MOLECULAR BIOLOGY

ZC3H4 safeguards genome integrity by preventing transcription-replication conflicts at noncoding RNA loci

Yann Frey[1], Liana Goehring[2], Majd Haj[1], Gergely Rona[2,3,4], Carel Fijen[2], Michele Pagano[2,4], Tony T. Huang[2], Eli Rothenberg[2], Yael Ziv[1], Yosef Shiloh[1]*

The cellular networks that maintain genome stability encompass numerous pathways involved in all aspects of nucleic acid metabolism. Through bioinformatic analysis, we identified the Zinc Finger CCCH-Type Containing 4 protein (ZC3H4), a suppressor of noncoding RNA (ncRNA) production, as a pivotal player in this system. Experimentally, ZC3H4 deficiency led to increased DNA damage, abnormal mitosis, and cellular senescence. Biochemical analysis and super-resolution microscopy revealed that the loss of ZC3H4 increased replication stress (RS)—a major driver of genome instability—by inducing a hypertranscription state that promoted R loop formation and transcription-replication conflicts (TRCs), both of which drive RS. Further bioinformatic analysis demonstrated that ZC3H4 preferentially binds to genomic regions prone to TRCs and R loops, where it suppresses ncRNA bursts, functioning as part of the Restrictor complex. Our findings identify ZC3H4 as a crucial factor in maintaining genome integrity, strategically positioned at the critical intersection of DNA and RNA synthesis.

INTRODUCTION

Maintaining genome integrity is essential for cellular homeostasis and the prevention of unscheduled cell death or neoplastic transformation. A substantial threat to genome stability is ongoing DNA damage, which is continually managed by the DNA damage response (DDR) signaling network (*1*, *2*). The loss or diminished efficiency of DDR pathways is a common cause of genome instability, with severe impacts on an organism's health and the rate of aging (*3*, *4*). Another critical source of genome instability is replication stress (RS), triggered by disruptions in DNA replication. RS can be caused by DNA lesions that stall replication fork progression, imbalances in nucleotide pools, genomic sequences that challenge the replication machinery, secondary DNA structures, oncogene-induced hyperproliferation, and transcription-replication conflicts (TRCs) (*5*). Prolonged RS can have profound consequences, including erroneous replication causing DNA sequence alterations, replication fork collapse leading to DNA double-strand breaks (DSBs), chromosomal aberrations, aneuploidies, and defective mitosis, ultimately leading to cellular neoplasia, senescence, or programmed cell death (*5*, *6*). Cells avoid or manage RS by activating damage tolerance pathways, mechanisms that stabilize the replication fork or facilitate fork reversal, and specific DDR branches, such as the elaborate DSB response (*5*, *7*, *8*).

TRCs are a significant source of RS (*9*, *10*). These conflicts arise during the S phase when the transcription and replication machinery collide, potentially leading to replication fork stalling. Fork obstruction can be caused by RNA polymerase itself or by R loops that occasionally form at transcription sites (*9*–*13*). These collisions are more likely to occur in highly transcribed regions and can be further aggravated by aberrant firing of replication origins (*9*, *14*, *15*) and pervasive transcription (*16*–*18*). The latter primarily generates a diverse array of noncoding RNAs (ncRNAs) (*19*–*21*). Consequently, mitigating pervasive transcription may help reduce TRCs, thereby decreasing RS and genome instability.

Several protein complexes are involved in the selective termination of pervasive transcription, including the Integrator complex, which also has broader roles in transcriptional regulation (*22*–*24*). Recently, another protein complex, known as Restrictor, has been identified as playing a crucial role in this process by negatively affecting the synthesis of a wide array of ncRNAs, including enhancer RNAs and promoter upstream transcripts (PROMPTs) (*25*–*30*). A key component of this complex is the zinc-finger protein ZC3H4, which interacts dynamically with various proteins such as the WD Repeat Domain 82 protein (WDR82), the Arsenate Resistance Protein (ARS2), Casein kinase 2 (CK2), and the Zinc Finger CCHC-Type Containing 8 protein (ZCCHC8)—a component of the nuclear exosome targeting (NEXT) complex (*25*, *26*, *28*, *31*–*34*). Restrictor acts synergistically with the protein phosphatase 1 regulatory subunit (PNUTS), Symplekin [part of the Cleavage and polyadenylation specificity factor (CPSF) complex], and the PAF complex, all of which are involved in transcription termination and RNA cleavage (*27*–*29*). Notably, ZC3H4 is consistently found in protein assemblies of varying compositions that counteract the pervasive transcription of ncRNAs (*25*, *26*, *28*–*34*), underscoring its pivotal role in this process, likely in conjunction with its closest interactor, WDR82 (*29*). ZC3H4 is essential for embryo development (*35*) and is conserved across metazoans (*25*), highlighting its crucial function.

Here, we demonstrate that ZC3H4, as a negative regulator of pervasive ncRNA transcription, plays a crucial role in maintaining genome stability by preventing TRCs at ncRNA sites. As a result, the loss of ZC3H4 induces RS and its associated cellular consequences: perturbed cellular growth, genome instability, severe mitotic defects, and premature senescence. Our findings position ZC3H4 as a previously unidentified and essential factor in the expanding network of mechanisms that safeguard genome stability, acting at the critical intersection between DNA replication and transcription.

[1]The David and Inez Myers Laboratory for Cancer Research, Department of Human Molecular Genetics and Biochemistry, Faculty of Medical & Health Sciences, Tel Aviv University, Tel Aviv 6997801, Israel. [2]Department of Biochemistry and Molecular Pharmacology, New York University Grossman School of Medicine, New York, NY 10016, USA. [3]Research Center for Natural Sciences, Institute of Molecular Life Sciences, Budapest 1117, Hungary. [4]Howard Hughes Medical Institute, New York University Grossman School of Medicine, New York, NY 10016, USA.
*Corresponding author. Email: yossih@tauex.tau.ac.il

RESULTS

Meta-analysis reveals ZC3H4 as a potential novel player in genome stability and cellular senescence

To identify potential novel DDR players, we recently performed a meta-analysis of 34 published screens that used functional, proteomic, or phosphoproteomic approaches to identifying these factors (*36*). This integrated database has previously been used to uncover unrecognized DDR components (*36*–*38*). Protein-coding genes were assigned a score reflecting the number of independent screens in which they were identified as hits (Fig. 1A). Most genes (94%) received a score below 5, while 5% scored between 6 and 10. Only the top 1% of genes had a score ranging from 10 to 15. To define our selection threshold, we analyzed the mutation rates of these genes in cancer using the Cancer Cell Line Encyclopedia (CCLE) Broad 2019 database (*39*). Ranking genes by their meta-analysis score (fig. S1A) revealed a positive correlation with their mutation frequency, with a marked increase in mutation rate for genes scoring above 10. Notably, these genes with a score of 10 or above included several well-established DDR factors (Fig. 1A). This reinforces the relevance of these high-scoring genes, suggesting that they may contribute to genome instability in cancer. Therefore, genes with a score of >10 were selected to be further investigated as potential previously unidentified DDR factors.

We used a similar approach using 22 RNA sequencing (RNA-seq) datasets representing the transcriptomic analysis of senescent cells (*40*). Briefly, we preprocessed the raw data from these studies and assessed differential gene expression between proliferative and senescent cells. Each gene in the compiled database was assigned a score based on the number of studies in which its expression was significantly deregulated in senescent cells (table S1). The distribution of these scores highlighted genes consistently up- or down-regulated during senescence, including known senescence markers (Fig. 1B, top). Notably, many genes that were down-regulated in senescent cells were also found to be down-regulated in human primary fibroblast lines derived from aging individuals (*40*) (RNA-seq dataset GSE113957). We observed that most of the top genes identified in our DDR meta-analysis were also down-regulated during senescence (Fig. 1B, bottom). This result aligns with our recent observation that DNA repair pathways are generally down-regulated during senescence (*40*). Consequently, this finding prompted us to focus on genes that are down-regulated during senescence, as potential candidates encoding DDR components. An empirical threshold was set at the top 3% of the most down-regulated genes.

The Venn diagram in Fig. 1C illustrates the overlaps among the top 1% of genes identified in the DDR meta-analysis (Fig. 1A) and the top 3% of genes down-regulated during senescence (Fig. 1B, top). To further refine our list and explore potential drivers of aging, we intersected these datasets with genes down-regulated in aging human fibroblasts (table S1, RNA-seq dataset GSE113957). Seven genes were common to all three groups (Fig. 1C). Within this intriguing small group, largely composed of genes known for their roles in the DDR and senescence, *ZC3H4* emerged as particularly notable. A closer examination of its expression dynamics in senescing cells revealed a significant decrease at the onset of overt senescence (Fig. 1D). At the time, ZC3H4 was recognized as a zinc-finger protein involved in repressing the expression of certain ncRNAs (*25*, *26*). To further explore its potential role in maintaining genome stability, we used the AnalyzeR tool (*41*) to identify genes coexpressed with *ZC3H4* (table S2). Protein-coding genes were ranked on the basis of their coexpression with *ZC3H4*, and gene set enrichment analysis (GSEA) was performed to identify the biological processes associated with these genes (Fig. 1E and table S2). This analysis highlighted pathways correlated with *ZC3H4* expression, which were regrouped and labeled using the Cytoscape AutoAnnotate app (*42*). Notably, these processes included "cell cycle checkpoint," "hallmark G_2-M checkpoint," "reactome regulation of p53," and "hallmark DNA repair," as visualized using the Enrichment Map network-based method (*43*) (Fig. 1F). These findings further suggest a potential role for ZC3H4 in cell cycle regulation and the maintenance of genome stability. This is further supported by previous observations of pronounced genome instability in early-stage *Zc3h4*$^{-/-}$ mouse embryos, which failed to develop further, leading to embryonic lethality of this genotype (*35*). Additional support comes from the frequent occurrence of *ZC3H4* mutations observed in the CCLE (*44*) (fig. S1B), as reported on the cBioPortal platform (*45*).

ZC3H4 loss impairs cellular proliferation and triggers senescence hallmarks

Building on the bioinformatic findings, we experimentally characterized the cellular phenotype of ZC3H4-deficient cells. We used CRISPR-Cas9 technology to ablate the *ZC3H4* gene in two human immortalized cell lines (the breast cancer cell line, CAL-51, and human embryonic kidney cell line, HEK-293) and in a human primary fibroblast line (fig. S2A). ZC3H4-deficient cells exhibited reduced proliferation rates, as measured by the water-soluble tetrazolium dye (WST-8) metabolic assay (Fig. 2A) (*46*), reduced rate of incorporation of 5-ethynyl-2′-deoxyuridine (EdU) into cellular DNA (Fig. 2B), and decreased colony formation ability (Fig. 2C). Cell lines lacking ZC3H4 also exhibited typical senescence markers: elevated senescence-associated β-galactosidase (SA-β-Gal) activity (Fig. 2D), decreased levels of Lamin B1 protein, and elevated levels of the p21$^{Cip1/Waf1}$ protein and phosphorylated histone H2AX (γH2AX) (Fig. 2E and fig. S2A). The levels of ZC3H4 in primary human fibroblasts that underwent replicative senescence or ionizing radiation (IR)–induced senescence were reduced compared to those in a proliferative state (Fig. 2F), which is consistent with the results from the transcriptomic meta-analysis (Fig. 1D). The depleted cells also displayed another hallmark of cellular senescence—a high level of basal, ongoing DNA damage evidenced by elevated numbers of nuclear foci of the p53-binding protein 1 (53BP1) protein (Fig. 2G) and phosphorylated form of H2A histone family member X (γH2AX) (Fig. 2H). This phenomenon was observed in both proliferating and senescing ZC3H4-deficient cells (Fig. 2, G and H), suggesting that genome instability characterized these cells before overt senescence appeared. Recently, we observed a broad reduction in DNA repair capacity in senescent cells, spanning most of the major DNA repair pathways (*40*). Accordingly, DSB repair was retarded in ZC3H4-depleted cells as evidenced by the lagging disappearance of 53BP1 nuclear foci after treatment with a moderate dose of IR (Fig. 2I).

Overall, ZC3H4 depletion significantly impaired cell proliferation across several cell lines, primarily by inducing senescence. While senescence itself can contribute to genome instability, as previously reported (*40*), our findings suggest that ZC3H4 loss first triggers cellular stress leading to senescence, likely as a consequence of underlying genome instability.

ZC3H4 loss disrupts cell cycle progression by triggering RS

A general inspection of the ZC3H4-deficient cell cultures revealed large, polynucleated cells, which were significantly more frequent

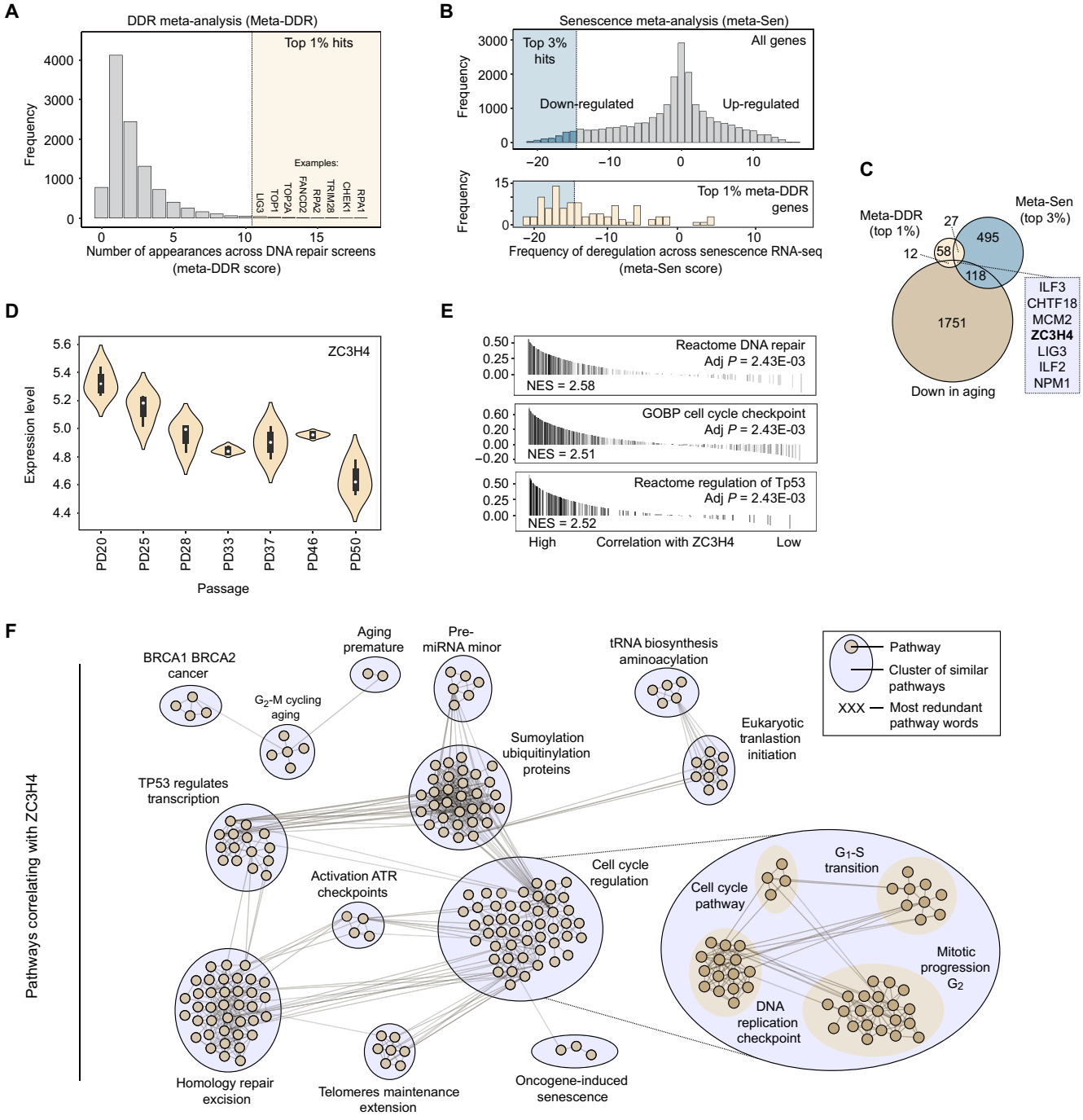

Fig. 1. Bioinformatic analysis identifies ZC3H4 as a potential previously unknown player in DNA repair and cell senescence. (**A**) Score distribution of genes identified in a meta-analysis of 40 screens for DDR factors. Each screen hit adds one point to the gene's score. (**B**) Score distribution of genes from a meta-analysis of 22 RNA-seq datasets obtained from senescent cells. Up-regulation during senescence in a screen adds +1 to the gene's score, while down-regulation adds −1. The top panel shows the distribution of all protein coding genes, and the bottom panel shows the distribution within the top 1% genes from the DDR meta-analysis (A). Most of these genes were down-regulated during senescence. (**C**) Venn diagram showing genes with the DDR score of >10, a senescence score of <−15, and genes significantly down-regulated in cell lines from elderly subjects (GSE113957 dataset). Down-regulated genes in aging were identified by comparing samples from individuals aged 90 to 96 to those aged 0 to 10, using a threshold of adjusted (Adj) $P < 0.01$ and $\log_2$ fold change < −0.58. Seven genes, including *ZC3H4*, are at the intersection of these groups. (**D**) Expression levels of *ZC3H4* across cell passages (GSE175533 dataset). (**E**) Gene set enrichment analysis (GSEA) plot for pathways significantly coexpressed with *ZC3H4*. NES, normalized enrichment score. The DNA repair, Cell Cycle Checkpoint and Regulation of the Tumor Protein P53 (Tp53) pathways were significantly enriched. (**F**) Network of GSEA pathways enriched for genes coexpressed with *ZC3H4*. A brown dot represents an enriched pathway (adjusted $P < 0.001$), and lines indicate group similarities based on shared genes. Highly similar pathways are grouped in light-beige circles (generated using Cytoscape and Enrichment Map) and labeled according to pathway content (computed with AutoAnnotate). The "cell cycle regulation" cluster is subclustered in the right-hand circle. miRNA, microRNA.

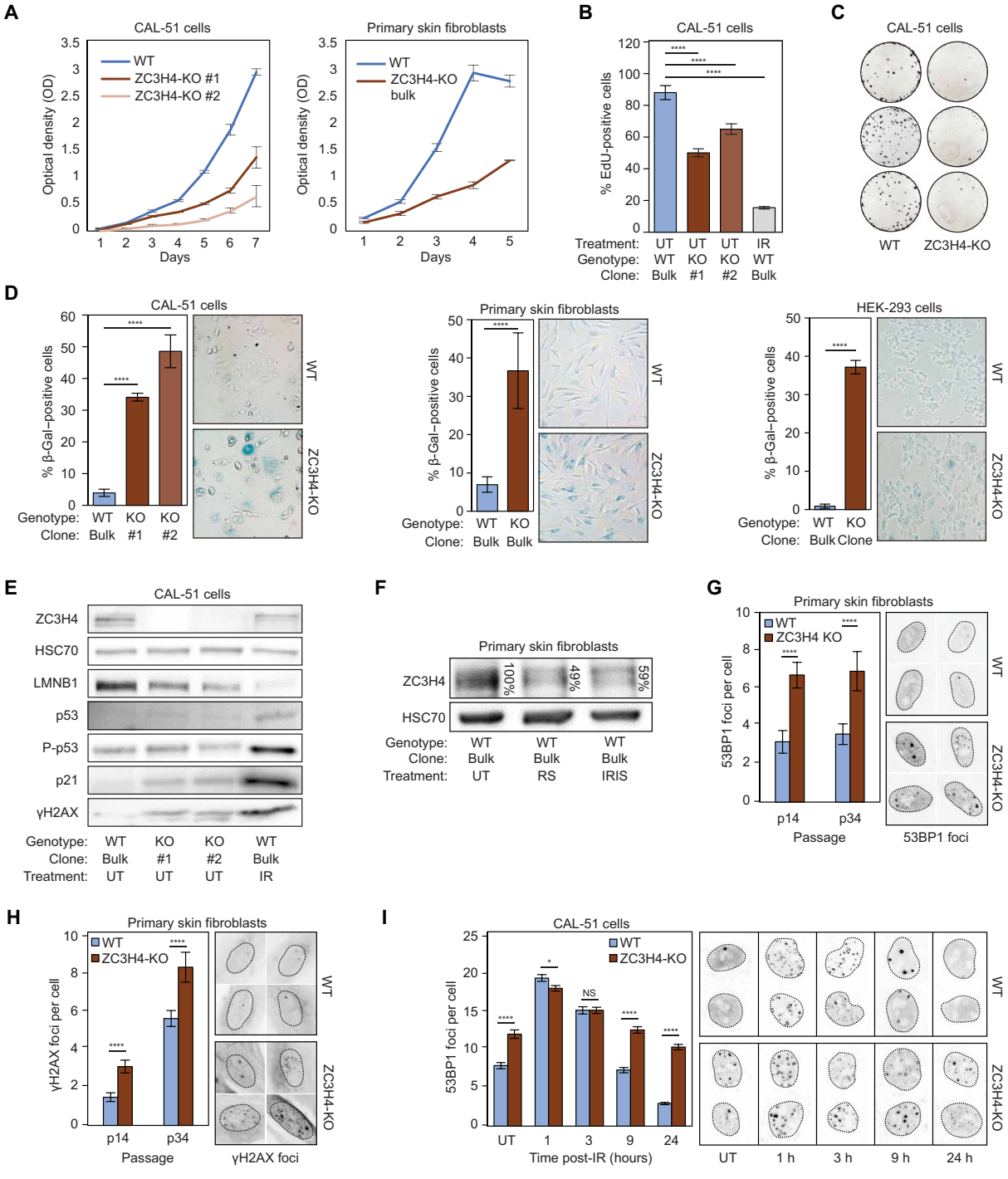

Fig. 2. ZC3H4 depletion induces senescence and genome instability hallmarks. (**A**) The WST-8 cellular viability assay reveals decreased proliferation rate in two single-cell clones of ZC3H4-depleted CAL-51 cells (left) and in a bulk population of ZC3H4-depleted human primary skin fibroblasts (right) when compared to control. (**B**) EdU incorporation assay shows reduced proliferation rates in ZC3H4-depleted CAL-51 cells. Cells irradiated (IR) with 10 Gy of x-rays and incubated for 10 days serve as a senescence positive control. UT, untreated. (**C**) ZC3H4 depletion significantly reduces the colony formation capability of CAL-51 cells. (**D**) ZC3H4-depleted cancerous and primary fibroblast lines exhibit a substantial increase in SA-β-Gal staining. (**E**) Western blotting analysis of senescence markers demonstrates robust induction of senescence in ZC3H4-deficient CAL-51 clones. (**F**) Western blotting analysis shows reduced ZC3H4 levels in senescent primary fibroblasts compared to proliferating ones. Senescence was either replicative or induced by IR. (**G**) Amounts of nuclear 53BP1 foci in ZC3H4-proficient and -deficient human primary fibroblasts at passage levels 14 and 34 suggest increased basal levels of DSBs in the absence of ZC3H4. (**H**) Amounts of nuclear γH2AX foci in ZC3H4-proficient and -deficient human primary fibroblasts at passage levels 14 and 34 suggest increased basal levels of DSBs in the absence of ZC3H4. (**I**) Quantification of 53BP1 foci in CAL-51 cells reveals elevated basal levels and delayed resolution of these foci in *ZC3H4*–knockout (KO) cells following irradiation with 1 Gy of x-rays. NS, not significant; h, hour. "ns" = P > 0.05; *$P \leq 0.05$; **$P \leq 0.01$; ***$P \leq 0.001$; ****$P \leq 0.0001$.

compared to the corresponding wild-type (WT) cells (Fig. 3A). EdU incorporation assays indicated that these polynucleated cells were not proliferating, with a reduction of 9.24-fold in EdU incorporation rate compared to depleted mononucleated cells (hypergeometric test, $P = 5.10 \times 10^{-19}$) (Fig. 3B). Similarly, these cells showed a 2.08-fold higher level of SA-β-Gal staining compared to mononucleated cells (hypergeometric test, $P = 1.84 \times 10^{-32}$), pointing to a higher senescence rate associated with polynucleation (Fig. 3C). Polynucleated cells typically indicate mitotic failures. Another common abnormality in cell division among genomically unstable cells is the presence of mitotic chromatin bridges, a hallmark of irregular chromosomal segregation (*47*). We observed a clear increase in these bridges in ZC3H4-depleted cells (Fig. 3D). To further investigate the role of ZC3H4 in cell cycle dynamics, we monitored its levels throughout the cell cycle phases following culture synchronization using a double thymidine block. ZC3H4 levels were high during the S, G_2, and M phases, with

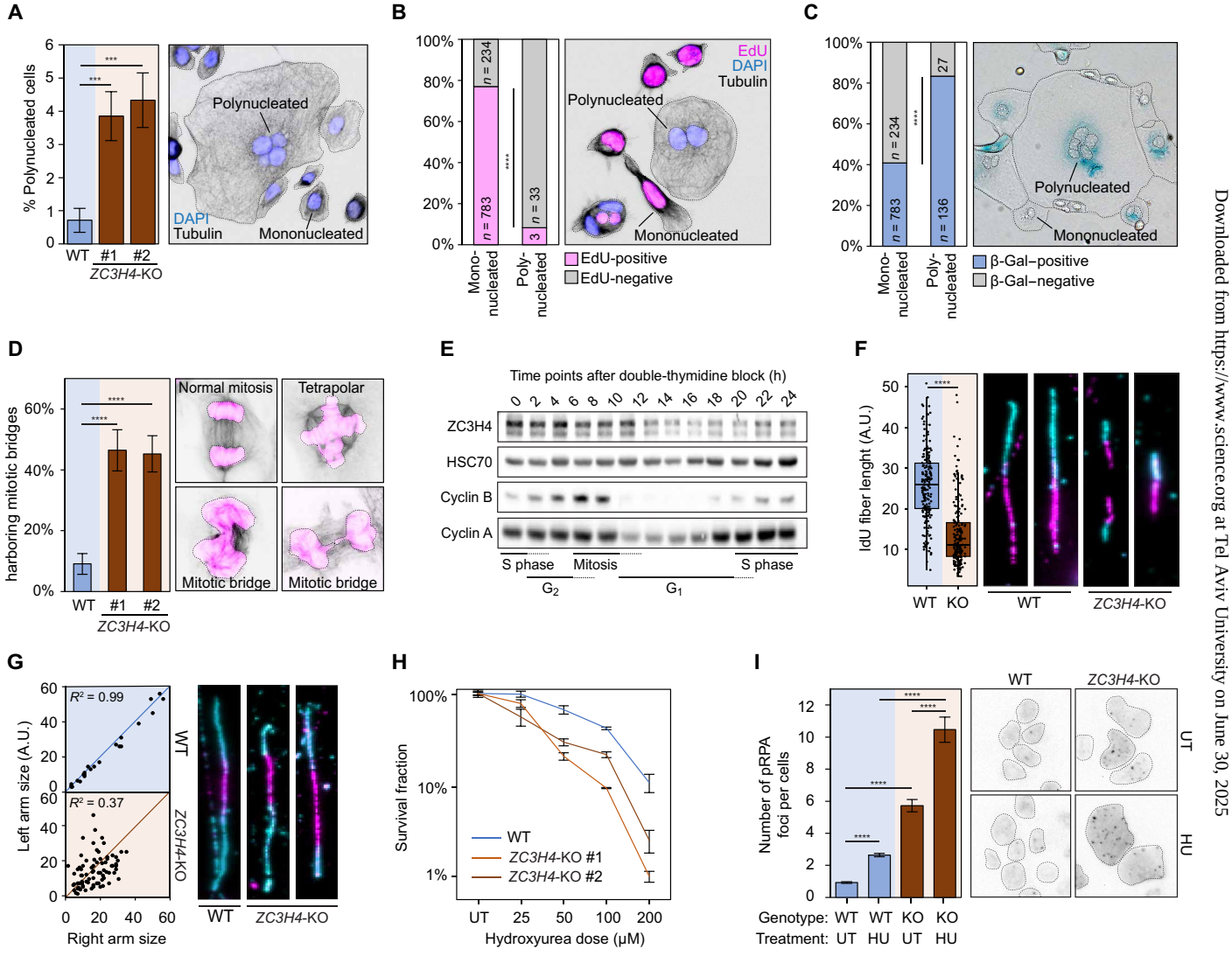

Fig. 3. ZC3H4 loss affects the cell cycle by inducing RS. (**A**) A marked increase in the frequency of multinucleated cells in ZC3H4-depleted CAL-51 cells compared to WT cells. (**B**) EdU incorporation assay indicates that the multinucleated cells within the ZC3H4-deficient cell population are largely nonproliferative. Hypergeometric test: Enrichment for EdU-positive cells in the multinucleated cell population ($P = 5.10 \times 10^{-19}$, enrichment factor = −9.24-fold). (**C**) SA-β-Gal staining shows that a high proportion of the *ZC3H4*$^{-/-}$ multinucleated cells are senescent. Hypergeometric test: Enrichment for SA-β-Gal–positive cells in the multinucleated cell population ($P = 1.84 \times 10^{-32}$, enrichment factor = 2.08-fold). (**D**) The frequency of mitotic bridges is fivefold higher in ZC3H4-deficient cells compared to controls. Representative microscopies display the DNA [4′,6-diamidino-2-phenylindole (DAPI)] in pink and tubulin in black. (**E**) Western blotting analysis of human primary fibroblasts synchronized through a double thymidine block reveals a cell cycle–dependent dynamics of ZC3H4 level, which is highest in S-G_2-M and gradually decrease during G_1. (**F**) DNA fiber assay shows a 2.4-fold decrease in replication fork progression rate in ZC3H4-depleted cells compared to controls, indicating RS. 5-Chloro-2′-deoxyuridine (CldU) is displayed in pink and 5-iodo-2′-deoxyuridine (IdU) is displayed in cyan. A.U., arbitrary units. (**G**) Fork asymmetry analysis shows decreased symmetry in ZC3H4-deficient cells, consistent with replication fork stalling. In WT cells, left and right fork arms were highly correlated ($R^2 = 0.99$), indicating equal progression. In contrast, ZC3H4-depleted cells showed a marked reduction in correlation ($R^2 = 0.37$), suggesting frequent fork blockage. (**H**) Cellular survival assay based on colony-forming ability demonstrates hypersensitivity of *ZC3H4*$^{-/-}$ CAL-51 cells to hydroxyurea (HU)—an RS inducer. (**I**) Immunofluorescence staining of phosphorylated RPA (pSer^{33})—an RS marker—reveals elevated RS in ZC3H4-deficient CAL-51 cells, compared to controls. HU was used as positive control for RS induction. "ns" = $P > 0.05$; *$P \leq 0.05$; **$P \leq 0.01$; ***$P \leq 0.001$; ****$P \leq 0.0001$.

a marked reduction during G_1 (Fig. 3E). Similar results were obtained after culture synchronization using serum starvation (fig. S3A), suggesting that the function of this transcription repressor is needed particularly during DNA replication and subsequent cell division. This observation may explain the decreased levels of ZC3H4 observed during senescence (Fig. 2F), as senescent cells experience prolonged cell cycle arrest. To further elucidate the potential effect of ZC3H4 absence on cell cycle progression, we removed senescent cells from the ZC3H4-depleted population by applying a nocodazole pulse, followed by mitotic shake-off to enrich for mitotic cells. We then assessed EdU incorporation in this population, finding that it was significantly reduced in ZC3H4-deficient cells compared to ZC3H4-proficient cells, indicating a defect in cell cycle progression (fig. S3B). Together, these results suggest that ZC3H4 is essential for the timely progression of the cell cycle through the S, G_2, and M phases.

Despite ZC3H4's role as a transcription regulator (*29*), our bioinformatic and experimental results suggested that its loss could lead to RS. To explore this possibility, we used a DNA fiber assay, which revealed a significant reduction in the rate of replication fork progression—an RS hallmark—in ZC3H4-deficient cells compared to the parental controls (Fig. 3F). To determine whether fork progression is impeded by obstacles in ZC3H4-depleted cells, we analyzed fork symmetry. In WT cells, replication forks were symmetrical, with both arms exhibiting equal length (Fig. 3G, top). However, in ZC3H4-depleted cells, the forks were asymmetrical, showing poor correlation between arm lengths (Fig. 3G, bottom). These results suggest that ZC3H4 loss introduces obstacles that hinder fork progression, leading to asymmetry at replication origins.

Further supporting the presence of RS in ZC3H4-depleted cells, clonogenic survival assays indicated increased cellular sensitivity of depleted cells to the RS-inducing agent hydroxyurea (HU) (Fig. 3H). Consistent with this, immunostaining for phosphorylated Ser^{33} of Replication Protein A2 (RPA2) subunit of the Replication Protein Complex (RPA complex)—another RS marker—showed a sixfold increase in the number of phosphorylated RPA2 (pRPA2) foci per cell in ZC3H4-deficient cells compared to control cells (Fig. 3I). Notably, ZC3H4 depletion resulted in a basal level of pRPA2 foci higher than that in HU-treated WT cells (Fig. 3I). Furthermore, HU treatment in ZC3H4-deficient cells caused notably higher numbers of pRPA2 foci compared to WT cells (Fig. 3I), aligning with the enhanced HU sensitivity observed in these cells (Fig. 3H). We concluded that ZC3H4 loss induced replication fork stalling, triggering a cascade of RS, followed by slow cellular proliferation, mitotic defects, chromatin bridges, polynucleation, and, ultimately, prolonged cell cycle arrest and cellular senescence.

ZC3H4 loss triggers transcription-dependent RS linked to TRCs

We investigated the mechanism underlying RS in the absence of ZC3H4, focusing on its previously described role in ncRNA suppression (*25*, *26*, *28*, *29*, *32*–*35*). Given this role, we hypothesized that ZC3H4 loss induces a hypertranscription state, a known driver of RS (*48*). To test this, we measured RNA synthesis rates in ZC3H4-deficient and -proficient cells using the 5-ethynyl uridine (5-EU) click reaction and epifluorescence microscopy to track the incorporation of 5-EU into nascent RNA. To confirm the signal specificity of this signal for transcription, we treated cells with the transcriptional elongation inhibitor, 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB). DRB effectively reduced the 5-EU signal in both WT and ZC3H4-depleted cells compared to controls (Fig. 4A). We observed a marked increase in RNA synthesis, in ZC3H4-deficient cells compared to controls (Fig. 4A). To assess whether ectopic expression of WT ZC3H4 could rescue this phenotype, we overexpressed ZC3H4 in both WT and knockout (KO) cells. This significantly reduced RNA synthesis in both genotypes (Fig. 4A). For further quantification, we used multicolor super-resolution localization microscopy [stochastic optical reconstruction microscopy (STORM)] combined with automated data analysis (*15*, *49*). STORM enabled single-molecule detection of 5-EU–labeled nascent RNA, revealing a significant increase in RNA synthesis in the absence of ZC3H4 (Fig. 4B). Notably, this approach allowed us to detect extensive RNA synthesis beyond the nucleoli (Fig. 4B), a phenomenon often obscured in epifluorescence microscopy due to strong nucleolar signals. Consistently, STORM-based quantification of active RNA polymerase II (RNAP II), marked by Ser^2 phosphorylation, showed elevated transcriptional activity in ZC3H4-deficient cells compared to WT cells (fig. S4A).

Since ZC3H4 levels increase during S phase (Fig. 3E), we investigated whether the hypertranscription induced by its depletion occurs predominantly during this cell cycle phase. To assess this, we pulse-labeled the cells with 5-EU and costained for cyclin A, a marker of the S-G_2 phases, allowing us to analyze transcriptional activity in relation to cyclin A levels (fig. S4B). ZC3H4-deficient cells exhibited increased 5-EU incorporation across all cell cycle phases compared to WT cells, suggesting that ZC3H4 acts as a transcriptional repressor throughout the cell cycle. Nevertheless, its loss may exert more pronounced effects during DNA replication.

To establish a causal link between hypertranscription and RS in the absence of ZC3H4, we treated cells with the transcription inhibitor DRB and assessed replication dynamics using a DNA fiber assay (Fig. 4C). Notably, DRB treatment significantly restored the slowed replication fork progression in *ZC3H4*-null cells, indicating that hypertranscription contributes directly to RS in these cells. Consistent with previous reports (*50*), DRB treatment in WT cells led to a reduction in fork progression (Fig. 4C). In ZC3H4-depleted cells, DRB likely mitigates excessive transcription, allowing stalled replication forks to resume progression. Conversely, in WT cells, where transcription is already balanced, DRB may suppress essential transcription, thereby disrupting normal replication dynamics and other cellular processes.

Further supporting the role of transcription in RS induced by ZC3H4 loss, DRB treatment also rescued the fork asymmetry observed in ZC3H4-depleted cells (fig. S4C). This suggests that excessive transcription in the absence of ZC3H4 obstructs fork progression, reinforcing the direct involvement of ZC3H4 in preventing RS.

TRCs are exacerbated under hypertranscriptional conditions (*5*, *9*, *48*). To quantify TRCs in ZC3H4-proficient and -deficient cells, we used the proximity ligation assay (PLA) (*51*), using proliferating cell nuclear antigen (PCNA) as a replication marker and phosphorylated RNA polymerase II (RNAPOL2)-Ser^2 as an indicator of active transcription (*15*). PLA analysis revealed a significantly elevated PCNA-RNAPOL2-Ser^2 signal in *ZC3H4*-KO cells compared to WT cells (Fig. 4D), suggesting an increased frequency of TRCs. DRB treatment reduced the PLA signal, indicating a corresponding decrease in TRC events. To further validate this finding, we used STORM microscopy coupled with automated data analysis (*15*, *49*), enabling single-molecule visualization of replication-transcription associations (*49*, *52*–*55*). CAL-51 cells of both *ZC3H4* genotypes were pulsed with 5-EU to label nascent RNA and immunostained for the PCNA heterotrimer, a replication fork marker. Quantification of 5-EU–PCNA

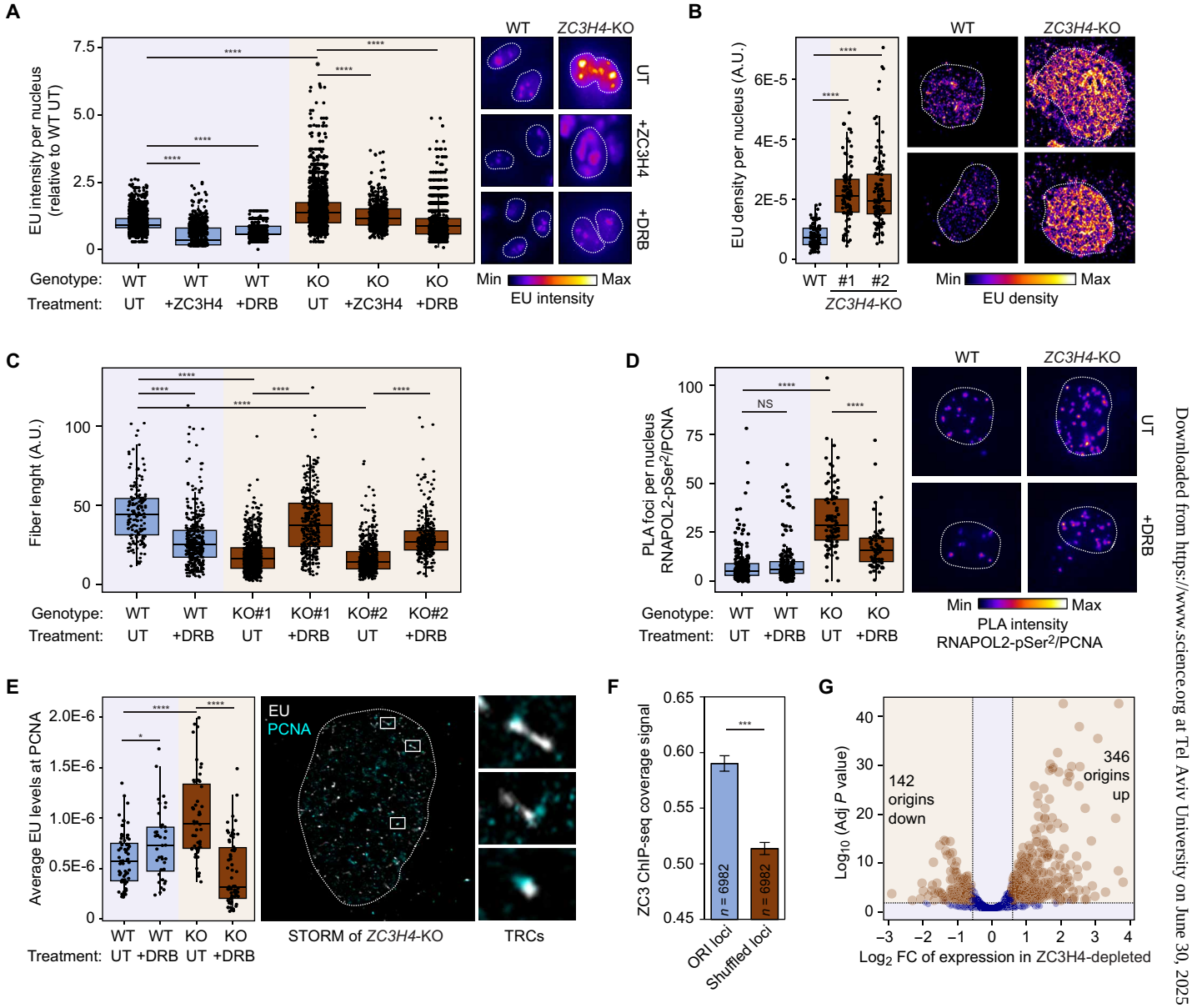

Fig. 4. RS caused by ZC3H4 absence is transcription-dependent. (**A**) 5-EU incorporation assay (30-min pulse) based on epifluorescence microscopy shows increased RNA synthesis in ZC3H4-deficient cells. Ectopic expression of ZC3H4 reduces RNA synthesis both in WT and *ZC3H4*$^{-/-}$ cells. DRB treatment significantly reduces the 5-EU signal. (**B**) STORM analysis of 5-EU incorporation (15-min pulse) demonstrates at super-resolution the increase in RNA synthesis upon ZC3H4 depletion. (**C**) DNA fiber analysis shows that DRB treatment alleviates the RS in ZC3H4-deficient cells. (**D**) Proximity ligation assay (PLA) demonstrates increased interaction between RNAPOL2-pSer^2 and PCNA in *ZC3H4*$^{-/-}$ cells. DRB treatment was used as a negative control. (**E**) STORM analysis shows that ZC3H4 loss increases the 5-EU density at PCNA sites, suggesting enhanced TRC rate, which is markedly reduced by DRB treatment. (**F**) ZC3H4 ChIP sequencing (ChIP-seq) coverage signal [accession ID: GSE207416 (*28*)] was quantified over 6982 origins of replication (ORI) [as defined in Liu *et al.* (*56*)] and compared to randomly shuffled loci to assess ZC3H4 binding specificity at these sites. (**G**) Volcano plot depicting fold change (FC) in RNA expression at individual replication origin loci upon ZC3H4 loss. A higher proportion of these loci exhibit RNA synthesis in *ZC3H4*$^{-/-}$ cells. Transcriptomic data were retrieved from a POINT-seq dataset [accession ID: GSE207417 (*28*)]. "ns" = $P > 0.05$; *$P \leq 0.05$; **$P \leq 0.01$; ***$P \leq 0.001$; ****$P \leq 0.0001$.

associations showed a marked increase in ZC3H4-deficient cells compared to WT cells (Fig. 4E), consistent with elevated TRC formation in the absence of ZC3H4. This trend was confirmed across two independent *ZC3H4*-null cell clones (fig. S4D). DRB treatment had no significant effect on TRC frequency in WT cells, likely due to their inherently low baseline levels in these cells. To further corroborate TRC elevation in ZC3H4-deficient cells, we pulsed the cells with EdU to label nascent DNA, followed by counterstaining for active RNAPOL2I and STORM imaging (fig. S4E). Together, these findings demonstrate that ZC3H4-deficient cells exhibit a markedly increased occurrence of TRCs—a major cause of RS.

ZC3H4 genomic associations support its role in preventing TRCs

We investigated whether the genomic distribution of ZC3H4 supports its emerging role in preventing unwanted encounters between

the replication and transcription machinery. Using a computational approach, we analyzed a high-quality chromatin immunoprecipitation (ChIP) dataset for ZC3H4, generated by Estell *et al.* (*28*) in K562 cells (raw data obtained from accession ID GSE207416). The data were preprocessed and aligned here to the human genome assembly GRCh37 (hg19) to match Okazaki fragment sequencing (Ok-seq) data (*56*).

First, we examined ZC3H4 enrichment at the 6982 origins of replication (ORCs) identified in K562 cells (*56*). Compared to randomly shuffled control sites, ZC3H4 ChIP sequencing (ChIP-seq) signals were significantly enriched at ORCs (Fig. 4F). Next, we integrated ZC3H4 genomic localizations with an Ok-seq dataset from the same cell line (*56*). The Ok-seq method quantitatively maps replication initiation and termination events (*57*). To assess ZC3H4 distribution relative to replication fork progression, we analyzed the replication fork direction (RFD) parameter (*56*). The genome was divided into 10-kb bins (~320,000 intervals), and RFD values were computed for each interval, with lower values indicating proximity to replication initiation sites and higher values reflecting proximity to termination sites. Ranking intervals by absolute RFD and grouping them into quartiles revealed a strong correlation between ZC3H4 occupancy and replication initiation zones, with the highest ZC3H4 signal detected in the quartile closest to ORCs (fig. S4F). These complementary analyses indicate that ZC3H4 preferentially binds genomic regions near ORCs, supporting its role in coordinating replication and transcription dynamics.

Given the above conclusions and ZC3H4's established role as a repressor of ncRNA synthesis (*25*, *26*, *28*, *32*–*35*), we investigated whether its depletion leads to ncRNA accumulation near ORCs, potentially contributing to TRCs at these critical sites. We computationally analyzed RNA synthesis levels around ORCs in ZC3H4-depleted cells using RNA-seq and Polymerase Intact Nascent Transcript RNA-seq (POINT-seq) datasets generated by Estell *et al.* (*28*) following degron-based depletion of ZC3H4. Raw RNA-seq data were preprocessed, and expression levels were specifically evaluated across the 6982 previously identified ORCs (*56*). We performed a differential expression analysis at ORC sites in ZC3H4-deficient cells, revealing a widespread increase in RNA synthesis at these regions. Notably, a significantly larger number of ORC sites exhibited elevated RNA levels upon ZC3H4 depletion in both POINT-seq ($n = 346$) and RNA-seq ($n = 750$) datasets (Fig. 4G and fig. S4G), whereas only a small subset showed reduced RNA production (POINT-seq, $n = 142$; RNA-seq, $n = 276$). These results indicate that, in most cases, ZC3H4 loss leads to increased RNA synthesis at ORC sites. Collectively, these findings suggest that ZC3H4 depletion promotes ncRNA accumulation at ORC regions, potentially contributing to TRC formation at these critical sites.

ZC3H4 loss–induced hypertranscription triggers R loop accumulation

Hypertranscription is expected to increase RNA-DNA hybrids (R loops) formation, which can act as physical barriers to replication forks and induce RS (*48*). Notably, the gene encoding ribonuclease H1 (RNase H1), which resolves R loops, exhibited strong coexpression with *ZC3H4* (fig. S5A), and ZC3H4 has been identified as an R loop interactor in a large-scale screen (*58*). To assess the contribution of R loops to RS in ZC3H4-deficient cells, we quantified R loops levels in WT and *ZC3H4*$^{-/-}$ cells using S9.6 immunostaining. We validated the specificity of the S9.6 antibody under our experimental conditions by overexpressing RNase H1 in WT cells (Fig. 5A). RNase H1 treatment markedly reduced the S9/6 signal, confirming R loops as its source. ZC3H4 depletion significantly increased R loop signals, while ectopic expression of ZC3H4 in *ZC3H4*$^{-/-}$ cells rescued the signal increase, reinforcing ZC3H4's role in preventing R loop accumulation (Fig. 5A). ZC3H4 overexpression strongly reduced R loops, raising the question of whether ZC3H4 directly degrades R loops or most likely primarily inhibits transcription (Fig. 4, A and B), thereby preventing R loop formation as a by-product. STORM imaging following nuclear staining with S9.6 confirmed an increase in R loops in ZC3H4-depleted cells (fig. S5B). The lack of detectable S9.6 signal in WT cells under STORM microscopy may result from experimental conditions, as permeabilization and stringent washing before fixation likely removed free-floating RNA-DNA hybrids and nonspecific RNA signals (*59*). This may also explain why RNase H1 treatment had no visible effect under STORM, contrasting with classical microscopy results (Fig. 5A). To evaluate whether R loops hinder replication fork progression, we used STORM imaging to quantify S9.6 signal density at replicating sites, labeled via a 10-min EdU pulse. ZC3H4 loss led to a notable accumulation of R loops localized at replication forks (Fig. 5B), demonstrating their direct impact on replication.

Further supporting our findings, we analyzed the correlation between ZC3H4 occupancy and R loop abundance across the genome. Using DNA-RNA immunoprecipitation sequencing (DRIPc-seq) data from Bayona-Feliu *et al.* (*60*) in K562 cells (GSE154631), we quantified R loop levels and compared them to ZC3H4 ChIP-seq data (GSE207416) (*28*). The human genome was divided into 10,000-nucleotide bins (~320,000 intervals), and R loop abundance (DRIPc-seq signal intensity) was calculated for each interval. Ranking all intervals by DRIPc-seq coverage and grouping them into quartiles revealed a strong correlation between ZC3H4 binding and R loop abundance (Fig. 5C). These findings suggest that ZC3H4 preferentially associates with genomic regions prone to R loop formation.

To further investigate the impact of ZC3H4 depletion on the expression of R loop–forming transcripts, we analyzed the raw RNA-seq and POINT-seq data from Estell *et al.* (*28*), focusing on gene expression changes at R loop–prone sites in ZC3H4-depleted K562 cells. We reprocessed the raw FASTQ data, performed alignment to the human genome (hg38), and quantified expression levels. Instead of focusing on protein-coding genes and promoter regions, we specifically targeted R loop–prone regions, as identified in a recent meta-analysis database (*61*). ZC3H4 depletion led to the up-regulation of expression at these R loop sites in the RNA-seq dataset (Fig. 5, D and E), which was further validated by the POINT-seq data (fig. S5C). Ectopic expression of ZC3H4 in ZC3H4-depleted cells successfully restored the repression of R loop formation at these sites (Fig. 6E).

To extend this observation to other cell lines, we analyzed publicly available transcriptomic data from ZC3H4-depleted cells (*25*, *26*, *29*, *32*–*34*) and assessed expression changes at R loop–prone sites. Consistently, we observed increased expression at R loops following ZC3H4 depletion in HCT116 (GSE163015; fig. S5D), HEK-293 (GSE186809; fig. S5E), and HeLa cells (GSE212208 and GSE133109; fig. S5, F and G, respectively). Notably, similar expression bursts at R loop sites were also observed in ZC3H4-depleted mouse embryonic stem cells (GSE199805; fig. S5H). For mice, R loop–prone sites were defined using RLBase (*62*), based on preprocessed DNA:RNA hybrid immunoprecipitation sequencing (DRIP-seq) data from mouse embryonic stem cells (SRX5290925), and expression at these sites

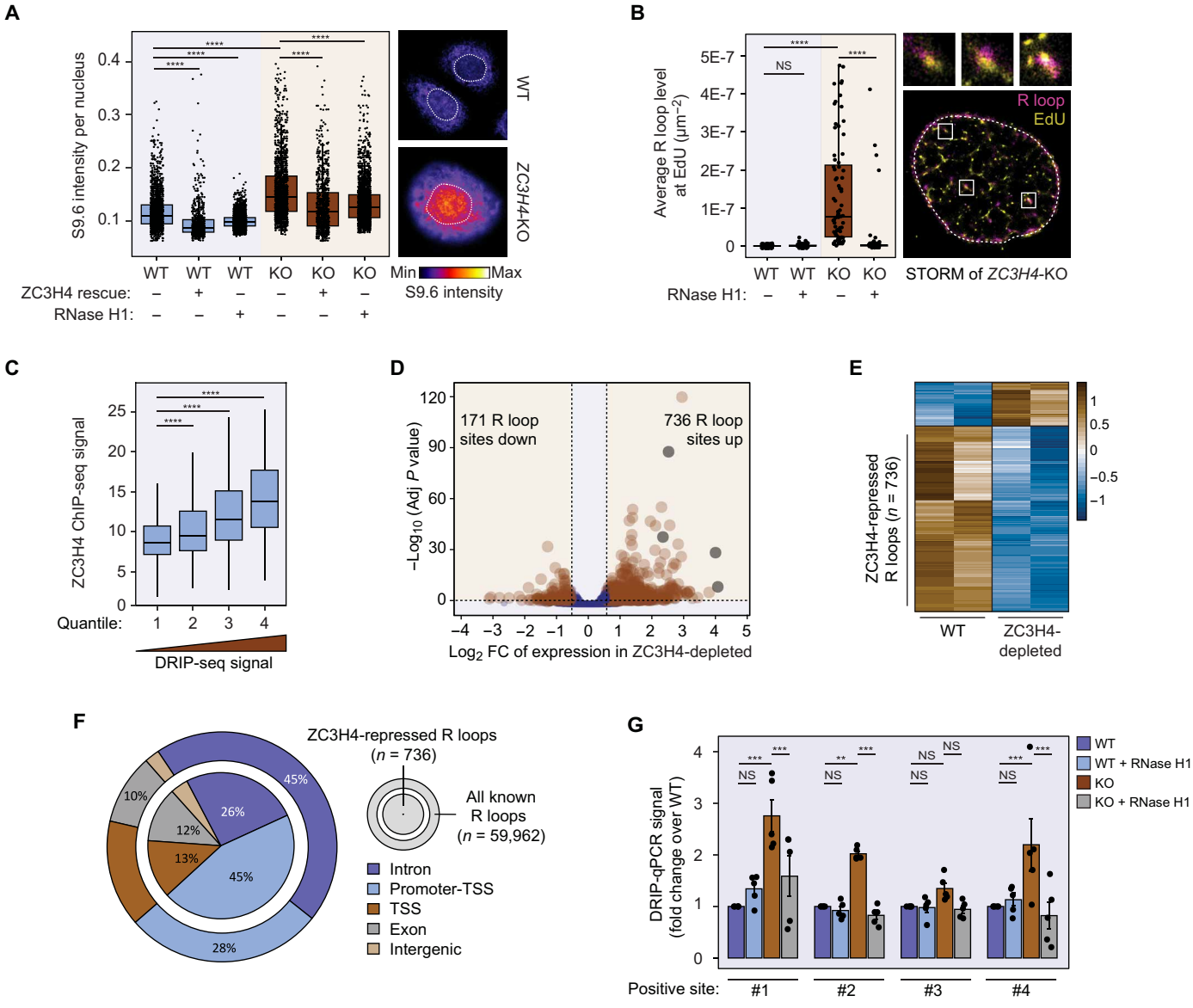

Fig. 5. R loops link the hypertranscription and the TRCs induced by ZC3H4 loss. (**A**) R loops were quantified using immunostaining with the S9.6 antibody and epifluorescence microscopy. RNase H1 treatment significantly reduced the S9.6 signal, confirming its specificity for DNA:RNA hybrids. ZC3H4 loss markedly increased the R loop signal, while ectopic expression of the protein reduced R loops in both WT and *ZC3H4*$^{-/-}$ cells. (**B**) Quantification of R loop density around PCNA foci using STORM. ZC3H4 deficiency induced a notable elevation of R loop signal around replicating sites, which was abolished by RNase H1. (**C**) ZC3H4 ChIP-seq coverage signal quantified across 10,000-nucleotide genomic segments distributed by their average DRIPc-seq signal [accession ID: GSE154631 (*60*)]. ZC3H4 signal shows a positive correlation with DRIPc-seq, suggesting preferential binding at potential R loop sites. (**D**) Volcano plot of RNA-seq data [accession ID: GSE207421 (*28*)] reveals a higher proportion of expressed R loop–prone sites (*61*) upon ZC3H4 loss. (**E**) Expression heatmap of significantly deregulated R loop sites in ZC3H4-depleted cells. A total of 736 up-regulated sites were found to be induced upon ZC3H4 loss. (**F**) A pie chart illustrating the distribution of R loop location in general [outside circle; data obtained from (*61*)] compared to those induced by ZC3H4 loss (inner circle), relative to gene features. A majority of these sites are found at gene promoters and TSSs. (**G**) DRIP-qPCR validates R loop increase in ZC3H4-deficient cells. Four R loop–prone sites with strong expression bursts in *ZC3H4*$^{-/-}$ K562 cells [(D), black circles] were selected. ZC3H4 loss in CAL-51 cells led to a strong increase in DRIP signal at these sites, confirming R loop induction. Signal specificity was validated by RNase H1 treatment, which significantly reduced the DRIP signal. "ns" = $P > 0.05$; *$P \leq 0.05$. **$P \leq 0.01$; ***$P \leq 0.001$; ****$P \leq 0.0001$.

was then assessed after ZC3H4 depletion. These analyses support the notion that ZC3H4-mediated repression of R loop–associated expression is a conserved phenomenon across species.

Further analysis of these R loop sites exhibiting elevated expression in ZC3H4-depleted cells revealed a distinct distribution: 58% were located at promoters or transcription start sites (TSSs), 38% within gene bodies, and 4% in intergenic regions (Fig. 5F, inner circle). Notably, ZC3H4-repressed R loops were particularly enriched at promoter-TSS regions, accounting for 45% of these sites, compared to 28% in the general R loop distribution (Fig. 5F, outer circle). To further characterize the transcriptional effect of ZC3H4 loss at R loop–prone sites, we performed a metagene analysis of expression

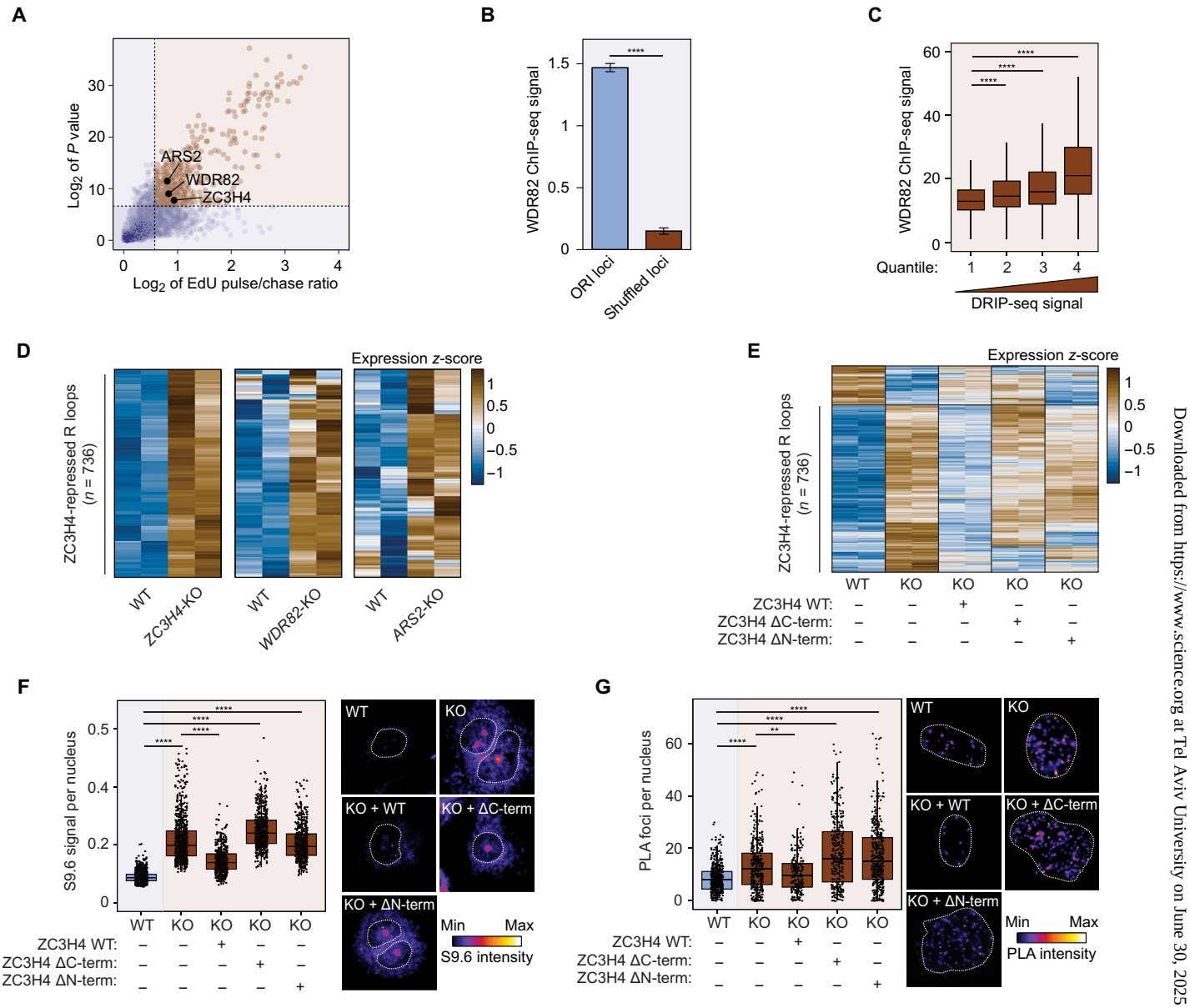

Fig. 6. ZC3H4's role in maintaining genome stability is carried out within the ZC3H4/Restrictor context. (**A**) Scatter plot showing significantly enriched proteins at nascent DNA, identified by iPOND (*64*) . Each point represents a protein. *x* axis: $\log_2$ fold change of the EdU pulse/chase ratio, where higher values indicate greater enrichment. *y* axis: -$\log_2$ *P* value, representing statistical significance. Restrictor complex members, ZC3H4, ARS2, and WDR82, are highlighted. (**B**) WDR82 ChIP-seq coverage signal [accession ID: GSE207416 (*28*)] quantified across 6982 ORCs in K562 cells (*56*) compared to randomly shuffled loci. (**C**) WDR82 ChIP-seq coverage signal [accession ID: GSE207416 (*28*)] quantified across 10,000-nucleotide genomic segments sorted by DRIPc-seq signal intensity [accession ID: GSE154631 (*60*)]. A positive correlation is observed between the WDR82 and the DRIPc-seq signals, suggesting preferential WDR82 binding at R loop–prone sites. (**D**) Expression heatmap at ZC3H4-repressed R loop sites, showing that depletion of WDR82 or ARS2 phenocopies ZC3H4 loss, with similar transcriptional up-regulation [accession ID: GSE207421 (*28*)]. (**E**) Expression heatmap at ZC3H4-repressed R loop sites [accession ID: GSE207421 (*28*)] reveals that ectopic expression of ZC3H4 in *ZC3H4*$^{-/-}$ cells represses the R loop burst, while truncated ZC3H4 mutants lacking the binding domains for WDR82 (d82-ZC3) or ARS2 (ARMS-ZC3) fail to rescue these bursts. These findings suggest that ZC3H4 interactions with WDR82 and ARS2 are necessary for suppressing the R loop bursts. (**F**) Elevated S9.6 signal in *ZC3H4*$^{-/-}$ is reduced upon ectopic expression of WT ZC3H4, while truncated versions of the protein missing the WDR82 or the ARS2 binding domains fail to do so. (**G**) PLA shows increased interaction between RNAPOL2-pSer^2 and PCNA upon ZC3H4 loss. Ectopic expression of WT ZC3H4 decreases the PLA signal, while truncated versions of the protein missing the WDR82 or the ARS2 binding domains fail to do so. "ns" = $P > 0.05$; *$P \leq 0.05$; **$P \leq 0.01$; ***$P \leq 0.001$; ****$P \leq 0.0001$.

changes spanning entire genes, from upstream of the TSS to the transcription end site (TES). R loop–prone regions were defined using publicly available DRIP-seq data in K562 cells (RLBase, SRX8122753), and transcriptional output was assessed using POINT-seq profiles following ZC3H4 depletion (GSE207421). This analysis revealed widespread transcriptional up-regulation across gene bodies, with the most prominent increase upstream of TSSs that extended well beyond canonical promoter boundaries (fig. S5I). These upstream signals are indicative of elevated expression of PROMPTs (*63*), consistent with ZC3H4's established role in repressing these ncRNAs.

These findings underscore ZC3H4's critical role in transcriptional regulation at these genomic loci. This pattern aligns with ZC3H4's

established genomic association and functional involvement in repressing PROMPT loci (*25*, *26*, *28*, *32*–*35*). Collectively, our data strongly support that ZC3H4's transcriptional repressor function is essential for mitigating a key RS driver—R loops—particularly around TSSs.

To experimentally validate the R loop induction at promoter-TSS sites upon ZC3H4 depletion and assess whether this phenomenon is conserved in our CAL-51 system, we performed DRIP–quantitative polymerase chain reaction (qPCR). We selected four R loop prone sites that exhibited strong expression increase in ZC3H4-depleted K562 cells (Fig. 5D, black circles, and table S3). Consistently, ZC3H4 depletion in CAL-51 cells led to a significant increase in DRIP signal at all tested sites (Fig. 5G), confirming R loop accumulation. Signal specificity was validated by RNase H1 treatment, which markedly reduced the DRIP signal. As a negative control, we performed DRIP-qPCR at genomic regions not prone to R loop formation (fig. S5J), where no signal increase was detected, further reinforcing the specificity of R loop detection at the selected positive sites.

To determine whether ZC3H4 modulates R loop formation specifically during S phase, we performed S9.6 immunostaining in combination with cell cycle markers: cyclin A (a marker of S-G_2) and a 15-min EdU pulse (to mark S phase). ZC3H4-depleted cells exhibited elevated S9.6 signals across all cell cycle phases compared to controls (fig. S5K), indicating that ZC3H4 influences R loop formation broadly, rather than being restricted to replication. Nevertheless, the accumulation of R loops may have more pronounced consequences during S phase, when TRCs are more likely to occur.

Our findings establish a direct link between ZC3H4 depletion and R loop accumulation, with a notable enrichment at promoter-TSS regions. This induction occurs independently of the cell cycle phase, suggesting that ZC3H4 plays a broad regulatory role rather than a replication-specific function. However, the accumulation of these R loops can create physical barriers to replication forks, increasing the likelihood of TRCs. These results underscore ZC3H4's critical role in preventing R loop–driven RS and safeguarding genome integrity.

ZC3H4 safeguards genome stability as a component of the restrictor complex

Given ZC3H4's involvement in restricting ncRNA expression as part of the Restrictor complex (*25*, *26*, *28*, *32*–*35*), it is plausible that its role in preventing RS operates within this complex. To investigate whether Restrictor is present at replication sites, we first queried a database of proteins associated with replicating DNA, identified using the isolation of proteins on nascent DNA (iPOND) method (*64*). Notably, all three Restrictor components were included in this list (Fig. 6A). Using the same approach that identified ZC3H4 around ORC sites, we analyzed ChIP-seq data for WDR82, a protein closely linked to ZC3H4 (*29*), and confirmed its enrichment at replication sites (Fig. 6B). Furthermore, we observed a similar correlation between WDR82 ChIP-seq signals and R loop abundance across the genome (Fig. 6C), mirroring the results seen with ZC3H4 and DRIPc-seq data (Fig. 5C).

Both ARS2 and WDR82 have been previously associated with R loops in a large screening effort, similarly to ZC3H4 (*58*). To assess whether depletion of ARS2 and WDR82 affected the expression of R loop–generating sites in a manner similar to ZC3H4 depletion, we analyzed transcriptomic data from cells depleted of these proteins (*28*). Loss of WDR82 and ARS2 resulted in an increased RNA-seq signal at ZC3H4-associated R loop sites (Fig. 6D). In addition, in the same study, ZC3H4-deficient cells were complemented with ectopic expression of truncated versions of ZC3H4, lacking its WDR82 and ARS2/NEXT binding sites. Reanalysis of this data revealed that while the WT version of ZC3H4 effectively restored the repression of RNA production at R loop–producing sites, the mutant versions failed to do so (Fig. 6E). These findings suggest that ZC3H4's ability to bind to its Restrictor partners is essential for its role in suppressing RNA production and preventing the formation of corresponding R loops.

To experimentally validate the role of the Restrictor complex in R loops and TRCs, we complemented ZC3H4-depleted cells with either a full WT ZC3H4 version or truncated versions lacking interactions with ARS2 (ΔN-term) or WDR82 (ΔC-term) (*28*). The expression of the plasmids was validated by Western blotting (fig. S5L). First, we quantified the R loop signal in complemented cells using S9.6 staining (Fig. 6F). While rescuing with the WT ZC3H4 version restored R loop repression, the truncated versions failed to reduce the S9.6 signal. These results support the requirement for the full Restrictor complex in R loop reduction. In addition, these findings align with the bioinformatics data presented in Fig. 6E, supporting the use of RNA-seq expression at R loop sites as a reliable proxy for studying R loop induction.

Next, we measured collision events in the complemented cells using the PLA (Fig. 6G). Rescue with WT ZC3H4 significantly reduced TRC events, whereas the truncated versions failed to reduce these events. This further reinforces the role of the Restrictor complex in mediating TRCs.

Collectively, these findings establish the Restrictor complex as a key safeguard of genome stability. Through integrated bioinformatic analyses and experimental validation—including RNA-seq, ChIP-seq, and functional assays—we demonstrate that the full Restrictor complex is required to restrict ncRNA production at R loop–prone sites and prevent TRCs, thereby preserving replication integrity (Fig. 7).

DISCUSSION

In our ongoing search for novel connections between genome instability, cellular senescence, and aging (*40*, *65*, *66*), we conducted computational analyses to identify previously unknown genes at the intersection of these processes. Through this approach, ZC3H4 emerged as a potential player in this link. Simultaneously, ZC3H4's role as a suppressor of pervasive production of ncRNA was uncovered through meticulous work in other laboratories (*25*, *26*, *28*–*30*, *32*, *33*), while our data revealed a previously unknown role of this protein in preventing RS. This led us to hypothesize that ZC3H4 may directly function to prevent the detrimental physical encounters of transcription and replication—TRCs, which are a major source of genome instability (*9*, *10*, *67*). Our experimental and computational results indeed support this hypothesis, showing that ZC3H4/Restrictor is strategically positioned to prevent TRC-driven RS by repressing ncRNA synthesis and subsequent R loop formation. The dynamic regulation of ZC3H4 across the cell cycle highlights its critical role during S phase. ZC3H4 and the Restrictor complex have not previously been linked to RS resolution, highlighting their previously undescribed role in maintaining genome stability. These findings provide an explanation for the genome instability observed in *Zc3h4*$^{-/-}$ mouse embryos (*35*).

The presence of R loops at sites of TRCs triggered by ZC3H4 depletion is noteworthy. R loop abundance increases in hypertranscription states (*13*, *48*), and their role in contributing to RS by acting as barriers to replication fork progression is well documented (*11*–*13*, *18*, *48*, *68*). We observed elevated R loop signals along with

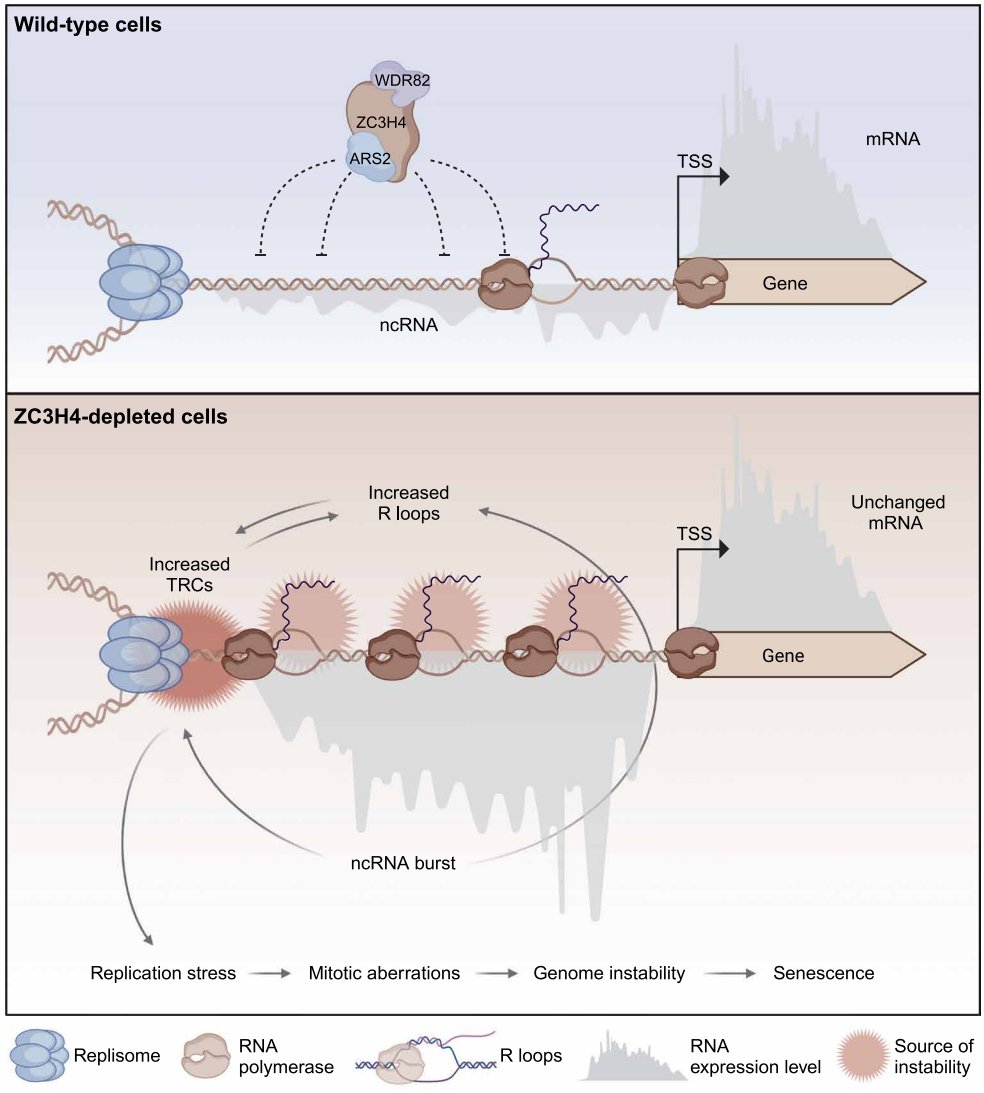

Fig. 7. Model of ZC3H4/Restrictor in maintenance of genome stability. In WT cells, ZC3H4/Restrictor represses ncRNA expression at promoter sites, facilitating proper replisome progression. In the absence of ZC3H4, uncontrolled ncRNA accumulation at its target sites leads to a hypertranscription state. This increases the likelihood of TRCs, triggering RS, genome instability, and, ultimately, cell senescence. Created in BioRender, M.H. (2025; https://BioRender.com/79gclaf).

ncRNA bursts in ZC3H4-depleted cells, and ZC3H4 was previously shown to be physically associated with R loops in an R loop proximity proteomics study (*58*).

Several mechanisms have been described for preventing TRCs. Chromatin-based regulation, such as H3K4 methylation, mitigates TRCs by modulating replication fork speed (*69*). Transcription termination factors, including Sen1, remove stalled RNA polymerase to prevent collisions with the replication machinery (*70*). R loop resolution mechanisms, such as those based on RNase H or the FA Complementation Group J (FANCJ) helicase, degrade RNA:DNA hybrids that obstruct fork progression (*68*), while RNA surveillance complexes such as TRAMP (Trf4/Air2/Mtr4p Polyadenylation complex) prevent stalled transcripts from obstructing replication (*71*). Depletion of SPT6, an RNAP II–associated elongation factor, leads to excessive long ncRNA production, promoting R loop formation, RS, and senescence (*72*). The Integrator complex, which attenuates transcription through direct interaction with RNAP II, has been implicated in mitigating TRCs (*22*–*24*). Another mechanism involves MYCN-dependent recruitment of the RNA-degrading exosome complex to transcribed regions, preserving genome stability (*73*). In addition, PNUTS and WDR82 have been shown to limit RS by promoting RNAP II degradation on chromatin, thereby preventing TRCs (*31*).

Our findings uncover another player in TRC prevention, involving the Restrictor complex (ZC3H4-WDR82-ARS2), which inhibits the early elongation of pervasive ncRNA transcripts, likely leading to their subsequent degradation (*29*, *30*). This prevents excessive ncRNA accumulation and subsequent stalling of replication forks caused by these roadblocks. This mechanism operates parallel to TRC resolution at gene bodies, with both mechanisms contributing to the maintenance of smooth DNA replication.

Many TRC studies have concentrated on gene bodies, active transcription units, or ribosomal DNA (rDNA) as primary conflict sites (*9*). Our study points to a regulatory layer at promoter regions, where

pervasive ncRNA production can hinder replication fork progression. Notably, ZC3H4 recruitment at pre-TSS regions is consistent with ZC3H4's role in repressing promoter-associated RNAs (*29*).

WDR82 appears to mediate TRCs through distinct mechanisms, depending on its interacting partners. Notably, it forms mutually exclusive complexes with either PNUTS or ZC3H4 (*29*), suggesting that the WDR82-PNUTS complex mediates TRCs independently of ZC3H4. Furthermore, the WDR82-PNUTS complex has been implicated in a ZC3H4-independent role in the regulation of pervasive RNA (*29*). This indicates that WDR82 may exert different regulatory functions depending on its binding partner. In this study, we specifically investigated WDR82 in the context of its interaction with ZC3H4. To achieve this, we used truncated ZC3H4 variants incapable of binding WDR82, enabling us to isolate and examine WDR82's function within the ZC3H4 complex while minimizing potential interference from its PNUTS-associated activity. Collectively, previous data and our findings suggest that WDR82 mediates TRCs through two distinct mechanisms: (i) When interacting with PNUTS, it facilitates RNAPOL2 degradation at gene bodies; and (ii) when interacting with ZC3H4 and ARS2, it suppresses ncRNA at promoters and R loop–prone sites via exosome-mediated RNA degradation (*25*, *29*, *33*). Moreover, WDR82 was even shown to exert an opposing effect to the ZC3H4-WDR82 complex when interacting with SET Domain Containing 1A (SET1), forming a WDR82-SET1 complex (*34*). In this context, WDR82-SET1 promotes the expression genes with low to moderate expression by counteracting the restrictor complex (*29*, *34*). In addition, WDR82 further enhances transcription by facilitating H3K4 methylation, a posttranslational modification associated with actively transcribed genes (*74*). Together, these findings suggest that WDR82 may have a more context-dependent and multifaceted role in the regulation of pervasive RNA and TRCs than ZC3H4.

An intriguing question concerns the specificity of ZC3H4/Restrictor in suppressing ncRNA synthesis (*29*). One proposed explanation is that this specificity arises from the low prevalence of CpG islands in extragenic sequences, which reduces the activity of CpG island–associated mechanisms that normally protect nascent RNAs from premature transcription termination (*34*). Another hypothesis suggests that U1 small nuclear RNA serves as a protective shield, preventing ZC3H4/Restrictor from targeting protein-coding genes (*28*). Notably, a key feature of pervasive ncRNA is its lack of intron-exon structure, which prevents spliceosome recruitment—an important factor in reducing R loop formation (*75*). Furthermore, it was recently shown that ZC3H4/Restrictor promiscuously suppresses early elongation by RNAPOL2. However, its activity is largely blocked at most mRNAs by the presence of the 5′ splice site, leaving ncRNAs more susceptible to this suppression (*30*). Thus, with regard to R loops and TRCs, ZC3H4/Restrictor's role is preventive, mitigating ncRNA bursts before they occur and cause R loop formation and the ensuing TRCs. ZC3H4 may be recruited to ORC sites to facilitate origin firing and prevent obstruction of this process by transcription. However, this enrichment must be interpreted with caution, as replication origins frequently coincide with promoters near active genes (*76*). Thus, ZC3H4 and WDR82 presence at these sites may reflect a correlation with transcriptionally active regions rather than direct recruitment to replication origins. Distinguishing between these models requires further experimental validation. Future studies tracking ZC3H4 dynamics during active replication under hypertranscription conditions could help determine whether its role in TRCs is direct or a secondary consequence of transcriptional dysregulation.

Similar to many genes encoding proteins involved in genome stability, *ZC3H4* is mutated in certain cancers. RS is a hallmark of cancer cells, driven primarily by oncogene activation (*77*–*79*), and is likely exacerbated by loss or inactivation of ZC3H4. The immediate consequence of ZC3H4 depletion is a hypertranscription state, a feature of some aggressive cancers that is closely linked to RS (*48*, *80*). Notably, TRC-prone sites have been associated with an increased frequency of somatic mutations in cancer cells (*81*). In this context, we found that ZC3H4 loss leads to the formation of polynucleated, senescent cells. These cells have been observed to escape senescence more rapidly than their single-nucleus counterparts (*82*, *83*). Notably, escape from therapy-induced senescence is a mechanism of tumor recurrence after remission (*82*, *84*).

From a technical perspective, our findings suggest that RNA-seq expression at R loop–prone loci can serve as a reliable proxy for identifying R loop repressors. We experimentally validated this approach using DRIP-qPCR, S9.6 staining, and STORM imaging, confirming that increased transcription at these sites correlates with R loop accumulation. This strong association underscores the potential of transcriptomic data as an indirect yet powerful tool for studying R loop regulation. Given the widespread availability of RNA-seq datasets across various conditions and cell types, this method could be used for in silico screening to identify novel R loop modulators. By systematically analyzing gene expression changes at known R loop sites, researchers could find previously undescribed candidates involved in R loop suppression, providing a foundation for future functional validation studies.

Our work highlights the previously underappreciated role of pervasive transcription in genome instability. These findings highlight the necessity of regulating pervasive RNA before RS arises, rather than solely resolving conflicts postoccurrence. Beyond its role in preventing TRCs and RS, our coexpression analysis and GSEA, along with ZC3H4's high score in our DDR meta-analysis, suggest that this protein may function within the broader system responsible for maintaining genome stability, likely interacting with additional partners. Hence, ZC3H4 may play an increasingly prominent role in the expanding networks that safeguard genome integrity.

MATERIALS AND METHODS

Experimental protocols

Cell culture, synchronization, and treatments with irradiation and chemicals

CAL-51 and HEK-293 cell lines, along with primary human skin fibroblast lines, were cultured in complete Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum at 37°C in a 5% CO_2 atmosphere. Irradiation-induced senescence (IRIS) was induced by exposing the cultures to 10 Gy of IR using an irradiation cabinet (RX-650, Faxitron Bioptics, Lincolnshire, IL, USA). Following irradiation, the cells were incubated for 14 days with medium replacement every 4 days. On day 14 postirradiation, cells were deemed senescent on the basis of the absence of proliferation and positive SA-β-Gal staining.

Cell cycle synchronization of human primary fibroblasts was achieved using either a double-thymidine block or serum starvation. For the double-thymidine block, ~2×10^5 per dish were seeded in 6-cm dishes and treated with 2 mM thymidine (Sigma-Aldrich,

St. Louis, MO, USA) for 18 hours. The cells were then released for 9 hours and retreated with thymidine for an additional 16 hours before harvesting aliquots every 2 hours over a 24-hour period. For serum starvation, the cells were cultured in medium containing 0.1% fetal bovine serum for 48 hours, followed by serum replenishment and subsequent harvesting every 2 hours over a 28-hour period. For complementation assays using truncated versions of ZC3H4, CAL-51 cells were grown to 70% confluency, trypsinized, resuspended in medium, and pelleted by centrifugation at 100*g* for 5 min. The pellet was resuspended in phosphate-buffered saline (PBS) and centrifuged again at 100*g* for 5 min, and the supernatant was removed. The cell pellet was resuspended in 20 μl of SE Cell Line 4D-Nucleofector buffer (Lonza, Basel, Switzerland) and ~5 μg of plasmid DNA along with 0.6 μl of Alt-R Cas9 Electroporation Enhancer V2 [Integrated DNA Technologies (IDT), Coralville, IA, USA] were added to ~300,000 cells. The mixture was transferred into a well of a 16-well Nucleocuvette (Lonza) and nucleofected using the 4D-Nucleofector System (Lonza), with the nucleofection program EO-100 for CAL-51 cells. Following nucleofection, 100 μl of warm medium was added to the cuvette, and the cells were incubated for 5 min before being plated in 12-well plates. After 2 days, Tetracycline-controlled transcriptional activation (TET-ON) plasmid expression was induced by adding doxycycline (1 μg/ml) for 24 hours, before phenotypic analysis.

ZC3H4 gene ablation

The ribonucleoprotein (RNP) nucleofection method was used for *ZC3H4* gene ablation. Single-guide RNAs (sgRNAs) were designed using the Knockout Guide Design tool (Synthego, Redwood City, CA, USA) and obtained along with Alt-R Cas9 RNPs from IDT (Coralville, IA, USA). CAL-51, HEK-293 cells, or human primary skin fibroblast lines were grown to 70% confluency, trypsinized, resuspended in medium, spun down for 5 min at 100*g*, and resuspended in PBS. Approximately 300,000 cells were used per nucleofection. The cells were spun down again for 5 min at 100*g*, while the RNP complex was formed by mixing 1 μl of Alt-R S.p. Cas9 Nuclease V3 (IDT) with 0.7 μl of 100 μM Alt-R CRISPR-Cas9 sgRNA in 5 μl of RNase-free water. The mixture was incubated for 10 min at room temperature (RT). After spinning, the PBS supernatant was removed, and the cell pellet was resuspended in 15 μl of SE Cell Line 4D-Nucleofector buffer (Lonza, Basel, Switzerland) for CAL-51 and HEK-293 cells and P3 Primary Cell 4D-Nucleofector buffer (Lonza) for primary skin fibroblasts. Approximately 6 μl of the RNP mix (sgRNA + Cas9) and 0.6 μl of Alt-R Cas9 Electroporation Enhancer V2 (IDT) were added. The mixture was transferred into a well of a 16-well Nucleocuvette (Lonza) and loaded onto the 4D-Nucleofector System (Lonza). The nucleofection programs used were EO-100 for CAL-51 cells, EN-400 for primary fibroblasts, and EN-100 for HEK-293 cells. Following nucleofection, 100 μl of warm medium was added to each cuvette, and the cells were incubated for 10 min before being plated in 12-well plates. After 5 days, the cells were replated in 96-well plates at an approximate dilution of one cell per well. Two weeks later, the resulting clones were transferred to 24-well plates in duplicates. DNA was extracted from one duplicate using the GenElute Mammalian Genomic DNA Miniprep Kit (Sigma-Aldrich). PCR was conducted to obtain products spanning 500 base pairs (bp) around the edited sites. The PCR products were sequenced, and the resulting Sanger trace data file was analyzed using the ICE CRISPR Analysis Tool (Synthego). Cell clones found to be homozygous for ZC3H4 inactive alleles were expanded, and ZC3H4 loss was verified using Western blotting analysis.

WST-8 cell viability assay

The WST-8 assay kit (Dojindo Molecular Technologies, Rockville, MD, USA) was used to assess cellular viability. Cells were resuspended in 100 μl of medium and seeded at a density of 1000 cells per well in 96-well plates. Samples were taken daily, in triplicate. Ten microliters of WST-8 reagent was added to each well, and the mixture was incubated for 4 hours at 37°C. The optical density of the mixture was measured at 460 nm using the Multiskan SkyHigh plate reader (Dutscher, Brumath, France). The average of the triplicates was calculated.

Measurement of cellular proliferation rate by EdU incorporation

EdU was added to the culture medium at a final concentration of 10 μM, and the cultures were incubated for 12 hours. The cultures were then washed three times with PBS, fixed with 4% paraformaldehyde (PFA) for 15 min, permeabilized in PBS containing 0.5% Triton X-100 for 20 min, washed twice with PBS, and blocked with PBS supplemented with 5% bovine serum albumin (BSA) and 1% normal donkey serum (NDS) for 1 hour. Incorporated EdU was visualized using the Click-iT Plus EdU Alexa Fluor 488 (AF488) Imaging Kit (Thermo Fisher Scientific, Waltham, MA, USA).

Clonogenic survival assay

Clonogenic survival was measured after treatment with HU (Sigma-Aldrich). Various amounts of cells (100 to 2000) were seeded in 6-cm plates, according to the increasing treatment doses. One day after seeding, HU was added at different concentrations, and the cultures were allowed to grow for 14 days. The resulting colonies were stained with 0.5% crystal violet in 30% ethanol for 30 min at RT. The staining solution was then washed off, and the colonies were imaged and counted using the FUSION FX imaging system (Vilber Lourmat, Eberhardzell, Germany).

SA-β-Gal imaging

A total of 5×10^4 cells were seeded in 12-well plates precoated with 0.2% gelatin and incubated for 24 hours. The cells were then fixed in 0.5% glutaraldehyde for 20 min at RT and washed three times with PBS (pH 6.0). A β-Gal staining solution was applied, consisting of X-Gal (1 mg/ml; Thermo Fisher Scientific), 5 mM potassium ferrocyanide (Sigma-Aldrich), 5 mM potassium ferricyanide (Sigma-Aldrich), and 1 mM $MgCl_2$ (Invitrogen) in PBS (pH 6.0). The plates were tightly sealed with Parafilm and incubated at 37°C in the absence of CO_2. The reaction was terminated by washing the cells twice with PBS (pH 7.4). Notably, the incubation time depended on the cell type: 3 hours for CAL-51 cells, 4 hours for HEK-293 cells, and 6 hours for primary skin fibroblasts. The cells were then incubated for 1 min in dimethyl sulfoxide to remove stain crystals and washed with PBS (pH 7.4). Bright-field images of random fields were captured using an inverted Eclipse Ti-S microscope (Nikon, Tokyo, Japan). SA-β-Gal–positive cells were scored using the ImageJ/Fiji software [National Institutes of Health (NIH), Bethesda, MD, USA].

Western blotting analysis and antibodies

A total of 5×10^5 to 10×10^5 cells were lysed in radioimmunoprecipitation assay buffer [50 mM tris-HCl (pH 7.4), 150 mM sodium chloride, 1 mM EDTA, 1% NP-40, 1% sodium deoxycholic acid, and 0.1% SDS] supplemented with a phosphatase inhibitor mix (Sigma-Aldrich), a protease inhibitor cocktail (Roche, Basel, Switzerland) and benzonase (Merck Millipore, Darmstadt, Germany). The lysates were centrifuged at 20,000*g* for 20 min at 4°C. The supernatant was transferred into a new test tube, and protein concentration was

determined using the Bradford Protein Quantification Assay (Bio-Rad, Hercules, CA, USA). The samples were denaturated in 3× loading sample buffer and boiled to 97°C for 10 min. Protein samples were resolved by SDS–polyacrylamide gel electrophoresis and then transferred onto a methanol-preactivated polyvinylidene difluoride membrane in ice-cold transfer buffer. The proteins were separated at 400 mA for 4 hours or 200 mA overnight at 4°C. The membranes were rinsed with double-distilled water, reactivated using 100% ethanol, and blocked with 5% BSA blocking buffer for 1 hour at RT. They were then incubated overnight in a solution of primary antibody in dilution buffer (1% BSA and 0.1% sodium azide) at 4°C. After washing three times for 5 min each with Tween-Tris buffered saline (TTBS), the membranes were incubated with species-appropriate horseradish peroxidase–conjugated secondary antibodies diluted in TTBS containing 5% nonfat skim milk for 1 hour at RT. This was followed by three washes with TTBS (5 min each). Immunoreactive proteins were detected using enhanced chemiluminescence and the FUSION FX imaging system (Vilber Lourmat, Eberhardzell, Germany). Immunoblot staining was performed using the antibodies against the following proteins in tris-buffered saline with Tween 20 containing 5% milk: ZC3H4 (1:1000 dilution; Atlas Antibodies, Stockholm, Sweden), Heat Shock Cognate 70 kDa protein (HSC70) (1:10,000; Santa Cruz Biotechnology), High Mobility Group Box 1 (HMGB1) (1:1000; Cell Signaling Technology, Danvers, MA, USA), Lamin B1 (LMNB1) (1:1000; Cell Signaling Technology), phospho-p53 (1:1000; Cell Signaling Technology), p53 (1:1000; Santa Cruz Biotechnology), p21 (1:1000; Santa Cruz Biotechnology), cyclin B (1:1000; Santa Cruz Biotechnology), and cyclin A (1:1000; Santa Cruz Biotechnology). For detecting γH2AX (1:1000 dilution; Merck Millipore) using Western blotting, cell lysis was performed with benzonase (Merck Millipore) to detach the histones from the DNA.

Immunostaining

A total of 5×10^4 cells per well were seeded in 24-well plates containing coverslips coated with 0.2% gelatin (Sigma-Aldrich) and, 24 hours later, were exposed, or not, to genotoxic agents. Cells were washed twice in PBS and fixed in 4% PFA for 20 min at RT, permeabilized in 0.5% Triton X-100 (Sigma-Aldrich) for 15 min, washed twice with PBS, and blocked in a blocking buffer (PBS containing 5% BSA and 10% NDS). The slides were subsequently incubated with primary antibody solutions for 1.5 hours at RT, with mild agitation. Primary antibodies against the following proteins were used: 53BP1 (1:500; Novus Biologicals, Centennial, CO, USA), γH2AX (1:500; Sigma-Aldrich), tubulin (1:5000; Sigma-Aldrich, T5168), and pRPA (1:500; Bethyl Laboratories, Montgomery, TX, USA). The slides were washed three times, 5 min each, in PBS containing 0.1% Triton X-100 and incubated for 30 min with secondary antibody solution [AF647-conjugated goat anti-rabbit immunoglobulin G (IgG) or AF488-conjugated goat anti-mouse IgG; 1:1000 in blocking buffer]. They were then washed three times, 5 min each, with PBS containing 0.1% Triton X-100. The coverslips were mounted on slides using a mounting medium containing 4′,6-diamidino-2-phenylindole (DAPI) (GBI Labs, Kentwood, MI, USA), dried at RT, and stored at 4°C, protected from light. Images of randomly selected fields were captured using an inverted microscope. For R loop immunostaining, cells were first incubated for 3 min in cytoskeleton (CSK) buffer [10 mM Hepes (pH 7.4), 300 mM sucrose, 100 mM NaCl, 3 mM $MgCl_2$, and 0.5% (v/v) Triton X-100] to remove the cellular cytoplasm. The cells were then fixed in ice-cold 100% methanol for 20 min at −20°C. Subsequently, they underwent the above process and stained with the S9.6 antibody (1:100; Sigma-Aldrich).

Microscopic images were analyzed using the CellProfiler software (*85*). For counting nuclear foci, the nuclear region of interest (ROI) was defined using DAPI image segmentation, and the corresponding images were used to quantify the number of foci per ROI. Foci counts were statistically analyzed using a nonparametric *t* test. For R loop signal intensity, the DAPI image was used to define the nuclear ROI, the corresponding S9.6 images were used to quantify and average the signal intensity across the nucleus, and the intensity data were statistically analyzed using a nonparametric *t* test. Microscopy images presented in the figures were processed using ImageJ solely for visualization purposes. Adjustments were limited to linear contrast enhancement and brightness optimization to improve image clarity. All quantitative analyses were performed on raw, unprocessed image data.

Evaluation of RNA synthesis via 5-EU incorporation (epifluorescence)

The Click-iT RNA AF488 Imaging Kit (Thermo Fisher Scientific) was used for evaluating RNA synthesis. A total of 1×10^5 cells per well were seeded in six-well plates precoated with 0.2% gelatin and left overnight to adhere. On the following day, cells were treated, or not, with 100 μM DRB (Sigma-Aldrich) for 1.5 hours in the incubator. Subsequently, incubation continued in the presence of 1 mM 5-EU for 30 min. The cultures were then washed with PBS, and the cells were fixed with 4% PFA for 15 min at RT. After two additional PBS washes, the fixed cells were permeabilized using 0.5% Triton X-100 (Sigma-Aldrich) in PBS for 15 min, washed twice with PBS, and blocked with a blocking buffer (PBS containing 5% BSA and 10% NDS). The incorporated 5-EU was stained using the Click-iT reaction. Coverslips were mounted on slides using a mounting medium containing DAPI (GBI Labs), dried at RT, and stored at 4°C, protected from light. Images of randomly selected fields were captured using an inverted microscope. The DAPI images were used to define nuclear ROIs via segmentation. The average 5-EU signal intensity per nuclear ROI was measured and statistically compared between conditions using a *t* test. Microscopy images displayed in the figures were processed using ImageJ for visualization purposes. Adjustments included linear contrast enhancement and brightness optimization to improve clarity for the reader. All quantifications, however, were performed on raw, unmodified images.

DNA fiber assay

DNA fiber analysis was performed as previously described (*86*). Briefly, cells at 70% confluency were sequentially labeled for 20 min each with the nucleotide analogs 5-iodo-2′-deoxyuridine (IdU) (Sigma-Aldrich) and then 5-chloro-2′-deoxyuridine (CldU) (Merck Millipore) at 37°C, with two PBS washes in between. Cells were harvested by trypsinization, washed with cold PBS, and lysed in 0.5% (v/v) SDS, 200 mM tris (pH 7.4), and 50 mM EDTA for 6 min. DNA lysates were spread on glass slides tilted at a 15° angle, dried at RT, and fixed in cold 3:1 methanol:acetic acid. Before staining, the DNA was denatured in 2.5 M HCl for 30 min at RT, washed with PBS, and blocked in 5% (w/v) BSA and 0.5% (v/v) Triton X-100 in PBS. Slides were then stained with primary antibodies (mouse anti-IdU, BD Biosciences, San Jose, CA, USA; and rat anti-CldU, Abcam) diluted in blocking buffer, followed by three PBS washes. Fluorescent secondary antibody staining was performed using AF647-conjugated goat anti-rabbit IgG and AF488-conjugated goat anti-mouse IgG (both 1:1000 in blocking buffer; Thermo Fisher Scientific). Fibers were visualized using a BZ-X710 microscope (Keyence, Osaka, Japan), and fiber track lengths were calculated using ImageJ (NIH, Bethesda,

MD, USA). Microscopy images displayed in the figures were processed using ImageJ for visualization purposes. Adjustments included linear contrast enhancement and brightness optimization to improve clarity for the reader. All quantifications, however, were performed on raw, unmodified images.

Proximity ligation assay

A total of 5×10^4 cells per well were seeded in 24-well plates containing coverslips coated with 0.2% gelatin (Sigma-Aldrich). After 24 hours, cells were fixed in 4% (v/v) PFA for 15 min at RT, followed by permeabilization with 0.5% (v/v) Triton X-100 for 5 min at RT. PLA was performed using the NaveniFlex Cell Red kit (Navinci Diagnostics AB, Uppsala, Sweden). Replication was stained using PCNA antibodies (1:2000; PC-10, Abcam), and transcription was assessed using RNAPOL2-Ser^2 antibody (1:2000; ab193468, Abcam). Imaging was performed using a BZ-X710 microscope (Keyence, Osaka, Japan), and the number of foci per nucleus was quantified using CellProfiler (The Imaging Platform, Broad Institute, Cambridge, MA, USA).

STORM imaging and data analysis

Multicolor single-molecule localization microscopy (SMLM)–STORM assays, including sample preparation, data acquisition, and application of automated data analysis, were performed using established protocols (*49*, *55*, *86*–*88*). Briefly, for the EdU staining, cells were seeded on coverslips and labeled with 10 μM EdU (Thermo Fisher Scientific) for 15 min at 37°C. For the 5-EU staining, cells were seeded on coverslips and labeled with 1 mM 5-EU (Thermo Fisher Scientific) for 10 min at 37°C. Treated or untreated cells were fixed in 4% (v/v) PFA for 15 min at RT and extracted with CSK buffer [10 mM Hepes (pH 7.4), 300 mM sucrose, 100 mM NaCl, 3 mM $MgCl_2$, and 0.5% (v/v) Triton X-100] for 3 min at RT. After washing three times in PBS, cells were blocked for 60 min at RT with a blocking buffer [5% (w/v) BSA, 2% (w/v) glycine, 0.2% (w/v) gelatin, and 50 mM NH_4Cl]. EdU and 5-EU were detected using Click-iT chemistry (Click-iT Plus EdU AF488 Imaging Kit or Click-iT RNA AF488 Imaging Kit, Thermo Fisher Scientific). Coverslips were stained for 1 hour at RT with antibodies against targets of interest: PCNA (1:5000; Santa Cruz Biotechnology, Dallas, TX, USA), R loop antibody S9.6 (1:5000; Sigma-Aldrich), and pSer^2/RNAP II (1:7500; Active Motif, Carlsbad, CA, USA) in blocking buffer. After mounting on microscope slides, imaging was conducted in an imaging buffer [glucose oxidase (1 mg/ml), catalase (0.02 mg/ml), 10% glucose, and 100 mM cysteamine]. SMLM was performed using a custom-built inverted microscope based on the Rapid Automated Modular Microscope System (ASI RAMM) platform. A 639-nm laser (MRL-FN-639-800, Ultralasers) and a 561-nm laser (Sapphire 561 LPX-500, Coherent) were aligned, expanded, collimated, and directed into a total internal reflection fluorescence objective (UApo N, 100×/1.49 numerical aperture, Olympus) using a pentaedged dichroic beam splitter (FF408/504/581/667/762-Di01, Semrock). Lasers were adjusted into highly inclined and laminated optical sheet mode before imaging with the illumination intensity at ~1.5 kW/cm^2 for the 639-nm laser and 1.0 kW/cm^2 for the 561-nm laser at the exit of the objective. A 405-nm laser (MDL-III-405-100, Ultralasers) was used to drive AF647 to its ground state, and AF647 and Janelia Fluor 549 (JF549) were excited sequentially by the 639- and 561-nm lasers, respectively. Both of their fluorescence was expanded to 1.67× and filtered by a single-band pass filter (FF01-676/37 for AF647 and FF01-607/36 for JF549, Semrock). Fluorescence emission was collected using a scientific complementary metal-oxide semiconductor camera (Prime 95B, Teledyne Photometrics) at 33 Hz/30 ms per frame for a 2000 frames for each color. A regular Dominion Astrophysical Observatory Stochastic Optical Reconstruction Microscopy (DAOSTORM) routine was used for single-molecule localization, as previously described (*49*). Each frame was stacked with a smaller (σ = ~143 nm) and a bigger (σ = ~ 286 nm) Gaussian kernel, and the local maximum was identified by subtracting the bigger from the smaller filtered images. A 9 × 9 square (1 pixel is about 65 nm) around each local maximum was cropped and submitted for multi-focus averaging (MFA) subpixel localization by fitting the data to one or more two-dimensional Gaussian point spread functions through maximum likelihood estimations. The fitting accuracy of each target of interest that was given by the Cramér-Rao lower bound was analyzed by fitting its distribution with a skew Gaussian distribution, using the center as the average localization precision. Localizations that appeared within 2.5 times of the average localization in consecutive frames were averaged and considered as a localization from one blinking event. The coordinates of these events were then submitted for pair correlation analysis. Representative SMLM images were generated from the coordinate list to the 10-nm pixel canvas and blurred with a Gaussian kernel (σ = ~143 nm) for display. SMLM data analysis, including the alignment of three-color channels, auto–pair correlation (PC), and cross-PC analysis, was performed as previously described (*49*, *55*, *86*–*88*).

DRIP-qPCR

DRIP-qPCR was performed as previously described (*15*) with modifications. Briefly, cell pellets corresponding to 5 to 10 million cells were harvested and resuspended in Tris and EDTA (TE) buffer. SDS was added to a final concentration of 0.5% (v/v), and proteinase K was added to a final concentration of 0.5 mg/ml, followed by incubation at 37°C overnight. DNA was extracted through two sequential phenol–chloroform–isoamyl alcohol extractions, precipitated with 2.5 M ammonium acetate and ethanol, washed with 70% (v/v) ethanol, and resuspended overnight at 4°C in nuclease-free TE buffer. DNA was sheared by sonication (8 cycles of 30-s on/30-s off) using a Bioruptor Pico (Diagenode, Denville, NJ, USA) to achieve a fragment size range of 200 to 1000 bp. Sonication efficiency was verified by running a 1% agarose gel. A total of 30 μg of DNA was immunoprecipitated overnight at 4°C with 3 μg of S9.6 antibody (Kerafast) in binding buffer [10 mM sodium phosphate (pH 7), 0.14 M NaCl, and 0.05% (v/v) Triton X-100] with gentle rotation. A slurry of 10 μl of Protein G Dynabeads (Thermo Fisher Scientific) and 10 μl of Protein A Dynabeads (Thermo Fisher Scientific), prewashed in binding buffer, was added to each immunoprecipitation and incubated at 4°C for 2 hours. Immunoprecipitates were washed twice with binding buffer for 15 min each at 4°C with gentle rotation. DNA was eluted in 50 mM tris (pH 8.0), 10 mM EDTA (pH 8.0), and 0.5% (v/v) SDS for 1 hour at 55°C with agitation. The eluate was recovered, and DNA was purified using phenol–chloroform–isoamyl alcohol in 2-ml 5PRIME phase-lock gel tubes (Quantabio) according to the manufacturer's instructions. The DNA was then precipitated and resuspended in RNase-free TE buffer. qPCR was performed on diluted immunoprecipitated DNA at the loci of interest (see table S3), as previously described (*15*), using a QuantStudio 6 Pro qPCR machine (Applied Biosystems, Foster City, CA, USA). Data were analyzed as a percent of input fold change over the control condition.

Computational analysis

Meta-analysis of protein-coding genes in DDR screens

A meta-analysis was conducted to rank protein-coding genes based on their recurrence in published DDR screens (*36*). Data from relevant

studies were manually curated [see Schlam-Babayov *et al.* (*36*) for the list of selected studies]. Next, the genes identified as hits were assigned a score based on their frequency of appearance. The resulting scores were visualized using a distribution histogram generated with the ggplot2 package (*89*) in R. For the correlation analysis between a gene's DDR score and its mutation rate in cancer (fig. S1A), the mutation frequencies for 18,736 protein-coding genes were extracted from the CCLE (*44*) database within the cBioPortal platform (*45*). Genes were subsequently categorized on the basis of their DDR score, and mutation rates were visualized using box plots generated in R Studio (*90*).

Meta-analysis of senescence-associated transcriptomes

This analysis was performed for a separate publication (*40*). Briefly, public RNA-seq data from senescence transcriptomic studies were selected according to two criteria: (i) availability of raw data for both senescent and proliferative states and (ii) experiments conducted in primary cell lines aiming at replicative senescence or senescence induced by DNA damaging agents. Studies involving senescence induced by genetic manipulation such as oncogene-induced senescence were not included. Raw data were acquired from the Short Read Archive (SRA) and were subjected to systematic preprocessing using the Galaxy platform (*91*). Quality control measures included the utilization of FASTQC (*92*), followed by reads trimming to remove 10 bp from the external regions and eliminating reads with a quality score (*Q*) less than 25, using the Trimmomatic tool (*93*) and FASTQC reassessment. Trimmed reads were aligned against the GRCh38/hg38 version of the human genome using the HISAT2 tool (*94*). The alignment quality was confirmed using the MultiQC tool (*95*). For protein-coding genes (GENCODE V44), gene expression quantification was carried out using FeatureCount (*96*). Count tables for each dataset were normalized to log counts per million using EdgeR (*97*) and used to validate a clear and distinct separation between the proliferative and senescent conditions through principal components analysis (R v3.5.3 "stats" package). Subsequently, differential gene expression analysis was performed using the DESeq2 package (*98*), with the following thresholds: adjusted $P < 0.05$ and absolute $\log_2$ fold change $> |0.58|$. To prioritize genes that were consistently deregulated across the selected studies, we assigned a score to each gene based on the number of studies in which its expression was significantly altered. Genes received a score of +1 for up-regulation and −1 for down-regulation in studies where their expression was significantly altered. Last, the scores were plotted in a distribution histogram using the ggplot2 package (*89*).

Coexpression analysis

To identify genes coexpressed with *ZC3H4*, correlation analysis was performed using the analyzeSingleGenes function from the correlationAnalyzeR R package (*49*). The analyzeSingleGenes function parameters were set to Tissue = "all" and the Sample Type = "normal." The implemented corGSEA analysis was subsequently used to predict the *ZC3H4* function and generate the pathway enrichment plots.

To generate the network of pathways correlating with *ZC3H4*, genes were ranked according to their correlation level with *ZC3H4*, as obtained from the previous correlationAnalyzeR analysis. The ranked list was uploaded to GSEA software (*99*) and used in the "Run GSEA-Preranked" tool with GSEA database set to "c2.all.v2023.2.Hs.symbols.gmt," and other settings set as default. Next, the "Enrichment Map Visualization" tool (*43*) was used to generate and export a graph to the Cytoscape software (*100*). Last, the graph was annotated using AutoAnnotate tool (*42*), and style and layout were organized within Cytoscape (*100*).

ChIP-seq signal at ORCs

Raw ChIP-seq data for ZC3H4, WDR82, and IgG input were downloaded from Gene Expression Omnibus (GEO) (*101*) (accession ID: GSE207416). The raw FASTA data were acquired from the SRA and preprocessed using the Galaxy platform (*91*). Quality control was performed using FASTQC tool (*92*), followed by reads trimming to remove 10 bp from the external regions, and reads with a quality score (*Q*) less than 25 were eliminated using the Trimmomatic tool (*93*). Trimmed reads were aligned against the GRCh37-hg19 version of the human genome using the HISAT2 tool (*94*). Aligned BAM files were filtered using "Filter BAM" from BAMtool tool suite (*102*) for a read mapping quality of >20. Next, BigWig files were generated by comparing the ChIP-seq data to the input IgG data using the "bamCompare" tool from deepTools2 suit (*103*). The ORCs were obtained from Liu *et al.* (*56*) (https://github.com/CL-CHEN-Lab/OK-Seq/tree/master/published_results/K562) as BED files. The ChIP-seq coverage (BigWig files) over the origin intervals (BED files) was calculated using the megadepth R package (*104*) and its get_coverage function. In parallel, a random origin BED file was generated using the "ShuffleBed" tool from BEDTools suite (*105*) and used to calculate the ChIP-seq coverage over random regions. Last, the ChIP-seq coverage signal distributions were compared between the origins and shuffled origins using a *t* test.

Correlation analysis between ChIP-seq and RFD

Raw ChIP-seq data for ZC3H4 and IgG Input were preprocessed from FASTA to BigWig files as described above. The RFDs were obtained from Liu *et al.* (*56*) (https://github.com/CL-CHEN-Lab/OK-Seq/tree/master/published_results/K562) as BigWig files. The two replicates within ChIP-seq ZC3H4 were merged in one using "bigwigAverage" tools from deepTools2 suite (*103*). Next, the averaged BigWig files were segmented into bins of 10,000 nucleotides using the "multiBigwigSummary" deepTools2 tool (*103*) with the "Save raw counts (scores) to file" option. The output table contains all the genomic intervals (bins) with the corresponding |RFD| value and ZC3H4 ChIP-seq coverage signal. These data were uploaded onto the R Studio. Bins were ordered by quantiles based on the |RFD| value, using the dplyr R package (*106*). The ZC3H4 ChIP-seq coverage signal was then plotted in a box plot using the ggplot2 package (*89*), with the *x* axis representing the ZC3H4 ChIP-seq coverage signal and the *y* axis representing the ascending |RFD| quantile.

Correlation analysis between ChIP-seq and DRIP-seq

Raw ChIP-seq data for ZC3H4, WDR82, and IgG input were preprocessed from FASTA to BigWig files as described above. For this analysis, the reads were aligned with the GRCh38/hg38 version of the human genome. The preprocessed DRIP-seq data (BigWig files) were retrieved from GEO (accession ID: GSE154631). The replicates within ChIP-seq ZC3H4, ChIP-seq WDR82, and DRIP-seq BigWig files were merged in one using bigwigAverage tools from deepTools2 suite (*103*). Next, the averaged BigWig files were segmented into a bin of 10,000 nucleotides using the multiBigwigSummary deepTools2 tool (*103*) with the Save raw counts (scores) to file option. The output table contains all the genomic intervals (bins) with the corresponding DRIP-seq, ZC3H4 ChIP-seq, and WDR82 ChIP-seq coverage signal. These data were uploaded for analysis in R Studio. Bins were ordered by quantile on the basis of the DRIP-seq value using the dplyr R package (*106*). Then, the ChIP-seq values (ZC3H4 or WDR82) were plotted in a box plot using the ggplot2 package (*89*), with the *x* axis representing the ChIP-seq coverage and the *y* axis representing the ascending DRIP-seq quantile.

RNA-seq public data analysis

RNA-seq and POINT-seq FASTA data were downloaded from GEO (accession ID: GSE207421) and subjected to systematic preprocessing using the Galaxy platform (*91*) as described above. Read quality control was performed using the FASTQC tool (*92*), followed by trimming using the Trimmomatic tool (*93*). Trimmed reads were aligned against the GRCh38/hg38 version of the human genome using the HISAT2 tool (*94*). The counting step was carried out with featureCount (*96*) to quantify the expression over specific ROIs: (i) genes, (ii) ORCs, and (iii) R loop–prone sites. Genes were defined using the hg38 built-in reference. ORCs were defined using the BED files from (*56*) (https://github.com/CL-CHEN-Lab/OK-Seq/tree/master/published_results/K562). R loop–prone sites were defined using Lin *et al.* (*61*) database (https://rloopbase.nju.edu.cn) with "level 5 R loop zones in humans." Differential region expression analysis was performed using the DESeq2 package (*98*), with the following thresholds: adjusted $P < 0.05$ and absolute $\log_2$ fold change > |0.58|. Differentially expressed regions were visualized in volcano plots using the ggplot2 package (*89*). Expression heatmaps were prepared using the pheatmap R package (*107*). Colored palettes were obtained from paletteer R package (*108*). The annotation of the identified ROIs (e.g., sites of R loops induced by ZC3H4 loss) was done using the online annotation tool GREAT (*109*).

Metagenomic profiling of ZC3H4-repressed R loops

R loop–prone sites were identified using publicly available DRIP-seq data from K562 cells [SRX8122753, RLBase; Miller *et al.* (*62*)]. Transcriptional changes at these regions following ZC3H4 depletion were analyzed via differential expression analysis using POINT-seq data (GSE207421). To associate R loop regions with genomic context, each ZC3H4-repressed R loop site was assigned to its nearest protein-coding gene using BEDtools closestBed (*105*), with gene annotations obtained from Ensembl BioMart (*110*). The resulting BED files generated a list of nearby protein-coding genes. Using deepTools2, we plotted the $\log_2$ fold change of expression across these genes, spanning from the promoter and TSS, through the gene body, to the TES. This gene-centric approach enabled us to assess the transcriptional changes surrounding R loop–enriched loci in the context of annotated genes, rather than analyzing R loops as isolated features.

Gene expression profiling interactive analysis

Expression correlation between *ZC3H4* and *RNASEH1* was calculated using the GEPIA2 (gene expression profiling interactive analysis 2) platform (*111*), exclusively analyzing The Cancer Genome Atlas normal tissue samples.

Supplementary Materials

The PDF file includes:
Figs. S1 to S5
Legends for tables S1 to S3

Other Supplementary Material for this manuscript includes the following:
Tables S1 to S3

REFERENCES AND NOTES

1. M. Goldstein, M. B. Kastan, The DNA damage response: Implications for tumor responses to radiation and chemotherapy. *Annu. Rev. Med.* **66**, 129–143 (2015).
2. A. Tubbs, A. Nussenzweig, Endogenous DNA damage as a source of genomic instability in cancer. *Cell* **168**, 644–656 (2017).
3. M. O'Driscoll, The pathological consequences of impaired genome integrity in humans; disorders of the DNA replication machinery. *J. Pathol.* **241**, 192–207 (2017).
4. K. W. Caldecott, DNA single-strand break repair and human genetic disease. *Trends Cell Biol.* **32**, 733–745 (2022).
5. S. Saxena, L. Zou, Hallmarks of DNA replication stress. *Mol. Cell* **82**, 2298–2314 (2022).
6. L. M. Herr, E. D. Schaffer, K. F. Fuchs, A. Datta, R. M. Brosh Jr., Replication stress as a driver of cellular senescence and aging. *Commun. Biol.* **7**, 616 (2024).
7. J. A. Nickoloff, A. S. Jaiswal, N. Sharma, E. A. Williamson, M. T. Tran, D. Arris, M. Yang, R. Hromas, Cellular responses to widespread DNA replication stress. *Int. J. Mol. Sci.* **24**, 16903 (2023).
8. M. B. Adolph, D. Cortez, Mechanisms and regulation of replication fork reversal. *DNA Repair* **141**, 103731 (2024).
9. L. Goehring, T. T. Huang, D. J. Smith, Transcription-replication conflicts as a source of genome instability. *Annu. Rev. Genet.* **57**, 157–179 (2023).
10. K. R. Browning, H. Merrikh, Replication-transcription conflicts: A perpetual war on the chromosome. *Annu. Rev. Biochem.* **93**, 21–46 (2024).
11. E. Petermann, L. Lan, L. Zou, Sources, resolution and physiological relevance of R-loops and RNA-DNA hybrids. *Nat. Rev. Mol. Cell Biol.* **23**, 521–540 (2022).
12. J. R. Brickner, J. L. Garzon, K. A. Cimprich, Walking a tightrope: The complex balancing act of R-loops in genome stability. *Mol. Cell* **82**, 2267–2297 (2022).
13. C. Kumar, D. Remus, Looping out of control: R-loops in transcription-replication conflict. *Chromosoma* **133**, 37–56 (2024).
14. F. J. Groelly, R. A. Dagg, M. Petropoulos, G. G. Rossetti, B. Prasad, A. Panagopoulos, T. Paulsen, A. Karamichali, S. E. Jones, F. Ochs, V. S. Dionellis, E. Puig Lombardi, M. J. Miossec, H. Lockstone, G. Legube, A. N. Blackford, M. Altmeyer, T. D. Halazonetis, M. Tarsounas, Mitotic DNA synthesis is caused by transcription-replication conflicts in BRCA2-deficient cells. *Mol. Cell* **82**, 3382–3397.e7 (2022).
15. L. Goehring, S. Keegan, S. Lahiri, W. Xia, M. Kong, J. Jimenez-Sainz, D. Gupta, R. Drapkin, R. B. Jensen, D. J. Smith, E. Rothenberg, D. Fenyo, T. T. Huang, Dormant origin firing promotes head-on transcription-replication conflicts at transcription termination sites in response to BRCA2 deficiency. *Nat. Commun.* **15**, 4716 (2024).
16. A. Jacquier, The complex eukaryotic transcriptome: Unexpected pervasive transcription and novel small RNAs. *Nat. Rev. Genet.* **10**, 833–844 (2009).
17. K. Ogami, H. I. Suzuki, Nuclear RNA exosome and pervasive transcription: Dual sculptors of genome function. *Int. J. Mol. Sci.* **22**, 13401 (2021).
18. R. Luna, B. Gomez-Gonzalez, A. Aguilera, RNA biogenesis and RNA metabolism factors as R-loop suppressors: A hidden role in genome integrity. *Genes Dev.* **38**, 504–527 (2024).
19. I. W. Deveson, S. A. Hardwick, T. R. Mercer, J. S. Mattick, The dimensions, dynamics, and relevance of the mammalian noncoding transcriptome. *Trends Genet.* **33**, 464–478 (2017).
20. F. J. Slack, A. M. Chinnaiyan, The role of non-coding RNAs in oncology. *Cell* **179**, 1033–1055 (2019).
21. J. S. Mattick, P. P. Amaral, P. Carninci, S. Carpenter, H. Y. Chang, L. L. Chen, R. Chen, C. Dean, M. E. Dinger, K. A. Fitzgerald, T. R. Gingeras, M. Guttman, T. Hirose, M. Huarte, R. Johnson, C. Kanduri, P. Kapranov, J. B. Lawrence, J. T. Lee, J. T. Mendell, T. R. Mercer, K. J. Moore, S. Nakagawa, J. L. Rinn, D. L. Spector, I. Ulitsky, Y. Wan, J. E. Wilusz, M. Wu, Long non-coding RNAs: Definitions, functions, challenges and recommendations. *Nat. Rev. Mol. Cell Biol.* **24**, 430–447 (2023).
22. E. J. Wagner, L. Tong, K. Adelman, Integrator is a global promoter-proximal termination complex. *Mol. Cell* **83**, 416–427 (2023).
23. I. Fianu, M. Ochmann, J. L. Walshe, O. Dybkov, J. N. Cruz, H. Urlaub, P. Cramer, Structural basis of Integrator-dependent RNA polymerase II termination. *Nature* **629**, 219–227 (2024).
24. R. Bhowmick, K. P. M. Mehta, M. Lerdrup, D. Cortez, Integrator facilitates RNAPII removal to prevent transcription-replication collisions and genome instability. *Mol. Cell* **83**, 2357–2366 (2023).
25. C. Estell, L. Davidson, P. C. Steketee, A. Monier, S. West, ZC3H4 restricts non-coding transcription in human cells. *eLife* **10**, e67305 (2021).
26. L. M. I. Austenaa, V. Piccolo, M. Russo, E. Prosperini, S. Polletti, D. Polizzese, S. Ghisletti, I. Barozzi, G. R. Diaferia, G. Natoli, A first exon termination checkpoint preferentially suppresses extragenic transcription. *Nat. Struct. Mol. Biol.* **28**, 337–346 (2021).
27. M. Russo, V. Piccolo, D. Polizzese, E. Prosperini, C. Borriero, S. Polletti, F. Bedin, M. Marenda, D. Michieletto, G. M. Mandana, S. Rodighiero, A. Cuomo, G. Natoli, Restrictor synergizes with Symplekin and PNUTS to terminate extragenic transcription. *Genes Dev.* **37**, 1017–1040 (2023).
28. C. Estell, L. Davidson, J. D. Eaton, H. Kimura, V. A. M. Gold, S. West, A restrictor complex of ZC3H4, WDR82, and ARS2 integrates with PNUTS to control unproductive transcription. *Mol. Cell* **83**, 2222–2239.e5 (2023).
29. C. Estell, S. West, ZC3H4/restrictor exerts a stranglehold on pervasive transcription. *J. Mol. Biol.* **437**, 168707 (2024).
30. C. A. Mimoso, H. Vlaming, N. P. de Wagenaar, K. Adelman, Restrictor slows early transcription elongation to render RNA polymerase II susceptible to termination at non-coding RNA loci. *Genes Dev*. doi: 10.1101/gad.352654.125 (2025).
31. H. B. Landsverk, L. E. Sandquist, L. T. E. Bay, B. Steurer, C. Campsteijn, O. J. B. Landsverk, J. A. Marteijn, E. Petermann, L. Trinkle-Mulcahy, R. G. Syljuasen, WDR82/PNUTS-PP1

prevents transcription-replication conflicts by promoting RNA polymerase II degradation on chromatin. *Cell Rep.* **33**, 108469 (2020).
32. K. Park, J. Zhong, J. S. Jang, J. Kim, H.-J. Kim, J.-H. Lee, J. Kim, ZWC complex-mediated SPT5 phosphorylation suppresses divergent antisense RNA transcription at active gene promoters. *Nucleic Acids Res.* **50**, 3835–3851 (2022).
33. J. O. Rouviere, A. Salerno-Kochan, S. Lykke-Andersen, W. Garland, Y. Dou, O. Rathore, E. S. Molska, G. Wu, M. Schmid, A. Bugai, L. Jakobsen, K. Zumer, P. Cramer, J. S. Andersen, E. Conti, T. H. Jensen, ARS2 instructs early transcription termination-coupled RNA decay by recruiting ZC3H4 to nascent transcripts. *Mol. Cell* **83**, 2240–2257 (2023).
34. A. L. Hughes, A. T. Szczurek, J. R. Kelley, A. Lastuvkova, A. H. Turberfield, E. Dimitrova, N. P. Blackledge, R. J. Klose, A CpG island-encoded mechanism protects genes from premature transcription termination. *Nat. Commun.* **14**, 726 (2023).
35. J. Su, X. Miao, D. Archambault, J. Mager, W. Cui, ZC3H4-a novel Cys-Cys-Cys-His-type zinc finger protein-is essential for early embryogenesis in micedagger. *Biol. Reprod.* **104**, 325–335 (2021).
36. S. Schlam-Babayov, A. Bensimon, M. Harel, T. Geiger, R. Aebersold, Y. Ziv, Y. Shiloh, Phosphoproteomics reveals novel modes of function and inter-relationships among PIKKs in response to genotoxic stress. *EMBO J.* **40**, e104400 (2021).
37. M. Gavish-Izakson, B. B. Velpula, R. Elkon, R. Prados-Carvajal, G. D. Barnabas, A. P. Ugalde, R. Agami, T. Geiger, P. Huertas, Y. Ziv, Y. Shiloh, Nuclear poly(A)-binding protein 1 is an ATM target and essential for DNA double-strand break repair. *Nucleic Acids Res.* **46**, 730–747 (2018).
38. R. D. Jachimowicz, F. Beleggia, J. Isensee, B. B. Velpula, J. Goergens, M. A. Bustos, M. A. Doll, A. Shenoy, C. Checa-Rodriguez, J. L. Wiederstein, K. Baranes-Bachar, C. Bartenhagen, F. Hertwig, N. Teper, T. Nishi, A. Schmitt, F. Distelmaier, H.-J. Lüdecke, B. Albrecht, M. Krüger, B. Schumacher, T. Geiger, D. S. B. Hoon, P. Huertas, M. Fischer, T. Hucho, M. Peifer, Y. Ziv, H. C. Reinhardt, D. Wieczorek, Y. Shiloh, UBQLN4 represses homologous recombination and is overexpressed in aggressive tumors. *Cell* **176**, 505–519.e22 (2019).
39. D. P. Nusinow, J. Szpyt, M. Ghandi, C. M. Rose, E. R. Mc Donald III, M. Kalocsay, J. Jané-Valbuena, E. Gelfand, D. K. Schweppe, M. Jedrychowski, J. Golji, D. A. Porter, T. Rejtar, Y. K. Wang, G. V. Kryukov, F. Stegmeier, B. K. Erickson, L. A. Garraway, W. R. Sellers, S. P. Gygi, Quantitative proteomics of the Cancer Cell Line Encyclopedia. *Cell* **180**, 387–402.e16 (2020).
40. Y. Frey, M. Haj, Y. Ziv, R. Elkon, Y. Shiloh, Broad repression of DNA repair genes in senescent cells identified by integration of transcriptomic data. *Nucleic Acids Res.* **53**, gkae1257 (2025).
41. H. E. Miller, A. J. R. Bishop, Correlation AnalyzeR: Functional predictions from gene co-expression correlations. *BMC Bioinformatics* **22**, 206 (2021).
42. M. Kucera, R. Isserlin, A. Arkhangorodsky, G. D. Bader, AutoAnnotate: A Cytoscape app for summarizing networks with semantic annotations. *F1000Res* **5**, 1717 (2016).
43. D. Merico, R. Isserlin, O. Stueker, A. Emili, G. D. Bader, Enrichment Map: A network-based method for gene-set enrichment visualization and interpretation. *PLOS ONE* **5**, e13984 (2010).
44. M. Ghandi, F. W. Huang, J. Jane-Valbuena, G. V. Kryukov, C. C. Lo, E. R. McDonald III, J. Barretina, E. T. Gelfand, C. M. Bielski, H. Li, K. Hu, A. Y. Andreev-Drakhlin, J. Kim, J. M. Hess, B. J. Haas, F. Aguet, B. A. Weir, M. V. Rothberg, B. R. Paolella, M. S. Lawrence, R. Akbani, Y. Lu, H. L. Tiv, P. C. Gokhale, A. de Weck, A. A. Mansour, C. Oh, J. Shih, K. Hadi, Y. Rosen, J. Bistline, K. Venkatesan, A. Reddy, D. Sonkin, M. Liu, J. Lehar, J. M. Korn, D. A. Porter, M. D. Jones, J. Golji, G. Caponigro, J. E. Taylor, C. M. Dunning, A. L. Creech, A. C. Warren, J. M. McFarland, M. Zamanighomi, A. Kauffmann, N. Stransky, M. Imielinski, Y. E. Maruvka, A. D. Cherniack, A. Tsherniak, F. Vazquez, J. D. Jaffe, A. A. Lane, D. M. Weinstock, C. M. Johannessen, M. P. Morrissey, F. Stegmeier, R. Schlegel, W. C. Hahn, G. Getz, G. B. Mills, J. S. Boehm, T. R. Golub, L. A. Garraway, W. R. Sellers, Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* **569**, 503–508 (2019).
45. I. de Bruijn, R. Kundra, B. Mastrogiacomo, T. N. Tran, L. Sikina, T. Mazor, X. Li, A. Ochoa, G. Zhao, B. Lai, A. Abeshouse, D. Baiceanu, E. Ciftci, U. Dogrusoz, A. Dufilie, Z. Erkoc, E. Garcia Lara, Z. Fu, B. Gross, C. Haynes, A. Heath, D. Higgins, P. Jagannathan, K. Kalletla, P. Kumari, J. Lindsay, A. Lisman, B. Leenknegt, P. Lukasse, D. Madela, R. Madupuri, P. van Nierop, O. Plantalech, J. Quach, A. C. Resnick, S. Y. A. Rodenburg, B. A. Satravada, F. Schaeffer, R. Sheridan, J. Singh, R. Sirohi, S. O. Sumer, S. van Hagen, A. Wang, M. Wilson, H. Zhang, K. Zhu, N. Rusk, S. Brown, J. A. Lavery, K. S. Panageas, J. E. Rudolph, M. L. LeNoue-Newton, J. L. Warner, X. Guo, H. Hunter-Zinck, T. V. Yu, S. Pilai, C. Nichols, S. M. Gardos, J. Philip, AACR Project GENIE BPC Core Team, AACR Project GENIE Consortium, K. L. Kehl, G. J. Riely, D. Schrag, J. Lee, M. V. Fiandalo, S. M. Sweeney, T. J. Pugh, C. Sander, E. Cerami, J. Gao, N. Schultz, Analysis and visualization of longitudinal genomic and clinical data from the AACR Project GENIE Biopharma Collaborative in cBioPortal. *Cancer Res.* **83**, 3861–3867 (2023).
46. M. Ishiyama, H. Tominaga, M. Shiga, K. Sasamoto, Y. Ohkura, K. Ueno, A combined assay of cell viability and in vitro cytotoxicity with a highly water-soluble tetrazolium salt, neutral red and crystal violet. *Biol. Pharm. Bull.* **19**, 1518–1520 (1996).
47. A. H. Bizard, I. D. Hickson, Anaphase: A fortune-teller of genomic instability. *Curr. Opin. Cell Biol.* **52**, 112–119 (2018).
48. A. Bowry, R. D. W. Kelly, E. Petermann, Hypertranscription and replication stress in cancer. *Trends Cancer* **7**, 863–877 (2021).
49. Y. Yin, W. T. C. Lee, D. Gupta, H. Xue, P. Tonzi, J. A. Borowiec, T. T. Huang, M. Modesti, E. Rothenberg, A basal-level activity of ATR links replication fork surveillance and stress response. *Mol. Cell* **81**, 4243–4257.e6 (2021).
50. R. Hand, I. Tamm, Inhibition of mammalian DNA replication by dichlorobenzimidazole riboside. *Exp. Cell Res.* **107**, 343–354 (1977).
51. O. Soderberg, M. Gullberg, M. Jarvius, K. Ridderstrale, K. J. Leuchowius, J. Jarvius, K. Wester, P. Hydbring, F. Bahram, L. G. Larsson, U. Landegren, Direct observation of individual endogenous protein complexes in situ by proximity ligation. *Nat. Methods* **3**, 995–1000 (2006).
52. F. Pessina, F. Giavazzi, Y. Yin, U. Gioia, V. Vitelli, A. Galbiati, S. Barozzi, M. Garre, A. Oldani, A. Flaus, R. Cerbino, D. Parazzoli, E. Rothenberg, F. di d'Adda Fagagna, Functional transcription promoters at DNA double-strand breaks mediate RNA-driven phase separation of damage-response factors. *Nat. Cell Biol.* **21**, 1286–1299 (2019).
53. A. M. Mas, E. Goni, R. de Los, I. Mozos, A. Arcas, L. Statello, J. Gonzalez, L. Blazquez, W. T. C. Lee, D. Gupta, A. Sejas, S. Hoshina, A. Armaos, G. G. Tartaglia, S. Waga, J. Ule, E. Rothenberg, M. Gomez, M. Huarte, ORC1 binds to cis-transcribed RNAs for efficient activation of replication origins. *Nat. Commun.* **14**, 4447 (2023).
54. G. D'Alessandro, D. R. Whelan, S. M. Howard, V. Vitelli, X. Renaudin, M. Adamowicz, F. Iannelli, C. W. Jones-Weinert, M. Lee, V. Matti, W. T. C. Lee, M. J. Morten, A. R. Venkitaraman, P. Cejka, E. Rothenberg, F. d'Adda di Fagagna, BRCA2 controls DNA:RNA hybrid level at DSBs by mediating RNase H2 recruitment. *Nat. Commun.* **9**, 5376 (2018).
55. Y. H. Chen, M. J. Jones, Y. Yin, S. B. Crist, L. Colnaghi, R. J. Sims III, E. Rothenberg, P. V. Jallepalli, T. T. Huang, ATR-mediated phosphorylation of FANCI regulates dormant origin firing in response to replication stress. *Mol. Cell* **58**, 323–338 (2015).
56. Y. Liu, X. Wu, Y. d'Aubenton-Carafa, C. Thermes, C.-L. Chen, OKseqHMM: A genome-wide replication fork directionality analysis toolkit. *Nucleic Acids Res.*, **51**, e22. (2023).
57. S. Kit Leng Lui, S. Keegan, P. Tonzi, M. Kahli, Y. H. Chen, N. Chalhoub, K. E. Coleman, D. Fenyo, D. J. Smith, T. T. Huang, Monitoring genome-wide replication fork directionality by Okazaki fragment sequencing in mammalian cells. *Nat. Protoc.* **16**, 1193–1218 (2021).
58. T. Mosler, F. Conte, G. M. C. Longo, I. Mikicic, N. Kreim, M. M. Mockel, G. Petrosino, J. Flach, J. Barau, B. Luke, V. Roukos, P. Beli, R-loop proximity proteomics identifies a role of DDX41 in transcription-associated genomic instability. *Nat. Commun.* **12**, 7314 (2021).
59. F. Chedin, C. J. Benham, Emerging roles for R-loop structures in the management of topological stress. *J. Biol. Chem.* **295**, 4684–4695 (2020).
60. A. Bayona-Feliu, S. Barroso, S. Munoz, A. Aguilera, The SWI/SNF chromatin remodeling complex helps resolve R-loop-mediated transcription-replication conflicts. *Nat. Genet.* **53**, 1050–1063 (2021).
61. R. Lin, X. Zhong, Y. Zhou, H. Geng, Q. Hu, Z. Huang, J. Hu, X. D. Fu, L. Chen, J. Y. Chen, R-loopBase: A knowledgebase for genome-wide R-loop formation and regulation. *Nucleic Acids Res.* **50**, D303–D315 (2022).
62. H. E. Miller, D. Montemayor, J. Li, S. A. Levy, R. Pawar, S. Hartono, K. Sharma, B. Frost, F. Chedin, A. J. R. Bishop, Exploration and analysis of R-loop mapping data with RLBase. *Nucleic Acids Res.* **51**, D1129–D1137 (2023).
63. P. Preker, J. Nielsen, S. Kammler, S. Lykke-Andersen, M. S. Christensen, C. K. Mapendano, M. H. Schierup, T. H. Jensen, RNA exosome depletion reveals transcription upstream of active human promoters. *Science* **322**, 1851–1854 (2008).
64. S. R. Wessel, K. N. Mohni, J. W. Luzwick, H. Dungrawala, D. Cortez, Functional analysis of the replication fork proteome identifies BET proteins as PCNA regulators. *Cell Rep.* **28**, 3497–3509 e4 (2019).
65. M. Haj, Y. Frey, A. Levon, A. Maliah, T. Ben-Yishay, R. Slutsky, R. Smoom, Y. Tzfati, U. Ben-David, C. Levy, R. Elkon, Y. Ziv, Y. Shiloh, The cGAS-STING, p38 MAPK, and p53 pathways link genome instability to accelerated cellular senescence in ATM-deficient murine lung fibroblasts. *Proc. Natl. Acad. Sci. U.S.A.* **122**, e2419196122 (2025).
66. M. Haj, A. Levon, Y. Frey, N. Hourvitz, J. Campisi, Y. Tzfati, R. Elkon, Y. Ziv, Y. Shiloh, Accelerated replicative senescence of ataxia-telangiectasia skin fibroblasts is retained at physiologic oxygen levels, with unique and common transcriptional patterns. *Aging Cell* **22**, e13869 (2023).
67. C. S. Germain, H. Zhao, J. H. Barlow, Transcription-replication collisions—A series of unfortunate events. *Biomolecules* **11**, 1249 (2021).
68. S. Kemiha, J. Poli, Y. L. Lin, A. Lengronne, P. Pasero, Toxic R-loops: Cause or consequence of replication stress? *DNA Repair* **107**, 103199 (2021).
69. S. Y. Chong, S. Cutler, J. J. Lin, C. H. Tsai, H. K. Tsai, S. Biggins, T. Tsukiyama, Y. C. Lo, C. F. Kao, H3K4 methylation at active genes mitigates transcription-replication conflicts during replication stress. *Nat. Commun.* **11**, 809 (2020).
70. U. Aiello, D. Challal, G. Wentzinger, A. Lengronne, R. Appanah, P. Pasero, B. Palancade, D. Libri, Sen1 is a key regulator of transcription-driven conflicts. *Mol. Cell* **82**, 2952–2966. e6 (2022).

71. R. E. Brown, X. A. Su, S. Fair, K. Wu, L. Verra, R. Jong, K. Andrykovich, C. H. Freudenreich, The RNA export and RNA decay complexes THO and TRAMP prevent transcription-replication conflicts, DNA breaks, and CAG repeat contractions. *PLoS Biol.* **20**, e3001940 (2022).
72. T. Nojima, M. Tellier, J. Foxwell, C. Ribeiro de Almeida, S. M. Tan-Wong, S. Dhir, G. Dujardin, A. Dhir, S. Murphy, N. J. Proudfoot, Deregulated expression of mammalian lncRNA through loss of SPT6 induces R-loop formation, replication stress, and cellular senescence. *Mol. Cell* **72**, 970–984.e7 (2018).
73. D. Papadopoulos, D. Solvie, A. Baluapuri, T. Endres, S. A. Ha, S. Herold, J. Kalb, C. Giansanti, C. Schulein-Volk, C. P. Ade, C. Schneider, A. Gaballa, S. Vos, U. Fischer, M. Dobbelstein, E. Wolf, M. Eilers, MYCN recruits the nuclear exosome complex to RNA polymerase II to prevent transcription-replication conflicts. *Mol. Cell* **82**, 159–176.e12 (2022).
74. M. Wu, P. F. Wang, J. S. Lee, S. Martin-Brown, L. Florens, M. Washburn, A. Shilatifard, Molecular regulation of H3K4 trimethylation by Wdr82, a component of human Set1/COMPASS. *Mol. Cell. Biol.* **28**, 7337–7344 (2008).
75. A. Bonnet, A. R. Grosso, A. Elkaoutari, E. Coleno, A. Presle, S. C. Sridhara, G. Janbon, V. Geli, S. F. de Almeida, B. Palancade, Introns protect eukaryotic genomes from transcription-associated genetic instability. *Mol. Cell* **67**, 608–621 (2017).
76. Y. H. Chen, S. Keegan, M. Kahli, P. Tonzi, D. Fenyo, T. T. Huang, D. J. Smith, Transcription shapes DNA replication initiation and termination in human cells. *Nat. Struct. Mol. Biol.* **26**, 67–77 (2019).
77. M. Macheret, T. D. Halazonetis, DNA replication stress as a hallmark of cancer. *Annu. Rev. Pathol.* **10**, 425–448 (2015).
78. H. Gaillard, T. Garcia-Muse, A. Aguilera, Replication stress and cancer. *Nat. Rev. Cancer* **15**, 276–289 (2015).
79. L. M. F. Primo, L. K. Teixeira, DNA replication stress: Oncogenes in the spotlight. *Genet. Mol. Biol.* **43**, e20190138 (2019).
80. M. Zatzman, F. Fuligni, R. Ripsman, T. Suwal, F. Comitani, L.-M. Edward, R. Denroche, G. H. Jang, F. Notta, S. Gallinger, S. P. Selvanathan, J. A. Toretsky, M. D. Hellmann, U. Tabori, A. Huang, A. Shlien, Widespread hypertranscription in aggressive human cancers. *Sci. Adv.* **8**, eabn0238 (2023).
81. A. Bayona-Feliu, E. Herrera-Moyano, N. Badra-Fajardo, I. Galvan-Femenia, M. E. Soler-Oliva, A. Aguilera, The chromatin network helps prevent cancer-associated mutagenesis at transcription-replication conflicts. *Nat. Commun.* **14**, 6890 (2023).
82. V. J. DeLuca, T. Saleh, Insights into the role of senescence in tumor dormancy: Mechanisms and applications. *Cancer Metastasis Rev.* **42**, 19–35 (2023).
83. G. Mosienia, E. Sikora, Polyploidy: The link between senescence and cancer. *Curr. Pharm. Des.* **16**, 734–740 (2010).
84. T. Saleh, S. Bloukh, M. Hasan, S. Al Shboul, Therapy-induced senescence as a component of tumor biology: Evidence from clinical cancer. *Biochim. Biophys. Acta Rev., Cancer* **1878**, 188994 (2023).
85. D. R. Stirling, M. J. Swain-Bowden, A. M. Lucas, A. E. Carpenter, B. A. Cimini, A. Goodman, CellProfiler 4: Improvements in speed, utility and usability. *BMC Bioinformatics* **22**, 433 (2021).
86. K. E. Coleman, Y. Yin, S. K. L. Lui, S. Keegan, D. Fenyo, D. J. Smith, E. Rothenberg, T. T. Huang, USP1-trapping lesions as a source of DNA replication stress and genomic instability. *Nat. Commun.* **13**, 1740 (2022).
87. D. R. Whelan, W. T. C. Lee, Y. Yin, D. M. Ofri, K. Bermudez-Hernandez, S. Keegan, D. Fenyo, E. Rothenberg, Spatiotemporal dynamics of homologous recombination repair at single collapsed replication forks. *Nat. Commun.* **9**, 3882 (2018).
88. P. Tonzi, Y. Yin, C. W. T. Lee, E. Rothenberg, T. T. Huang, Translesion polymerase kappa-dependent DNA synthesis underlies replication fork recovery. *eLife* **7**, e41426 (2018).
89. H. Wickham, https://ggplot2.tidyverse.org (2016).
90. Team R Studio, www.rstudio.com/ (2020).
91. E. Afgan, D. Baker, B. Batut, M. van den Beek, D. Bouvier, M. Cech, J. Chilton, D. Clements, N. Coraor, B. A. Gruning, A. Guerler, J. Hillman-Jackson, S. Hiltemann, V. Jalili, H. Rasche, N. Soranzo, J. Goecks, J. Taylor, A. Nekrutenko, D. Blankenberg, The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* **46**, W537–W544 (2018).
92. S. Andrews, www.bioinformatics.babraham.ac.uk/projects/fastqc/ (2010).
93. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
94. D. Kim, J. M. Paggi, C. Park, C. Bennett, S. L. Salzberg, Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915 (2019).
95. P. Ewels, M. Magnusson, S. Lundin, M. Kaller, MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).
96. Y. Liao, G. K. Smyth, W. Shi, featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
97. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
98. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
99. A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub, E. S. Lander, J. P. Mesirov, Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550 (2005).
100. P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
101. T. Barrett, S. E. Wilhite, P. Ledoux, C. Evangelista, I. F. Kim, M. Tomashevsky, K. A. Marshall, K. H. Phillippy, P. M. Sherman, M. Holko, A. Yefanov, H. Lee, N. Zhang, C. L. Robertson, N. Serova, S. Davis, A. Soboleva, NCBI GEO: Archive for functional genomics data sets--update. *Nucleic Acids Res.* **41**, D991–D995 (2013).
102. D. W. Barnett, E. K. Garrison, A. R. Quinlan, M. P. Stromberg, G. T. Marth, BamTools: A C++ API and toolkit for analyzing and managing BAM files. *Bioinformatics* **27**, 1691–1692 (2011).
103. F. Ramirez, D. P. Ryan, B. Gruning, V. Bhardwaj, F. Kilpert, A. S. Richter, S. Heyne, F. Dundar, T. Manke, deepTools2: A next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
104. C. Wilks, O. Ahmed, D. N. Baker, D. Zhang, L. Collado-Torres, B. Langmead, Megadepth: Efficient coverage quantification for BigWigs and BAMs. *Bioinformatics* **37**, 3014–3016 (2021).
105. A. R. Quinlan, I. M. Hall, BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
106. H. Wickham, R. François, L. Henry, K. Müller, D. Vaughan, https://dplyr.tidyverse.org (2023).
107. R. Kolde, 10.32614/CRAN.package.pheatmap (2019).
108. E. Hvitfeldt, https://github.com/EmilHvitfeldt/paletteer (2021).
109. C. Y. McLean, D. Bristor, M. Hiller, S. L. Clarke, B. T. Schaar, C. B. Lowe, A. M. Wenger, G. Bejerano, GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* **28**, 495–501 (2010).
110. S. C. Dyer, O. Austine-Orimoloye, A. G. Azov, M. Barba, I. Barnes, V. P. Barrera-Enriquez, A. Becker, R. Bennett, M. Beracochea, A. Berry, J. Bhai, S. K. Bhurji, S. Boddu, P. R. Branco Lins, L. Brooks, S. B. Ramaraju, L. I. Campbell, M. C. Martinez, M. Charkhchi, L. A. Cortes, C. Davidson, S. Denni, K. Dodiya, S. Donaldson, B. El Houdaigui, T. Naboulsi, O. Falola, R. Fatima, T. Genez, J. G. Martinez, T. Gurbich, M. Hardy, Z. Hollis, T. Hunt, M. Kay, V. Kaykala, D. Lemos, D. Lodha, N. Mathlouthi, G. A. Merino, R. Merritt, L. P. Mirabueno, A. Mushtaq, S. N. Hossain, J. G. Perez-Silva, M. Perry, I. Pilizota, D. Poppleton, I. Prosovetskaia, S. Raj, A. I. A. Salam, S. Saraf, N. Saraiva-Agostinho, S. Sinha, B. Sipos, V. Sitnik, E. Steed, M. M. Suner, L. Surapaneni, K. Sutinen, F. F. Tricomi, I. Tsang, D. Urbina-Gomez, A. Veidenberg, T. A. Walsh, N. L. Willhoft, J. Allen, J. Alvarez-Jarreta, M. Chakiachvili, J. Cheema, J. B. da Rocha, N. H. De Silva, S. Giorgetti, L. Haggerty, G. R. Ilsley, J. Keatley, J. E. Loveland, B. Moore, J. M. Mudge, G. Naamati, J. Tate, S. J. Trevanion, A. Winterbottom, B. Flint, A. Frankish, S. E. Hunt, R. D. Finn, M. A. Freeberg, P. W. Harrison, F. J. Martin, A. D. Yates, Ensembl 2025. *Nucleic Acids Res.* **53**, D948–D957 (2025).
111. Z. Tang, B. Kang, C. Li, T. Chen, Z. Zhang, GEPIA2: An enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res.* **47**, W556–W560 (2019).

Acknowledgments: We thank S. West (University of Exeter, UK) for discussions during the early stage of this research. **Funding:** Research in the Shiloh laboratory was supported by grants from the Dr. Miriam and Sheldon G. Adelson Medical Research Foundation (AMRF) (to Y.S.) and the Horizon 2020 Marie Curie ITN "HealthAge" program (GA 812830) (to Y.F.). Work in the Pagano laboratory was funded by an NIH grant (GM136240) (to M.P.). M.P. is a Howard Hughes Medical Institute Investigator. Research in the Huang laboratory was supported by the Goldberg Family Foundation and grants from the NIH: GM139610, ES031658, and CA288368 (to T.T.H.). **Author contributions:** Y.F. conceived the study, conducted the experiments, curated the data, performed validation and formal analysis, developed software and visualizations, and wrote the original draft. Y.S. initiated and supervised the study and contributed to conceptualization, investigation, methodology development, funding acquisition, validation, formal analysis, project administration, software, visualization, and both writing—original draft and review and editing. M.H. contributed to conceptualization, investigation, methodology development, validation, and writing—review and editing. Y.Z. contributed to conceptualization, investigation, methodology development, validation, resources, supervision, formal analysis, visualization, and writing—review and editing. G.R. contributed to methodology development and writing—review and editing. C.F. contributed to conceptualization, methodology, supervision, formal analysis, and software. L.G. contributed to resources and writing—review and editing. M.P. contributed to conceptualization and writing—review and editing. E.R. contributed to conceptualization, writing—review and editing, methodology, resources, funding acquisition, supervision, formal analysis, and project administration. T.T.H. contributed to conceptualization, writing—review and editing, funding acquisition, validation, and formal analysis. **Competing interests:** M.P. is a scientific cofounder of SEED Therapeutics and holds financial interests in the company. He

also serves as a consultant and/or a member of the scientific advisory boards for CullGen, Sibylla Biotech, and Triana Biomedicines, with associated financial interests. In addition, he has received research funding from and holds shares in Kymera Therapeutics. The findings presented in this manuscript were not discussed with any individuals associated with these companies. The other authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The public dataset used across the paper are listed below. RNA-seq and POINT-seq data in K562 cells following ZC3H4 depletion are available under GEO accession (GSE207421, used in Figs. 4G; 5, D and E; and 6, D and E; and figs. S4G and S5I). ChIP-seq data for ZC3H4, WDR82, and Input control are available under GEO accession (GSE207416, used in Figs. 4F, 5C, and 6, B and C; and fig. S4F). Replication origin coordinates in K562 cells were obtained from Liu *et al.* (*56*) (https://github.com/CL-CHEN-Lab/OK-Seq/tree/master/published_results/K562) (used in Fig. 4, F and G, and fig. S4G). DRIPc-seq data in K562 cells are available under GEO accession (GSE154631), samples GSM4679345 and GSM4676155 (used in Fig. 6C and fig. S4F). Public universal R loop–prone regions in human cells were obtained from R-loopBase [Lin *et al.* (*61*); https://rloopbase.nju.edu.cn, files "level-5_human_R-loops.gz"] (used in Fig. 5, D and F, and fig. S5, C to H). Preprocessed DRIP-seq data used for selected ROIs in metagene R loop profiling in K562 cells are available under SRX accession (SRX8122753) [from RLBase (*62*)] (used in fig. S5I). RNA-seq data from aging human dermal fibroblasts are available under GEO accession (GSE113957, used in Fig. 1C, aging circle). RNA-seq data across passages until replicative senescence in human fibroblasts are available under GEO accession (GSE175533, used in Fig. 1D). Transcriptomic datasets for ZC3H4 depletion in various cell lines were used for cross-species R loop–prone site analysis: HCT116 cells (GSE163015, used in fig. S5D), HEK-293 cells (GSE186809, used in fig. S5E), HeLa cells (GSE212208, used in fig. S5F; and GSE133109, used in fig. S5G), and mouse embryonic stem cells (GSE199805, used in fig. S5H), with R loop–prone regions defined using DRIP-seq dataset (SRX5290925, from RLBase).

Submitted 13 October 2024
Accepted 19 May 2025
Published 18 June 2025
10.1126/sciadv.adt8346