The SWI/SNF chromatin remodeling complex helps resolve R-loop-mediated transcription-replication conflicts

Aleix Bayona-Feliu[1,2], Sonia Barroso[1], Sergio Muñoz[1] and Andrés Aguilera[1,2,] *

[1] *Centro Andaluz de Biología Molecular y Medicina Regenerativa CABIMER, Universidad de Sevilla-CSIC-Universidad Pablo de Olavide, Sevilla 41092, Spain*

[2] *Departamento de Genética, Facultad de Biología, Universidad de Sevilla, Seville 41012, Spain*

* *Corresponding author*. *E-mail:* aguilo@us.es

Word count: 4972

ABSTRACT

ATP-dependent chromatin remodelers are commonly mutated in human cancer. Mammalian SWI/SNF complexes comprise three conserved multi-subunit chromatin remodelers (cBAF, ncBAF and PBAF) that share the BRG1 (also known as SMARCA4) subunit responsible for the main ATPase activity. BRG1 is the most frequently mutated Snf2-like ATPase in cancer. Here we have investigated the role of SWI/SNF in genome instability, a hallmark of cancer cells, given its role in transcription, DNA replication and DNA damage repair. We show that depletion of BRG1 increases R-loops and R-loop-dependent DNA breaks, as well as transcription-replication conflicts. BRG1 colocalizes with R-loops and replication fork blocks, as determined by FANCD2 foci, with BRG1 depletion being epistatic to FANCD2 silencing. Our study, extended to other components of SWI/SNF, uncovers a key role of the SWI/SNF complex, in particular cBAF, in helping resolve R-loop-mediated transcription-replication conflicts; thus, unveiling a novel mechanism by which chromatin remodeling protects genome integrity.

Keywords: SWI/SNF, chromatin remodelers, R-loops, transcription-replication conflicts, BRG1, tumor suppressors, cBAF, Fanconi Anemia

INTRODUCTION

Transcription is a process that requires tight regulation, as it can become an obstacle to replication fork progression. Evidence indicates that transcription-replication (T-R) conflicts are a major spontaneous source of genome instability[1] that in many cases is linked to R-loops, structures containing a DNA-RNA hybrid and the displaced single-stranded DNA (ssDNA) [2,3]. Cells use different strategies to counteract harmful R-loop accumulation and the associated genetic instability by: i) preventing R-loop formation via factors that assemble onto the nascent RNA [4–9] or control topological changes associated with transcription[10],[11]; ii) removing R-loops via nucleases such as RNase H[12] or DNA-RNA helicases such as SETX or UAP56 [7,13,14], and iii) repairing the damage or the T-R conflicts generated by R-loops, thus helping to resolve them [15–18].

Chromatin factors have emerged as relevant players in R-loop homeostasis and its associated genome instability. These factors include the FACT chromatin reorganizing complex and the histone deacetylase complex mSin3A in yeast and human cells[19,20], INO80 and sirtuins in yeast[21,22], histone H1 in *Drosophila* and human cells[23,24], the ATRX chromatin remodeling complex at telomeric repeats[25], and the Tip60-p400 histone acetyltransferase complex that is associated with genes harboring promoter-proximal R-loops and controls the genome-wide occupancy of the PRC2 histone methyl-transferase[26]. However, despite the number of reports relating chromatin factors with R-loop homeostasis, we do not know the molecular basis of this connection or its biological significance.

ATP-dependent chromatin remodeler genes are among the most commonly mutated in cancer[27]. SWI/SNF is the most mutated with ~20% of human malignancies presenting alterations in this complex[28] and its core subunit, BRG1, the most frequently mutated ATPase subunit among the 4 main chromatin remodeling families (SWI/SNF, ISWI, CHD, INO80). SWI/SNF defines a family of highly conserved multisubunit remodelers, originally discovered in *Saccharomyces cerevisiae*[29,30]. Mammalian SWI/SNF complexes include the BRG1-associated factor (BAF), in its canonical (cBAF) and non-canonical (ncBAF) subtypes, and the polybromo BRG1-associated factor (PBAF), all consisting of a central ATPase subunit, a multimeric conserved core and variant subunits different for each complex. BRG1 is the main ATPase in these complexes, even though BRM (also known as SMARCA2) can also exert this function[31,32]. The complexes remodel chromatin by mobilizing nucleosomes through sliding and by ejection and insertion of histone octamers[33], and they also contribute to high-order chromatin structures[34].

Here we show that depletion of BRG1 increases R-loops and R-loop-dependent DNA breaks, as well as T-R conflicts. Our study unveils a new function of SWI/SNF, and in particular cBAF, as a key player in the maintenance of genome stability by helping resolve R-loop-mediated T-R conflicts.

RESULTS

BRG1 suppresses R-loops and associated genome instability

A total of 1,722 mutations has been identified in *BRG1* that accumulate preferentially at the ATPase and helicase domains and *BRG1* mutation frequencies are close to 90% in some cancer subtypes (The cBio Cancer Genomics Portal)[35]. Thus, we assayed the role of BRG1 in DNA damage response (DDR) and genome integrity. Transient depletion of BRG1 by siRNA in HeLa cells (Extended Data Fig. 1a-c) causes DNA break accumulation, as seen by alkaline single-cell electrophoresis (comet assay) and γH2AX foci by immunofluorescence (IF), in comparison with siC control cells (Fig. 1a,b; Extended Data Fig. 1d).

Since chromatin structure has been shown to control R-loop accumulation and its associated genome instability, we wondered whether BRG1 depletion increased R-loops. Analysis of R-loops by IF using the S9.6 anti-DNA-RNA monoclonal antibody revealed a significant increase of the S9.6 nuclear signal in siBRG1 cells compared to control siC HeLa cells, either using a siRNA pool or individually (Fig. 1c; Extended Data Fig. 1e). S9.6 foci could be clearly detected in BRG1-depleted cells accumulating up to 4.5-fold above siC levels. Moreover, a strong increase of S9.6 staining was observed at nucleoli of siBRG1 cells (Extended Data Fig. 1f). In all cases, the increase in S9.6 signals was suppressed upon RNH1 overexpression (Fig.1c, Extended Data Fig. 1f), confirming that they detected DNA-RNA hybrids.

Next, we confirmed the IF result by the more accurate method of DNA-RNA immunoprecipitation followed by qPCR (DRIP-qPCR). DNA-RNA hybrids accumulate up to 2.5 times more in BRG1-depleted cells than in the siC control in all analyzed genes (Fig. 1d, Extended Data Fig. 1g). The hybrid signals were removed by *in vitro* RNH treatment as a confirmation of the specificity of the assay. The specificity of the observed effects on siBRG1 was shown by expressing siRNA-resistant wild-type and catalytically-dead versions of BRG1 in BRG1-depleted cells and analysis of S9.6 reactivity. Consistently, expression of the wild-type, but not mutant, form of BRG1 complemented siBRG1 depletion, seen by a significant decrease in S9.6 nuclear signal (Extended Data Fig. 1h).

We asked next whether DNA break increase in siBRG1 cells was R-loop-dependent. We overexpressed RNH1 in siBRG1 cells and found that the comet tail moments and γH2AX foci were significantly reduced (Fig. 1b,c; Extended Data Fig. 1d). Importantly, RNaseH1 overexpression did not alter cell cycle progression either in siC or siBRG1 cells, excluding the possibility that rescue of the DNA damage phenotypes could be attributed to S-phase exit (Extended Data Fig. 1i). Therefore, DNA damage generated in siBRG1 cells is R-loop-dependent, a conclusion further supported by other genome instability phenotypes. Anaphase bridges and micronuclei were increased in siBRG1 cells, consistent with previous data[36]; both increases were suppressed by RNH1 overexpression (Fig. 1e,f). These bridges frequently presented a clear S9.6 staining signal observed as bright dots (Fig. 1f). Moreover, nucleoli presented aberrant shape and ectopic nucleolin foci, a fraction of which colocalized with S9.6 foci and were sensitive to RNH1 overexpression (Fig. 1g). Altogether, the results indicate that BRG1 protects cells from unscheduled R-loops, R-loop-mediated DNA damage and genome instability throughout the nucleus, including the nucleolus.

BRG1 depletion-induced R-loops block RF progression

R-loop-mediated DNA breaks arise mainly as a consequence of replication fork stalling, as they pose an obstacle to fork progression[1]. Consequently, we analyzed whether R-loops impair DNA replication in siBRG1 cells as a major mechanism for damage. First, EdU (5-ethynyl-2'-desoxyuridine) incorporation, used as a measure of DNA synthesis, was decreased in siBRG1 cells versus siC control cells (Fig. 2a). The percentage of siBRG1 cells incorporating EdU and the EdU intake were diminished; both phenotypes were partially but significantly rescued by RNH1 overexpression (Fig. 2a; Extended Data Fig. 2a). Further, DNA replication analysis by DNA combing using a double pulse-labelling with two consecutive thymidine analogues (IdU and CldU), revealed significant decreases in replication fork velocity and increases in replication fork asymmetry in siBRG1 cells versus the siC control (Fig. 2b). Importantly, both effects were suppressed by RNH1 overexpression, so that fork velocity and asymmetry recovered values closer to siC levels, indicating that R-loops accumulated in siBRG1 cells impair replication fork progression.

To confirm that replication forks stalled at R-loops, as shown by fork asymmetry, we analyzed the appearance of FANCD2 and BLM[37,38] foci, which are known to accumulate at replication fork stalls. Cells with FANCD2 foci increased from 5.9% in siC control cells to 13.8% in siBRG1 cells (Fig. 2c, Extended Data Fig. 2b), and cells with BLM foci increased from 7.8% in siC cells to 46.5% in siBRG1 cells (Fig. 2d, Extended Data Fig. 2c). In both cases, cells accumulating foci were undergoing DNA

replication, as most of them were EdU-positive (Extended Data Fig. 2d,e). Upon RNH1 overexpression, both FANCD2 or BLM foci and the percentage of cells containing those foci were significantly reduced in siBRG1 cells to levels close to the siC control (Fig. 2c,d, Extended Data Fig. 2b,c). FANCD2 and BLM foci were often found at lighter DAPI areas in some cells depleted for BRG1 (Extended Data Fig. 2f,g). These areas correspond mainly to nucleoli, suggesting that DNA replication is also strongly affected at nucleoli.

Given the strong R-loop-dependent phenotype of DNA replication, we assayed R-loop and BRG1 levels throughout the cell cycle by high-throughput microscope imaging and flow cytometry. We performed IF using S9.6 and anti-BRG1 antibodies plus DAPI to identify the specific cell cycle-phase of each cell. The high number of processed cells allowed us to conclude that the main increase in S9.6 signal occurs during S-phase in BRG1-depleted cells versus control cells (Fig. 2e, Extended Data Fig. 2h). Consistently, an increase in S9.6-dependent fluorescence from G1 to S/G2 was also observed for BRG1-depleted cells by flow cytometry (Extended Data Fig. 2i). We noted that the BRG1 nuclear content was higher during the S-phase in the siC control cells (Extended Data Fig. 2j). This, together with the S-phase-specific increase of the S9.6 signal and the R-loop-mediated replication fork stalling caused by BRG1 depletion, prompted us to test whether BRG1 was enriched at replicating sites by proximity ligation assay (PLA) between BRG1 and PCNA, a key component of the replisome. Indeed, the use of anti-BRG1 and anti-PCNA antibodies resulted in nuclear PLA foci (Extended Data Fig. 2k), strengthening the conclusion that BRG1 is enriched at fork stalls.

Altogether, our results indicate that BRG1 depletion causes an increase in R-loops during S-phase, causing replication fork stalls that subsequently lead to DNA breaks throughout the nucleus.

Genome-wide accumulation of DNA-RNA hybrids in siBRG1 cells

To evaluate how general the role of BRG1 in protecting the genome from R-loop accumulation is, we investigated the R-loop profile along the genome in BRG1-depleted K562 cells, a cell line used as a standard for genome-wide studies[39]. First, we tested that immunoprecipitation with S9.6 in these analyses was specific for DNA-RNA hybrids, i.e. RNH-sensitive, via DRIP-seq and DRIP-qPCRs at different genomic loci (Fig. 3a, Extended Data Fig. 3a,b). Once we confirmed the reliability of our immunoprecipitation method, we performed in-depth analysis of R-loop distribution by DRIPc-seq (DRIP followed by RNA purification and stranded cDNA sequencing), a

more informative method than DRIP-seq because it provides strand-specific information, which we recently validated in our systems[14].

DRIPc-seq data in siC control cells identified the R-loop-prone regions along the genome that largely correlate with RNH-sensitive DRIP-seq data (Fig. 3a; Extended Data Fig. 3c). Consistent with previous results[40], the median size of R-loop peaks was ~1kb and corresponded largely to genes (Extended Data Fig. 3d,e), concentrating mainly along the gene body, although they are also present at promoters and downstream regions. A comparative analysis of our DRIPc-seq data with RNA-seq[14] and ChIP-seq data of BRG1 from ENCODE[41] shows that R-loop-prone genes co-localize with BRG1 and expressed sequences, consistent with the conclusion that R-loops are co-transcriptional (Fig. 3b,c).

Comparison of DRIPc-seq of siBRG1-depleted cells with siC cells (Fig. 3d) identified 3,200 genomic sites with higher R-loop-content in siBRG1 cells versus siC cells (Extended Data Fig. 3f). 257 regions showed apparent lower R-loop-content in BRG1-depleted cells due to the high presence of sequenced fragments from the enriched regions in the total reads, as the seq analysis is based on the same number or reads per sample. Therefore, *BRG1* knock-down induces a genome-wide increase of R-loops. Among R-loop-enriched sites in siBRG1 cells, 8% appear *de novo* and 92% were sites presenting R-loops in the siC control, but at a lower level (Extended Data Fig. 3g).

The sites where R-loops are increased in siBRG1 cells compared to siC control (R-loop-gain sites) corresponded to protein-coding genes (73.7%), non-codingRNA-coding sequences (16.8%) and pseudogenes (6.6%), or these are not assigned to a known class (2.9%) (Fig. 3e; Extended Data Fig. 3e). Gene-metaplot analysis showed that R-loop-enriched sites mainly corresponded to sense transcripts of gene bodies, although they were also formed by antisense transcripts at promoter and downstream regions (Fig. 3e,f; Extended Data Fig. 3e). Such differences were persistent when Alu sequences were removed, further validating our analysis and DRIPc-seq signals at genes in siBRG1 cells, regardless of Alu signals (Extended Data Fig. 3h,i).

R-loop-gain genes in siBRG1 cells strongly correlate with BRG1-binding genes observed in K562 control cells (Extended Data Fig. 3j,k). This suggests that BRG1 functions directly at these genes preventing R-loop accumulation, and excludes the possibility that the genome-wide R-loop increase seen in siBRG1 cells is a consequence of an off-target effect.

Genes containing R-loop-gain peaks upon BRG1 depletion are not different in length from the whole-genome average and their expression is not significantly affected by BRG1 depletion (Extended Data Fig. 3l,m), implying that R-loop increases

are not determined by changes in transcription levels. However, these genes are significantly enriched in GC-content and have higher GC-skew and expression levels (Extended Data Fig. 3n-p). Consistently, substantial R-loop accumulation in siBRG1 cells occurs towards subtelomeric regions, where GC skew is high, as shown by metachromosome analysis (Extended Data Fig. 3q, upper panel). R-loops also accumulated at rDNA (Extended Data Fig. 3q, lower panel), consistent with the IF and DRIP-qPCR data from HeLa cells (Fig. 1c,d; Extended Data Fig. 1f,g).

We note that most genes accumulating R-loops in siBRG1 cells (2,480) do not show higher R-loops in cells depleted of the UAP56 RNA-binding and helicase factor (6,035) (Extended Data Fig. 3r)[14]. Only 727 genes accumulated R-loops in both conditions, with 701 genes also being R-loop-prone in control cells. This suggests that BRG1 contributes to R-loop homeostasis via a different mechanism from UAP56. Thus, BRG1 induces unique genome-wide accumulation of R-loops, with a large impact on genes at particular chromatin environments, including rDNA and subtelomeric regions.

BRG1 resolves R-loop-mediated T-R conflicts

R-loop homeostasis is regulated by three main mechanisms involving proteins that prevent the nascent RNA from hybridizing back with the DNA (RNA binding factors THOC1 and UAP56), proteins that resolve R-loops (RNH1 and SETX), or DNA repair factors (Fanconi Anemia (FA) pathway)[2]. To explore the mechanism by which BRG1 controls R-loop homeostasis, we did double depletion of BRG1 and a representative factor of each of these mechanisms (UAP56, THOC1, SETX, FANCD2) (Extended Data Fig. 4a) and assessed R-loop levels and DNA damage by IF with S9.6 and γH2AX antibodies, respectively. Single depletion of either factor increased R-loops and DNA breaks (Fig. 4a,b), as expected. Co-depletion of BRG1 with THOC1 or UAP56 resulted in a further increase of both R-loops and DNA breaks (Fig. 4a,b; Extended Data Fig. 4b,c). Consistent with previous data (Figure 2e), R-loops increased in THOC1/UAP56 and BRG1 co-depletions from G1 to S-phase when compared to single depletions, as determined by high throughput microscope imaging at different cell cycle stages (Extended Data Fig. 4d,e). In contrast, when BRG1 was co-depleted with FANCD2 or SETX, R-loops and DNA breaks did not change significantly with respect to the single depletions (Fig 4a,b; Extended Data Fig. 4b,c). Therefore, BRG1 seems to function in the same class of mechanisms as FANCD2 and SETX to prevent R-loop accumulation and associated DNA damage.

Given that R-loop increase in siBRG1 cells occurs preferentially during S-phase (Fig. 2e, Extended Data Fig. 2h,i), we wondered whether BRG1's role in R-loop protection was associated with a putative role in managing T-R conflicts, and whether

BRG1 functionally interacted with FANCD2, used as a marker of replication fork stalls. To evaluate T-R conflicts, PLA was performed using anti-PCNA antibody, as a marker of replication[42], and the elongating form of RNA polymerase II (RNAPII) phosphorylated at Ser2 (P-Ser2-RNAP)[43], in pre-extracted cells (Extended Data Fig. 4f). BRG1 depletion significantly increased PLA foci. This was rescued by overexpressing siRNA-resistant wild-type BRG,1 but not a catalytically-dead BRG1 (Extended Data Fig. 4g). We also observed that PLA foci increased synergistically when BRG1 was co-depleted together with UAP56 or THOC1, but not when it was co-depleted with FANCD2 or SETX (Fig. 4c, Extended Data Fig. 4h). Thus, these results suggest that BRG1 helps resolve R-loop-mediated T-R collisions together with the FA pathway.

BRG1 accumulates at R-loop-dependent fork blocks

Given that BRG1 and FANCD2 function in the same pathway to prevent R-loop accumulation and T-R conflicts (Fig. 4a-c), we assayed whether BRG1 colocalized with FANCD2 in an R-loop-dependent manner via PLA using anti-BRG1 and anti-FANCD2 antibodies. PLA-positive foci were enriched in siC cells (Fig. 5a; Extended Data Fig. 5a). Notably, the number of BRG1-FANCD2 PLA foci per cell increased significantly when R-loop-mediated T-R conflicts were induced by UAP56 depletion, and were significantly reduced after RNH1 overexpression (Fig. 5a, Extended Data Fig. 5b). Therefore, BRG1 and FANCD2 associate with each other in an R-loop-dependent manner.

To assay whether FANCD2-BRG1 association reflected replication fork blocks that could lead to double-stranded breaks (DSBs), we performed PLA between BRG1 and marks of either replication fork stalling [RPA phosphorylated at serine 4/8 (RPA-S4/8P)][44] or DNA breakage (γH2AX)[45]. Consistently, we detected PLA foci in siC control cells for both BRG1-RPA-S4/8P and BRG1-γH2AX interactions (Fig. 5b,c, Extended Data Fig. 5c,d). Importantly, PLA foci were significantly increased upon UAP56 depletion (Fig. 5b,c) and disappeared when siUAP56 was combined with siBRG1, confirming specificity of the signal (Extended Data Fig. 5e,f). This is consistent with genome-wide BRG1 accumulation at R-loop-prone regions (Fig. 3b,c). Further confirmation of this was obtained by PLA with anti-BRG1 and S9.6 antibodies (Fig. 5d; Extended Data Fig. 5g). BRG1-S9.6 PLA foci were seen in siC cells, but they were significantly enhanced upon UAP56 depletion (Fig. 5d, Extended Data Fig. 5h) and significantly reduced by RNH1 overexpression. Therefore, BRG1 accumulates at stalled replication forks at T-R conflict regions harboring DNA-RNA hybrids.

Genome-wide association of BRG1 with T-R conflict regions

To explore the relationship of BRG1 with R-loop-dependent replication fork blocks, we analyzed the correlation of BRG1 and R-loop-enriched sites with those of DNA replication, damage and replication fork stalling genome-wide. For this, we integrated our siBRG1 R-loop data with data of active replication determined by Okazaki fragment sequencing (OK-seq)[46], data of replication fork stalling detected by FANCD2 ChIP-seq[47], and data of DNA breaks detected by γH2AX ChIP-seq[48] in normal K562 cells. Correlation was observed between R-loops and BRG1, FANCD2 and γH2AX ChIP-seq data in normal K562 cells (Fig. 6a). Indeed, 5,182 out of 5,419 genes enriched for FANCD2 marks and 6,095 out of 7,203 of genes presenting γH2AX marks were BRG1-target genes (Fig. 6b,c), with 2,964 commonly enriched in both features (Extended Data Fig. 6a). 3,731 of the FANCD2-BRG1 common target genes and 4,771 of the BRG1-target genes enriched in γH2AX were also R-loop-prone genes.

Next, we analyzed those sites where R-loops encountered replication forks coming from only one direction, in order to maximize the probability of an S-phase-specific effect. Notably, BRG1 and γH2AX distribution around R-loop-peaks was asymmetric, with a sharp decrease upstream and a gradual and extended decrease downstream of the replication fork. In contrast, FANCD2 distribution was more symmetric (Fig. 6d, Extended Data Fig. 6b). Analysis of the ratio of codirectional vs head-on T-R collisions showed that fork blocks were preferentially associated with head-on collisions (Extended Data Fig. 6c). FANCD2 and BRG1 presented a sharp asymmetric accumulation at head-on collisions (Fig. 6e, Extended Data Fig. 6d), but a weaker accumulation at co-directional encounters (Fig. 6f, Extended Data Fig. 6e), whereas R-loop levels were similar for the head-on and co-directional collisions analyzed (Extended Data Fig. 6f). This excludes the possibility that the presence of BRG1 and FANCD2 at head-on collisions is an indirect consequence of higher R-loop levels. Indeed, BRG1 and FANCD2 enrichment at head-on versus co-directional collisions occurs non-specifically at regions with both high and low R-loop levels (Extended Data Fig. 6g,h), suggesting that their action is linked to resolving T-R collisions regardless of whether these are mediated by R- loops.

Thus, BRG1, together with the FA factors, has a major function in resolving T-R conflicts responsible for replication fork blockage and associated DNA breaks genome-wide.

BRG1 controls chromatin accessibility at T-R collisions

Several studies have confirmed that *BRG1* knockout (KO) results in a global loss of chromatin accessibility, but accessibility gain is also observed at certain regions[49,50]. To investigate the role of BRG1 in R-loop-associated chromatin structure, we integrated our DRIPc-seq data with genome-wide chromatin accessibility and nucleosome occupancy data from K562 cells and other similar cell types with BRG1 ChIP-seq data (Extended Data Fig. 7a-c). Analysis of chromatin accessibility at R-loop-gain sites upon BRG1 depletion in K562 using the ATAC-seq and DNase-seq data of ENCODE[41] and other publicly available FAIRE-seq[48] data in normal K562 cells shows that R-loop-gain sites have high chromatin accessibility (Fig. 7a). Analysis of MNase-seq data from ENCODE [41] showed that nucleosome occupancy drops considerably at R-loop-enriched regions in BRG1-depleted cells. Therefore, siBRG1 R-loop-gain sites present an accessible chromatin structure and low nucleosome occupancy in control cells.

To determine chromatin accessibility at those R-loop-gain sites in siBRG1 cells, we used the publicly available genome-wide ATAC-seq from HAP1[49] and BIN67[51] cells and MNase-seq data from CD36 cells[52]. BRG1 ChIP-seq data from these cell lines correlated strongly with those of K562 (Extended Data Fig. 7a-c), validating the comparative analysis between them. *BRG1*-KO HAP1 cells showed higher chromatin accessibility values than control wild type cells at these sites, whereas shBRG1-treated CD36 cells showed decreased nucleosome occupancy upon BRG1 depletion compared to shLuc control cells (Fig. 7b). Consistently, BRG1-deficient BIN67 cells presented higher chromatin accessibility values that were rescued by wild type BRG1, whereas partially inactive (T910M) and fully inactive (K785R) BRG1 had either poor or no effect (Fig. 7c).

Finally, meta-genomic analysis in control K562 cells showed an open chromatin structure with low nucleosome occupancy at T-R collision regions (Fig. 7d). BRG1-deficiency enhanced chromatin accessibility and caused lower nucleosome occupancy at these regions, as seen in HAP1 and CD36 cells, respectively (Fig. 7e). Again, chromatin accessibility was rescued by wild type BRG1 but not by the mutant variants, as shown in BRG1-deficient BIN67 cells (Fig. 7f).

These results suggest an active role of BRG1 in limiting chromatin accessibility at the regions that upon BRG1 depletion are enriched in R-loops and T-R collisions.

The SWI/SNF complex helps prevent R-loop-mediated DNA breaks

Next, we asked whether BRG1 protects against R-loops when part of any SWI/SNF subcomplex, whether BAF or PBAF. Thus, we analyzed R-loop-dependent DNA damage when depleting BRM and the PBRM1 and ARID1A members specific to each

subcomplex subtype. DNA damage detected by γH2AX foci was increased in cells depleted of the three subunits assayed. These increases were suppressed by RNH1 overexpression in all cases, with the exception of siBRM (Fig. 8a, Extended Data Fig. 7d). Consistently, R-loops were also significantly increased in siPBRM1 and siARID1A cells, as detected by IF by S9.6 staining intensity (Fig. 8b, Extended Data Fig. 7e).

We next analyzed genome-wide ChIP-seq data of the three SWI/SNF subunits in HAP1 cells[49], given their high BRG1 ChIP-seq data correlation (Extended Data Fig. 7a-c), together with our DRIPc-seq data. Notably, siBRG1 R-loop-gain genes identified in our study presented high abundancy of ARID1A, similar to BRG1, and some amounts of PBRM1 (Extended Data Fig. 7f,g). The same results were observed near R-loop-forming sites (Extended Data Fig. 7h). In contrast, BRM was absent from these genes and the from the vicinity of R-loop-forming sites (Extended Data Fig. 7f-h). ARID1A was also highly prevalent downstream of T-R conflicts, preferentially at head on collisions, similar to BRG1. PBRM1 was only minimally detected, and BRM was absent (Fig. 8c). Consistently, chromatin accessibility was also enhanced in *ARID1A*-KO HAP1 cells at siBRG1 R-loop-gain and T-R collision sites identified in our K562 control cells (Extended Data Fig. 7i,j). Therefore, cBAF-specific subunit ARID1A impacts chromatin and prevents R-loop-dependent DNA damage similarly to BRG1. PBRM1 also does, but to a lesser extent, suggesting that subunits from other complex subtypes could also protect against genome instability via similar mechanisms.

Finally, as discussed above, SWI/SNF genes are frequently mutated in cancer, with BRG1 being the most highly mutated chromatin remodeling ATPase (Extended Data Fig. 8a,b), and BRG1's ATPase domain being mutated in various cancers (Extended Data Fig. 8c). The ATPase domain K785R mutation used in our study was unable to rescue R-loop-dependent phenotypes (Extended Data Fig. 1h,4g). Interestingly, BRG1-deficient C-33 A cancer cells[53] show higher R-loops levels than BRG1-proficient HeLa cells, despite both being cervical carcinoma cells (Extended Data Fig. 8d). R-loops are significantly decreased by overexpressing wild type BRG1 in C-33 A cells, further supporting the link between BRG1 deficiency and R-loops (Extended Data Fig. 8d). Notably, BRG1 and ARID1A, the two subunits showing higher impact on R-loop-dependent genome instability, are among the most frequently altered SWI/SNF genes in cancer (Extended Data Fig. 8e,f).

Therefore, SWI/SNF protects cells against R-loops and R-loop-mediated DNA damage via the BRG1 ATPase, which could thus be a tumor suppressor.

DISCUSSION

Here we show that BRG1, the main ATPase from the SWI/SNF chromatin remodeling complexes, plays an active role in suppressing R-loop accumulation and genome instability derived from R-loop-dependent T-R collisions, a result extended to other members of SWI/SNF subunits, PBRM1 and ARID1A. We show that chromatin remodeling is a key player in T-R conflict resolution, through a process involving the FA pathway of repair. As BRG1 is the most frequently mutated chromatin remodeling ATPase activity in cancer, our study not only helps to understand T-R conflict resolution, but suggests a link between unresolved R-loop-mediated T-R conflicts and cancer propensity.

BRG1's contribution to transcription regulation has been extensively analyzed[31], but less studied for its role in DNA repair and replication[54–58], for which its mechanistic involvement is unclear. BRG1 depletion in HeLa cells significantly increases R-loop-dependent DNA damage and causes accumulation of R-loops (Fig. 1; Extended Data Fig. 1). These are accumulated genome-wide when *BRG1* is knocked-down in K562 cells, and most R-loop-gain sites map in coding genes (73.7%) (Fig. 3, Extended Data Fig. 3). This leads to high genome instability detected by micronuclei, anaphase bridges and ectopic nucleoli (Fig. 1). The observation that R-loops and T-R conflicts cause DNA damage in siBRG1 cells highlights the importance of remodeling ATPases in genome integrity, but does not support a direct role of SWI/SNF in DSB repair as previously suggested[57,58].

Despite evidence showing a role for specific nucleases in the occurrence of DNA breaks at R-loops[59], R-loop-induced DNA breaks derive mainly from replication fork progression stalling[1]. R-loops induced by BRG1 depletion cause replication fork slow-down and stalling, and factors that process replication fork stalls (e.g. FANCD2 and BLM) accumulate in an R-loop and transcription-dependent manner (Fig. 2). Notably, BRG1s role in R-loop protection is not related to RNA biogenesis itself, since co-depletion of BRG1 with THOC1 and UAP56 synergistically increases R-loop-dependent DNA damage. Also, R-loop-accumulating genes differ between siBRG1 and siUAP56 cells. In contrast, BRG1 depletion is epistatic with FANCD2 or SETX silencing (Fig. 4), which supports the model that the role of BRG1, and by extension SWI/SNF, in R-loop protection is linked to replication rather than transcription. Indeed, most BRG1 accumulation at chromatin and the higher R-loop increase in siBRG1 cells occur in S-phase (Fig. 2), consistent with previous reports indicating that R-loop homeostasis depends on the DDR[60]. In contrast, SETX functions in transcription termination[7,15,61], but also in DNA repair and T-R conflicts[62–64]. Our results support that SETX protects

against R-loops during replication. This together with the epistasis of BRG1 depletion with FANCD2 silencing suggests that BRG1 functions together with the FA pathway to resolve T-R conflicts.

BRG1 co-IP with members of the FA pathway such as FANCA, FANCS/BRCA1and FANCD2[55,56,65], which we confirmed by PLA, importantly, depends on R-loops (Fig. 5, 6). Consistently, FANCD2 and BRG1 largely colocalize genome-wide, particularly at R-loop-prone sites. Similarly, BRG1+S9.6 PLA is also positive and significantly enhanced in an R-loop-dependent manner upon depletion of the DNA-RNA helicase UAP56/DDX39B. Interestingly, BRG1 is an interacting player in the DNA-RNA hybrid interactome[66] and colocalizes with DNA damage or stalled replication fork markers such as γH2AX and RPA-S4/8P. These data support the view that BRG1 is enriched at R-loop-mediated replication fork stalling sites.

Head-on T-R conflicts are much more harmful than co-directional collisions[67-69]. It has been proposed that this is due to a higher accumulation of hybrids at head-on conflicts and a more efficient removal of co-directional collisions[68–70] . Genome-wide analysis shows that for head-on conflicts, the damage is accumulated at higher frequency, as it happens for BRG1 recruitment; but this does not relate to higher R-loop abundancy, as our analysis of a subset of preselected collision sites show that they accumulate at similar levels regardless of orientation (Fig. 6, Extended Data Fig. 6), consistent with the view that R-loops do not form preferentially at head-on collision sites[70,71]. Importantly, the increased co-localization of BRG1 and FANCD2 at head-on sites even in the absence of R-loops strengthens the view that head-on conflicts are harmful, regardless of R-loops.

Either BRG1 or BRM may perform the ATPase function in SWI/SNF, but their expression is generally anti-correlated and tissue-specific[72,73]. Importantly, BRM genome-wide binding data did not correlate with siBRG1 R-loop-gain sites, nor did its depletion cause R-loop-dependent DNA damage. However, depletion of ARID1A, and to a lesser extent PBRM1, did cause R-loop-mediated damage (Fig. 8), consistent with a role during DNA repair as has been described for ARID1A and PBRM1[74–77].

Even though *BRG1*-KO causes a global loss of chromatin accessibility[49,50], certain regions gain accessibility, as observed for T-R conflict sites when knocking-out *BRG1* (Fig. 7). Indeed, BRG1 functions in chromatin packaging through interaction with HP1a or the mSin3a/HDAC complex[78,79]. Alternatively, ectopic remodeling activities may also act in BRG1-deficient cells, increasing chromatin accessibility; however, additional R-loops may also contribute. Interestingly, ARID1A is also enriched at T-R collisions and its deletion makes chromatin more accessible (Fig. 8, Extended Data

Fig. 7). This BRG1-ARID1A correlation suggests that the cBAF subcomplex prevents R-loop-dependent genome instability, even though PBRM1 may also play a role.

We propose a model in which the joint action of SWI/SNF and FA factors at T-R conflict sites modulate chromatin structure to allow T-R conflict resolution and repair (Fig. 8d). An accessible chromatin state might facilitate binding of DNA repair and R-loop resolution factors such as FA factors, BRCA2, SETX or RNH1. This would be consistent with reports indicating that other chromatin remodelers (Sin3A, FACT) regulate R-loop homeostasis[19,20]. Such activities would help resolve T-R conflicts, especially those that are R-loop-mediated, preventing them from causing DNA damage, replication stress and genome instability, hallmarks of cancer cells[80], and consequently causing retention of R-loops. This impact of SWI/SNF alterations on T-R conflicts could help explain the prevalence of its mutations in human malignancies[27] and why SWI/SNF factors are more broadly mutated than any other tumor suppressor or oncogene, except for TP53[28].

ACKNOWLEDGEMENTS

We thank J.C. Reyes (CABIMER, Seville) for reagents provided. Research was funded by grants from the European Research Council (ERC2014 AdG669898 TARLOOP), the Spanish Ministries of Economy and Competitiveness (BFU2016-75058-P) and Science and Innovation (PDI2019-104270GB-I00/BMC), the European Union (FEDER) and Foundation “Vencer el Cancer”. A.B-F. was supported in part by a Juan de la Cierva postdoctoral contract from the Spanish Ministry of Science and Innovation.

AUTHOR CONTRIBUTIONS

A.B.-F. and A.A. designed the study and the experiments; A.B.-F. performed most of the experiments and all the bioinformatic analysis. S. B. and S. M. contributed with specific experiments. A.B.-F and A.A. wrote the manuscript. All authors read, discussed and agreed with the final version of this manuscript.

DECLARATION OF INTERESTS

All authors, AA, A.B-F., S.B. and S. M., declare no competing interests, whether non-financial or financial, in relation to the work described here.

REFERENCES

1. Gómez-González, B. & Aguilera, A. Transcription-mediated replication hindrance: A major driver of genome instability. *Genes and Development* (2019) doi:10.1101/gad.324517.119.
2. García-Muse, T. & Aguilera, A. R Loops: From Physiological to Pathological Roles. *Cell* (2019) doi:10.1016/j.cell.2019.08.055.
3. Crossley, M. P., Bocek, M. & Cimprich, K. A. R-Loops as Cellular Regulators and Genomic Threats. *Molecular Cell* (2019) doi:10.1016/j.molcel.2019.01.024.
4. Domínguez-Sánchez, M. S., Barroso, S., Gómez-González, B., Luna, R. & Aguilera, A. Genome instability and transcription elongation impairment in human cells depleted of THO/TREX. *PLoS Genet.* (2011) doi:10.1371/journal.pgen.1002386.
5. Huertas, P. & Aguilera, A. Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. *Mol. Cell* (2003) doi:10.1016/j.molcel.2003.08.010.
6. Li, X. & Manley, J. L. Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. *Cell* (2005) doi:10.1016/j.cell.2005.06.008.
7. Skourti-Stathaki, K., Proudfoot, N. J. & Gromak, N. Human Senataxin Resolves RNA/DNA Hybrids Formed at Transcriptional Pause Sites to Promote Xrn2-Dependent Termination. *Mol. Cell* (2011) doi:10.1016/j.molcel.2011.04.026.
8. Paulsen, R. D. *et al.* A Genome-wide siRNA Screen Reveals Diverse Cellular Processes and Pathways that Mediate Genome Stability. *Mol. Cell* (2009) doi:10.1016/j.molcel.2009.06.021.
9. Wahba, L., Amon, J. D., Koshland, D. & Vuica-Ross, M. RNase H and Multiple RNA Biogenesis Factors Cooperate to Prevent RNA:DNA Hybrids from Generating Genome Instability. *Mol. Cell* (2011) doi:10.1016/j.molcel.2011.10.017.
10. Tuduri, S. *et al.* Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription. *Nat. Cell Biol.* (2009) doi:10.1038/ncb1984.
11. Aguilera, A. & Gómez-González, B. DNA-RNA hybrids: The risks of DNA breakage during transcription. *Nature Structural and Molecular Biology* (2017) doi:10.1038/nsmb.3395.
12. Cerritelli, S. M. & Crouch, R. J. Ribonuclease H: The enzymes in eukaryotes. *FEBS Journal* vol. 276 1494–1505 (2009).
13. Sreerama Chaitanya Sridhara, A., Carvalho, S., Rita Grosso, A., Marcela

Gallego-Paez, L. & Carmo-Fonseca, M. Transcription Dynamics Prevent RNA-Mediated Genomic Instability through SRPK2-Dependent DDX23 Phosphorylation. *CellReports* **18**, 334–343 (2017).

14. Pérez-Calero, C. *et al.* UAP56/DDX39B is a major cotranscriptional RNA–DNA helicase that unwinds harmful R loops genome-wide. *Genes Dev.* (2020) doi:10.1101/gad.336024.119.

15. Hatchi, E., Skourti-Stathaki, K., Proudfoot, N. J. & Livingston Correspondence, D. M. BRCA1 Recruitment to Transcriptional Pause Sites Is Required for R-Loop-Driven DNA Damage Repair. *Mol. Cell* **57**, 636–647 (2015).

16. Bhatia, V. *et al.* BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. *Nature* **511**, 362–365 (2014).

17. García-Rubio, M. L. *et al.* The Fanconi Anemia Pathway Protects Genome Integrity from R-loops. *PLOS Genet.* **11**, e1005674 (2015).

18. Schwab, R. A. *et al.* The Fanconi Anemia Pathway Maintains Genome Stability by Coordinating Replication and Transcription. *Mol. Cell* (2015) doi:10.1016/j.molcel.2015.09.012.

19. Herrera-Moyano, E., Mergui, X., García-Rubio, M. L., Barroso, S. & Aguilera, A. The yeast and human FACT chromatinreorganizing complexes solve R-loopmediated transcription-replication conflicts. *Genes Dev.* (2014) doi:10.1101/gad.234070.113.

20. Salas-Armenteros, I. *et al.* Human THO–Sin3A interaction reveals new mechanisms to prevent R-loops that cause genome instability. *EMBO J.* **36**, (2017).

21. Poli, J. *et al.* Mec1, INO80, and the PAF1 complex cooperate to limit transcription replication conflicts through RNAPII removal during replication stress. *Genes Dev.* (2016) doi:10.1101/gad.273813.115.

22. Feldman, J. L. & Peterson, C. L. Yeast Sirtuin Family Members Maintain Transcription Homeostasis to Ensure Genome Stability. *Cell Rep.* (2019) doi:10.1016/j.celrep.2019.05.009.

23. Bayona-Feliu, A., Casas-Lamesa, A., Reina, O., Bernués, J. & Azorín, F. Linker histone H1 prevents R-loop accumulation and genome instability in heterochromatin. *Nat. Commun.* **8**, (2017).

24. Almeida, R. *et al.* Chromatin conformation regulates the coordination between DNA replication and transcription. *Nat. Commun.* (2018) doi:10.1038/s41467-018-03539-8.

25. Nguyen, D. T. *et al.* The chromatin remodelling factor ATRX suppresses R loops in transcribed telomeric repeats . *EMBO Rep.* (2017)

doi:10.15252/embr.201643078.

26. Chen, P. B., Chen, H. V., Acharya, D., Rando, O. J. & Fazzio, T. G. R loops regulate promoter-proximal chromatin architecture and cellular differentiation. *Nat. Struct. Mol. Biol.* (2015) doi:10.1038/nsmb.3122.

27. Hodges, C., Kirkland, J. G. & Crabtree, G. R. The many roles of BAF (mSWI/SNF) and PBAF complexes in cancer. *Cold Spring Harb. Perspect. Med.* (2016) doi:10.1101/cshperspect.a026930.

28. Kadoch, C. *et al.* Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nat. Genet.* (2013) doi:10.1038/ng.2628.

29. Neigeborn, L. & Carlson, M. Genes affecting the regulation of SUC2 gene expression by glucose repression in Saccharomyces cerevisiae. *Genetics* (1984).

30. Stern, M., Jensen, R. & Herskowitz, I. Five SWI genes are required for expression of the HO gene in yeast. *J. Mol. Biol.* (1984) doi:10.1016/0022-2836(84)90315-2.

31. Wilson, B. G. & Roberts, C. W. M. EPIGENETICS AND GENETICS SWI / SNF nucleosome remodellers and cancer. *Nat. Rev. Cancer* (2011) doi:10.1038/nrc3068.

32. Centore, R. C., Sandoval, G. J., Soares, L. M. M., Kadoch, C. & Chan, H. M. Mammalian SWI/SNF Chromatin Remodeling Complexes: Emerging Mechanisms and Therapeutic Strategies. *Trends in Genetics* (2020) doi:10.1016/j.tig.2020.07.011.

33. Phelan, M. L., Sif, S., Narlikar, G. J. & Kingston, R. E. Reconstitution of a core chromatin remodeling complex from SWI/SNF subunits. *Mol. Cell* (1999) doi:10.1016/S1097-2765(00)80315-9.

34. Euskirchen, G. M. *et al.* Diverse roles and interactions of the SWI/SNF chromatin remodeling complex revealed using global approaches. *PLoS Genet.* (2011) doi:10.1371/journal.pgen.1002008.

35. Cerami, E. *et al.* The cBio Cancer Genomics Portal: An open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* (2012) doi:10.1158/2159-8290.CD-12-0095.

36. Dykhuizen, E. C. *et al.* BAF complexes facilitate decatenation of DNA by topoisomerase IIα. *Nature* (2013) doi:10.1038/nature12146.

37. Schlacher, K., Wu, H. & Jasin, M. A Distinct Replication Fork Protection Pathway Connects Fanconi Anemia Tumor Suppressors to RAD51-BRCA1/2. *Cancer Cell* (2012) doi:10.1016/j.ccr.2012.05.015.

38. Pichierri, P., Franchitto, A. & Rosselli, F. BLM and the FANC proteins collaborate in a common pathway in response to stalled replication forks. *EMBO J.* (2004) doi:10.1038/sj.emboj.7600277.
39. Zhou, B. *et al.* Comprehensive, integrated, and phased whole-genome analysis of the primary ENCODE cell line K562. *Genome Res.* (2019) doi:10.1101/gr.234948.118.
40. Sanz, L. A. *et al.* Prevalent, Dynamic, and Conserved R-Loop Structures Associate with Specific Epigenomic Signatures in Mammals. *Mol. Cell* (2016) doi:10.1016/j.molcel.2016.05.032.
41. Dunham, I. *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* (2012) doi:10.1038/nature11247.
42. Mailand, N., Gibbs-Seymour, I. & Bekker-Jensen, S. Regulation of PCNA-protein interactions for genome stability. *Nature Reviews Molecular Cell Biology* (2013) doi:10.1038/nrm3562.
43. Zaborowska, J., Egloff, S. & Murphy, S. The pol II CTD: New twists in the tail. *Nature Structural and Molecular Biology* vol. 23 771–777 (2016).
44. Ashley, A. K. *et al.* DNA-PK phosphorylation of RPA32 Ser4/Ser8 regulates replication stress checkpoint activation, fork restart, homologous recombination and mitotic catastrophe. *DNA Repair (Amst).* (2014) doi:10.1016/j.dnarep.2014.04.008.
45. Ward, I. M. & Chen, J. Histone H2AX Is Phosphorylated in an ATR-dependent Manner in Response to Replicational Stress. *J. Biol. Chem.* (2001) doi:10.1074/jbc.C100569200.
46. Wu, X. *et al.* Developmental and cancer-associated plasticity of DNA replication preferentially targets GC-poor, lowly expressed and late-replicating regions. *Nucleic Acids Res.* (2018) doi:10.1093/nar/gky797.
47. Richardson, C. D. *et al.* CRISPR–Cas9 genome editing in human cells occurs via the Fanconi anemia pathway. *Nat. Genet.* (2018) doi:10.1038/s41588-018-0174-0.
48. Pang, B., De Jong, J., Qiao, X., Wessels, L. F. A. & Neefjes, J. Chemical profiling of the genome with anti-cancer drugs defines target specificities. *Nat. Chem. Biol.* (2015) doi:10.1038/nchembio.1811.
49. Schick, S. *et al.* Systematic characterization of BAF mutations provides insights into intracomplex synthetic lethalities in human cancers. *Nat. Genet.* (2019) doi:10.1038/s41588-019-0477-9.
50. Barisic, D., Stadler, M. B., Iurlaro, M. & Schübeler, D. Mammalian ISWI and SWI/SNF selectively mediate binding of distinct transcription factors. *Nature*

(2019) doi:10.1038/s41586-019-1115-5.

51. Pan, J. *et al.* The ATPase module of mammalian SWI/SNF family complexes mediates subcomplex identity and catalytic activity–independent genomic targeting. *Nature Genetics* (2019) doi:10.1038/s41588-019-0363-5.

52. Hu, G. *et al.* Regulation of nucleosome landscape and transcription factor targeting at tissue-specific enhancers by BRG1. *Genome Res.* (2011) doi:10.1101/gr.121145.111.

53. Dunaief, J. L. *et al.* The retinoblastoma protein and BRG1 form a complex and cooperate to induce cell cycle arrest. *Cell* (1994) doi:10.1016/0092-8674(94)90405-7.

54. Cohen, S. M. *et al.* BRG1 co-localizes with DNA replication factors and is required for efficient replication fork progression. *Nucleic Acids Res.* (2010) doi:10.1093/nar/gkq559.

55. Wang, H. *et al.* BRCA1/FANCD2/BRG1-Driven DNA Repair Stabilizes the Differentiation State of Human Mammary Epithelial Cells. *Mol. Cell* (2016) doi:10.1016/j.molcel.2016.05.038.

56. Bochar, D. A. *et al.* BRCA1 is associated with a human SWI/SNF-related complex: Linking chromatin remodeling to breast cancer. *Cell* (2000) doi:10.1016/S0092-8674(00)00030-1.

57. Park, J. H. *et al.* Mammalian SWI/SNF complexes facilitate DNA double-strand break repair by promoting γ-H2AX induction. *EMBO J.* (2006) doi:10.1038/sj.emboj.7601291.

58. Lee, H. S., Park, J. H., Kim, S. J., Kwon, S. J. & Kwon, J. A cooperative activation loop among SWI/SNF, γ-H2AX and H3 acetylation for DNA double-strand break repair. *EMBO J.* (2010) doi:10.1038/emboj.2010.27.

59. Sollier, J. *et al.* Transcription-Coupled Nucleotide Excision Repair Factors Promote R-Loop-Induced Genome Instability. *Mol. Cell* (2014) doi:10.1016/j.molcel.2014.10.020.

60. Barroso, S. *et al.* The DNA damage response acts as a safeguard against harmful DNA–RNA hybrids of different origins. *EMBO Rep.* (2019) doi:10.15252/embr.201847250.

61. Yanling Zhao, D. *et al.* SMN and symmetric arginine dimethylation of RNA polymerase II C-terminal domain control termination. *Nature* (2016) doi:10.1038/nature16469.

62. Alzu, A. *et al.* Senataxin associates with replication forks to protect fork integrity across RNA-polymerase-II-transcribed genes. *Cell* (2012) doi:10.1016/j.cell.2012.09.041.

63. Brambati, A. *et al.* Dormant origins and fork protection mechanisms rescue sister forks arrested by transcription. *Nucleic Acids Res.* (2018) doi:10.1093/nar/gkx945.
64. Mischo, H. E. *et al.* Cell-Cycle Modulation of Transcription Termination Factor Sen1. *Mol. Cell* (2018) doi:10.1016/j.molcel.2018.03.010.
65. Otsuki, T. *et al.* Fanconi anemia protein, FANCA, associates with BRG1, a component of the human SWI/SNF complex. *Hum. Mol. Genet.* **10**, 2651–2660 (2001).
66. Cristini, A., Groh, M., Kristiansen, M. S. & Gromak, N. RNA/DNA Hybrid Interactome Identifies DXH9 as a Molecular Player in Transcriptional Termination and R-Loop-Associated DNA Damage. *Cell Rep.* (2018) doi:10.1016/j.celrep.2018.04.025.
67. Prado, F. & Aguilera, A. Impairment of replication fork progression mediates RNA polII transcription-associated recombination. *EMBO J.* (2005) doi:10.1038/sj.emboj.7600602.
68. Hamperl, S., Bocek, M. J., Saldivar, J. C., Swigut, T. & Cimprich, K. A. Transcription-Replication Conflict Orientation Modulates R-Loop Levels and Activates Distinct DNA Damage Responses. *Cell* (2017) doi:10.1016/j.cell.2017.07.043.
69. Lang, K. S. *et al.* Replication-Transcription Conflicts Generate R-Loops that Orchestrate Bacterial Stress Survival and Pathogenesis. *Cell* (2017) doi:10.1016/j.cell.2017.07.044.
70. García-Rubio, M. *et al.* Yra1-bound RNA–DNA hybrids cause orientation-independent transcription– replication collisions and telomere instability. *Genes Dev.* (2018) doi:10.1101/gad.311274.117.
71. Achar, Y. J., Adhil, M., Choudhary, R., Gilbert, N. & Foiani, M. Negative supercoil at gene boundaries modulates gene topology. *Nature* (2020) doi:10.1038/s41586-020-1934-4.
72. Reyes, J. C. *et al.* Altered control of cellular proliferation in the absence of mammalian brahma (SNF2α). *EMBO J.* (1998) doi:10.1093/emboj/17.23.6979.
73. Muchardt, C., Bourachot, B., Reyes, J. C. & Yaniv, M. ras transformation is associated with decreased expression of the brm/SNF2α ATPase from the mammalian SWI-SNF complex. *EMBO J.* (1998) doi:10.1093/emboj/17.1.223.
74. Watanabe, R. *et al.* SWI/SNF factors required for cellular resistance to dna damage include arid1a and arid1b and show interdependent protein stability. *Cancer Res.* (2014) doi:10.1158/0008-5472.CAN-13-3608.
75. Niimi, A., Chambers, A. L., Downs, J. A. & Lehmann, A. R. A role for chromatin

remodellers in replication of damaged DNA. *Nucleic Acids Res.* (2012) doi:10.1093/nar/gks453.

76. Kakarougkas, A. *et al.* Requirement for PBAF in Transcriptional Repression and Repair at DNA Breaks in Actively Transcribed Regions of Chromatin. *Mol. Cell* (2014) doi:10.1016/j.molcel.2014.06.028.

77. Meisenberg, C. *et al.* Repression of Transcription at DNA Breaks Requires Cohesin throughout Interphase and Prevents Genome Instability. *Mol. Cell* (2019) doi:10.1016/j.molcel.2018.11.001.

78. Nielsen, A. L. *et al.* Selective interaction between the chromatin-remodeling factor BRG1 and the heterochromatin-associated protein HP1α. *EMBO J.* (2002) doi:10.1093/emboj/cdf560.

79. Sif, S., Saurin, A. J., Imbalzano, A. N. & Kingston, R. E. Purification and characterization of mSin3A-containing Brg1 and hBrm chromatin remodeling complexes. *Genes Dev.* (2001) doi:10.1101/gad.872801.

80. Halazonetis, T. D., Gorgoulis, V. G. & Bartek, J. An oncogene-induced DNA damage model for cancer development. *Science* (2008) doi:10.1126/science.1140735.

FIGURE LEGENDS

Figure 1 – Analysis of DNA damage and genome instability in siBRG1 cells.

(a) Percentage of cells containing >5 γH2AX foci in control (siC) and BRG1-depleted (siBRG1) HeLa cells with (+) and without (-) overexpression of RNH1. Data are plotted as mean + SEM (n=4). Scale bar, 2.5μm. (Paired Student's t-test, one-tailed).

(b) Alkaline comet assay tail moment quantification in cells treated as in (a). Data are plotted as mean of the medians + SEM (n=3). Scale bar, 40μm. (Unpaired Student's t-test, two-tailed)

(c) Quantification of nuclear S9.6 signal intensity in cells treated as in (a). Data presented as scatter plot (n=3). Median values are indicated by red lines and printed in red. Scale bar, 10 μm. (Mann-Whitney U test, two-tailed).

(d) DRIP-qPCR using S9.6 antibody in siC and siBRG1 HeLa cells. Signal values normalized with respect to the siC control and plotted as mean ± SEM (n=3). (Paired Student's t-test, one-tailed).

(e) Percentage of cells presenting micronuclei in cells treated as in (a). Data are plotted as mean + SEM (n=3). Scale bar, 5μm. (Unpaired Student's t-test, one-tailed).

(f) Percentage of DNA bridge occurrence in cells treated as in (a). Data are plotted as mean + SEM (n=3). Scale bar, 5μm. (Unpaired Student's t-test, one-tailed).

(g) Percentage of cells containing >10 nucleolin ectopic foci in cells treated as in (a). Data are plotted as mean + SEM (n=3). nucl.:nucleolin. Scale bar, 5μm (left) and 1μm (right). (Paired Student's t-test, one-tailed).

Representative images, with nuclear perimeter highlighted (yellow dashed line) are shown. A.U.: Arbitrary Units. P-values are indicated. See also Extended Data Fig. 1.

Figure 2 – Effect of BRG1 depletion on DNA replication dynamics.

(a) Quantification of nuclear EdU signal intensity (left) and percentage of cells incorporating EdU (EdU^+) (right) in control (siC) and BRG1-depleted (siBRG1) with (+) and without (-) overexpression of RNH1. Signal intensity data are plotted as box plot (n=3). Center line indicate median value. Boxes and whiskers indicate 25^{th} to 75^{th} and 10^{th} to 90^{th} percentiles, respectively. Points below and above the whiskers are drawn as individual points. (Mann-Whitney U test, two-tailed). For cell percentages, data are shown as mean + SEM (n=3). (Paired Student's t-test, one-tailed). Scale bar, 10 μm.

(b) Quantification of replication fork velocity and asymmetry in cells treated as in (a). Data are shown as box plots (n=2). Box plot details as in (a). Scale bar, 10μm.

(c) Percentage of cells containing >10 FANCD2 foci in cells treated as in (a). Data are plotted as mean + SEM (n=3). (Paired Student's t-test, one-tailed). Scale bar, 2.5μm.

(d) Percentage of cells showing BLM foci in cells treated as in (a). Data are plotted as mean + SEM (n=2). (Paired Student's t-test, one-tailed). Scale bar, 5μm.

(e) Quantification of nuclear S9.6 signal intensity through cell cycle in siC and siBRG1 HeLa cells. Data presented as scatter plot (n=3). Median values are indicated by red lines and printed in red. Scale bar, 5μm. (Mann-Whitney U test, two-tailed).

Other details as in Figure 1. P-values are indicated. See also Extended Data Fig. 2.

Figure 3. Genome-wide analysis of R-loop accumulation upon BRG1 depletion.

(a) Representative screenshot of an R-loop-prone chromosome region in control K562 cells. DRIPc-seq (green) (n=2) and untreated (dark blue) and RNH-treated (light blue) DRIP-seq data are shown.

(b) Representative genome-wide screenshot showing colocalization of R-loop, BRG1 and expression in control K562 cells. DRIPc-seq (green) (n=2), BRG1 ChIP-seq (yellow) (n=2) and RNA-seq (purple) (n=2) data are shown.

(c) Venn diagram showing genome-wide co-occurrence between R-loop-prone genes (DRIPc-seq), BRG1 target genes (ChIP-seq) and expressed genes (RNA-seq) in control K562 cells.

(d) Representative screenshot showing R-loop accumulation upon BRG1 depletion in K562 cells (n=2). Data are shown according to DNA strand (W:Watson or C:Crick).

(e) Examples of R-loop-accumulating genes from W (left) and C (right) strands in siBRG1-transfected K562 cells. DRIPc-seq data in siC (blue) and siBRG1-transfected (vermillion) K562 cells (n=2) is presented.

(f) Metagene analysis. DRIPc-seq mean coverages in siC (blue) and siBRG1-transfected (vermillion) K562 cells along the gene body (+/- 2kb) are shown. Average coverages from 2 biological replicates are shown. TSS, transcription start site; TTS, transcription termination site.

Genome localization and scale bars are indicated at top left and right corners, coverage scale at top left in each track and gene tracks (black) below for the corresponding genome windows. Replicates are also indicated (R1; R2). See also Extended Data Fig. 3.

Figure 4 – Epistatic analysis between BRG1 and known R-loop-preventing factors.

(a) Quantification of nuclear S9.6 signal intensity in HeLa cells double-transfected with the indicated siRNAs and siC (-) or siBRG1 (+). Data are plotted as scatter plot (n=3; except for siC, siTHOC1, siTHOC1+siBRG1, siUAP56 and siUAP56+siBRG1 where n=4). Median values are indicated by red lines and printed in red. Scale bar, 5μm. (Mann-Whitney U test, two-tailed).

(b) Percentage of cells containing >5 γH2AX foci in cells treated as in (a). Data are plotted as mean + SEM (n=3; except for siC and siFANCD2 where n=4). Scale bar, 10μm. (Paired Student's t-test, one-tailed).

(c) Quantification of RNAPIIS2P+PCNA PLA in cells treated as in (a). Foci/cell are plotted as box plot (n=3; except for siC (n=6) and siSETX (n=4)). Box plot details as in Fig. 2a. Scale bar, 10μm. (Mann-Whitney U test, two-tailed).

Other details as in Figure 1. P-values are indicated. See also Extended Data Fig. 4.

Figure 5 – Evaluation of BRG1 occurrence at replication fork stalling sites.

(a) Quantification of BRG1+FANCD2 PLA in control (siC) and UAP56-depleted (siUAP56) HeLa cells with (+) and without (-) overexpression of RNH1.

(b) Quantification of BRG1+γH2AX PLA in siC and siUAP56-transfected HeLa cells.

(c) Quantification of BRG1+ RPA S4/8P PLA in cells treated as in (b).

(d) Quantification of BRG1+S9.6 PLA in cells treated as in (a).

Foci number/cell are plotted as box plot (n=3). Box plot details as in Fig. 2a. Scale bar, 5μm. (Mann-Whitney U test, two-tailed). P-values are indicated. See also Extended Data Fig. 5.

Figure 6 – BRG1 genome-wide co-localization analysis with R-loop, Replication fork (RF) stalling and DNA damage sites.

(a) Representative screenshot of a genome region showing accumulation of FANCD2 (red) and γH2AX (purple) at sites where R-loops (green) and BRG1 (yellow) are present. Arrows indicate RF directionality. RFD: Replication Fork Directionality.

(b) Venn diagram showing genome-wide co-occurrence between R-loop-prone genes (DRIPc-seq; green), BRG1 (yellow) and FANCD2 (red) target genes.

(c) Venn diagram showing genome-wide co-occurrence between R-loop-prone genes (DRIPc-seq; green), BRG1 (yellow) and γH2AX (purple) target genes.

(d) DRIPc-seq, BRG1, FANCD2 and γH2AX ChIP-seq mean coverage around transcription-replication (T-R) conflicts (+/-1Mb). Data plotted as heatmap according to RF direction. Red arrow indicates the site where R-loop accumulate.

(e) BRG1 and FANCD2 ChIP-seq mean signal intensity around head-on (HO) T-R conflicts (+/-1Mb). Data plotted as heatmap according to RF direction. Red arrow indicates the site where R-loop accumulate.

(f) As in (e), but around co-directional (CD) T-R conflicts (+/-1Mb).

Average coverages from two DRIPc-seq and BRG1 ChIP-seq biological replicates are shown. Color scales indicate protein abundancy. Trx:Trancription. Other details as in Figure 3. See also Extended Data Fig. 6.

Figure 7 – Chromatin accessibility and nucleosome occupancy analysis at R-loop-gain sites in BRG1-deficient cells.

(a) ATAC-seq, DNase-seq, FAIRE-seq and MNase-seq mean coverages in control K562 cells around R-loop-gain peaks (+/- 10kb). Data presented as metaplot.

(b) Average coverages from ATAC-seq in wild-type (WT) or *BRG1*-KO HAP1 cells and MNase-seq in shluc or shBRG1-treated CD36 cells at R-loop-gain peaks. Data plotted as box plot. (Mann-Whitney U test, two-tailed).

(c) ATAC-seq average coverages in BIN67 (BRG1-deficient) cells untreated (-), expressing WT BRG1 and partially (T910M) or totally inactive (K785R) versions of BRG1 at R-loop-gain peaks. Data plotted as box plot. (Mann-Whitney U test, two-tailed).

(d) ATAC-seq, DNase-seq, FAIRE-seq and MNase-seq mean coverages in control K562 cells at transcription-replication (T-R) collisions (+/- 1Mb). Data plotted as heatmap.

(e) Mean coverages from ATAC-seq in WT or *BRG1*-KO HAP1 cells and MNase-seq in shluc or shBRG1-treated CD36 cells at T-R collisions. Data plotted as heatmap.

(f) ATAC-seq mean coverages in BIN67 (BRG1-deficient) cells untreated (-), expressing wild-type (WT) BRG1 and partially (T910M) or totally inactive (K785R) versions of BRG1 at T-R collisions (+/- 1Mb). Data plotted as heatmap.

Color scales indicate signal intensity. Box plot details as in Fig. 2a. P-values are indicated. RF, replication fork; HO, head-on; CD, co-directional, Trx, transcription. Other details as in Figure 6. See also Extended Data Fig. 7.

Figure 8 – R-loop and DNA damage study in cells depleted of different SWI/SNF subunits.

(a) Percentage of cells containing >5 γH2AX foci in HeLa cells transfected with siRNAs against indicated SWI/SNF subunits with (+) and without (-) overexpression of RNH1. Data are plotted as mean + SEM (n=3). (Paired Student's t-test, one-tailed).

(b) Quantification of nuclear S9.6 signal intensity in HeLa cells treated as in (a). Data are plotted as scatter plot (n=3). Median values are indicated by red lines and printed in red. (Mann-Whitney U test, two-tailed).

(c) BRG1, BRM, ARID1A and PBRM1 ChIP-seq mean coverage around head-on (HO; left) and co-directional (CD; right) transcription-replication (T-R) collisions (+/- 1Mb). Data presented as heatmap according to replication fork (RF) direction. Red arrow indicates the site where R-loop accumulate.

(d) Working model. Naturally occurring or unscheduled R-loop formation represents an obstacle to RF advance that needs to be properly addressed. RF blockage is sensed by FA pathway and SWI/SNF activity is required at these sites. SWI/SNF activity might be necessary to modulate chromatin structure and facilitate DNA damage response. In wild-type conditions, this mechanism would ensure maintenance of genome integrity. However, this pathway is strongly affected in SWI/SNF-deficient cells, where genome instability may arise and eventually result in a transformation phenotype.

Color scales indicate protein abundance. Scale bars, 5μm. P-values are indicated. Trx, transcription. Other details as in Figure 6. See also Extended Data Fig. 7.

METHODS

Cell lines

Human female HeLa and K562 cell lines were retrieved from American Type Culture Collection (ATCC). HeLa (ATCC CCL-2) cells were cultured in Dulbecco's modified Eagle's medium (DMEM; GIBCO) supplemented with 10% heat-inactivated fetal bovine serum (Sigma Aldrich, Merck KGaA) and 1% antibiotic-antimycotic (BioWEST) at 37ºC (5% CO_2). K562 (ATCC, CCL-243) cells were cultured in Iscove's Modified Dulbecco's medium (IMDM; GIBCO) supplemented with 10% heat-inactivated fetal bovine serum (Sigma Aldrich) and 1% antibiotic-antimycotic (BioWEST) at 37ºC (5% CO_2). C-33 A cells were kindly provided by Dr. José Carlos Reyes and cultured in Dulbecco's modified Eagle's medium (DMEM; GIBCO) supplemented with 10% heat-inactivated fetal bovine serum (Sigma Aldrich, Merck KGaA) and 1% antibiotic-antimycotic (BioWEST) at 37ºC (5% CO_2).

Protein knock-down

Cells were transfected with 50nM siRNA against desired target using DharmaFECT 1 (Dharmacon), according to the manufacturer's instructions. When co-depletion was intended, 25nM siRNA against each target were used. ON-TARGET SMARTpool siRNAs from Dharmacon against BRG1(L-010431-00), UAP56 (L-003805-00), SETX (L-021420-00), FANCD2 (L-016376-00), THOC1 (L-016376-00), BRM (L-017253-00), PBRM1(L-008692-01), ARID1A (L-017263-00) were used to achieve protein depletion. ON-TARGETplus Non-targeting Control Pool was used as control (siC). ON-TARGETplus Human SMARCA4 siRNA (Set of 4) (LQ-010431-00-0005) were used to test siRNA depletion of BRG1 individually. For siRNA-resistant expression of BRG1, 5'-CAGUGUCACUGGAUGUCAA-3' siRNA was used. A detailed list of siRNAs used in this study is supplied as Supplementary Table 2.

Plasmid transfection

Protein overexpression was performed transfecting cells with expression plasmids at 1µg/mL final concentration using Lipofectamine 2000 (Invitrogen), according to the manufacturer's instructions. pcDNA3-RNaseH1, containing the full-length RNH1 cloned into pcDNA3[81] and pEGFP-M27-H1, containing the GFP-fused RNH11 lacking the first 26 amino acids responsible for its mitochondrial localization cloned into pEGFP for GFP-RNH11 overexpression[82] were used to overexpress RNH11. pcDNA3 (Invitrogen) and the pEGFP (Clontech) empty vectors were used as controls.

For IF, western blot, EdU incorporation and Proximity Ligation assays, cells from each condition were transfected either with pEGFP or pEGFP-M27-H1 after 48h of siRNA treatment. For single-cell electrophoresis and DNA combing, transfection was performed either with pcDNA3 or pcDNA3-RNaseH1 48h after siRNA transfection. Once transfected with the plasmid, cells were cultured 24h further in all cases to allow protein expression and assayed. GFP signal was used to determine transfected cells in IF experiments.

For BRG1 rescue experiments, pSV2-hSNF2[83] was used to overexpress wild-type BRG1 and pTS-CeBRG1 K-R[84] was used to overexpress catalytically dead (K785R mutation) BRG1. pcDNA3 was used as empty plasmid control.

mRNA quantification

RNA purification was performed using RNeasy Mini kit (Qiagen) according to manufacturer's conditions. Then, cDNA synthesis was achieved using QuantiTect Rev. Transcription kit (Qiagen), according to manufacturer's guidelines. Finally, quantitative PCR were performed on a 7500 FAST Real-Time PCR system (ThermoFisher Scientific) and mRNA expression values calculated using the ΔΔCt method and HPRT housekeeping gene as control. A detailed list with primers used in this study is provided in Supplementary Table 3.

Western Blot

Chromatin fraction was obtained as described[85] and subjected to Western blot following standard procedures. Membranes were incubated with rabbit anti-BRG1 (1:500), anti-THOC1 (1:1000), anti-UAP56 (1:1000), anti-FANCD2 (1:500) or anti-SETX (1:500). Anti-GAPDH (1:4000) or anti-vinculin (1:5000) antibodies were used as loading controls. Pounceau S was used to determine loading amount.

Immunofluorescence

S9.6 immunofluorescence was performed essentially as described[86]. Briefly, cells were fixed with 100% ice-cold methanol, blocked with PBS-BSA 2% overnight at 4ºC and incubated with S9.6 (1:1000) and anti-nucleolin (1:2000) antibodies overnight at 4ºC. Then, coverslips were washed three times in PBS1X, and then incubated with secondary antibodies (1:1000) for 1 hour at room temperature. Finally, cells were washed again, stained with DAPI and mounted in ProLong Gold AntiFade reagent (Invitrogen).

DNA damage assessment by γH2AX immunostaining was performed mainly as previously described[14] with minor modifications. Briefly, cells were pre-extracted and

fixed with Triton X-100 0.1% + PBS 1X + formaldehyde (methanol-free) 4% for 10 minutes at RT, washed with PBS, permeabilized with PBS + 0.5% Triton X-100 for 5 minutes at RT and blocked with TBS 1X+ BSA 3% + Tween-20 0.1% for 30 minutes at RT. Then, cells were incubated overnight at 4ºC with anti-γH2AX (1:1000) (Merck Cat#05-636 in Fig.1; Abcam Cat#ab2893 in others) in blocking solution, washed, and incubated again with the corresponding secondary antibodies (1:1000) for 1 hour at RT. Finally, coverslips were washed again, stained with DAPI and mounted in ProLong Gold AntiFade reagent (Invitrogen). Same methodology was also used when performing IFs against BRG1, FANCD2 and BLM. In this case, mouse anti-BRG1 (1:100), anti-FANCD2 (1:100) and anti-BLM (1:250) antibodies and PBS 1X+ BSA 3% + Tween-20 0.1% as blocking solution were used.

IF images were acquired with a Leica DM6000 microscope equipped with a DFC390 camera (Leica) at 63X magnification and LAS AX image acquisition software (Leica). FIJI (ImageJ) image processing package[87] was used for IF analysis. Nuclear mean grey value for S9.6, after subtraction of nucleolar signal, was measured for each condition. When indicated, S9.6 nuclear foci were also quantified. In the case of γH2AX, FANCD2 and BLM, foci per cell were quantified. IF signal profile along specified sketches was obtained using RGB profiler plug-in for FIJI[87].

Single-cell electrophoresis

Alkalyne single-cell electrophoresis or comet assay was performed using a commercial kit (Trevigen) following manufacturer's protocol. Comet slides were stained with SYBRGreen, and images were captured with a Leica DM6000 microscope equipped with a DFC390 camera (Leica) at 10X magnification. Analysis of comet assay images was performed using OpenComet plug-in[88] for FIJI[87]. Tail moment was measured for each cell in each condition.

DNA:RNA immunoprecipitation (DRIP) assays

DNA–RNA hybrids were immunoprecipitated using the S9.6 antibody on purified genomic DNA enzymatically digested with HindIII, EcoRI, XbaI, SspI and BsrGI restriction enzymes. As control, samples were in vitro treated with RNH1 (New England Biolabs) as described[89]. Finally, eluted DNA was subjected to quantitative PCR (or sequencing) and immunoprecipitation rate expressed as Input %. When considered, relative values respect siC without RNH were also calculated and plotted. A detailed list with primers used in this study is provided in Supplementary Table 3.

Library construction and sequencing

DRIPc-seq was performed by sequencing RNA moieties from K562 DRIP samples purified as described[89]. RNA was fragmented and size checked on a 2100 Bioanalyzer (Agilent) and libraries constructed using TruSeq Stranded Total RNA Library Prep (Illumina) as determined by the company.

DRIP-seq was performed by sequencing DNA from K562 DRIP samples as described[14]. DNA elutes were sonicated and size checked on a 2100 Bioanalyzer (Agilent) and libraries constructed using the ThruPLEX DNA-Seq 6S kit (Rubicon Genomics) according to manufacturer's guidelines.

Stranded total RNA-seq was performed from K562 purified RNA as reported[14]. RNA was purified, fragmented and size checked on 2100 Bioanalyzer (Agilent). Libraries were built using TruSeq Stranded Total RNA Library Prep (Illumina) after ribosomal RNA depletion according to producer protocol as in previous reports[14].

In all cases, samples were 2x75bp pair-end (PE) sequenced on NextSeq500 platform (Illumina).

Genome-wide data downstream analysis

Paired-end sequencing reads were subjected to quality control pipeline using the FASTQ Toolkit v.1.0.0 software (Illumina) and uploaded to the Galaxy web platform for data analysis[90]. We used the public servers at usegalaxy.org and usegalaxy.eu., as well as a locally installed version. First, reads were mapped to the canonical version of human reference genome hg38 using Bowtie2[91] except for DRIP-seq and RNA-seq where BWA[92] and HISAT2[93] were used, respectively. Then, PCR duplicates were removed and reads assigned to Watson and Crick strand when stranded libraries were built (DRIPc-seq and RNA-seq) using SAMTools[94]. For DRIPc-seq analysis without Alus, reads mapping to Alu sequences were filtered out using BAM filter option from NGSUtils[95].

DRIPc-seq peak calling was performed using MACS2 package[96] without input file with a FDR<0.01 and allowing broad region detection with a 0.1 cutoff. Regions covered by peaks in both replicates were retained and merged when closer than 5kb using BEDtools[97]. For comparative purposes, siC and siBRG1 resulting peaks were merged using BEDtools[97] and counts per peak calculated using FeatureCounts[98] for each condition and replicate. Then, counts were RPKM normalized and values subjected to differential analysis with limma-voom[99]. Differentially R-loop accumulating peaks upon BRG1 depletion were determined as those with a |FC|>1.5 and p-value<0.01. Those with a FC>1.5 and p-value<0.01 were designated as R-loop-gain peaks. Inside this category, peaks with an average RPKM value from both replicates

<=0 in siC were established as *de novo* appearing peaks upon *BRG1* knock-down, whereas those presenting an average RPKM>0 in siC were designated as increasing peaks.

K562 BRG1 ChIP-seq, ATAC-seq, DNase-seq and MNase-seq data were obtained from ENCODE database. K562 FANCD2 ChIP-seq, γH2AX ChIP-seq, FAIRE-seq and OK-seq, HAP1 ChIP-seq and ATAC-seq, CD36 MNase-seq and BIN67 ATAC-seq reads were gathered from publicly available data. Reads were aligned to the human reference genome hg38 using Bowtie2[91] and PCR duplicates removed using SAMTools[94]. For OK-seq, reads were also assigned to Watson and Crick strand using SAMTools[94]. BRG1 ChIP-seq peaks for both replicates were retrieved from ENCODE project[41]. Only those peaks appearing in both replicates were further considered for this study. FANCD2 and γH2AX ChIP-seq peak calling was performed using MACS2 with a p-value<0.01 and allowing broad region detection with a 0.001 cutoff.

Genome and gene annotation of peaks were performed with ChIPseeker[100] and genes retrieved from Ensembl release 94 2018[101]. Gene annotation was performed only on protein coding genes considering as promoter the region ranging from TSS to 2kb upstream and termination from TTS to 2kb downstream. With regard to RNA-seq, counts were assigned to every gene as mentioned and RPKM normalized for each replicate. After entire population analysis, those ones with an RPKM average from both replicates >0.001 were considered as expressed genes.

DeepTools2[102] was used to calculate average coverages, generate RPKM-normalized coverage profiles and metaplot images. RF directionality (RFD) track values were calculated as described previously for OK-seq data[103]. RFD values were measured for each R-loop peak to stablish R-loop (or gene when considered) orientation with respect to the RF. For our purposes, assays involving RFD were performed only those sites (R-loops or genes) were |RFD|>0.75 ensuring a high chance of collision in a specific orientation.

EdU incorporation and detection

Click-iT™ EdU Cell Proliferation Kit for Imaging (ThermoFisher Scientific) was used to assay DNA replication through EdU incorporation. First, cells were cultured in complete medium supplemented with 10μM EdU for 30 minutes. Then, samples were fixed, permeabilized and Click-iT reaction performed according to manufacturer's guidelines. Finally, nuclei were stained with DAPI and mounted in ProLong Gold AntiFade reagent (Invitrogen). Images were acquired with a Leica DM6000 microscope equipped with a DFC390 camera (Leica) at 63X magnification and LAS AX image acquisition software (Leica). FIJI image processing package[87] was used for image analysis and

quantification. EdU entire population nuclear intensity and % of cells incorporating EdU were determined. EdU intensity was also determined only for those cells that incorporated EdU.

DNA combing

DNA combing was performed as previously described[14,20]. Thymidine analogues (iododeoxyuridine (IdU) and chlorodeoxyuridine (CldU)) were added in two consecutive 15 min pulses. DNA molecules were counterstained with an anti-ssDNA antibody (1:500) and an anti-mouse IgG coupled to Alexa 647 (1:50). CldU and IdU were detected using BU1/75 (1:20) and BD44 (1:20) anti-BrdU antibodies, respectively. Goat anti-mouse IgG Alexa 546 (1:50) and chicken anti-rat Alexa 488 (1:50) were used as secondary antibodies. DNA fiber images were acquired using a Leica DM6000 microscope equipped with an automated plate, a DFC390 camera and LAS AX software (Leica). Large-field images were ensembled using LAS AX software (Leica) and processed as described[104]. RF velocity (kb/min) was determined by measuring second analogue track length on individual ongoing RFs and multiplying this value by 2 and dividing by analogue incubation time to convert it into kb/min. Replication asymmetry was calculated by dividing the shortest second analogue track by the longest in divergent forks.

Flow Cytometry

For cell cycle analysis after RNH1 overexpression cells were treated and analyzed as described in[14]. Briefly, cells were harvested, fixed with PBS 1X+4% formaldehyde for 10 min at room temperature (RT) and permeabilized with PBS 1X+0,2% Triton X-100 during 10 min at RT. Finally, DNA was stained with DAPI (1 µg/ml) 4ºC overnight in PBS. Cells were analyzed in a BD influx sorter. Plasmid-transfected cells were identified with GFP signal and cell cycle was evaluated with DAPI signal.

For S9.6 analysis, samples were processed as described in[60] with minor modifications. Prior to harvesting, cells were treated with EdU (10 µM) for 30 min, then cells were fixed with methanol 100% at -20ºC for 7 minutes, washed with PBS 1X and treated with RNAse III (40 U/ml) for 30 min at 37°C. Samples were washed with PBS 1X and blocked for 1 hour at RT with PBS 1X+3% BSA+0,05% Tween 20. After that, Click-it reaction (ThermoFisher Scientific) was performed following manufacturer instructions. Finally, cells were stained with S9.6 antibody in suspension as previously describe[60]. DNA was stained with 1 µg/ml of DAPI at 4ºC overnight in PBS 1X. Cells were analyzed in a BD influx sorter and data were analyzed in FlowJo 9.3.2 (Tree Star).

Proximity Ligation Assay

Proximity Ligation Assay was performed using Duolink PLA Technology (Merck) as in previous reports[20]. First, samples were pre-extracted, fixed, permeabilized and incubated with primary antibodies as described for immunofluorescence assays. Then, secondary antibody binding, ligation and amplification reactions were performed according to manufacturer guidelines. Duolink in situ PLA probe anti-rabbit PLUS, Duolink in situ PLA probe anti-mouse MINUS and Duolink-Detection Reagents Red (Merck) were used to perform PLA reaction. Finally, nuclei were stained with DAPI and mounted in ProLong Gold AntiFade reagent (Invitrogen). For PLA reactions requiring rabbit BRG1, PCNA or RNAPII S2P antibodies, 1:500 dilution of these antibodies was used. Finally, images were acquired with a Leica DM6000 microscope equipped with a DFC390 camera (Leica) at 63X magnification and LAS AX image acquisition software (Leica). FIJI image processing package (Schindelin, J et al, Nat Medicine 2012) was used for image analysis and quantification. PLA foci number per cell were quantified for all conditions.

High-throughput immunofluorescence analysis

Cell cycle distribution of S9.6 reactivity was achieved using wide-field images acquired using automated plate on microscope Leica DM6000 as recently reported[14]. Cells were immunostained using S9.6 antibody as described and large images containing high amounts of cells (>1000) to ensure enough cell quantity in all phases of cell cycle were obtained with the mentioned microscope equipped with a DFC390 camera (Leica) at 63X magnification and LAS AX image acquisition software (Leica). Then, cells were associated to the corresponding cell cycle phase according to its DAPI content using DNA cell cycle plug-in (MBF collection) on FIJI[87]. In vivo validations of this plug-in with EdU staining and H3S10P immunofluorescence were already reported[14]. Finally, S9.6 nuclear intensity was determined for each cell in every condition as for immunofluorescence.

The same procedure was used to determine BRG1 signal distribution through cell cycle (>800 cells).

Image processing

Images were acquired with a Leica DM6000 microscope equipped with an automated plate, a DFC390 camera (Leica) and LAS AX image acquisition software (Leica). FIJI image processing package[87] was used for image analysis and quantification. DNA cell cycle (MBF collection for ImageJ; https://imagej.nih.gov/ij/plugins/mbf/index.html), RGB profiler (ImageJ Plug-in site; https://imagej.nih.gov/ij/plugins/index.html) and

OpenComet[88] plug-ins were added to the package to perform cell cycle-dependent IF and comet assay analysis. Nuclear mean intensity and foci quantification were performed using FIJI homemade generated macros, available upon request. Mean nuclear S9.6 signal was quantified after subtracting nucleolar signal, considered as that signal colocalizing with nucleolin. When indicated, S9.6 nuclear foci or nucleolar mean S9.6 reactivity were also determined. In the case of γH2AX, FANCD2 and BLM, foci per cell were always quantified. Regarding genome instability phenotypes, % of cells presenting micronuclei, bridges/cell expressed as % or nucleolin foci number were calculated. Mean EdU nuclear intensity was also measured when cells were treated with EdU. PLA foci/cell were measured when PLA were performed. In all cases, nuclear mean GFP signal was obtained and used to determine transfected cells. In representative images, DAPI (4′,6-Diamidino-2-phenylindole dihydrochloride) nuclear staining is shown in blue, GFP in green and IF/EdU/PLA in red, except in BLM IF and γH2AX IF in Fig. 4 and Extended Data Fig. 4 were IF is shown in green.

Statistics & Reproducibility

Statistical parameters including the number of biological replicates (n), standard error of the mean (SEM) and statistical significance are reported in the figure legends. All results presented in this manuscript were obtained from a minimum of 3 independent biological replicates, except for genome-wide analysis and BLM IF. BLM IF, DRIPc-seq, DRIP-seq, siC RNA-seq and BRG1 ChIP-seq data resulted from 2 independent biological replicates. siBRG1 RNA-seq data comes from one biological replicate. γH2AX and FANCD2 ChIP-seq and OK-seq data were retrieved from external database repositories and one replicate was used. Control PLAs were realized only once.

For data represented as histograms, Student's t-test was used. When data was presented as box plot or scatter plot, Mann–Whitney U-test was performed. Test details indicated in figure legends. For box plots, boxes and whiskers indicate 25–75 and 10–90 percentiles, respectively, and median values are indicated. In scatter plots and DNA combing box plots, median values are indicated and printed in red. Hypergeometric test and Pearson correlation were calculated for 2D Venn diagram and correlation analysis, respectively. In Venn diagrams, numbers represent genes co-occurring between conditions. In screenshots from genome-wide experiments, scales are adjusted so that background signal is low for better visualization of the results.

For IF experiments, >100 cells per replicate were measured (S9.6; BRG1), while at least 50 cells were considered when selecting subpopulations expressing GFP (γH2AX; FANCD2; BLM). A minimum of 50 cells was also measured in each comet

assay replicate. In EdU experiments, >100 cells per replicate were analyzed. For DNA combing analysis, 100 tracks per replicate were measured to determine fork velocity, while >45 asymmetries per replicate were analyzed. In RNAPIIS2P+PCNA PLA box plots >240 total cells were measured in double knockdown experiments and >340 cells in BRG1 rescue experiment. In FANCD2+BRG1 and S9.6+BRG1 PLA box plots >350 total cells are represented, while in RPA S4/8P+BRG1 and γH2AX+BRG1 PLA box plots >200 total cells are represented from at least 35 cells per replicate. In high-throughput IF analysis >1200 (S9.6) and >800 (BRG1) total cells were measured. >500 cells were analyzed when it was applied to compare single and double knockdowns. When comparing S9.6 IF in HeLa and C-33 A cancer cells, box plots were built using 50 cells per replicate.

Graphs were generated with Prism (GraphPad Software, Inc.), genome-wide screenshots obtained from Integrative Genome Viewer (IGV)[105] and coverage profile graphs were generated using Galaxy platform[106].

P-values are indicated in figures, and statistical tests applied described in figure legends. Analyzed samples were randomly chosen and data acquisition automatically performed by analysis software to ensure unbiased results.

Data and Code Availability

siBRG1 DRIPc-seq and RNA-seq, the first siC DRIPc-seq replicate and one DRIP-seq replicate datasets have been deposited at Gene Expression Omnibus (GEO) repository and are available under accession code GSE154631. Original data for another DRIP-seq replicate, siC RNA-seq and the second siC DRIPc-seq replicate datasets are available at the same database under accession code GSE127979[14], even though all experiments were performed in parallel. Other publicly available genome-wide data used in this study are listed in Supplementary Table 1.

Cancer-related information for SWI/SNF genes was retrieved from cBio Cancer Genomics Portal (www.cbioportal.org)[35,107] and Integrative Onco Genomics (www.intogen.org)[108,109] databases.

Software and algorithms source and links are listed in Supplementary Table 1.

Resources details

For more information on reagents and resources used, see Supplementary Table 1.

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Andrés Aguilera (aguilo@us.es).

METHODS REFERENCES

81. Ten Asbroek, A. L. M. A., Van Groenigen, M., Nooij, M. & Baas, F. The involvement of human ribonucleases H1 and H2 in the variation of response of cells to antisense phosphorothioate oligonucleotides. *Eur. J. Biochem.* (2002) doi:10.1046/j.0014-2956.2001.02686.x.
82. Cerritelli, S. M. *et al.* Failure to produce mitochondrial DNA results in embryonic lethality in Rnaseh1 null mice. *Mol. Cell* (2003) doi:10.1016/S1097-2765(03)00088-1.
83. Chiba, H., Muramatsu, M., Nomoto, A. & Kato, H. Two human homologues of saccharomyces cerevisiae SWI2/SNF2 and Drosophila brahma are transcriptional coactivators cooperating with the estrogen receptor and the retinoic acid receptor. *Nucleic Acids Res.* (1994) doi:10.1093/nar/22.10.1815.
84. de la Serna, I. L. *et al.* Mammalian SWI-SNF Complexes Contribute to Activation of the hsp70 Gene. *Mol. Cell. Biol.* (2000) doi:10.1128/mcb.20.8.2839-2851.2000.
85. Méndez, J. & Stillman, B. Chromatin Association of Human Origin Recognition Complex, Cdc6, and Minichromosome Maintenance Proteins during the Cell Cycle: Assembly of Prereplication Complexes in Late Mitosis. *Mol. Cell. Biol.* (2000) doi:10.1128/mcb.20.22.8602-8612.2000.
86. García-Rubio, M., Barroso, S. I. & Aguilera, A. Detection of DNA-RNA hybrids in vivo. in *Methods in Molecular Biology* (2018). doi:10.1007/978-1-4939-7306-4_24.
87. Schindelin, J. *et al.* Fiji: An open-source platform for biological-image analysis. *Nature Methods* (2012) doi:10.1038/nmeth.2019.
88. Gyori, B. M., Venkatachalam, G., Thiagarajan, P. S., Hsu, D. & Clement, M. V. OpenComet: An automated tool for comet assay image analysis. *Redox Biol.* (2014) doi:10.1016/j.redox.2013.12.020.
89. Sanz, L. A. & Chédin, F. High-resolution, strand-specific R-loop mapping via S9.6-based DNA–RNA immunoprecipitation and high-throughput sequencing. *Nat. Protoc.* (2019) doi:10.1038/s41596-019-0159-1.
90. Jalili, V. *et al.* The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2020 update. *Nucleic Acids Res.* (2020) doi:10.1093/nar/gkaa434.
91. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* (2012) doi:10.1038/nmeth.1923.
92. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-

Wheeler transform. *Bioinformatics* (2009) doi:10.1093/bioinformatics/btp324.

93. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* (2019) doi:10.1038/s41587-019-0201-4.

94. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* (2009) doi:10.1093/bioinformatics/btp352.

95. Breese, M. R. & Liu, Y. NGSUtils: A software suite for analyzing and manipulating next-generation sequencing datasets. *Bioinformatics* (2013) doi:10.1093/bioinformatics/bts731.

96. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* (2008) doi:10.1186/gb-2008-9-9-r137.

97. Quinlan, A. R. & Hall, I. M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* (2010) doi:10.1093/bioinformatics/btq033.

98. Liao, Y., Smyth, G. K. & Shi, W. FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* (2014) doi:10.1093/bioinformatics/btt656.

99. Ritchie, M. E. *et al.* Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* (2015) doi:10.1093/nar/gkv007.

100. Yu, G., Wang, L. G. & He, Q. Y. ChIP seeker: An R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* (2015) doi:10.1093/bioinformatics/btv145.

101. Zerbino, D. R. *et al.* Ensembl 2018. *Nucleic Acids Res.* (2018) doi:10.1093/nar/gkx1098.

102. Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* (2016) doi:10.1093/nar/gkw257.

103. Petryk, N. *et al.* Replication landscape of the human genome. *Nat. Commun.* (2016) doi:10.1038/ncomms10208.

104. Bianco, J. N. *et al.* Analysis of DNA replication profiles in budding yeast and mammalian cells using DNA combing. *Methods* (2012) doi:10.1016/j.ymeth.2012.04.007.

105. Robinson, J. T. *et al.* Integrative genomics viewer. *Nature Biotechnology* (2011) doi:10.1038/nbt.1754.

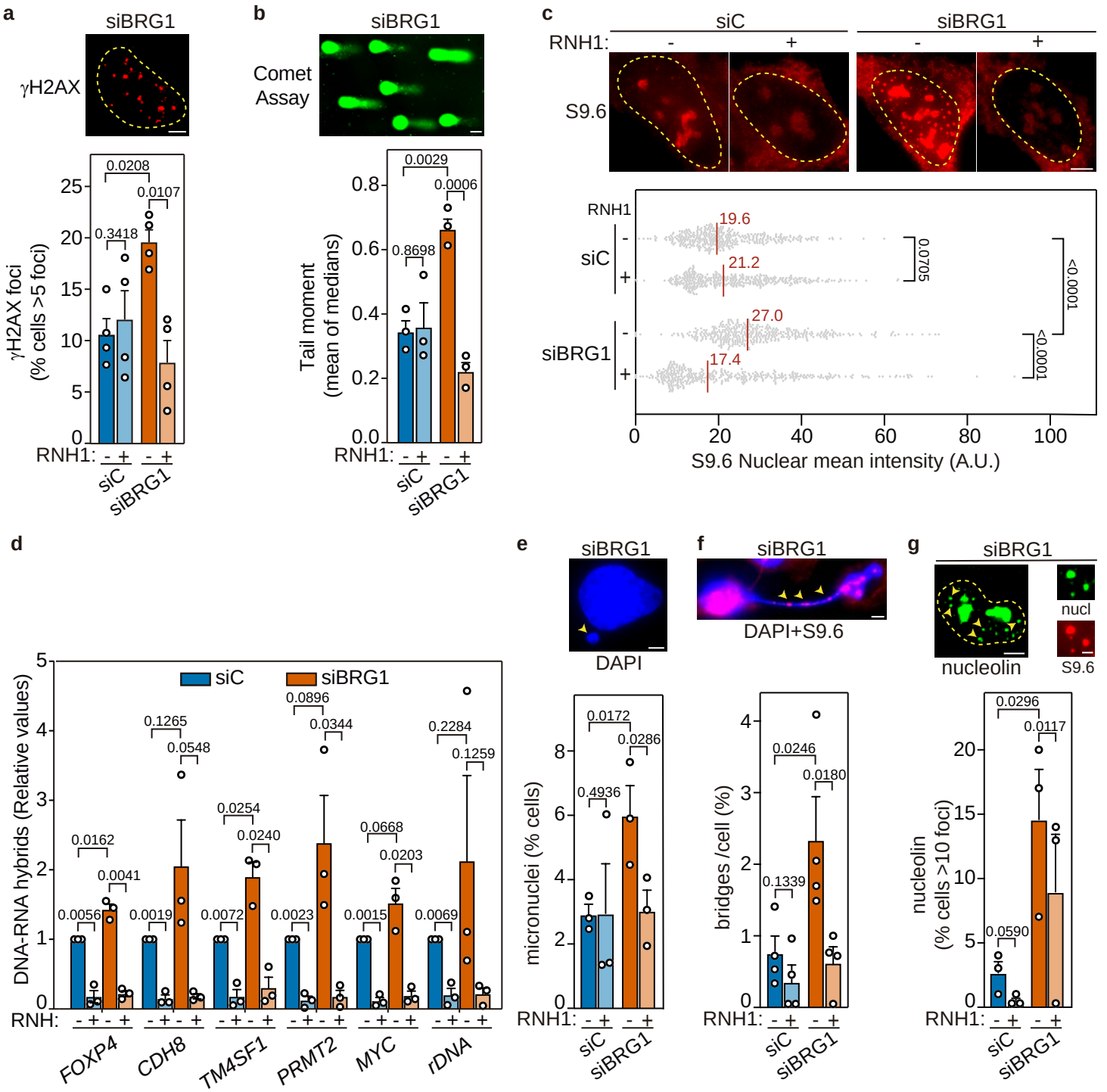
106. Afgan, E. *et al.* The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res.* (2016) doi:10.1093/nar/gkw343.

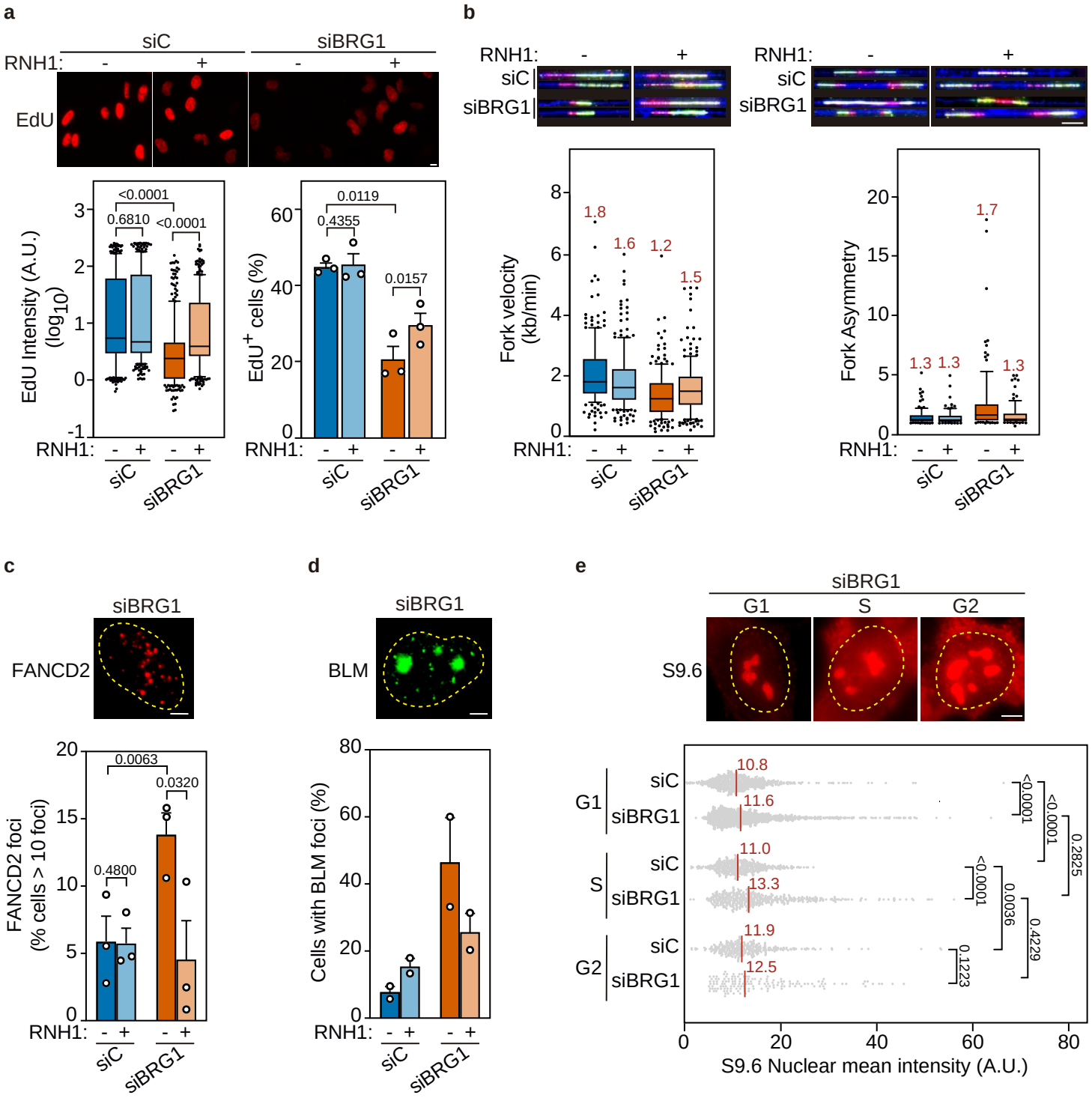
107. Gao, J. *et al.* Integrative analysis of complex cancer genomics and clinical

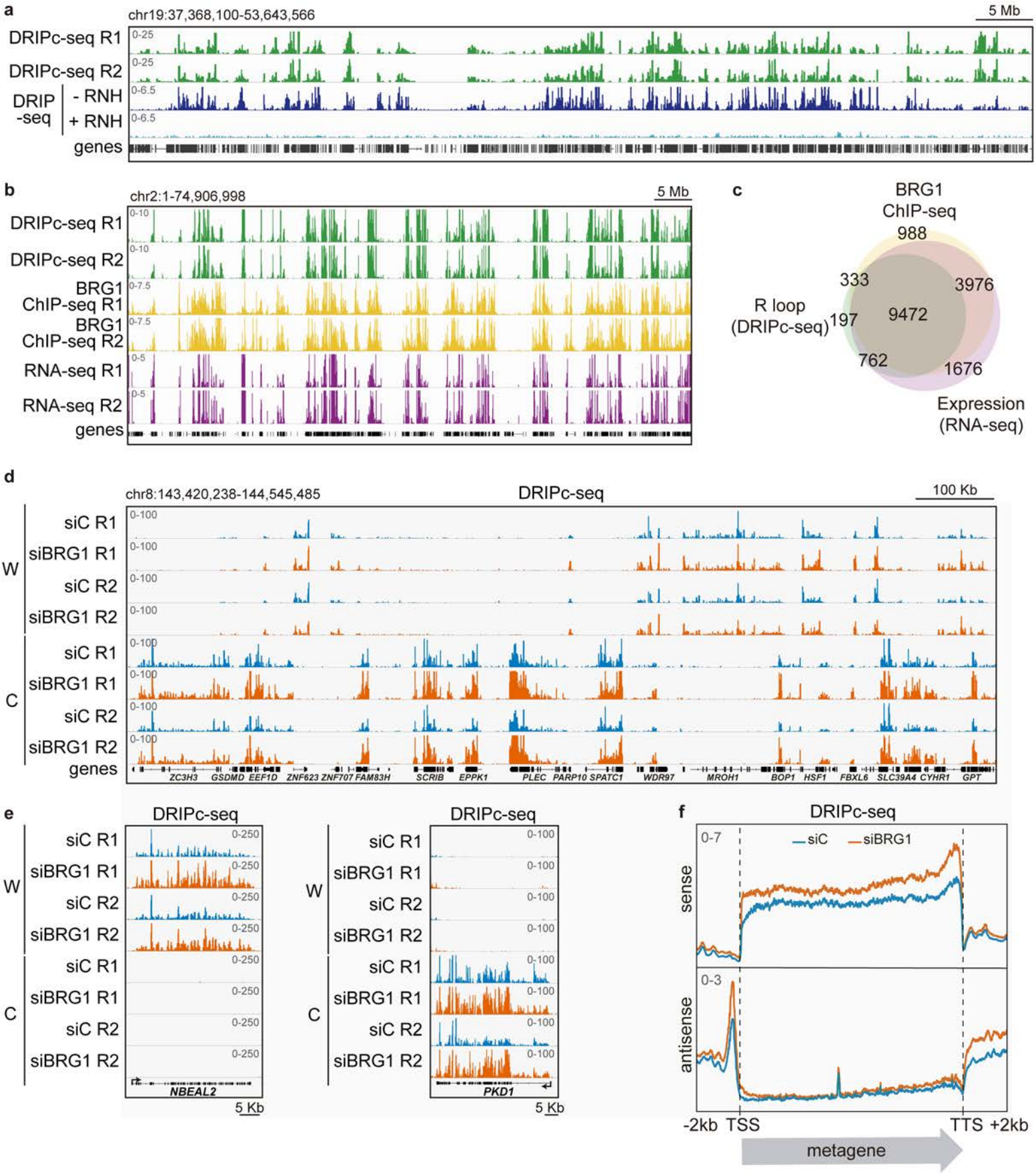
profiles using the cBioPortal. *Sci. Signal.* (2013) doi:10.1126/scisignal.2004088.

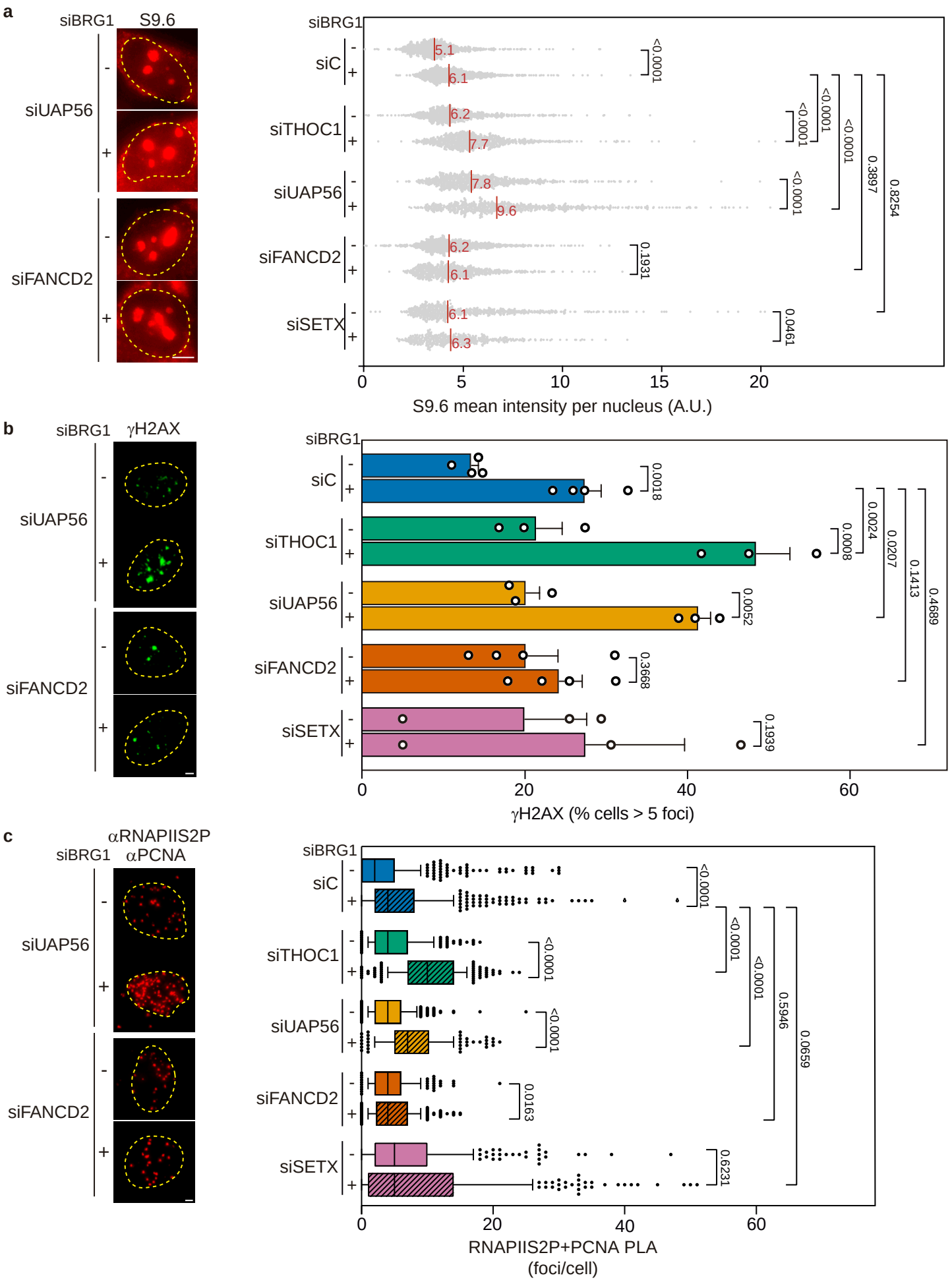
108. Gonzalez-Perez, A. *et al.* IntOGen-mutations identifies cancer drivers across tumor types. *Nat. Methods* (2013) doi:10.1038/nmeth.2642.

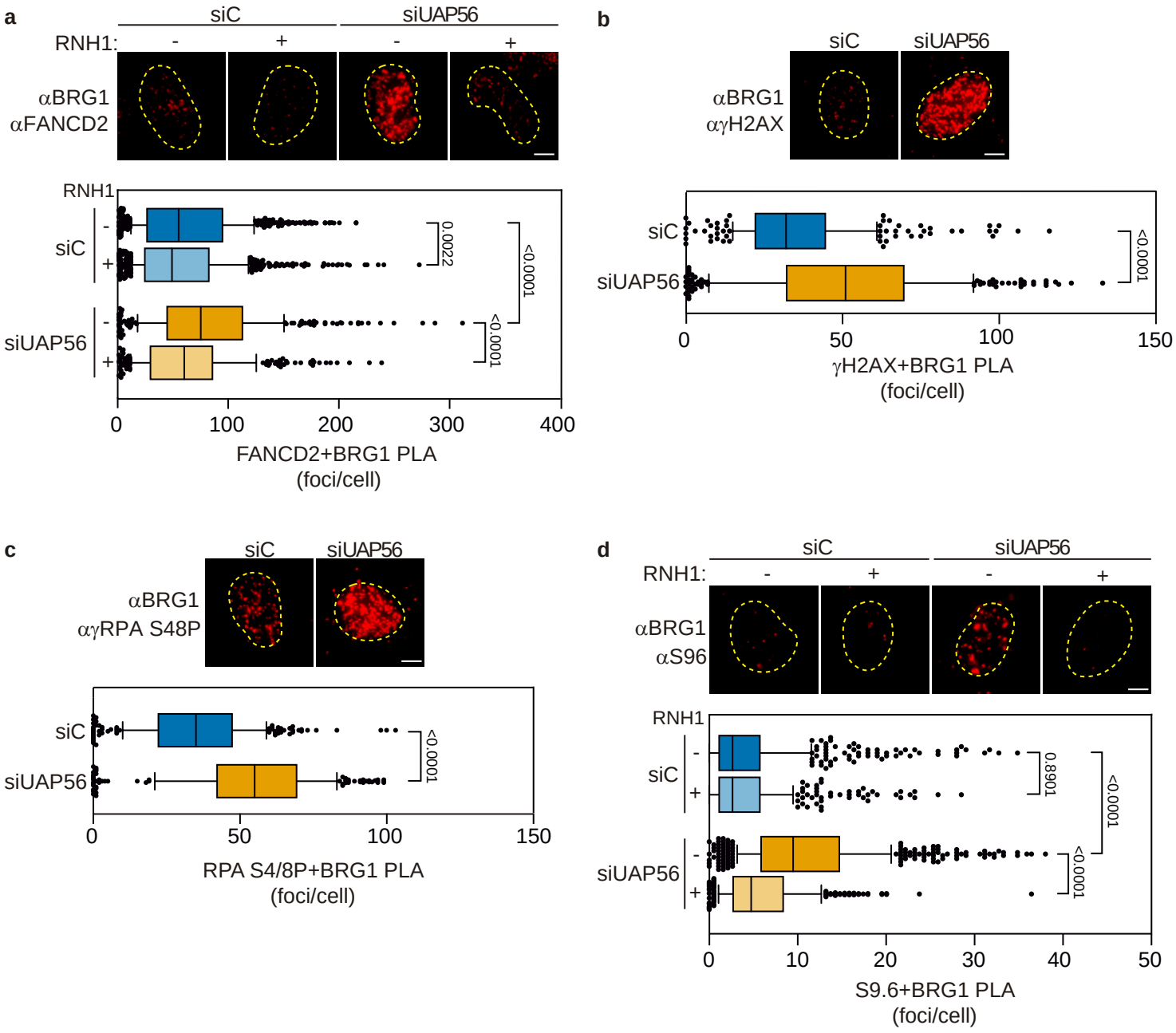
109. Martínez-Jiménez, F. *et al.* A compendium of mutational cancer driver genes. *Nature Reviews Cancer* (2020) doi:10.1038/s41568-020-0290-x.

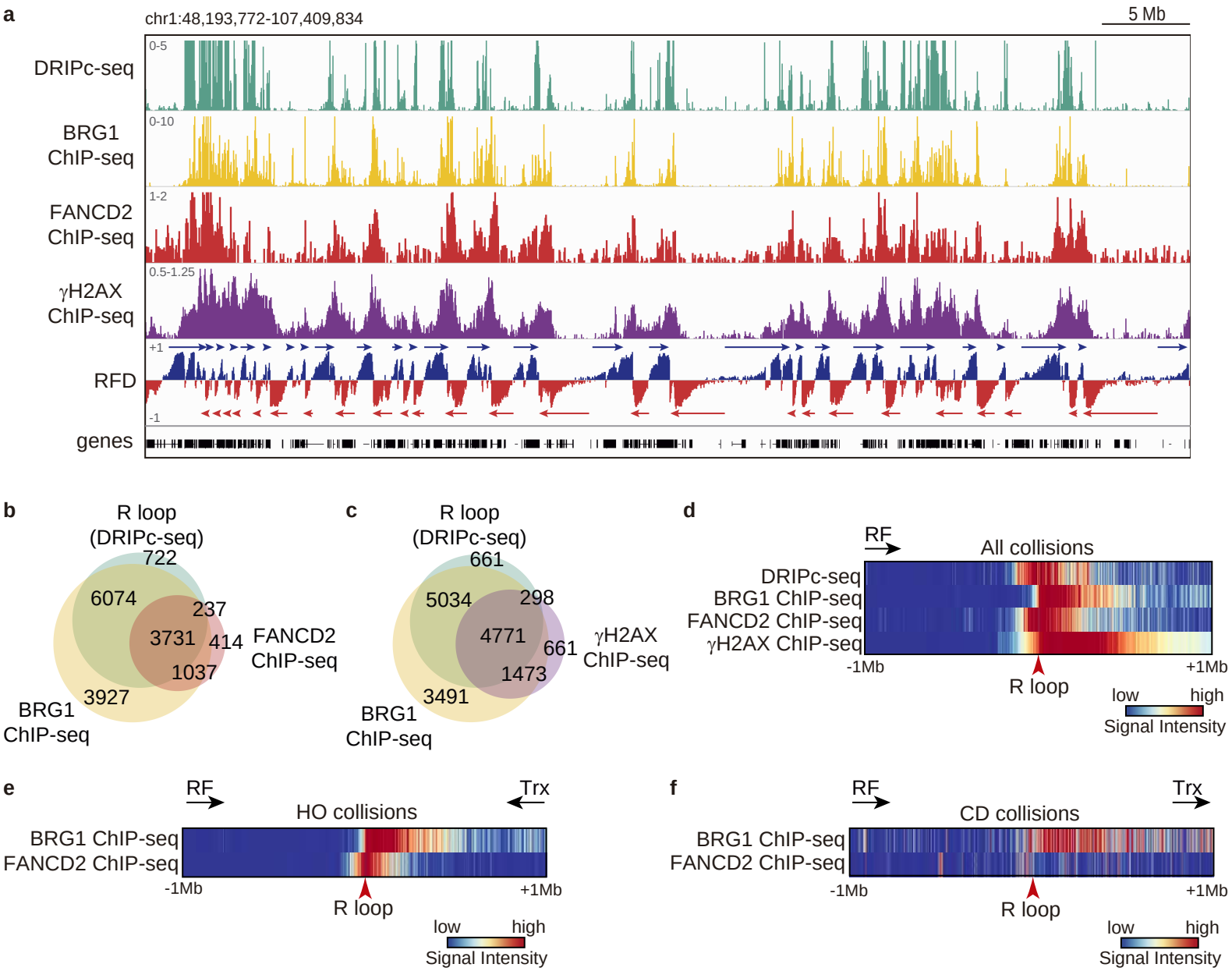

a
siBRG1
γH2AX
γH2AX foci (% cells >5 foci)
0.0208
0.0107
0.3418
RNH1: - + - +
siC
siBRG1
b
siBRG1
Comet Assay
Tail moment (mean of medians)
0.0029
0.0006
0.8698
RNH1: - + - +
siC
siBRG1
c
siC
siBRG1
RNH1: - + - +
S9.6
RNH1
siC
siBRG1
19.6
21.2
27.0
17.4
0.0705
<0.0001
<0.0001
S9.6 Nuclear mean intensity (A.U.)
d
siC
siBRG1
DNA-RNA hybrids (Relative values)
0.0162
0.0041
0.0056
0.1265
0.0548
0.0019
0.0254
0.0240
0.0072
0.0896
0.0344
0.0023
0.0668
0.0203
0.0015
0.2284
0.1259
0.0069
RNH: - + - + - + - + - + - + - + - + - + - + - + - +
FOXP4
CDH8
TM4SF1
PRMT2
MYC
rDNA
e
siBRG1
DAPI
micronuclei (% cells)
0.0172
0.0286
0.4936
RNH1: - + - +
siC
siBRG1
f
siBRG1
DAPI+S9.6
bridges /cell (%)
0.0246
0.0180
0.1339
RNH1: - + - +
siC
siBRG1
g
siBRG1
nucleolin
nucl
S9.6
nucleolin (% cells >10 foci)
0.0296
0.0117
0.0590
RNH1: - + - +
siC
siBRG1

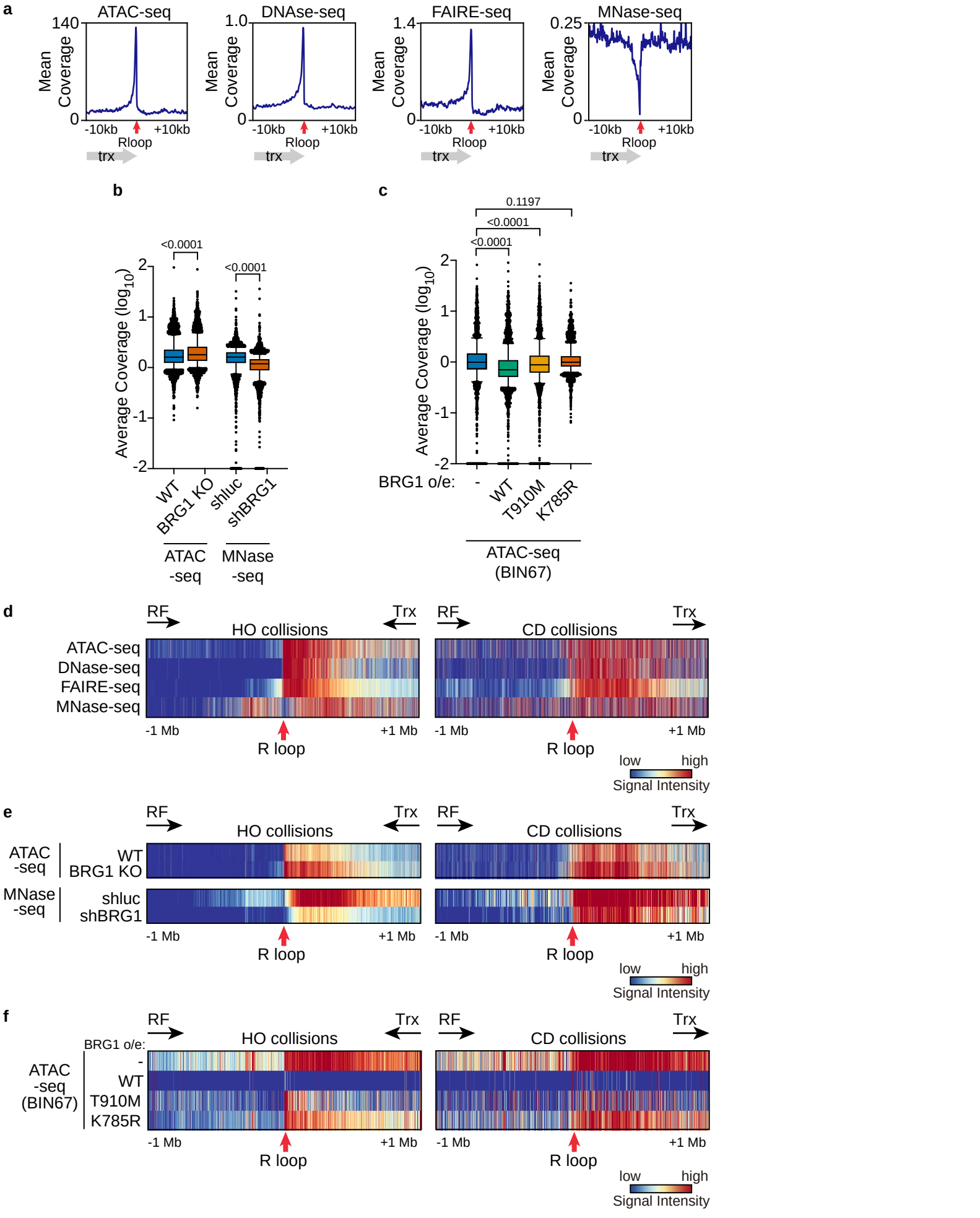

a
siC
siBRG1
RNH1:
EdU
EdU Intensity (A.U.) (log10)
<0.0001
0.6810
<0.0001
EdU+ cells (%)
0.0119
0.4355
0.0157
b
Fork velocity (kb/min)
1.8
1.6
1.2
1.5
Fork Asymmetry
1.7
1.3
1.3
1.3
c
FANCD2
FANCD2 foci (% cells > 10 foci)
0.0063
0.0320
0.4800
d
BLM
Cells with BLM foci (%)
e
G1
S
G2
S9.6
10.8
11.6
11.0
13.3
11.9
12.5
<0.0001
<0.0001
0.2825
<0.0001
0.0036
0.4229
0.1223
S9.6 Nuclear mean intensity (A.U.)

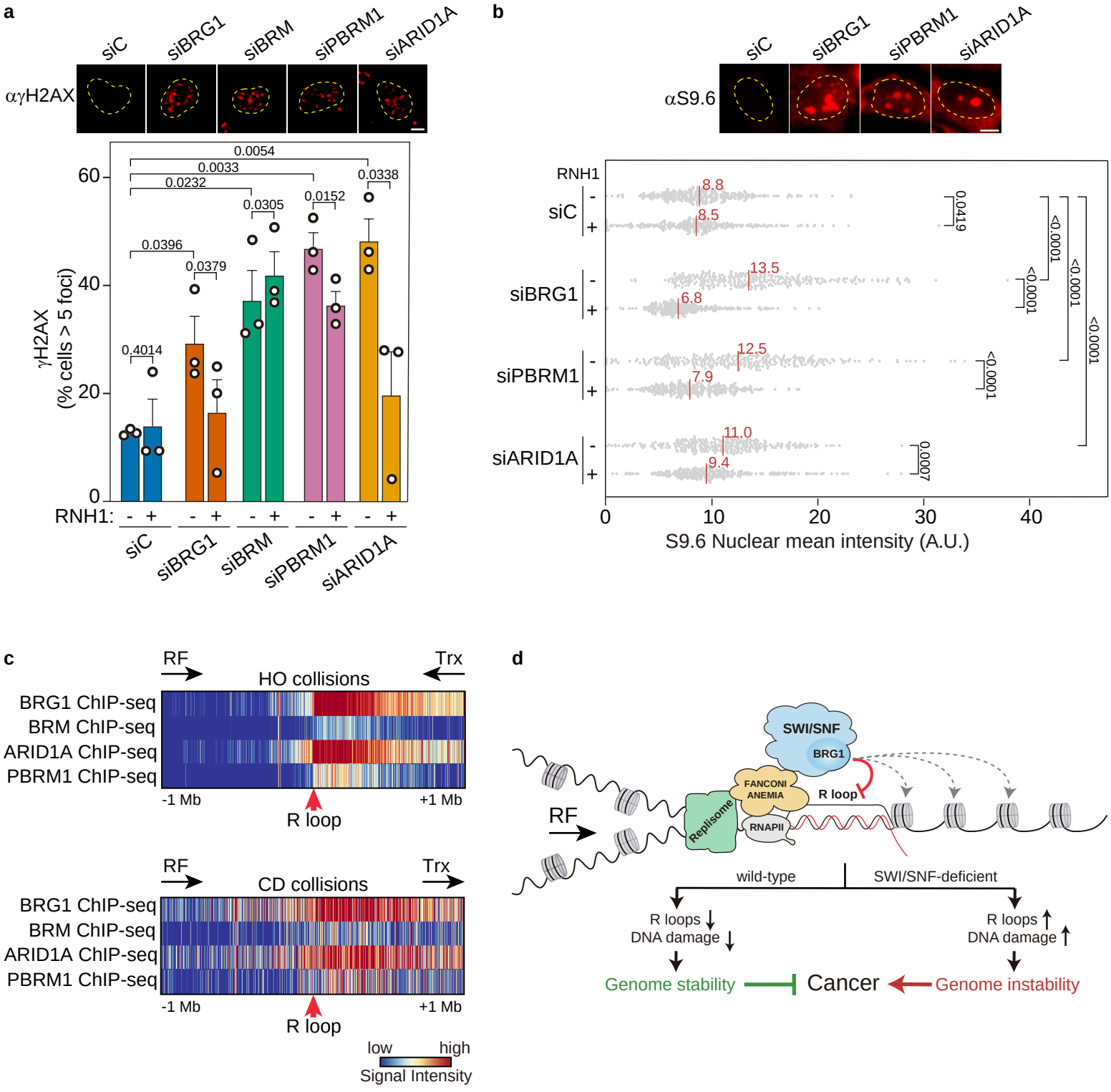

a
chr19:37,368,100-53,643,566
5 Mb
DRIPc-seq R1
DRIPc-seq R2
DRIP -seq
- RNH
+ RNH
genes
b
chr2:1-74,906,998
5 Mb
DRIPc-seq R1
DRIPc-seq R2
BRG1 ChIP-seq R1
BRG1 ChIP-seq R2
RNA-seq R1
RNA-seq R2
genes
c
BRG1 ChIP-seq
988
333
3976
R loop (DRIPc-seq)
197
9472
762
1676
Expression (RNA-seq)
d
chr8:143,420,238-144,545,485
DRIPc-seq
100 Kb
W
C
siC R1
siBRG1 R1
siC R2
siBRG1 R2
genes
ZC3H3 GSDMD EEF1D ZNF623 ZNF707 FAM83H SCRIB EPPK1 PLEC PARP10 SPATC1 WDR97 MROH1 BOP1 HSF1 FBXL6 SLC39A4 CYHR1 GPT
e
DRIPc-seq
NBEAL2
5 Kb
PKD1
5 Kb
f
DRIPc-seq
siC
siBRG1
sense
antisense
-2kb TSS
TTS +2kb
metagene

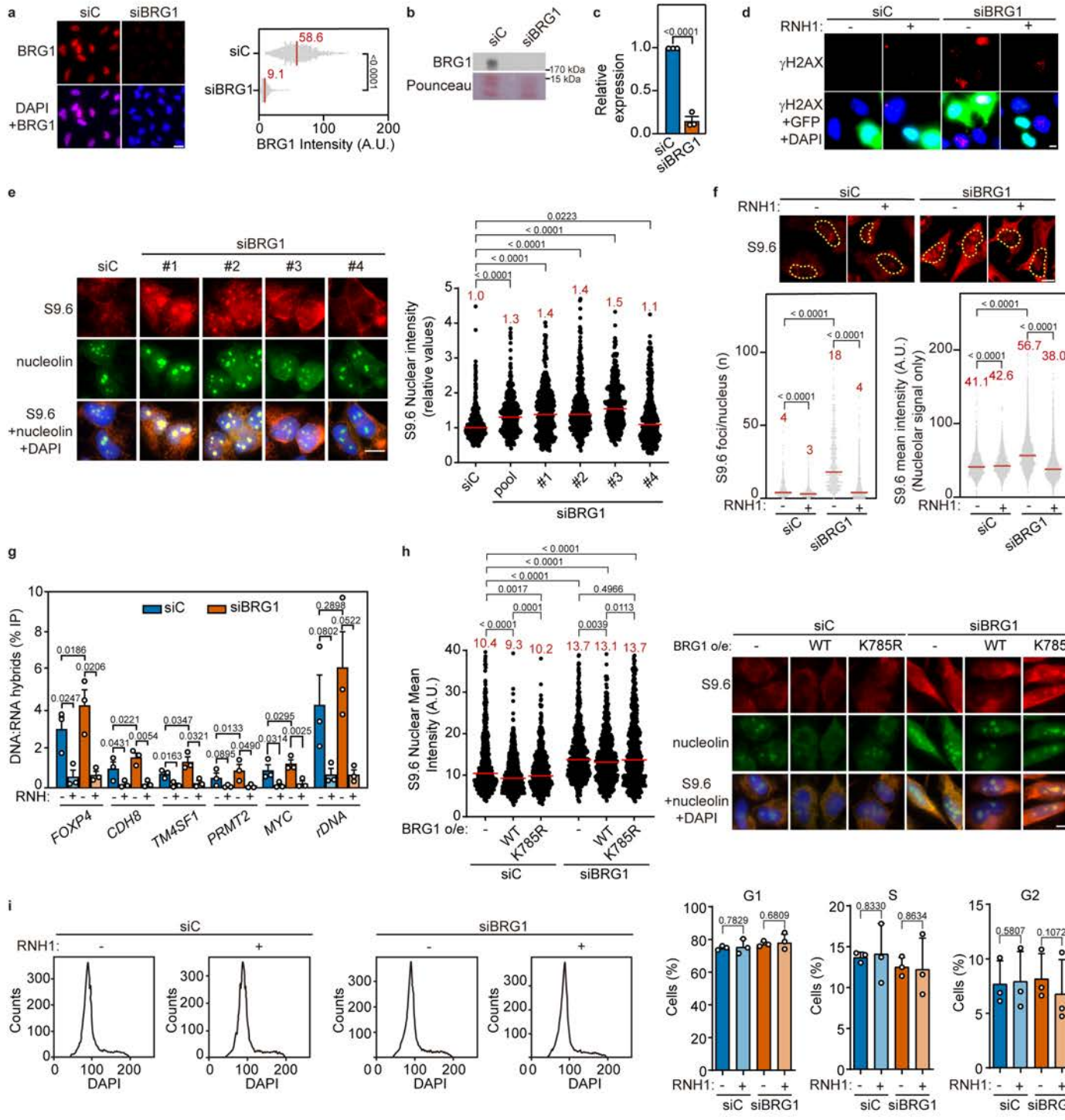

a
siBRG1
S9.6
siUAP56
siFANCD2
siBRG1
siC
siTHOC1
siUAP56
siFANCD2
siSETX
5.1
6.1
6.2
7.7
7.8
9.6
6.2
6.1
6.1
6.3
<0.0001
<0.0001
<0.0001
<0.0001
<0.0001
0.3897
0.8254
0.1931
0.0461
0
5
10
15
20
S9.6 mean intensity per nucleus (A.U.)
b
siBRG1
γH2AX
siUAP56
siFANCD2
siBRG1
siC
siTHOC1
siUAP56
siFANCD2
siSETX
0.0018
0.0008
0.0024
0.0207
0.1413
0.4689
0.0052
0.3668
0.1939
0
20
40
60
γH2AX (% cells > 5 foci)
c
αRNAPIIS2P
siBRG1
αPCNA
siUAP56
siFANCD2
siBRG1
siC
siTHOC1
siUAP56
siFANCD2
siSETX
<0.0001
<0.0001
<0.0001
<0.0001
<0.0001
0.5946
0.0659
0.0163
0.6231
0
20
40
60
RNAPIIS2P+PCNA PLA
(foci/cell)

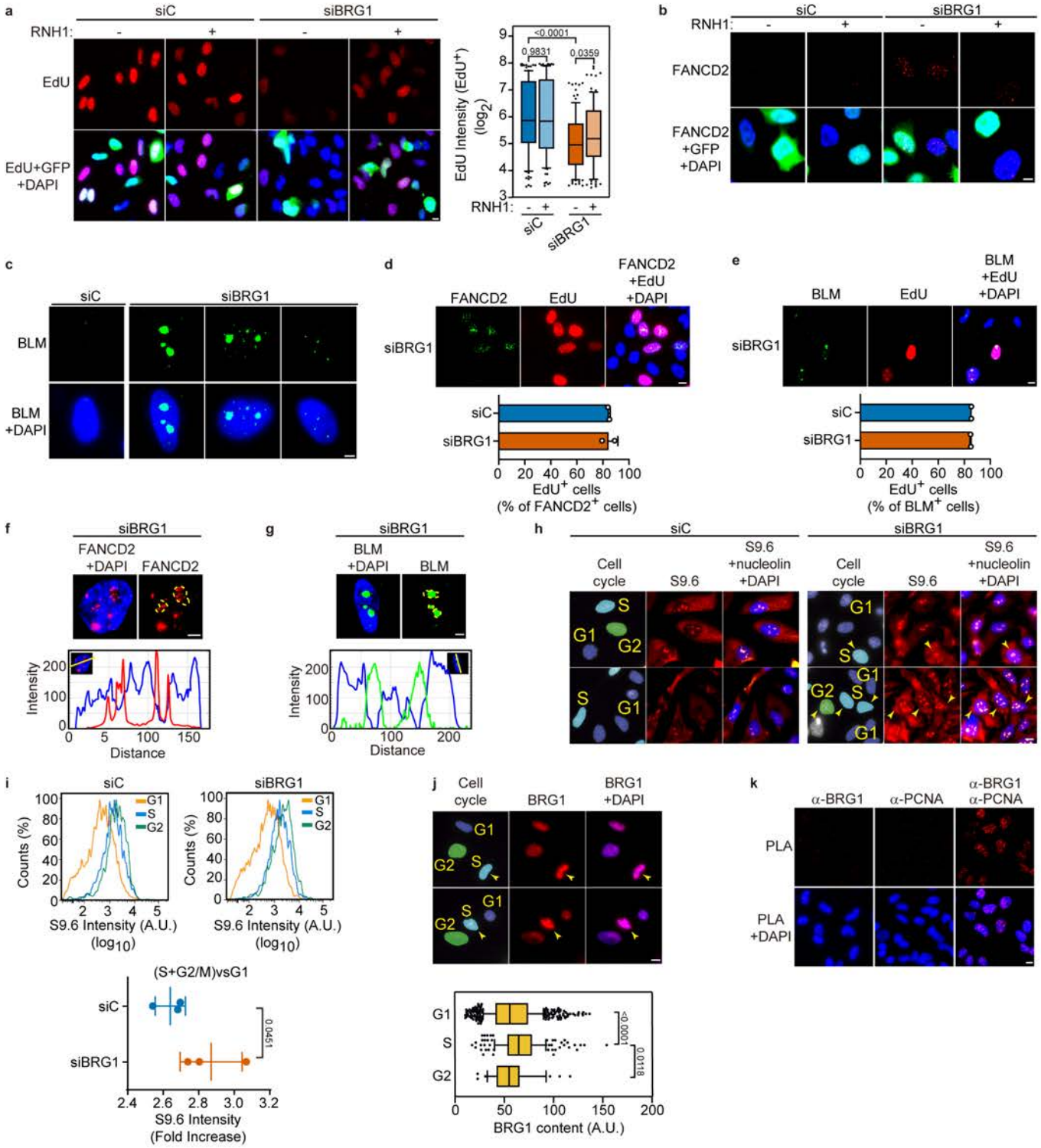

a
siC
siUAP56
RNH1:
-
+
-
+
αBRG1
αFANCD2
RNH1
siC
siUAP56
0.0022
<0.0001
<0.0001
0
100
200
300
400
FANCD2+BRG1 PLA
(foci/cell)
b
siC
siUAP56
αBRG1
αγH2AX
siC
siUAP56
<0.0001
0
50
100
150
γH2AX+BRG1 PLA
(foci/cell)
c
siC
siUAP56
αBRG1
αγRPA S48P
siC
siUAP56
<0.0001
0
50
100
150
RPA S4/8P+BRG1 PLA
(foci/cell)
d
siC
siUAP56
RNH1:
-
+
-
+
αBRG1
αS96
RNH1
siC
siUAP56
0.8901
<0.0001
<0.0001
0
10
20
30
40
50
S9.6+BRG1 PLA
(foci/cell)

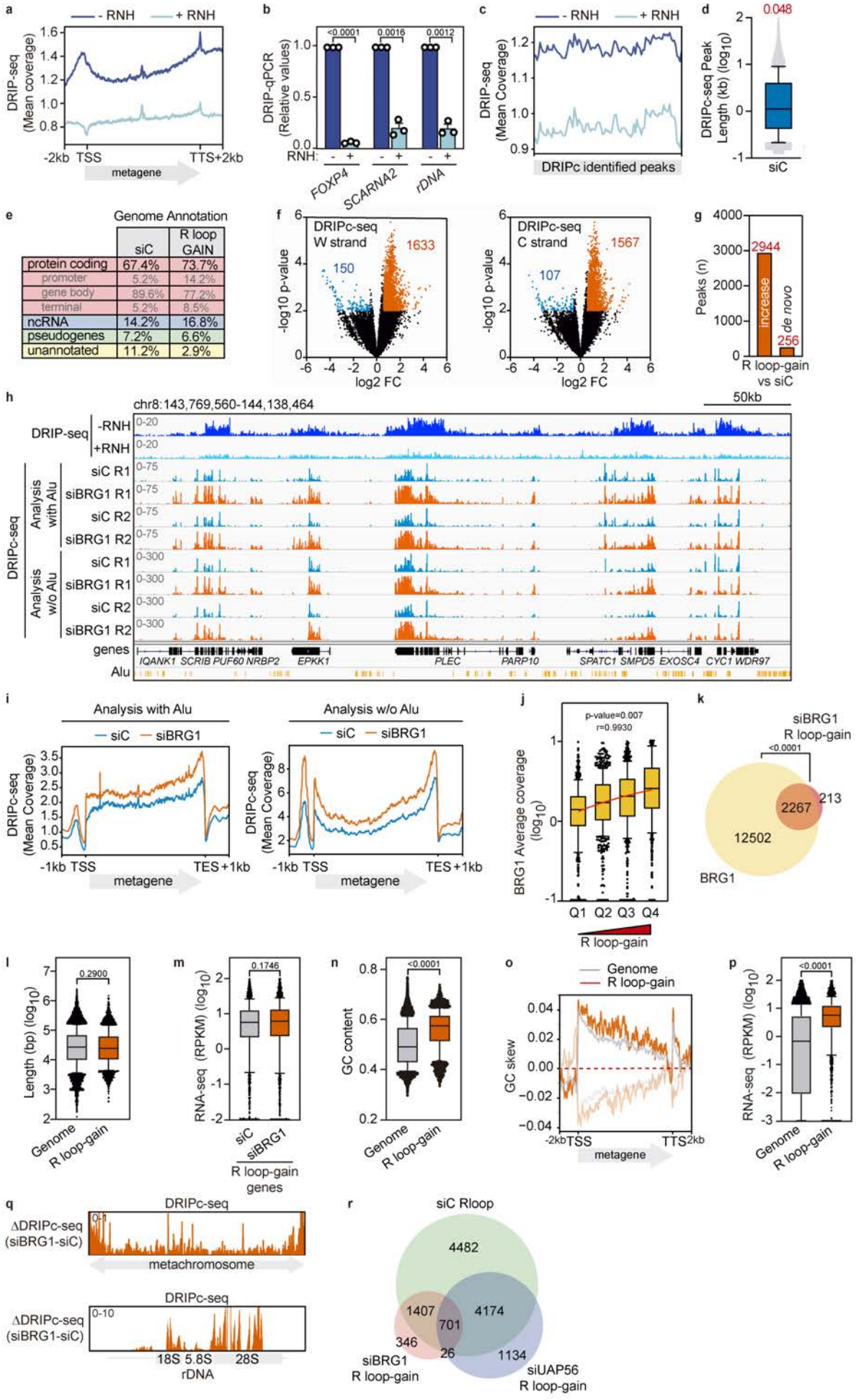

a
chr1:48,193,772-107,409,834
5 Mb
DRIPc-seq
0-5
BRG1 ChIP-seq
0-10
FANCD2 ChIP-seq
1-2
γH2AX ChIP-seq
0.5-1.25
RFD
+1
-1
genes
b
R loop (DRIPc-seq)
722
6074
237
3731
414
1037
3927
FANCD2 ChIP-seq
BRG1 ChIP-seq
c
R loop (DRIPc-seq)
661
5034
298
4771
661
1473
3491
γH2AX ChIP-seq
BRG1 ChIP-seq
d
RF
All collisions
DRIPc-seq
BRG1 ChIP-seq
FANCD2 ChIP-seq
γH2AX ChIP-seq
-1Mb
+1Mb
R loop
low
high
Signal Intensity
e
RF
Trx
HO collisions
BRG1 ChIP-seq
FANCD2 ChIP-seq
-1Mb
+1Mb
R loop
low
high
Signal Intensity
f
RF
Trx
CD collisions
BRG1 ChIP-seq
FANCD2 ChIP-seq
-1Mb
+1Mb
R loop
low
high
Signal Intensity

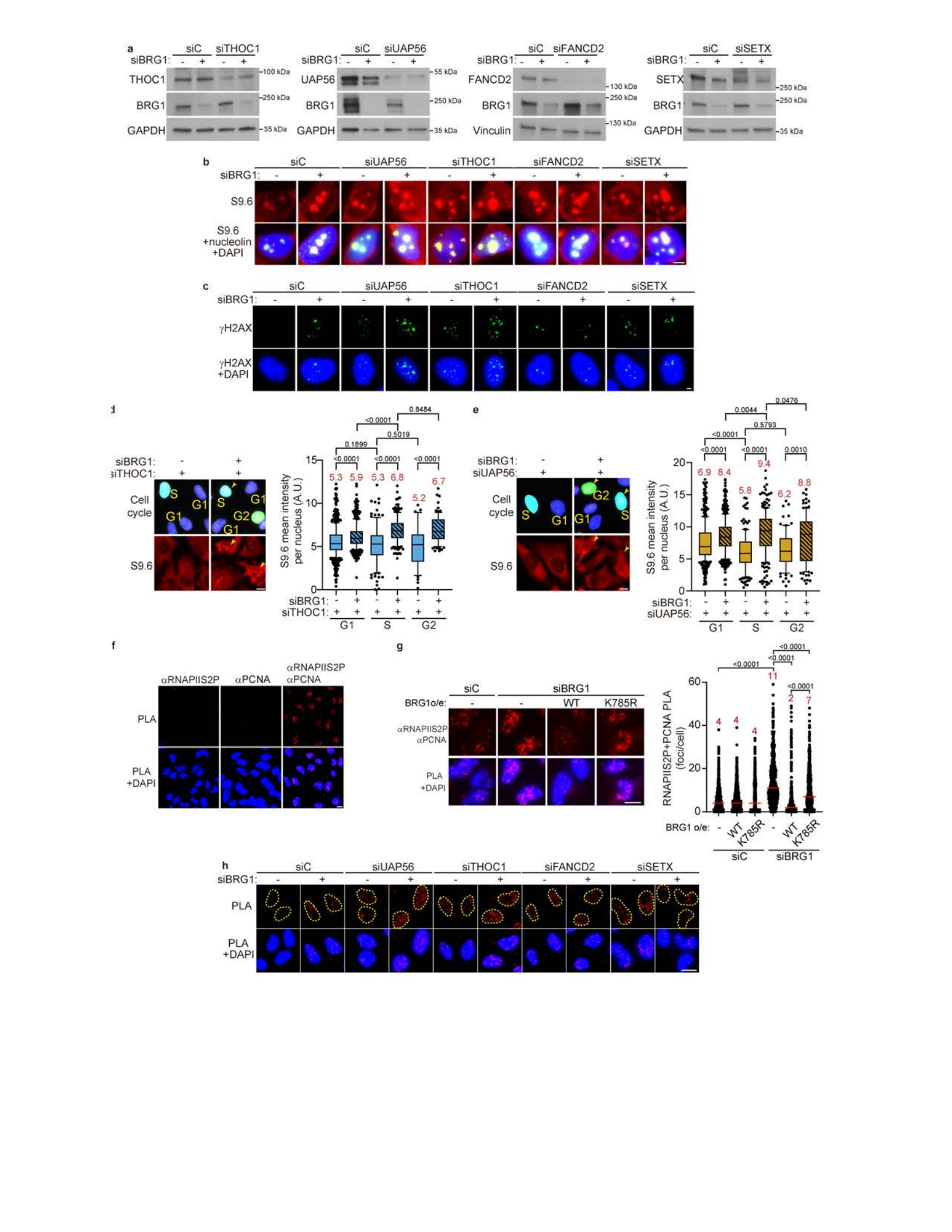

a
ATAC-seq
DNAse-seq
FAIRE-seq
MNase-seq
Mean Coverage
140
1.0
1.4
0.25
0
-10kb
+10kb
Rloop
trx
b
Average Coverage (log10)
<0.0001
<0.0001
WT
BRG1 KO
shluc
shBRG1
ATAC -seq
MNase -seq
c
0.1197
<0.0001
<0.0001
BRG1 o/e:
-
WT
T910M
K785R
ATAC-seq (BIN67)
d
RF
Trx
HO collisions
CD collisions
ATAC-seq
DNase-seq
FAIRE-seq
MNase-seq
-1 Mb
+1 Mb
R loop
low
high
Signal Intensity
e
ATAC -seq
MNase -seq
WT
BRG1 KO
shluc
shBRG1
f
BRG1 o/e:
ATAC -seq (BIN67)

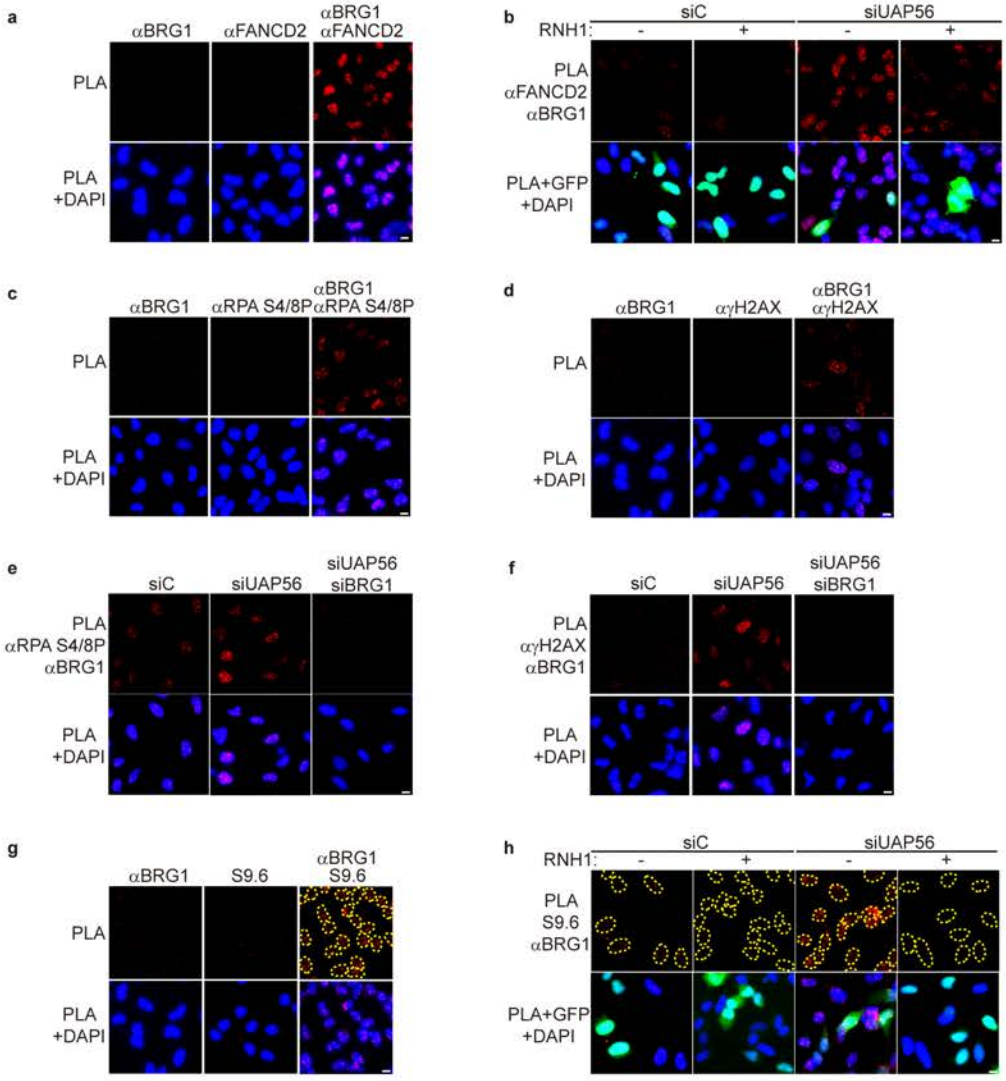

a
siC
siBRG1
siBRM
siPBRM1
siARID1A
αγH2AX
γH2AX (% cells > 5 foci)
0.0054
0.0033
0.0232
0.0338
0.0305
0.0152
0.0396
0.0379
0.4014
RNH1: - + - + - + - + - +
b
αS9.6
RNH1
siC
siBRG1
siPBRM1
siARID1A
8.8
8.5
13.5
6.8
12.5
7.9
11.0
9.4
0.0419
<0.0001
0.007
S9.6 Nuclear mean intensity (A.U.)
c
RF
Trx
HO collisions
BRG1 ChIP-seq
BRM ChIP-seq
ARID1A ChIP-seq
PBRM1 ChIP-seq
-1 Mb
+1 Mb
R loop
CD collisions
low
high
Signal Intensity
d
SWI/SNF
BRG1
FANCONI ANEMIA
R loop
Replisome
RNAPII
wild-type
SWI/SNF-deficient
R loops
DNA damage
Genome stability
Cancer
Genome instability

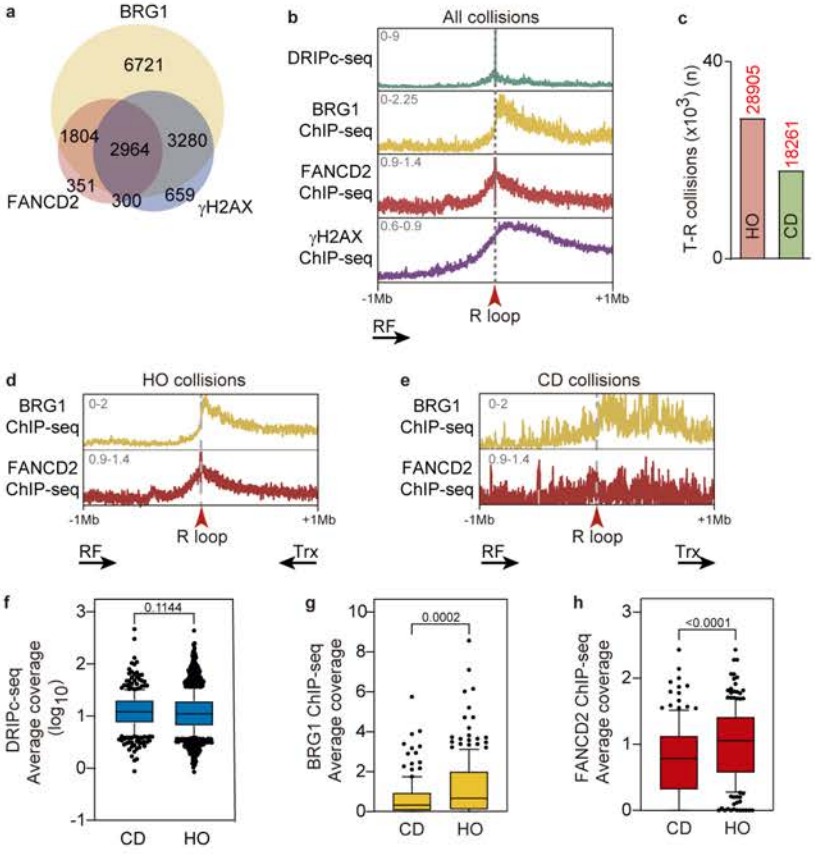

a
siC
siBRG1
BRG1
DAPI
+BRG1
58.6
9.1
siC
siBRG1
0
100
200
BRG1 Intensity (A.U.)
b
siC
siBRG1
BRG1
Pounceau
170 kDa
15 kDa
c
<0.0001
Relative expression
1.0
0.5
0.0
siC
siBRG1
d
siC
siBRG1
RNH1:
-
+
-
+
γH2AX
γH2AX
+GFP
+DAPI
e
siBRG1
siC
#1
#2
#3
#4
S9.6
nucleolin
S9.6
+nucleolin
+DAPI
0.0223
< 0.0001
< 0.0001
< 0.0001
< 0.0001
1.0
1.3
1.4
1.4
1.5
1.1
S9.6 Nuclear intensity
(relative values)
5
4
3
2
1
0
siC
pool
#1
#2
#3
#4
siBRG1
f
siC
siBRG1
RNH1:
-
+
-
+
S9.6
< 0.0001
< 0.0001
< 0.0001
18
4
4
3
S9.6 foci/nucleus (n)
100
50
0
RNH1:
siC
siBRG1
< 0.0001
< 0.0001
< 0.0001
56.7
38.0
< 0.0001
41.1
42.6
S9.6 mean intensity (A.U.)
(Nucleolar signal only)
200
100
0
RNH1:
siC
siBRG1
g
siC
siBRG1
DNA:RNA hybrids (% IP)
10
8
6
4
2
0
0.0186
0.0206
0.0247
0.0221
0.0431
0.0054
0.0347
0.0321
0.0163
0.0133
0.0490
0.0895
0.0295
0.0314
0.0025
0.2898
0.0522
0.0802
RNH:
- + - +
FOXP4
CDH8
TM4SF1
PRMT2
MYC
rDNA
h
< 0.0001
< 0.0001
< 0.0001
0.0017
0.4966
0.0001
0.0113
< 0.0001
0.0039
10.4
9.3
10.2
13.7
13.1
13.7
S9.6 Nuclear Mean
Intensity (A.U.)
40
30
20
10
0
BRG1 o/e:
-
WT
K785R
-
WT
K785R
siC
siBRG1
siC
siBRG1
BRG1 o/e:
-
WT
K785R
-
WT
K785R
S9.6
nucleolin
S9.6
+nucleolin
+DAPI
i
siC
siBRG1
RNH1:
-
+
-
+
Counts
300
200
100
0
0
100
200
DAPI
G1
S
G2
0.7829
0.6809
0.8330
0.8634
0.5807
0.1072
Cells (%)
100
80
60
40
20
0
20
15
10
5
0
15
10
5
0
RNH1:
-
+
-
+
siC
siBRG1

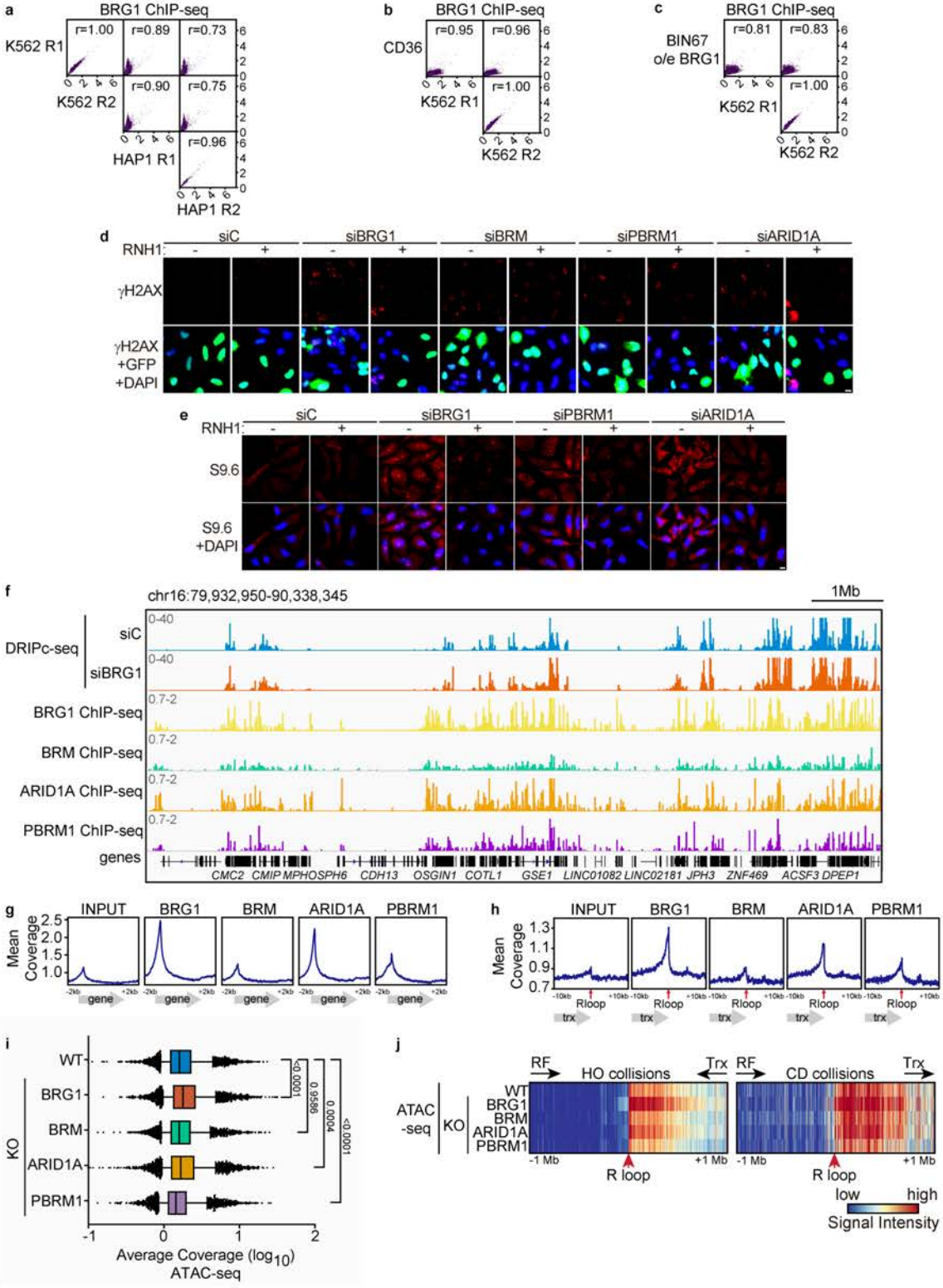

a
siC
siBRG1
RNH1:
EdU
EdU+GFP +DAPI
EdU Intensity (EdU+) (log2)
<0.0001
0.9831
0.0359
b
FANCD2
FANCD2 +GFP +DAPI
c
BLM
BLM +DAPI
d
FANCD2
EdU
FANCD2 +EdU +DAPI
siBRG1
EdU+ cells (% of FANCD2+ cells)
e
BLM
BLM +EdU +DAPI
EdU+ cells (% of BLM+ cells)
f
FANCD2 +DAPI
Intensity
Distance
g
BLM +DAPI
h
Cell cycle
S9.6
S9.6 +nucleolin +DAPI
G1
G2
S
i
Counts (%)
S9.6 Intensity (A.U.) (log10)
(S+G2/M)vsG1
0.0451
S9.6 Intensity (Fold Increase)
j
BRG1
BRG1 +DAPI
BRG1 content (A.U.)
0.0118
k
α-BRG1
α-PCNA
α-BRG1 α-PCNA
PLA
PLA +DAPI

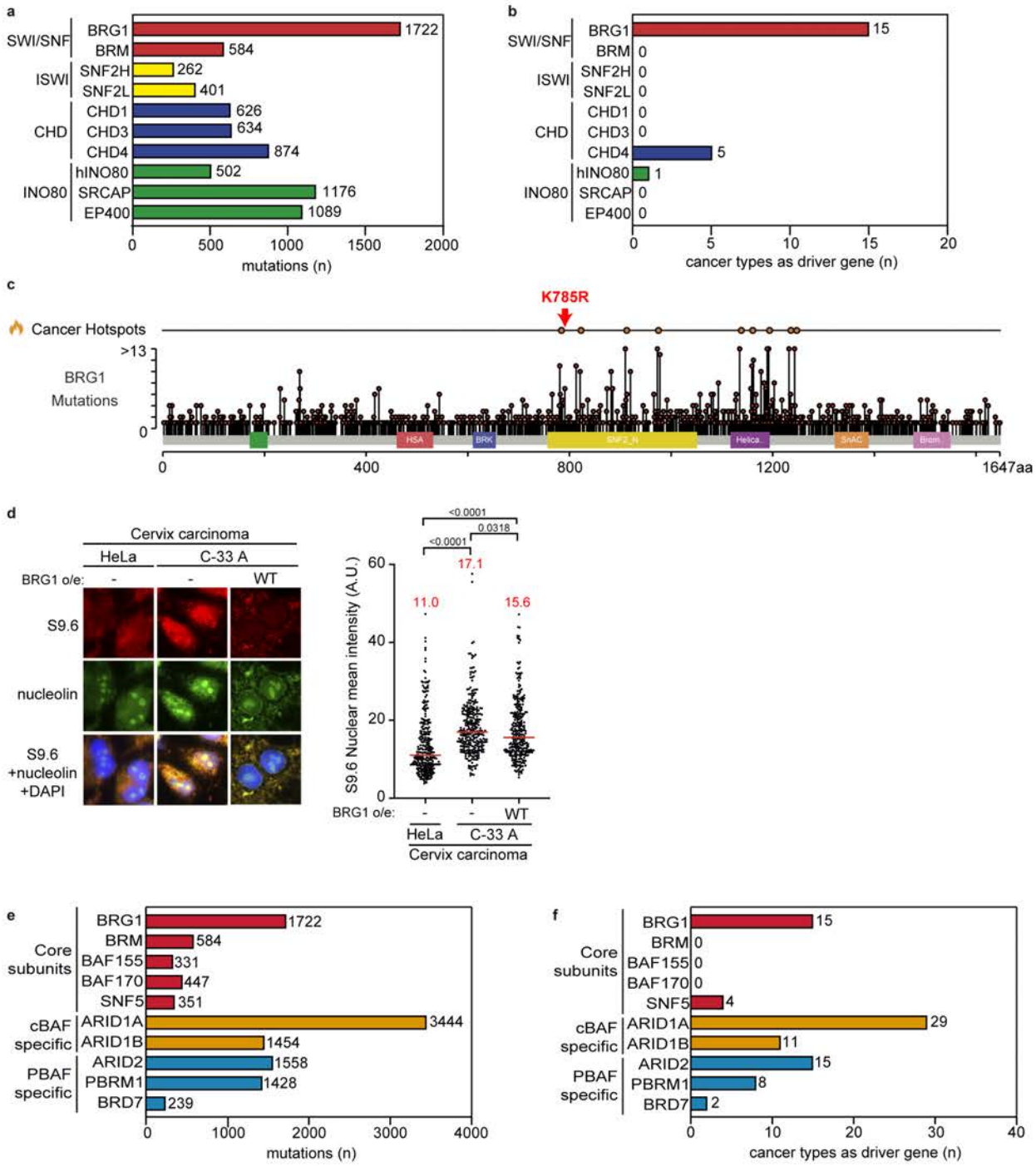

a
- RNH
+ RNH
DRIP-seq (Mean coverage)
-2kb TSS
TTS+2kb
metagene
b
DRIP-qPCR (Relative values)
<0.0001
0.0016
0.0012
RNH:
FOXP4
SCARNA2
rDNA
c
- RNH
+ RNH
DRIP-seq (Mean Coverage)
DRIPc identified peaks
d
0.048
DRIPc-seq Peak Length (kb) (log10)
siC
e
Genome Annotation
siC
R loop GAIN
protein coding 67.4% 73.7%
promoter 5.2% 14.2%
gene body 89.6% 77.2%
terminal 5.2% 8.5%
ncRNA 14.2% 16.8%
pseudogenes 7.2% 6.6%
unannotated 11.2% 2.9%
f
DRIPc-seq W strand
1633
150
DRIPc-seq C strand
1567
107
-log10 p-value
log2 FC
g
2944
256
increase
de novo
Peaks (n)
R loop-gain vs siC
h
chr8:143,769,560-144,138,464
50kb
DRIP-seq
-RNH
+RNH
DRIPc-seq
Analysis with Alu
Analysis w/o Alu
siC R1
siBRG1 R1
siC R2
siBRG1 R2
genes
Alu
IQANK1 SCRIB PUF60 NRBP2 EPKK1 PLEC PARP10 SPATC1 SMPD5 EXOSC4 CYC1 WDR97
i
Analysis with Alu
Analysis w/o Alu
siC siBRG1
DRIPc-seq (Mean Coverage)
-1kb TSS
TES +1kb
metagene
j
p-value=0.007
r=0.9930
BRG1 Average coverage (log10)
Q1 Q2 Q3 Q4
R loop-gain
k
siBRG1 R loop-gain
<0.0001
2267
213
12502
BRG1
l
0.2900
Length (bp) (log10)
Genome
R loop-gain
m
0.1746
RNA-seq (RPKM) (log10)
siC
siBRG1
R loop-gain genes
n
<0.0001
GC content
Genome
R loop-gain
o
Genome
R loop-gain
GC skew
-2kb TSS
TTS 2kb
metagene
p
<0.0001
RNA-seq (RPKM) (log10)
Genome
R loop-gain
q
DRIPc-seq
ΔDRIPc-seq (siBRG1-siC)
metachromosome
DRIPc-seq
ΔDRIPc-seq (siBRG1-siC)
18S 5.8S 28S
rDNA
r
siC Rloop
4482
1407
4174
701
346
26
1134
siBRG1 R loop-gain
siUAP56 R loop-gain


a
siC
siTHOC1
siBRG1:
THOC1
BRG1
GAPDH
100 kDa
250 kDa
35 kDa
siUAP56
UAP56
55 kDa
siFANCD2
FANCD2
130 kDa
Vinculin
siSETX
SETX
b
S9.6
S9.6 +nucleolin +DAPI
c
γH2AX
γH2AX +DAPI
d
Cell cycle
S
G1
G2
S9.6
S9.6 mean intensity per nucleus (A.U.)
0.8484
<0.0001
0.5019
0.1899
5.3
5.9
6.8
6.7
5.2
siTHOC1:
e
siUAP56:
0.0476
0.0044
0.5793
0.0010
6.9
8.4
9.4
5.8
6.2
8.8
f
αRNAPIIS2P
αPCNA
αRNAPIIS2P αPCNA
PLA
PLA +DAPI
g
BRG1o/e:
WT
K785R
αRNAPIIS2P αPCNA
RNAPIIS2P+PCNA PLA (foci/cell)
11
h


a
αBRG1
αFANCD2
αBRG1
αFANCD2
PLA
PLA
+DAPI
b
siC
siUAP56
RNH1:
-
+
-
+
PLA
αFANCD2
αBRG1
PLA+GFP
+DAPI
c
αBRG1
αRPA S4/8P
αBRG1
αRPA S4/8P
PLA
PLA
+DAPI
d
αBRG1
αγH2AX
αBRG1
αγH2AX
PLA
PLA
+DAPI
e
siC
siUAP56
siUAP56
siBRG1
PLA
αRPA S4/8P
αBRG1
PLA
+DAPI
f
siC
siUAP56
siUAP56
siBRG1
PLA
αγH2AX
αBRG1
PLA
+DAPI
g
αBRG1
S9.6
αBRG1
S9.6
PLA
PLA
+DAPI
h
siC
siUAP56
RNH1:
-
+
-
+
PLA
S9.6
αBRG1
PLA+GFP
+DAPI


a
BRG1
6721
1804
2964
3280
351
300
659
FANCD2
γH2AX
b
All collisions
DRIPc-seq
0-9
BRG1 ChIP-seq
0-2.25
FANCD2 ChIP-seq
0.9-1.4
γH2AX ChIP-seq
0.6-0.9
-1Mb
+1Mb
R loop
RF
c
T-R collisions (x10^3) (n)
40
0
28905
18261
HO
CD
d
HO collisions
BRG1 ChIP-seq
0-2
FANCD2 ChIP-seq
0.9-1.4
-1Mb
+1Mb
R loop
RF
Trx
e
CD collisions
BRG1 ChIP-seq
0-2
FANCD2 ChIP-seq
0.9-1.4
-1Mb
+1Mb
R loop
RF
Trx
f
DRIPc-seq Average coverage (log10)
0.1144
CD
HO
g
BRG1 ChIP-seq Average coverage
0.0002
CD
HO
h
FANCD2 ChIP-seq Average coverage
<0.0001
CD
HO


a
BRG1 ChIP-seq
K562 R1
K562 R2
HAP1 R1
HAP1 R2
r=1.00
r=0.89
r=0.73
r=0.90
r=0.75
r=0.96
b
BRG1 ChIP-seq
CD36
K562 R1
K562 R2
r=0.95
r=0.96
r=1.00
c
BRG1 ChIP-seq
BIN67 o/e BRG1
K562 R1
K562 R2
r=0.81
r=0.83
r=1.00
d
siC
siBRG1
siBRM
siPBRM1
siARID1A
RNH1
γH2AX
γH2AX +GFP +DAPI
e
siC
siBRG1
siPBRM1
siARID1A
RNH1
S9.6
S9.6 +DAPI
f
chr16:79,932,950-90,338,345
1Mb
DRIPc-seq
siC
siBRG1
BRG1 ChIP-seq
BRM ChIP-seq
ARID1A ChIP-seq
PBRM1 ChIP-seq
genes
CMC2 CMIP MPHOSPH6 CDH13 OSGIN1 COTL1 GSE1 LINC01082 LINC02181 JPH3 ZNF469 ACSF3 DPEP1
g
INPUT
BRG1
BRM
ARID1A
PBRM1
Mean Coverage
gene
h
INPUT
BRG1
BRM
ARID1A
PBRM1
Mean Coverage
RLoop
trx
i
KO
WT
BRG1
BRM
ARID1A
PBRM1
0.0001
0.9956
0.0004
<0.0001
Average Coverage (log10)
ATAC-seq
j
ATAC -seq
KO
WT
BRG1
BRM
ARID1A
PBRM1
RF
HO collisions
Trx
CD collisions
-1 Mb
+1 Mb
R loop
low
high
Signal Intensity


a
SWI/SNF
BRG1
BRM
ISWI
SNF2H
SNF2L
CHD
CHD1
CHD3
CHD4
INO80
hINO80
SRCAP
EP400
1722
584
262
401
626
634
874
502
1176
1089
0
500
1000
1500
2000
mutations (n)
b
15
0
5
1
0
10
20
cancer types as driver gene (n)
c
K785R
Cancer Hotspots
>13
BRG1
Mutations
0
HSA
BRK
SNF2_N
Helica
SnAC
Bromo
400
800
1200
1647aa
d
Cervix carcinoma
HeLa
C-33 A
BRG1 o/e:
WT
S9.6
nucleolin
S9.6
+nucleolin
+DAPI
<0.0001
0.0318
<0.0001
17.1
11.0
15.6
S9.6 Nuclear mean intensity (A.U.)
60
40
20
0
e
Core
subunits
cBAF
specific
PBAF
specific
BRG1
BRM
BAF155
BAF170
SNF5
ARID1A
ARID1B
ARID2
PBRM1
BRD7
1722
584
331
447
351
3444
1454
1558
1428
239
0
1000
2000
3000
4000
mutations (n)
f
15
0
0
0
4
29
11
15
8
2
0
10
20
30
40
cancer types as driver gene (n)