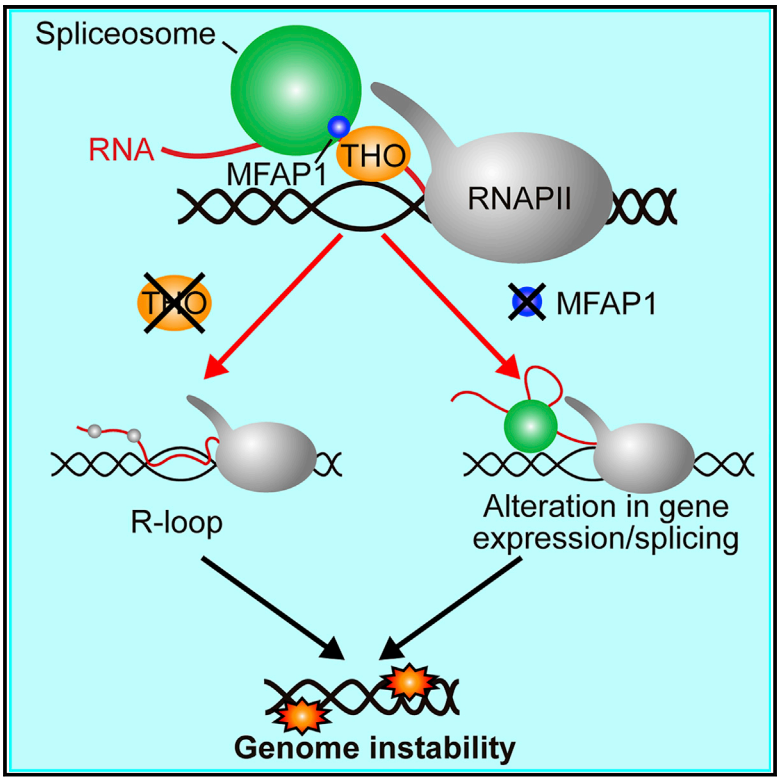
Article

Cell Reports

Depletion of the MFAP1/SPP381 Splicing Factor Causes R-Loop-Independent Genome Instability

Graphical Abstract



Authors

Irene Salas-Armenteros, Sonia I. Barroso, Ana G. Rondón, Mónica Pérez, Eloisa Andújar, Rosa Luna, Andrés Aguilera

Correspondence

rlvaro@us.es (R.L.), aguilo@us.es (A.A.)

In Brief

THO, an mRNA biogenesis factor, interacts with MFAP1, a conserved spliceosome-associated protein. Salas-Armenteros et al. show that MFAP1/SPP381 depletion alters splicing and gene expression and increases genome instability in an RNA-DNA hybrid-independent manner. Therefore, RNA-DNA hybrid accumulation is not an intrinsic consequence of splicing defects.

Highlights

- Human THO interacts physically with MFAP1, a spliceosome-associated factor
- MFAP1 depletion indirectly impairs cell proliferation and genome integrity
- MFAP1/SPP381 defects lead to transcription-independent genome instability
- MFAP1/SPP381 has a wide effect on gene expression and splicing

Salas-Armenteros et al., 2019, Cell Reports *28*, 1551–1563
August 6, 2019
https://doi.org/10.1016/j.celrep.2019.07.010

CellPress

Cell Reports
Article

Depletion of the MFAP1/SPP381 Splicing Factor Causes R-Loop-Independent Genome Instability

Irene Salas-Armenteros,[1] Sonia I. Barroso,[1] Ana G. Rondón,[1] Mónica Pérez,[1] Eloisa Andújar,[1] Rosa Luna,[1,*] and Andrés Aguilera[1,2,*]
[1]Centro Andaluz de Biología Molecular y Medicina Regenerativa-CABIMER, Universidad de Sevilla-CSIC-Universidad Pablo de Olavide, Seville 41092, Spain
[2]Lead Contact
*Correspondence: rlvaro@us.es (R.L.), aguilo@us.es (A.A.)
https://doi.org/10.1016/j.celrep.2019.07.010

SUMMARY

THO/TREX is a conserved complex with a role in messenger ribonucleoprotein biogenesis that links gene expression and genome instability. Here, we show that human THO interacts with MFAP1 (microfibrillar-associated protein 1), a spliceosome-associated factor. Interestingly, MFAP1 depletion impairs cell proliferation and genome integrity, increasing γH2AX foci and DNA breaks. This phenotype is not dependent on either transcription or RNA-DNA hybrids. Mutations in the yeast orthologous gene *SPP381* cause similar transcription-independent genome instability, supporting a conserved role. MFAP1 depletion has a wide effect on splicing and gene expression in human cells, determined by transcriptome analyses. MFAP1 depletion affects a number of DNA damage response (DDR) genes, which supports an indirect role of MFAP1 on genome integrity. Our work defines a functional interaction between THO and RNA processing and argues that splicing factors may contribute to genome integrity indirectly by regulating the expression of DDR genes rather than by a direct role.

INTRODUCTION

Transcription plays a key role not only in gene expression, but also can affect other aspects of DNA metabolism. It has been shown to induce genome instability by increasing the rate of mutagenesis and recombination from prokaryote to eukaryotes (Aguilera and García-Muse, 2013; Gaillard and Aguilera, 2016; Jinks-Robertson and Bhagwat, 2014). A collateral effect of transcription is the formation of R-loops, as a result of the co-transcriptional hybridization of the nascent RNA with the DNA template (Aguilera and García-Muse, 2012; Hamperl and Cimprich, 2014). These R-loops can play specific physiological roles, but in most cases they can constitute a source of genome instability by their ability to halt replication and subsequently cause DNA breaks (García-Muse and Aguilera, 2016; Skourti-Stathaki and Proudfoot, 2014; Sollier and Cimprich, 2015; Yu et al., 2003). Cells have evolved factors and mechanisms, such as RNaseH1, RNA-DNA helicases, topoisomerases, and mRNA ribonucleoprotein (mRNP) biogenesis factors, to prevent harmful R-loops (Santos-Pereira and Aguilera, 2015; Skourti-Stathaki et al., 2011; Tuduri et al., 2009).

The different steps of mRNP biogenesis are tightly connected, so that there is a physical and functional link between transcription, RNA processing (5′ end-capping, splicing, and 3′ end-formation), mRNA export, and nuclear mRNA surveillance to guarantee export of functionally competent RNAs (Bentley, 2014; Schmid and Jensen, 2013). Cumulative evidence supports the idea that a suboptimal mRNP particle can favor the formation of harmful R-loops. The THO complex, involved in mRNP biogenesis, is a paradigm for the connection between mRNP biogenesis with the maintenance of genome integrity. Yeast and human cells lacking a functional THO complex show transcription and R-loop-mediated hyperrecombination, as well as DNA break accumulation, that associate with an increase in transcription-replication conflicts (Gómez-González et al., 2011a; Huertas and Aguilera, 2003; Salas-Armenteros et al., 2017; Wellinger et al., 2006). SRSF1, a splicing factor that regulates early splicing steps, also prevents R-loops as shown in DT40 chicken cells and human cells (Li and Manley, 2005). Genome-wide searches in yeast and human cells identified additional RNA binding and processing factors and other proteins involved in mRNP assembly that increased genome instability (Paulsen et al., 2009; Stirling et al., 2012; Wahba et al., 2011). In addition, RNA-DNA helicases such as senataxin, DDX23, or DDX19, which function in different steps of mRNA biogenesis, also protect cells from R-loops that compromise genome instability (Hodroj et al., 2017; Skourti-Stathaki et al., 2011; Sridhara et al., 2017). Loss of such factors induces γH2AX, DNA breaks, hyperrecombination, chromosome loss, and/or genetic rearrangements. In most cases, overexpression of RNase H1, which specifically degrades the RNA moiety in RNA-DNA hybrids, suppresses these phenotypes and the mutant cells accumulate RNA-DNA hybrids (Aguilera and García-Muse, 2012).

Chromatin may also contribute to protecting cells from R-loop accumulation. Thus, the yeast and human FACT chromatin reorganizing complex and the human mSin3A histone deacetylase prevent R-loop accumulation and/or transcription and R-loop-mediated genome instability (Herrera-Moyano et al., 2014; Salas-Armenteros et al., 2017). Interestingly, mSin3A functions together with THO, which suggests a co-transcriptional crosstalk between THO and chromatin modification factors to

promote histone deacetylation and, consequently, a transient chromatin closing as a coordinated mechanism to prevent R-loop formation.

To explore further the mechanisms by which THO prevents co-transcriptional R-loops, we searched for factors interacting physically with THO. Thus, we identified MFAP1 (microfibrillar-associated protein 1), a putative mRNA splicing factor. Interestingly, MFAP1 depletion also leads to genome instability phenotype, but this cannot be suppressed by either transcription inhibition or RNase H1 overexpression. Similar results were obtained for the yeast ortholog *SPP381* gene. Functional analysis and transcriptomics of small interfering RNA (siRNA) MFAP1-depleted HeLa cells and RNA sequencing (RNA-seq) of an *spp381* yeast mutant support that MFAP1/SPP381 affects splicing and expression of a broad number of genes including DDR genes. Our study thus supports that THO is a conserved mRNA biogenesis factor that interacts with co-transcriptional splicing factors with a distinctive role in preventing R-loop formation that do not share with other RNA-binding and processing factors.

RESULTS

MFAP1 Is a Spliceosome-Associated Protein that Interacts with the THO Complex

To explore further how RNA binding factors could protect genome integrity during transcription, we searched for THO-interacting proteins in human cells. We carried out a two-hybrid screening using as bait the THOC1 subunit of THO fused to the DNA binding domain of Gal4 (GAL4BD-THOC1). We found a clone encoding MFAP1, presumably implicated in mRNA splicing (Andersen and Tapon, 2008; Makarov et al., 2002; Salas-Armenteros et al., 2017; Yeh et al., 1994). To validate this interaction in human cells, we first performed co-immunoprecipitation (coIP) assays in HEK293T cells transiently transfected with a GFP-MFAP1 fusion or GFP. THOC1 was detected by immunoblot after immunoprecipitation (IP) using GFP-trap (Figure 1A). MFAP1-THOC1 interaction was confirmed by coIP with anti-MFAP1 antibody (Figure 1A, bottom panel) and by proximity ligation assay (PLA), which detects protein-protein association *in situ* (Figure 1B). A strong PLA signal was detected, confirming a close association between these factors in human cells. As expected, such association is observed exclusively in the nucleus, because both are nuclear proteins.

Because different genome-wide approaches suggested that MFAP1 was a splicing factor (Hegele et al., 2012; Makarov et al., 2002), we tested the interaction of MFAP1 with the core spliceosome protein SF3B1, a subunit of the U2 small nuclear ribonucleoprotein (snRNP). SF3B1 was enriched in the IP from cells bearing GFP-MFAP1 fusion (Figure 1C). The association between MFAP1 and SF3B1 was confirmed by PLA with anti-MFAP1 and anti-SF3B1 antibodies (Figures 1D). Interestingly, SF3B1 was shown to interact with THOC1 by coIP (Figure S1A). Next, we performed immunofluorescence (IF) of MFAP1 and SC35, commonly used as a marker of splicing speckles. As can be seen in Figure 1E, MFAP1 and SC35 colocalize, indicating a cytological association of MFAP1 with the spliceosome. Human THO/TREX has also been shown to localize at speckles, and there is evidence that the recruitment of this complex to chromatin is partially mediated by the capping and the splicing machineries (Cheng et al., 2006; Custódio et al., 2004; Masuda et al., 2005).

The interaction of the spliceosome-associated factor MFAP1 with THOC1 is consistent with the predicted role of THO in mRNA biogenesis. Next, we asked whether MFAP1 could bind to active chromatin and contribute to the loading of the THO complex or its stabilization in a chromatin-bound stage. Chromatin-IP (ChIP) analysis in highly transcribed genes such as *β-actin* and GAPDH indicated that MFAP1 is present at transcribed chromatin (Figures 1F and S1B), but neither MFAP1 nor SF3B1 seem to be required for THOC1 loading to chromatin. Indeed, a slightly high enrichment of THOC1 was detected in siMFAP1 cells (Figures S1C and S1D).

MFAP1 Depletion Leads to Genome Instability Independent of Transcription and RNA-DNA Hybrids

To investigate whether MFAP1 has a role in the maintenance of genome integrity, as it is the case of THO, we determined the levels of DNA damage, as detected by γH2AX foci, in HeLa cells depleted of MFAP1. A 3.3-fold increase in the percentage of cells with γH2AX foci was observed in siMFAP1 cells with respect to the siC control (Figure 2A). Direct analysis of DNA breaks by single-cell electrophoresis confirmed a significant increase in tail moment in MFAP1 siRNA-depleted cells (5.5-fold and 2.1-fold in alkaline or neutral single-cell electrophoresis assays, respectively) consistent with an accumulation of DNA breaks (Figures 2B and 2C). The levels of sister chromatin exchange (SCE) events resulting from the recombinational repair of replication-born DNA breaks, however, showed no differences between siMFAP1 and siC control cells (Figure S2). Finally, we assayed whether DNA breaks would lead to anaphase bridges resulting from interlinked sister-chromatid intermediates resulting from unfinished replication or incomplete DNA repair (Aguilera and García-Muse, 2013). The percentage of anaphase cells containing bridges was significantly increased after MFAP1 depletion (Figure 2D). We conclude that the MFAP1 splicing factor is required for the maintenance of genome integrity, mainly by preventing the formation and/or accumulation of DNA breaks.

Next, we wondered whether similar to THO-depleted cells, genome instability observed in MFAP1-depleted cells was dependent on transcription and R-loops. We performed alkaline single-cell electrophoresis in the presence of the transcription inhibitor cordycepin and γH2AX IF under overexpression of RNase H1. No significant reduction of DNA breaks was observed in cells depleted of MFAP1 after transcription inhibition (Figure 3A) or after RNase H1 overexpression (Figure 3B), in contrast with the suppression observed in cells depleted of THOC1 and other factors that prevent R-loop accumulation (Domínguez-Sánchez et al., 2011; Herrera-Moyano et al., 2014; García-Rubio et al., 2015; Salas-Armenteros et al., 2017), indicating that DNA breaks induced by MFAP1 depletion were not transcription dependent or mediated by R-loops. Therefore, we conclude that the effect of MFAP1 on genome instability may not be mediated by THO, because the genome instability phenotype is of a different nature. Our data suggest that the impact of MFAP1 depletion in genome integrity has to be indirect.

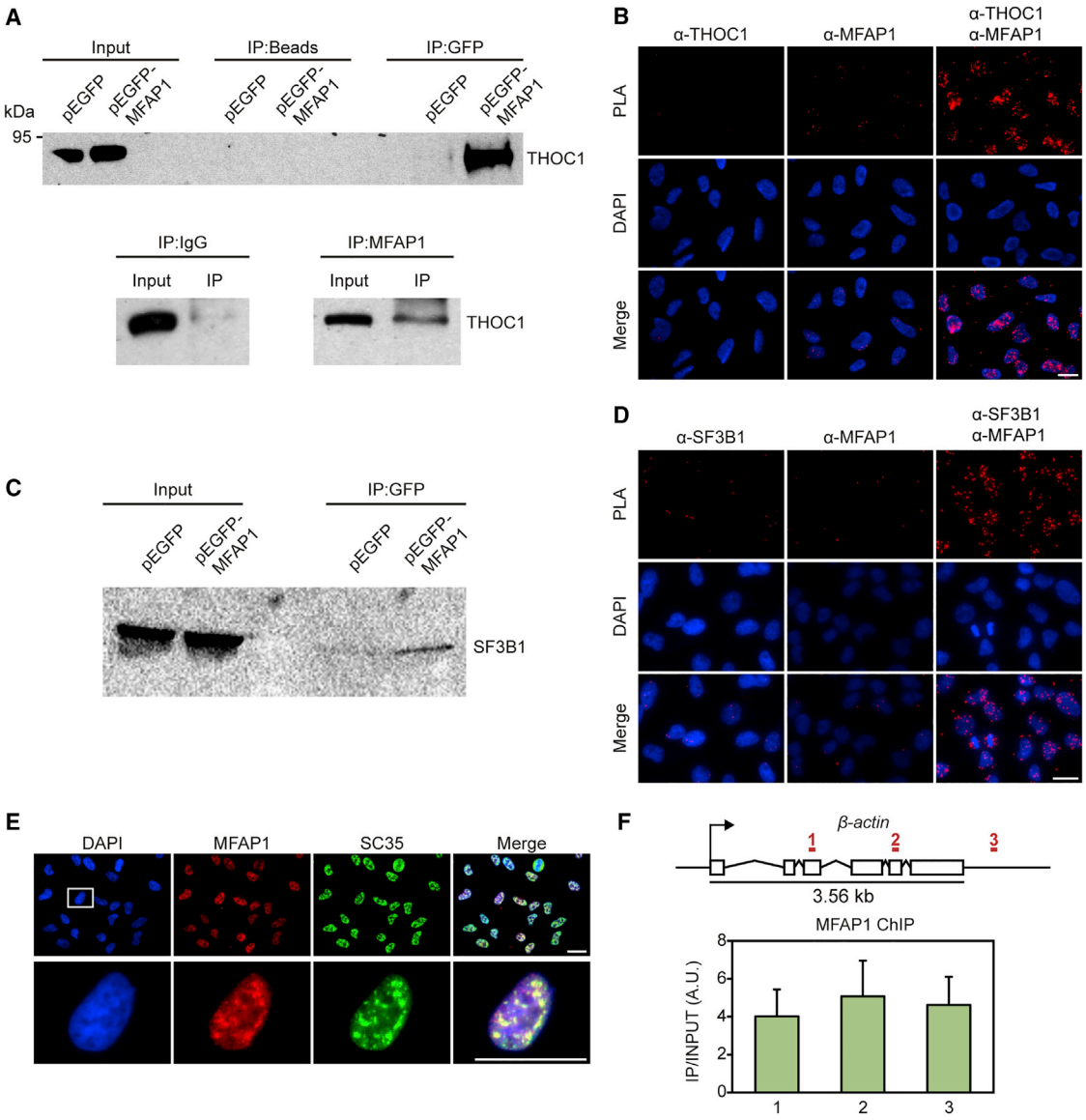


Figure 1. MFAP1 Physically Interacts with THOC1, and SF3B1 Binds to Actively Transcribed Genes and Is Localized at Nuclear Speckles

(A) MFAP1 and THOC1 protein interaction detected by co-immunoprecipitation with anti-GFP antibody (GFP-trap) from whole-cell extracts of HEK293T cells transfected with pEGFP-MFAP1 for 24 h (top panel), and co-immunoprecipitation with anti-MFAP1 antibody from whole-cell extracts of HEK293T (bottom panel). Input extract and total immunoprecipitate (IP) were analyzed by western blot with anti-THOC1 antibody. Input lanes represent 1% of the amount of whole-cell extract used in each experiment (n = 2).

(B) Proximity ligation assay (PLA) in HeLa cells showing specific *in situ* association of THOC1 and MFAP1 endogenous proteins. PLA signal (red spots; n = 2). Negative controls with only one of the antibodies are also shown.

(C) GFP-trap assay from whole extract of cells transfected with pGEFP-MFAP1 for 24 h and immunoblot with anti-SF3B1 antibody (n = 2).

(D) PLA assay in HeLa cells showing specific *in situ* association of MFAP1 and SF3B1 proteins. PLA signal (red spots) (n = 2) is shown.

(E) Colocalization of endogenous MFAP1 with splicing speckles. HeLa cells were labeled by double immunofluorescence using specific antibodies against MFAP1 and the nuclear speckle marker SC35 (n = 2). Bottom panels are a magnification of the region indicated in the first top panel.

(F) ChIP analysis of MFAP1 within the *β-actin* in HeLa cells. Values represent the ratios of precipitated DNA (IP) to input DNA (INPUT). The positions of amplified regions are indicated on the x axis. Mean ± SEM of three independent experiments are shown. Schematic diagram of the gene is depicted. Open boxes represent exons, and the arrow indicates the start transcription site. Numbered red lines indicate the regions where ChIP-qPCR analyses were performed.

Nuclei were stained with DAPI (B, D, and E). Scale bars, 20 μm (B, D, and E).

See also Figure S1.

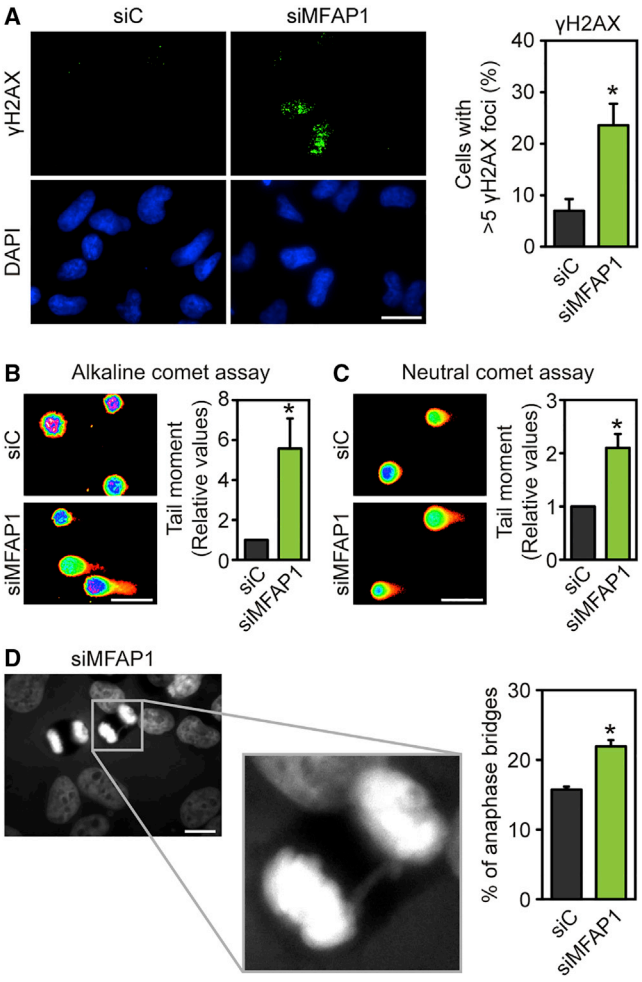


Figure 2. MFAP1 Depletion Leads to Genome Instability

(A) Detection of γH2AX foci by immunofluorescence in siC- (control) and siMFAP1-transfected HeLa cells. The percentage of cells with >5 γH2AX foci is shown. Nuclei were stained with DAPI.

(B) Alkaline single-cell electrophoresis in control and MFAP1-depleted HeLa cells.

(C) Neutral single-cell electrophoresis of HeLa cells transfected with the indicated siRNAs. Relative comet-tail moments are plotted (B and C).

(D) Anaphase bridges in siC- and siMFAP1-transfected HeLa cells. Microphotography showing anaphase bridges in MFAP1-depleted cells. The percentages of anaphase cells with anaphase bridges are plotted.

Data are plotted as mean ± SEM (n = 3). Scale bars, 20 μm (A and D) or 125 μm (B and C). *p < 0.05 as determined by Student's t test (A and D) or Mann-Whitney U test (B and C).

See also Figure S2.

Proliferation and Cell Cycle Are Impaired in MFAP1-Depleted Cells

To gain further insight into the functional role of MFAP1, we analyzed cell proliferation and cell-cycle progression in siRNA MFAP1-depleted cells. As can be seen in Figure 4A, cell proliferation was severely affected in siMFAP1 cells. FACS analysis of asynchronously growing MFAP1-depleted HeLa cells revealed an accumulation of cells in G2 together with a decrease in S phase in MFAP1-depleted cells (Figure 4B). Cell-cycle progression analysis in G1 synchronized cells by double-thymidine block after subsequent release shows S-phase progression slowed down and an increase in the percentage of G2-arrested cells after MFAP1 depletion (Figure 4C). These phenotypes were accompanied by an increase in apoptotic cells, as measured by sub-G1 DNA content (Figure 4D). The results are consistent with the effect of MFAP1 depletion in G2/M transition described in *Drosophila* and the identification of MFAP1 in a siRNA screen for genes affecting cell division (Andersen and Tapon, 2008; Kittler et al., 2004). Therefore, our results show that depletion of MFAP1 leads to defective proliferation and cell cycle in human cells.

MFAP1 Depletion Impairs Gene Expression and Alternative Splicing

Next, we wondered whether the effect of MFAP1 depletion on genome integrity could be related with a differential role on gene expression in comparison with THO-depleted cells. We analyzed the genome-wide transcription and alternative splicing profiles of control and MFAP1-depleted cells, using the Human Transcriptome Array (HTA) 2.0 from Affymetrix. Microarray analyses were performed with three different biological replicates of HeLa cells transfected with siC and siMFAP1 for 72 h. From a total of 67,528 genes analyzed, gene expression analysis identified 904 differentially expressed genes ($p < 0.05$ and $|\text{linear fold change}| > 1.5$), of which 359 (39.7%) were downregulated and 545 (60.3%) were upregulated in siMFAP1 cells (Figure 5A; Table S2). Most downregulated transcripts were mRNAs (49%) or non-coding RNA (ncRNAs; 36.2%), whereas a majority of upregulated transcripts belonged to the category of U snRNAs (40%) (Figure 5B; Table S2). Upregulation of U snRNAs, which are part of the spliceosome and participate in splicing, is consistent with a putative role of MFAP1 in pre-mRNA splicing (Andersen and Tapon, 2008). Gene ontology (GO) analyses reveal a high variety of biological processes enriched ($p < 0.05$ and fold enrichment > 1.5 times), such as those related to chromatin, gene silencing, apoptosis, DNA repair, and cellular response to stress and DNA damage (Table S2). Genes involved in different stress responses were altered in siMFAP1 cells. Thus, the transcriptome analysis reveals an increase of p21 expression, involved in the p53-dependent G1 arrest, and MDM2, an E3 ubiquitin ligase that regulates p53, among others (Table S2). Transcript levels of some DDR genes were also altered, such as FANCD2, a component of the Fanconi anemia pathway, and DRAM2, a DNA damage-regulated autophagy modulator (Table S2). Western blots also showed a reduction at the protein level in some of the cases (Figure S3A). Altogether, our results suggest that siMFAP1 affects genes responding to stress and DNA damage, among others.

Given the interaction between MFAP1 and THOC1, we also included siTHOC1 cells in our analysis. Only a few genes were deregulated (84 out of 67,528), and no biological process was significantly enriched in siTHOC1 cells (Table S3).

Next, we analyzed whether the splicing patterns were altered in siMFAP1 cells. For that, the splicing index (SI) was used to detect splicing changes between MFAP1-depleted cells and siC control cells. The SI represents the change in probe-set inclusion normalized to the change of gene-level expression. We

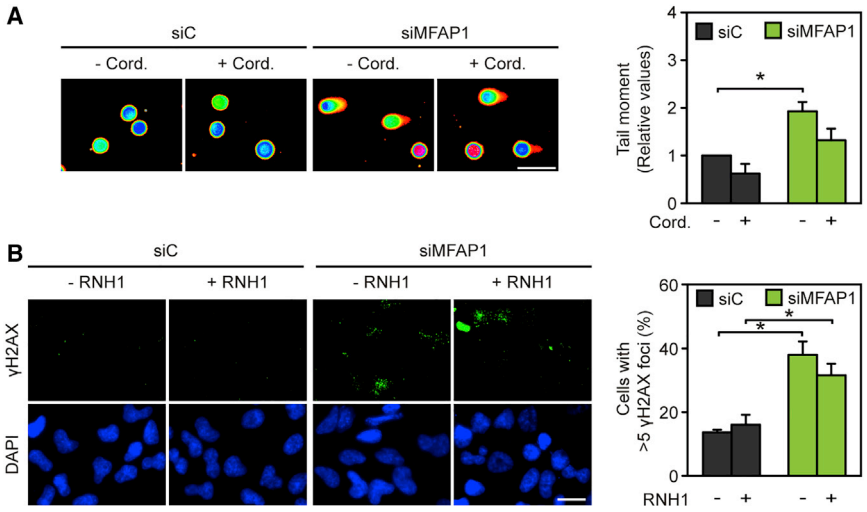


Figure 3. Genome Instability Caused by MFAP1 Depletion Is Not Associated with Transcription or R-Loop Formation

(A) Alkaline comet assay in siC- and siMFAP1-transfected HeLa cells untreated or treated with 50 μM cordycepin for 4 h. Relative comet-tail moments are plotted as mean ± SEM (n = 3).

(B) Immunofluorescence of γH2AX in HeLa cells after siC and siMFAP1 transfection with or without RNase H1 (RNH1) overexpression. Cells were transfected with empty pcDNA3 vector (−RNH1) or pcDNA3-RNaseH1 (+RNH1) for 24 h. Nuclei were stained with DAPI. The percentage of cells with >5 γH2AX foci is shown. Data are plotted as mean ± SEM (n = 3).

Scale bars, 125 μm (A) or 20 μm (B). *p < 0.05 as determined by Mann-Whitney U test (A) or Student's t test (B).

identified 5,490 alternative splicing events that belonged to 3,689 genes in siMFAP1 cells (|splicing index| ≥ 2 and $p < 0.01$) (Table S4). A lower effect in gene splicing was observed after THOC1 depletion (1,022 alternative splicing events corresponding to 921 genes) (Table S5). Comparative analysis of data reveals only a minor overlap of genes with significant changes in splicing between siTHOC1 and siMFAP1 cells (299 genes out 3,390 genes with different alternative splicing in siMFAP1 cells) (Figure S3B).

Taken together, these data suggest that MFAP1 depletion has a high impact on splicing, but not THOC1. Further analysis of the transcriptome of siMFAP1 cells revealed that MFAP1 depletion caused different types of alternative splicing being the skipping or inclusion of a particular exon (cassette exon), the most common type (49%), whereas no case of mutually exclusive exons was found (Figure 5C; Figure S3C). Alterations in splicing patterns in siMFAP1 cells seem not to be due to changes in transcription, because no relationship was found between gene expression levels and SIs (Figure 5D). To assay whether MFAP1 depletion impaired splicing, we performed qRT-PCR of the *β-actin* gene and measured the efficiency of splicing as the ratio of unspliced versus spliced RNAs for different introns. A lower splicing efficiency was observed in siMFAP1 cells (Figure S3D), supporting that MFAP1 functions in splicing.

Interestingly, even though a varied number of biological processes are covered by the differentially spliced genes in siMFAP1 cells, a large proportion of the genes were involved in cell-cycle regulation, DNA damage response (DDR), DNA repair, nuclear export, or chromatin organization and modification (Figure 5E; Table S6) ($p < 0.001$ for a fold enrichment > 1.5). These data indicate that MFAP1 has a broad effect on splicing, and that the different phenotypes of siMFAP1 cells, including genome instability, cell-cycle proliferation defect, or apoptosis, could be the consequence of the altered expression and splicing profile of a broad range of genes, rather than a subset of genes, as it has been shown for other splicing factors (Wang et al., 2016; Vohhodina et al., 2017).

A Conserved Role for the Yeast Orthologous *SPP381* Gene

MFAP1 is conserved in the eukaryotes analyzed (Andersen and Tapon, 2008; Ma et al., 2012); so to identify the putative MFAP1 ortholog in *Saccharomyces cerevisiae*, the most appropriate model organism to validate our study in another eukaryote, we performed an *in silico* search based on sequence comparison. Thus, we identified Spp381, a protein that displays a 13.3% amino acid identity and 22.3% similarity with human MFAP1, as determined by BLAST analysis of *S. cerevisiae* proteome using as a query Saf3, the putative ortholog in the fission yeast *Schizosaccharomyces pombe* (Ren et al., 2011) (Figure S4A). Indeed, recent bioinformatics and biochemical analysis have shown that MFAP1 and Spp381 proteins share characteristic structural features, favoring the conclusion that these genes are conserved from yeast to humans (Ulrich and Wahl, 2017; Ulrich et al., 2016). Interestingly, *SPP381* was identified as a suppressor of a mutant of the splicing factor *PRP38* (Lybarger et al., 1999), consistent with the idea that *SPP381* could also be involved in splicing.

Consequently, we explored whether *SPP381* had any effect on genome instability in *S. cerevisiae*. For this, we assayed the sensitivity of *ssp381-ts* to different genotoxic agents such as UV radiation; the alkylating agent methylmethanesulfonate (MMS); hydroxyurea (HU), which inhibits replication; and Camptothecin (CPT), an inhibitor of DNA topoisomerase I (Figure 6A). *ssp381-ts* was highly sensitive to UV and HU, suggesting an accumulation of DNA damage. Next, we analyzed whether lack of a functional *SPP381* conferred an increase of spontaneous recombination as an indirect measure of DNA breaks, using the L and LYΔNS direct-repeat recombination systems based on direct repeats of a 600-bp region of the *LEU2* gene containing 31 bp and 3.7 kb of intervening sequence, respectively (Prado et al., 1997). The *spp381-ts* mutant shows a 3.9- and 7.8-fold increase in the two systems, respectively (Figure 6B, left panel), suggesting that the longer regions being transcribed in LYΔNS versus L had little accumulation of DNA damage, in contrast

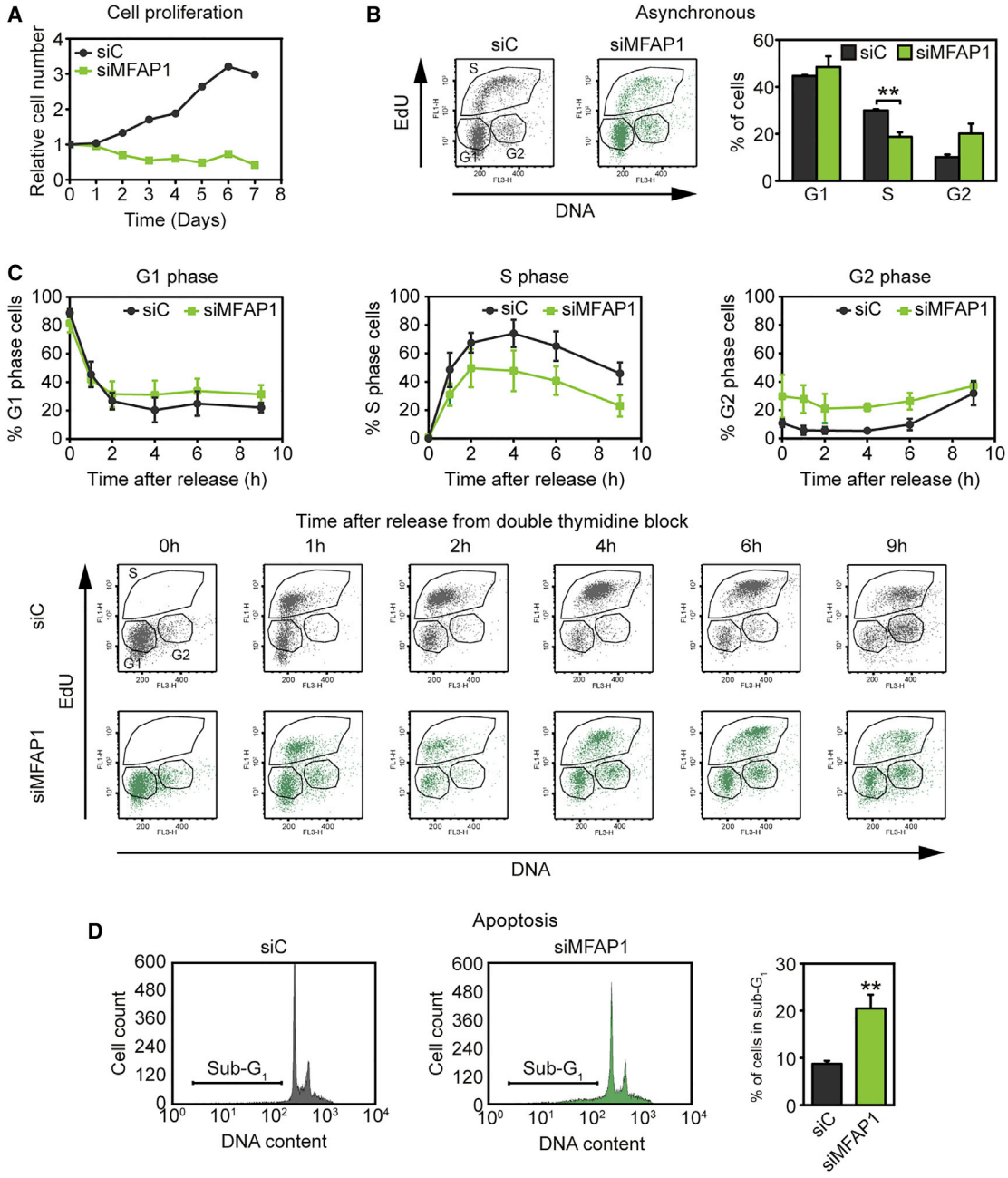


Figure 4. Defects in Cell Proliferation and Cell-Cycle Progression in MFAP1-Depleted Cells

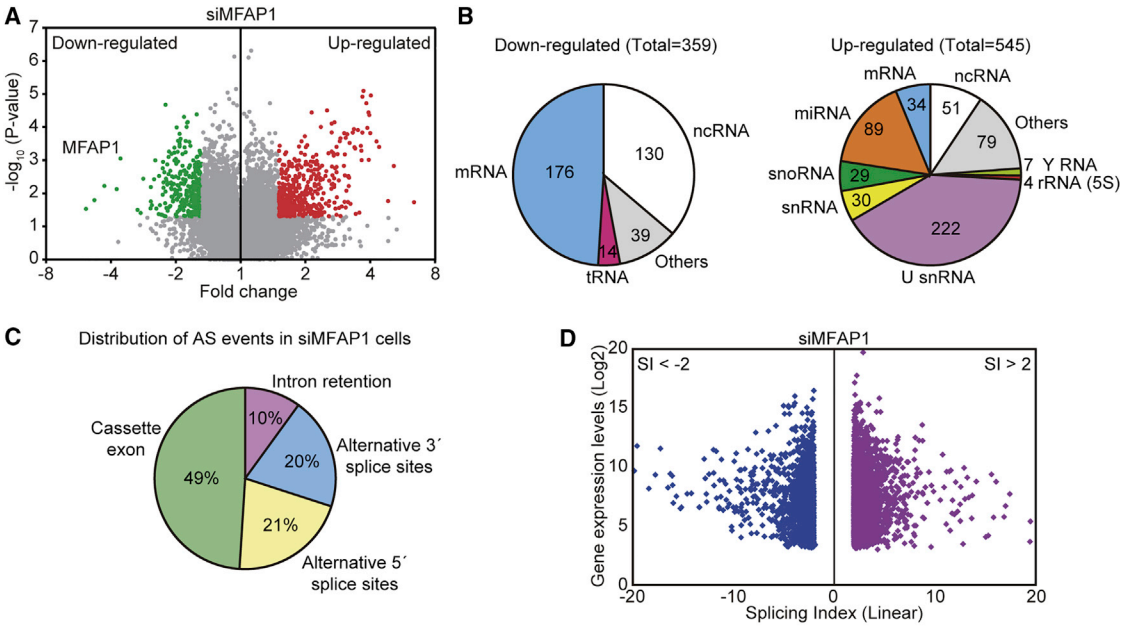
(A) Cell proliferation assays of HeLa cells after MFAP1 depletion. Cells were plated 48 h after siRNA depletion (day 0). Viable cells were those showing metabolic activity of mitochondrial dehydrogenases. Relative values were normalized to the day 0 value. The plotted values correspond to a representative experiment (n = 2).

(B) FACS analysis of asynchronously growing HeLa cells depleted of MFAP1 after 20 min of incubation with 20 μM 5′-ethynyl-2-deoxyuridine (EdU). Cells at different stages of the cell cycle were distinguished based on their DNA content (7-amino-actinomycin D [7-AAD] staining), and replication activity was analyzed by EdU incorporation. Data are plotted as mean ± SEM (n = 3).

(C) FACS analysis of cell-cycle progression in control and MFAP1-depleted HeLa cells after G1 synchronization by double-thymidine block. Samples were collected at the indicated times (hours) after release and processed for flow cytometry. G1, S, and G2 cell populations were distinguished based on their DNA content and replication activity as in (B). Data are plotted as mean ± SEM (n = 3). The percentages of cells in G1, S, and G2 phases are depicted (B and C).

(D) Analysis of apoptosis in siC- and siMFAP1-transfected HeLa cells as determined by FACS analysis of cells displaying subG1-DNA content. Data are plotted as mean ± SEM (n = 4).

**p < 0.01 (Student's t test) (B and D).



E

GO - Biological process	CF	BF	FE	*P*-value
GO:0022402 - Cell cycle process	12.4% (307/2470)	8.09% (1359/16787)	1.54	6.31E-16
GO:1903047 - Mitotic cell cycle process	8.7% (215/2470)	5.28% (887/16787)	1.65	1.83E-14
GO:0000278 - Mitotic cell cycle	9.19% (227/2470)	5.75% (965/16787)	1.60	9.33E-14
GO:0006974 - Cellular response to DNA damage stimulus	7.81% (193/2470)	4.82% (809/16787)	1.62	2.22E-12
GO:0006259 - DNA metabolic process	9.07% (224/2470)	5.95% (999/16787)	1.52	2.18E-11
GO:0006281 - DNA repair	5.51% (136/2470)	3.15% (529/16787)	1.75	2.69E-11
GO:0044770 - Cell cycle phase transition	5.51% (136/2470)	3.23% (542/16787)	1.71	1.60E-10
GO:0018205 - Peptidyl-lysine modification	4.37% (108/2470)	2.41% (405/16787)	1.81	3.86E-10
GO:0044772 - Mitotic cell cycle phase transition	5.14% (127/2470)	3.04% (511/16787)	1.69	1.31E-09
GO:0016570 - Histone modification	4.53% (112/2470)	2.63% (441/16787)	1.73	3.73E-09

Figure 5. HTA 2.0 Gene Expression and Splicing Analysis of MFAP1-Depleted Cells

(A) Volcano plot of differential gene expression in MFAP1-depleted HeLa cells compared with siC control cells (n = 3). Green dots correspond to downregulated genes, and red dots correspond to upregulated genes ($p < 0.05$ and |linear fold change| > 1.5) (see Table S2).

(B) Pie charts showing the distribution of the different downregulated (left) and upregulated (right) transcript types. Transcript numbers in each group are also indicated.

(C) Pie chart showing the alternative splicing-type distribution in siMFAP1-transfected HeLa cells. See also Figure S3B. Alternative splicing events (AS events) in siMFAP1 cells with $p < 0.01$ and |splicing index| > 2 were considered (n = 3). This distribution was calculated from those alternative splicing events with annotations for the type of alternative splicing in the Ensembl database (n = 2491). See also Table S4.

(D) Plot showing the levels of gene expression of the corresponding gene against the splicing indexes (SIs) < −2 or >2 ($p < 0.01$) of probe sets within siMFAP1 cells.

(E) Table showing the top 10 most significant gene ontology (GO) biological processes affected.

BF, background frequency; CF, cluster frequency; FE, fold enrichment. See also Table S6.

with an expected transcription-dependent phenotype as previously shown for THO mutants (Prado et al., 1997). To directly assay whether the hyperrecombination phenotype was transcription dependent, we analyzed recombination in the L-lacZ and GL-lacZ systems carrying the same 0.6-kb *leu2* direct repeats flanking the lacZ open reading frame (ORF), under the *LEU2* and *GAL1* promoter, respectively, under no transcription (GL-lacZ in 2% glucose) or active transcription (L-lacZ in 2% glucose). As can be seen in Figure 6B (right panel), a similar increase of recombination was observed in both systems in the *spp381* mutant. These results suggest that the hyperrecombination phenotype of *spp381* mutant is not associated with transcription. Consistent with the increased genomic instability, high levels of Rad52-YFP foci were detected in *SPP381*-deficient cells, both in the conditional mutant *spp381-ts* (BY4741 background) and in the null mutant *spp381Δ* (viable in the CEN.PK background) (Figure 6C). Importantly, such an increase in Rad52-YFP foci percentage was not suppressed by RNase H1 overexpression in any of the mutants, confirming that R-loops are not responsible for the genetic instability of *SPP381*-deficient cells (Figure 6C). Altogether, our data suggest that lack of MFAP1 and Spp381 proteins confers a similar

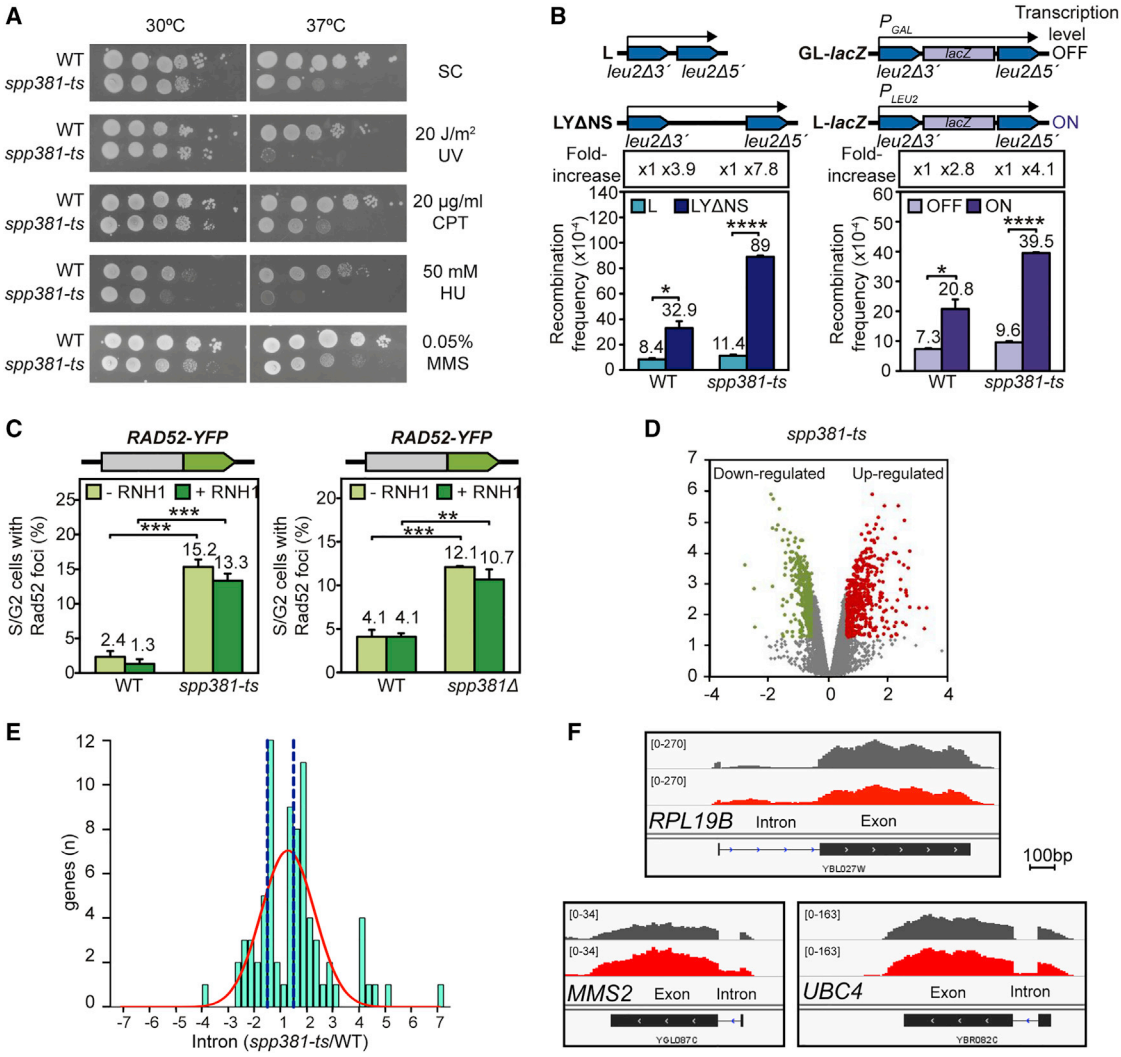


Figure 6. Genome Instability and Expression Defects Caused by Lack of *SPP381* in *S. cerevisiae*

(A) Analysis of sensitivity to different genotoxic agents (UV, methylmethanesulfonate [MMS], hydroxyurea [HU], and camptothecin [CPT]) of wild-type (WT; YDR007W) and *spp381-ts* (YPH2622-2D-4B) strains. Ten-fold serial dilutions of *spp381-ts* and WT cells were plated on minimal complete medium (SC). Photographs were taken after 3 days of growth at 30°C and 37°C (n = 3).

(B) Recombination analysis of WT (Y07202) and *spp381-ts* (YPH2622-2D-4B) cells carrying L and LYΔNS plasmid system (left panel), and cells carrying the GL-*lacZ* system under the control of the inducible *GAL* promoter in glucose (OFF), or the L-*lacZ* system expressed under the control of the constitutive *LEU2* promoter (ON) (right panel).

(C) Percentage of spontaneous Rad52-YFP foci formation in WT and *spp381-ts* strains (BY4741 background) (left panel) with or without RNase H1 (RNH1) overexpression from the pCM184-RNH1 plasmid. Percentage of spontaneous Rad52-YFP foci formation in WT (CEN.PK) and *spp381Δ* (CEN.RO20-3A) strains with or without RNase H1 (RNH1) overexpression from the pCM189-RNH1 plasmid (right panel). Mean ± SEM of three independent experiments are depicted (B and C). A scheme of recombination system (B) or Rad52-YFP plasmid (C) is shown on top of each panel.

(D) Volcano plot of differential gene expression as determined by RNA-seq in *spp381-ts* cells compared with WT control cells after temperature shift (n = 3). Green dots correspond to downregulated genes, and red dots correspond to upregulated genes ($p < 0.05$ and |linear fold change| > 1.5) (see Table S7).

(E) Distribution of intron-containing genes versus the relative intron signal of the *spp381-ts* with respect to wild type. The data for each intron were normalized by host gene expression and pairwise between WT and spp381-ts as described in the STAR Methods (Transcription analysis by RNA-seq) (Table S7). Data are adjusted to a Gaussian curve plotted in red ($R^2 = 0.59$). Dashed lines in blue represent the threshold of significant fold-changes of intron signals (>1.5 and p value < 0.05).

(F) Visualization of RNA-seq coverage across different regions of *Saccharomyces cerevisiae*. Zoomed view of *RPL19B* gene, *MMS2* gene, and *UBC4* gene. The y axis shows the number of reads. Genes and other features are represented according to the *Saccharomyces* Genome Database (SGD).

****p < 0.0001, ***p < 0.001, **p < 0.01, *p < 0.05 (Student's t test) (B and C).

genome-instability phenotype that is not co-transcriptional or R-loop dependent, implying an evolutionarily conserved role in the maintenance of genome integrity of MFAP1 in eukaryotes that is not related with that of THO.

We wondered whether *SPP381* could have a role in gene expression and splicing similar to its human homolog MFAP1. For this purpose, we did RNA-seq of *spp381-ts* and wild-type cells after 1 h of shift from permissive (26°C) to restrictive temperature (37°C). RNA-seq analysis revealed that from a total of 6,692 genes, mRNAs from 500 genes were significantly increased and mRNAs from 378 genes were decreased in *spp381-ts* (|linear fold change| > 1.5 and p value < 0.05) (Table S7; Figure 6D). The functional classes of genes related with metabolic process and with the response to different stresses were enriched in upregulated transcripts (Table S7). Classes including endoplasmic reticulum, protein glycosylation, ribosome, microtubules, and cell-cycle genes, among others, were overrepresented in downregulated transcripts. Although DDR genes were not found as an enriched category, RNA levels of some DDR proteins were downregulated (Alk1, Alk2, and Cdc5 kinases, Rtt109 histone acetyltransferase, and Rtt107), in agreement with the sensitivity to genotoxic stress (Table S7).

Finally, we wondered whether other mRNA biogenesis processes such as mRNA export were also affected. For this purpose, oligo dT-FISH experiments were performed in *spp381-ts* cells, using *mex67-ts* as an export mutant-positive control. No retention of bulk polyA$^+$ mRNA was observed in *spp381-ts* cells (Figure S4B). This is consistent with the result that an RNA export defect per se does not cause genome instability (García-Rubio et al., 2018). Next, we analyzed the impact of *spp381-ts* in splicing by measuring the intron signal in the mutant versus the wild-type from the RNA-seq reads of the intron-containing genes. Significant differences (|linear fold change| > 1.5 and p value < 0.05) were found in 54 out of 297 RNA polymerase II intron-containing genes, consistent with a global defect in splicing (Table S7; Figure 6E) and with the identification of *SPP381* in a screening of splicing mutants in yeast (Dreumont and Séraphin, 2013). Interestingly, despite that in *S. cerevisiae* only a few DDR genes contain introns, an effect in splicing of some of them (*MMS2*, *REC114*, *UBC4*, and *DMC1*) was found, apart from ribosomal protein-coding genes (Table S7; Figure 6F). Altogether, our data support a conserved role of *SPP381*/MFAP1 in gene expression and splicing.

DISCUSSION

We have found an interaction between THOC1, a subunit of the conserved THO complex, and MFAP1, a spliceosome-associated protein whose depletion alters gene expression and splicing, and affects genome integrity. This interaction highlights the key role of THO in the connection of transcription with mRNA metabolism. We show evidence that MFAP1 is a conserved spliceosome component with a broad role in splicing (Figures 1 and 5) and whose depletion affects cell cycle and proliferation in human cells (Figure 4). Importantly, functional analysis in human and yeast cells shows that lack of MFAP1/Spp381 confers a genome instability that, in contrast with THO mutants, is not mediated by R-loops (Figures 2, 3, and 6). Our results suggest that MFAP1 contributes to genome integrity by regulating gene expression and splicing, including that of genes involved in DDR. It is likely that the effect of the yeast Spp381 ortholog in genome instability is also due to its impact on gene expression, provided that many yeast genes do not contain introns. Indeed, we show an RNA-seq analysis of the *spp381-ts* mutant that supports a wider role of *SPP381* in gene expression and mRNA splicing (Figures 6D–6F; Table S7).

MFAP1 was initially identified as a component of the extracellular matrix (Horrigan et al., 1992; Yeh et al., 1994), but later was found in proteomic analysis of the spliceosome bound to the B complex, an intermediate complex in spliceosome assembly (Hegele et al., 2012). Here we show that MFAP1 interacts with SF3B1, a core spliceosome protein (Figures 1C and 1D), and is located at nuclear splicing speckles in human cells (Figure 1E). Altogether our results support that MFAP1 associates with the spliceosome in human cells. Consistently, hMFAP1 has been recently characterized as a B complex-associated spliceosome protein, as determined by biochemical and structural data (Ulrich and Wahl, 2017; Ulrich et al., 2016). Because the B complex carries out the first *trans*-esterification reaction of a splicing event, it is possible that some B complex-specific components have a role in alternative splicing. Indeed, we provide evidence that MFAP1 is a spliceosome-associated factor that affects both transcription and splicing, as determined by qRT-PCR of pre-mRNA (Figure S3D) and transcriptome analysis of siMFAP1 cells (Figure 5; Tables S2, S4, and S6). This may explain the varied number of phenotypes of MFAP1-depleted cells in cell proliferation (Figure 4), cell-cycle regulation (Andersen and Tapon, 2008; Kittler et al., 2004), or sister chromatid cohesion, in this latter case via the control of splicing of CDC5A (Sundaramoorthy et al., 2014). Our transcriptome analysis indeed reveals alternative splicing alterations of a broad number of genes, which are enriched in those encoding factors involved in cell-cycle regulation, cellular response to DNA damage, or DNA repair among deregulated spliced genes (Figure 5E; Table S6). This can explain the pleiotropic effects of MFAP1 depletion on cell proliferation, apoptosis, and genome instability (Figures 2 and 4).

Our data suggest that MFAP1 has a broad effect on gene expression and splicing, being unlikely that only a subset of splicing events requires MFAP1. In contrast, a different scenario has been proposed for other splicing factors, for which a link between gene expression and changes in splicing of specific genes involved in genome instability has been described (Wang et al., 2016; Prados-Carvajal et al., 2018; Vohhodina et al., 2017). Thus, mutations of SF3B1 have been associated with alterations in the DDR in cancer cells (Wang et al., 2016), and mRNA processing factors THRAP3 and BCLAF1 have been shown to play a role in the selective mRNA splicing and export of transcripts of DDR proteins. Interestingly, a recent study has reported a potential control of gene expression by spliceosome components including MFAP1 (Papasaikas et al., 2015).

THO/TREX subunits were identified in proteomic analysis as factors associated with the spliceosome (Rappsilber et al., 2002; Zhou et al., 2002), and some of them (THOC1, THOC2, UAP56, and ALY) were found to colocalize with splicing factors at nuclear speckle domains (Gatfield et al., 2001; Masuda et al., 2005; Zhou et al., 2000). In addition, ALY and UAP56

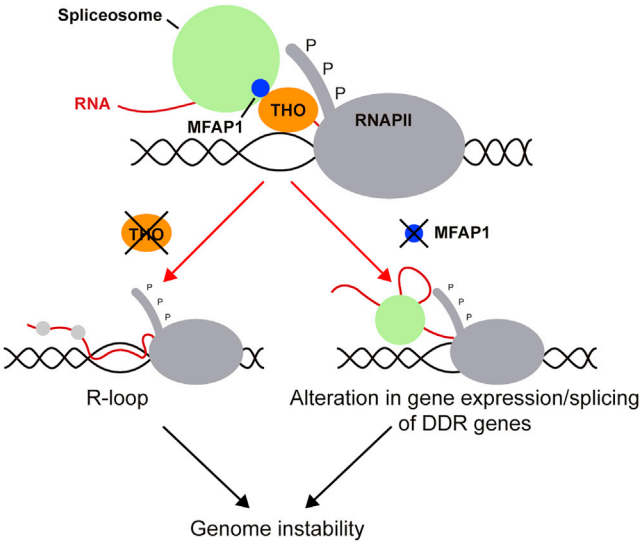


Figure 7. A Model to Explain the Differential Role of RNA Processing Factors in Genome Integrity

Among the many mRNP biogenesis and processing factors, specific RNA-binding factors, such as THO/TREX, would actively contribute to prevent co-transcriptional R-loop formation, thus playing a direct role in the prevention of genome instability, whereas other factors, such as MFAP1 and likely others, would play an indirect role on genome integrity independent of transcription and R-loops, resulting from their impact on the expression and/or splicing of genes directly controlling genome integrity.

proteins have been shown to be necessary for efficient export of spliced mRNA from the nuclear speckles to the cytoplasm (Dias et al., 2010). THOC1 interaction with MFAP1 strengthens the conclusions of other studies showing a relationship between THO/TREX and the splicing machinery (Chanarat and Sträßer, 2013; Cheng et al., 2006; Hurt et al., 2004; Masuda et al., 2005), which may explain why a subset of genes shows similar splicing alterations in siMFAP1 and siTHOC1 cells (Figure S3B). In any case, our study indicates that THOC1 recruitment to active genes is not dependent on MFAP1 or the SF3B1 splicing factor (Figure S1D).

Several RNA binding, heterogeneous nuclear ribonucleoproteins (hnRNPs) and RNA processing proteins, among them splicing factors, have been identified in different global screenings in yeast and humans as factors involved in the maintenance of genome instability (Chan et al., 2014; Paulsen et al., 2009; Stirling et al., 2012). A link between splicing and R-loop-mediated genome instability was suggested after a genome-wide screening revealing that knockdown of a number of splicing factors lead to an increase in DNA damage, as determined by γH2AX foci, which in some cases was partially suppressed by RNase H overexpression (Paulsen et al., 2009). It has also been proposed that blocked transcription after DNA damage promotes selective chromatin displacement of late-stage spliceosomes, thus favoring R-loop formation, even though it is unclear how this would occur (Tresini et al., 2015). However, lack of MFAP1 in human cells compromises genome stability associated with an increase in DNA breaks and anaphase bridges, but this is not associated with transcription or R-loops (Figures 2 and 3). Similar phenotypes are observed in yeast *spp381* mutants (Figure 6). The data indicate that the role of MFAP1 in the maintenance of genome is conserved in eukaryotes and not related with R-loop prevention, arguing that splicing by itself has no direct protective role on genome integrity, consistent with the observation that different splicing-related factors have distinct effects in the DDR (Shkreta and Chabot, 2015). Some examples have been provided on a putative indirect effect of splicing mutations in genome integrity in yeast (Tam et al., 2019; Dahan and Kupiec, 2004). Moreover, it has just been reported that downregulation of the human splicing factor SLU7 is associated with an increase in R-loops and loss of sister chromatid cohesion, but these effects could be caused by the regulatory role of SLU7 in the expression of splicing factors such as SRSF3 and SRSF1 and the protein soronin, which is involved in sister chromatid cohesion (Jiménez et al., 2019). Indeed, a cytological screen of RNA processing factors and chromatin modification mutants in yeast has shown that several splicing factors do not affect R-loops or R-loop-mediated genome instability (Chan et al., 2014). Altogether, these results strengthen the conclusion that only a subset of RNA binding factors, such as, for instance, THO/TREX and SRSF1, has a direct role in preventing R-loop accumulation that compromises genome integrity (Figure 7).

In summary, the spliceosome-associated protein MFAP1 is a partner of THOC1, with a broad effect in alternative splicing. Despite the fact that MFAP1 contributes to the maintenance of genome integrity, this is not functionally related to THO but rather an indirect consequence of the effect of MFAP1 depletion on alternative splicing and gene expression of a number of genes, including DDR genes, a phenomenon that may explain the effect of numerous mutations in splicing factors on genome integrity.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cell culture, siRNA and Plasmid Transfection
- METHOD DETAILS
  - siRNAs, plasmids and transfections
  - Co-immunoprecipitation (coIP)
  - Chromatin immunoprecipitation (ChIP)
  - Immunofluorescence
  - Proximity ligation assay (PLA)
  - Single-cell electrophoresis
  - Anaphase bridges and sister chromatin exchange (SCE) events assays
  - Microscopy images acquisition and data analysis
  - Cell proliferation, cell cycle and apoptosis assays
  - mRNA quantification
  - Microarray analyses of gene expression and alternative splicing
  - Microarray data analysis
  - Gene set enrichment analysis

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.celrep.2019.07.010.

ACKNOWLEDGMENTS

We thank C. Tous for technical assistance with two-hybrid analysis, A. Panek for a preliminary analysis of yeast mutants, and A. Bayona-Feliu for help in the bioinformatics analysis. Research was funded by the European Research Council (grant ERC2014 AdG669898 TARLOOP), the Spanish Ministry of Economy and Competitiveness (grant BFU2016-75058-P), the Junta de Andalucía Spain (grant BIO1238), and the European Union Regional Funds (FEDER). I.S.-A. was a recipient of a FPU pre-doctoral training grant from the Spanish Ministry of Education, Culture and Sports.

AUTHOR CONTRIBUTIONS

I.S.-A., S.I.B., A.G.R., and R.L. performed the experiments. M.P. and E.A. did the sequencing and bioinformatics analyses of the transcriptome and RNA-seq experiments. I.S.-A., R.L., and A.A. designed the experiments, analyzed the results, and wrote the manuscript. All authors read, discussed, and agreed with the final version of the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: September 5, 2018
Revised: March 18, 2019
Accepted: July 1, 2019
Published: August 6, 2019

REFERENCES

Afgan, E., Baker, D., Batut, B., van den Beek, M., Bouvier, D., Cech, M., Chilton, J., Clements, D., Coraor, N., Grüning, B.A., et al. (2018). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Res. *46* (*W1*), W537–W544.

Aguilera, A., and García-Muse, T. (2012). R loops: from transcription byproducts to threats to genome stability. Mol. Cell *46*, 115–124.

Aguilera, A., and García-Muse, T. (2013). Causes of genome instability. Annu. Rev. Genet. *47*, 1–32.

Andersen, D.S., and Tapon, N. (2008). Drosophila MFAP1 is required for pre-mRNA processing and G2/M progression. J. Biol. Chem. *283*, 31256–31267.

Barnett, D.W., Garrison, E.K., Quinlan, A.R., Strömberg, M.P., and Marth, G.T. (2011). BamTools: a C++ API and toolkit for analyzing and managing BAM files. Bioinformatics *27*, 1691–1692.

Ben-Aroya, S., Coombes, C., Kwok, T., O'Donnell, K.A., Boeke, J.D., and Hieter, P. (2008). Toward a comprehensive temperature-sensitive mutant repository of the essential genes of Saccharomyces cerevisiae. Mol. Cell *30*, 248–258.

Bentley, D.L. (2014). Coupling mRNA processing with transcription in time and space. Nat. Rev. Genet. *15*, 163–175.

Bhatia, V., Barroso, S.I., García-Rubio, M.L., Tumini, E., Herrera-Moyano, E., and Aguilera, A. (2014). BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. Nature *511*, 362–365.

Castellano-Pozo, M., Santos-Pereira, J.M., Rondón, A.G., Barroso, S., Andújar, E., Pérez-Alegre, M., García-Muse, T., and Aguilera, A. (2013). R loops are linked to histone H3 S10 phosphorylation and chromatin condensation. Mol. Cell *52*, 583–590.

Chan, Y.A., Aristizabal, M.J., Lu, P.Y., Luo, Z., Hamza, A., Kobor, M.S., Stirling, P.C., and Hieter, P. (2014). Genome-wide profiling of yeast DNA:RNA hybrid prone sites with DRIP-chip. PLoS Genet. *10*, e1004288.

Chanarat, S., and Sträßer, K. (2013). Splicing and beyond: the many faces of the Prp19 complex. Biochim. Biophys. Acta *1833*, 2126–2134.

Cheng, H., Dufu, K., Lee, C.S., Hsu, J.L., Dias, A., and Reed, R. (2006). Human mRNA export machinery recruited to the 5′ end of mRNA. Cell *127*, 1389–1400.

Custódio, N., Carvalho, C., Condado, I., Antoniou, M., Blencowe, B.J., and Carmo-Fonseca, M. (2004). In vivo recruitment of exon junction complex proteins to transcription sites in mammalian cell nuclei. RNA *10*, 622–633.

Dahan, O., and Kupiec, M. (2004). The Saccharomyces cerevisiae gene CDC40/PRP17 controls cell cycle progression through splicing of the ANC1 gene. Nucleic Acids Res. *32*, 2529–2540.

Dias, A.P., Dufu, K., Lei, H., and Reed, R. (2010). A role for TREX components in the release of spliced mRNA from nuclear speckle domains. Nat. Commun. *1*, 97.

Domínguez-Sánchez, M.S., Barroso, S., Gómez-González, B., Luna, R., and Aguilera, A. (2011). Genome instability and transcription elongation impairment in human cells depleted of THO/TREX. PLoS Genet. *7*, e1002386.

Dreumont, N., and Séraphin, B. (2013). Rapid screening of yeast mutants with reporters identifies new splicing phenotypes. FEBS J. *280*, 2712–2726.

Gaillard, H., and Aguilera, A. (2016). Transcription as a Threat to Genome Integrity. Annu. Rev. Biochem. *85*, 291–317.

García-Muse, T., and Aguilera, A. (2016). Transcription-replication conflicts: how they occur and how they are resolved. Nat. Rev. Mol. Cell Biol. *17*, 553–563.

García-Rubio, M.L., Pérez-Calero, C., Barroso, S.I., Tumini, E., Herrera-Moyano, E., Rosado, I.V., and Aguilera, A. (2015). The Fanconi Anemia Pathway Protects Genome Integrity from R-loops. PLoS Genet. *11*, e1005674.

García-Rubio, M., Aguilera, P., Lafuente-Barquero, J., Ruiz, J.F., Simon, M.N., Geli, V., Rondón, A.G., and Aguilera, A. (2018). Yra1-bound RNA-DNA hybrids cause orientation-independent transcription-replication collisions and telomere instability. Genes Dev. *32*, 965–977.

Gatfield, D., Le Hir, H., Schmitt, C., Braun, I.C., Köcher, T., Wilm, M., and Izaurralde, E. (2001). The DExH/D box protein HEL/UAP56 is essential for mRNA nuclear export in Drosophila. Curr. Biol. *11*, 1716–1721.

Gómez-González, B., García-Rubio, M., Bermejo, R., Gaillard, H., Shirahige, K., Marín, A., Foiani, M., and Aguilera, A. (2011a). Genome-wide function of THO/TREX in active genes prevents R-loop-dependent replication obstacles. EMBO J. *30*, 3106–3119.

Gómez-González, B., Ruiz, J.F., and Aguilera, A. (2011b). Genetic and molecular analysis of mitotic recombination in Saccharomyces cerevisiae. Methods Mol. Biol. *745*, 151–172.

Hamperl, S., and Cimprich, K.A. (2014). The contribution of co-transcriptional RNA:DNA hybrid structures to DNA damage and genome instability. DNA Repair (Amst.) *19*, 84–94.

Hegele, A., Kamburov, A., Grossmann, A., Sourlis, C., Wowro, S., Weimann, M., Will, C.L., Pena, V., Lührmann, R., and Stelzl, U. (2012). Dynamic protein-protein interaction wiring of the human spliceosome. Mol. Cell *45*, 567–580.

Herrera-Moyano, E., Mergui, X., García-Rubio, M.L., Barroso, S., and Aguilera, A. (2014). The yeast and human FACT chromatin-reorganizing complexes solve R-loop-mediated transcription-replication conflicts. Genes Dev. *28*, 735–748.

Hodroj, D., Recolin, B., Serhal, K., Martinez, S., Tsanov, N., Abou Merhi, R., and Maiorano, D. (2017). An ATR-dependent function for the Ddx19 RNA helicase in nuclear R-loop metabolism. EMBO J. *36*, 1182–1198.

Horrigan, S.K., Rich, C.B., Streeten, B.W., Li, Z.Y., and Foster, J.A. (1992). Characterization of an associated microfibril protein through recombinant DNA techniques. J. Biol. Chem. *267*, 10087–10095.

Huertas, P., and Aguilera, A. (2003). Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. Mol. Cell *12*, 711–721.

Hurt, E., Luo, M.J., Röther, S., Reed, R., and Strässer, K. (2004). Cotranscriptional recruitment of the serine-arginine-rich (SR)-like proteins Gbp2 and Hrb1 to nascent mRNA via the TREX complex. Proc. Natl. Acad. Sci. USA *101*, 1858–1862.

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U., and Speed, T.P. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics *4*, 249–264.

Jiménez, M., Urtasun, R., Elizalde, M., Azkona, M., Latasa, M.U., Uriarte, I., Arechederra, M., Alignani, D., Bárcena-Varela, M., Álvarez-Sola, G., et al. (2019). Splicing events in the control of genome integrity: role of SLU7 and truncated SRSF3 proteins. Nucleic Acids Res. *47*, 3450–3466.

Jimeno, S., Rondón, A.G., Luna, R., and Aguilera, A. (2002). The yeast THO complex and mRNA export factors link RNA metabolism with transcription and genome instability. EMBO J. *21*, 3526–3535.

Jinks-Robertson, S., and Bhagwat, A.S. (2014). Transcription-associated mutagenesis. Annu. Rev. Genet. *48*, 341–359.

Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: a fast spliced aligner with low memory requirements. Nat. Methods *12*, 357–360.

Kittler, R., Putz, G., Pelletier, L., Poser, I., Heninger, A.K., Drechsel, D., Fischer, S., Konstantinova, I., Habermann, B., Grabner, H., et al. (2004). An endoribonuclease-prepared siRNA screen in human cells identifies genes essential for cell division. Nature *432*, 1036–1040.

Law, C.W., Chen, Y., Shi, W., and Smyth, G.K. (2014). voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol. *15*, R29.

Li, X., and Manley, J.L. (2005). Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. Cell *122*, 365–378.

Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930.

Lisby, M., Mortensen, U.H., and Rothstein, R. (2003). Colocalization of multiple DNA double-strand breaks at a single Rad52 repair centre. Nat. Cell Biol. *5*, 572–577.

Lybarger, S., Beickman, K., Brown, V., Dembla-Rajpal, N., Morey, K., Seipelt, R., and Rymond, B.C. (1999). Elevated levels of a U4/U6.U5 snRNP-associated protein, Spp381p, rescue a mutant defective in spliceosome maturation. Mol. Cell. Biol. *19*, 577–584.

Ma, L., Gao, X., Luo, J., Huang, L., Teng, Y., and Horvitz, H.R. (2012). The Caenorhabditis elegans gene mfap-1 encodes a nuclear protein that affects alternative splicing. PLoS Genet. *8*, e1002827.

Makarov, E.M., Makarova, O.V., Urlaub, H., Gentzel, M., Will, C.L., Wilm, M., and Lührmann, R. (2002). Small nuclear ribonucleoprotein remodeling during catalytic activation of the spliceosome. Science *298*, 2205–2208.

Masuda, S., Das, R., Cheng, H., Hurt, E., Dorman, N., and Reed, R. (2005). Recruitment of the human TREX complex to mRNA during splicing. Genes Dev. *19*, 1512–1517.

Mumberg, D., Müller, R., and Funk, M. (1994). Regulatable promoters of Saccharomyces cerevisiae: comparison of transcriptional activity and their use for heterologous expression. Nucleic Acids Res. *22*, 5767–5768.

Papasaikas, P., Tejedor, J.R., Vigevani, L., and Valcárcel, J. (2015). Functional splicing network reveals extensive regulatory potential of the core spliceosomal machinery. Mol. Cell *57*, 7–22.

Paulsen, R.D., Soni, D.V., Wollman, R., Hahn, A.T., Yee, M.C., Guan, A., Hesley, J.A., Miller, S.C., Cromwell, E.F., Solow-Cordero, D.E., et al. (2009). A genome-wide siRNA screen reveals diverse cellular processes and pathways that mediate genome stability. Mol. Cell *35*, 228–239.

Piruat, J.I., and Aguilera, A. (1998). A novel yeast gene, THO2, is involved in RNA pol II transcription and provides new evidence for transcriptional elongation-associated recombination. EMBO J. *17*, 4859–4872.

Prado, F., Piruat, J.I., and Aguilera, A. (1997). Recombination between DNA repeats in yeast hpr1delta cells is linked to transcription elongation. EMBO J. *16*, 2826–2835.

Prados-Carvajal, R., López-Saavedra, A., Cepeda-García, C., Jimeno, S., and Huertas, P. (2018). Multiple roles of the splicing complex SF3B in DNA end resection and homologous recombination. DNA Repair (Amst.) *66-67*, 11–23.

Rappsilber, J., Ryder, U., Lamond, A.I., and Mann, M. (2002). Large-scale proteomic analysis of the human spliceosome. Genome Res. *12*, 1231–1245.

Ren, L., McLean, J.R., Hazbun, T.R., Fields, S., Vander Kooi, C., Ohi, M.D., and Gould, K.L. (2011). Systematic two-hybrid and comparative proteomic analyses reveal novel yeast pre-mRNA splicing factors connected to Prp19. PLoS ONE *6*, e16719.

Salas-Armenteros, I., Pérez-Calero, C., Bayona-Feliu, A., Tumini, E., Luna, R., and Aguilera, A. (2017). Human THO-Sin3A interaction reveals new mechanisms to prevent R-loops that cause genome instability. EMBO J. *36*, 3532–3547.

Santos-Pereira, J.M., and Aguilera, A. (2015). R loops: new modulators of genome dynamics and function. Nat. Rev. Genet. *16*, 583–597.

Santos-Pereira, J.M., Herrero, A.B., García-Rubio, M.L., Marín, A., Moreno, S., and Aguilera, A. (2013). The Npl3 hnRNP prevents R-loop-mediated transcription-replication conflicts and genome instability. Genes Dev. *27*, 2445–2458.

Schmid, M., and Jensen, T.H. (2013). Transcription-associated quality control of mRNP. Biochim. Biophys. Acta *1829*, 158–168.

Shkreta, L., and Chabot, B. (2015). The RNA Splicing Response to DNA Damage. Biomolecules *5*, 2935–2977.

Skourti-Stathaki, K., and Proudfoot, N.J. (2014). A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. Genes Dev. *28*, 1384–1396.

Skourti-Stathaki, K., Proudfoot, N.J., and Gromak, N. (2011). Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. Mol. Cell *42*, 794–805.

Söderberg, O., Gullberg, M., Jarvius, M., Ridderstråle, K., Leuchowius, K.J., Jarvius, J., Wester, K., Hydbring, P., Bahram, F., Larsson, L.G., and Landegren, U. (2006). Direct observation of individual endogenous protein complexes in situ by proximity ligation. Nat. Methods *3*, 995–1000.

Sollier, J., and Cimprich, K.A. (2015). Breaking bad: R-loops and genome integrity. Trends Cell Biol. *25*, 514–522.

Sridhara, S.C., Carvalho, S., Grosso, A.R., Gallego-Paez, L.M., Carmo-Fonseca, M., and de Almeida, S.F. (2017). Transcription Dynamics Prevent RNA-Mediated Genomic Instability through SRPK2-Dependent DDX23 Phosphorylation. Cell Rep. *18*, 334–343.

Stirling, P.C., Chan, Y.A., Minaker, S.W., Aristizabal, M.J., Barrett, I., Sipahimalani, P., Kobor, M.S., and Hieter, P. (2012). R-loop-mediated genome instability in mRNA cleavage and polyadenylation mutants. Genes Dev. *26*, 163–175.

Sundaramoorthy, S., Vázquez-Novelle, M.D., Lekomtsev, S., Howell, M., and Petronczki, M. (2014). Functional genomics identifies a requirement of pre-mRNA splicing factors for sister chromatid cohesion. EMBO J. *33*, 2623–2642.

Tam, A.S., Sihota, T.S., Milbury, K.L., Zhang, A., Mathew, V., and Stirling, P.C. (2019). Selective defects in gene expression control genome instability in yeast splicing mutants. Mol. Biol. Cell *30*, 191–200.

ten Asbroek, A.L., van Groenigen, M., Nooij, M., and Baas, F. (2002). The involvement of human ribonucleases H1 and H2 in the variation of response of cells to antisense phosphorothioate oligonucleotides. Eur. J. Biochem. *269*, 583–592.

Trcek, T., Chao, J.A., Larson, D.R., Park, H.Y., Zenklusen, D., Shenoy, S.M., and Singer, R.H. (2012). Single-mRNA counting using fluorescent in situ hybridization in budding yeast. Nat. Protoc. *7*, 408–419.

Tresini, M., Warmerdam, D.O., Kolovos, P., Snijder, L., Vrouwe, M.G., Demmers, J.A., van IJcken, W.F., Grosveld, F.G., Medema, R.H., Hoeijmakers, J.H., et al. (2015). The core spliceosome as target and effector of non-canonical ATM signalling. Nature *523*, 53–58.

Tuduri, S., Crabbé, L., Conti, C., Tourrière, H., Holtgreve-Grez, H., Jauch, A., Pantesco, V., De Vos, J., Thomas, A., Theillet, C., et al. (2009). Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription. Nat. Cell Biol. *11*, 1315–1324.

Tumini, E., Barroso, S., Perez-Calero, C., and Aguilera, A. (2016). Roles of human POLD1 and POLD3 in genome stability. Sci. Rep. *6*, 38873.

Ulrich, A.K., and Wahl, M.C. (2017). Human MFAP1 is a cryptic ortholog of the Saccharomyces cerevisiae Spp381 splicing factor. BMC Evol. Biol. *17*, 91.

Ulrich, A.K.C., Seeger, M., Schütze, T., Bartlick, N., and Wahl, M.C. (2016). Scaffolding in the Spliceosome via Single α Helices. Structure *24*, 1972–1983.

Vohhodina, J., Barros, E.M., Savage, A.L., Liberante, F.G., Manti, L., Bankhead, P., Cosgrove, N., Madden, A.F., Harkin, D.P., and Savage, K.I. (2017). The RNA processing factors THRAP3 and BCLAF1 promote the DNA damage response through selective mRNA splicing and nuclear export. Nucleic Acids Res. *45*, 12816–12833.

Wahba, L., Amon, J.D., Koshland, D., and Vuica-Ross, M. (2011). RNase H and multiple RNA biogenesis factors cooperate to prevent RNA:DNA hybrids from generating genome instability. Mol. Cell *44*, 978–988.

Wang, L., Brooks, A.N., Fan, J., Wan, Y., Gambe, R., Li, S., Hergert, S., Yin, S., Freeman, S.S., Levin, J.Z., et al. (2016). Transcriptomic Characterization of SF3B1 Mutation Reveals Its Pleiotropic Effects in Chronic Lymphocytic Leukemia. Cancer Cell *30*, 750–763.

Wellinger, R.E., Prado, F., and Aguilera, A. (2006). Replication fork progression is impaired by transcription in hyperrecombinant yeast cells lacking a functional THO complex. Mol. Cell. Biol. *26*, 3327–3334.

Yeh, H., Chow, M., Abrams, W.R., Fan, J., Foster, J., Mitchell, H., Muenke, M., and Rosenbloom, J. (1994). Structure of the human gene encoding the associated microfibrillar protein (MFAP1) and localization to chromosome 15q15-q21. Genomics *23*, 443–449.

Yu, K., Chedin, F., Hsieh, C.L., Wilson, T.E., and Lieber, M.R. (2003). R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. Nat. Immunol. *4*, 442–451.

Zhou, Z., Luo, M.J., Straesser, K., Katahira, J., Hurt, E., and Reed, R. (2000). The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans. Nature *407*, 401–405.

Zhou, Z., Licklider, L.J., Gygi, S.P., and Reed, R. (2002). Comprehensive proteomic analysis of the human spliceosome. Nature *419*, 182–185.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse monoclonal anti-THOC1	Abcam	Cat# ab487, RRID:AB_304696
Rabbit polyclonal anti-MFAP1	Abcam	Cat# ab175508
Mouse monoclonal anti-GFP	Roche	Cat# 11814460001, RRID:AB_390913
Mouse monoclonal anti-SC35	Abcam	Cat# ab11826,RRID:AB_298608
Mouse monoclonal anti-γH2AX	Millipore	Cat# 05-636, RRID:AB_309864
Rabbit monoclonal anti-p21	Abcam	Cat# ab109520, RRID:AB_10860537
Mouse monoclonal anti-FANCD2	Santa Cruz	Cat# sc-20022, RRID:AB_2278211
Mouse monoclonal anti-vinculin	Sigma-Aldrich	Cat# V9264, RRID:AB_10603627
Mouse monoclonal anti-SF3B1	Antibodies online	Cat# ABIN1449241
Rabbit polyclonal anti-beta-actin	Abcam	Cat# ab8227, RRID:AB_2305186
Alexa fluor 594 Rabbit	Thermofisher	Cat# A21201, RRID:AB_2535787
Alexa fluor 488 Mouse	Thermofisher	Cat# A21200, RRID:AB_2535786
Anti-rabbit IgG peroxidase antibody produced in goat	Sigma	Cat# A6154, RRID:AB_258284
Anti-mouse IgG peroxidase antibody produced in goat	Sigma	Cat# A4416, RRID:AB_258167
Chemicals, Peptides, and Recombinant Proteins		
Cordycepin	Sigma	Cat# C3394
methylmethanesulfonate (MMS)	Fluka	Cat# 64294
Hydroxyurea	USBiological	Cat# C11082252
Camptothecin	Sigma	Cat# C9911
Nocodazole	Sigma	Cat# 487928
Thymidine	Sigma	Cat# T9250
Critical Commercial Assays		
Duo-link *in situ* Red started Kit (mouse-rabbit). Proximity ligation assay	Sigma	Cat# DUO92101
Comet assay Kit	Trevigen	Cat# 4250-050-K
GeneChip Human Transcriptome array 2.0 (HTA 2.0)	Thermofisher	Cat# 902162
TruSeq® Stranded mRNA Library prep kit	Illumina	Cat# 20020594
Deposited Data		
Gene expression and alternative splicing profile of MFAP1- and THOC1-depleted human cells	Gene Expression Omnibus database https://www.ncbi.nlm.nih.gov/geo/	accession number GEO: GSE119412
Transcriptome analysis by RNA-Seq of spp381-ts mutant	Gene Expression Omnibus database https://www.ncbi.nlm.nih.gov/geo/	accession number GEO: GSE128010
Experimental Models: Cell Lines		
HeLa	ECACC	Cat# 93021013
HEK293T	ATCC	Cat# CRL-11268
Experimental Models: Organisms/Strains		
Saccharomyces cerevisiae: Strain background: BY4741 *MATa can1Δ::MFA1*pr-*HIS3*::*LEU2 his3Δ1 met15Δ0 spp381-ts::URA3*	Ben-Aroya et al., 2008	YPH2622
Saccharomyces cerevisiae: Strain background: BY4741 *MATα his3Δ1 leu2Δ0 trp1Δ::KanMX4 spp381 ts::URA3*	This study	(YPH2622-2D-4B) This study
Saccharomyces cerevisiae: Strain background: BY4741 *MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; YDR007w::kanMX4*	Euroscarf	Y07202

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Saccharomyces cerevisiae: Strain background: CEN.PK *MATa his3Δ1 leu2-3,112 ura3-52 trp1-289 MAL2 8c SUC2*	Euroscarf	CEN.PK2
Saccharomyces cerevisiae: Strain background: CEN.PK *MATα his3Δ1 leu2-3,112 ura3-52 trp1 289 spp381Δ::LEU2*	Euroscarf	CEN.RO20-3A
Saccharomyces cerevisiae: Strain background: W303-1A *MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 mex67-5*	Jimeno et al., 2002	WMC1-1A
Oligonucleotides		
siRNA and DNA oligonucleotides are listed in Table S1	N/A	N/A
oligodT-Cy3 probes	Sigma-Aldrich	N/A
Recombinant DNA		
pEGFP-C2	Clontech	Cat# 6083-1
pEGFP-MFAP1	This study	N/A
pcDNA3	Adgene	N/A
pcDNA3-RNaseH1	(ten Asbroek et al., 2002)	N/A
pRS314L	(Prado et al., 1997)	N/A
pRS314-LYΔNS	(Prado et al., 1997)	N/A
pRS314L-lacZ	(Mumberg et al., 1994; Piruat and Aguilera, 1998)	N/A
pRS314GL-lacZ	(Mumberg et al., 1994; Piruat and Aguilera, 1998)	N/A
pCM184-RNH1	(Santos-Pereira et al., 2013)	N/A
pCM189-RNH1	(Castellano-Pozo et al., 2013)	N/A
pWJ1344	(Lisby et al., 2003)	N/A
Software and Algorithms		
Adophe Photoshop	Adobe	N/A
PRISM software	Graphpad	N/A
Leica DM6000 microscope	Leica	N/A
Leica DFC365FX camera	Leica	N/A
MetaMorph v7.5.1.0	Molecular Probes	N/A
Comet-score (version 1.5) software/ TriTek CometScore Professional (version 1.0.1.36) software	TriTek Corporation	N/A

LEAD CONTACT AND MATERIALS AVAILABILITY

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

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell culture, siRNA and Plasmid Transfection

HeLa and HEK293T cells were cultured in Dulbecco's modified Eagle's medium (DMEM; GIBCO, Thermo Scientific, Waltham, MA) supplemented with 10% heat-inactivated fetal bovine serum at 37°C (5% CO2).

Yeast strains and plasmid used in this study are listed in Key Resources Table. All yeast strains were grown at 30°C, except for RNA Seq analysis, in which wild-type and *spp381-ts* cells were grown in YPAD media at mid-log phase at 26°C. After 1h of temperature shift from 26°C to 37°C cells were collected for RNA extraction.

METHOD DETAILS

siRNAs, plasmids and transfections

Transient transfection of siRNA was performed using DharmaFECT 1 (Dharmacon) according to the manufacturer's instructions. siRNA duplexes were obtained from Dharmacon (Table S1). Fugene (Promega, Madison, WI, USA) was used for plasmid transfection. Assays were performed 72 h after siRNA transfection and 24 h after plasmid transfection.

Plasmids used to transfect cells are listed in Key Resources Table. pEGFP-MFAP1, contains the full-length human MFAP1 cDNA (from IMAGE clone 6200464, LifeSciences) cloned into pEGFP-C2 fused to EGFP; pcDNA3-RNaseH1, contains the full-length RNase H1 cloned into pcDNA3 (ten Asbroek et al., 2002).

Co-immunoprecipitation (coIP)

For co-immunoprecipitation assays, whole-cell extracts from a 10 cm Petri dish of HEK293T cells at 80% confluence were used and coIPs were performed as described (Salas-Armenteros et al., 2017). When it is indicated, HEK293T cells were transfected with pEGFP-C2 or pEGFP-MFAP1 plasmids for 24 h. Briefly, cells were lysed in lysis buffer (10 mM Tris–HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 0.5% (vol/vol) NP-40, 1 mm PMSF, and protease inhibitor cocktail) for 30 min on ice and the lysate was centrifuged for 15 min at 12,000 g at 4°C. 1:10 of cell extract was used for the input. Then, the lysate was diluted with dilution buffer (10 mM Tris–HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 1 mM PMSF, and protease inhibitor cocktail) to obtain a 0.25% (vol/vol) final concentration of NP-40. For the immunoprecipitation, protein extract was incubated for 2 h at 4°C with 50 μl of Dynabeads Protein A (Invitrogen), previously conjugated with the specific antibody, or 40 μl of magnetic anti-GFP beads (GFP-trap, Chromotek). Beads were then washed twice with PBS and three times with dilution buffer with 0.2% NP-40. Finally, the precipitate was eluted by boiling the beads for 10 min in 2 × Laemmli loading buffer and analyzed by western blot. Incubation of one sample with IgG or protein extract from cells transfected with pEGFP-C2 were use as control.

Chromatin immunoprecipitation (ChIP)

HeLa cells (5×10^6) plated on 10-cm Petri dishes were crosslinked for 10 min with 1% formaldehyde and processed for ChIP as described (Salas-Armenteros et al., 2017).

Cells were resuspended in 2.5 mL of cell lysis buffer (5 mM PIPES pH 8, 85 mM KCl, 0.5% NP-40, 1 mM PMSF, and protease inhibitor cocktail). The lysate was then centrifuged and resuspended in 1 mL of nuclei lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris–HCl pH 8, 1 mM PMSF, and protease inhibitor cocktail). Chromatin fragmentation was performed on ice-cold water bath with fifteen pulses of 30 s on and 30 s off on the maximum intensity setting in Bioruptor (Diagenode). 1300 μl containing 25 μg of chromatin diluted with IP buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris–HCl pH 8, 167 mM NaCl) were used for input (100 μl) and immunoprecipitation (1200 μl).Chromatin was incubated overnight at 4°C with 7.5 μg of MFAP1 or THOC1 specific antibodies or IgG to calculate the background signal. After incubation with 30 μl of Dynabeads Protein G (Invitrogen) for 2 h at 4°C, beads were washed once with wash buffer 1 (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris–HCl pH 8, 150 mM NaCl), once with wash buffer 2 (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris–HCl pH 8, 500 mM NaCl), once with wash buffer 3 (0.25 M LiCl, 1% NP-40, 1% sodium deoxycholate, 1 mM EDTA, 10 mM Tris–HCl pH 8), and twice with 1 × TE. The precipitated was eluted in TE-1% SDS for 20 min at 65°C. Input and immunoprecipitate were then un-crosslinked overnight at 65°C in TE-1% SDS. After proteinase K treatment for 1 h at 45°C, NucleoSpin Gel and PCR Clean-up kit (Macherey- Nagel) was used to isolate the DNA.

The amount of DNA was determined by real-time qPCR. DNA amount expressed in arbitrary units (A. U.) was calculated based on a standard DNA curve. Signal values in the different regions were plotted as the ratio between the immunoprecipitated DNA subtracting the background signal (IP) and the total amount of DNA (input) of each region. Signal values in the different regions were plotted as the ratio between the DNA amount immunoprecipitated subtracting the background signal (IP) and the total amount of DNA (input) of each region. Primers used for real-time quantitative PCR (qPCR) are described in Table S1.

Immunofluorescence

Cells were cultured on glass coverslips, pre-permeabilized with cold 0.1% triton in PBS on ice for 1 min and then fixed in 2% formaldehyde in PBS for 20 min at RT and permeabilized with 70% ethanol for 5 min at −20°C, 5 min at 4°C, and washed twice in PBS. PBS-3% BSA was used as blocking solution and for antibodies dilution. The following antibody dilutions were used: anti-MFAP1 (1:250), anti-SC35 (1:500), anti-THOC1 (1:250), anti-SF3B1 (1:200), and secondary antibodies conjugated with Alexa Fluor (1:500) (Key Resources Table). DNA was stained with DAPI. Images were captured at 63 × magnification. For γH2AX foci, measurements were analyzed and processed with the MetaMorph v7.5.1.0 (Molecular Probes) image analysis software.

Proximity ligation assay (PLA)

The proximity ligation assay was performed as described (Bhatia et al., 2014; Söderberg et al., 2006) with reagents from Duolink *In Situ* Red Starter Kit (Sigma) in accordance with the manufacturer's instructions. Cells were cultured on glass coverslips pre-permeabilized with cold 0.1% Triton in PBS on ice for 1 min, fixed in 2% formaldehyde in PBS for 20 min and permeabilized with 70% ethanol for 5 min at 20°C. Coverslips were blocked with PBS-3% BSA for 1 h at RT, incubated with primary antibodies diluted in PBS-3%BSA

for 2 h at RT, washed three times in PBS for 5 min and incubated with PLA probes for 1 h at 37°C. After two washed in wash buffer A for 5 min, ligation reaction was performed for 30 min at 37°C. Then, cells were washed twice in wash buffer A, incubated with the amplification reaction for 100 min at 37°C in darkness, washed twice in buffer B for 10 min and once in 0.01X wash buffer B for 1 min. Finally, coverslips were dried in darkness, mounted with mounting medium with DAPI. For negative controls, everything was performed identically, except that only one of the primary antibodies was added. Images were captured at 63 × magnification.

Single-cell electrophoresis

Alkaline and neutral single-cell electrophoresis or comet assays were performed as described (Domínguez-Sánchez et al., 2011; Salas-Armenteros et al., 2017) using a commercial kit (Trevigen, Gaithersburg, MD, USA) following the manufacturer's protocol. In brief, cells were collected using accutase, washed and resuspended in ice cold 1X PBS, combined with low melting agarose, immobilized on CometSlides (30 min at 4°C, until agarose is solidified). For alkaline single-cell electrophoresis, lysis was performed for 30 min at 4°C in lysis buffer. Then, DNA was unwound and denatured in freshly prepared alkaline unwinding solution pH>13 (200 mM NaOH, 1 mM EDTA) for 30 min at RT and electrophoresis was performed in prechilled alkaline electrophoresis solution pH>13 (200 mM NaOH, 1 mM EDTA) at 21 V for 30 min. Next, slides were immersed twice in dH_2O for 5 min each, then in 70% ethanol for 5 min and dried at RT. DNA was stained with SYBR Green at 4°C for 5 min. For neutral single-cell electrophoresis, cells were lysed for 1 h at 4°C and immersed in prechilled 1X neutral electrophoresis buffer (0.1M Tris, 0.3M Sodium acetate, pH 9 adjusted with glacial acetic acid) for 30 min at 4°C. Electrophoresis was performed in prechilled 1X neutral electrophoresis buffer at 35 V for 15 min and then immersed in DNA precipitation solution for 30 min at RT. Finally, slides were immersed in 70% ethanol for 30 min at RT and dried. DNA was stained with SYBR Green at 4°C for 30 min. When it is indicated, 50 μM cordycepin was added to the culture 4 h before starting the experiment. Comet slides were stained with SYBR Green. Images were captured at 10 × magnification. Comet-score (version 1.5) software or TriTek CometScore Professional (version 1.0.1.36) software were used to analyze comet tail moments. At least 100 cells were scored in each experiment and the median of tail moments was calculated.

Anaphase bridges and sister chromatin exchange (SCE) events assays

Analysis of anaphase bridges and SCE events were performed 72 h after siRNA transfection, as previously described (Tumini et al., 2016). To score anaphase bridges, HeLa cells were cultured on glass coverslips and transfected with siRNA for 72 h. In order to increase the proportion of cells in mitosis, cells were treated with 50 ng/ml nocodazol (Sigma) in complete medium for 4 h at 37°C, washed twice carefully in PBS and incubated for 1 h in 2 mL complete medium before to be harvested. Then, cells were fixed and permeabilized directly in the well adding 2 mL of 4% formaldehyde – 0.5% Triton X-100 in PBS during 20 min at RT. After immersing the coverslips carefully in PBS vectashield mounting medium with DAPI was used for the mounting. The percentage of cells in anaphase with anaphase bridges was calculated. At least 100 anaphases from each experiment were analyzed. For SCE assay, 24 hours after siRNA transfection HeLa cells were incubated with 10 μM BrdU for 42 hours followed by 3 hours treatment with 0.1 μg/ml of KaryoMAX colcemid solution (Invitrogen). Cells were harvested by tripsinization, resuspended in 0.075 M KCl hypotonic solution 10 min at 37°C and fixed with Carnoy's fixative (3:1 methanol:acetic acid, 3 changes). Cells were then dropped onto slides and baked overnight at 65°C. For the differential staining of sister chromatids, chromosome preparation were immersed in Hoechst solution (20 μg/ml) for 20 min, then exposed for 1 hour to UVA irradiation in 2X SSC (0.3 M NaCl, 0.03 M trisodium citrate, pH 7). and incubated in 2X SSC at 60°C for 20 min. Slides were stained with Giemsa (1:20) in buffer Weise pH 6.8 (Millipore). At least 20 metaphases from each experiment were analyzed. Metaphases were scored using a 100 × objective.

Microscopy images acquisition and data analysis

A Leica DM6000 microscope equipped with a DFC390 camera (Leica) and the LAS AF software (Leica) were used for fluorescence microscopy images and data acquisition, respectively. Software used for image analysis are listed in Key Resources Table.

Cell proliferation, cell cycle and apoptosis assays

Cell proliferation, cell cycle and apoptosis assays were performed as previously described (Salas-Armenteros et al., 2017). For proliferation assays, after 48 h of siRNA transfection cells were detached using accutase, counted and plated in wells of 96-well plates (2000 cells/well) in a final volume of 100 μl. Once cells were attached on the plate (7 h later), the first quantification was performed and was considered the day 0. Cells were transfected again 3 days after first transfection with 2/5 of siRNA standard amount. For colorimetric determination of viable cell numbers the cell proliferation reagent WST-1 reagent (Roche) was used following the manufacture's protocol. Cells were incubated with WST-1 reagents (10 μl/well) for 2 h at 37°C and 5% CO_2 in darkness. After this incubation period, absorbance was measured at 450 nm with a reference wavelength of 690 nm. For this, the scanning multi-well spectrophotometer ELISA reader VARIOSKAN FLASH (Thermo) was used. The measured absorbance directly correlates to the number of viable cells. Measurements were performed each 24 hours. Absorbance values were normalized to the value at the day 0 and represented as arbitrary units (A.U.). For cell cycle and FACS analysis, HeLa cells were incubated with EdU (20 μM) for 20 min 72 h after siRNA transfection to allow EdU incorporation in replicating cells. Cells were harvested with accutase, washed with PBS and centrifuged (10000 rpm 10 s). Pellet was resuspended in 300 μL cold PBS and cells were fixed by adding 700 μL cold 96% ethanol drop by drop and incubated at least 1 h on ice. After centrifugation (10000 rpm) ethanol was removed and 800 μL of 0.5% Triton X-100 in PBS was added to the pellet, mixed by vortexing and incubated on ice for 15 min. Then, cells were centrifuged (10000 rpm), washed

in 800 μL of 1% BSA, 0.5% Tween-20 in PBS and centrifuged again. Click-iT assay kit (Thermo Fisher Scientific) was used for EdU incorporation detection, using the manufacturer's protocol. For this, pellet was resuspended in 100 μL Click-iT reaction, incubated for 30 min at RT in darkness and washed three times in 800 μL of 1% BSA, 0.5% Tween-20 in PBS. For RNA degradation and DNA stain, cells were incubated with 700 μL of PBS with 7 μL RNase A 10 mg/ml and 5 μL of 7-Amino-Actinomycin D (7-AAD) at least for 30 min before being scored by FACS. G1, S and G2-phase cell were analyzed base on their DNA content and EdU signal by FACS. When it is indicated, cells were synchronized using a double thymidine block. For this, 24 h after siRNA transfection, cells were incubated with fresh medium containing 2 mM thymidine for 19 h. Then, cells were washed three times with PBS and were released in fresh complete medium for 7 h. Afterward, medium was replaced with fresh medium containing 2 mM thymidine and cells were incubated for 17–19 h. After three washed in PBS, cell were released in fresh complete medium and time points were collected. 20 min before collect the samples, cells were incubated with EdU (20 μM) for 20 min and processed as described above. To measure apoptotic cell population by analysis of sub-G1 DNA content, floating and adherent siRNA-depleted HeLa cells from two well of a 6-well plate were harvested, washed with cold PBS, centrifuged at 3500 rpm for 2 min and resuspended in 100 μL PBS. For fixation, 900 μL of cold 70% ethanol was added drop by drop during vortexing and cells were incubated at least 5 minutes at 4°C. Afterward, cells were centrifuged at 3500 rpm for 2 min, washed with 2 mL PBS, resuspended in 250 μL PBS and 250 μL of DNA extraction solution (0.2 M Na_2HPO_4, 4 mM Citric Acid, pH 7.8) was added. After incubation for 10 min at 37°C, cells were centrifuged and pellet was resuspended in 300 μL of PI/RNase A solution (100 μg/ml RNase A, 5 μg/ml Propidium Iodide in PBS) and incubated for 30 min in darkness. Quantitative analyses of different cell populations were carried out in a FACScan cytometer using the Cell Quest software (BD Biosciences). At least 5000 cells were scored by FACS and the percentage of cells in G1, S and G1-phase or cells with sub-G1 DNA content was calculated.

mRNA quantification

Total RNA was extracted using RNeasy Mini Kit (QIAGEN) and cDNA was synthesized by reverse transcription using QuantiTect Reverse transcription (QIAGEN). Relative mRNA levels in human cells were determined using relative qPCRs. A 7500 Fast Real-Time PCR system (Applied Biosystems, Carlsbad, CA) was used to perform real-time qPCRs (qPCRs). mRNA expression values of MFAP1 gene was normalized to the expression of the HPRT housekeeping gene. Mean and SEM of relative signal values normalized with respect to the siC control are provided. Primers used in real-time qPCR for mRNA quantification are listed in Table S1. The ratio of β-actin unspliced RNA (exon-intron primers) versus the spliced RNA (exon-exon primers) for different introns was analyzed by reverse transcription-quantitative PCR (RT-qPCR) in siC and siMFAP1 cells. Primers used are listed in Table S1.

Microarray analyses of gene expression and alternative splicing

Gene expression and alternative splicing microarray was performed using the Affymetrix platform GeneChip Human Transcriptome array 2.0 (HTA 2.0) (Key Resources Table) containing a median of 10 probes per exon or ncRNA and 4 probes per junction (> 6.0 million distinct probes in total covering coding and non-coding transcript) . 70% of the probes on this array cover exons for coding transcripts, and the remaining 30% of probes on the array cover exon-exon junctions, splice junctions and non-coding transcripts.

Total RNA was isolated in triplicate from HeLa siRNA transfected cells for 72 h by using the RNeasy Mini Kit (QIAGEN). The purity and quality of isolated RNA was confirmed by 2100 Bioanalyzer (Agilent Technologies). 100 ng of total RNA from each sample was used for amplification and production of biotinylated sense-strand cDNA from the entire expressed genome according to the GeneChip® WT PLUS Reagent Kit User Manual (P/N 703147 Rev. 3). Labeled ssDNA was hybridized to the GeneChip HTA Arrays (GeneChipTM Human Transcriptome Array 2.0) and scanned using the GeneChip Scanner 3000 7G according to the manufacture's recommendations.

Microarray data analysis

Data are available from the Gene Expression Omnibus database (accession number GSE119412) (Key Resources Table).

For each condition, microarray analysis was conducted in triplicate from three different biological replicates. Raw data were extracted from the scanned images and analyzed with the Affymetrix GeneChip Command Console® (AGCC) Software 2.0. Raw data were pre-processed and normalized using the SST (Signal Space transformation)-RMA (Robust Microarray Average) method (Irizarry et al., 2003). Data were further processed using Transcriptome Analysis Console (TAC) Software 3.0 from Affymetrix, which performs a gene-level analysis or an alternative splicing analysis.

For gene expression analysis, genes showing expression changes with P value < 0.05 (ANOVA) and a |linear fold change| > 1.5 were considered significant.

For alternative splicing analysis, splicing index (SI) was calculated as follows:

SI = Log2 [(Gene-set 1 Condition 1 /Gene 1 Condition 1 intensity)/(Gene-set 1 Condition Control Intensity/Gene1 Condition Control Intensity)] Condition 1 refers to siMFAP1 or siTHOC1 depleted cells and Condition control to siC depleted cells.

Splicing changes were considered to be significant when a gene is expressed in both conditions, A PRS/Junction must be expressed in at least one condition, a gene must contain at least one PRS, |Splicing Index (SI)| was > 2 and P value < 0.01 (ANOVA).

Gene set enrichment analysis

Gene Ontology (GO) analyses of enriched biological processes (BPs) or cellular component were performed using DAVID (Database for Annotation, Visualization and Integrated Discovery) bioinformatics resources 6.8 tool (https://david-d.ncifcrf.gov/summary.jsp). A P value < 0.01 or P value < 0.001 and fold enrichment > 1.5 was established to consider GO terms as significantly enriched in gene expression (microarrays and RNA-seq) and alternative splicing analyses, and was indicated in each case.

Transcription analysis by RNA-seq

Wild-type and spp381-ts cells were grown in YPAD media at mid-log phase at 26°C. After 1h of temperature shift from 26°C to 37°C cells were collected for RNA extraction. RNA was extracted with the RNeasyMini Kit (QIAGEN, Hilden, Germany) and treated with the RNase-Free DNase Set (QIAGEN) to remove any contaminating genomic DNA, according to the manufacturer's protocols. RNA samples were analyzed for quality with the RNA 6000 Nano assay on a 2100 Bioanalyzer (Agilent Technologies), and quantified with the Qubit RNA HS Assay (ThermoFisher Scientific).

RNA-Seq was performed by Genomics Core Facility of CABIMER. Sequencing libraries were performed from 150 ng RNA samples using the TruSeq® Stranded mRNA Library prep kit (Illumina) (Key Resources Table), and subsequently pooled and sequenced on Illumina NextSeq500 using 2 × 75-bp paired-end chemistry. Raw reads were filtered and trimmed with FASTQ toolkit 1.0.0 by BaseSpace Sequence Hub Illumina website and Fastq data were submitted to the Galaxy web platform for further analyses (Afgan et al., 2018). So, reads were mapped to the *S. cerevisiae* genome version sacCer3 with the HISAT2 v2.1.0 aligner (Kim et al., 2015), and filtered for high quality mapping (MapQuality > 30) using BAM tools (Barnett et al., 2011). Next, aligned reads were counted at gene and intron level by Feature Counts v1.6.2 (Liao et al., 2014) and differential expression analyses were performed using limma Bioconductor package (Law et al., 2014). The differentially expressed genes (DEGs) were selected as those with q-value threshold < 0.05 and a linear fold change value > 1.5. To identify intron retention (IR) event, the data for each intron was normalized by host gene expression and pairwise comparised between wt and spp381-ts within limma pipeline. The cutoff for filtering IR event was p value < 0.05 and fold-change value > 1.5

Data are available from the Gene Expression Omnibus database (accession number GSE128010).

Yeast strains and analysis of Rad52-foci and recombination frequencies

Yeast strains and plasmids used in this study are described in the Key Resources Table. YPH2622-2D-4B carries a termosensitive allele of SSP381 gene. This strain has a BY4741 genetic background and it was obtained by genetic crosses between YDR007W and YPH2622 strains. CEN.RO20-3A, is a deletion mutant (Euroscarf collection), that is viable, but has a lower growth rate compared to the wild-type. This deletion strain has a different genetic background CEN.PK, in which a number of mutations inviable in the BY genetic background are viable in this one.

Plasmids containing recombination systems were: pRS314L, pRS314-LYΔNS, pRS314L-lacZ, pRS314GL-lacZ. Plasmid for expression of RNH1: pCM184-RNH1 and pCM189-RNH1. Plasmid for DNA damage analysis: pWJ1344 containing the tagged RAD52-YFP fusion (Key Resources Table).

Analysis of Rad52-foci and recombination frequencies in *S. cerevisiae*

For Rad52 analysis, spontaneous Rad52-YFP foci in S/G2 cells from mid-log growing cultures carrying plasmid pWJ1344 or pWJ1213 were visualized and counted by fluorescence microscopy, as previously described (Lisby et al., 2003). More than 200 S/G2 cells from at least three independent transformants were analyzed. More than 200 S/G2 cells were analyzed in each experiment.

Recombination frequencies were determined by fluctuation tests (Gómez-González et al., 2011b). Yeast cells were transformed with the indicated plasmids and were grown in selective synthetic medium for 3-5 days at 30°C. For each strain, the recombination frequencies are given as the average and standard deviation of the median recombination frequency value obtained from fluctuation tests performed in 3–4 different transformants using 6 independent colonies per transformant. Recombinants were selected as Leu+ colonies for the plasmid containing LEU2 truncated repeat systems.

FISH in *S. cerevisiae*

mRNA export was determined by *in situ* hybridization in yeast cells with oligodT-Cy3 probes (Key Resources Table) as previously described (Trcek et al., 2012) with the following modifications. Cells were spheroplasted with 1mg 20T zymoliase at 30°C 10 min. washed and placed on the coated slides. Samples were hybridized with a 100 ng/sample oligodT (40xT) labeled with the Cy3 fluorochrome (Sigma) in hybridization buffer with 20% formamide (2X SSC, 20% formamide /2.5 mg/ml BSA/10 mM VRC containing a total of 100 ng of Cy3-oligodT, 100 μg g of sonicated salmon sperm DNA and 100 μg g of E.coli tRNA) for at least 3h at 30°C. Immunofluorescence images were acquired using a Leica DM6000 wide-field microscope and Leica DFC365FX camera.

Miscellanea

Analysis of sensitivity to genotoxic agents, Western, and yeast cultures were performed using standard procedures.

QUANTIFICATION AND STATISTICAL ANALYSIS

For γH2AX foci, anaphase bridges, cell cycle, apoptosis, recombination and Rad52 foci analyses, paired Student's t test (parametric) was used. For single-cell electrophoresis assays, nonparametric Mann–Whitney U-test was used. Data were analyzed with EXCEL (Microsoft) or GraphPad Prism software. The statistical test used in each experiment is mentioned in the figure legend. In general, a P value < 0.05 was considered as statistically significant (****$p < 0.0001$; ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$). Specific replicate numbers (n) for each experiment can be found in the corresponding figure legends. In all figures, means are plotted and SEM is represented as error bars.

DATA AND CODE AVAILABILITY

The accession number for the microarray analysis of gene expression and alternative splicing of MFAP1- and THOC1-depleted human cells reported in this paper is GEO: GSE119412. The accession number for the transcription analysis by RNA-Seq of spp381-ts mutant reported in this paper is GEO: GSE128010.