RIF1 regulates replication origin activity and early replication timing in B cells

Daniel Malzl[1,6], Mihaela Peycheva[1,6], Ali Rahjouei[2], Stefano Gnan[3], Kyle N. Klein[4], Mariia Nazarova[1], Ursula Schoeberl[1], David M. Gilbert[4], Sara C. B. Buonomo[3], Michela Di Virgilio[2], Tobias Neumann[1,5,7] and Rushad Pavri[1,7]

1. Research Institute of Molecular Pathology (IMP), Vienna Biocenter, 1030 Vienna, Austria
2. Max-Delbruck Center for Molecular Medicine in the Helmholtz Association (MDC), 13125 Berlin, Germany
3. Institute of Cell Biology, University of Edinburgh, Edinburgh EH9 3FF, UK
4. Dept. of Biological Science, Florida State University, Tallahassee FL 32306, USA
5. Present address: Quantro Therapeutics, Vienna Biocenter, 1030 Vienna, Austria.
6. Equal contribution
7. Corresponding authors (tobias.neumann@imp.ac.at, rushad.pavri@imp.ac.at)

ABSTRACT

The mammalian DNA replication timing (RT) program is crucial for the proper functioning and integrity of the genome. The best-known mechanism for controlling RT is the suppression of late origins of replication in heterochromatin by RIF1. Here, we report that in antigen-activated, hypermutating B lymphocytes, RIF1 binds predominantly to early-replicating active chromatin, regulates early origin firing and promotes early replication. RIF1 has a minor role in gene expression and genome organization in B cells. Furthermore, we find that RIF1 functions in a complementary and non-epistatic manner with minichromosome maintenance (MCM) proteins to establish early RT signatures genome-wide and, specifically, to ensure the early replication of highly transcribed genes. These findings reveal a new layer of regulation within the B cell RT program, driven by the coordinated activity of RIF1 and MCM proteins.

51 **INTRODUCTION**

52 The faithful and timely replication of the genome is essential for inheriting genetic information and
53 avoiding chromosomal abnormalities. To ensure this, large metazoan genomes initiate replication
54 from several discrete loci, termed origins of replication. Origins are not activated simultaneously
55 across the genome, but rather, in an asynchronous manner referred to as the DNA replication timing
56 (RT) program[1]. A hallmark of the RT program is that genomic A compartments enriched in
57 transcriptionally active genes generally replicate earlier in S phase whereas B compartments
58 harboring silent heterochromatin typically replicate later in S phase[2,3]. Deregulation of RT has been
59 correlated with defects in chromosome condensation, sister chromatid cohesion, gene expression and
60 genome instability [4-6]. Altered RT can disrupt the distribution of active and repressive epigenetic
61 marks causing major alterations in genome architecture[7]. In genetic diseases and cancer, defects in
62 RT have been correlated with deleterious chromosomal translocations[8-11]. Recently, RT has been
63 directly implicated in the biogenesis of oncogenic translocations found in B cell lymphomas and other
64 leukemias[12]. Yet, despite its key role in maintaining cellular physiology and genome integrity, our
65 understanding of the mechanisms regulating RT remains incomplete.

66 Origins are specified in the late G2/M and early G1 phases of the cell cycle via the loading of the
67 origin recognition complex (ORC) and associated factors[13-17]. The subsequent recruitment of the
68 hexameric ring-shaped MCM complex helicase (MCM2-7) licenses these sites for activation[14],
69 resulting in the formation of the pre-replication complex (pre-RC)[13-17]. In S phase, a subset of
70 licensed origins is activated by a set of proteins collectively called replication firing factors, resulting in
71 the initiation of replication[13-16]. A subset of licensed, MCM-bound origins can be activated by firing
72 factors at any given point in S phase[18]. This has led to the idea that firing factors must be recycled to
73 activate subsequent sets of origins later in S phase, thus ensuring timely completion of replication. In
74 support, overexpression of firing factors has been shown to advance the RT of late-replicating
75 chromatin[19-21]. These studies imply that the time of recruitment of the firing factors to a licensed origin
76 determines its RT [22]. In essence, the probability of origin activation in S phase determines the RT of a
77 region, meaning that early RT domains harbor origins that typically fire earlier in S phase whereas late
78 RT domains contain origins that tend to fire later in S phase.

79 The best-studied mechanism of RT control is the suppression of late origin firing by the
80 multifunctional protein, RIF1[23-25]. RIF1 associates with late-replicating chromatin in large, mega-base
81 domains called RIF1-associated domains (RADs) and represses origin firing via protein phosphatase
82 1 (PP1)-mediated dephosphorylation of MCM4[26-28]. Ablation of RIF1 (*Rif1*$^{-/-}$) in human and mouse
83 embryonic stem cells (hESCs and mESCs) and in primary mouse embryonic fibroblasts [23] resulted in
84 a genome-wide loss of early and late RT domain distinction associated with alterations in epigenetic
85 marks and genome compartmentalization[7,27,28]. However, in some *Rif1*$^{-/-}$ cell lines, this RT phenotype
86 was considerably weaker [7], suggesting the existence of additional modes of RT control.

87 We recently showed that in MCM-depleted CH12 cells, a murine B cell line, the RT program was
88 globally deregulated without major changes in transcription or genome architecture[12]. Since RIF1 is
89 the only other factor whose loss could lead to such a major phenotype[7,23,27], we investigated the role
90 of RIF1 in in the B cell RT program. We report the surprising findings that RIF1 is predominantly
91 bound to active chromatin in B cells, regulates early origin activity in these regions and promotes their
92 early replication. In addition, we find that RIF1 acts in concert with MCM complexes to drive early
93 replication of highly transcribed regions in B cells. In sum, our study reveals an additional regulatory
94 layer within the global RT program and a new role for RIF1 in promoting early replication.

95

96 **RESULTS**

97 **RIF1 regulates early replication in B cells**

98 To measure RT, we performed Repli-seq from early (E) and late (L) S phase fractions in normal and
99 *Rif1*$^{-/-}$ CH12 B cells[29,30] (Fig. S1A) MCM complexes were depleted by infection with lentiviruses
100 expressing short hairpin RNAs (shRNAs) targeting Mcm6 (shMcm6), and an shRNA against LacZ
101 (shLacZ) served as a control for infection and shRNA expression, as described previously[12,31] (Fig.
102 S1B). RT was calculated as the (log2) E/L ratio[32] (Fig. 1A). RT values showed the expected bimodal
103 distribution in shLacZ cells reflecting distinct early (E; positive RT) and late (L; negative RT) domains
104 (Fig. 1A), which was also reflected in the genomic RT profiles (Fig. 1B). In shMcm6 cells, this
105 distinction was globally weakened with all early and late RT values approaching zero (Fig. 1A-B). In
106 *Rif1*$^{-/-}$ shLacZ cells, many early-replicating domains underwent a delay in replication (Fig. 1B). In

107 comparison, late RT domains showed a mixed phenotype in *Rif1*$^{-/-}$ shLacZ cells with some, typically
108 smaller, domains showing advanced replication signatures, and other, typically larger, domains
109 undergoing delayed replication relative to shLacZ cells (Fig. 1A-B). Direct comparison of RT in 20 kb
110 genomic bins between the different conditions showed that most early- and late-replicating bins were
111 strongly shifted towards zero in shMcm6 cells indicative of a deregulation of the RT program (Fig.
112 1C). In *Rif1*$^{-/-}$ shLacZ cells, most early-replicating regions underwent delayed replication to varying
113 degrees (note the red stripe in Fig. 1C), albeit to a lesser extent than in shMcm6 cells. Although late-
114 replicating bins were mildly affected in *Rif1*$^{-/-}$ shLacZ cells, many of these bins underwent slightly
115 delayed RT (Fig. 1C). The latter observation is reminiscent of reports in other *Rif1*$^{-/-}$ cell lines where
116 the RT of many late-replicating domains that were not associated with RIF1 was unaffected by the
117 loss of RIF1 [7].

118 We also performed Repli-seq in a more physiological system, namely, primary activated splenic B
119 cells from *Rif1*$^{+/+}$, RIF1 heterozygous (*Rif1*$^{-/+}$) and *Rif1*$^{-/-}$ mice[33] (Fig. S1C). WT primary B cells had
120 relatively fewer early replicating domains and more mid-replicating domains than shLacZ CH12 cells,
121 resulting in different RT density profiles (Fig. 1D, compare with Fig. 1A). Strikingly, we observed a
122 nearly exclusive effect on early-replicating genomic bins in heterozygous *Rif1*$^{-/+}$ cells indicating that
123 reduced levels of RIF1 protein was sufficient to delay replication of early-replicating regions without
124 majorly affecting late replicating ones (Fig. 1 D-E). This phenotype was exacerbated in homozygous
125 *Rif1*$^{-/-}$ cells as seen by the further delay in replication of early RT domains, which was accompanied
126 by the advanced replication of several late RT domains albeit to varying degrees (Fig. 1 B and D-E).
127 These results, especially from heterozygous cells, highlight a new role for RIF1 in regulating early
128 replication in B cells.

129 To systematically measure changes in RT upon loss of RIF1, we generated three RT states, early
130 (E), middle (M) and late (L), in shLacZ CH12 cells using a Hidden Markov Model (HMM) to segment
131 the genome into 20 kb bins based on their RT value (Fig. 1F). We split the M state at the zero RT
132 value into early-like (E-like) having positive RT values and late-like (L-like) having negative RT values,
133 which resulted in four RT classes (Fig. 1F). To allow direct comparison of RT changes between
134 conditions, we created density plots for each of the four RT states in shLacZ cells displaying the RT of
135 all other experimental conditions in those genomic bins. The results showed that E and E-like bins in

136 shLacZ replicated later in *Rif1*$^{-/-}$ cells, but, importantly, that this was not accompanied by
137 commensurate advances in the RT of L bins suggesting that loss of RIF1 does not globally deregulate
138 RT (Fig. 1G). Indeed, most L bins retained their RT values with only some bins showing advanced RT
139 and some others showing further delays in RT relative to shLacZ cells (Fig. 1G). In contrast, in
140 shMcm6 cells, all shLacZ RT bins shifted towards zero RT values implying that these cells have
141 undergone a global deregulation of the RT program (Fig. 1G). The same HMM-based analysis in
142 primary, splenic B cells revealed that in *Rif1*$^{-/+}$ heterozygotes, there was a strong delay in the RT of E
143 bins associated with a much weaker advance in the RT of L bins and virtually no changes in L-like
144 bins, which further demonstrates that RIF1 primarily regulates early replication in B cells (Fig. 1H-I).
145 These changes were exacerbated in *Rif1*$^{-/-}$ B cells as seen by the further delay in the RT of E bins
146 which as accompanied by the advanced RT of L bins relative to *Rif1*$^{-/+}$ cells (Fig. 1H-I).

147 We conclude that RIF1 functions as modulator of early replication in B cells rather than as a global
148 regulator of RT.

149 **RIF1 plays a minor role in gene expression and genome compartmentalization in B cells**

150 In comparison to the RT phenotype of *Rif1*$^{-/-}$ CH12 cells, where nearly all early replicating domains
151 were affected, only 55 and 102 genes were two-fold up- or downregulated in these cells (Fig. S2A and
152 Table S1). A gene ontology (GO) analysis revealed no enrichment of pathways associated with DNA
153 replication or gene expression (Fig. S2B).

154 RIF1 has been implicated in regulating genome compartmentalization, in part, through
155 deregulation of RT [7]. However, in *Rif1*$^{-/-}$ shLacZ CH12 cells, Hi-C analysis revealed no gross changes
156 in compartment profiles based on evaluation of Hi-C heatmaps where compartments manifest as the
157 checkerboard patterns off the diagonal (Fig. 2A). We quantified the compartment signals based on the
158 first principal component (PC1) of the Hi-C contact matrix where positive PC1 values denote active,
159 early-replicating A compartments and negative PC1 values denote silent, late-replicating B
160 compartments (Fig. 2B-C)[34]. Comparative PC1 analysis showed that although the majority of the bins
161 in *Rif1*$^{-/-}$ shLacZ cells underwent minor changes in their PC1 values relative to shLacZ cells, a few
162 bins did shift substantially, including PC1 sign flips in both directions, which gave a more dispersed
163 appearance to the distribution relative to control cells (Fig. 2B, right). Visualization of PC1 profiles in

164 various genomic regions confirmed that small changes in PC1 occurred in many locations in *Rif1*$^{-/-}$
165 shLacZ cells, but importantly, that these did not always correlate with the changes in RT (Fig. 2C).
166 However, the magnitude of PC1 changes and the frequency of compartment switching (PC1 sign
167 flips) is considerably milder than that reported in other cell types where loss of RIF1 led to substantial
168 compartment switching[7,27]. In comparison, PC1 values in shMcm6 cells were comparable to those in
169 shLacZ cells (Fig. 2B, left and Fig. 2C)[12].

170 We also generated PC1 versus RT density contour plots which allowed us to simultaneously
171 compare the changes in compartmentalization and RT within the same set of genomic 20 kb bins.
172 The density contour distribution profile in *Rif1*$^{-/-}$shLacZ cells showed a marked shift of early RT bins
173 towards delayed RT but without major changes in their PC1 values (Fig. 2D). This implies that the
174 roles of RIF1 in RT and genome architecture are separable in B cells, as they are in other cells[7,27].
175 However, *Rif1*$^{-/-}$ shLacZ cells showed a gain of A-B inter-compartmental interactions with a
176 corresponding decrease in A-A and B-B contacts (Fig. 2E-F). This effect was clearly manifested in
177 fold-change saddle plots where contacts between regions of highest PC1 (early-replicating) and
178 lowest PC1 (late-replicating) were increased in *Rif1*$^{-/-}$ shLacZ cells (Fig. 2F). Thus, RIF1 is important
179 for maintaining normal compartmentalization in B cells by preventing the mixing of A and B
180 compartments. Importantly, this compartmentalization phenotype was observed in mESCs and
181 hESCs where, in contrast to B cells, loss of RIF1 caused a severe deregulation of the RT program[7,27].
182 This suggests that the changes in compartmentalization in *Rif1*$^{-/-}$ cells are conserved between
183 diverse cell lineages, but that these are unrelated to the changes in RT. Moreover, in shMcm6 cells,
184 where RT is globally disrupted to a similar degree as in *Rif1*$^{-/-}$ mESCs and hESCs [7,12,27], we observed
185 relatively minor changes in compartmentalization, as seen by the mild increase in A-A contacts and a
186 similar decrease in B-B contacts (Fig. 2E-F).

187 In sum, the uncoupling of RT from genome organization in B cells and other cells [7,12,27] leads us to
188 conclude that although RIF1 contributes to the normal spatial separation of A and B compartments in
189 B cells, it's role in regulating early RT is unlikely to be directly linked to these structural changes.

190 **RIF1 is predominantly located in early-replicating transcribed chromatin in B cells**

191 We next investigated the genomic occupancy of RIF1 in B cells. We performed chromatin
192 immunoprecipitation (ChIP) from primary, mature splenic B cells derived from the *Rif1*$^{FH/FH}$ mouse line

193 wherein RIF1 is endogenously tagged at its C terminus with the Flag and Hemagluttinin (HA)
194 epitopes[23,28]. RIF1 chromatin occupancy was determined via ChIP sequencing (ChIP-seq) with an
195 anti-HA antibody in *Rif1*$^{FH/FH}$ and *Rif1*$^{+/+}$ cells (Fig. S3A).

196 Visual analysis on the genome browser revealed RIF1 binding predominantly in early replicating
197 domains in B cells (Fig. 3A). Comparison with ChIP-seq profiles from *Rif1*$^{FH/FH}$ mouse embryonic
198 fibroblasts (MEFs)[28] showed that while RIF1 localized to broad domains in late-replicating regions in
199 MEFs, in B cells, enrichments were mostly in early-replicating domains (Fig. 3B). We next identified
200 peaks of RIF1 occupancy [35] as well as RADs[36] in both B cells and MEFs. Peaks in both cell types
201 were enriched in early-replicating regions (Fig. 3C-D). However, although RADs in MEFs were mostly
202 late-replicating, as described[28], in B cells, they were largely early-replicating (Fig. 3C-D).

203 We next determined the overlap of RIF1 peaks with nascent transcription measured by precision
204 run-on sequencing (PRO-seq)[37], chromatin accessibility, measured by assay for transposase-
205 accessible chromatin (ATAC-seq)[38], and ChIP-seq of histone H3 acetylated at lysine 27 (H3K27Ac), a
206 mark of active transcription start sites (TSSs) and enhancers. The resulting heatmaps, centered on
207 the RIF1 peak summits, revealed an enrichment of RIF1 in transcribed, H3K27ac-rich accessible
208 chromatin (Fig. 3E). In agreement, ~70% of RIF1 peaks and ~50% of RADs overlapped TSSs or gene
209 bodies (Fig. 3F), and a metagene analysis showed that RIF1 was relatively enriched at TSSs (Fig.
210 3G).

211 We conclude that RIF1 predominantly occupies early-replicating, transcribed chromatin in B cells,
212 suggesting that RIF1 promotes their early replication via direct association.

213 **RIF1 regulates the activity of a subset of origins of replication in active chromatin**

214 The localization of RIF1 in active chromatin coupled with the delayed replication of these domains in
215 *Rif1*$^{-/-}$ cells led us to investigate whether RIF1 played a role in promoting origin firing in activated B
216 cells. To address this, we measured origin activity with short nascent strand sequencing (SNS-seq) in
217 WT and *Rif1*$^{-/-}$ CH12 cells (Fig. S3B). SNS-seq maps the location and relative usage of origins
218 (termed replication initiation sites; ISs) in a population of cells by quantifying the levels of nascent
219 leading strands[39,40]. As described in our previous work [12], ISs and initiation zones were defined based

220 on peak-calling and peak-clustering, respectively, and only ISs within initiation zones were used for
221 analysis.

222 We generated heatmaps of IS read densities in a 4 kb window centered at the IS peak summit
223 and ordered by the IS density in WT cells (Fig. 4B). To visualize how IS density correlated with
224 various chromatin features in the genomic neighborhood, we also generated heatmaps showing the
225 densities of DNase hypersensitive sites (DHSs) which mark accessible chromatin at active promoters
226 and enhancers, H3K36me3, which is enriched in the bodies of transcriptionally active genes, and the
227 repressive mark, H3K9me3 (Fig. 4B). In WT cells, the most active ISs were embedded in active
228 chromatin and gene bodies whereas the least active ISs were associated with H3K9me3 (Fig. 4B). In
229 *Rif1*$^{-/-}$ cells, origin activity was repressed in active chromatin but increased in H3K9me3 domains (Fig.
230 4B). This phenotype is reminiscent of what we previously reported in shMcm6 B cells [12]. Importantly,
231 however, only ~5% of ISs are deregulated (1.5-fold) upon loss of RIF1 compared to ~38% in shMcm6
232 cells (Fig. S3C), which correlates with the magnitude of their respective RT phenotypes (Fig. 1).

233 To quantify the changes in origin efficiency, we calculated fold-changes of ISs (log2 WT/ *Rif1*$^{-/-}$),
234 ranked them from highest to lowest and created five equal classes (quintiles) such that class 1
235 contained the most downregulated ISs in *Rif1*$^{-/-}$ cells and class 5 contained the most upregulated ISs
236 in *Rif1*$^{-/-}$cells. We next generated violin plots displaying the density of ISs (Fig. 4C) or RT values (Fig.
237 4D) in WT and *Rif1*$^{-/-}$ cells within each IS class. The results showed that in WT cells, class 1 ISs were
238 the most active (Fig. 4C) and were mostly early-replicating (Fig. 4D) whereas class 5 ISs were the
239 least active with the majority being late-replicating (Fig. 4C-D). Importantly, class 5 IS densities in
240 *Rif1*$^{-/-}$ cells were comparable to class 1 IS densities in WT cells, suggesting that the upregulated
241 origins in *Rif1*$^{-/-}$ cells (class 5), which are normally the weakest, fire at similar efficiencies as the most
242 active origins in WT cells (class 1) (Fig. 4C).

243 We conclude that RIF1 is required for the optimal activity of a subset of early replicating origins in
244 B cells.

245 **RIF1 and MCM complexes act in a complementary manner to regulate the B cell RT program**

246 The differing RT phenotypes in *Rif1*$^{-/-}$ and shMcm6 cells led us to hypothesize that they may function
247 in a non-epistatic manner to drive early replication in B cells. To address this, we infected *Rif1*$^{-/-}$ cells

248 with lentiviruses expressing shMcm6 (*Rif1*$^{-/-}$ shMcm6) or shLacZ (*Rif1*$^{-/-}$ shLacZ) (Fig. S5A) and
249 performed Repli-seq (Fig. S4A). Of note, the converse experiment, that is, depletion of RIF1 in
250 shMcm6 cells was precluded by the fact that the viability of shMcm6 cells was severely compromised
251 upon viral infection with *Rif1*-specific sgRNAs or shRNAs, in line with the observation that cells with
252 limiting MCM proteins are sensitive to stress[41-43].

253 Repli-seq revealed that there was a further loss of early and late RT domain distinction in *Rif1*$^{-/-}$
254 shMcm6 cells compared to *Rif1*$^{-/-}$ shLacZ cells or shMcm6 cells, with a shift of both E and L RT
255 values towards zero (Fig. 5A). An HMM-based classification of RT values, as in Fig. 1D, showed that
256 E bins replicated later, and L bins replicated earlier in *Rif1*$^{-/-}$ shMcm6 compared to shMcm6 cells (Fig.
257 5B). However,

258 Direct comparison of RT values in 20 kb genomic bins showed that this exacerbation of the RT
259 phenotype in *Rif1*$^{-/-}$ shMcm6 relative to *Rif1*$^{-/-}$ shLacZ and shMcm6 cells was observed in most of the
260 RT bins (Fig. 5C). Additionally, in *Rif1*$^{-/-}$ shMcm6 cells, many large early-replicating domains showed
261 considerable fluctuation in the RT values resulting in a highly fragmented RT profile (Fig. 5D).
262 However, this fragmentation was considerably lower in large late-replicating domains (Fig. 5D). The
263 RT profiles in *Rif1*$^{-/-}$ shMcm6 cells were also marked by switching of RT signatures (both E to L and L
264 to E) and loss of clear domain boundaries (Fig. 5D). These changes, and especially the global
265 fluctuations of early RT values, result in *Rif1*$^{-/-}$ shMcm6 cells acquiring an RT signature distinct from
266 shMcm6 cells (Fig. 5A, C). In sum, these findings reveal an additive effect of MCM depletion and loss
267 of RIF1 on RT in B cells.

268 Despite the changes in RT, gene expression (Fig. S4B-C and Table S2), compartment identity
269 (Fig. 5D and Fig. S4D) and compartment strength (Fig. S4E-F) were not majorly altered in *Rif1*$^{-/-}$
270 shMcm6 cells relative to *Rif1*$^{-/-}$ shLacZ cells, although slight gains in A-B, B-A and A-A interactions
271 were observed, akin to the changes seen in shMcm6 cells (Fig. S4F, compare with Fig. 2F). However,
272 the major changes in RT led to a distinct pattern of PC1 versus RT profiles in in *Rif1*$^{-/-}$ shMcm6 cells
273 relative to in *Rif1*$^{-/-}$ shLacZ cells (Fig. S4G). Thus, the deregulation of the RT program by MCM
274 depletion does not majorly impact genome architecture in normal (Fig. 2) or in *Rif1*$^{-/-}$ cells.

275 To determine whether the increased deregulation of RT in *Rif1*$^{-/-}$ shMcm6 cells was associated
276 with changes in underlying replication origin activity, we performed SNS-seq in *Rif1*$^{-/-}$ shLacZ and
277 *Rif1*$^{-/-}$ shMcm6 cells (Fig. 6A). We further classified ISs based on fold-changes
278 (*Rif1*$^{-/-}$ shLacZ/*Rif1*$^{-/-}$ shMcm6) into quintiles and analyzed the distribution of IS densities (Fig. 6B)
279 and RT values (Fig. 6C) within the IS classes. The most downregulated ISs (class 1) were the most
280 active in *Rif1*$^{-/-}$ shLacZ and were predominantly early-replicating, whereas the most upregulated ISs
281 (class 5) were the least active in *Rif1*$^{-/-}$ shMcm6 cells and were mostly late-replicating (Fig. 6B-C).
282 This is consistent with the phenotype we reported previously in shMcm6 cells[12]. Indeed, SNS-seq
283 profiles showed a characteristic downregulation of origin activity in A compartments accompanied by
284 upregulation in B compartments in both shMcm6 and *Rif1*$^{-/-}$ shMcm6 cells (Fig. 6D). We infer that the
285 additional delay in early replication in *Rif1*$^{-/-}$ shMcm6 cells relative to *Rif1*$^{-/-}$ shLacZ cells is likely due
286 to the further decrease in the efficiency of early-firing origins.

287 We conclude that MCM proteins and RIF1 regulate RT in a complementary and non-epistatic
288 manner in B cells.

289 **Within early RT domains, the RT of highly transcribed regions is most sensitive to the**
290 **depletion of MCM complexes and, to a weaker extent, the loss of RIF1**

291 A closer examination of the RT profiles revealed that that the extensive fragmentation seen in *Rif1*$^{-/-}$
292 shMcm6 cells was also visible to a lesser extent in shMcm6 and to the weakest extent in *Rif1*$^{-/-}$
293 shLacZ cells (Fig. 5C-D). Furthermore, we observed a high degree of similarity in the fragmentation
294 patterns between the different conditions. This suggested that these were not random fluctuations of
295 the RT signals but reflected an underlying mechanism supporting early replication that is reliant on
296 MCM proteins and RIF1. We therefore investigated whether the levels of nascent transcription within
297 genomic RT bins could explain these altered RT profiles. This reasoning was based on the
298 observations that fragmentation was typically seen in early RT domains where most transcribed
299 genes are located, and secondly, that in *Rif1*$^{-/-}$, shMcm6 and *Rif1*$^{-/-}$ shMcm6 cells, early origins in
300 active chromatin were downregulated despite there being no major changes in transcription.

301 To address this, we divided the genome into 20 kb bins and extracted all early-replicating (E)
302 bins in shLacZ cells using the HMM-based approach described in Fig. 1C. Within this group, we

303 defined transcribed bins as those having a PRO-seq density (RPM) ≥ 10 and these were divided into
304 three equal sub-groups (tertiles) based on their RPM values, termed High (199-3358 RPM), Medium
305 (78-198.9 RPM) and Low (10-77.9 RPM) (Fig. S4H). The remaining early-replicating bins were termed
306 Untranscribed (0-9.9 RPM) (Fig. S4H). To compare changes in RT values, we generated RT
307 heatmaps for the four transcription-based groups and ranked each of them by decreasing WT PRO-
308 seq read density such that, effectively, the entire set of bins were ranked from highest to lowest PRO-
309 seq density in shLacZ cells (Fig. 7A). In addition, we generated density plots to visualize the
310 differences in the distribution of RT values between the four groups in all experimental conditions (Fig.
311 7B).

312 In shLacZ cells, all four transcription-based groups showed nearly identical distributions of RT
313 values, indicating that transcribed and non-transcribed regions normally have similar probabilities of
314 origin activation and early replication (Fig. A-B). The High and Medium groups appeared very similar
315 in terms of the RT values between bins (Fig. 7A) and the overall distribution profiles (Fig. 7B).
316 However, in shMcm6 and *Rif1*$^{-/-}$ shMcm6 cells, the High and Medium transcribed bins showed the
317 strongest delays in RT whereas the Untranscribed group showed the least delays (Fig. 7A-B). These
318 differences were observed across all bins (Fig. 7A). The shift in RT between the High and
319 Untranscribed groups was ~2-fold (log2 RT ~1), as measured by the distance between the modes of
320 the distributions (marked by dotted lines in Fig. 7B). The lowly transcribed regions had an
321 intermediate distribution of RT values between the High/Medium and Untranscribed groups in both
322 shMcm6 and *Rif1*$^{-/-}$ shMcm6 cells (Fig. 7A-B). Thus, the strongest delays in RT are associated with
323 regions harboring higher levels of transcription. In *Rif1*$^{-/-}$ shLacZ cells, however, all groups had a
324 broad spread of RT values with similar distribution profiles (Fig. 7A-B). However, a closer inspection
325 of the heatmap and the modes of the distributions revealed a slightly greater delay in RT for the high
326 and medium transcribed regions compared to the lowly transcribed and non-transcribed bins (Fig.
327 7B).

328 Taken together, these results suggest that the RT of transcribed regions, and hence the
329 underlying activity of early origins, is more sensitive to the reduction of MCM proteins than to the loss
330 of RIF1, which correlates with the magnitude of the fragmentation observed in the early RT domains
331 (Fig. 5C-D). We conclude that the fragmented appearance of early RT profiles is due to underlying

332 transcription, with stronger delays correlating with higher levels of transcription. These results suggest
333 that transcriptional strength negatively regulates the probability of early origin firing at active genes,
334 and that this is normally overcome by the presence of a normal complement of MCM proteins and, to
335 a lesser extent, RIF1.

336

337 **DISCUSSION**

338 In this study, we show that the B cell RT program is driven by the complementary and non-epistatic
339 functions of RIF1 and MCM complexes. We identify a new role for RIF1, namely, the regulation of
340 early replication in B cells. RIF1 executes this function by directly associating with early-replicating
341 chromatin and transcribed regions and modulating the activity of a subset of early origins of
342 replication. Finally, we show that transcriptional strength impedes origin establishment and early RT,
343 and that overcoming this requires the normal quantum of MCM proteins with minor contribution from
344 RIF1.

345 How might transcription negatively impact the efficiency of replication origins? There is evidence
346 from *in vitro* studies that RNA polymerase (RNAP) complexes can push MCM complexes along DNA
347 [44,45]. Indeed, MCM and ORC proteins were found to be relatively depleted in gene bodies[46] and
348 Okazaki fragment sequencing (OK-seq) studies have reported that replication initiation zones occur
349 largely in the intergenic regions flanking transcribed genes[47]. A recent single molecule imaging study,
350 where replication-transcription encounters were reconstituted using purified proteins, may be
351 instructive in this regard [45]. This study found that T7 RNAP could efficiently push DNA-bound ORC,
352 OCCM (an intermediate of the pre-RC consisting of ORC, Cdc6, Cdt1 and a single MCM hexamer)
353 and MCM double hexamers (the configuration within the fully assembled pre-RC). However, whereas
354 OCCM and double hexamers were rarely ejected by RNAP, ORC alone was frequently evicted by
355 RNAP. Moreover, most of the ORC molecules repositioned by RNAP were unstable [45]. Since ORC
356 loading is the first step in origin licensing [17], the inference is that the labile binding of ORC makes
357 licensing in transcribed regions inherently less efficient than in non-transcribed regions, but that this is
358 overcome by the association of MCM complexes, which minimizes the loss of single ORC complexes
359 by RNAP.

360 Importantly, our data show that, normally, the distribution of early RT values is similar between
361 transcribed and non-transcribed regions. Hence, we propose that, in WT cells, the large pool of MCM
362 complexes ensures that the loading of ORC is rapidly followed by the assembly of the OCCM and
363 pre-RC such that the loss of single ORC complexes by RNAP is minimized. This ensures that both
364 transcribed and non-transcribed regions have a similar probability of early origin activation and early
365 replication in WT cells. When MCM complexes are limiting, licensing in both transcribed and non-
366 transcribed regions is reduced. However, given the inhibitory effect of transcription on ORC stability,
367 the formation of OCCM and pre-RC formation in highly transcribed regions will be inefficient, allowing
368 for more eviction of ORC by RNAP. This will reduce early origin licensing efficiencies in highly
369 transcribed chromatin to a greater degree than in non-transcribed chromatin, leading to our
370 observation of a higher probability of early replication in non-transcribed regions relative to transcribed
371 regions.

372 Although the mechanism by which RIF1 regulates early origin firing in B cells remains to be
373 deciphered, RIF1 can interact with MCM complexes in activated B cells [30]. Intriguingly, a previous
374 study in human cells reported a role for RIF1 in promoting origin licensing by protecting ORC1 from
375 phosphorylation-mediated degradation [48] . Thus, it is plausible that RIF1 prevents premature
376 disassembly of OCCM and pre-RC complexes in B cells. In addition, although RIF1 binds broadly
377 across early-replicating chromatin domains in B cells, ~70% of RIF1 peaks and ~50% of RADs in B
378 cells are in genes (Fig. 3). This suggests that the protective function of RIF1 in origin licensing may be
379 more pronounced in highly transcribed regions. According to this model, the loss of RIF1 would lead
380 to fewer licensed origins in highly transcribed regions compared to neighboring non-transcribed
381 regions within the same early-replicating domains, resulting in non-transcribed regions replicating
382 slightly later than highly transcribed regions, as we observe.

383 Given that RIF1 has not been previously implicated in regulating early replication, our study
384 raises the question of why B cells have functionally repurposed RIF1 in this manner. Antigen-
385 activated B cells are amongst the fastest proliferating cells in the body with a cell cycle duration of 6-8
386 h in vivo within germinal centers and 8-12 h in culture[49-51]. A unique feature of these cells is that they
387 undergo extremely high rates of somatic hypermutation at the highly transcribed immunoglobulin
388 genes, a necessary event in antibody maturation upon infection or vaccination. Genome instability is
389 further elevated by the fact that somatic hypermutation also occurs at many other transcribed loci,

390 including proto-oncogenes like *BCL6* and *MYC* [52-55], resulting in oncogenic translocations typical of
391 most mature B cell cancers [56-58]. Moreover, the DNA repair pathways activated by somatic
392 hypermutation result in single- and double-strand breaks as well as singe-strand patches, all of which
393 are impediments for replication[59,60]. Therefore, these B cells proliferate in a highly genotoxic
394 environment where replication stress is elevated. Under such conditions, it is plausible that B cells
395 have functionally repurposed RIF1 as an additional mechanism to safeguard its genome by enforcing
396 the early replication of active genes in order to ensure their faithful replication prior to cell division. It is
397 also possible that this function of RIF1 is masked in other cell types where the dominant mode of RT
398 regulation is that of RIF1-mediated suppression of origins in late-replicating chromatin[7,26,27].

399

400 **METHODS**

401 Cell culture

402 CH12 cells were maintained in RPMI medium with 10% fetal bovine serum (FBS), glutamine, sodium
403 pyruvate, Hepes and antibiotic/antimycotic mix. LentiX packaging cells were maintained in DMEM
404 medium with 10% FBS and antibiotics.

405 Transfections and infections

406 Transfection of LentiX cells with shRNAs against LacZ and Mcm6 followed by infection of CH12 cells
407 with lentiviral supernatants was done exactly as described in our previous study[31].

408 Sample preparation for replication timing (RT) analysis

409 Repli-seq was performed as previously described[32]. In brief, two million asynchronously dividing cells
410 were seeded and incubated with 100μM BrdU (Sigma) for 2h in a light protected environment to
411 maintain BrdU stability. Cells were fixed and incubated with a mix of RNase A (Invitrogen) and
412 propidium iodide (Sigma) for 30 min (light protected). For each sample, three fractions were sorted:
413 G1 phase, early S phase and late S phase cells, and for each fraction, two independent samples of
414 50,000 cells (technical replicates) were sorted on a Sony SH800S Cell Sorter. Sorted cells were lysed
415 with Proteinase K buffer overnight. Extracted DNA was sonicated for 9 min in a Diagenode Bioruptor
416 resulting in 100-500 bp DNA fragments as determined on an agarose gel. Sonicated DNA was

417 subjected to end-repair and adapter ligation using the NEBNext® Ultra™ II DNA Library Prep Kit
418 (NEB) following the NEB protocol. Adapter-ligated DNA was incubated with 25 μg/ml of anti-BrdU
419 antibody (BD Pharmingen) for 4h with rotation followed by incubation with 40 µg of anti-mouse IgG
420 antibody (Sigma) for 1h with rotation (light protected). DNA was precipitated via
421 Centrifugation at 16,000*g* for 5 min at 4°C. Pellet was resuspended in 200 µl of digestion buffer(for 50
422 ml of digestion buffer, combine 44 ml of autoclaved double-distilled water, 2.5 ml of 1M Tris-HCl, pH
423 8.0, 1 ml of 0.5M EDTA and 2.5 ml of 10% SDS) with freshly added 0.25 mg/ml proteinase K and
424 incubated overnight at 37°C. The immunoprecipitated DNA was used for Repli-qPCR or next
425 generation sequencing (Repli-seq). Libraries that were successfully validated by Repli-qPCR were
426 sequenced on an Illumina HiSeq 2500 machine (50bp, single-end). Up to 12 barcoded samples were
427 pooled per lane.

428 Chromatin immunoprecipitation for RIF1 occupancy in primary B cells

429 Chromatin immunoprecipitation was performed as described [28]. In summary, primary B cells were
430 isolated from Rif1$^{FH/FH}$ and C57BL/6 (Rif1$^{WT/WT}$) mice spleens using anti-CD43 MicroBeads (Miltenyi
431 Biotec) and expanded in complete RPMI containing 5 µg/ml LPS (Sigma-Aldrich) and 5 ng/ml mouse
432 recombinant IL-4 (Sigma-Aldrich) to allow B cell activation and class switch recombination to IgG1. B
433 cells were harvested 72h after activation and 4 x 10^7 cells were cross-linked by using 2 mM
434 disuccinimidyl glutarate (ThermoFisher 20593) in PBS for 45 min and 1% formaldehyde for 5 min
435 (Thermo Scientific 28908). The reaction was quenched with 0.125 M glycine. Cells were washed
436 thrice with ice-cold PBS and lysed in SDS lysis buffer. Chromatin fragmentation was performed using
437 a Covaris E220 sonicator to obtain fragments between 200 and 600 bp. Chromatin was quantified
438 with a ND-1000 NanoDrop spectrophotometer. Immunoprecipitation was performed with 0.5 µg of
439 anti-HA antibody (Santa Cruz sc-7392) and 50 µl of Dynabeads protein G (Thermofisher 10003D) and
440 25 µg chromatin. ChIP libraries were prepared by NEBNext ultra II DNA library preparation kit (NEB
441 E7645L) and sequenced on one lane of a NovaSeq 6000 (Illumina) machine.

442 RNA-seq from CH12 cells

443 RNA-seq was performed on CH12 and two independent RIF1-deficient (RIF1$^{-/-}$) CH12 clonal
444 derivatives [29] with three replicates. Cells were expanded in complete RPMI and 1 million cells were
445 collected by centrifugation. RNA was extracted with TRIzol (Invitrogen) according to manufacturer's

446 instructions. TruSeq RNA Library Prep Kit v2 (Illumina) was used to prepare a whole-transcriptome
447 sequencing library and sequenced on one lane of a NovaSeq 6000 SP (Illumina) machine.

448 Isolation of short nascent strands (SNSs)

449 SNSs were isolated following an established protocol[39] kindly provided by Dr. Maria Gomez (CBMSO,
450 Madrid) and also described in our previous study (Wiedemann et al, 2016). In brief, 200 million cells
451 per replicate from asynchronous cell cultures were harvested and genomic DNA was extracted. DNA
452 was denatured at 95°C for 10 minutes and then subjected to size fractionation via 5-20% neutral
453 sucrose gradient centrifugation (24,000g for 20h). Fractions were analyzed by alkaline agarose gel
454 electrophoresis and those in the 500-2000 nt range were pooled. Prior to all following enzymatic
455 treatments, ssDNA was heat denatured for 5 min at 95°C. DNA was phosphorylated for 1h at 37°C
456 with T4 Polynucleotide kinase (NEB or made in-house by the Molecular Biology Service, IMP). To
457 enrich for nascent DNA strands, the phosphorylated ssDNA was digested overnight at 37°C with
458 Lambda exonuclease (NEB or made in-house by the Molecular Biology Service, IMP). Both T4 PNK
459 and Lambda exonuclease steps were repeated twice for a total of three rounds of phosphorylation
460 and digestion. After the final round of digestion, the DNA was treated with RNaseA/T1 mix (Thermo
461 Scientific) to remove 5' RNA primers and genomic RNA contamination. DNA was purified via phenol-
462 chloroform extraction and ethanol precipitation This material was either used directly for qPCR or
463 further processed for library preparation.

464 SNS library preparation for SNS-seq

465 SNSs prepared as described above were converted to double stranded DNA (dsDNA) via random
466 priming with random hexamer primer phosphate (Roche) and ligation with Taq DNA ligase (NEB)[39].
467 DNA was checked on a fragment analyzer and 50ng was used for library preparation with the
468 NEBNext® Ultra™ II DNA Library Prep Kit (NEB) following the manufacturer's protocol. Libraries were
469 barcoded using the NEBNext® Singleplex Oligos (NEB) as per the NEB protocol which allowed 4-8
470 libraries to be pooled per run. Sequencing was performed on an Illumina HiSeq 2500 machine (50 bp,
471 single-end).

472

473

474 In situ Hi-C

475 Hi-C was performed as previously described[34] with minor modifications. In brief, 5 million cells were
476 crosslinked with 1% formaldehyde (Sigma) for 10 min and quenched with 0.6M glycine for 5 min.
477 Cells were lysed with Hi-C lysis buffer for 1h on ice and nuclei were collected by centrifugation. Nuclei
478 were digested with 375U of MboI (NEB) overnight at 37°C with rotation. Biotin-14-dATP (Life
479 Technologies) was incorporated for 1h at 37°C with rotation. Ligation of overhangs was performed
480 with 20,000U of T4 DNA ligase (NEB) for 4h at room temperature with rotation. Nuclei were pelleted
481 and reverse-crosslinked overnight. Purified DNA was sonicated for 14 min in a Diagenode Bioruptor
482 to obtain a size range of 200-700bp. This material was purified using Agencourt AMPure XP beads
483 (Beckman Coulter). Between 8-15 μg of DNA was incubated with 100μl (~10 mg) of Dynabeads
484 MyOne Streptavidin C1 (Invitrogen) for 15 min with rotation. End repair and adapter ligation using the
485 NEBNext® Ultra™ II DNA Library Prep Kit was performed on-beads following the kit manual. The
486 adapter-ligated DNA was washed, eluted and PCR-amplified with KAPA 2X HiFi HotStart ReadyMix
487 (Kapa Biosystems) and the NEBNext Multiplex Oligos for Illumina® (Dual Index Primers Set 1). Four
488 pooled, barcoded samples were sequenced on an Illumina NovaSeq 6000 machine (50 bp, paired-
489 end).

490 PRO-seq

491 PRO-seq was performed as described previously[61] with minor modifications. To isolate nuclei, CH12
492 cells and *Drosphila* S2 cells were resuspended in cold Buffer IA (160mM sucrose, 10mM Tris-Cl pH 8,
493 3mM CaCl2, 2mM MgAc2, 0.5% NP-40, 1mM DTT added fresh), incubated on ice for 3 min and
494 centrifuged at 700g for 5 min. The pellet was resuspended in nuclei resuspension buffer NRB (50mM
495 Tris-Cl pH 8, 40% glycerol, 5mM MgCl2, 0.1mM EDTA). For each run-on, 10 million CH12 nuclei were
496 spiked with 10% *Drosophila* S2 nuclei in a total of 100μL NRB and incubated at 30°C for 3 min with
497 100 μL 2x NRO buffer including 5μl of each 1mM NTP (biotinylated ATP and GTP, and unlabelled
498 UTP and CTP). The following steps were performed as described[61] with the following changes: (1) we
499 used different adapters, namely, 3` RNA adapter
500 5Phos/NNNNNNNGAUCGUCGGACUGUAGAACUCUGAAC/3InvdT-3' and 5'RNA adapter: 5'-
501 CCUUGGCACCCGAGAAUUCCANNNN-3'. (2) 3' and 5' ligations which were done at 16°C overnight,
502 and (3) CapClip pyrophosphatase (Cellscript) used for 5' decapping. RNA was reverse transcribed by

503 SuperScript III RT (Invitrogen) with RP1 Illumina primer to generate cDNA libraries. Libraries were
504 amplified using barcoding Illumina RPI-x primers and the universal RP1 and KAPA HiFi Real-Time
505 PCR Library Amplification Kit. Amplified libraries were subjected to electrophoresis on 2.5% low
506 melting agarose gel and amplicons from 150-350 bp were extracted from the gel, multiplexed and
507 sequenced on Illumina platform NextSeq 550 SR75 high. Bioinformatics analyses were performed as
508 described[61] but additionally the random 8-mer was used to exclude PCR duplicates and only
509 deduplicated reads were aligned.

510 Biological and technical replicates

511 For biological replicates, different frozen vials of CH12 cells were thawed and kept separate
512 throughout the course of the experiment. Replicates for all next-generation sequencing experiments
513 were derived in this manner. Technical replicates, where used (such as in RT-qPCR or ChIP-qPCR),
514 were subsets of the biological replicate.

515 Statistical analyses

516 For correlation scatter plots, the Pearson correlation coefficient was calculated to determine the
517 degree of correlation. In all other cases, the two-tailed Student's t test was used for statistical
518 significance.

519

520 **Bioinformatics**

521 Main next generation sequencing data analysis workflows for SNS-seq, Repli-seq and Hi-C have
522 been wrapped with *Nextflow* workflow language [62] and are centrally available at
523 https://github.com/pavrilab [12]. The usage of Docker and Singlularity containers ensures portability,
524 reproducibility and reliability for all workflows. Along with integration of GitHub repositories for self-
525 contained pipelines all workflows are easily ported to all major HPC computing platforms such as
526 SGE, SLURM, AWS. Furthermore, continuous checkpoints for pipeline execution allow for resuming
527 and automatic retrial of failed steps. All workflows produce elaborate QC-reports and out-of-the box,
528 resources consumption reports to allow tailoring resources requirements to your datasets which
529 especially for Hi-C datasets can vary by several orders of magnitudes.

530 Processing of raw Repli-seq, Hi-C, SNS-seq and RNA-seq data was done as described
531 previously.[12]

532 Generating heatmaps for IS peak summit neighborhoods

533 Sequencing coverage normalized bigWig tracks were generated from raw aligned reads of NGS data
534 sets with deeptools bamCoverage v3.3.0 [63] using command-line parameters --normalizeUsing CPM --
535 exactScaling and --ignoreDuplicates. Next, we computed the position of the replication start site
536 (RSS) as the genomic coordinate of the maximum pile-up of the merged tracks of compared
537 conditions. The generated summit positions were then grouped by their associated peak class and
538 sorted in decending order on the log2-ratio of CPM values of treatment versus control. The sorted
539 summits were then used as reference points for deeptools computeMatrix v3.3.0 [63] to compute signal
540 distributions for previously generated CPM-normalized bigWig tracks within a 4kb region centered on
541 the summit using a binsize of 50 bp and setting missing values to zero. The results were then plotted
542 using the associated plotHeatmap command.

543 PRO-seq analysis

544 3' adaptors were removed using cutadapt v1.4.2 [64] with -O 5 -a NNNNTGGAATTCTCGGTGCC) and
545 9 nucleotides from their 5′ were removed. The trimmed reads were reverse complemented using
546 fastx_reverse_complement (http://hannonlab.cshl.edu/fastx_toolkit; version 0.0.13). Trimmed reads
547 longer than 18bp were aligned to a hybrid mouse drosophila genome (mouse genome build NCBI
548 m37, GCA_000001635.18, drosophila FlyBase release 5) using bowtie v1.0.0 [65] with -v 2 --best --
549 strata --tryhard -m 1 --chunkmbs 256. Unique mappers from the resulting BAM file were used to
550 create bigwigs with deeptools bamCoverage 3.3.0.

551 RT domain analysis

552 20kb binned RT tracks for all conditions were segmented into three states using hmm_bigwigs
553 (https://github.com/gspracklin/hmm_bigwigs). Fitted states were then mapped to either early (E), mid
554 (M) or late (L) replication timing based on the RT value distribution of each state (E for RT > 0, L for
555 RT < 0 and M for RT ~ 0). M state bins were split at 0 into L-like (values < 0) and E-like (values > 0)
556 state bins.

557

558 Generating saddle plots from Hi-C data

559 Saddle plots were computed using cooltools v0.3.2. In brief, we subdivided all bins of a 20Kb KR-
560 normalized contact matrix into 50 equal-sized groups based on the bins compartment signal as
561 derived from the eigenvector of the WT data, where group 1 has the lowest signal (i.e. most B) and
562 group 50 has the highest signal (i.e. most A). Subsequently,we compute the mean observed/expected
563 value for each pair of groups and plot it as a 50 x 50 matrix. Similarly, replication timing and H3K9me3
564 data can be used for bin group assignment.

565 RIF1 peak and RIF1 associated domains (RADs) analyses

566 Peaks and domains (RADs) were called as described previously [66] additionally calling narrow peaks
567 with MACS2. In brief, raw ChIP-seq and input reads in MEFs and primary B cells were trimmed with
568 trim_galore v0.6.4 (https://github.com/FelixKrueger/TrimGalore) setting minimum length to 18 bases
569 and default arguments otherwise and aligned to the mm9 reference genome with bowtie v1.0.0 [65] with
570 -S --trim5 0 --trim3 0 -v 2 --best --strata --tryhard -m 1 --phred33-quals --chunkmbs 256.
571 Subsequently, RIF1 narrow and broad peaks were called for each replicate with MACS v2.2.6 [67], RIF1
572 associated domains (RADs) were called using EDD v1.1.19 [68]. RADs, narrow and broad peaks were
573 merged using bedtools v2.27.1 [69] and only those peaks called in both replicates were retained for
574 downstream analysis. Genome coverage tracks were computed with deeptools bamCoverage v3.3.0
575 [63] using command-line parameters --normalizeUsing CPM --exactScaling –binSize 50 and –
576 ignoreDuplicates. Log2 ratio tracks were computed with deeptools bamCompare v3.3.0 [63] using
577 command-line parameters --scaleFactorsMethod readCount --operation log2 --pseudocount 1 –
578 binSize 50 –ignoreDuplicates.

579 Summary of all NGS data used

580 *Table 1: Summary of used data sets in this study including public Accession numbers*

Sample Name	Data type	Source	Accession
CH12 shLacZ	SNS-seq	GEO	GSE161822
CH12 shMcm6	SNS-seq	GEO	GSE161822
CH12 WT	SNS-seq	This study	-
CH12 Rif1$^{-/-}$	SNS-seq	This study	-

CH12 H3K4me1	ChIP-seq	GEO/UCSC	GSM946546/wgEncodeEM002370
CH12 H3K4me3	ChIP-seq	GEO/UCSC	GSM946528/wgEncodeEM002366
CH12 H3K9me3 shLacZ	ChIP-seq	GEO	GSE161822
CH12 H3K9me3 shMcm6	ChIP-seq	GEO	GSE161822
CH12 H3K27ac	ChIP-seq	GEO/UCSC	GSM1000117/wgEncodeEM003167
CH12 H3K36me3	ChIP-seq	GEO/UCSC	GSM946530/wgEncodeEM002364
CH12 DHS-seq	DHS-seq	GEO/UCSC	GSM1014153/wgEncodeEM003416
CH12 shLacZ	Hi-C	GEO	GSE161822
CH12 shMcm6	Hi-C	GEO	GSE161822
CH12 Rif1$^{-/-}$ shLacZ	Hi-C	This study	-
CH12 Rif1$^{-/-}$ shMcm6	Hi-C	This study	-
CH12 shLacZ	Repli-seq	GEO	GSE161822
CH12 shMcm6	Repli-seq	GEO	GSE161822
CH12 Rif1$^{-/-}$ shLacZ	Repli-seq	This study	-
CH12 Rif1$^{-/-}$ shMcm6	Repli-seq	This study	-
CH12 shLacZ	RNA-seq	GEO	GSE161822
CH12 shMcm6	RNA-seq	GEO	GSE161822
CH12 Rif1$^{-/-}$ shLacZ	RNA-seq	This study	-
CH12 Rif1$^{-/-}$ shMcm6	RNA-seq	This study	-
CH12 shLacZ	PRO-seq	This study	-
primary B/CH12 input track	ChIP-seq	GEO	GSM2058441
primary B Rif1$^{FH/FH}$	ChIP-seq	This study	-
primary B Rif1$^{WT/WT}$	ChIP-seq	This study	-
primary B WT Repli-seq	Repli-seq	This study	-
primary B Rif1$^{-/+}$ Repli-seq	Repli-seq	This study	-
primary B Rif1$^{-/-}$ Repli-seq	Repli-seq	This study	-
primary B WT ATAC-seq	ATAC-seq	GEO	GSE132029
primary B WT PRO-seq	PRO-seq	This study	-
primary B WT H3K9me3	ChIP-seq	This study	-

581 Summary of software used in the study

582 All custom Python scripts and tools, except for scripts using ClusterScan, were developed, tested and
583 executed using Python v3.7.3. ClusterScan and related scripts were executed using Python v2.7.13.
584 The used software and packages are summarized in Table , Table 3 and Table 4.

585

586 *Table 2: Summary of standalone software*

Software	**Version**	**Usage**
bedtools	2.27.1	general operations on sets of genomic regions
bowtie	1.0.0/1.2.3	Alignment of reads
bowtie2	2.2.9	Alignment of reads
bwa	0.7.17	Alignment of Repli-seq data
ClusterScan	0.2.1	Clustering of initiation sites
cutadapt	2.6	Adapter trimming
deeptools	3.3.0	Read counting in regions and heatmap generation
DESeq2	1.22.2	Differential expression analysis
EDD	1.1.19	Calling of Rif1 associated domains
epic2	0.0.41	Peak calling for broad peak type data
FastQC	0.11.5	Quality control of sequencing reads
featureCounts	2.0.0	Computing RNA-seq data counts
HICUP	0.7.3	Processing of Hi-C data
hmm_bigwig	1.3	Segmentation of genome into RT states
HOMER	4.10	Hi-C Eigenvector and TAD annotation
UCSCkentUtils	356	Conversion between file formats
MACS2	2.2.6	Peak calling for narrow peak type data
Python2	2.7.13	Execution of ClusterScan and associated scripts
Python3	3.7.3	General script execution

R	3.5.1	Execution of R code
samtools	1.9	General operations on aligned reads
STAR	2.4.2a	Alignment of RNA-seq data
Trim_galore	0.6.4	Quality control of sequencing data

587

588 *Table 3: Summary of R packages*

Package	***Version***	***Usage***
ashr	2.2-47	Differential expression analysis

589

590 *Table 4: Summary of Python packages*

Package	**Version**	**Usage**
argparse	1.1	Passing command-line arguments to scripts
BioPython	1.77	Mapping of EntrezID to gene symbol
cooler	0.8.3	Hi-C matrix generation and storage
cooltools	0.3.2	Generic Hi-C downstream analysis
genomepy	0.8.3	Genome masking in Python
krbalancing	0.0.5	HiCExplorer implementation of KR for matrix correction
logging	0.5.1.2	Printing information to stdout
matplotlib	3.1.0	Generation of general plots
matplotlib_venn	0.11.5	Generation of Venn diagrams
numpy	1.16.4	General computations on data in matrices
pandas	1.1.0	General computations on tabulated data
pybedtools	0.8.0	Using bedtools within Python
pyBigWig	0.3.16	Generation of bigWig files
pysam	0.15.2	Using samtools in Python
re	2.2.1	Regular expressions for efficient text parsing
rpy2	3.0.5	Using R packages in Python

scipy	1.3.0	General statistical computation in Python
seaborn	0.11.1	Kernel density estimation of scatter plots
tables	3.5.2	Reading of HDF5 files

591

592 **ACKNOWLEDGEMENTS**

593 We are grateful to the Vienna Biocenter Core Facilities (VBCF) for next-generation sequencing and
594 the IMP/IMBA core facilities especially the animal house, bio-optics, molecular biology service and
595 proteomics service. We are grateful to Roman Stocsits and Dr. Anton Goloborodko for advice on Hi-C
596 data analysis. This work was funded by Boehringer Ingelheim, The Austrian Industrial Research
597 Promotion Agency (Headquarter Grant FFG-834223), grants from the Austrian Science Fund to RP
598 (FWF P 29163-B26) and UES (FWF T 795-B30), and the Helmholtz-Gemeinschaft Zukunftsthema
599 "Immunology and Inflammation" ZT810 0027 to MDV and AR. Please add grants from your side.

600

601 **AUTHOR CONTRIBUTIONS**

602 DM performed most of the bioinformatics analyses, established pipelines and analyzed data. MP
603 performed Repli-seq, Hi-C, RNA-seq and SNS-seq, and analyzed data. AR performed RIF1 ChIP-seq
604 from primary B cells and RNA-Seq in CH12 cells. SG and KNK performed Repli-seq from primary B
605 cells. US performed PRO-seq. MN collaborated on Repli-seq from CH12 cells. DMG, SCBB and MDV
606 contributed personnel, resources, extensive discussions and critical reading of the manuscript. TN
607 and RP co-supervised the project and analyzed data. RP conceived the project and wrote the
608 manuscript with inputs from all authors.

609

610

611

612

613

614

615 **FIGURE LEGENDS**

616 **Figure 1: RIF1 regulates early RT in activated B cells**

617 A. RT histograms in 5 kb genomic bins from shLacZ, shMcm6 and *Rif1*$^{-/-}$ shLacZ CH12 cells. RT is
618 calculated as the ratio of the read densities of early (E) and late (L) Repli-seq fractions (log2 E/L).
619 B. UCSC genome browser view of RT for the indicated conditions in CH12 cells (top panel) and
620 primary, activated splenic B cells (bottom panel). Positive and negative values correspond to early
621 and late-replicating regions.
622 C. RT histograms as in (A) from WT, *Rif1*$^{-/+}$ and *Rif1*$^{-/-}$ primary, activated splenic B cells.
623 D. Comparison of RT values in 20 kb genomic bins between shLacZ and shMcm6 CH12 cells (left)
624 and shLacZ and *Rif1*$^{-/-}$ shLacZ cells (right).
625 E. Comparison of RT values in 20 kb genomic bins between WT, *Rif1*$^{-/+}$ and *Rif1*$^{-/-}$ in primary B cells.
626 F. Classification of shLacZ RT bins into four states using a Hidden Markov Model (HMM). Three
627 states E, M (mid) and L were called, and the middle state was split into into E-like and L-like at the
628 log2 (E/L) = 0 boundary.
629 G. RT histograms showing how the bins in each shLacZ RT state (blue) called in (D) change in
630 shMcm6 (red) and *Rif1*$^{-/-}$ shLacZ (green) CH12 cells.
631 H. HMM analysis as in (D) but using WT RT values from primary B cells to call RT states.
632 I. RT histograms showing how bins in each WT RT state (blue) change in *Rif1*$^{-/+}$ (ochre) and
633 *Rif1*$^{-/-}$ (green) primary B cells.

634 **Figure 2: RIF1 has a minor role in B cell genome architecture**

635 A. Left: Hi-C contact matrix of chromosome 11 showing shLacZ contacts (above the diagonal) and
636 shMcm6 contacts (below the diagonal). The PC1 compartment signals (Hi-C eigenvector
637 eigenvalues) and RT (log2 E/L) tracks are shown above (for shLacZ) and on the left (for
638 shMcm6) for each matrix. Right: Same as before but comparing shLacZ and *Rif1*$^{-/-}$ shLacZ cells.
639 B. Density-contour plot of PC1 compartment signals per 20 kb genomic bin. The left plot compares
640 the PC1 signals in shLacZ and shMcm6 cells, and the right plot compares PC1 values between
641 shLacZ and *Rif1*$^{-/-}$ shLacZ cells. the numbers within the plots are the percentage of bins in that
642 quadrant.

643 C. A representative UCSC genome browser view comparing Hi-C PC1 profiles with RT (log2 E/L)
644 profiles in shLacZ, shMcm6 and *Rif1*$^{-/-}$ shLacZ cells.
645 D. PC1 versus RT density-contour plots in 20 kb genomic bins to compare changes in
646 compartmental identities (PC1) with RT in shLacZ, shMcm6 and *Rif1*$^{-/-}$ shLacZ CH12 cells.
647 E. Saddle plot from shLacZ, shMcm6 and *Rif1*$^{-/-}$ shLacZ CH12 cells showing long-range (> 2Mb)
648 intra-chromosomal contact enrichments between bins of varying compartment signal strength
649 (PC1). The values were computed from 20 kb KR-normalized contact matrices.
650 F-G Fold-change saddle plots highlighting the changes in compartmentalization in shMcm6 and *Rif1*$^{-/-}$
651 shLacZ cells relative to shLacZ cells.
652

653 **Figure 3: RIF1 localizes predominantly to active chromatin in activated B cells**

654 A. A representative UCSC genome browser view of RIF1 occupancy in *Rif1*$^{FH/FH}$ primary, activated
655 splenic B cells. ChIP-seq was performed with an anti-HA antibody and *Rif1*$^{WT/WT}$ cells were used
656 as a negative control. The RT track from WT, primary B cells provides a reference for early and
657 late-replicating domains.
658 B. UCSC browser snapshot comparing ChIP enrichments of RIF1 in primary B and MEFs from
659 *Rif1*$^{FH/FH}$ mice. To allow direct comparison, ChIP signal was normalized to input signal and the
660 ratio (ChIP/Input) tracks are shown. RT tracks are from WT cells.
661 C. Box plots comparing RT values of RIF1 peaks and RADs in MEFs and primary B cells. 16,043
662 peaks and 289 RADs were identified in primary B cells and 862 peaks and 332 RADs were called
663 in MEFs.
664 D. Peaks and RADs were called as in C above and classified into the four RT states described in Fig.
665 1D.
666 E. Heatmap analysis of the chromatin locale in a 4 kb window surrounding RIF1 peaks called in
667 *Rif1*$^{FH/FH}$ cells. The heatmap is centered on the RIF1 peak summit and ordered by decreasing RIF1
668 read density (reads per million, RPM). All other genomic tracks shown (PRO-seq, ATAC-seq,
669 H3K27Ac and Repli-seq) are from WT primary, activated splenic B cells.
670 F. Distribution of RIF1 peaks and RADs at TSSs, gene bodies and intergenic regions.

671 G. Metagene analysis showing average RIF1 occupancy patterns at genes in *Rif1*$^{FH/FH}$ and *Rif1*$^{WT/WT}$
672 cells from two replicates each (rep1 and rep2). The TSS and TTS are the transcription start site
673 and transcription termination site.

674 **Figure 4: RIF1 regulates the activity of early origins of replication in B cells**

675 A. Scatter plot of IS density in log2 RPM (reads per million) in WT and *Rif1*$^{-/-}$ cells identified from
676 SNS-seq.
677 B. Heatmaps spanning a 4 kb region centered on the IS peak summit were generated to visualize the
678 profiles of SNS-seq, DNase hypersensitivity (DHS)-seq and histone modifications (H3K36me3
679 and H3K9me3) around ISs. All heatmaps display read densities as RPM. Within each heatmap,
680 the ordering is based on decreasing SNS-seq fold-change (*Rif1*$^{-/-}$/WT).
681 C. ISs were split into five equal classes (quintiles) based on their fold-change (log2 *Rif1*$^{-/-}$/WT) such
682 that class 1 contained the most downregulated ISs and class 5 contained the most upregulated
683 ISs, respectively, in *Rif1*$^{-/-}$ cells. The violin plots show the IS density within each class in WT and
684 *Rif1*$^{-/-}$ cells.

685 **Figure 5: RIF1 and MCM complexes act in an additive manner to regulate early RT in B cells**

686 A. RT histogram comparing *Rif1*$^{-/-}$ shLacZ and *Rif1*$^{-/-}$ shMcm6 CH12 cells.
687 B. Identical to Fig. 1E showing how HMM-based RT states called in shLacZ cells change in the other
688 three conditions.
689 C. A representative UCSC genome browser view comparing Hi-C PC1 profiles with RT (log2 E/L)
690 profiles in shLacZ, shMcm6, *Rif1*$^{-/-}$ shLacZ and *Rif1*$^{-/-}$ shMcm6 CH12 cells.
691 D. A magnified view of the boxed region in C. The PRO-seq track shows nascent transcription profiles
692 in shLacZ cells.

693 **Figure 6: MCM depletion in *Rif1*$^{-/-}$ cells leads to a further deregulation of early origin activity**

694 A. Scatter plot of IS density in log2 RPM (reads per million) in *Rif1*$^{-/-}$ shLacZ and *Rif1*$^{-/-}$ shMcm6 cells
695 identified from SNS-seq.
696 B. ISs were split into five equal classes (quintiles) based on their fold-change (log2 *Rif1*$^{-/-}$ shMcm6/
697 *Rif1*$^{-/-}$ shLacZ) such that class 1 contained the most downregulated ISs and class 5 contained the

698 most upregulated ISs, respectively, in *Rif1*$^{-/-}$ shMcm6 cells. The violin plots show the IS density
699 within each class in *Rif1*$^{-/-}$ shLacZ and *Rif1*$^{-/-}$ shMcm6 cells.
700 C. ISs were classified into five classes as in C above. The violin plots show the RT values within each
701 IS class.
702 D. Representative genomic snapshot of SNS-seq and HiC PC1 profiles. Data from shLacZ (blue) and
703 shMcm6 (red) are overlaid with black being the overlap between them. *Rif1*$^{-/-}$ shLacZ and *Rif1*$^{-/-}$
704 shMcm6 tracks are in green and purple, respectively, with black indicating the overlap between
705 them.

706 **Figure 7: Within early-replicating domains, the RT of highly transcribed regions is more**
707 **sensitive to reduction in MCM proteins and loss of RIF1 than poorly transcribed regions**

708 A. Heatmap of RT values in 20 kb E bins from shLacZ cells that were split into four groups based on
709 PRO-seq RPM. The range of PRO-seq RPMs in each group is indicated on the left. All heatmaps
710 are ranked from highest to lowest PRO-seq density such that the entire set of heatmaps is
711 effectively ranked from highest to lowest PRO-seq density for all shLacZ E bins.
712 B. Density distribution of the RT (top) values in the four transcription-based groups from A above. The
713 dashed black lines indicate the modes of the distributions in the High and Untranscribed classes.

714

715 **Supplemental figure legends**

716 **Figure S1: Analysis of Repli-seq replicates**

717 A. Pearson correlation matrix of Repli-seq data from two replicates (rep1 and rep2) of shLacZ,
718 shMcm6, *Rif1*$^{-/-}$ shLacZ and *Rif1*$^{-/-}$ shMcm6 CH12 cells. E and L indicate the early and late
719 fractions, respectively. The color gradient (bottom bar) corresponds to the Pearson correlation
720 coefficient.
721 B. Western blot for Mcm6 protein levels from nuclear extracts obtained from shLacZ and shMcm6
722 cells. Data from two replicates are shown with γ-Tubulin serves as the loading control. In all cases,
723 20 μg of extract was loaded.
724 C. Repli-seq correlation analysis as in A from WT (two replicates), *Rif1*$^{-/+}$ (four replicates) and *Rif1*1$^{-/-}$
725 (four replicates) primary, activated splenic B cells.

726 **Figure S2: RNA-seq analysis from *Rif1*$^{-/-}$ CH12 cells**

727 A. Scatter plot of RNA-seq read densities in transcripts per million (TPM) versus fold-change (FC;
728 *Rif1*$^{-/-}$ shLacZ/shLacZ) at 8,606 expressed genes (defined as TPM >5 in shLacZ cells). The dotted
729 lines mark log2 FC 1 or -1 corresponding to 2-fold upregulated and 2-fold downregulated genes,
730 respectively.
731 B. GO term analysis of the upregulated and downregulated genes from above.

732 **Figure S3: Analysis of ChIP-seq and SNS-seq replicates**

733 A. Pearson correlation analysis of RIF1 ChIP-seq datasets. Two replicates from *Rif1*$^{FH/FH}$ and
734 *Rif1*$^{WT/WT}$ primary B cells were analyzed.
735 B. Pearson correlation analysis of SNS-seq data. Two replicates from each indicated condition were
736 analyzed.
737 C. Bar plots showing the percent of ISs upregulated (1.5-fold), downregulated (1.5-fold) or unchanged
738 in shMcm6 (relative to shLacZ) or *Rif1*$^{-/-}$ (relative to WT) cells. The shMcm6 data is from our
739 previous study[12].

740 **Figure S4: Gene expression and genome compartmentalization analyses in *Rif1*$^{-/-}$ shMcm6**
741 **cells**

742 A. Western blot for Mcm6 protein levels from nuclear extracts obtained from two replicates of *Rif1*$^{-/-}$
743 shLacZ and *Rif1*$^{-/-}$ shMcm6 cells. In all cases, 20 μg of extract was loaded. γ-Tubulin serves as the
744 loading control.
745 B. Scatter plot of RNA-seq read densities in TPM versus fold-change (*Rif1*$^{-/-}$ shMcm6/*Rif1*$^{-/-}$ shLacZ)
746 at 8,591 expressed genes (defined as TPM >5 in *Rif1*$^{-/-}$ shLacZ cells). The dotted lines mark log2
747 FC 1 and -1 corresponding to 2-fold upregulated and 2-fold downregulated genes, respectively.
748 C. GO term analysis of the 50 downregulated genes from above.
749 D. Density-contour plot comparing PC1 compartment signals in 20 kb genomic bins from *Rif1*$^{-/-}$
750 shLacZ and *Rif1*$^{-/-}$ shMcm6 cells.
751 E. Saddle plots comparing compartmental interactions between *Rif1*$^{-/-}$ shLacZ and *Rif1*$^{-/-}$ shMcm6
752 cells.
753 F. Fold-change saddle plots based on data from D.

754 G. PC1 versus RT density-contour plots in 20 kb genomic bins comparing the changes in either
755 feature in *Rif1*$^{-/-}$ shLacZ and *Rif1*$^{-/-}$ shMcm6 cells.
756 H. Histogram showing the distribution of PRO-seq densities (RPM) in 20 kb genomic bins from
757 shLacZ cells. High, Medium, Low and Untranscribed groups are color-coded to match the analysis
758 in Fig. 7B. The dotted red line indicates RPM 10 which marks the boundary between transcribed
759 and untranscribed regions.

760

761

762

763

764

765

766

767

768

769

770

771

772

773

774

775

776

777

778 **REFERENCES**

779 1. Marchal, C., Sima, J., and Gilbert, D.M. (2019). Control of DNA replication timing in the 3D
780 genome. Nat Rev Mol Cell Biol *20*, 721-737. 10.1038/s41580-019-0162-y.
781 2. Pope, B.D., Ryba, T., Dileep, V., Yue, F., Wu, W., Denas, O., Vera, D.L., Wang, Y., Hansen, R.S.,
782 Canfield, T.K., et al. (2014). Topologically associating domains are stable units of replication-
783 timing regulation. Nature *515*, 402-405. 10.1038/nature13986.
784 3. Sima, J., Chakraborty, A., Dileep, V., Michalski, M., Klein, K.N., Holcomb, N.P., Turner, J.L.,
785 Paulsen, M.T., Rivera-Mulia, J.C., Trevilla-Garcia, C., et al. (2019). Identifying cis Elements for
786 Spatiotemporal Control of Mammalian DNA Replication. Cell *176*, 816-830 e818.
787 10.1016/j.cell.2018.11.036.
788 4. Rivera-Mulia, J.C., Desprat, R., Trevilla-Garcia, C., Cornacchia, D., Schwerer, H., Sasaki, T.,
789 Sima, J., Fells, T., Studer, L., Lemaitre, J.M., and Gilbert, D.M. (2017). DNA replication timing
790 alterations identify common markers between distinct progeroid diseases. Proc Natl Acad Sci
791 U S A *114*, E10972-e10980. 10.1073/pnas.1711613114.
792 5. Platt, E.J., Smith, L., and Thayer, M.J. (2018). L1 retrotransposon antisense RNA within ASAR
793 lncRNAs controls chromosome-wide replication timing. J Cell Biol *217*, 541-553.
794 10.1083/jcb.201707082.
795 6. Sasaki, T., Rivera-Mulia, J.C., Vera, D., Zimmerman, J., Das, S., Padget, M., Nakamichi, N.,
796 Chang, B.H., Tyner, J., Druker, B.J., et al. (2017). Stability of patient-specific features of
797 altered DNA replication timing in xenografts of primary human acute lymphoblastic
798 leukemia. Exp Hematol *51*, 71-82.e73. 10.1016/j.exphem.2017.04.004.
799 7. Klein, K.N., Zhao, P.A., Lyu, X., Sasaki, T., Bartlett, D.A., Singh, A.M., Tasan, I., Zhang, M.,
800 Watts, L.P., Hiraga, S.I., et al. (2021). Replication timing maintains the global epigenetic state
801 in human cells. Science *372*, 371-378. 10.1126/science.aba5545.
802 8. Du, Q., Bert, S.A., Armstrong, N.J., Caldon, C.E., Song, J.Z., Nair, S.S., Gould, C.M., Luu, P.L.,
803 Peters, T., Khoury, A., et al. (2019). Replication timing and epigenome remodelling are
804 associated with the nature of chromosomal rearrangements in cancer. Nat Commun *10*, 416.
805 10.1038/s41467-019-08302-1.
806 9. Janoueix-Lerosey, I., Hupe, P., Maciorowski, Z., La Rosa, P., Schleiermacher, G., Pierron, G.,
807 Liva, S., Barillot, E., and Delattre, O. (2005). Preferential occurrence of chromosome
808 breakpoints within early replicating regions in neuroblastoma. Cell Cycle *4*, 1842-1846.
809 10.4161/cc.4.12.2257.
810 10. Shugay, M., Ortiz de Mendibil, I., Vizmanos, J.L., and Novo, F.J. (2012). Genomic hallmarks of
811 genes involved in chromosomal translocations in hematological cancer. PLoS Comput Biol *8*,
812 e1002797. 10.1371/journal.pcbi.1002797.
813 11. Sima, J., and Gilbert, D.M. (2014). Complex correlations: replication timing and mutational
814 landscapes during cancer and genome evolution. Curr Opin Genet Dev *25*, 93-100.
815 10.1016/j.gde.2013.11.022.
816 12. Peycheva, M., Neumann, T., Malzl, D., Nazarova, M., Schoeberl, U.E., and Pavri, R. (2022).
817 DNA replication timing directly regulates the frequency of oncogenic chromosomal
818 translocations. Science *377*, eabj5502. 10.1126/science.abj5502.
819 13. Bleichert, F. (2019). Mechanisms of replication origin licensing: a structural perspective. Curr
820 Opin Struct Biol *59*, 195-204. 10.1016/j.sbi.2019.08.007.
821 14. Bochman, M.L., and Schwacha, A. (2009). The Mcm complex: unwinding the mechanism of a
822 replicative helicase. Microbiol Mol Biol Rev *73*, 652-683. 10.1128/MMBR.00019-09.
823 15. Burgers, P.M.J., and Kunkel, T.A. (2017). Eukaryotic DNA Replication Fork. Annu Rev Biochem
824 *86*, 417-438. 10.1146/annurev-biochem-061516-044709.
825 16. Fragkos, M., Ganier, O., Coulombe, P., and Mechali, M. (2015). DNA replication origin
826 activation in space and time. Nat Rev Mol Cell Biol *16*, 360-374. 10.1038/nrm4002.

827 17. Hu, Y., and Stillman, B. (2023). Origins of DNA replication in eukaryotes. Molecular Cell *83*,
828 352-372. https://doi.org/10.1016/j.molcel.2022.12.024.
829 18. Hyrien, O., Marheineke, K., and Goldar, A. (2003). Paradoxes of eukaryotic DNA replication:
830 MCM proteins and the random completion problem. Bioessays *25*, 116-125.
831 10.1002/bies.10208.
832 19. Collart, C., Allen, G.E., Bradshaw, C.R., Smith, J.C., and Zegerman, P. (2013). Titration of four
833 replication factors is essential for the Xenopus laevis midblastula transition. Science *341*,
834 893-896. 10.1126/science.1241530.
835 20. Tanaka, S., Nakato, R., Katou, Y., Shirahige, K., and Araki, H. (2011). Origin association of
836 Sld3, Sld7, and Cdc45 proteins is a key step for determination of origin-firing timing. Curr Biol
837 *21*, 2055-2063. 10.1016/j.cub.2011.11.038.
838 21. Mantiero, D., Mackenzie, A., Donaldson, A., and Zegerman, P. (2011). Limiting replication
839 initiation factors execute the temporal programme of origin firing in budding yeast. Embo j
840 *30*, 4805-4814. 10.1038/emboj.2011.404.
841 22. Douglas, Max E., and Diffley, John F.X. (2012). Replication Timing: The Early Bird Catches the
842 Worm. Current Biology *22*, R81-R82. https://doi.org/10.1016/j.cub.2011.12.028.
843 23. Cornacchia, D., Dileep, V., Quivy, J.P., Foti, R., Tili, F., Santarella-Mellwig, R., Antony, C.,
844 Almouzni, G., Gilbert, D.M., and Buonomo, S.B. (2012). Mouse Rif1 is a key regulator of the
845 replication-timing programme in mammalian cells. EMBO J *31*, 3678-3690.
846 10.1038/emboj.2012.214.
847 24. Yamazaki, S., Ishii, A., Kanoh, Y., Oda, M., Nishito, Y., and Masai, H. (2012). Rif1 regulates the
848 replication timing domains on the human genome. EMBO J *31*, 3667-3677.
849 10.1038/emboj.2012.180.
850 25. Hayano, M., Kanoh, Y., Matsumoto, S., Renard-Guillet, C., Shirahige, K., and Masai, H. (2012).
851 Rif1 is a global regulator of timing of replication origin firing in fission yeast. Genes Dev *26*,
852 137-150. 10.1101/gad.178491.111.
853 26. Hiraga, S., Alvino, G.M., Chang, F., Lian, H.Y., Sridhar, A., Kubota, T., Brewer, B.J., Weinreich,
854 M., Raghuraman, M.K., and Donaldson, A.D. (2014). Rif1 controls DNA replication by
855 directing Protein Phosphatase 1 to reverse Cdc7-mediated phosphorylation of the MCM
856 complex. Genes Dev *28*, 372-383. 10.1101/gad.231258.113.
857 27. Gnan, S., Flyamer, I.M., Klein, K.N., Castelli, E., Rapp, A., Maiser, A., Chen, N., Weber, P.,
858 Enervald, E., Cardoso, M.C., et al. (2021). Nuclear organisation and replication timing are
859 coupled through RIF1-PP1 interaction. Nat Commun *12*, 2910. 10.1038/s41467-021-22899-2.
860 28. Foti, R., Gnan, S., Cornacchia, D., Dileep, V., Bulut-Karslioglu, A., Diehl, S., Buness, A., Klein,
861 F.A., Huber, W., Johnstone, E., et al. (2016). Nuclear Architecture Organized by Rif1
862 Underpins the Replication-Timing Program. Mol Cell *61*, 260-273.
863 10.1016/j.molcel.2015.12.001.
864 29. Delgado-Benito, V., Rosen, D.B., Wang, Q., Gazumyan, A., Pai, J.A., Oliveira, T.Y.,
865 Sundaravinayagam, D., Zhang, W., Andreani, M., Keller, L., et al. (2018). The Chromatin
866 Reader ZMYND8 Regulates Igh Enhancers to Promote Immunoglobulin Class Switch
867 Recombination. Mol Cell *72*, 636-649.e638. 10.1016/j.molcel.2018.08.042.
868 30. Balasubramanian, S., Andreani, M., Andrade, J.G., Saha, T., Sundaravinayagam, D., Garzon, J.,
869 Zhang, W., Popp, O., Hiraga, S.I., Rahjouei, A., et al. (2022). Protection of nascent DNA at
870 stalled replication forks is mediated by phosphorylation of RIF1 intrinsically disordered
871 region. Elife *11*. 10.7554/eLife.75047.
872 31. Wiedemann, E.M., Peycheva, M., and Pavri, R. (2016). DNA Replication Origins in
873 Immunoglobulin Switch Regions Regulate Class Switch Recombination in an R-Loop-
874 Dependent Manner. Cell Rep *17*, 2927-2942. 10.1016/j.celrep.2016.11.041.
875 32. Marchal, C., Sasaki, T., Vera, D., Wilson, K., Sima, J., Rivera-Mulia, J.C., Trevilla-Garcia, C.,
876 Nogues, C., Nafie, E., and Gilbert, D.M. (2018). Genome-wide analysis of replication timing

877 by next-generation sequencing with E/L Repli-seq. Nat Protoc *13*, 819-839.
878 10.1038/nprot.2017.148.
879 33. Enervald, E., Powell, L.M., Boteva, L., Foti, R., Blanes Ruiz, N., Kibar, G., Piszczek, A., Cavaleri,
880 F., Vingron, M., Cerase, A., and Buonomo, S.B.C. (2021). RIF1 and KAP1 differentially regulate
881 the choice of inactive versus active X chromosomes. The EMBO Journal *40*, e105862.
882 https://doi.org/10.15252/embj.2020105862.
883 34. Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T.,
884 Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., and Aiden, E.L. (2014). A 3D map of the
885 human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*,
886 1665-1680. 10.1016/j.cell.2014.11.021.
887 35. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C.,
888 Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq
889 (MACS). Genome Biol *9*, R137. 10.1186/gb-2008-9-9-r137.
890 36. Lund, E., Oldenburg, A.R., and Collas, P. (2014). Enriched domain detector: a program for
891 detection of wide genomic enrichment domains robust against local variations. Nucleic Acids
892 Res *42*, e92. 10.1093/nar/gku324.
893 37. Kwak, H., Fuda, N.J., Core, L.J., and Lis, J.T. (2013). Precise maps of RNA polymerase reveal
894 how promoters direct initiation and pausing. Science *339*, 950-953.
895 10.1126/science.1229386.
896 38. Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013).
897 Transposition of native chromatin for fast and sensitive epigenomic profiling of open
898 chromatin, DNA-binding proteins and nucleosome position. Nat Methods *10*, 1213-1218.
899 10.1038/nmeth.2688.
900 39. Lombrana, R., Almeida, R., Revuelta, I., Madeira, S., Herranz, G., Saiz, N., Bastolla, U., and
901 Gomez, M. (2013). High-resolution analysis of DNA synthesis start sites and nucleosome
902 architecture at efficient mammalian replication origins. EMBO J *32*, 2631-2644.
903 10.1038/emboj.2013.195.
904 40. Cayrou, C., Coulombe, P., Vigneron, A., Stanojcic, S., Ganier, O., Peiffer, I., Rivals, E., Puy, A.,
905 Laurent-Chabalier, S., Desprat, R., and Mechali, M. (2011). Genome-scale analysis of
906 metazoan replication origins reveals their organization in specific but flexible sites defined
907 by conserved features. Genome Res *21*, 1438-1449. 10.1101/gr.121830.111.
908 41. Ge, X.Q., Jackson, D.A., and Blow, J.J. (2007). Dormant origins licensed by excess Mcm2-7 are
909 required for human cells to survive replicative stress. Genes Dev *21*, 3331-3341.
910 10.1101/gad.457807.
911 42. Woodward, A.M., Gohler, T., Luciani, M.G., Oehlmann, M., Ge, X., Gartner, A., Jackson, D.A.,
912 and Blow, J.J. (2006). Excess Mcm2-7 license dormant origins of replication that can be used
913 under conditions of replicative stress. J Cell Biol *173*, 673-683. 10.1083/jcb.200602108.
914 43. Ibarra, A., Schwob, E., and Mendez, J. (2008). Excess MCM proteins protect human cells from
915 replicative stress by licensing backup origins of replication. Proc Natl Acad Sci U S A *105*,
916 8956-8961. 10.1073/pnas.0803978105.
917 44. Gros, J., Kumar, C., Lynch, G., Yadav, T., Whitehouse, I., and Remus, D. (2015). Post-licensing
918 Specification of Eukaryotic Replication Origins by Facilitated Mcm2-7 Sliding along DNA. Mol
919 Cell *60*, 797-807. 10.1016/j.molcel.2015.10.022.
920 45. Scherr, M.J., Wahab, S.A., Remus, D., and Duderstadt, K.E. (2022). Mobile origin-licensing
921 factors confer resistance to conflicts with RNA polymerase. Cell Reports *38*, 110531.
922 https://doi.org/10.1016/j.celrep.2022.110531.
923 46. Kirstein, N., Buschle, A., Wu, X., Krebs, S., Blum, H., Kremmer, E., Vorberg, I.M.,
924 Hammerschmidt, W., Lacroix, L., Hyrien, O., et al. (2021). Human ORC/MCM density is low in
925 active genes and correlates with replication time but does not delimit initiation zones. eLife
926 *10*, e62161. 10.7554/eLife.62161.

47. Petryk, N., Kahli, M., d'Aubenton-Carafa, Y., Jaszczyszyn, Y., Shen, Y., Silvain, M., Thermes, C., Chen, C.L., and Hyrien, O. (2016). Replication landscape of the human genome. Nat Commun *7*, 10208. 10.1038/ncomms10208.
48. Hiraga, S.I., Ly, T., Garzon, J., Horejsi, Z., Ohkubo, Y.N., Endo, A., Obuse, C., Boulton, S.J., Lamond, A.I., and Donaldson, A.D. (2017). Human RIF1 and protein phosphatase 1 stimulate DNA replication origin licensing but suppress origin activation. EMBO Rep *18*, 403-419. 10.15252/embr.201641983.
49. Dowling, M.R., Kan, A., Heinzel, S., Zhou, J.H.S., Marchingo, J.M., Wellard, C.J., Markham, J.F., and Hodgkin, P.D. (2014). Stretched cell cycle model for proliferating lymphocytes. Proceedings of the National Academy of Sciences *111*, 6377-6382. doi:10.1073/pnas.1322420111.
50. Zhang, J., MacLennan, I.C.M., Liu, Y.-J., and Lane, P.J.L. (1988). Is rapid proliferation in B centroblasts linked to somatic mutation in memory B cell clones? Immunology Letters *18*, 297-299. https://doi.org/10.1016/0165-2478(88)90178-2.
51. Liu, Y.-J., Zhang, J., Lane, P.J.L., Chan, E.Y.-T., and Maclennan, I.C.M. (1991). Sites of specific B cell activation in primary and secondary responses to T cell-dependent and T cell-independent antigens. European Journal of Immunology *21*, 2951-2962. https://doi.org/10.1002/eji.1830211209.
52. Liu, M., Duke, J.L., Richter, D.J., Vinuesa, C.G., Goodnow, C.C., Kleinstein, S.H., and Schatz, D.G. (2008). Two levels of protection for the B cell genome during somatic hypermutation. Nature *451*, 841-845.
53. Alvarez-Prado, A.F., Perez-Duran, P., Perez-Garcia, A., Benguria, A., Torroja, C., de Yebenes, V.G., and Ramiro, A.R. (2018). A broad atlas of somatic hypermutation allows prediction of activation-induced deaminase targets. J Exp Med *215*, 761-771. 10.1084/jem.20171738.
54. Pasqualucci, L., Neumeister, P., Goossens, T., Nanjangud, G., Chaganti, R.S., Kuppers, R., and Dalla-Favera, R. (2001). Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. Nature *412*, 341-346. 10.1038/35085588 35085588 [pii].
55. Shen, H.M., Peters, A., Baron, B., Zhu, X., and Storb, U. (1998). Mutation of BCL-6 gene in normal B cells by the process of somatic hypermutation of Ig genes. Science *280*, 1750-1752.
56. Nussenzweig, A., and Nussenzweig, M.C. (2010). Origin of chromosomal translocations in lymphoid cancer. Cell *141*, 27-38. S0092-8674(10)00289-8 [pii]10.1016/j.cell.2010.03.016.
57. Casellas, R., Basu, U., Yewdell, W.T., Chaudhuri, J., Robbiani, D.F., and Di Noia, J.M. (2016). Mutations, kataegis and translocations in B cells: understanding AID promiscuous activity. Nat Rev Immunol *16*, 164-176. 10.1038/nri.2016.2.
58. Basso, K., and Dalla-Favera, R. (2015). Germinal centres and B cell lymphomagenesis. Nat Rev Immunol *15*, 172-184. 10.1038/nri3814.
59. Methot, S.P., and Di Noia, J.M. (2017). Molecular Mechanisms of Somatic Hypermutation and Class Switch Recombination. Adv Immunol *133*, 37-87. 10.1016/bs.ai.2016.11.002.
60. Zeman, M.K., and Cimprich, K.A. (2014). Causes and consequences of replication stress. Nature Cell Biology *16*, 2-9. 10.1038/ncb2897.
61. Mahat, D.B., Kwak, H., Booth, G.T., Jonkers, I.H., Danko, C.G., Patel, R.K., Waters, C.T., Munson, K., Core, L.J., and Lis, J.T. (2016). Base-pair-resolution genome-wide mapping of active RNA polymerases using precision nuclear run-on (PRO-seq). Nat Protoc *11*, 1455-1476. 10.1038/nprot.2016.086.
62. Di Tommaso, P., Chatzou, M., Floden, E.W., Barja, P.P., Palumbo, E., and Notredame, C. (2017). Nextflow enables reproducible computational workflows. Nature Biotechnology *35*, 316-319. 10.1038/nbt.3820.
63. Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Research *44*, W160-W165. 10.1093/nar/gkw257.

977 64. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing
978 reads. 2011 *17*, 3. 10.14806/ej.17.1.200.
979 65. Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient
980 alignment of short DNA sequences to the human genome. Genome Biology *10*, R25.
981 10.1186/gb-2009-10-3-r25.
982 66. Foti, R., Gnan, S., Cornacchia, D., Dileep, V., Bulut-Karslioglu, A., Diehl, S., Buness, A., Klein,
983 Felix A., Huber, W., Johnstone, E., et al. (2016). Nuclear Architecture Organized by Rif1
984 Underpins the Replication-Timing Program. Molecular Cell *61*, 260-273.
985 https://doi.org/10.1016/j.molcel.2015.12.001.
986 67. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C.,
987 Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based Analysis of ChIP-Seq
988 (MACS). Genome Biology *9*, R137. 10.1186/gb-2008-9-9-r137.
989 68. Lund, E., Oldenburg, A.R., and Collas, P. (2014). Enriched domain detector: a program for
990 detection of wide genomic enrichment domains robust against local variations. Nucleic Acids
991 Research *42*, e92-e92. 10.1093/nar/gku324.
992 69. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing
993 genomic features. Bioinformatics *26*, 841-842. 10.1093/bioinformatics/btq033.

994

Figure 1

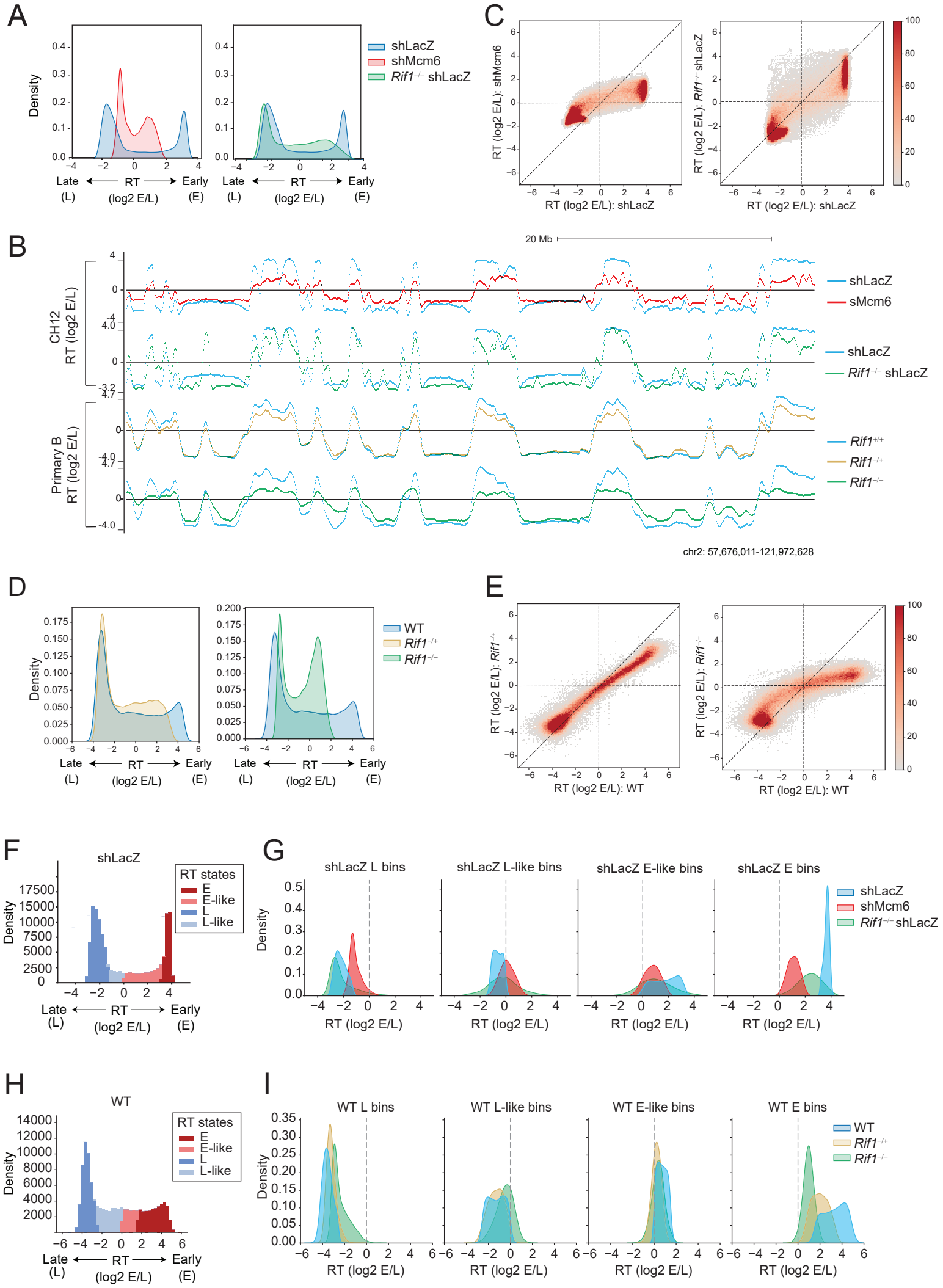

A
Density
shLacZ
shMcm6
Rif1−/− shLacZ
Late (L) ← RT (log2 E/L) → Early (E)
C
RT (log2 E/L): shMcm6
RT (log2 E/L): Rif1−/− shLacZ
RT (log2 E/L): shLacZ
B
20 Mb
CH12 RT (log2 E/L)
Primary B RT (log2 E/L)
shLacZ
sMcm6
Rif1−/− shLacZ
Rif1+/+
Rif1−/+
Rif1−/−
chr2: 57,676,011-121,972,628
D
WT
Rif1−/+
Rif1−/−
E
RT (log2 E/L): Rif1−/+
RT (log2 E/L): Rif1−/−
RT (log2 E/L): WT
F
shLacZ
RT states
E
E-like
L
L-like
G
shLacZ L bins
shLacZ L-like bins
shLacZ E-like bins
shLacZ E bins
RT (log2 E/L)
H
WT
I
WT L bins
WT L-like bins
WT E-like bins
WT E bins

Figure 2

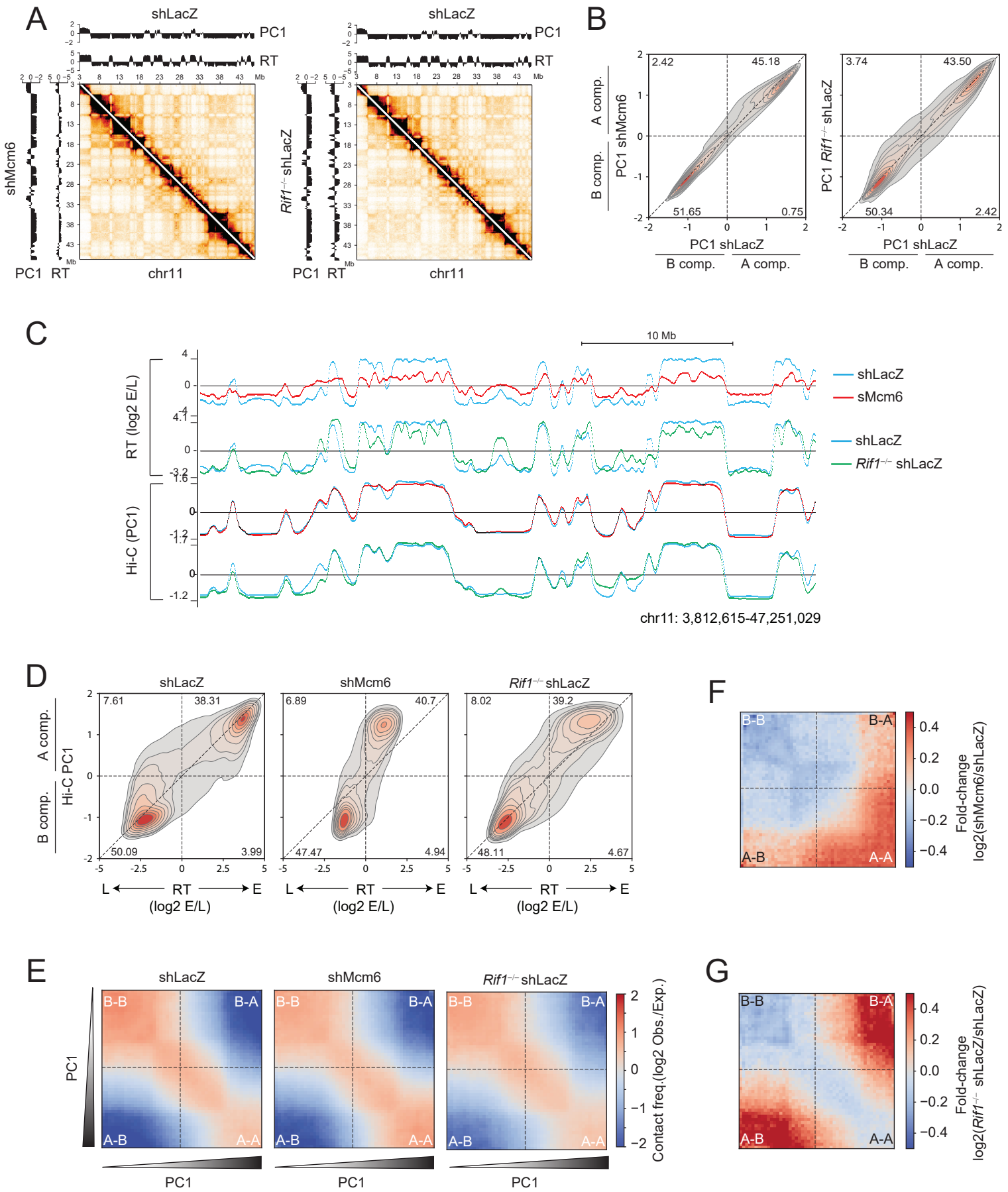

A
shLacZ
PC1
RT
shMcm6
Rif1−/− shLacZ
chr11
B
A comp.
B comp.
PC1 shMcm6
PC1 shLacZ
PC1 Rif1−/− shLacZ
2.42
45.18
51.65
0.75
3.74
43.50
50.34
2.42
C
10 Mb
RT (log2 E/L)
Hi-C (PC1)
shLacZ
sMcm6
Rif1−/− shLacZ
chr11: 3,812,615-47,251,029
D
shLacZ
shMcm6
Rif1−/− shLacZ
Hi-C PC1
7.61
38.31
50.09
3.99
6.89
40.7
47.47
4.94
8.02
39.2
48.11
4.67
L ← RT → E
(log2 E/L)
E
B-B
B-A
A-B
A-A
PC1
Contact freq.(log2 Obs./Exp.)
F
Fold-change
log2(shMcm6/shLacZ)
G
Fold-change
log2(Rif1−/− shLacZ/shLacZ)

Figure 3

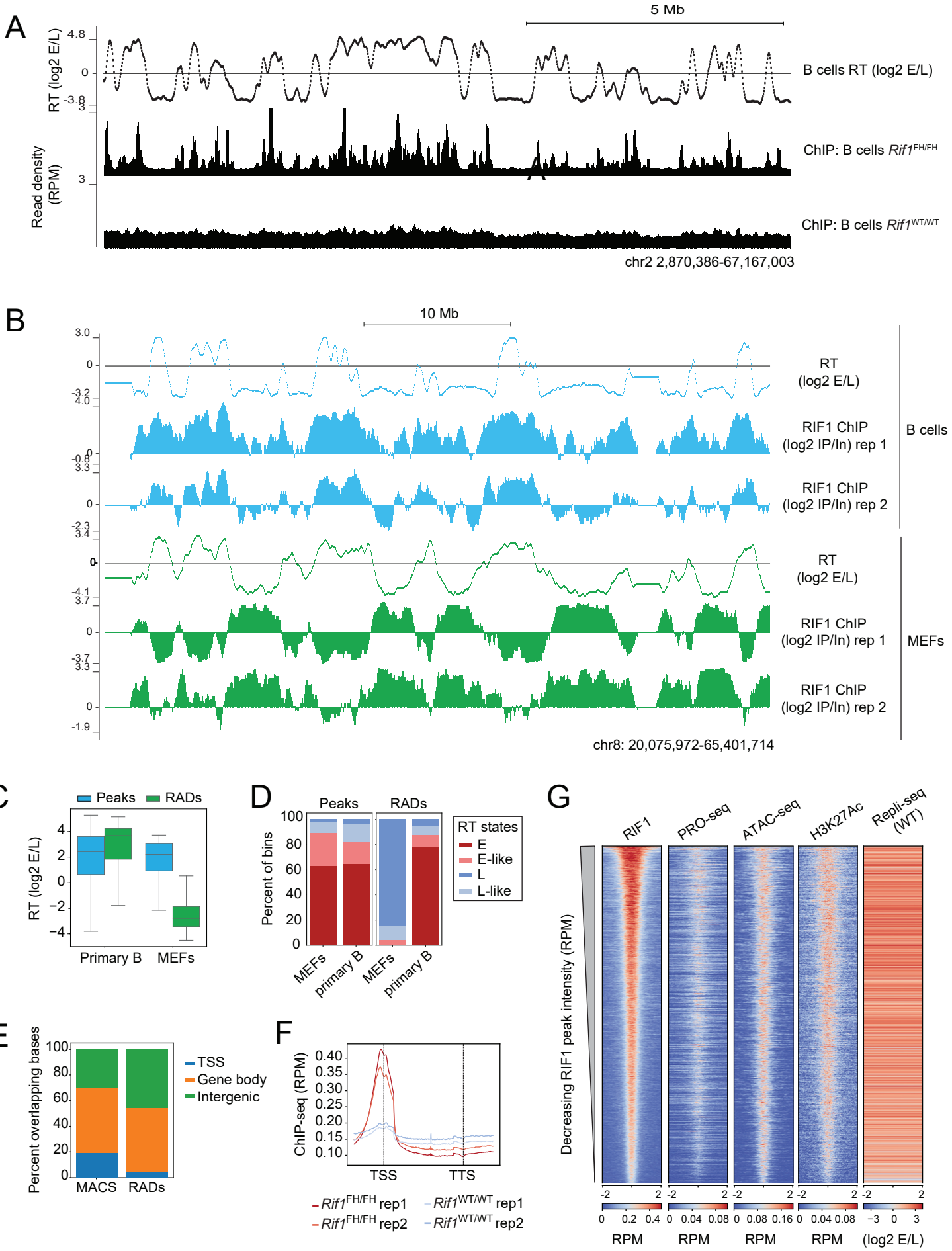

A
5 Mb
RT (log2 E/L)
4.8
0
-3.8
B cells RT (log2 E/L)
Read density (RPM)
3
ChIP: B cells Rif1FH/FH
ChIP: B cells Rif1WT/WT
chr2 2,870,386-67,167,003
B
10 Mb
RT (log2 E/L)
RIF1 ChIP (log2 IP/In) rep 1
RIF1 ChIP (log2 IP/In) rep 2
B cells
MEFs
chr8: 20,075,972-65,401,714
C
Peaks
RADs
RT (log2 E/L)
Primary B
MEFs
D
Peaks
RADs
Percent of bins
RT states
E
L-like
MEFs
primary B
E
Percent overlapping bases
TSS
Gene body
Intergenic
MACS
RADs
F
ChIP-seq (RPM)
TSS
TTS
Rif1FH/FH rep1
Rif1FH/FH rep2
Rif1WT/WT rep1
Rif1WT/WT rep2
G
RIF1
PRO-seq
ATAC-seq
H3K27Ac
Repli-seq (WT)
Decreasing RIF1 peak intensity (RPM)
RPM
(log2 E/L)

Figure 4

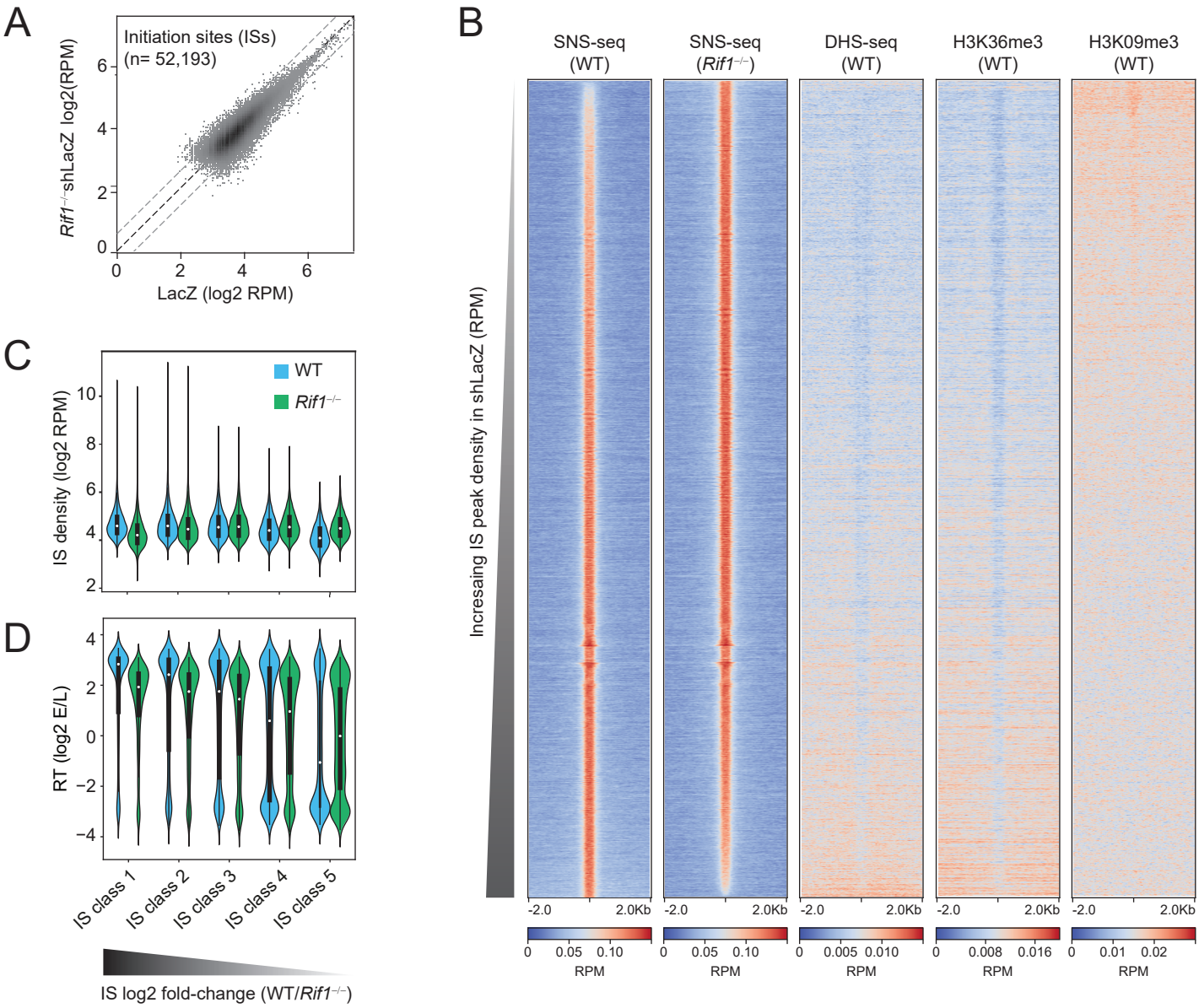

A
Initiation sites (ISs)
(n= 52,193)
Rif1−/−shLacZ log2(RPM)
LacZ (log2 RPM)
B
SNS-seq (WT)
SNS-seq (Rif1−/−)
DHS-seq (WT)
H3K36me3 (WT)
H3K09me3 (WT)
Increasing IS peak density in shLacZ (RPM)
-2.0
2.0Kb
RPM
C
WT
Rif1−/−
IS density (log2 RPM)
D
RT (log2 E/L)
IS class 1
IS class 2
IS class 3
IS class 4
IS class 5
IS log2 fold-change (WT/Rif1−/−)

Figure 5

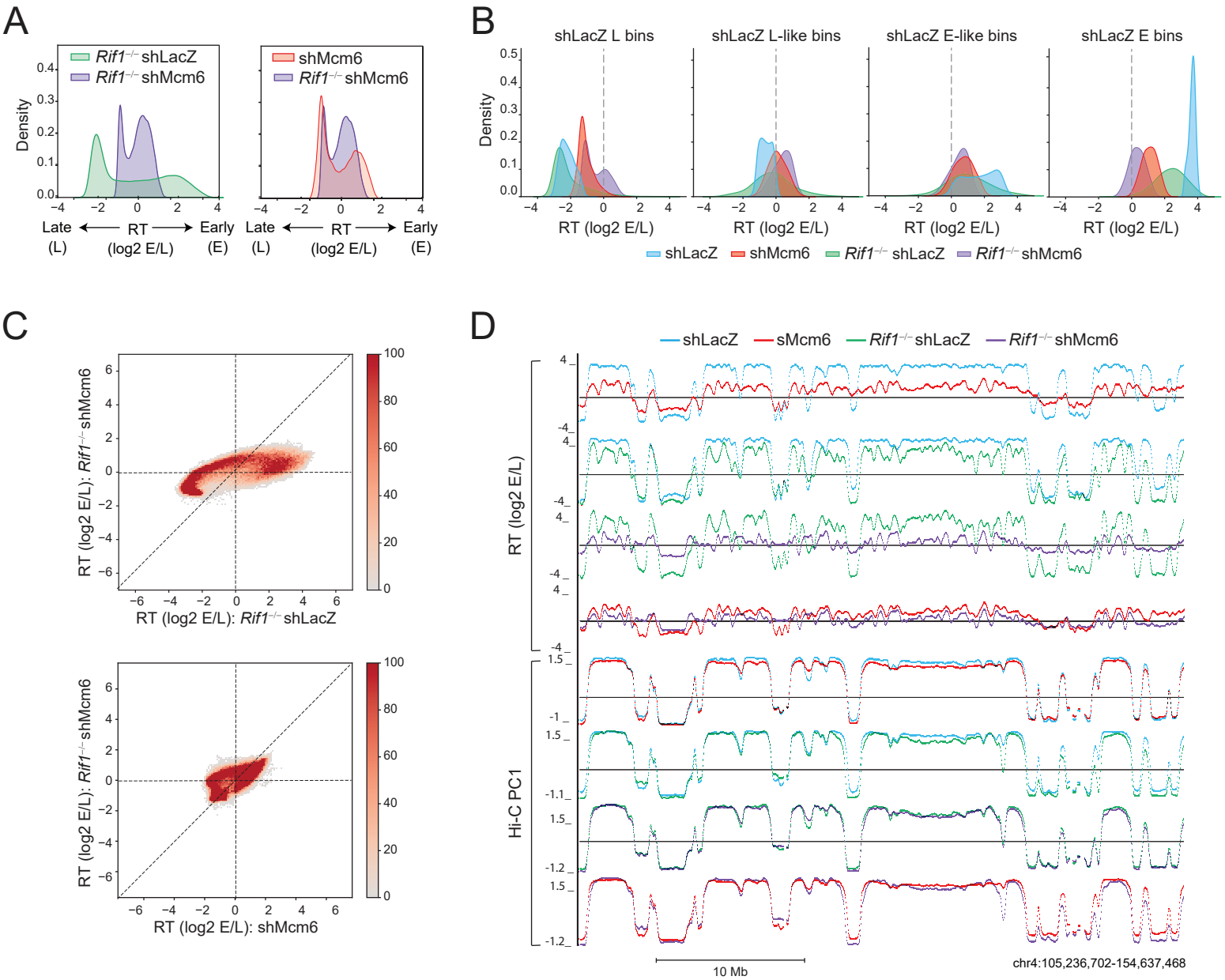

A
Rif1−/− shLacZ
Rif1−/− shMcm6
shMcm6
Rif1−/− shMcm6
Density
Late (L) ← RT (log2 E/L) → Early (E)
B
shLacZ L bins
shLacZ L-like bins
shLacZ E-like bins
shLacZ E bins
Density
RT (log2 E/L)
shLacZ
shMcm6
Rif1−/− shLacZ
Rif1−/− shMcm6
C
RT (log2 E/L): Rif1−/− shMcm6
RT (log2 E/L): Rif1−/− shLacZ
RT (log2 E/L): shMcm6
D
shLacZ
sMcm6
Rif1−/− shLacZ
Rif1−/− shMcm6
RT (log2 E/L)
Hi-C PC1
10 Mb
chr4:105,236,702-154,637,468

Figure 6

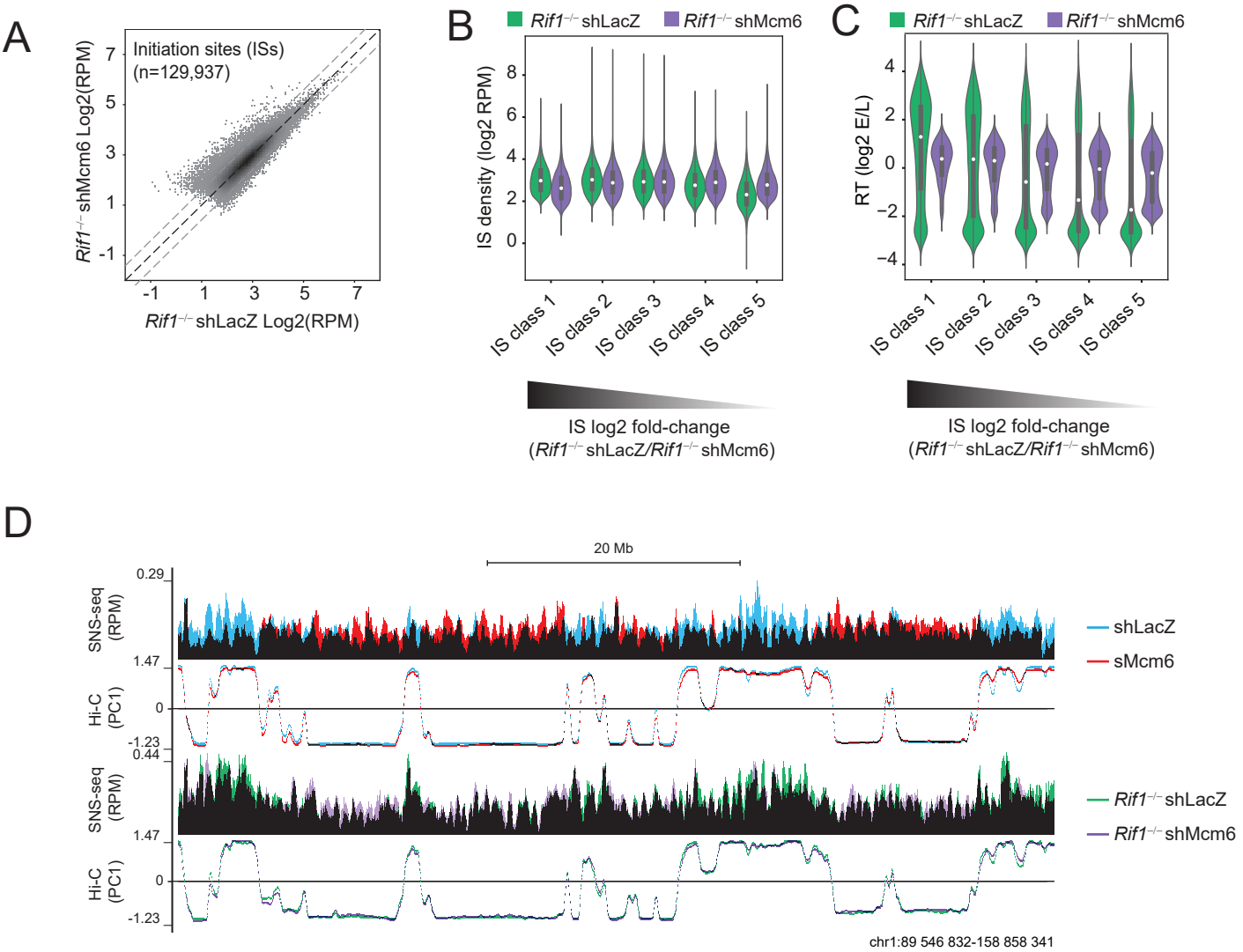

A
Initiation sites (ISs)
(n=129,937)
Rif1−/− shMcm6 Log2(RPM)
Rif1−/− shLacZ Log2(RPM)
B
Rif1−/− shLacZ
Rif1−/− shMcm6
IS density (log2 RPM)
IS class 1
IS class 2
IS class 3
IS class 4
IS class 5
IS log2 fold-change
(Rif1−/− shLacZ/Rif1−/− shMcm6)
C
Rif1−/− shLacZ
Rif1−/− shMcm6
RT (log2 E/L)
IS class 1
IS class 2
IS class 3
IS class 4
IS class 5
IS log2 fold-change
(Rif1−/− shLacZ/Rif1−/− shMcm6)
D
20 Mb
SNS-seq (RPM)
Hi-C (PC1)
shLacZ
sMcm6
Rif1−/− shLacZ
Rif1−/− shMcm6
chr1:89 546 832-158 858 341

Figure 7

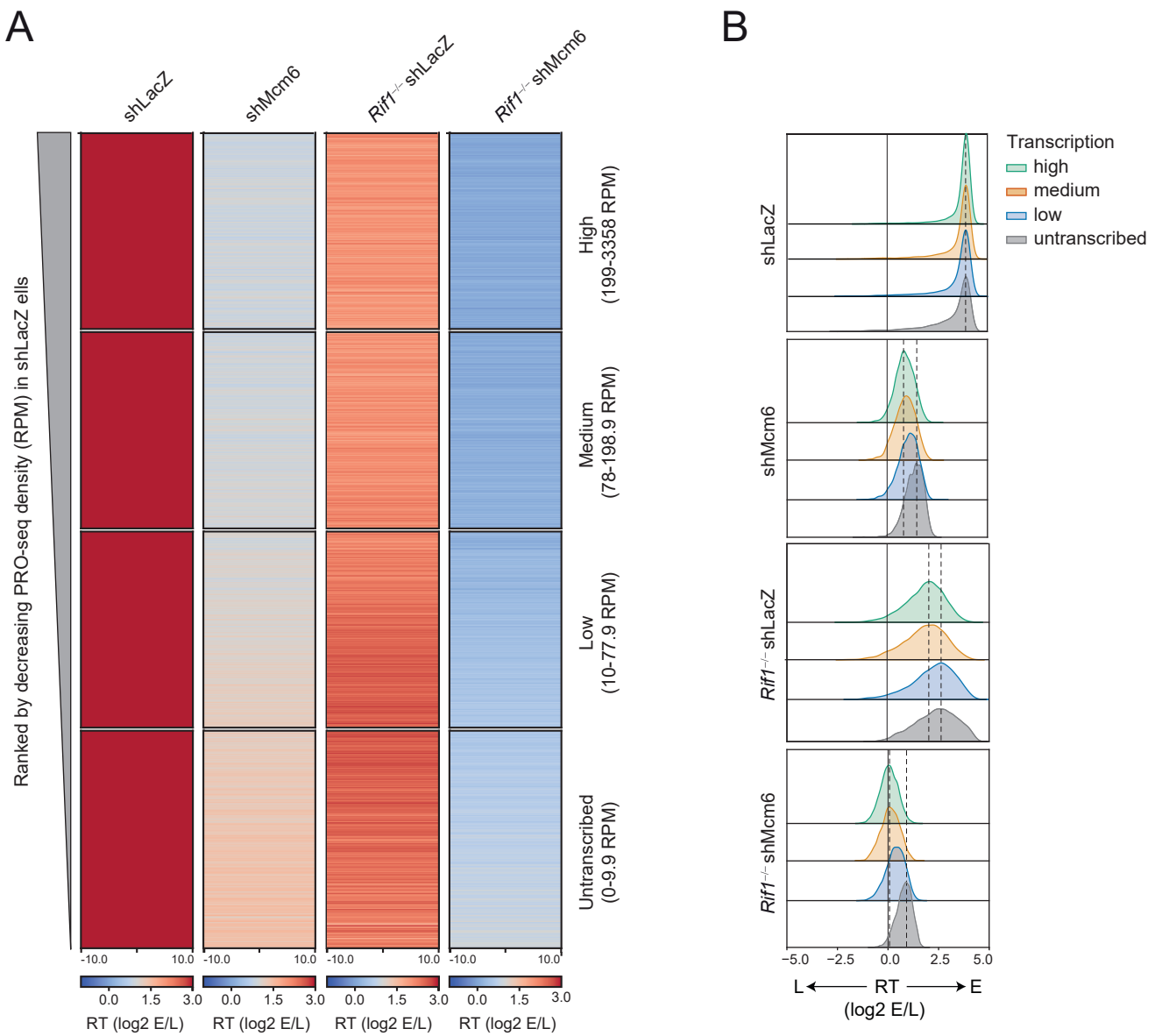

A
shLacZ
shMcm6
Rif1−/− shLacZ
Rif1−/− shMcm6
Ranked by decreasing PRO-seq density (RPM) in shLacZ ells
High (199-3358 RPM)
Medium (78-198.9 RPM)
Low (10-77.9 RPM)
Untranscribed (0-9.9 RPM)
-10.0
10.0
0.0
1.5
3.0
RT (log2 E/L)
B
Transcription
high
medium
low
untranscribed
shLacZ
shMcm6
Rif1−/− shLacZ
Rif1−/− shMcm6
−5.0
−2.5
0.0
2.5
5.0
L
RT
E
(log2 E/L)

Figure S1

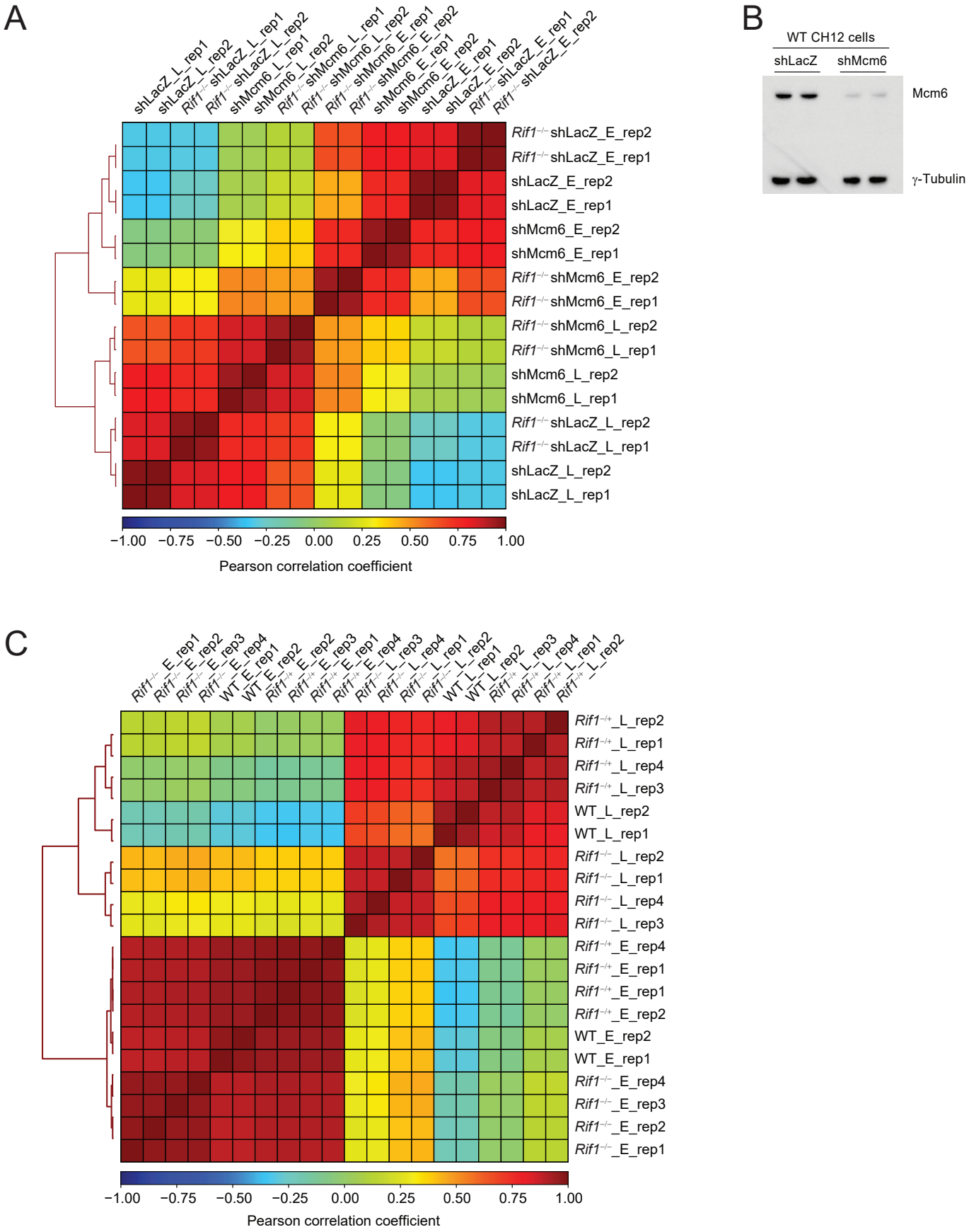
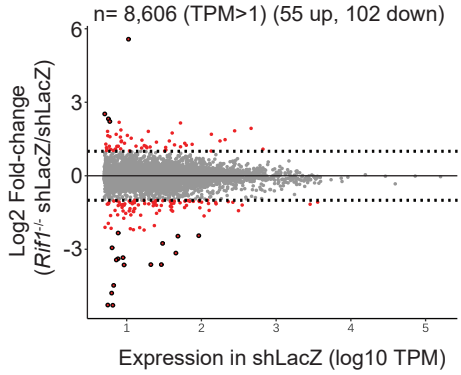
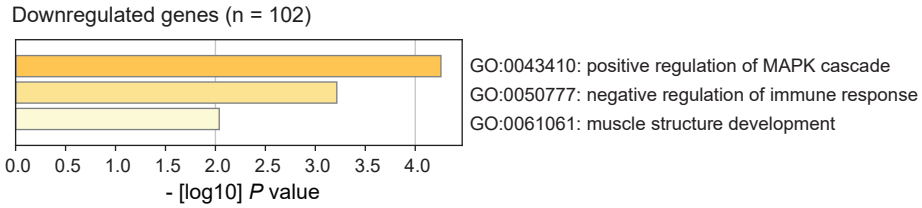

A
shLacZ_L_rep1
shLacZ_L_rep2
Rif1−/− shLacZ_L_rep1
Rif1−/− shLacZ_L_rep2
shMcm6_L_rep1
shMcm6_L_rep2
Rif1−/− shMcm6_L_rep1
Rif1−/− shMcm6_L_rep2
shMcm6_E_rep1
shMcm6_E_rep2
Rif1−/− shMcm6_E_rep1
Rif1−/− shMcm6_E_rep2
shLacZ_E_rep1
shLacZ_E_rep2
Rif1−/− shLacZ_E_rep1
Rif1−/− shLacZ_E_rep2
Rif1−/− shLacZ_E_rep2
Rif1−/− shLacZ_E_rep1
shLacZ_E_rep2
shLacZ_E_rep1
shMcm6_E_rep2
shMcm6_E_rep1
Rif1−/− shMcm6_E_rep2
Rif1−/− shMcm6_E_rep1
Rif1−/− shMcm6_L_rep2
Rif1−/− shMcm6_L_rep1
shMcm6_L_rep2
shMcm6_L_rep1
Rif1−/− shLacZ_L_rep2
Rif1−/− shLacZ_L_rep1
shLacZ_L_rep2
shLacZ_L_rep1
−1.00 −0.75 −0.50 −0.25 0.00 0.25 0.50 0.75 1.00
Pearson correlation coefficient
B
WT CH12 cells
shLacZ
shMcm6
Mcm6
γ-Tubulin
C
Rif1−/− E_rep1
Rif1−/− E_rep2
Rif1−/− E_rep3
Rif1−/− E_rep4
WT_E_rep1
WT_E_rep2
Rif1−/+ E_rep2
Rif1−/+ E_rep3
Rif1−/+ E_rep1
Rif1−/+ E_rep4
Rif1−/− L_rep3
Rif1−/− L_rep4
Rif1−/− L_rep1
Rif1−/− L_rep2
WT_L_rep1
WT_L_rep2
Rif1−/+ L_rep3
Rif1−/+ L_rep4
Rif1−/+ L_rep1
Rif1−/+ L_rep2
Rif1−/+_L_rep2
Rif1−/+_L_rep1
Rif1−/+_L_rep4
Rif1−/+_L_rep3
WT_L_rep2
WT_L_rep1
Rif1−/−_L_rep2
Rif1−/−_L_rep1
Rif1−/−_L_rep4
Rif1−/−_L_rep3
Rif1−/+_E_rep4
Rif1−/+_E_rep1
Rif1−/+_E_rep1
Rif1−/+_E_rep2
WT_E_rep2
WT_E_rep1
Rif1−/−_E_rep4
Rif1−/−_E_rep3
Rif1−/−_E_rep2
Rif1−/−_E_rep1
−1.00 −0.75 −0.50 −0.25 0.00 0.25 0.50 0.75 1.00
Pearson correlation coefficient

Figure S2

A


n= 8,606 (TPM>1) (55 up, 102 down)
Log2 Fold-change
(Rif1-/- shLacZ/shLacZ)
6
3
0
-3
1
2
3
4
5
Expression in shLacZ (log10 TPM)

B


Downregulated genes (n = 102)
GO:0043410: positive regulation of MAPK cascade
GO:0050777: negative regulation of immune response
GO:0061061: muscle structure development
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
- [log10] P value

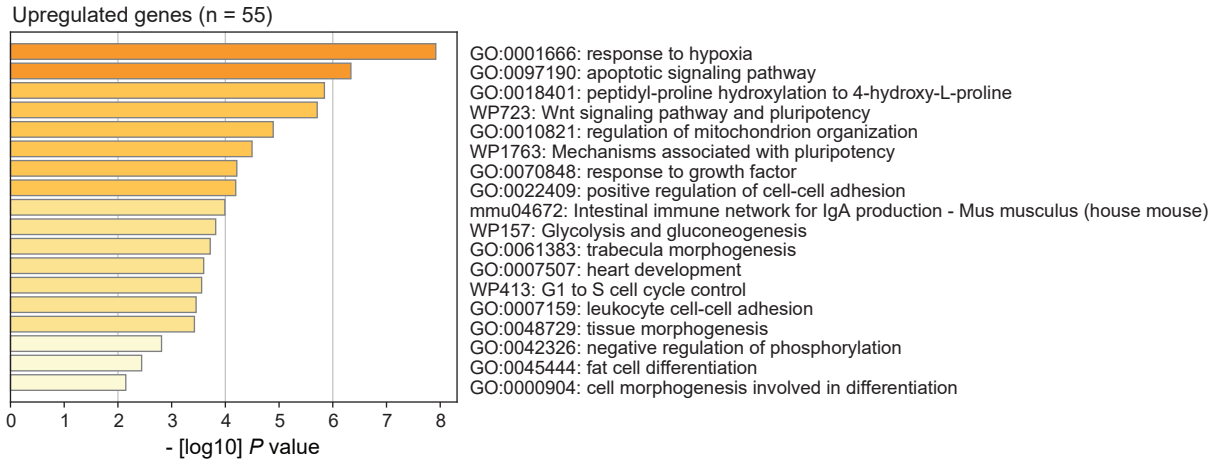

Upregulated genes (n = 55)
GO:0001666: response to hypoxia
GO:0097190: apoptotic signaling pathway
GO:0018401: peptidyl-proline hydroxylation to 4-hydroxy-L-proline
WP723: Wnt signaling pathway and pluripotency
GO:0010821: regulation of mitochondrion organization
WP1763: Mechanisms associated with pluripotency
GO:0070848: response to growth factor
GO:0022409: positive regulation of cell-cell adhesion
mmu04672: Intestinal immune network for IgA production - Mus musculus (house mouse)
WP157: Glycolysis and gluconeogenesis
GO:0061383: trabecula morphogenesis
GO:0007507: heart development
WP413: G1 to S cell cycle control
GO:0007159: leukocyte cell-cell adhesion
GO:0048729: tissue morphogenesis
GO:0042326: negative regulation of phosphorylation
GO:0045444: fat cell differentiation
GO:0000904: cell morphogenesis involved in differentiation
0
1
2
3
4
5
6
7
8
- [log10] P value

Figure S3

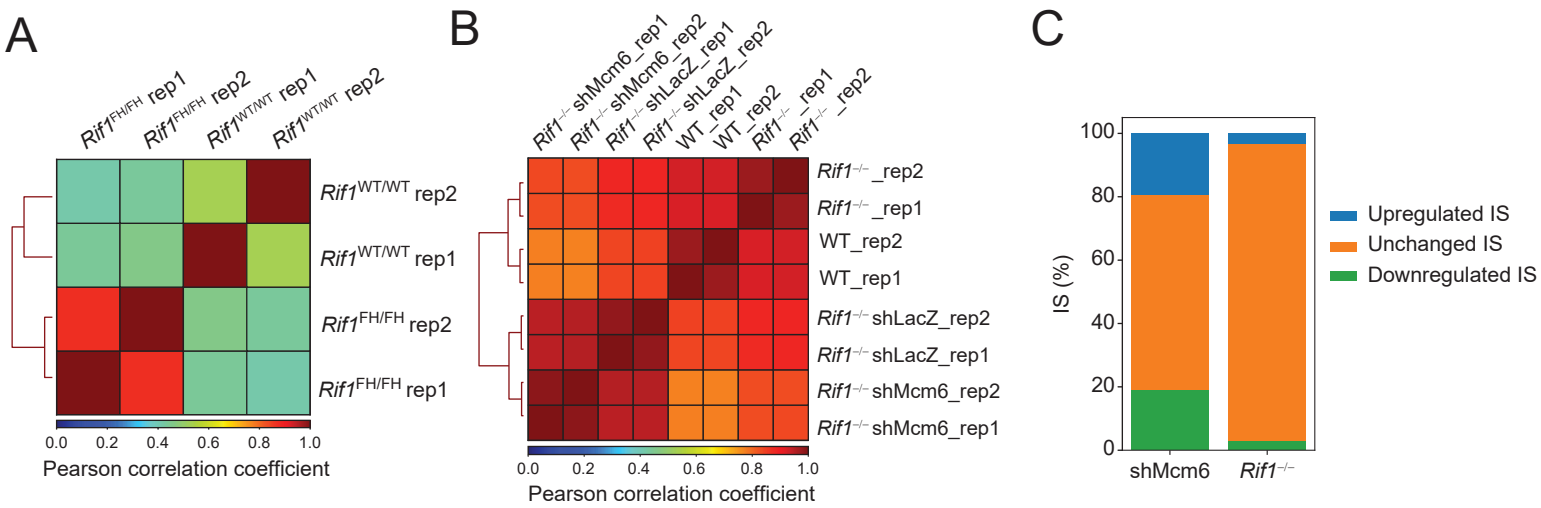
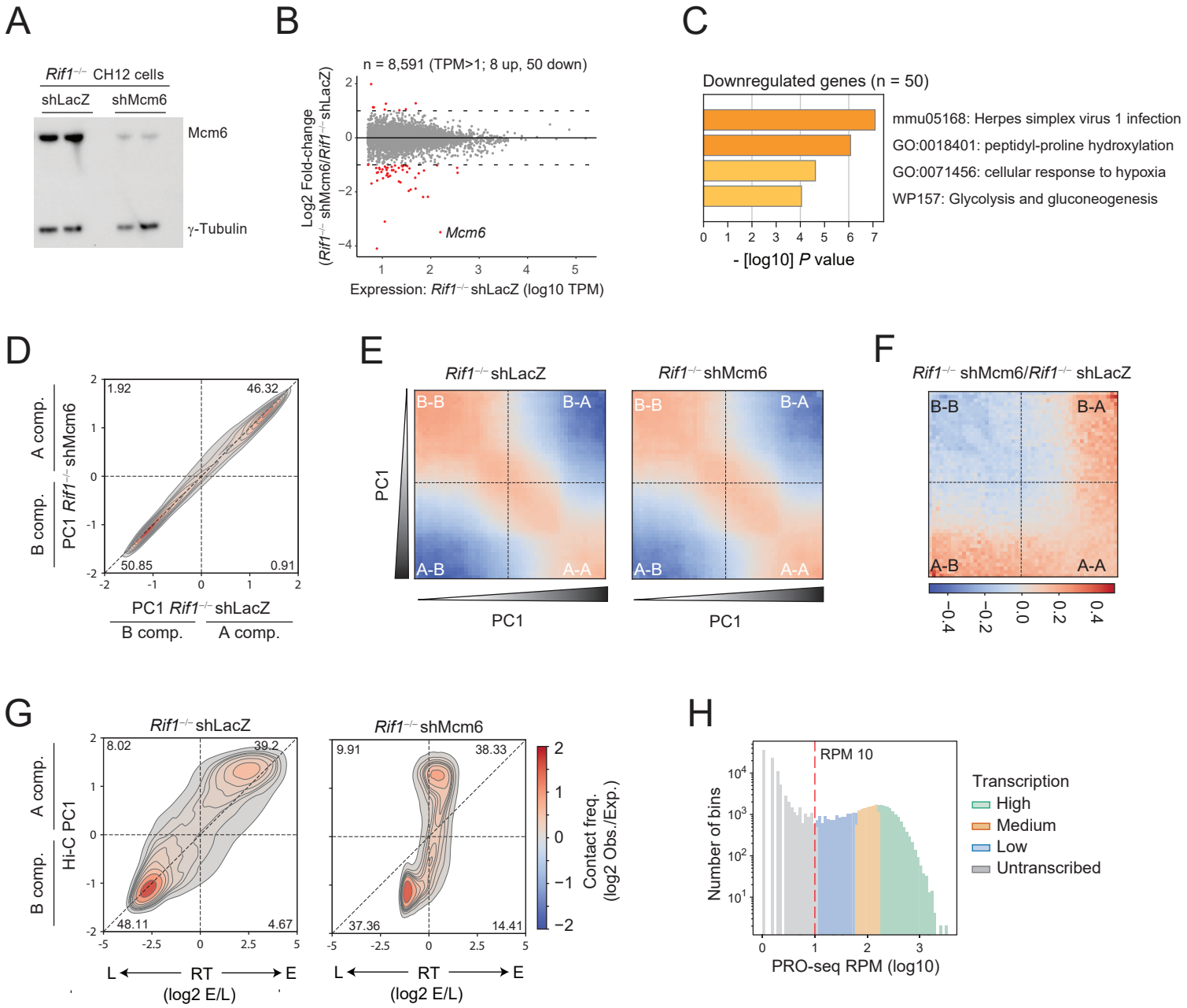

A
Rif1FH/FH rep1
Rif1FH/FH rep2
Rif1WT/WT rep1
Rif1WT/WT rep2
Rif1WT/WT rep2
Rif1WT/WT rep1
Rif1FH/FH rep2
Rif1FH/FH rep1
0.0 0.2 0.4 0.6 0.8 1.0
Pearson correlation coefficient
B
Rif1−/− shMcm6_rep1
Rif1−/− shMcm6_rep2
Rif1−/− shLacZ_rep1
Rif1−/− shLacZ_rep2
WT_rep1
WT_rep2
Rif1−/− _rep1
Rif1−/− _rep2
Rif1−/− _rep2
Rif1−/− _rep1
WT_rep2
WT_rep1
Rif1−/− shLacZ_rep2
Rif1−/− shLacZ_rep1
Rif1−/− shMcm6_rep2
Rif1−/− shMcm6_rep1
0.0 0.2 0.4 0.6 0.8 1.0
Pearson correlation coefficient
C
IS (%)
100
80
60
40
20
0
shMcm6
Rif1−/−
Upregulated IS
Unchanged IS
Downregulated IS

Figure S4


A
Rif1−/− CH12 cells
shLacZ
shMcm6
Mcm6
γ-Tubulin
B
n = 8,591 (TPM>1; 8 up, 50 down)
Log2 Fold-change (Rif1−/− shMcm6/Rif1−/− shLacZ)
Mcm6
Expression: Rif1−/− shLacZ (log10 TPM)
C
Downregulated genes (n = 50)
mmu05168: Herpes simplex virus 1 infection
GO:0018401: peptidyl-proline hydroxylation
GO:0071456: cellular response to hypoxia
WP157: Glycolysis and gluconeogenesis
- [log10] P value
D
A comp.
B comp.
PC1 Rif1−/− shMcm6
PC1 Rif1−/− shLacZ
1.92
46.32
50.85
0.91
E
Rif1−/− shLacZ
Rif1−/− shMcm6
B-B
B-A
A-B
A-A
PC1
F
Rif1−/− shMcm6/Rif1−/− shLacZ
G
Rif1−/− shLacZ
Rif1−/− shMcm6
Hi-C PC1
8.02
39.2
48.11
4.67
9.91
38.33
37.36
14.41
Contact freq. (log2 Obs./Exp.)
L ← RT → E
(log2 E/L)
H
RPM 10
Number of bins
PRO-seq RPM (log10)
Transcription
High
Medium
Low
Untranscribed