 1hr at RT and washed away. Goat 488 anti-mouse (1:100) in PBS plus 10% goat serum was added for 1hr at RT and washed away. Nuclei were stained with DAP1 and cells were imaged on the DeltaVision (GE Life Sciences) microscope. Patterns were categorized as early, middle, late, or blended as in Fig S3A.

High resolution Repli-seq profiling and analysis: High resolution Repli-seq was performed and analyzed as previously described (*10*). Briefly, asynchronously growing cells were labeled with BrdU (400 uM) for 30 minutes followed by ethanol fixation. PI staining for DNA content was carried out and 16 fractions of S phase were sorted on the FACS. Gates were set by marking G1 and G2 peaks and dividing S phase into 16 equal fractions between the peaks. At least 80k cells were sorted for each S phase fraction. BrdU incorporated mouse mitochondrial DNA and BrdU negative mouse DNA (as used in *10*) was spiked into the purified genomic DNA to serve as a BrdU immunoprecipitation control. BrdU incorporated DNA was immunoprecipitated using mouse anti-BrdU (BD 55567). Libraries were prepared for Illumina sequencing with NEBNext Ultra DNA Library Prep Kit for Illumina (New England Biolabs E7370). The reads were aligned to hg38 using bowtie2 with parameters --no-mixed --no-discordant --reorder. Replication heatmap matrices were constructed from RPM (read per million) bedgraph files at 50kb window size ranked from S1 to S16. The matrices were Gaussian smoothed with lambda=1 and column-wise scaled. Sex chromosomes were excluded from the heatmap matrix construction.

Calculating RT indices from high resolution Repli-seq: For each 50kb genomic bin the RT indices computed according to the equation below and graphically illustrated in Supplementary Figure 4b:

$$RTindex=log2(Sum(S1*8,S2*7,S3*6,...S8*1)/Sum(S9*1,S10*2,S11*3...S16*8))$$

(Equation II)

RTdiff is computed by subtracting RTindexKO from RTindexWT.

Calling unaffected and affected regions using high resolution Repli-seq: We fitted a Gaussian Mixture model probability distribution on RTdiff between WT and KO calculated for all genomic bins. The model is composed of three distribution components: one that contains large negative values representing the bins that on a population level replicate later in WT and earlier in KO, a second distribution of positive values representing bins that on a population level replicate earlier in WT and later in KO, a third one that contains values close to zero representing the bins that on a population level replicate at times statistically indistinguishable from WT. The validity of the model is checked through the minimization of Bayesian information criterion. Continuous 50kb genomic bins that were associated with negative RT indices in WT and assigned to the first distribution were called as affected late regions. Continuous 50kb genomic bins that were associated with negative RT indices in WT but assigned to the third distribution were called as unaffected late regions.

Twidth from high resolution Repli-seq: The degree of heterogeneity of RT was estimated by performing a sigmoidal fitting on the column wise cumulative replication percentage as previously described (*10*). Briefly, the sigmoidal function below:

$$f(x) = \frac{e^{-kx}}{1 + e^{-kx}}$$

(Equation III)

was fitted to bin-wise cumulative replication percentage using curve_fit function in scipy, which aims to minimize the mean squared error. Tolerance is set at 0.0001. Trep used in the timing-variation measurement is f(x) when x = 0.5 which means the genomic bin is 50% replicated in the cell population and Twidth used in the timing-variation measurement is f(0.75) – f(0.25) which is the time difference between 75% replicated and 25% replicated for any genomic bin.

Single cell Repli-seq: As previously described (*11, 16*) with slight modifications. Briefly, single cells were sorted from five evenly spaced windows throughout S phase directly into individual wells of a 96 well plate containing single cell lysis buffer and lysed at 50°C for 1 hour followed by 4 minutes at 99°C. Whole genome amplification (WGA) was performed as previously described (*16*) with minor modifications, namely the oligonucleotides used in the Degenerate Oligonucleotide Priming PCR (DOP-PCR) were modified to introduce a unique index in each WGA reaction. Subsequent to WGA, each individual WGA DNA product was purified and pooled for standard NEBNext Ultra DNA Library preparation.

Single cell Repli-seq analysis: Single cell Repli-seq fastqs were first demultiplexed according to the barcodes added during library making using custom script. The reads were trimmed 100bp from the 5' end using CutAdapt (https://github.com/marcelm/cutadapt) and subsequently aligned to hg38 using bowtie2 with the parameters --no-mixed --no-discordant --reorder. Cells with fewer than 25,000 reads aligned were filtered out. Subsequent analysis was carried out as previously described (*11*). Briefly, mappability correction using G1 cells and smoothing were

applied, RPM was calculated in 50kb bins for single cells. Binarization was subsequently carried out on smoothed RPM signal files by applying a 2-component mixed Gaussian model where bins assigned to the distribution with lower mean RPM were binarized to 0 (unreplicated) and those that were assigned to the distribution with higher mean RPM were binarized to 1 (replicated). Heatmaps were generated by sorting according to copy number with the cell with the fewest bins replicated at the top and that with the most bins replicated at the bottom. Twidth calculation was performed as previously described (*11*). Briefly, for individual bins, an index called 'time from population average replication' was calculated, which represents the time in hours passed from the time of replication indicated from E/L repli-seq. A negative number indicates that according to E/L repli-seq the genomic bin should not have undergone replication on a population level whereas a positive number indicates that according to E/L repli-seq the bin should have finished replicating. The 2-D distributions of such indices against % replicated were plotted as heatmaps. A sigmoidal curve was fitted to the distribution as described above in 'Twidth from high resolution Repli-Seq'.

ChIP-seq experiments and analyses: ChIP-seq experiments in HCT116 cells and H9 hESCs were carried out as previously described (*17*). For all H3K27ac and H3K9me3 ChIP-seq experiments, Drosophila S2 cells were used as spike-in for pull down normalization at a 1:2 ratio to human cells as previously described (*18*). All H3K27me3 and H3K4me3 ChIP-seq were performed in the absence of spike-in. After removal of medium, cells were crosslinked in 1% formaldehyde in PBS at room temperature for 10 min and quenched with glycine. PBS rinsed cell pellets were flash frozen in liquid nitrogen and stored at -80°C or continue with cell and nuclear lysis steps. Nuclear lysates were precleared with protein A/G beads followed by incubation with proper antibodies (Table S5). After washing with high salt buffer, LiCl buffer, and TE, chromatin was eluted and reverse crosslinked. Purified DNA was ethanol precipitated followed by Illumina Truseq library preparation. Libraries for Illumina sequencing were constructed using the following standard protocol. Fragment ends were repaired using the NEBNext End Repair Module and adenosine was added at the 3' ends of fragments using Klenow fragment (3' to 5' exo minus, New England Biolabs), universal adaptors were ligated to the A-tailed DNA fragments at room temperature for 1 h with T4 DNA ligase (New England Biolabs) and amplified with Illumina barcoded primers using KAPA SYBR FAST qPCR Master Mix for 5~12 PCR cycles to obtain enough DNA for sequencing. Generated libraries were paired-end sequenced on Illumina HiSeq2500 v4. ChIP-seq datasets were aligned to hg38 using bowtie-2 with the parameters --no-mixed --no-discordant --reorder. Resulting bam files for each histone mark were sorted and deduplicated using samtools. Deduplicated bam files for ChIPed and input libraries were passed onto the peak calling algorithm MACS2 for peak calling and generating fold enrichment signal tracks. For spike-in ChIP-seq datasets, fastq reads were aligned to dm6 in addition to hg38 using the same parameters as stated above. The numbers of aligned dm6 reads were used to normalise between spike-in ChIP-seq samples. Let Nd-i be the number of aligned dm6 reads for the i-th ChIP-seq sample and Nd-min be the lowest number of aligned dm6 reads among all ChIP-seq samples. Reads aligned to hg38 for the i-th ChIP-seq sample are normalised by factor α where α is Nd-i / Nd-min. The heatmaps centered at features were constructed by filling a matrix where the rows represent ChIP-seq signals around individual features.

RIF1 Cut&Run: Cut&Run was performed as previously described (*19*) with antibody against eGFP (Table S5) in eGFP-RIF1 tagged H9 hESCs and HCT116. Briefly, cells were washed and bound to Concanavalin-A-coated magnetic beads, and permeabilized in wash buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM spermidine and one Roche Complete protein inhibitor tablet per 50 mL) containing 0.03% digitonin and 20 mM EDTA (Dig-wash). Antibody was added to a final concentration of 1:100 and incubated overnight at 4°C. Cells were washed with Dig-wash buffer and incubated with Protein-A-MNase (pA-MN) for 1 hour at 4°C. Cells were washed three times with Dig-wash buffer to remove unbound pA-MN before placing cells in an ice-cold block and activating cleavage with the addition of CaCl2 to 2mM final concentration. Cleavage was stopped by the addition of 2x Stop Buffer (340 mM NaCl, 20 mM EDTA, 4 mM EGTA, 0.05% Digitonin, 0.05 mg/mL glycogen, 5 µg/mL RNase A), and fragments were released by 30 min incubation at 37°C. Supernatant was recovered and DNA was purified with phenol:chloroform extraction and precipitation with ethanol, before being used as input for library preparation.

Calling H3K9me3 domains: In addition to applying MACS2 (*20*) and epic2 (*21*) on H3K9me3 ChIP-seq datasets to call peaks, we also called large H3K9me3 domains. H3K9me3 fold enrichment was calculated as a log2 ratio over input in 50kb genomic bins. Considering any genomic bin can assume one of two states: inside or outside of a H3K9me3 domain, a two-state hidden markov model was fitted to the fold enrichment distribution with tolerance set at 0.0001. Implementation was carried out in Python using the hmmlearn package (https://github.com/hmmlearn/hmmlearn). Each genomic bin was assigned a state using the Viterbi algorithm. The state associated with the higher signal mean was inside an H3K9me3 domain. 'H3K9me3 domain' was therefore defined as continuous genomic bins (>500kb in H9 hESCs and >1mb in HCT116) associated with the state of higher mean signal.

H3K9me3 knockdown: Depletion of H3K9me3 from HCT116 RIF1 KO cells was performed similar to (*22*). On-Target SMART pool siRNA SUV39H1 (Horizon Dharmacon L-009604-00-005), On-Target SMART pool siRNA SUV39H2 (Horizon Dharmacon L-008512-00-005), On-Target SMART pool siRNA SETDB1 (Horizon Dharmacon L-020070-00-005), and negative control On-Target plus Non-targeting pool (Horizon Dharmacon D-001810-10-005) siRNAs were diluted in siRNA buffer (Horizon Dharmacon B-002000-UB-100) at 20 µM stock solutions. HCT116 RIF1 KO cells were plated in 10 cm dishes at a confluence of 30% and 24 hours later media was refreshed, and cells were transfected with 30 nM siRNAs against SUV39H1, SUV39H2, and SETDB1 or Non-targeting control using OptiMEM (Invitrogen 31985070) and Lipofectamine RNAi:MAX reagent (Invitrogen 13778075) according to manufacturer's instructions. 72 hours after the initial transfection a second siRNA transfection was performed on the same cells using the same protocol. 48 hours after the second siRNA transfection cells were labeled and collected for E/L Repli-seq and Western blot.

Hi-C procedure and sequence processing: Two subclones of H9 and HCT116 RIF1 KO cells were prepared and processed for Hi-C as previously described (*23*) using DpnII for digestion. ~1 billion 50bp paired-end reads were obtained for each subclone for both H9 and HCT116. Reads were processed using HiCUP pipeline available from Babraham Institute (https://www.bioinformatics.babraham.ac.uk). Briefly the reads were truncated at DpnII recognition sites and mapped to hg38 using bowtie2. The uniquely mapped reads were further

filtered to remove common Hi-C artefacts e.g. self-ligated fragments and same frag and subsequently deduplicated. Bam files were converted to cooler files using 4DN DCIC utility bam2pairs and cooler load at 5kb, 50kb, 100kb and 250kb. Hi-C maps used in this work were either normalized using iterative correction (ICE) (*24*) or distance normalized as log2 (observed/expected) where expected is computed according to the equation below:

$$\frac{\sum_i^j diagonal(|i-j|)}{(|i-j|)}$$

(Equation IV)

Eigendecomposition is performed using cooltools package and eigenvectors are ranked using gene density.

First S phase HCT116 RIF1-AID ChIP-seq and Hi-C: RIF1 degraded cells struggled to transition from early S phase to middle and late S phase, leaving ~20% more cells at the G1/S transition than control (Fig S15E). To address this discrepancy, we computationally combined G1 reads with middle and late S phase ChIP-seq and Hi-C reads in control samples for all analyses so that G1 reads constituted 20% of the final total read count for these samples. This step ensured that the epigenetic and structural changes we called in RIF1 degraded cells were not due to G1 cell contamination relative to control cells.

Domainograms showing adjusted p-values for tag count difference between KD ChIP-seq and control ChIP-seq: The subtraction track of control ChIP-seq tag count track from that of KD ChIP-seq was fitted using distfit (https://github.com/erdogant/distfit) for common univariant probability distributions and the closest distribution was selected as one with minimal residual sum of squares. P-values calculated using the cumulative density function of the chosen probability distribution were further corrected using the Benjamini-Hochberg approach implemented through the python packages statsmodel. To generate domainograms, subtraction tracks were calculated at 1, 2.5, 5, 25, 30, 40 and 50 kb bin sizes.

TAD calling: TADs and insulating boundaries in this work were called using DI (*25*) and insulation score (*26*) on raw and ICE normalized contact matrices respectively at 50kb bins with a sliding window of 1Mb.

Calling differential interacting pairs: Differential interacting pairs were called using diffHiC (*27*). Briefly, read counts matrices were generated from bam files resulting from HiCUP for WT and RIF1 KO. Subsequently, edgeR (*28*) was applied to count matrices to estimate technical noise between replicates and call differential pairs by quasi-F test through generalized linear model (glm) fitting. Pairs with ≥2-fold difference and <0.05 FDR were called to be up- or down-regulated depending on the sign of fold change.

RNA-seq: Libraries were prepared with NEBNEXT rRNA Depletion kit (human/mouse/rat) (New England Biolabs E6310) and NEBNEXT Ultra II Directional RNA Library Prep kit for Illumina (New England Biolabs E7760) according to manufacturer's instructions. Sample input was 900ng total RNA (determined by Qubit RNA HS reagents, ThermoFisher Q32852) with a RIN >8 (TapeStation High Sensitivity RNA ScreenTape Assay, Agilent 5067-5579). RNA was fragmented for 15 minutes and libraries were constructed with a 1/5th dilution of NEB adaptor and 12 cycles of PCR amplification with dual-indexing primers. Amplified libraries were initially quantified by Qubit DNA HS reagents (ThermoFisher Q32851) and checked for size and artifacts using TapeStation DNA HS reagents (Agilent 5067-5584). Excess adaptors were removed using an additional 0.9x AMPure bead purification. KAPA qPCR (KAPA Biosystems) was used to determine molar quantities of each library and individual libraries were diluted and pooled at equimolar concentrations. The final library pool was again checked by TapeStation and KAPA qPCR before submission for sequencing. Fastq reads were aligned to hg38 using STAR aligner with the options --quantMode GeneCounts and --bamRemoveDuplicatesType UniqueIdentical turned on. Gene count files were generated using HTseq (https://github.com/simon-anders/htseq) from bam files produced by STAR aligner. DEseq2 (https://bioconductor.org/packages/release/bioc/html/DESeq2.html) was subsequently used to call differentially expressed transcripts using gene count files as input. Differential expression events are defined as q-value < 0.1.

For HERV analysis of RNA-seq datasets, a custom alignment approach was used to generate RPKM for each HERV subfamily with both uniquely aligned and multi-aligned reads (*29*). Briefly, sequencing reads were aligned to human reference genome hg19 using STAR with setting --outFilterMultimapNmax as 100 (*30*). The output BAM file was subsequently separated into files containing uniquely aligned and multi-aligned reads by SAMtools (*31*), followed by conversion into BED files. The resulted uniquely aligned and multi-aligned reads were overlapped with the coordinates of repetitive elements, extracted from Repeatmasker definitions from UCSC using BEDTools (*32*). For multi-aligned reads, only those uniquely "anchored" to a repetitive subfamily were included for further analysis. The fragments Per Kilobase per Million reads (RPKM) of each repetitive subfamily and its fold change in RIF1 KO versus WT cells were calculated from the aligned reads. T-test was used to measure the significance.

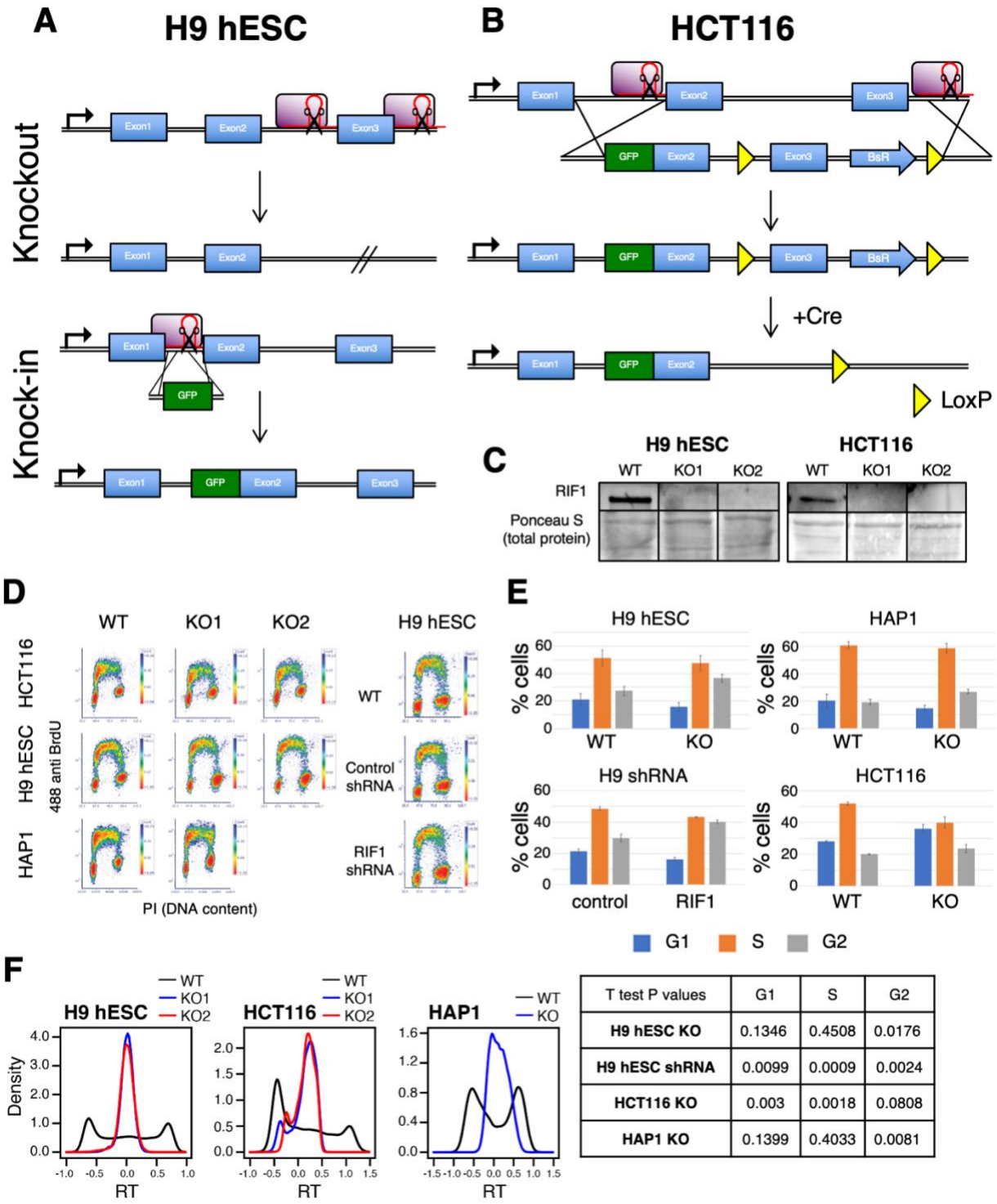
Single cell RNA-seq: HCT116 and HAP1 cells were harvested with 0.05% trypsin-0.442 nM EDTA and H9 hESCs were harvested with Gentle Cell Dissociation Reagent (StemCell Technologies 07174). Single cell suspension was adjusted to 1000 cells/ uL for generating single cell RNA-seq libraries using 10X Genomics Chromium Next GEM Single Cell 3' Reagent Kits v3.1 per manufacturer's instruction. Cell ranger 3.1.0 developed by 10X Genomics was subsequently applied to the fastq reads. Briefly, reads were aligned using STAR aligner to GRCh38 transcriptome and assigned to cells by barcodes. Cells with ≥ 10% mitochondrial reads were filtered out. Subsequently, single-cell count matrices where rows represent genes and columns represent individual cells were used as input for Seurat R package to perform tSNE dimension reduction and variable feature selection.

Random Forest Classifier: Random forest classification for differentially expressed genes was performed on arrays where each row represented the TSS of a differentially expressed gene and

each column represented the RPM signal of H3K27ac, H3K4me3, H3K27me3 and H3K9me3 according to the algorithm presented in (*33*) using scikit-learn.

Multiple regression coefficients for describing the correlation between histone marks and Hi-C PC1: Ordinary Least Squares Regression was fitted to describe the relationship between histone marks and Hi-C PC1 using the equation $Y \sim aX1+bX2+cX3\ldots+zXn+k$ where Y denotes Hi-C PC1 values, X(1…n) denote histone marks of interest, a…z denote regression coefficients and k denotes the unknown intercept. The implementation was carried out using python module statsmodels (www.statsmodels.org).

Fig. S1.



T test P values	G1	S	G2
H9 hESC KO	0.1346	0.4508	0.0176
H9 hESC shRNA	0.0099	0.0009	0.0024
HCT116 KO	0.003	0.0018	0.0808
HAP1 KO	0.1399	0.4033	0.0081

Fig. S1. CRISPR/Cas9 RIF1 KO does not cause cell cycle arrest.
(A) Schematic representation of RIF1 KO strategy: RIF1 KO in H9 hESCs was achieved by CRISPR/Cas9 cutting at regions up and down stream of exon 3 in the RIF1 coding sequence (see Table S4 for all sgRNA sequences). Knock in of an eGFP tag was achieved by CRISPR/Cas9 cutting at a region near the start codon and providing a repair vector containing the eGFP coding sequence and flanking homology arms. (B) RIF1 GFP tagging in HCT116 was achieved by CRISPR/Cas9 cutting at regions upstream of exon 2 and downstream of exon 3 and providing a repair vector that contained an eGFP tagged version of exon 2 and a loxP flanked exon 3 as well as a blasticidin resistance gene. Cre expression was then used to remove exon 3. In both cell lines removal of exon 3 caused a premature stop codon to be added to the RIF1 mRNA and thus a nonfunctional, truncated protein produced. (C) Immunoblot of RIF1 protein in WT and two KO clones of HCT116 and H9 hESCs. Ponceau S total protein stain shown as loading control. (D) 2-dimensional (2D) cell cycle analysis using FACS to detect incorporation of 488 anti BrdU (y axis) and DNA content by propidium iodide (x axis) in WT and RIF1 KO cells. (E) Bar plots showing mean cell cycle quantification from 2D FACS analyses of three replicates for each cell line with standard deviation error bars; table of p values by t-test between comparators for each cell line. (F) Genome wide probability density of E/L RT values in H9 hESCs (left), HCT116 (middle), and HAP1 (right) WT (black) and RIF1 KO (blue and red).

Fig. S2.

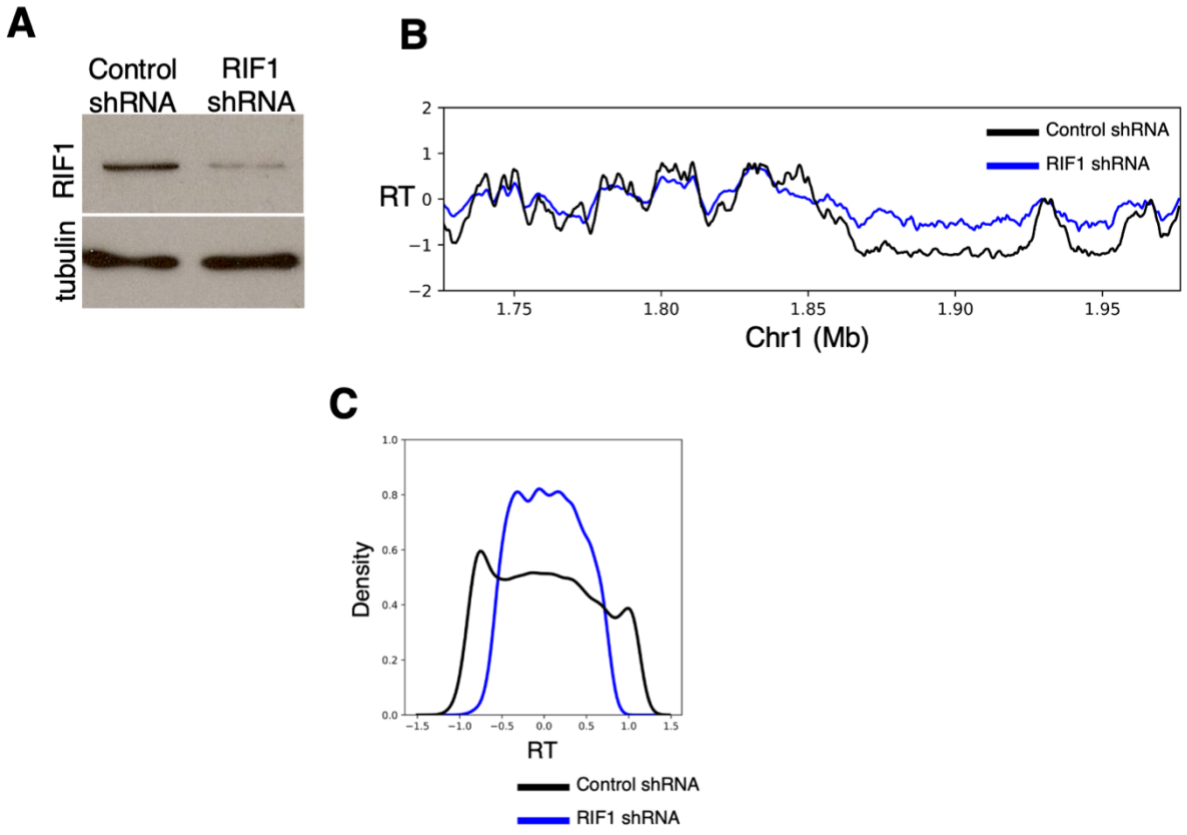
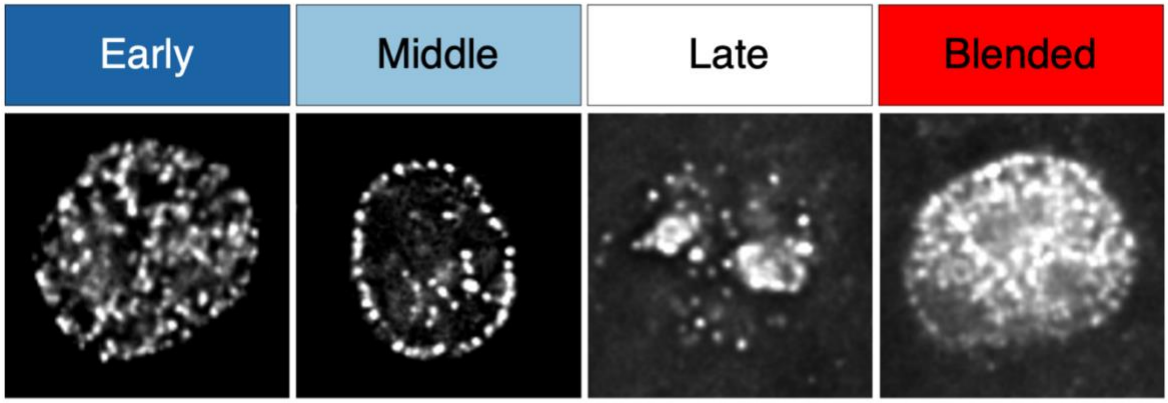


Fig. S2. RIF1 KD in H9 hESCs causes intermediate RT changes.
(A) Immunoblot of RIF1 protein in control and RIF1 shRNA in H9 hESCs. Tubulin used as loading control. (B) E/L Repli-seq data of Chr1 172.6-197.6 Mb in control shRNA (black) and RIF1 shRNA (blue) H9 hESCs. Same region as Fig 1a. (C) Genome wide probability density of E/L Repli-seq RT values in control (black) and RIF1 shRNA (blue) H9 hESCs.

Fig. S3.

A


Early
Middle
Late
Blended

B

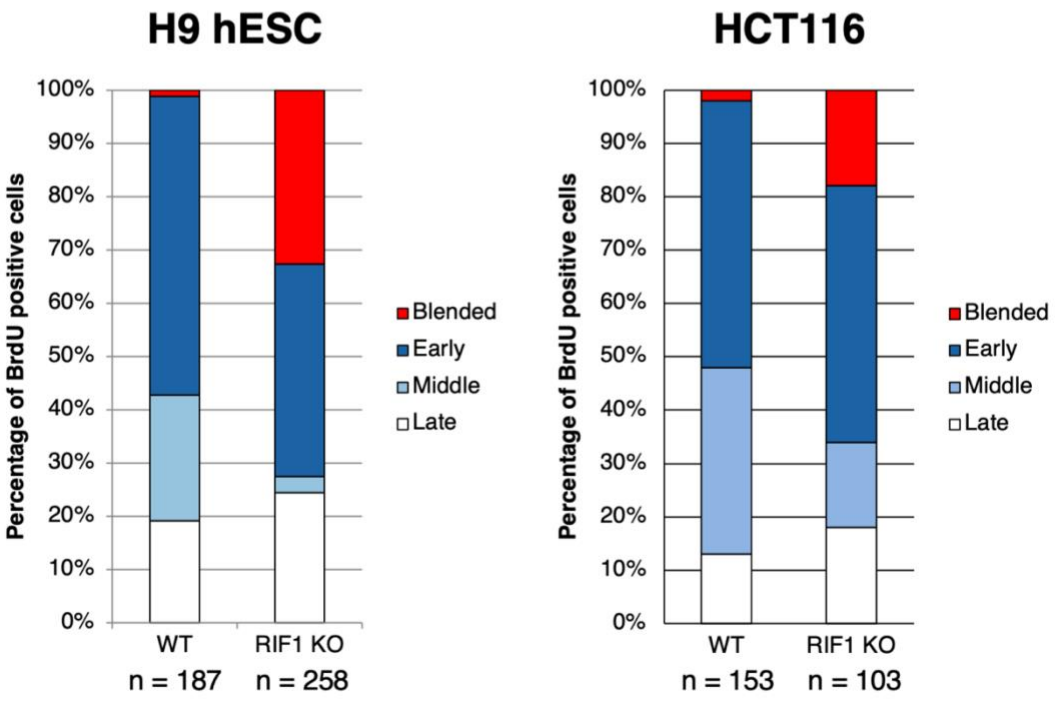
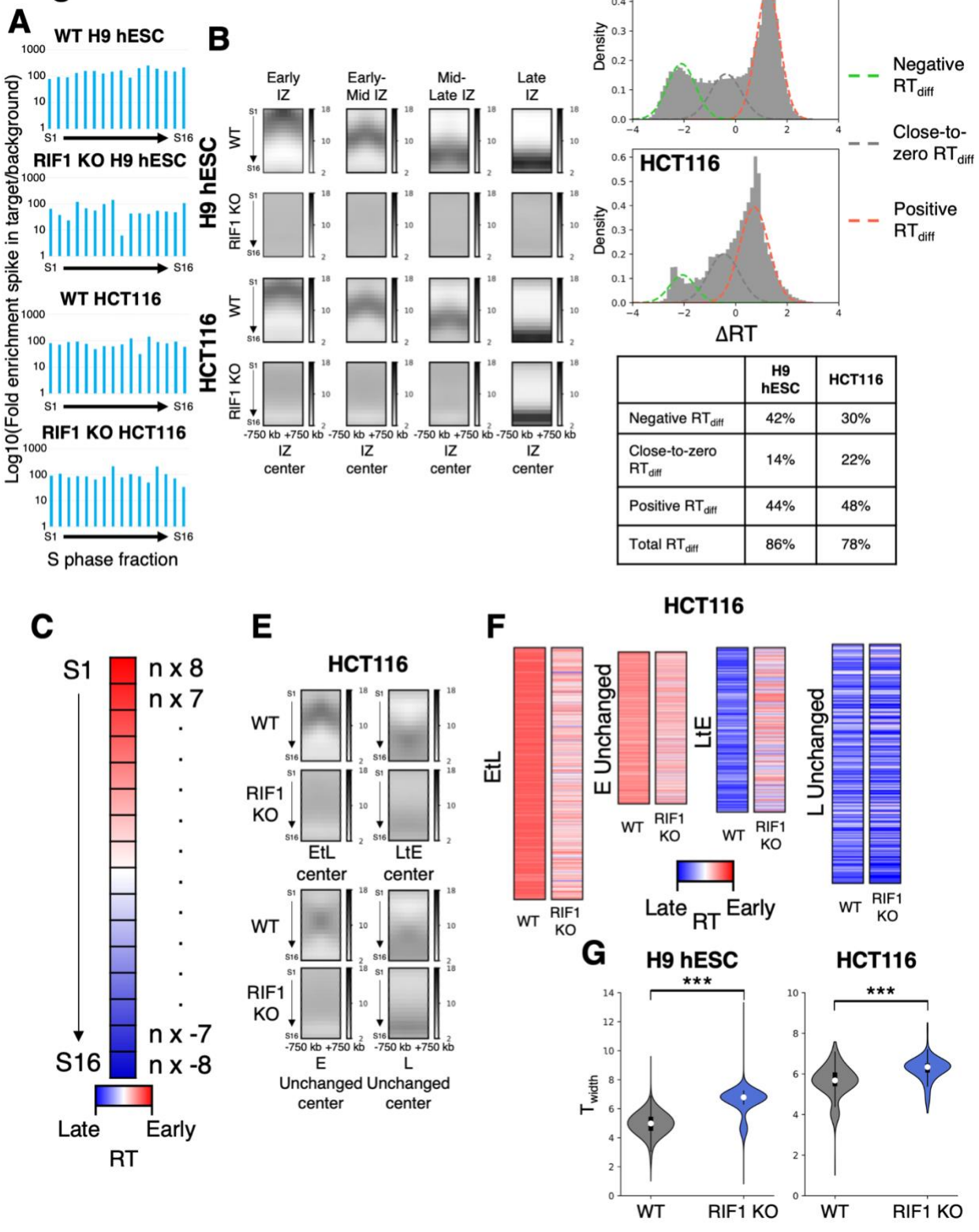

H9 hESC
HCT116
Percentage of BrdU positive cells
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
WT
RIF1 KO
n = 187
n = 258
n = 153
n = 103
Blended
Early
Middle
Late

Fig. S3. Replication foci patterns are disrupted in RIF1 KO cells.
(A) Example BrdU incorporation patterns of S phase human nuclei. (B) Quantification of percentage of S phase BrdU incorporation patterns in WT and RIF1 KO H9 hESCs and HCT116. Blended patterns are likely due to replication of normally early and middle replicating chromatin at the same time causing the reduced percentage of middle S phase patterns in RIF1 KO cells. Late patterns maintained in RIF1 KO H9 hESCs are likely the result of highly heterochromatic telomeres and centromeres.

Fig. S4.



	H9 hESC	HCT116
Negative RT_{diff}	42%	30%
Close-to-zero RT_{diff}	14%	22%
Positive RT_{diff}	44%	48%
Total RT_{diff}	86%	78%

Fig. S4. High resolution and single cell Repli-seq reveal major RT heterogeneity in RIF1 KO cells.
(A) Log10 fold enrichment qPCR results of BrdU incorporated mouse DNA spike in target over BrdU negative mouse DNA spike in background (see Methods) after BrdU immunoprecipitation of each S phase fraction in WT and RIF1 KO H9 hESCs and HCT116 cells to approximate pull-down efficiency. WT and RIF1 KO of both cell lines show consistent and similar pull-down efficiencies. (B) High resolution Repli-seq pile up plots in WT and RIF1 KO H9 hESCs and HCT116 centered on IZ timing categories called in WT cells ±750 Kb. (C) Schematic representation of RT indices calculation method used in Fig S4D (see Methods). (D) Genome wide fitted multivariant Gaussian probability density plots of RT differences ($RT_{diff,}$ See Methods) between WT and KO for each cell line calculated from high resolution Repli-seq by subtracting RIF1 KO RT indices from WT RT indices. Dotted lines indicate three fitted distributions: negative RT_{diff} (green), close-to-zero RTdiff (grey), or positive RTdiff (red). Negative RT_{diff} indicates regions that are later replicating in WT and earlier replicating in KO, positive RT_{diff}, indicates regions that are earlier replicating in WT and later replicating in KO, and close-to-zero RT_{diff} indicates regions that showed limited RT difference between WT and KO or could not be called as significantly different. Table shows the percentage of the genome within each category. (E) High resolution Repli-seq pile ups centered on RT changed regions called in E/L Repli-seq data (top) and RT unchanged regions called in E/L Repli-seq (bottom) in WT and RIF1 KO HCT116 ±750 Kb. (F) RT indices calculated from high resolution Repli-seq for RT changed regions (top) and RT unchanged regions (bottom) called in E/L Repli-seq in WT and RIF1 KO HCT116. Early replicating regions that were not called as EtL switches (E Unchanged) in the HCT116 E/L Repli-seq showed noticeable changes in their RT indices similar to called EtL switches. (G) Violin plots of genome wide T_{width} of high resolution Repli-seq data in WT and RIF1 KO H9 hESCs and HCT116.

Fig. S5.

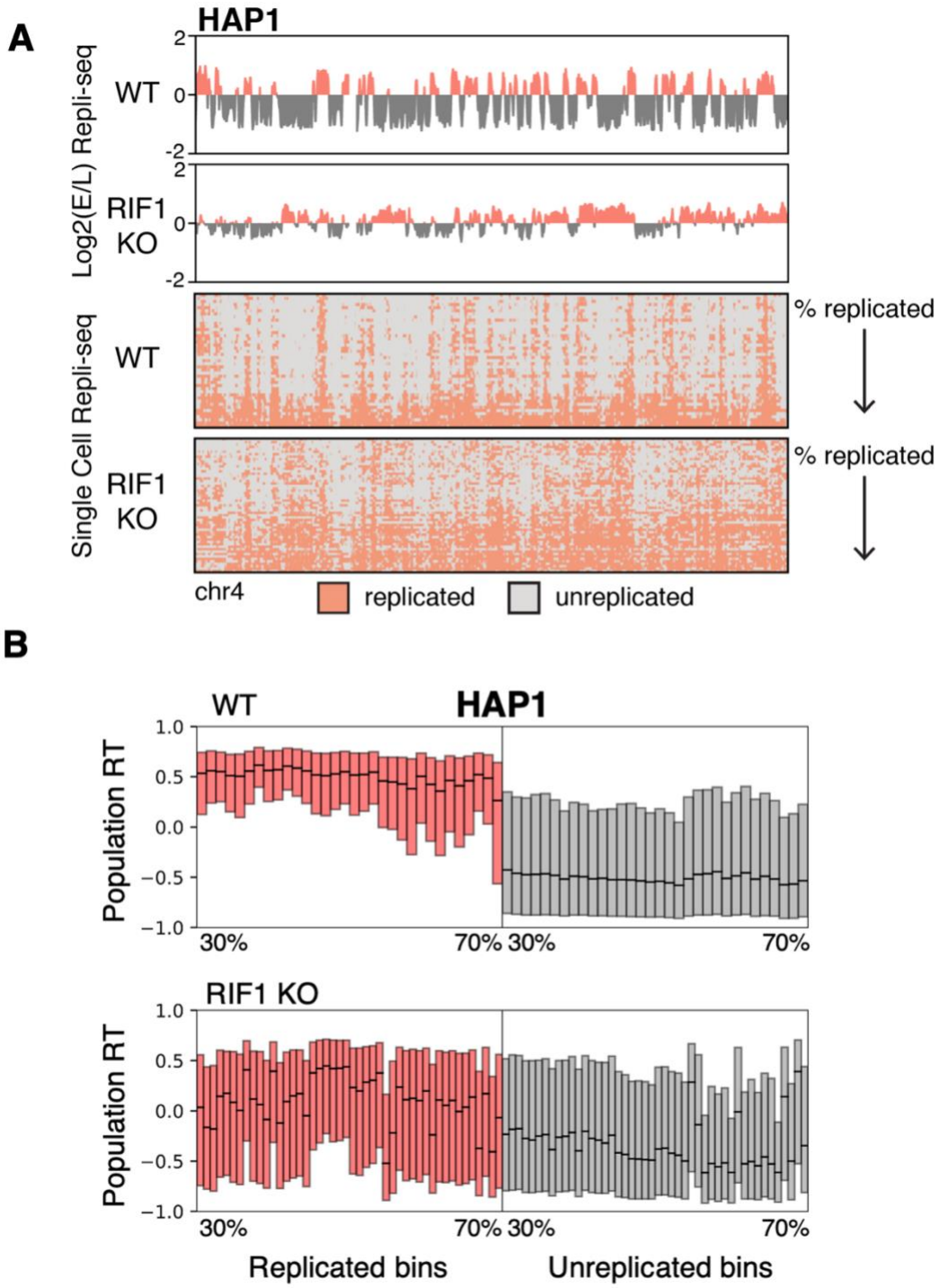

A
HAP1
Log2(E/L) Repli-seq
WT
RIF1 KO
2
0
-2
Single Cell Repli-seq
WT
RIF1 KO
% replicated
% replicated
chr4
replicated
unreplicated
B
WT
HAP1
RIF1 KO
Population RT
1.0
0.5
0.0
−0.5
−1.0
30%
70%
30%
70%
Replicated bins
Unreplicated bins

Fig. S5. Single cell repli-seq confirms RT heterogeneity in RIF1 KO cells.
(A) Log2(E/L) Repli-seq plots of Chr4 in HAP1 WT and RIF1 KO (top). Binarized single-cell Repli-seq heatmaps of Chr4 (bottom) in HAP1 WT and RIF1 KO sorted by percentage replicated where the top row represents the cell with the lowest percentage of genome replicated and the last row represents the cell with the most percentage of genome replicated. (B) Boxplots showing distribution of population-based E/L Repli-seq RT values for replicated (red) and unreplicated (grey) bins for each single cell in both WT (top) and RIF1 KO (bottom) HAP1 ranked between 30 and 70% S-phase progression.

Fig. S6.

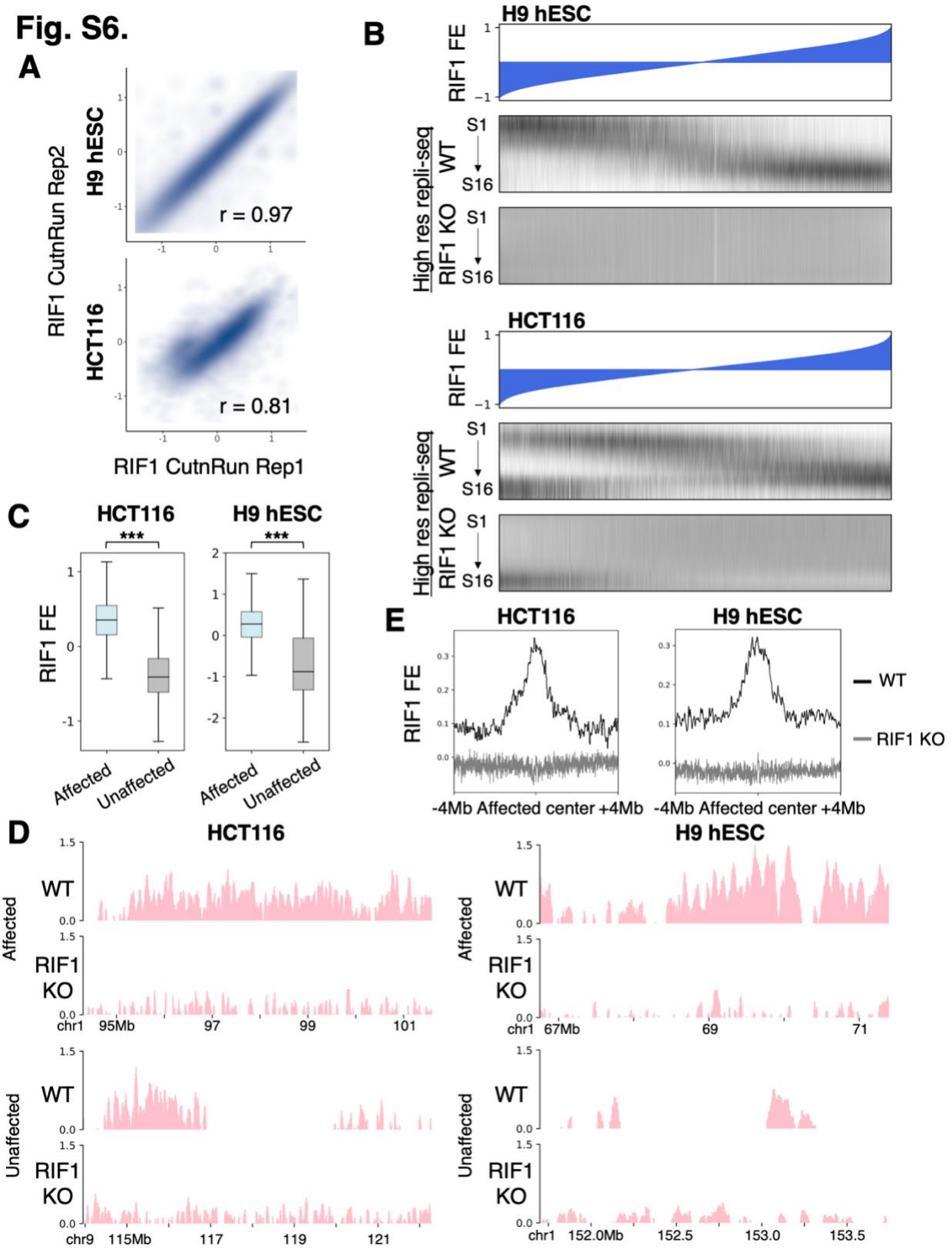

A
RIF1 CutnRun Rep2
H9 hESC
r = 0.97
HCT116
r = 0.81
RIF1 CutnRun Rep1
B
H9 hESC
RIF1 FE
High res repli-seq
WT
RIF1 KO
S1
S16
HCT116
C
HCT116
H9 hESC
***
RIF1 FE
Affected
Unaffected
D
HCT116
H9 hESC
Affected
Unaffected
WT
RIF1 KO
chr1 95Mb 97 99 101
chr1 67Mb 69 71
chr9 115Mb 117 119 121
chr1 152.0Mb 152.5 153.0 153.5
E
HCT116
H9 hESC
RIF1 FE
WT
RIF1 KO
-4Mb Affected center +4Mb

Fig. S6. RIF1 binds late replicating chromatin in large domains.
(A) Smoothed scatter plots and Pearson's correlation coefficients showing correlation between RIF1 Cut&Run replicates in GFP tagged RIF1 H9 hESC and HCT116. (B) RIF1 log2 fold enrichment (FE) over input ranked from lowest (left) to highest (right) and accordingly column-wise sorted high resolution Repli-seq data in WT and RIF1 KO H9 hESCs and HCT116. (C) Boxplot of RIF1 log2 FE at all affected (blue) and unaffected (grey) late regions in WT H9 hESCs and HCT116; ***$p < 0.0005$. (D) RIF1 CutnRun tracks in WT and RIF1 KO (as negative control) H9 hESC and HCT116 cell lines at regions shown in Fig 2a, c. (E) Pile up line plot showing average enrichment of RIF1 centered at affected late regions ±4Mb in WT (black) and RIF1 KO (grey, as negative control) HCT116 and H9 hESCs.

Fig. S7.

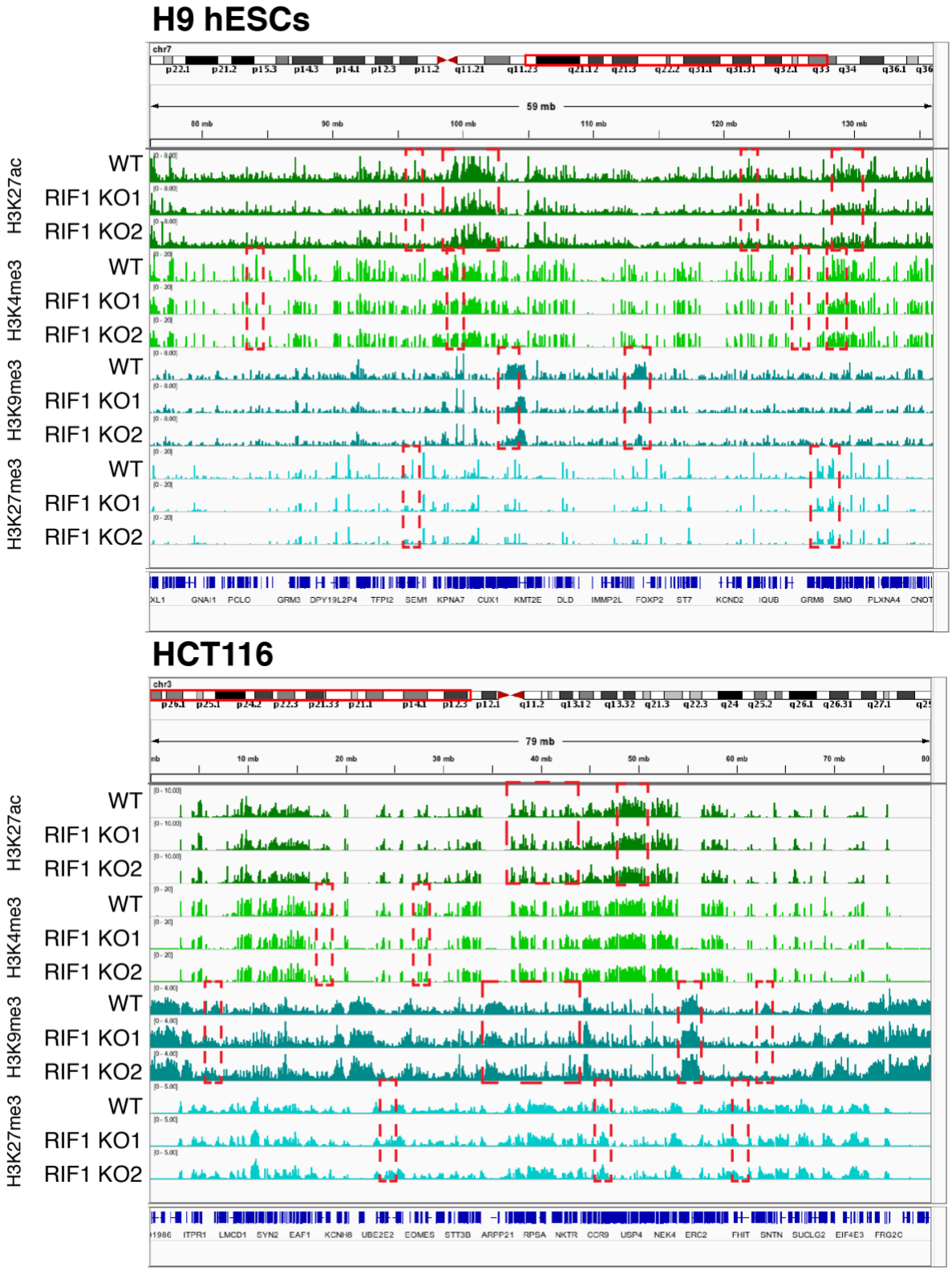

H9 hESCs
chr7
59 mb
80 mb
90 mb
100 mb
110 mb
120 mb
130 mb
H3K27ac
H3K4me3
H3K9me3
H3K27me3
WT
RIF1 KO1
RIF1 KO2
HCT116
chr3
79 mb
10 mb
20 mb
30 mb
40 mb
50 mb
60 mb
70 mb

Fig. S7. RIF1 KO causes histone mark changes.
Example tracks of indicated histone modifications in WT and two RIF1 KO replicates at Chr7 75-137 Mb in H9 hESCs (top) and Chr3 0-80.5 Mb in HCT116 (bottom). Red boxes indicate selected regions of differential enrichment.

Fig. S8.

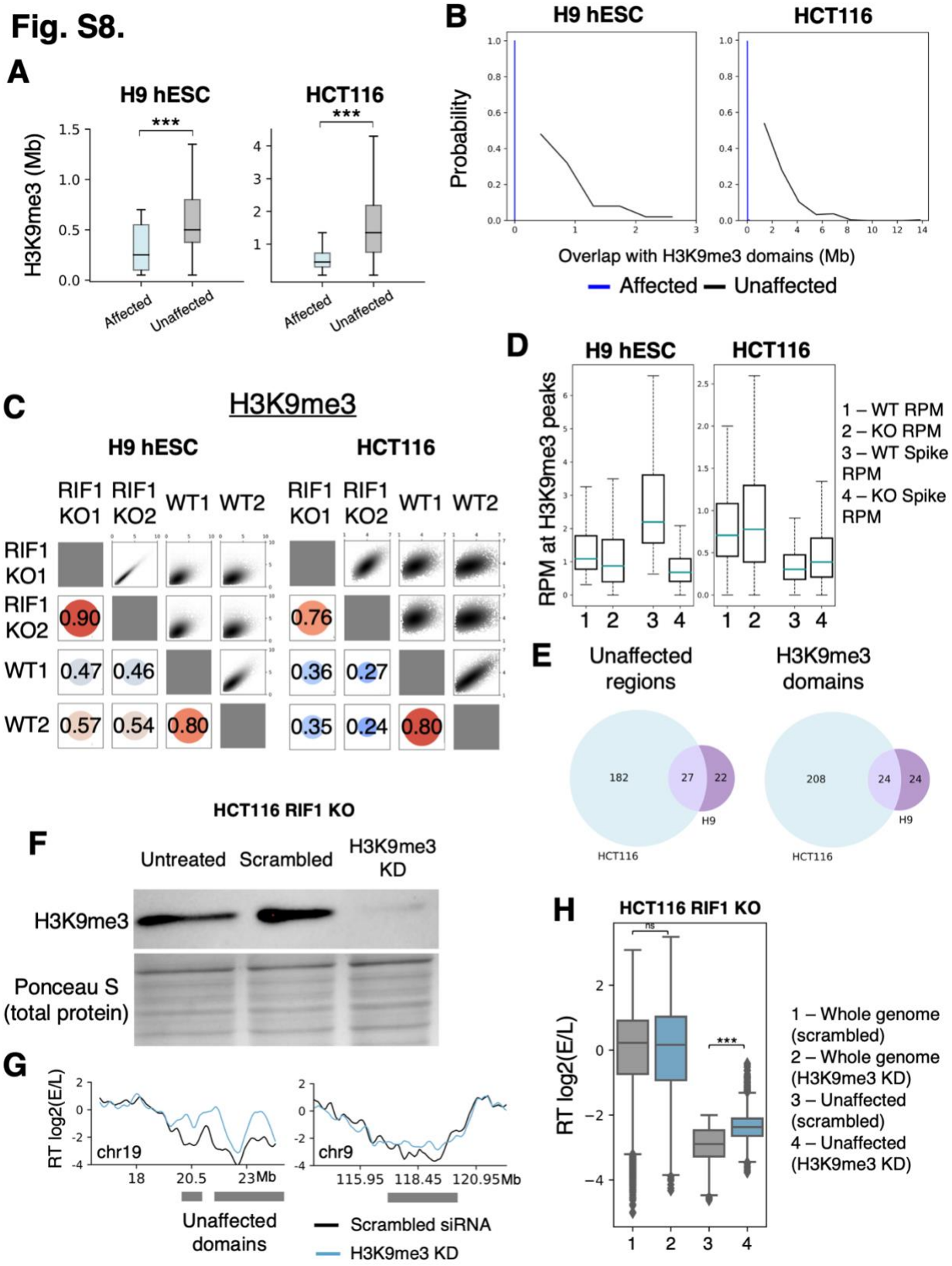

A
H9 hESC
HCT116
H3K9me3 (Mb)
***
Affected
Unaffected
B
H9 hESC
HCT116
Probability
Overlap with H3K9me3 domains (Mb)
Affected
Unaffected
C
H3K9me3
H9 hESC
HCT116
RIF1 KO1
RIF1 KO2
WT1
WT2
0.90
0.47
0.46
0.57
0.54
0.80
0.76
0.36
0.27
0.35
0.24
0.80
D
H9 hESC
HCT116
RPM at H3K9me3 peaks
1 – WT RPM
2 – KO RPM
3 – WT Spike RPM
4 – KO Spike RPM
E
Unaffected regions
H3K9me3 domains
182
27
22
208
24
24
H9
HCT116
F
HCT116 RIF1 KO
Untreated
Scrambled
H3K9me3 KD
H3K9me3
Ponceau S (total protein)
G
RT log2(E/L)
chr19
18
20.5
23Mb
chr9
115.95
118.45
120.95Mb
Unaffected domains
Scrambled siRNA
H3K9me3 KD
H
HCT116 RIF1 KO
ns
***
RT log2(E/L)
1 – Whole genome (scrambled)
2 – Whole genome (H3K9me3 KD)
3 – Unaffected (scrambled)
4 – Unaffected (H3K9me3 KD)

Fig. S8. Repressive histone marks at affected and unaffected regions.
(A) Boxplots of size distributions of H3K9me3 peaks or domains in affected (blue) and unaffected (grey) late regions in WT H9 hESCs and HCT116; ***p < 0.0005. (B) Line plots showing the distribution of overlap between affected (blue) and unaffected (black) regions with H3K9me3 domains in Mbs (see Methods for domain definition) in WT H9 hESCs and HCT116. (C) Smoothed scatter plots and Pearson's correlation coefficients (mapped onto a red-blue divergent colormap) of -log10q values associated with H3K9me3 broad peaks identified using MACS2 in two WT and RIF1 KO clones in H9 hESCs and HCT116. (D) Boxplot of RPM at H3K9me3 peaks and domains in WT and RIF1 KO HCT116 and H9 hESCs using only sequencing depth normalization (RPM columns) or RPM normalization to Drosophila spike-in (Spike RPM columns). T test p values are H9 hESC RPM p = 1.55x10-4; H9 hESC Spike RPM p = 1.13x10-63; HCT116 RPM p = 1.99x10-5; HCT116 Spike RPM p = 8.73x10-10. (E) Venn Diagrams of shared H3K9me3 domains (left) and unaffected RT regions (right) between RIF1 KO HCT116 and H9 hESCs showing that many H3K9me3 unaffected regions (excluding repetitive regions that cannot be effectively mapped to the genome) do not overlap between cell types, consistent with the model that H3K9me3 domains are developmentally regulated and implicated in silencing of lineage inappropriate genes. (F) Immunoblot of H3K9me3 in HCT116 RIF1 KO cells with indicated treatments. Ponceau S total protein stain used as loading control. (G) E/L Repli-seq of example unaffected domains in HCT116 RIF1 KO cells treated with either scrambled siRNA (black) or H3K9me3 KD siRNAs (blue). Grey bars indicate regions called as unaffected late domains in RIF1 KO HCT116 E/L Repli-seq data. (H) Boxplot of genome wide (columns 1 and 2) and unaffected late region (columns 3 and 4) E/L RT values in HCT116 RIF1 KO cells treated with either scrambled siRNA (grey) or H3K9me3 KD siRNAs (blue) (***p < 0.0005).

Fig. S9.

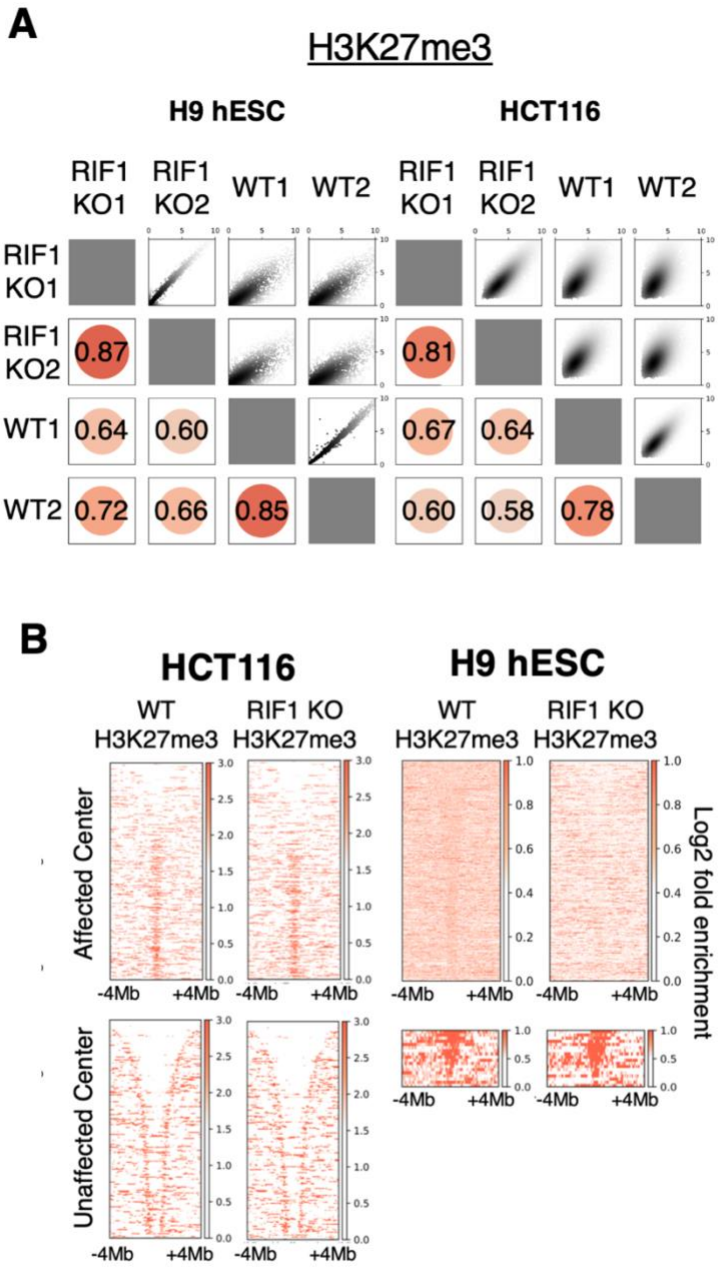

A
H3K27me3
H9 hESC
HCT116
RIF1 KO1
RIF1 KO2
WT1
WT2
0.87
0.64
0.60
0.72
0.66
0.85
0.81
0.67
0.64
0.60
0.58
0.78
B
HCT116
H9 hESC
WT H3K27me3
RIF1 KO H3K27me3
Affected Center
Unaffected Center
-4Mb
+4Mb
Log2 fold enrichment

Fig. S9. RIF1 KO causes cell-type specific changes to H3K27me3.
(A) Smoothed scatter plots and Pearson's correlation coefficients (mapped onto a red-blue divergent colormap) of -log10q values associated with H3K27me3 broad peaks identified using MACS2 in two WT and RIF1 KO clones in H9 hESCs and HCT116. (B) Log2 Fold enrichment over input pile ups of signal from H3K27me3 ChIP-seq in HCT116 and H9 hESC WT or RIF1 KO cells centered on affected late regions (top) or unaffected late regions (bottom) ±4 Mb and sorted by region size. H3K27me3 peaks were enriched at most affected regions but depleted from unaffected regions in WT HCT116 cells and these distributions were not significantly altered by RIF1 KO. H3K27me3 was similar to H3K9me3 in H9 hESCs in that RIF1 loss resulted in the depletion of H3K27me3 peaks at affected regions signal at the few unaffected regions was maintained.

Fig. S10.

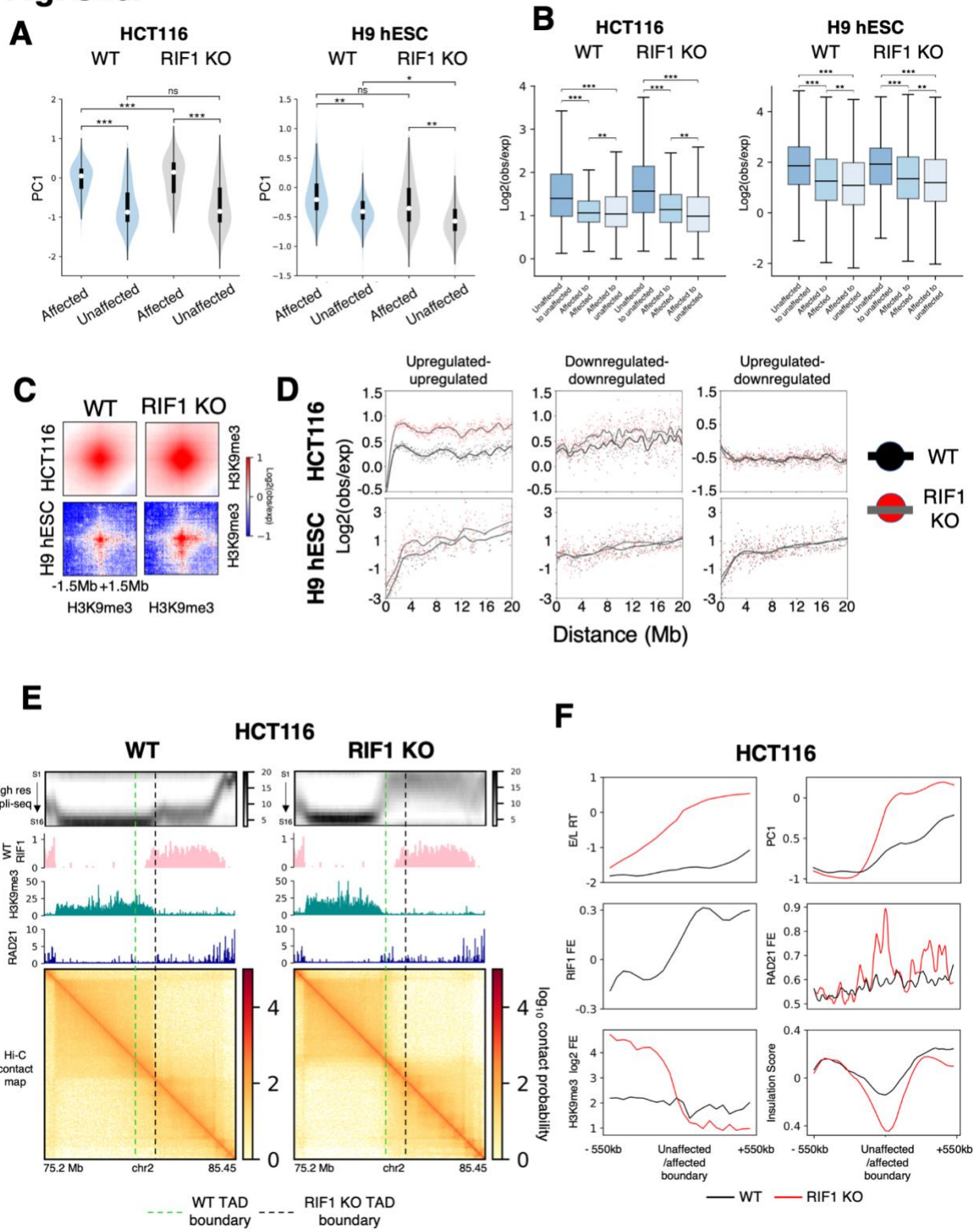

A
HCT116
WT
RIF1 KO
H9 hESC
PC1
Affected
Unaffected
B
Log2(obs/exp)
C
H3K9me3
-1.5Mb +1.5Mb
D
Upregulated-upregulated
Downregulated-downregulated
Upregulated-downregulated
Distance (Mb)
E
High res repli-seq
Hi-C contact map
75.2 Mb
chr2
85.45
WT TAD boundary
RIF1 KO TAD boundary
F
E/L RT
PC1
RIF1 FE
RAD21 FE
Insulation Score
- 550kb
Unaffected /affected boundary
+550kb

Fig. S10. Affected and Unaffected late regions form separate interaction hubs.
(A) Violin plots of PC1 Eigenvector values for affected and unaffected late regions in WT (blue) and RIF1 KO (grey) HCT116 (left) and H9 hESCs (right); (***$p < 0.0005$, **$p < 0.005$, *$p < 0.05$). (B) Box plots of log2(obs/exp) interactions between and within affected and unaffected late regions in WT and RIF1 KO HCT116 (left) and H9 hESCs (right); (***$p < 0.0005$, **$p < 0.005$). (C) Log2(obs/exp) aggregate Hi-C interactions between H3K9me3 domains genome wide ±1.5 Mb in WT and RIF1 KO HCT116 and H9 hESCs. (D) Log2(obs/exp) interactions between indicated H3K9me3 peaks/domains over genomic distance in WT (black dots, black line) and RIF1 KO (red dots, grey line) cells. Interactions between upregulated H3K9me3 peaks (left). Interactions between downregulated H3K9me3 peaks (middle). Interactions between upregulated and downregulated H3K9me3 peaks (right). (E) ICE normalized Hi-C contact map of Chr2 75.2-85.45 Mb in WT and RIF1 KO HCT116 showing a new TAD boundary formation at the juxtaposition of an unaffected late replicating domain and an affected LtE region. Tracks of WT RIF1 Cut&Run, H3K9me3 ChIP-seq, and RAD21 ChIP-seq are shown to illustrate the concordant changes of these marks with RT changes. Black dotted line indicates the TAD boundary in WT cells. Green dotted line indicates new TAD boundary in RIF1 KO cells. In individual cases (n = 7) where RIF1 KO caused a large late replicating domain to break into an unaffected region and an affected region, a new TAD boundary formed coinciding with the new RT boundary and demarcated by enriched H3K9me3 and the cohesin subunit RAD21, showing that changes in chromatin state correlated with loss of RT control also correlate with the formation of new TAD boundaries. These regions also exhibited a pronounced shift in PC1 Eigenvector indicating the formation of a new compartment boundary. (F) Pile up line plots showing indicated features at newly formed unaffected/affected timing boundaries as exemplified in Fig. S10E in WT (black) and RIF1 KO (red) HCT116 cells.

Fig. S11.

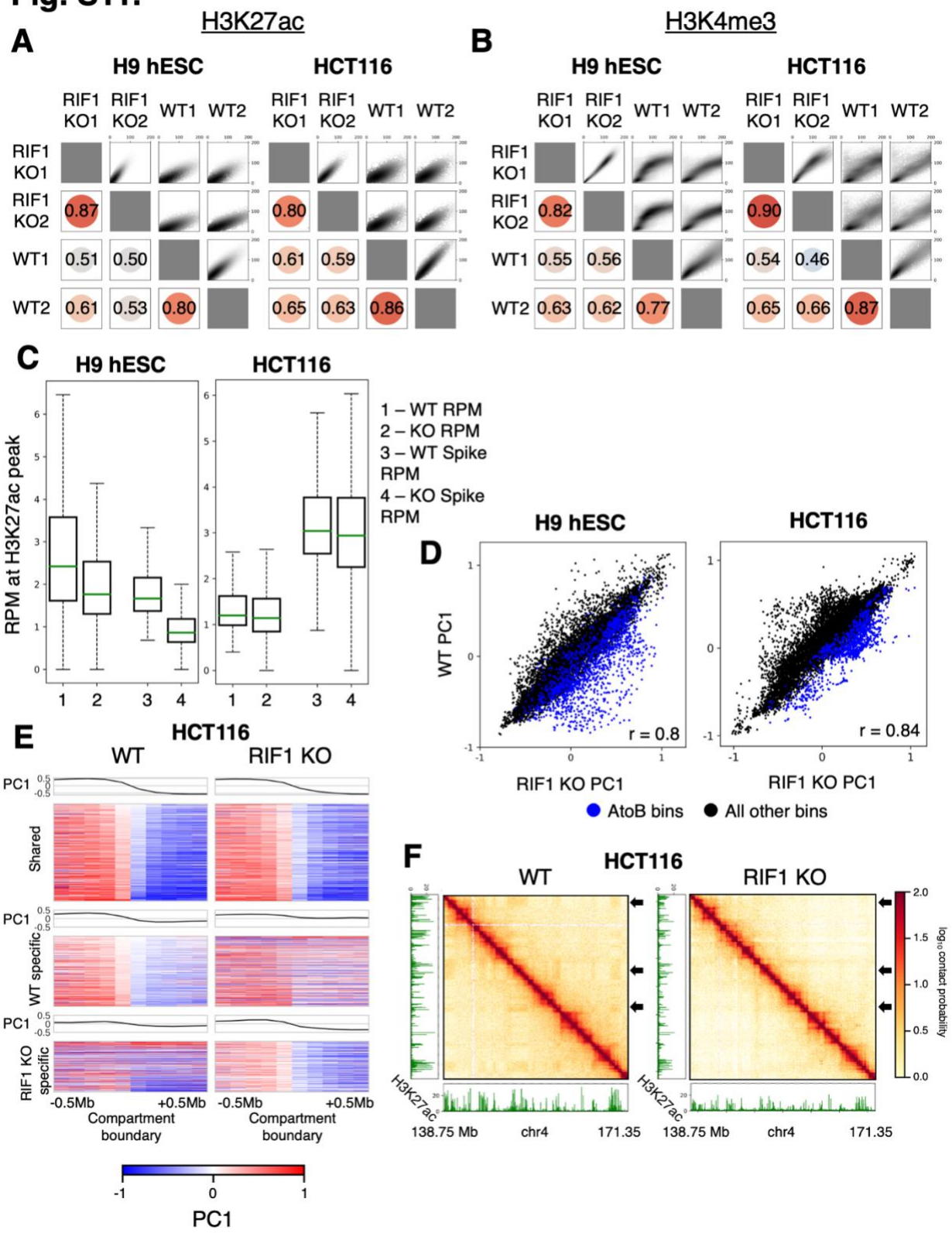

A
H3K27ac
H9 hESC
HCT116
RIF1 KO1
RIF1 KO2
WT1
WT2
0.87
0.51
0.50
0.61
0.53
0.80
0.80
0.61
0.59
0.65
0.63
0.86
B
H3K4me3
0.82
0.55
0.56
0.63
0.62
0.77
0.90
0.54
0.46
0.65
0.66
0.87
C
RPM at H3K27ac peak
1 – WT RPM
2 – KO RPM
3 – WT Spike RPM
4 – KO Spike RPM
D
WT PC1
RIF1 KO PC1
r = 0.8
r = 0.84
AtoB bins
All other bins
E
WT
RIF1 KO
PC1
Shared
WT specific
RIF1 KO specific
-0.5Mb
+0.5Mb
Compartment boundary
PC1
F
WT
RIF1 KO
H3K27ac
138.75 Mb
chr4
171.35
log10 contact probability

Fig. S11. Active histone marks are depleted and correlate with compartmental changes in RIF1 KO cells.
(A) Smoothed scatter plots and Pearson's correlation coefficients (mapped onto a red-blue divergent colormap) of -log10q values associated with H3K27ac narrow peaks identified by MACS2 in two WT and RIF1 KO clones in H9 hESCs and HCT116. (B) Smoothed scatter plots and Pearson's correlation values (mapped onto a red-blue divergent colormap) of -log10q values associated with H3K4me3 narrow peaks identified by MACS2 in WT and two RIF1 KO clones in H9 hESCs and HCT116. (C) Boxplot of RPM at H3K27ac peaks in WT and RIF1 KO HCT116 and H9 hESCs using only sequencing depth normalization (RPM columns) or RPM normalization to Drosophila spike-in (Spike RPM columns). T test p values are H9 hESC RPM p = 3.368x10-21; H9 hESC Spike RPM p = 5.692x10-79; HCT116 RPM p = 1.843x10-17; HCT116 Spike RPM p = 0.00014. (D) Scatterplots of genome wide PC1 eigenvector values for 250kb bins between RIF1 KO (x axis) and WT (y axis) samples in H9 hESCs and HCT116 with Pearson's correlation coefficients. Blue points represent bins contained within AtoB compartment switching regions. (E) Heatmaps of PC1 values centered on shared (top), WT specific (middle), and RIF1 KO specific (bottom) compartment boundaries ±0.5 Mb in WT and RIF1 KO HCT116. Each heatmap column represents 100kb. (F) ICE normalized Hi-C contact map of Chr4 138.75-171.35 Mb in HCT116 WT and RIF1 KO cells with accompanying H3K27ac ChIP-seq plots. Arrows indicate A compartments and H3K27ac peaks that are lost upon RIF1 KO.

Fig. S12.

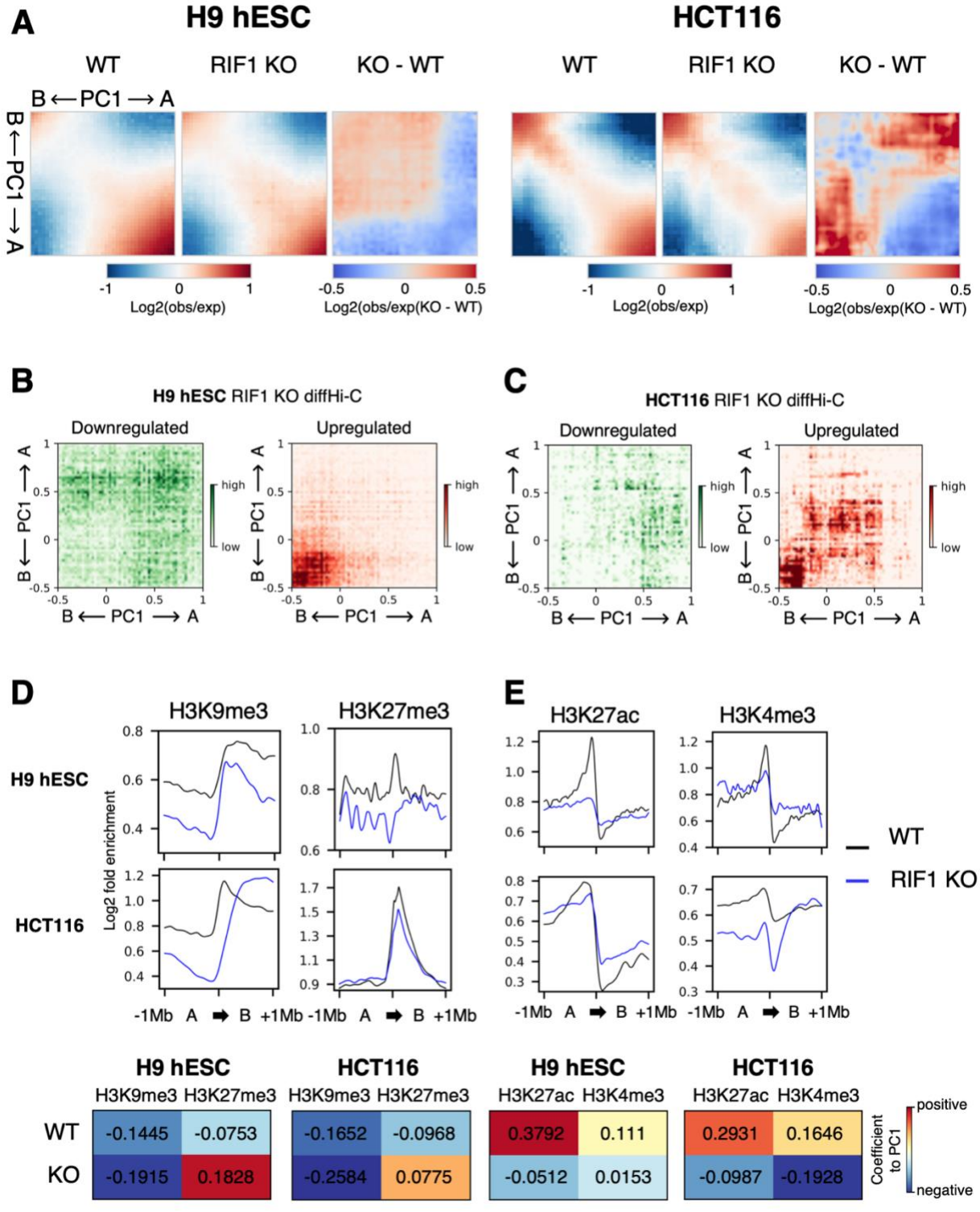

A
H9 hESC
HCT116
WT
RIF1 KO
KO - WT
B ⟵ PC1 ⟶ A
Log2(obs/exp)
Log2(obs/exp(KO - WT)
B
H9 hESC RIF1 KO diffHi-C
Downregulated
Upregulated
C
HCT116 RIF1 KO diffHi-C
high
low
D
H3K9me3
H3K27me3
H9 hESC
HCT116
Log2 fold enrichment
-1Mb A ➡ B +1Mb
E
H3K27ac
H3K4me3
WT
RIF1 KO
H9 hESC
HCT116
H3K9me3 H3K27me3
WT -0.1445 -0.0753
KO -0.1915 0.1828
H3K9me3 H3K27me3
WT -0.1652 -0.0968
KO -0.2584 0.0775
H3K27ac H3K4me3
WT 0.3792 0.111
KO -0.0512 0.0153
H3K27ac H3K4me3
WT 0.2931 0.1646
KO -0.0987 -0.1928
Coefficient to PC1
positive
negative

Fig. S12. RIF1 KO causes disruption of compartment identities.
(A) Genome wide saddle plots of log2(obs/exp) chromatin interactions in WT, RIF1 KO, and subtractions in H9 hESCs and HCT116 sorted by increasing PC1 values. (B) 2D histogram of PC1 eigenvectors of downregulated (left) and upregulated (right) Hi-C interactions in H9 hESCs identified through diffHiC. (C) 2-D histogram of PC1 eigenvector values of downregulated (left) and upregulated (right) Hi-C interactions identified through diffHiC in HCT116 cells. (D) (top) Pile up line plots of log2 FE for H3K9me3 and H3K27me3 at A to B compartment boundary ±1 Mb in WT (black) and RIF1 KO (blue) HCT116 and H9 hESCs. (bottom) Multiple linear regression coefficient values for indicated histone marks to PC1 values of the same bins in respective cell lines, where larger positive values indicate greater positive correlation and larger negative values indicate greater inverse correlation. (E) (top) Pile up line plots of log2 FE for H3K27ac and H3K4me3 at A to B compartment boundary ±1 Mb in WT (black) and RIF1 KO (blue) HCT116 and H9 hESCs. (bottom) Multiple linear regression coefficient values for indicated histone marks to PC1 values of the same bins in respective cell lines.

Fig. S13.

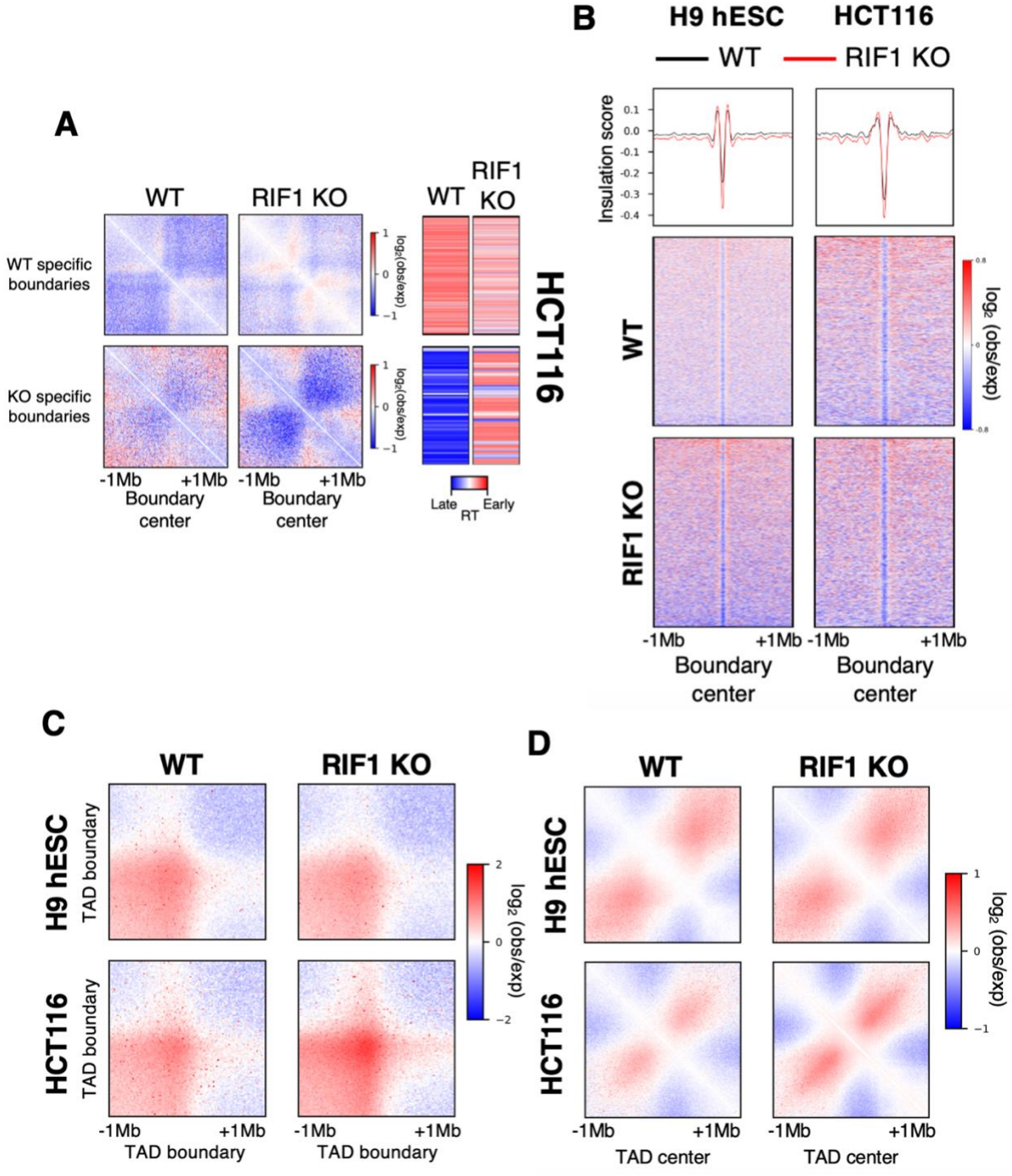

A
WT
RIF1 KO
WT
RIF1 KO
WT specific boundaries
KO specific boundaries
log2(obs/exp)
HCT116
-1Mb
+1Mb
Boundary center
Late
Early
RT
B
H9 hESC
HCT116
WT
RIF1 KO
Insulation score
WT
RIF1 KO
log2 (obs/exp)
C
WT
RIF1 KO
H9 hESC
HCT116
TAD boundary
log2 (obs/exp)
TAD boundary
D
WT
RIF1 KO
H9 hESC
HCT116
log2 (obs/exp)
TAD center

Fig. S13. TAD positions are maintained, and boundaries are strengthened in RIF1 KO cells.
(A) Log2(obs/exp) interaction pile ups centered at WT and RIF1 KO specific TAD boundaries ±1 Mb in WT and RIF1 KO HCT116 (left). RT indices of WT and RIF1 KO specific TAD boundaries in WT and RIF1 KO HCT116 (right) show that WT specific boundaries are enriched in EtL regions whereas KO specific boundaries are associated with LtE RT changes. (B) Insulation score pile ups centered on all TAD boundaries genome wide ±1 Mb in WT and RIF1 KO H9 hESCs and HCT116. Line plots represent mean scores in WT (black) and RIF1 KO (red). Genome wide, the number (~4000 in H9 hESC and ~2600 in HCT116) and positioning of TAD boundaries remained similar between WT and RIF1 KO. (C) Log2(obs/exp) interaction pile ups centered on TAD boundary ±1 Mb in WT and RIF1 KO H9 hESCs and HCT116. (D) Log2(obs/exp) interaction pile ups centered on TAD center ±1 Mb in WT and RIF1 KO H9 hESCs and HCT116.

Fig. S14.

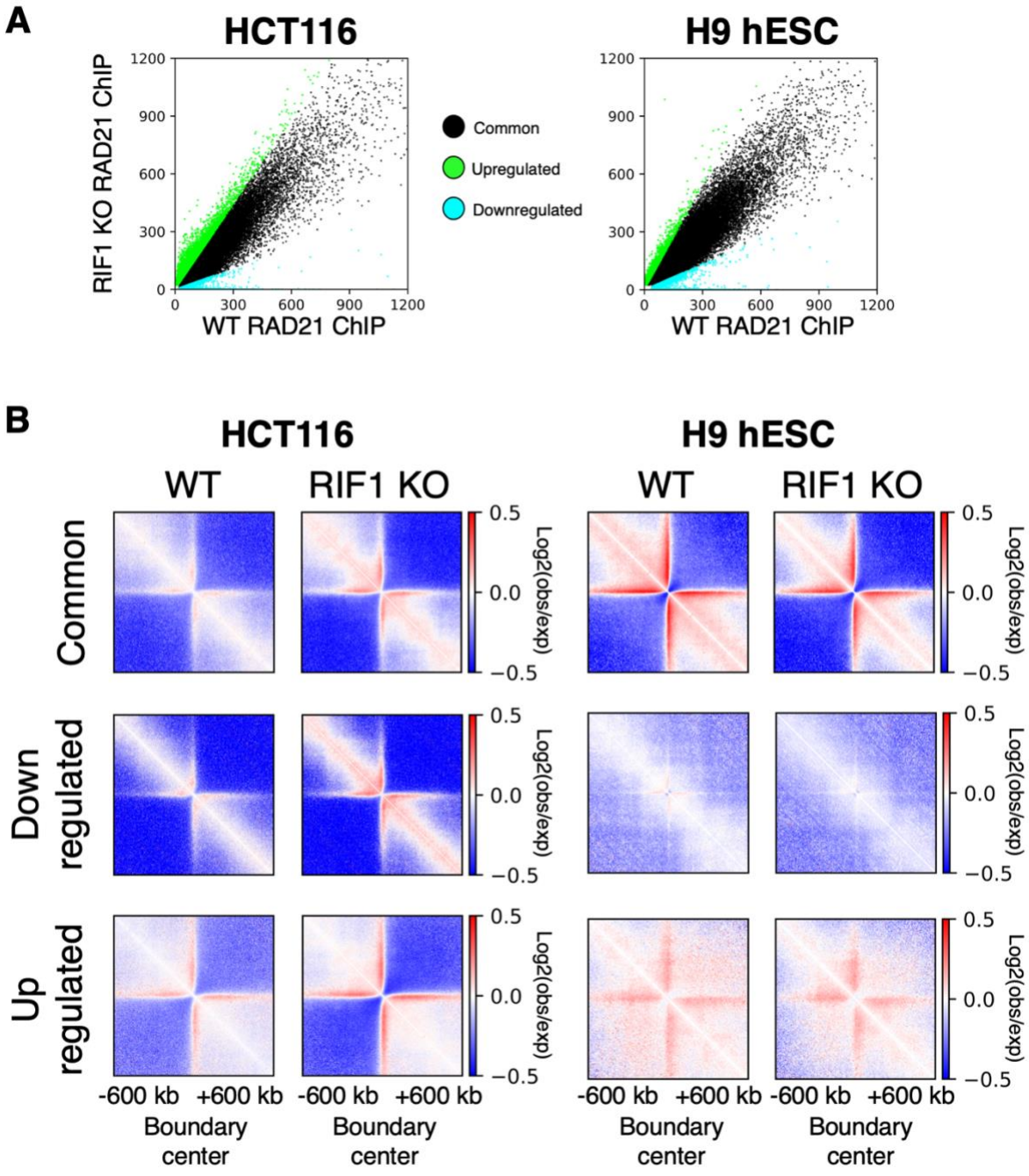
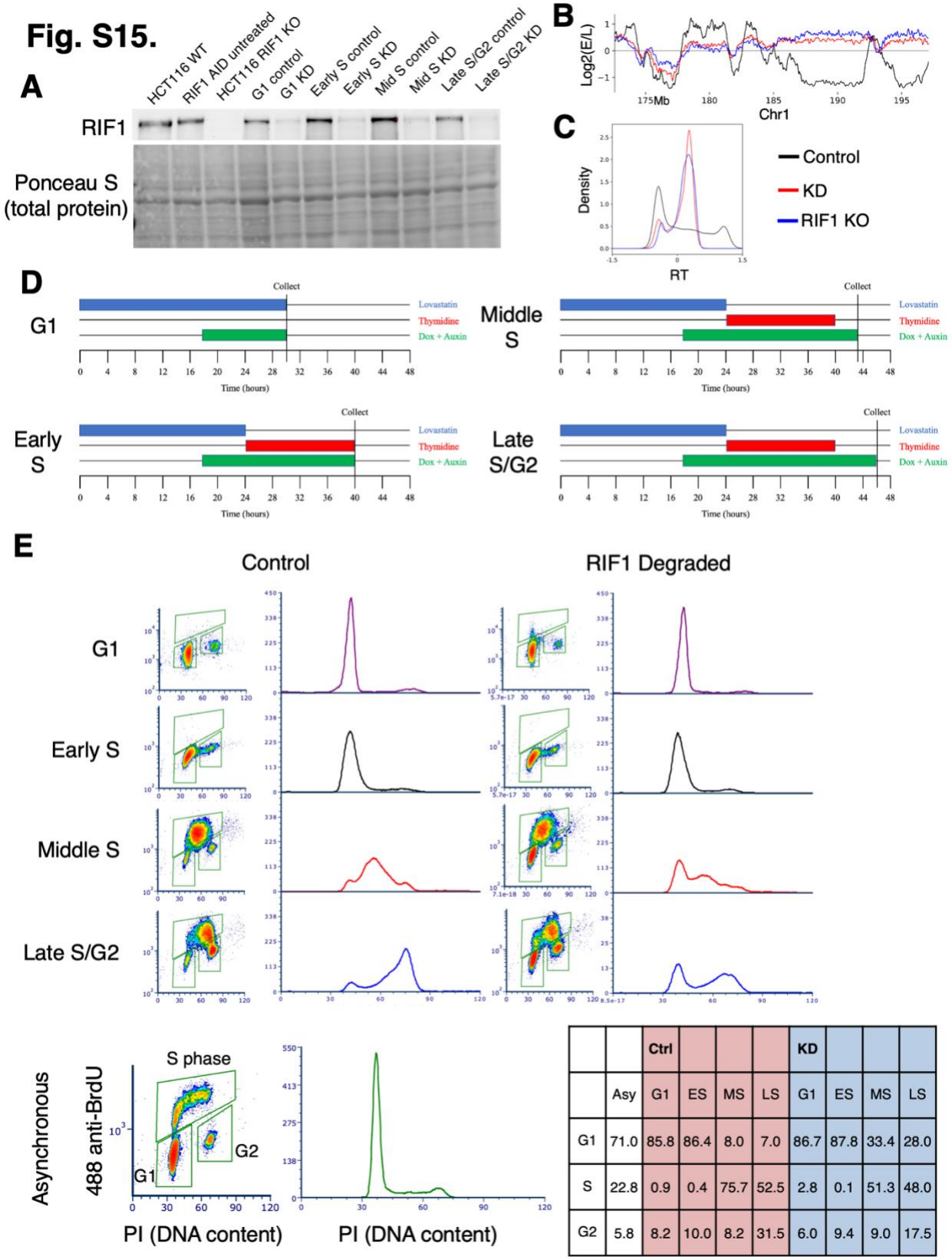

A
HCT116
H9 hESC
RIF1 KO RAD21 ChIP
WT RAD21 ChIP
Common
Upregulated
Downregulated
0
300
600
900
1200
B
HCT116
H9 hESC
WT
RIF1 KO
Common
Down regulated
Up regulated
Log2(obs/exp)
0.5
0.0
−0.5
-600 kb
+600 kb
Boundary center

Fig. S14. TAD boundary strengthening is not caused by changes in RAD21 binding in RIF1 KO cells.
(A) Scatterplots of RAD21 ChIP-seq peak read counts in WT (x axis) versus RIF1 KO (y axis). Significant up or down regulated peaks are those with at least 2-fold difference and FDR < 0.05. Common peaks colored in black, upregulated peaks colored in green, and downregulated peaks colored in cyan. RAD21's binding remained largely unchanged with a small number of up and down regulated peaks in both RIF1 KO cell lines. (B) Log2(obs/exp) interaction pile ups centered at RAD21 peaks that are common (top), downregulated (middle), or upregulated (bottom) ±600kb between WT and RIF1 KO cells in HCT116 (left) and H9 hESCs (right) indicating a lack of correlation between RAD21 peak change and boundary strength.

Fig. S15.



		Ctrl				KD			
	Asy	G1	ES	MS	LS	G1	ES	MS	LS
G1	71.0	85.8	86.4	8.0	7.0	86.7	87.8	33.4	28.0
S	22.8	0.9	0.4	75.7	52.5	2.8	0.1	51.3	48.0
G2	5.8	8.2	10.0	8.2	31.5	6.0	9.4	9.0	17.5

Fig. S15. RIF1-AID degron synchronization.
(A) Immunoblot of RIF1 protein in HCT116 WT, untreated HCT116 RIF1-AID, HCT116 RIF1 KO, and control and KD RIF1-AID samples at indicated synchronization time point. Ponceau S total protein stain used as loading control. (B) E/L Repli-seq track showing Chr1 172.6-197.6 Mb in control (black) and KD RIF1-AID (red) cells in the first S phase after synchronization in G1 compared to RIF1 KO HCT116 (blue) showing the full capitulation of RT disruption in the first S phase post RIF1 KD. (C) Genome wide histogram of E/L Repli-seq values in control (black) and KD RIF1-AID (red) cells in the first S phase after synchronization in G1 compared to RIF1 KO HCT116 (blue). (D) Schematic representation of synchronization protocol for each synchronization time point (See Methods). Collection point indicates when cells were harvested for ChIP-seq or Hi-C preparation. (E) Cell cycle FACS plots of control and KD RIF1-AID samples at each synchronization time point showing the synchronized release of cells into S phase. Table shows the percentage of cells within each cell cycle gate at each time point based on 2D BrdU versus PI FACS. KD RIF1-AID samples contained fewer cells in S phase at middle S and late S/G2 time points due to a subpopulation of cells stuck in G1 so for subsequent genomic experiments, a proportional number of reads from G1 synchronized samples were added to control samples at these time points to make fair comparisons between control and RIF1 KD samples (See Methods).

Fig. S16.

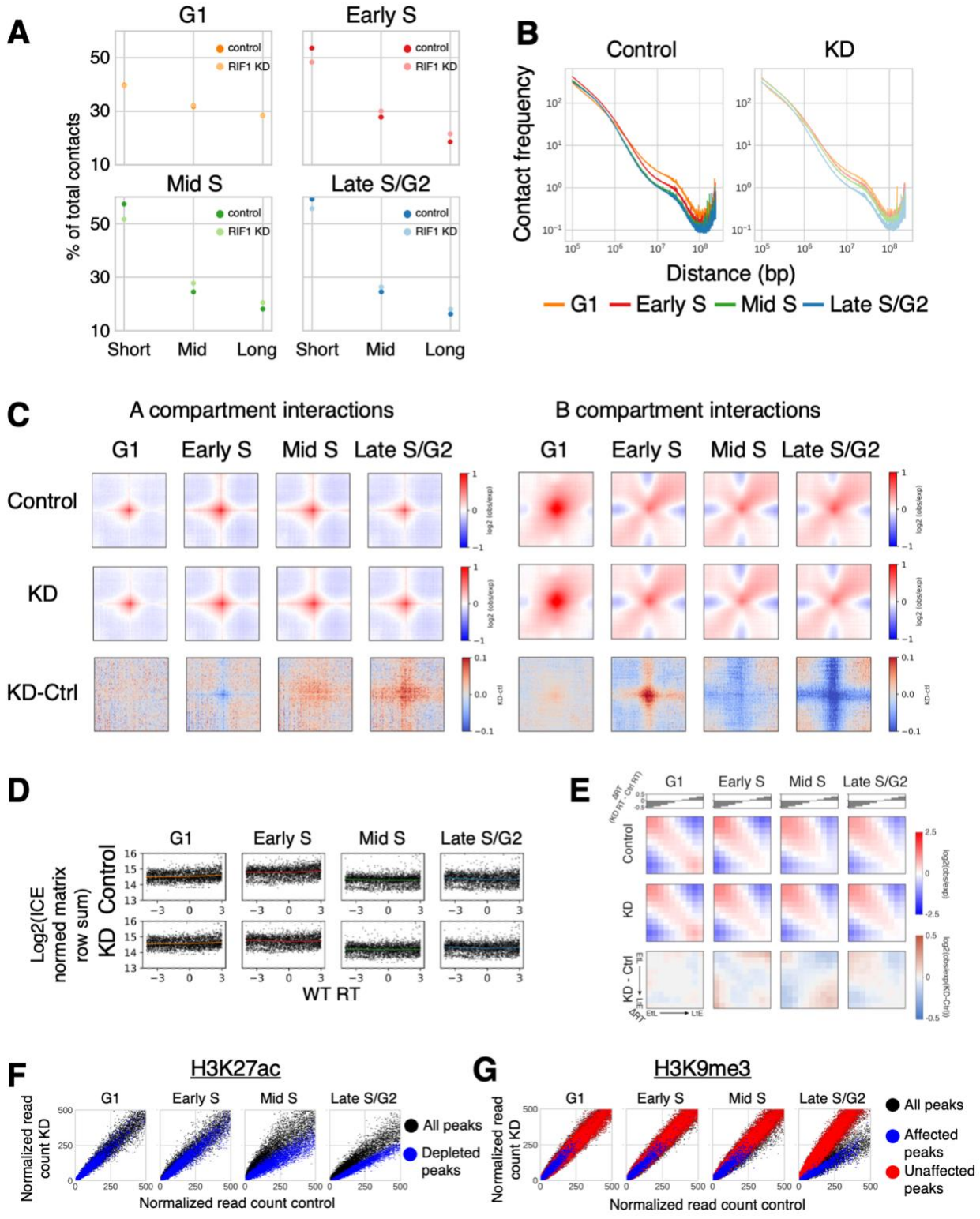

A
G1
Early S
Mid S
Late S/G2
control
RIF1 KD
% of total contacts
Short
Mid
Long
B
Control
KD
Contact frequency
Distance (bp)
G1
Early S
Mid S
Late S/G2
C
A compartment interactions
B compartment interactions
Control
KD
KD-Ctrl
log2 (obs/exp)
KD-ctrl
D
Log2(ICE normed matrix row sum)
Control
KD
WT RT
E
ΔRT (KD RT - Ctrl RT)
KD - Ctrl
ΔRT
E1L
L1E
log2(obs/exp)
log2(obs/exp)(KD-Ctrl)
F
H3K27ac
Normalized read count KD
Normalized read count control
All peaks
Depleted peaks
G
H3K9me3
Normalized read count KD
Normalized read count control
All peaks
Affected peaks
Unaffected peaks

Fig. S16. Epigenomic changes are an indirect result of RIF1 loss and require S phase. (A) Dot plot of the percentage of total Hi-C contacts constituted by short (< 1Mb), mid (1-10Mb), and long (10-100Mb) range contacts in control and RIF1 KD samples at G1, early S, middle S, and late S/G2 synchronization time points. Darker dots indicate control samples and lighter dots indicate KD samples. Medium and long-range chromatin contacts were depleted, and short-range contacts were enriched in S phase time points compared to G1 in control cells RIF1 degraded cells showed more medium and long-range interactions and fewer short-range interactions during S phase than control cells. (B) Distance decay curves of Hi-C contact frequency in control and KD RIF1-AID samples at each synchronization time point showing greater similarity at mid to long range interactions between KD curves through S phase than between control curves. (C) Log2(obs/exp) aggregate Hi-C interactions and subtraction (KD - control) plots of top 10% A compartment interactions (left) and top 10% B compartment interactions (right) in control and KD RIF1-AID samples at each synchronization time point. (D) Normalized coverage of genomic bins calculated as row-wise sum of genomic contacts in ICE normalized Hi-C matrices of control and KD RIF1-AID samples at each synchronization time point plotted against WT HCT116 RT showing that the whole genome has equal coverage at each time point regardless of RT. (E) Saddle plots of cis log2 (obs/exp) contacts of control (top), RIF1 KD (middle), and subtraction (KD – Control; bottom) sorted by ΔRT (KD RT – Control RT) at G1, early S, middle S, and late S/G2 synchronization time points. (F) Scatterplots of normalized read counts at H3K27ac peaks in control (x axis) and KD RIF1-AID (y axis) samples at G1, early S, middle S, and late S/G2 synchronization time points. Blue dots indicate significantly depleted H3K27ac peaks as called in late S/G2 time point and black dots indicate all H3K27ac peaks. (G) Scatterplots of normalized read counts at H3K9me3 peaks and domains in control (x axis) and KD RIF1-AID (y axis) samples at G1, early S, middle S, and late S/G2 synchronization time points. Red dots indicate H3K9me3 domains at unaffected late regions, blue dots indicate H3K9me3 peaks at affected late regions, and black dots indicate all H3K9me3 peaks/domains.

Fig. S17.

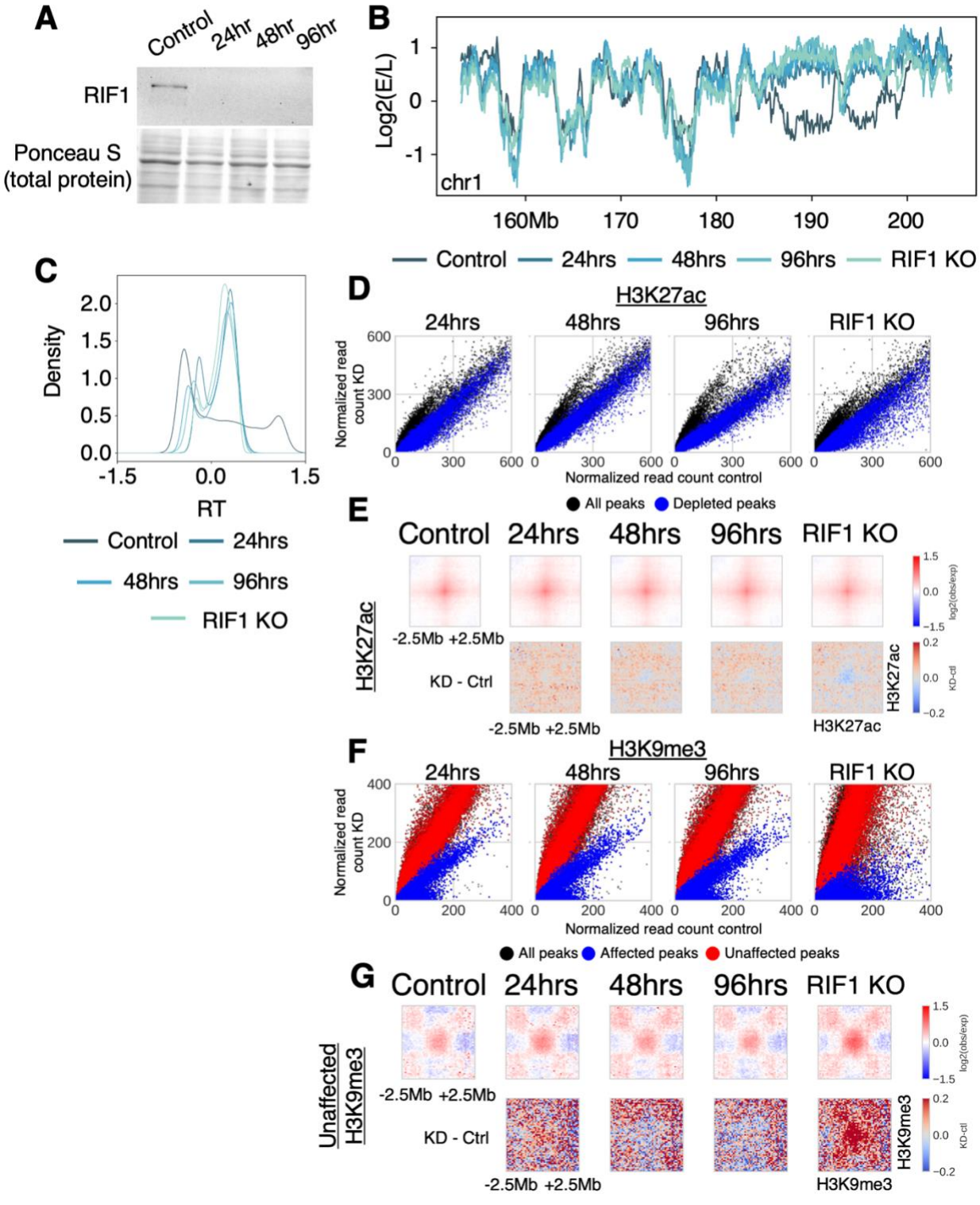

A
Control
24hr
48hr
96hr
RIF1
Ponceau S
(total protein)
B
Log2(E/L)
chr1
160Mb
170
180
190
200
Control
24hrs
48hrs
96hrs
RIF1 KO
C
Density
RT
Control
24hrs
48hrs
96hrs
RIF1 KO
D
H3K27ac
24hrs
48hrs
96hrs
RIF1 KO
Normalized read count KD
Normalized read count control
All peaks
Depleted peaks
E
Control
24hrs
48hrs
96hrs
RIF1 KO
H3K27ac
-2.5Mb +2.5Mb
KD - Ctrl
log2(obs/exp)
F
H3K9me3
Affected peaks
Unaffected peaks
G
Unaffected H3K9me3

Fig. S17. Epigenomic changes are exacerbated as cells continuously cycle without RIF1. (A) Immunoblot of RIF1 protein in untreated HCT116 RIF1-AID (control) and KD RIF1-AID samples at each degradation time point. Ponceau S total protein stain used as loading control. (B) E/L Repli-seq data of Chr1 155-205Mb in control RIF1-AID, RIF1 KO HCT116, 24 hour, 48 hour, and 96 hour RIF1-AID degradation time points. (C) Genome wide histogram of E/L Repli-seq values in control RIF1-AID, RIF1 KO HCT116, 24 hour, 48 hour, and 96 hour RIF1-AID degradation time points. (D) Log2(obs/exp) aggregate Hi-C interactions and subtraction (KD/KO – Control) plots between H3K27ac peaks genome wide ±2.5 Mb in control RIF1-AID, RIF1 KO HCT116, 24 hour, 48 hour, and 96 hour RIF1-AID degradation time points showing gradual decrease in interaction frequency as a function of time. (E) Log2(obs/exp) aggregate Hi-C interactions and subtraction (KD/KO – Control) plots between H3K9me3 domains genome wide ±2.5 Mb in control RIF1-AID, RIF1 KO HCT116, 24 hour, 48 hour, and 96 hour RIF1-AID degradation time points showing interaction increase only in RIF1 KO.

Fig. S18.

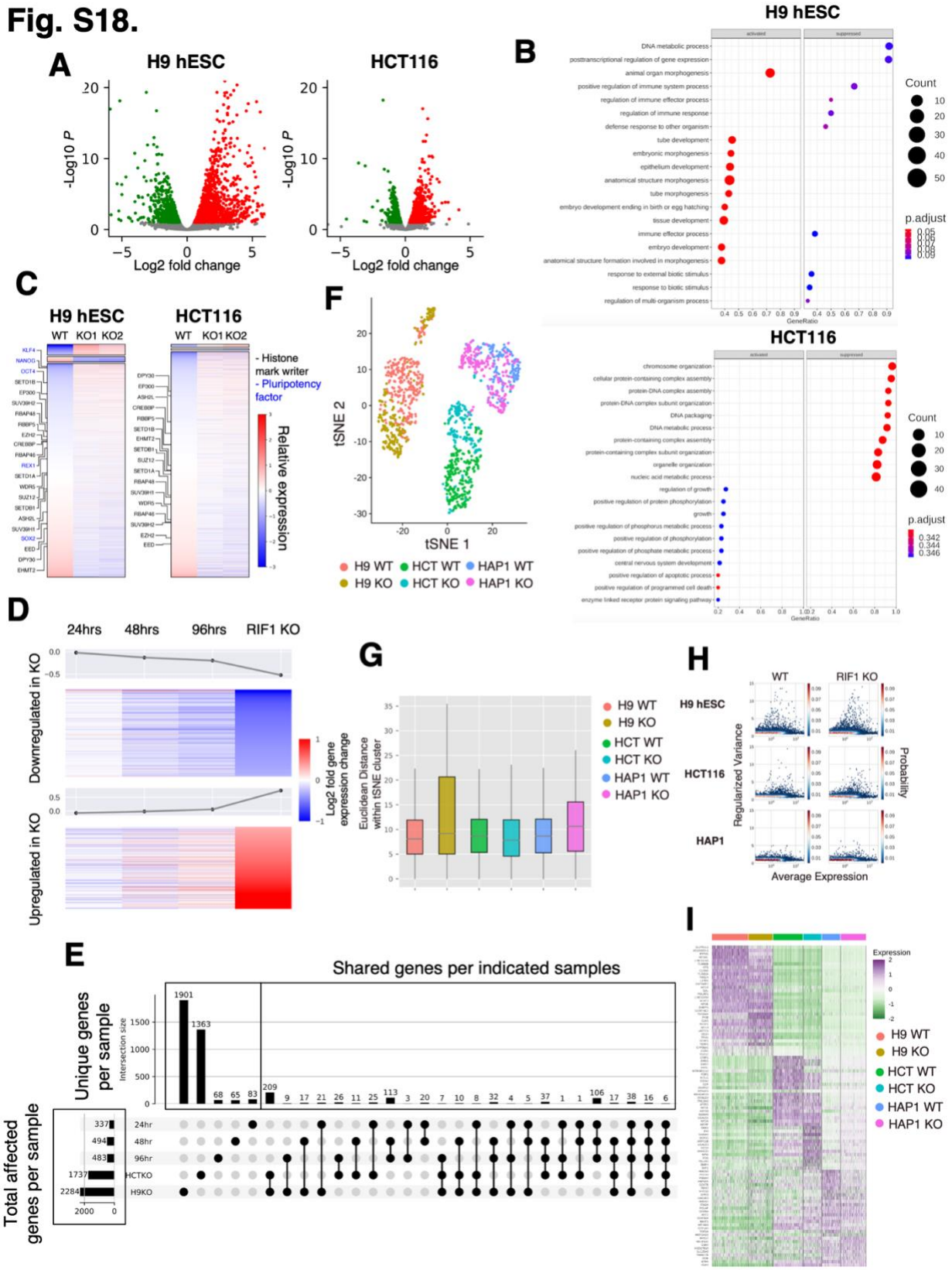

A
H9 hESC
HCT116
-Log10 P
Log2 fold change
B
H9 hESC
HCT116
Count
p.adjust
GeneRatio
C
H9 hESC
HCT116
WT KO1 KO2
- Histone mark writer
- Pluripotency factor
Relative expression
D
24hrs 48hrs 96hrs RIF1 KO
Downregulated in KO
Upregulated in KO
Log2 fold gene expression change
E
Shared genes per indicated samples
Unique genes per sample
Intersection size
Total affected genes per sample
24hr
48hr
96hr
HCTKO
H9KO
F
tSNE 1
tSNE 2
H9 WT
H9 KO
HCT WT
HCT KO
HAP1 WT
HAP1 KO
G
Euclidean Distance within tSNE cluster
H
WT
RIF1 KO
H9 hESC
HCT116
HAP1
Regularized Variance
Probability
Average Expression
I
Expression

Fig. S18. RIF1 KO causes sporadic gene expression changes.
(A) Volcano plots showing gene expression changes upon RIF1 KO in H9 hESC (1378 upregulated, 906 downregulated) (left) and HCT116 (818 upregulated, 919 downregulated) (right). Gene expression changes with an FDR< 0.1 were called as differential expression events (up-regulated genes: red, down-regulated genes: green, genes with non-significant changes: grey). (B) GO analysis dot plot of differentially expressed genes divided into activated and suppressed classes in RIF1 KO H9 hESCs and HCT116 where the sizes of dots are proportional to the number of genes and the color coding indicates the adjusted p value (FDR) of each GO term (red = more significant, blue = less significant). GO analysis revealed expression changes for genes important for cancer progression in HCT116 and developmentally regulated genes in H9 hESCs. (C) Heat maps of regularized log2(count) – log2(row-mean) of all genes in WT and two RIF1 KO clones in both cell lines. Top row contains significantly upregulated genes. Middle row contains significantly downregulated genes. Bottom row contains genes with no significant expression change. Genes in all rows are ranked by relative expression in WT cells. Positions of key pluripotency factors (blue text) and histone modification writers (black text) are indicated. (D) Heatmaps showing log2 fold change of genes downregulated (top) or upregulated (bottom) in HCT116 RIF1 KO at each time point over the extended RIF1 degradation time course in RIF1-AID HCT116 cells. Connected dots representing the average log2 fold change for downregulated or upregulated genes at each time point are shown on top of the heatmap. (E) Upset plot of significantly affected genes in RIF1 KO H9 hESCs, RIF1 KO HCT116, and each RIF1 degradation time point. Horizontal columns on the left indicate the total number of significantly affected genes in each sample. Vertical columns on top indicate the number of significantly affected genes that are specific to the category indicated by the dots below. (F) tSNE clusters of single cell RNA-seq of WT and RIF1 KO H9 hESCs, HCT116, and HAP1 cells. (G) Boxplots quantifying Euclidean distances within tSNE clusters in WT and RIF1 KO H9 hESCs ($p = 4.7736x10-102$), HCT116 ($p = 0.12552$), and HAP1 ($p = 4.5141x10-63$) cells. (H) Scatter density plots of regularized variance (y axis) versus average expression (x axis) of all genes in single cell RNA-seq data of WT and RIF1 KO H9 hESCs ($p = 0.01689$), HCT116 ($p = 0.17527$), and HAP1 ($p = 0.00037$). (I) Heatmap of log2 fold change of top 20 most differentially expressed genes for each cell type showing that expression differences between cell types are more pronounced than those between WT and KO of the same cell type. Differences in gene expression between WT and RIF1 KO within the same cell type were much less significant than those between distinct cell types.

Fig. S19.

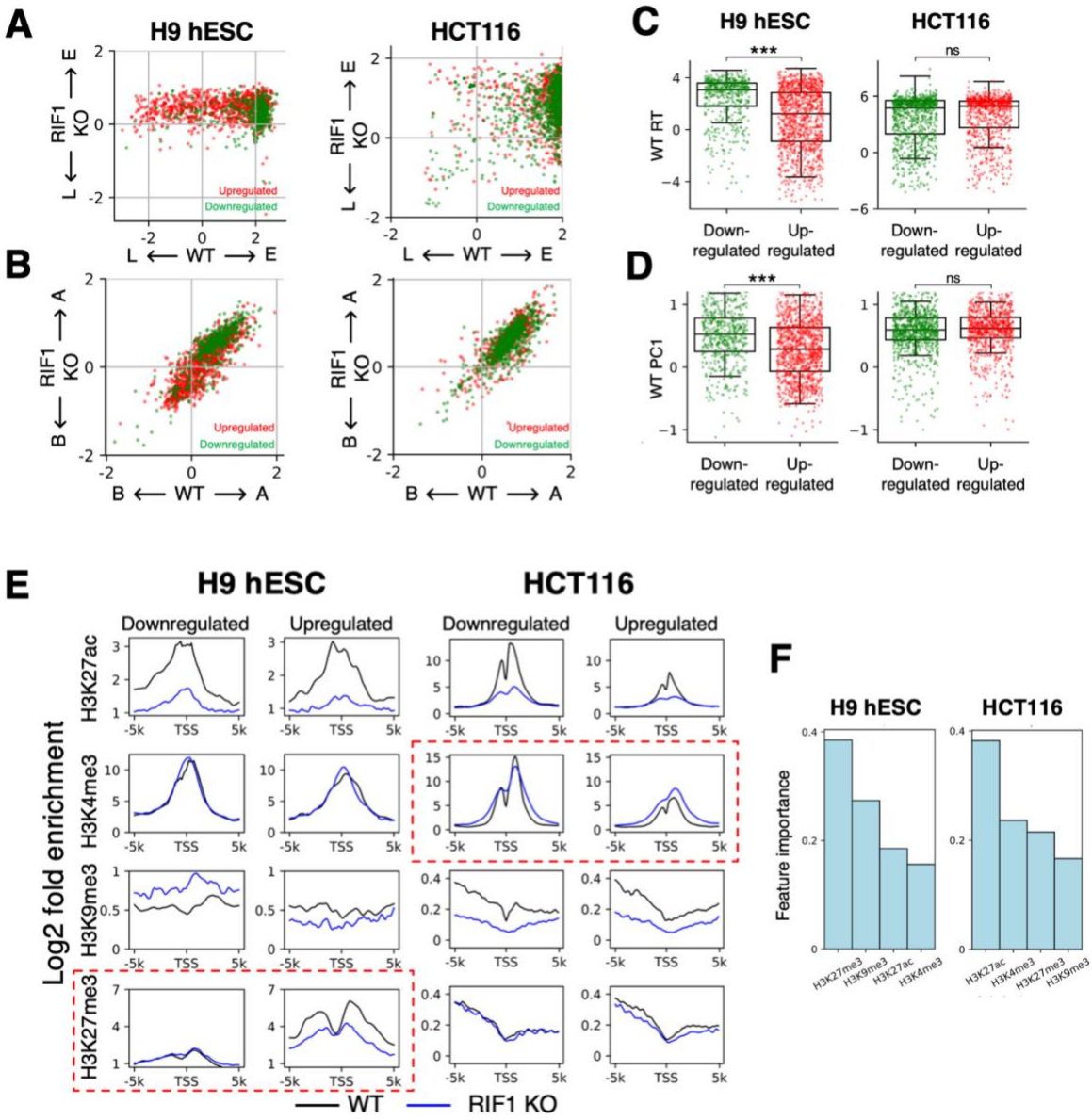

A
H9 hESC
HCT116
RIF1 KO
L ⟵ WT ⟶ E
Upregulated
Downregulated
B
B ⟵ WT ⟶ A
C
WT RT
***
ns
Down-regulated
Up-regulated
D
WT PC1
E
Log2 fold enrichment
H3K27ac
H3K4me3
H3K9me3
H3K27me3
-5k
TSS
5k
WT
RIF1 KO
F
Feature importance
H3K27me3
H3K9me3
H3K27ac
H3K4me3
H3K27ac
H3K4me3
H3K27me3
H3K9me3

Fig. S19. Histone modifications, not RT or compartments, correlate with gene expression alterations in RIF1 KO.
(A) Scatterplots of RT values at upregulated (red) and downregulated (green) genes in WT (x axis) and RIF1 KO (y axis) H9 hESCs and HCT116. Loss of late timing was not correlated with increased gene expression nor was loss of earlier timing correlated with decreased expression. (B) Scatterplot of PC1 eigenvectors at upregulated (red) and downregulated (green) genes in WT (x axis) and RIF1 KO (y axis) H9 hESCs and HCT116. Shifts from the B to A compartment were not correlated with increased gene expression nor were shifts from A to B correlated with downregulated expression. (C) Box plots showing distribution of WT RT values at down (green) and up (red) regulated genes in both cell lines. Each point represents an individual gene. (D) Box plots of distribution of WT PC1 eigenvector values at down (green) and up (red) regulated genes in both cell lines. Each point represents an individual gene. In H9 hESCs earlier replicating genes with strong A compartment association in WT cells were likely to be downregulated while later replicating genes with weak A compartment association in WT cells were likely to be upregulated upon RIF1 KO. In HCT116 cells, only genes in early replicating, A compartment chromatin were differentially expressed (both up and downregulated) upon RIF1 KO. (E) Average fold enrichment pile up line plots of indicated histone modification ±5kb around TSS of either up (right) or down (left) regulated genes in WT (black) and RIF1 KO (blue) H9 hESCs and HCT116. Plots highlighted with a red dashed square indicate histone marks that change in the same direction as gene expression changes from WT to RIF1 KO. In H9 hESCs H3K27me3 was high in WT cells around TSSs of upregulated genes and decreased upon RIF1 KO. In HCT116, changes in H3K4me3 levels correlate with expression changes. (F) Feature importance of random forest regression modelling of histone modification in predicting direction of gene expression changes in both RIF1 KO cell lines. Changes to heterochromatic marks were most correlated with changing gene expression in H9 hESCs while euchromatic marks were most correlated in HCT116.

Ensembl ID	Gene name	GO Gene Class
ENSG00000002587	**HS3ST1**	**Sulfotransferase**
ENSG00000107249	**GLIS3**	**Zinc Finger Protein**
ENSG00000162772	**ATF3**	**Basic Leucine Zipper Transcription Factor**
ENSG00000169174	**PCSK9**	**Serine Protease**
ENSG00000263934	**SNORD3A**	**Small Nucleolar RNA**
ENSG00000264940	**SNORD3C**	**Small Nucleolar RNA**

Table S1.

All significantly affected genes shared between RIF1 KO/KD conditions.

Plasmid name	Addgene number
pAAV-Ef1a-mCherry-IRES-Cre	**55632**
px458	**48138**
px330	**42230**

Table S2.

Information table of plasmids used in this study.

CRISPR use	5’ to 3’ target sequence
H9 RIF1 KO	**CCTGCTTTTTAAGTCGTATGACT**
	CCTCGGCTGTACAAAGTTTTAAA
H9 RIF1 eGFP KI	**CCGACATGACGGCCAGGGGTCAG**
HCT RIF1 KI	**CCCATGGAGGGCAGACTGACGCT**
	CCTAAAGTTGCCTAGGGGCACTT

Table S3.

Information table of CRISPR sgRNA target sequences used in this study. PAM sequences underlined.

Primer use	5' to 3' primer sequence
HCT RIF1 KI	**T*C*AGGGTGGCCGACATGAC**
	A*C*CTCACCTGGTCAGAGTCAG
H9 RIF1 eGFP KI	**CGGTCACTGGATTTTCTCCTCTTCCGGTTCGGGCCTCAGG GTGGCCGACAatggtgagcaagggcgag**
	TTCCAAAGTCTCCAACAGCGGCGCGAGGGGGCTCTGACC CCTGGCCGTCATgctaccgctgccgctacccttgtacagctcgtccatgcc

Table S4.

Information table of PCR primers used in this study. Asterixis (*) indicate phosphorothioate bonds.

	Supplier	Product Number	Research Resource Identifier
RIF1	**Bethyl Labs**	**A300-568A**	**RRID:AB_669806**
H3K9me3	**Abcam**	**ab8898**	**RRID:AB_306848**
H3K27me3	**Millipore**	**07-449**	**RRID:AB_310624**
H3K27ac	**Abcam**	**ab4729**	**RRID:AB_2118291**
H3K4me3	**Abcam**	**ab8580**	**RRID:AB_306649**
GFP	**MBL**	**598**	**RRID:AB_591819**
BrdU	**BD Biosciences**	**555627**	**RRID:AB_395993**
Rad21	**Abcam**	**ab992**	**RRID:AB_21176601**

Table S5.

Information table of antibodies used in this study.

Data S1. (separate file)

aba5545_Suppl. Excel_seq1_v1

Data S2. (separate file)

aba5545_Suppl. Excel_seq2_v1

Data S3. (separate file)

aba5545_Suppl. Excel_seq3_v1