R-loop-mediated genome instability in mRNA cleavage and polyadenylation mutants

Peter C. Stirling,[1,3] Yujia A. Chan,[1,3] Sean W. Minaker,[1,3] Maria J. Aristizabal,[2] Irene Barrett,[1] Payal Sipahimalani,[1] Michael S. Kobor,[2] and Philip Hieter[1,4]

[1]Michael Smith Laboratories, University of British Columbia, Vancouver, British Columbia V6T1Z4, Canada; [2]Centre for Molecular Medicine and Therapeutics, Vancouver, British Columbia V5Z4H4, Canada

Genome instability via RNA:DNA hybrid-mediated R loops has been observed in mutants involved in various aspects of transcription and RNA processing. The prevalence of this mechanism among essential chromosome instability (CIN) genes remains unclear. In a secondary screen for increased Rad52 foci in CIN mutants, representing ~25% of essential genes, we identified seven essential subunits of the mRNA cleavage and polyadenylation (mCP) machinery. Genome-wide analysis of fragile sites by chromatin immunoprecipitation (ChIP) and microarray (ChIP–chip) of phosphorylated H2A in these mutants supported a transcription-dependent mechanism of DNA damage characteristic of R loops. In parallel, we directly detected increased RNA:DNA hybrid formation in mCP mutants and demonstrated that CIN is suppressed by expression of the R-loop-degrading enzyme RNaseH. To investigate the conservation of CIN in mCP mutants, we focused on FIP1L1, the human ortholog of yeast *FIP1*, a conserved mCP component that is part of an oncogenic fusion in eosinophilic leukemia. We found that truncation fusions of yeast *FIP1* analogous to those in cancer cause loss of function and that siRNA knockdown of FIP1L1 in human cells increases DNA damage and chromosome breakage. Our findings illuminate how mCP maintains genome integrity by suppressing R-loop formation and suggest that this function may be relevant to certain human cancers.

Genome instability is seen in most human cancers in the form of either increased mutation rate (microsatellite instability [MIN]) or increased chromosome instability (CIN) (Cahill et al. 1999; Schvartzman et al. 2010; Loeb 2011). The maintenance of genome integrity may play the role of a tumor suppressor that prevents the accumulation of mutations that ultimately drive uncontrolled cell proliferation (Schvartzman et al. 2010; Loeb 2011). Since CIN can be thought of as an early event in the oncogenic process, understanding the molecular basis of this phenotype has implications for cancer therapy, detection, and classification. CIN describes the phenomenon in which daughter cells inherit different complements of genetic material; i.e., loss or gain of whole chromosomes or chromosomal fragments. The result of CIN, when not lethal, can be aneuploidy or polyploidy, which have been linked to tumor formation (Fujiwara et al. 2005). Alterations in diverse cellular pathways can lead to CIN and tumor predisposition, especially disruption of normal DNA repair (e.g., BRCA1, ATM, BLM, and ATAD5) or mitotic functions (e.g., BUB1 and CENP-A) (Jackson and Bartek 2009; Schvartzman et al. 2010; Bell et al. 2011). Despite considerable effort, the spectrum of mutations that cause CIN in tumors is only partly known. Even as next-generation sequencing technology rapidly expands the known repertoire of somatic variants in tumors, ascribing functions to these variants remains a daunting task.

In the model eukaryote Baker's yeast (*Saccharomyces cerevisiae*), genome-wide screens have culminated in an extensive compilation of nearly 700 CIN genes across an array of cellular processes (Stirling et al. 2011). This list is a valuable resource for the study of candidate CIN genes in humans and also reveals many cellular components for which the mechanism of genome instability is obscure (Smith et al. 2004; Kanellis et al. 2007; Yuen et al. 2007; Andersen et al. 2008; Stirling et al. 2011). A widely recognized mechanism leading to CIN is excessive DNA damage or defective DNA repair (Jackson and Bartek 2009). DNA damage occurs primarily due to the perils of normal

DNA replication, where it tends to occur at fragile sites marked by H2A phosphorylation in yeast or H2AX phosphorylation in mammalian cells (i.e., mapped recently in Szilard et al. 2010). However, DNA damage can also occur as a result of a host of environmental factors, mutations, and intracellular stresses. This damage is repaired effectively by an extensive DNA repair machinery (Jackson and Bartek 2009). Many diverse pathways seem to converge on DNA damage, and it may be that a significant number of CIN mutations act through inappropriate DNA damage or repair (Alvaro et al. 2007; Jackson and Bartek 2009).

One more recently appreciated group of CIN genes—associated with increases in DNA damage, mutation, and hyperrecombination—are subsets of genes involved in transcription and RNA processing. Mutations in topoisomerase I, Sen1/SENATAXIN, THO/TREX, Sin3, and the SR protein-splicing factor ASF/SF2 have each been linked to genome instability via a common mechanism (Li and Manley 2005; Gomez-Gonzalez et al. 2009; El Hage et al. 2010; Mischo et al. 2011; Wahba et al. 2011). These mutants induce the formation of persistent, transcription-associated RNA:DNA hybrids that form R loops. The R-loop structures expose damage-prone ssDNA on the nonsense strand and may act as a block for replication fork progression, consistent with observations that mammalian transcription-associated recombination requires DNA replication and that R-loop-mediated genome instability in *Escherichia coli* is caused by replication fork collisions (Prado and Aguilera 2005; Gottipati et al. 2008; Gomez-Gonzalez et al. 2009; Gan et al. 2011). At present, the extent of cellular processes that contribute to R-loop-based genome instability is unclear.

To identify CIN processes that increase cellular demands on the DNA repair/recombination machinery, we performed a visual screen for Rad52-marked recombination centers in mutants of 305 essential CIN genes. This represents direct tests of mutants in >25% of essential genes. Remarkably, of 44 strains with increased Rad52 foci, we identified seven subunits of the mRNA cleavage and polyadenylation (mCP) machinery. These mCP proteins have been implicated in transcription elongation and termination and mRNA export due to their role in RNA processing (Brodsky and Silver 2000; Luna et al. 2005; Tous et al. 2011). Chromatin immunoprecipitation (ChIP) and microarray (ChIP–chip) using phosphorylation of H2A as a marker of DNA damage revealed fragile site differences between mCP mutants and wild type that map to a set of transcribed ORFs linked to replication origins, supporting a transcription-dependent mechanism for DNA damage. We directly detected RNA:DNA hybrids in mCP mutants and found that expression of RNaseH, which degrades RNA:DNA hybrids, rescues the chromosome transmission defect phenotype of these mutants. Finally, we showed that mutations of the mCP component *FIP1*, analogous to those seen in cancer, cause genome instability and that siRNA knockdown of the human ortholog FIP1L1 causes DNA damage and chromatid breaks in the colorectal cancer cell line HCT116. Together, our findings support a mechanism for CIN in mCP mutants involving transcription-coupled R-loop formation and suggest that this mechanism may be operating when human mCP genes are mutated in cancer.

Results

Rad52 foci screening of essential CIN genes

Rad52p is essential for homologous recombination and organizes into repair centers in response to double-strand breaks (DSBs) and other recombination events (Lisby et al. 2001; Mortensen et al. 2009). Increases in Rad52p foci can therefore indicate a number of genome-destabilizing conditions, including increased DSBs, inefficient resolution of recombination intermediates, or hyperrecombination (Alvaro et al. 2007). To determine which CIN mutants cause increased or prolonged engagement of the homologous recombination machinery, we introduced a *RAD52-YFP* fusion under the control of its native promoter by synthetic genetic array (SGA) (Tong et al. 2004). The resultant strains expressing a mutant CIN gene and *RAD52-YFP* were screened visually for increased levels of Rad52 foci (see the Materials and Methods). We screened 360 alleles of 305 essential CIN genes, including 306 ts (temperature-sensitive) and 54 DAmP alleles (Stirling et al. 2011). Similar to previous studies, we retested mutants in which ≥15% of cells had Rad52 foci in the primary screen (Alvaro et al. 2007). Triplicate retesting produced a list of 46 alleles in 44 unique genes whose mutation elicits an increased level of Rad52 foci (Fig. 1A). Eighteen additional mutants had increased but variable levels of Rad52 foci that did not meet our threshold across replicates (Supplemental Table S1). The largest functional group of mutants with increased Rad52 foci affected DNA replication, consistent with the S-phase function of Rad52 and its role in repairing damage caused by collapsed replication forks (Lisby et al. 2001). Interestingly, mutations in multiple genes involved in the proteasome, Smc5/6 complex, early secretion, transcription, and mRNA processing all caused increased rates of Rad52 foci (Fig. 1). These data stress the value of screening essential gene mutant collections, since a systematic screen of nonessential gene deletions uncovered a nonoverlapping set of biological processes (Alvaro et al. 2007).

Published genetic interactions exist between *rad52Δ* and 12 of 44 genes identified in our screen (i.e., BioGrid/DRYGIN databases). The presence of Rad52 foci in a mutant may indicate the dependence of that mutant on *RAD52* function (Alvaro et al. 2007). To further this hypothesis, we directly tested selected mutants from different biological pathways for genetic interactions with *rad52Δ*. For *pol31*-ts, *sld7*-ts, *rpn5-1*, *orc6*-ts, and *rna15-58* there were clear growth defects in the absence of *RAD52* (Fig. 1B). For *ssl1*^{T242I} and *pcf11-2*, a growth defect was apparent when a low level of hydroxyurea (HU) was present, whereas *nse3*-ts and *sec12-4* did not show obvious growth defects in this assay (Fig. 1B).

Overall, our screen indicates that a large proportion of essential CIN genes (i.e., 14%; 44 of 306 genes) exhibit significant levels of mutant-induced Rad52 foci and thus that disruption of unexpected cellular pathways create
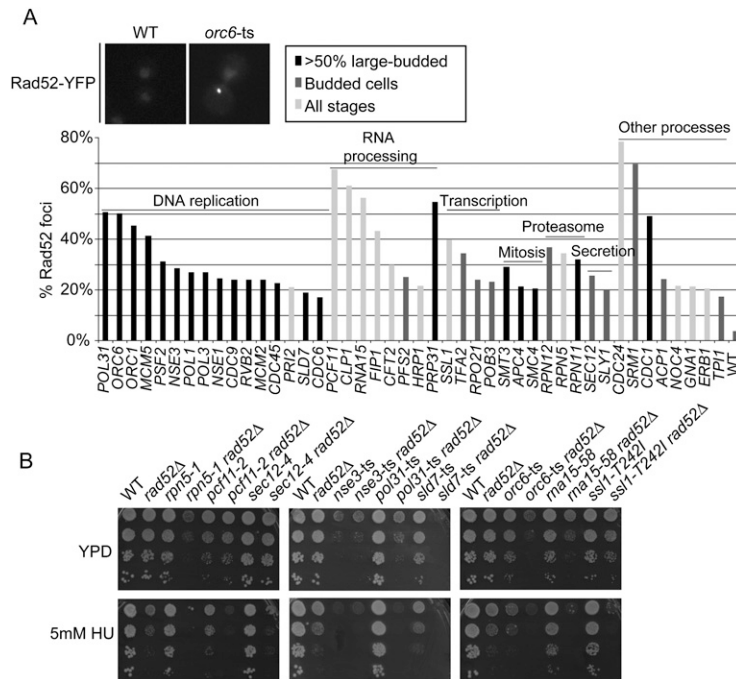
Figure 1. A screen for DNA damage foci in essential CIN genes. (*A*) Percentage of cells with Rad52 foci. Bars are color-coded to denote cell cycle arrest as large-budded cells (black), foci formation only in budded cells (dark gray), or foci formation at all stages (light gray). Multimember biological groups are labeled *above*. (*B*) Genetic interactions between foci-generating mutants and *rad52*Δ. Equal ODs of the indicated strains were serially diluted and spotted on YPD ± 5 mM HU at 30°C.

a requirement for homology-directed DNA repair. Of particular interest was the identification of seven mCP genes in the Rad52 foci screen (Fig. 1A). The mCP machinery is essential for processing nascent RNAs to free polyadenylated species (Gross and Moore 2001). Interestingly, unlike the majority of mutants identified in the screen, which arrested as large-budded G2/M cells, most mCP mutants exhibited Rad52 foci at all cell cycle stages, including G1. By comparison with those Rad52 foci-forming mutants involved in DNA replication, the role of mCP genes in genome integrity is not well characterized, and we decided to characterize it further.

Fragile sites in mCP mutants differ in transcribed ORFs and near replication origins

Phosphorylation of H2A-Ser129 is a mark of DNA damage analogous to human γ-H2AX modification in response to DNA damage and is critical for DSB repair (Downs et al. 2000). Recent analysis of sites of H2A-S129 phosphorylation in the yeast genome revealed a set of fragile loci or "γ sites" using ChIP–chip (Szilard et al. 2010). It is known that *h2a*-S129A mutants accumulate Rad52 foci and that both Rad52 and H2A phosphorylation contribute to efficient DNA repair (Downs et al. 2000; Mortensen et al. 2009; Szilard et al. 2010). We reasoned that mapping differences in fragile sites between wild type and mCP mutants could suggest a mechanism for CIN in the mutants. We performed γ-H2A ChIP–chip on the two mCP alleles with the highest levels of Rad52 foci: *pcf11-2* and *clp1-ts*. We applied double-T7-amplified anti-phospho-Ser129-H2A chromatin immunoprecipitates from log phase wild-type, mCP mutant, or *h2a*-S129A control cells to a high-density tiling microarray containing 3.2 million probes

with an average 5-base-pair (bp) resolution and 20-bp overlap between probes (Schulze et al. 2009). Normalizing our data to the *h2a*-S129A control produced a profile of yeast γ sites (Fig. 2A; see also Supplemental Table S2 for correlations; Supplemental Table S3 for all ChIP–chip peak coordinates).

Overall, our wild-type ChIP–chip profiles confirmed the findings of Szilard et al. (2010); for example, identifying γ sites enriched in subtelomeric regions, replication origins, long terminal repeats, repressed ORFs, MAT, and near rDNA and centromeres (Fig. 2A; Supplemental Figs. S1, S2; Supplemental Material). Initial examination of the mutant data revealed similar γ-site peak profiles in replicates of *clp1*-ts, *pcf11-2*, and wild-type strains (Fig. 2A; Supplemental Table S2). For example, Figure 2B shows similar mutant and wild-type γ sites associated with a repressed gene, *HXT10*, and a replication origin in the indicated 50-kb segment of chromosome VI. To highlight regions of the mutant-specific γ-H2A signal, which could be linked to mutant-induced DNA damage, we normalized the mutant profiles to that of the wild type (Fig. 2C). This analysis revealed reproducible and widespread differences in γ sites adjacent to replication origins (e.g., ARS603.5 in Fig. 2C).

To quantify the observed connection between mCP mutant γ-site enhancement and autonomous replicating sequence (ARS)-linked sites, we scanned a 2-kb window on either side of all ARSs for enriched regions ≥500 bp in length. This analysis found significantly more sites than would be expected at random for both *clp1*-ts and *pcf11-2* mutants (i.e., *P* = 0.99997) (Fig. 3A compares the observed values with the predicted number of peaks if ARS start coordinates were randomized). Sixty-four ARSs met these stringent criteria for three of four mutant replicates (i.e.,
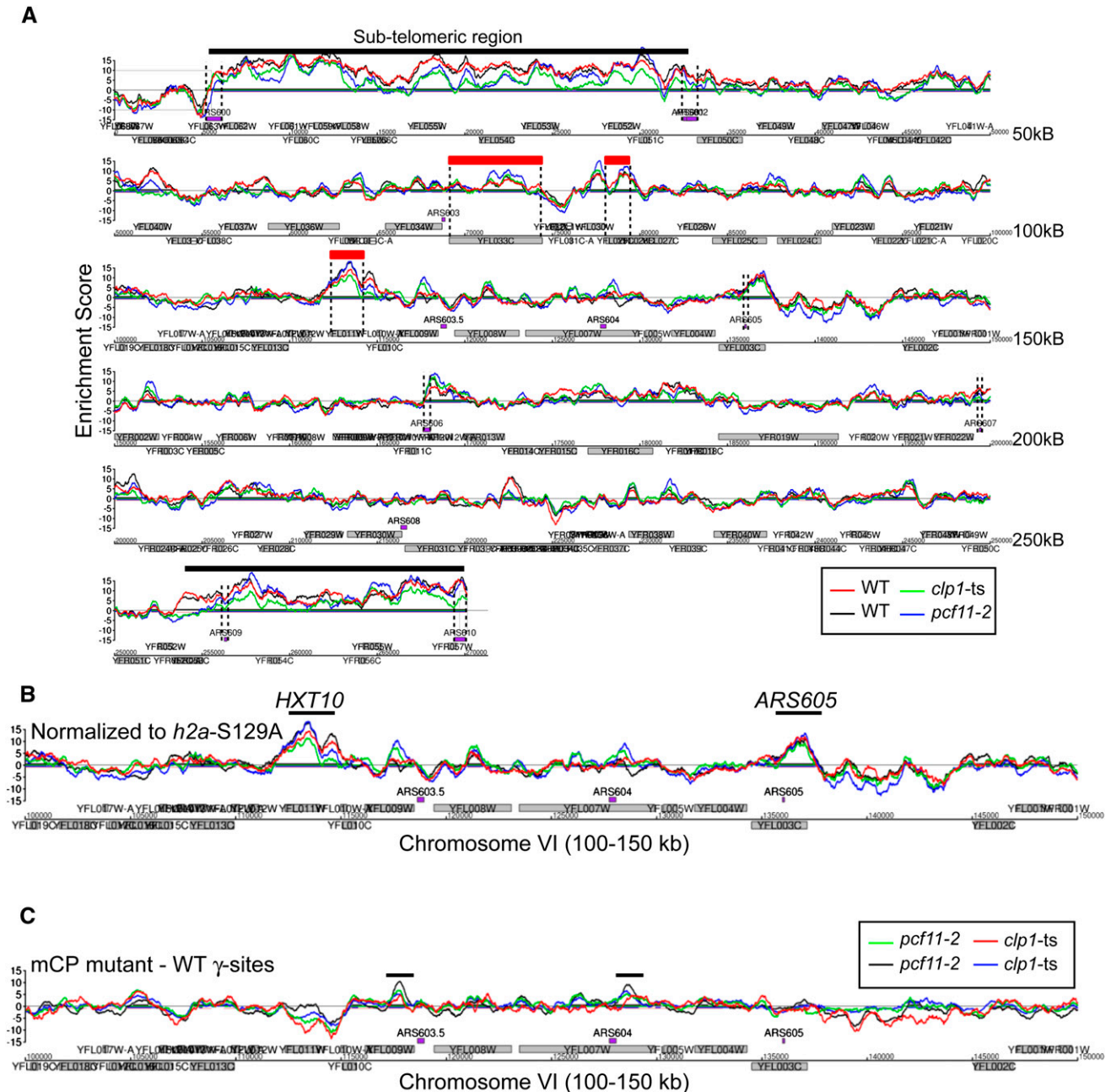
Figure 2. High-resolution mapping of yeast γ sites. (*A*) Overview of γ sites for Chromosome VI (ChrVI), indicating chromosomal features. (Purple box) ARS; (gray box) ORF. For *A* and *B*, the colored traces represent duplicates of wild type (red and black) and one replicate each of *clp1*-ts (green) and *pcf11-2* (blue). In *A*, large subtelomeric γ-site regions are noted with black bars, seven enriched ARS are noted with vertical dotted lines, and three examples of repressed genes (less than one mRNA per hour) are noted with dotted lines connecting to red boxes to define the gene boundaries. For a detailed comparison with published γ sites, see the Supplemental Material. (*B*) Representative *pcf11-2* and *clp1*-ts γ-site profiles for a 50-kB segment of ChrVI from the third row in *A*. Enriched γ sites at *HXT10*, which is repressed in glucose, and *ARS605* are indicated. (*C*) Replicates of *pcf11-2* (black, green) and *clp1*-ts (blue, red) γ-site profiles normalized to wild type to identify regions of enhanced signal common to both mutants. The same 50-kB region of ChrVI from *B* is shown after normalization to the wild-type γ-site profile. Common differences are marked with a black bar.

the sites are found at least once in both *clp1-ts* and *pcf11-2* and are found in two replicates of either *clp1-ts* or *pcf11-2*) (Fig. 3A; Supplemental Table S4). Very few ARSs were themselves enriched in the mCP-specific γ sites; instead, the peaks appeared in ORFs in close proximity to the ARS.

The peaks of 60 of the 64 ARS-linked sites could be assigned to single genes. Sixty-five percent of these genes were oriented toward the ARS, similar to the genome-wide frequency of ORF–ARS collision (60%). Notably, ARS-linked γ-H2A-enriched ORFs oriented in such a colliding
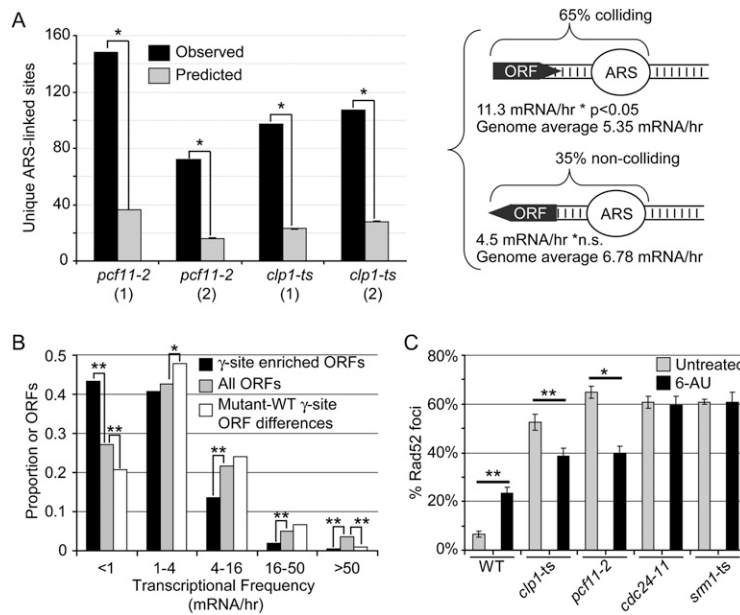
Figure 3. *clp1*-ts and *pcf11-2* mutant γ sites link DNA damage to transcription. (*A*) Linkage of mutant-specific peaks to replication origins. (*Left* panel) Peaks within 2 kb of replication origins were identified in the mutant-WT difference profiles and compared with predicted values generated by Monte Carlo simulation of randomized ARS positions. (*) $P < 0.0001$. The *right* panel shows the relative orientation, with respect to ARS, of ORFs encompassed by ARS-linked γ-site differences across three of four mutant replicates (pointed end indicates the direction of transcription). The average transcriptional frequency of ARS-colliding ORFs in γ sites was significantly higher than for all ARS-colliding ORFs across the entire genome ($P < 0.05$). (*B*) Transcriptional frequency of ORFs with an enhanced γ-H2A signal in wild-type cells and mCP mutant normalized to wild-type samples. The distribution of transcription frequencies (Holstege et al. 1998) is shown for all ORFs (gray bars), ORFs covered at least 25% by a wild-type γ site (black bars), and ORFs covered at least 10% by a γ-site difference in *clp1*-ts and *pcf11-2* ChIP–chip profiles. (*C*) Effect of the transcription inhibitor 6-azauracil on Rad52 foci formation in wild-type and mutant strains. Asterisks in *B* indicate significant deviations from all ORFs within a transcription frequency category (hypergeometric test), and asterisks in *C* indicate significant differences in Rad52 foci levels; Student's *t*-test, (*) $P < 0.05$; (**) $P < 0.005$.

fashion had a higher average transcriptional frequency than the complete set of colliding ORF–ARS pairs, while this trend was not seen for noncolliding ORF–ARS pairs (Fig. 3A, right panel; Supplemental Table S4). These observations are consistent with a mechanism contingent on R-loop formation, since they implicate the collision of transcription and DNA replication as a source of DNA damage in mCP mutants.

In addition to the ARS-linked γ sites, there was mutant-specific γ-H2A enrichment at a set of 918 ORFs throughout the genome. When we examined the transcriptional frequency of these 918 ORFs, we found significantly fewer of the lowest transcription category (i.e., less than one mRNA per hour) compared with all ORFs (Fig. 3B; Holstege et al. 1998). This represents a dramatic shift from the wild-type γ sites, which strongly enriched repressed or weakly transcribed genes and underrepresented higher transcriptional categories (Fig. 3B; Szilard et al. 2010). If we map the relative position of γ-site differences within the 918 ORFs, there is a slight bias toward the 3′ end of the gene for both *pcf11* and *clp1*-ts mutants compared with enriched ORFs in the global γ-site profile of wild-type cells (Supplemental Fig. S3). Moreover, examining the 55 genes whose 3′ untranslated region (UTR) contains a γ site in both *clp1*-ts and *pcf11-2* profiles (i.e., three of four replicates covered by >50% of a 500-bp window downstream) reveals that the associated genes tend to have higher transcriptional frequencies (Supplemental Fig. S3). Together, these data suggest that at least some transcription occurs at ORFs specifically enriched for γ-H2A in mCP mutants and that damage may be biased to the site of mCP function at 3′ ends. One potential reason that the bias is not more pronounced is that various 3′ end processing factors, including the mCP, also affect transcription

elongation and could be causing damage within genes as well as downstream (Tous et al. 2011).

If DNA damage in *pcf11-2* and *clp1*-ts mutants occurs at transcribed ORFs, it is plausible that reducing transcription could mitigate the increase in Rad52 foci. To test this hypothesis, we treated cells with 6-azauracil (6AU), an inhibitor of transcription elongation, and scored the presence of Rad52 foci. Wild-type cells showed an increase in Rad52 foci when treated with 6AU (Fig. 3C). Remarkably, both *clp1*-ts and *pcf11-2* strains showed a significant decrease in Rad52 foci when treated with 6AU (Fig. 3C). This effect was not seen in the strong hits *srm1*-ts or *cdc24-11*, which both retained high levels of Rad52 foci in 6AU (Fig. 3C). We conclude from this experiment that at least some of the DNA damage in mCP mutants is occurring within transcribed regions.

CIN in mCP mutants occurs via R loops

Our data show that mutants in the mCP pathway accumulate Rad52 foci, and our ChIP–chip data suggest a connection between DNA replication and the role of mCP in transcription. It is known that transcription can act as a replication fork barrier (Deshpande and Newlon 1996; Takeuchi et al. 2003) and that mutations in mCP components can cause transcription defects (Birse et al. 1998; Prado and Aguilera 2005). To investigate the connection of these phenotypes to CIN, we first confirmed reported mCP phenotypes in three standard CIN assays, investigating primarily chromosome loss (chromosome transmission fidelity [Ctf]), loss of heterozygosity (BiMater [BiM]), and rearrangement/gene conversion (a-like faker [ALF]). Mutants in *CFT2*, *CLP1*, *FIP1*, *PCF11*, and *RNA15* had a robust Ctf phenotype and a weak ALF phenotype; *RNA15*,

CLP1, and *PCF11* mutants also had a detectable BiM phenotype (Supplemental Table S5). Each mCP mutant was hypersensitive to the genotoxic agents HU, cisplatin, bleomycin, and UV (Fig. 4A; data not shown). These agents may synergize with the DNA damage that occurs in mCP mutants; for example, failed transcription termination in mCP mutants could act as a block to replication forks, thereby sensitizing cells to replication inhibitors like HU or cisplatin (Birse et al. 1998; Prado and Aguilera 2005). Indeed, HU and certain mCP mutations both increase recombination rates; thus, the effects may be additive (Luna et al. 2005).

Our data and the literature support the notion that transcriptional defects underlie genome instability in mCP mutants. One common way by which this occurs is through formation of transcription-coupled RNA:DNA hybrids called R loops. The identification of *pcf11-2* and *clp1*-ts γ-site differences in ARS-proximal genes, combined with the suppression of Rad52 foci by 6AU, makes this an especially attractive model (Fig. 3). Consistent with this explanation of our CIN data, we found that known R-loop-forming mutations in the helicase *SEN1* and the THO components *MFT1* and *THP2* lead to strong CTF and ALF phenotypes (Supplemental Table S5). To test the R-loop hypothesis for mCP mutants, we expressed a recombinant RNaseH enzyme that specifically degrades RNA:DNA hybrids. Analysis of Ctf in *pcf11*, *clp1*, *fip1*, and *rna15* mutant strains showed a remarkable suppression by RNaseH, which should reduce R loops, supporting a direct role for R loops in genome instability (Fig. 4B). Suppression of Ctf by RNaseH was not seen in Ctf control strains representing

other biological pathways (i.e., *rpf1-1* or *scc2-4* in Fig. 4B; quantified in Supplemental Fig. S4). Thus, RNaseH suppression of Ctf is highly specific to mCP mutants.

S. cerevisiae encodes at least two RNaseH enzymes encoded by *RNH1* and *RNH201*. Deletion of both RNaseH genes, which increases the frequency of R loops (see below), caused synthetic growth defects at a semipermissive temperature (32°C) in *CFT2*, *CLP1*, *PCF11*, or *RNA15* mutants (Fig. 4C). While we did not identify a growth defect associated with RNaseH deletion in *fip1-ts*, the triple-mutant strains did show enhanced sensitivity to low concentrations of bleomycin (Supplemental Fig. S5). These data indicate that mCP mutants require RNaseH activity under certain conditions to promote viability, consistent with a detrimental accumulation of RNA:DNA hybrids.

To confirm that the mCP mutants accumulate R loops, we transiently grew cells at the nonpermissive temperature of 37°C and harvested them for chromosome spreads. We performed immunofluorescent detection of RNA:DNA hybrids using the S9.6 monoclonal antibody in chromosome spreads of wild type; mCP mutants; known or predicted R-loop-forming mutants, including *rnh1Δrnh201Δ* double mutants; and a panel of Rad52 foci-forming strains from other biological pathways (Fig. 4D). For each allele of the mCP machinery tested, we observed an increased number of chromosome spreads with RNA:DNA hybrid staining. The *rnh1Δrnh201Δ* double mutant also showed high levels of RNA:DNA hybrids, confirming that the role of endogenous RNaseH is actively suppressing R-loop formation (Fig. 4D). Indeed, the *rnh1Δrnh201Δ* strain exhibits
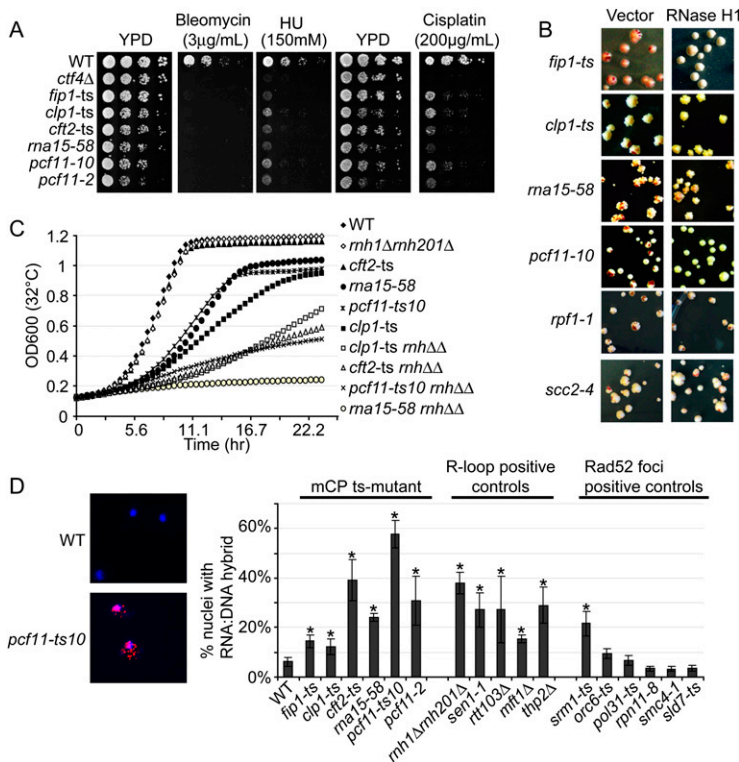

Figure 4. Transcription-coupled R loops are the likely cause of CIN in mCP mutants. (*A*) Tenfold serial dilution spot assays of mCP mutant strains on indicated medium. *ctf4Δ* is included as a sensitive control strain. (*Right*) An additional YPD control is included for cisplatin sensitivity because these plates have a different pH. (*B*) Plate images of Ctf phenotypes for selected mCP mutants and a control Ctf mutant expressing an empty vector or human RNaseH1. Significantly fewer colonies with red sectors (*P* < 0.05) are seen when RNaseH1 is expressed in mCP but not control Ctf mutants (see Supplemental Fig. S4). (*C*) Growth curves of mCP mutants with and without the yeast RNaseH genes *RNH1* and *RNH201* (*rnhΔΔ*). (*D*) Immunofluorescence of RNA:DNA hybrids in chromosome spreads. Representative spreads from wild-type and *pcf11-ts10* cells ([blue] DNA; [red] RNA:DNA hybrid) (*left* panel) and quantification of the immunofluorescence data for the indicated mutants (*right* panel) are shown. "R-loop-positive controls" indicates strains known or predicted to form R loops, and "Rad52 foci-positive controls" indicates strains that form Rad52 foci (Fig. 1) but are not predicted to form R loops. Significant differences from wild type (*P* < 0.05) are indicated by an asterisk.

a strong CIN phenotype (Supplemental Table S5). As predicted, mutants in the THO subunits *MFT1* and *THP2* and in *SEN1* caused RNA:DNA hybrids (Fig. 4D; Gomez-Gonzalez et al. 2009; Mischo et al. 2011). Deletion of the transcription termination factor *RTT103* also increased RNA:DNA hybrids (Fig. 4D). Among the panel of Rad52 foci-forming control strains, only *srm1*-ts showed an increase in RNA:DNA hybrids (Fig. 4D). *SRM1* plays a role in nucleocytoplasmic transport of mRNAs, which could account for the accumulation of RNA:DNA hybrids in *srm1*-ts mutants. As an antibody specificity control, we show that pretreatment of chromosome spreads with recombinant RNaseH in vitro significantly reduces the R-loop signal (Supplemental Fig. S6). Together, these data suggest that mCP mutants are similar to THO and *SEN1* mutants in that the mechanism of genome instability requires the accumulation of RNA:DNA hybrids.

Comparative analysis of mCP mutants with known R-loop-forming mutants

Known R-loop-forming mutants (e.g., THO complex and *SEN1*) work in parallel with mCP in the continuum of transcription termination and mRNA processing and export. Since mCP mutants have defects in transcription elongation and termination and mRNA processing and export, we wanted to address their similarity to known R-loop-forming mutants to probe the mechanism by which mCP mutant-induced R loops cause genome instability (Brodsky and Silver 2000; Luna et al. 2005; Tous et al. 2011). Mutants in THO and the *SEN1* helicase exhibit CIN phenotypes and RNA:DNA hybrid formation in our cytological assay (Fig. 4D; Supplemental Table S5). Moreover, representative alleles in mCP (i.e., *fip1-ts*, *clp1-ts*, *rna15-58*, and alleles of *pcf11*) exhibit synthetic growth defects or lethality with deletion of the THO subunit *THP2* or the termination factor *RTT103* or with the *sen1-1* ts allele, as assessed by tetrads and growth curve analysis (Supplemental Fig. S7). This analysis confirms that the cellular defects caused by our mCP mutants are buffered by the action of THO and *SEN1* and vice versa.

Another phenotype common to R-loop-forming mutations is hyperrecombination, and a weak phenotype has been reported for certain components of mCP (Chavez et al. 2000; Luna et al. 2005; Mischo et al. 2011). To clarify these observations, we subjected the mCP mutants to a well-characterized direct repeat recombination assay (Prado et al. 1997). At 30°C, the mCP mutants *clp1*-ts, *rna15-58*, and *pcf11-ts10* exhibited significant hyperrecombination (Supplemental Fig. S7). In this assay, mutations in *SEN1* or the THO complex lead to strong hyperrecombination phenotypes; however, the introduction of a transcription terminator after the first homology segment blocks hyperrecombination in THO mutants but stimulates recombination in the *sen1-1* allele (Mischo et al. 2011; Supplemental Fig. S7). Thus, while THO and *SEN1* mutants both induce R loops, this reporter system mechanistically differentiates RNA packaging defects (i.e., THO mutants) from *SEN1* helicase defects. Unlike *sen1-1*, hyperrecombination in mCP mutants is not increased by the tran-

scription terminator, but rather is generally reduced (Supplemental Fig. S7). Thus, mCP mutants behave like THO mutants, consistent with the known interdependence of THO and mCP function in mRNA processing and export (Luna et al. 2005).

Truncation fusion of yeast *FIP1* causes genome instability

Our data suggest that the mCP machinery prevents genome instability by suppressing the formation of RNA:DNA hybrids in a manner at least partially similar to the THO complex. The human ortholog of the mCP gene *FIP1*, FIP1L1, comprises the N-terminal moiety of a fusion protein with PDGFRα, which plays a causative role in 10%–20% of hypereosinophilic syndrome/eosinophilic leukemia (Cools et al. 2003; Gotlib and Cools 2008). These oncogenic fusion proteins truncate the C terminus of FIP1L1 and liberate the tyrosine kinase domain of PDGFRα, which becomes dysregulated. While the PDGFRα moiety is responsible for the proliferative phenotype, the impact of the truncation on FIP1L1 function is not clear (Stover et al. 2006). However, the fusion is under the control of the FIP1L1 promoter, and in vitro colony-forming assays indicate that FIP1L1 does have an effect on the fitness of transformed cells (Buitenhuis et al. 2007; Gotlib and Cools 2008). Thus, the role of FIP1L1 may be important to the etiology of this leukemia.

To model the putative effects of *FIP1* truncation fusions, we fused a commonly used globular domain (GFP) to two orthologous breakpoints in the yeast *FIP1* gene. The fusions Fip1$^{1\text{-}213}$-GFP and Fip1$^{1\text{-}279}$-GFP correspond to human exons 7 and 9, respectively, in the most closely related human FIP1L1 isoform, isoform 4 (Fig. 5A, left panel). Tetrad dissection of heterozygous diploids revealed that each truncation fusion affected fitness; Fip1$^{1\text{-}213}$-GFP did not support viability, whereas Fip1$^{1\text{-}279}$-GFP produced small colonies (Fig. 5A, right panel). A control in which GFP was fused to the C terminus of full-length *FIP1* did not appear to affect fitness. Since the Fip1$^{1\text{-}279}$-GFP construct was viable, we focused our subsequent characterization on this allele. Similar to the *fip1-ts* allele, we found that Fip1$^{1\text{-}279}$-GFP caused a Ctf phenotype not seen in the *FIP1*-GFP control (Figs. 4B, 5B). In addition, the truncation fusion protein increased formation of Rad52 foci and RNA:DNA hybrids, although in this case, the full-length *FIP1*-GFP also increased focus and hybrid formation above wild-type levels (Fig. 5C,D). This indicates that perturbations of *FIP1* function, even those that do not grossly affect fitness, can induce R-loop-mediated DNA damage and supports the idea that human FIP1L1, in the context of a PDGFRα fusion, may be only partly functional. Our data also predict that more severe truncations would further inactivate FIP1L1.

Human FIP1L1 may be required for genome stability

Our examination of CIN phenotypes in yeast is ultimately directed at predicting human genome integrity pathways (Yuen et al. 2007; Barber et al. 2008; Stirling et al. 2011). Therefore, an important confirmation of our data is a
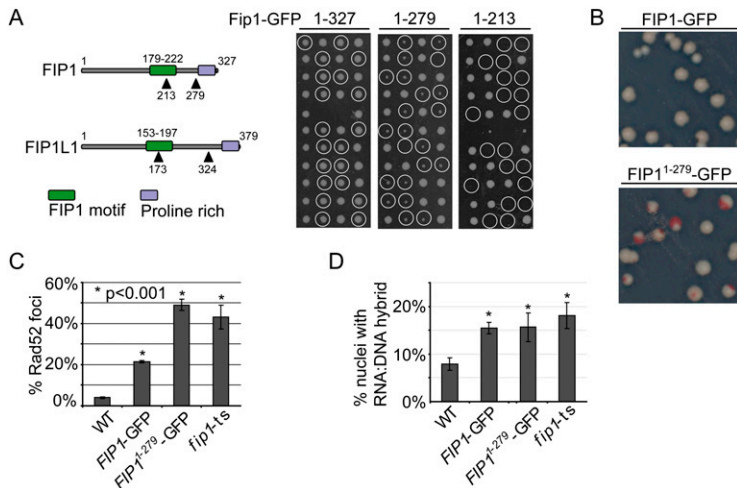
Figure 5. Truncation fusion of *FIP1* causes genome instability and DNA damage. (*A*) Tetrad dissection of truncation fusions of *FIP1* at two orthologous sites to breakpoints seen in human FIP1L1-PDGFRα-driven eosinophilic leukemia. (*Left*) The sites of truncation fusion are indicated in the schematic. (*Right*) Fusion-bearing His⁺ colonies are circled. (*B*) Ctf assay of full-length FIP1-GFP and FIP1¹⁻²⁷⁹-GFP. (*C*) Rad52-YFP foci measurements from *FIP1* mutant strains. (*D*) Quantification of RNA:DNA hybrid formation in *FIP1* mutant strains.

cross-species validation of the CIN phenotype. Partial reduction of yeast *FIP1* activity by a ts allele or truncation fusion leads to Rad52 foci and CIN in yeast cells (Figs. 1, 5; Supplemental Table S5). Thus, we predicted that reduction of human FIP1L1 function would also lead to genome instability in human cells.

We transiently reduced the protein levels of FIP1L1 in the near-diploid colorectal cell line HCT116 by transfection with pooled or, as a specificity control, individual siRNAs (Fig. 6A). We directly assessed chromosome integrity in cells with reduced FIP1L1 by analysis of mitotic metaphase chromosome spreads. When cells were treated with FIP1L1 siRNA, we reproducibly observed a significant increase in the appearance of chromosome breaks and chromosome fragments as compared with cells treated with a GAPDH control siRNA ($P < 0.05$) (Fig. 6A,B). The aberrant chromosomal structures resemble the damage seen in chromosome spreads from bleomycin-treated cells (data not shown). Consistently, high-content analysis of viability suggested that FIP1L1 siRNA treatment was toxic to cells (data not shown). In yeast, the CIN phenotype of *fip1-ts* is related to DNA damage (Figs. 1, 4). Therefore, reduction of FIP1L1 should also lead to DNA damage. In support of this, we found that FIP1L1 siRNA treatment increases the number of 53BP1-mCherry foci (Fig. 6C,D). 53BP1 foci mark sites of DNA damage and H2AX phosphorylation and have recently been shown to increase in human cells depleted for THO (Dominguez-Sanchez et al. 2011). These data support the hypothesis that, at least for FIP1L1, the essential roles of yeast mCP genes in suppressing DNA damage and genome instability are conserved to human cells.

Discussion

Rad52 foci screening of essential CIN genes

Focused secondary screens are a powerful way to define specific functions within primary functional genomic data (e.g., Doheny et al. 1993; Michaelis et al. 1997; Carroll et al. 2009; Stirling et al. 2011). By direct screening of more than one-quarter of essential genes, we linked 44 essential

CIN mutants to increased Rad52 foci formation, including seven mCP factors with an unappreciated role in genome integrity (Fig. 1). Rad52 foci indicate organization of the homologous recombination machinery into a repair center in response to DNA damage (Lisby et al. 2001). Our genetic data support the idea that increased Rad52 foci often indicate a requirement for Rad52 function (Fig. 1B; Alvaro et al. 2007).

Aside from the mCP mutants that were the focus of follow-up studies, the Rad52 foci screen identified mutants with diverse functions. The highest levels of foci were seen in *srm1-ts* and *cdc24-11* alleles (Fig. 1). *SRM1* strongly interacts with nucleosomes, and mutants cause gross chromosomal defects, although we also found evidence that *srm1*-ts may cause RNA:DNA hybrids, potentially due to its role in RNA export (Fig. 4; Koerber et al. 2009). *CDC24* encodes a guanine–nucleotide exchange factor for the small GTPase Cdc42, and while the mechanism of *cdc24-11* induced Rad52 foci is unknown, a recent high-content screen identified foci of the DNA damage marker Ddc2-GFP in alleles of *CDC24* (Li et al. 2011). Mutations affecting DNA replication were the class that most frequently exhibited enhanced Rad52 foci and usually arrested as large-budded G2/M cells (Fig. 1). DNA replication mutants were not significantly identified in the Rad52 foci screen of nonessential genes because so many of the components are essential. The same is true of essential transcription/mRNA processing, and proteasome subunits that were not identified in the literature (Alvaro et al. 2007). Conversely, Alvaro et al. (2007) identified DNA repair, mitotic checkpoint, and other genes not found in our survey of essential CIN mutants. Screening of both essential and nonessential gene mutants therefore provided complementary, nonoverlapping data sets.

Profiling fragile site changes in Rad52 foci-forming mRNA processing mutants

By comparing fragile sites in wild-type and mCP mutant yeast, we formulated a putative mechanism involving transcription-coupled R loops, which we confirmed by direct tests (Figs. 2–4). The robustness of the wild-type
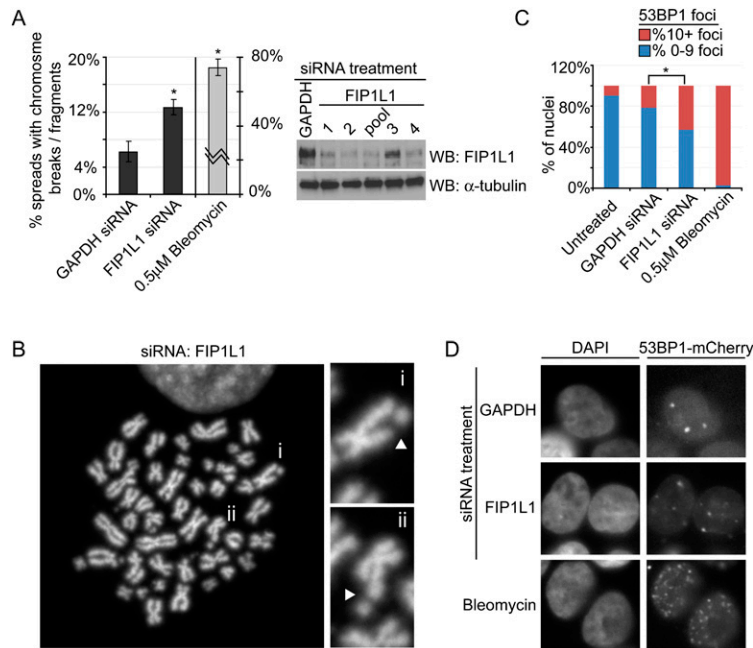
Figure 6. Loss of human FIP1L1 function causes CIN. (*A*, *left* panel) Chromatid breaks/fragments in mitotic chromosome spreads from siRNA- or bleomycin-treated cells (bleomycin value is shown on the right axis). The average value of replicate experiments is shown (error bars, SEM; [*] *P* < 0.05). (*Right* panel) Western blot of FIP1L1 in siRNA-treated HCT116 cells by single (FIP1L1) or pooled (FIP1L1 and GAPDH control) siRNAs. α-Tubulin was blotted as a loading control. (*B*) Representative chromosome spread indicating DNA breaks and fragments (white arrowheads in *inset*) associated with FIP1L1 siRNA. (*C*) 53BP1-mCherry foci per nucleus in bleomycin and GAPDH or FIP1L1 siRNA-treated cells ([*] mean foci/nucleus is significantly different, *P* < 0.001). Values from a representative of three data sets are shown. (*D*) Representative images of 53BP1-mCherry used to generate data in *C*.

γ-site profiles to perturbations by different mutants was somewhat surprising but similar to findings in strains lacking *RRM3*, a helicase that facilitates replisome processivity (Szilard et al. 2010). Despite the similarities in wild-type and mutant γ sites, we were able to extract meaningful differences within transcribed ORFs adjacent to ARSs and with a slight bias toward the 3′ end of the ORF (Figs. 2, 3; Supplemental Fig. S3). These observations are consistent with the function of mCP in transcription termination and 3′ end processing of pre-mRNAs and suggest that DNA damage may occur near the site of mCP function. Given the role of mCP in transcription, we were surprised that we did not recover the most highly transcribed genes; indeed, this category was underrepresented (Fig. 3B). It is possible that highly transcribed genes may have mechanisms in place to protect the genome from the deleterious effects of high transcription. This idea has been supported by work in mammalian cells showing that highly expressed genes orient codirectionally with replication origins (Huvet et al. 2007; Tuduri et al. 2009). Based on the success of our approach, we predict that mapping changes in fragile sites (e.g., with ChIP–chip of γ-H2A or another DNA repair protein) in other CIN mutant backgrounds could be an informative method to suggest basic mechanisms.

Mechanism of genome instability in mCP mutants

R loops underlie genome instability in a growing number of transcription and RNA processing mutants, including *SPT2*, *TOP1*, *SEN1*, the THO complex, and ASF/SF2 (Li and Manley 2005; Sikdar et al. 2008; Gomez-Gonzalez et al. 2009; Tuduri et al. 2009; El Hage et al. 2010; Mischo et al. 2011; Wahba et al. 2011). R loops have been linked to various genome instability phenotypes, including

hyperrecombination, replication defects, DNA damage, and chromosome loss. Our Rad52 screening and ChIP–chip studies suggested a mechanism for genome instability in mCP mutants that involves DNA damage induced by the formation of transcription-coupled R loops (Figs. 1–3). The fact that Rad52-foci are suppressed by reducing transcription with 6AU (Fig. 3D) shows that recombination is linked to transcription in the mCP mutants. We confirm the R-loop mechanism by rescuing Ctf with RNaseH, immunofluorescent detection of RNA:DNA hybrids, and demonstrating genetic interactions of mCP mutants with endogenous RNaseH mutants and known R-loop-forming THO/SEN1 mutants (Fig. 4; Supplemental Fig. S7).

mCP mutants have reported defects in transcription termination, elongation, and mRNA export (Birse et al. 1998; Brodsky and Silver 2000; Tous et al. 2011). More recently, mRNA 3′ end processing has been shown to directly influence transcription reinitiation (Mapendano et al. 2010). Therefore, several potential molecular mechanisms could explain the R-loop formation underlying genome instability in mCP mutants. Importantly, by comparing hyperrecombination phenotypes with mutants in THO and *SEN1*, we found that mCP mutants behave similarly to THO mutants (Supplemental Fig. S7). This is consistent with the established strong interdependence of THO and mRNA 3′ end processing factors (Luna et al. 2005). It is known that THO occupancy at gene 3′ ends is decreased by ~50% when mCP genes are mutated, and mCP occupancy at gene 3′ ends is significantly increased in THO mutants (Luna et al. 2005, Rougemaille et al. 2008). Thus, mCP acts upstream of THO in the biogenesis of mRNA export complexes, but THO is required to disassemble mCP machinery at 3′ ends. Therefore, our data best support a mechanism where mCP mutants facilitate RNA:DNA hybrid formation by reducing THO recruitment and/or

increasing the time that transcripts are present at gene 3′ ends due to inefficient termination.

While the molecular mechanism of R-loop formation and propagation is complex even for better-characterized mutants in THO and *SEN1*, we can also make a statement globally about the impact of R-loop-induced DNA damage in mCP mutants. Our ChIP–chip data show that mCP mutant-induced DNA damage occurs within transcribed genes and that the damaged 3′ UTRs tend to be associated with higher transcriptional frequencies (Fig. 3; Supplemental Fig. S3). More importantly, transcribed genes colliding with adjacent replication origins are enriched in DNA damage in mCP mutants, supporting a model in which collision with DNA replication is a major source of DNA damage in mCP mutants (Fig. 3; Supplemental Table S4). This is consistent with literature showing that collision of R loops with the replisome is commonly associated with genome instability (Prado and Aguilera 2005; Sikdar et al. 2008; Tuduri et al. 2009; Gan et al. 2011). In contrast, the fact that mCP mutants can increase Rad52 foci in G1 cells suggests that DNA replication is not strictly required to elicit foci, although it may play a major role (Fig. 1). It may be that mCP mutants cause lesions requiring Rad52 in more than one way, and it will be interesting in future studies to determine the relative contributions of transcription and replication to genome stability in these mutants. One confounding factor for the interpretation of these mechanisms is that we are studying a set of partial loss-of-function mutations in essential genes whose protein products form a multifunctional and dynamic protein complex. Therefore, despite their phenotypic overlap, there are certain to be gene and allele-specific effects within the mCP mutants described here that modulate the phenotypes observed.

Conservation of a CIN phenotype in the human cancer gene FIP1L1

Our data suggest that truncation fusion of human FIP1L1 in eosinophilic leukemia may cause a loss of FIP1L1 function based on homology with yeast (Fig. 5). By demonstrating that reduced FIP1L1 levels lead to increased DNA damage foci and chromosome breaks in mitosis (Fig. 6), we suggest a role for FIP1L1 in genome integrity that is conserved from yeast to humans. These data raise the possibility that genome stability in cancers with FIP1L1 truncation fusion proteins may be impacted by the partial loss of FIP1L1 function. FIP1L1 is dispensable for activation of the PDGFRα kinase activity, and therefore it has been proposed that some aspect of its promoter or the presence of an adjacent fragile site is responsible for the frequent observation of FIP1L1-PDGFRα gene fusions (Stover et al. 2006). Recent confirmation of FIP1L1-RARA gene fusions in acute promyelocytic leukemia shows that oncogenic translocation with FIP1L1 is not limited or specific to PDGFRα (Menezes et al. 2011). In addition, FIP1L1-PDGFRα fusions have been found in patients presenting with acute myeloid leukemia, which may have progressed from eosinophilic leukemia, and T-cell non-Hodgkins lymphoma, both of which are sometimes associated with an-

euploidy (Metzgeroth et al. 2007). Translocation may impact the function of another mCP component, human CLP1, which is associated with an oncogenic translocation in mixed-lineage leukemia (MLL) (Tanabe et al. 1996; Cools et al. 2003). We speculate that one outcome of interrupting FIP1L1 or CLP1 could be to destabilize the genome of these leukemias and, when left untreated, in combination with the proliferative phenotype, facilitate further evolution of the cancer genome. It is worth noting that heterozygous mutation of other CIN genes has been shown to increase tumor incidence in mouse models (e.g., Bub1, Aurora kinase, and CENP-E) (for review, see Schvartzman et al. 2010).

A genome-wide compilation of yeast CIN genes identified dozens of transcription and RNA processing-related CIN mutants whose R-loop status is unknown (Stirling et al. 2011). We predict that other CIN mutants in these pathways also induce R-loop formation. Interestingly, a high-throughput study of γ-H2AX DNA damage foci in mammalian cells identified many RNA processing genes, suggesting that this pathway is generally important for genome integrity (Paulsen et al. 2009). Moreover, these investigators found that overexpression of RNaseH could suppress γ-H2AX foci in some instances, supporting the R-loop model for diverse RNA processing pathways (Paulsen et al. 2009). The prevalence of R-loop-mediated CIN in cancer is unclear, although it is tempting to speculate that it could occur in cancers where the orthologs of known R-loop-forming mutants are disrupted (e.g., *TOP1* or *FIP1*) (Cools et al. 2003; Iwase et al. 2003).

Materials and methods

Yeast strains and plasmids

Yeast strains were grown in YPD or synthetic complete medium lacking the appropriate amino acid where nutritional selection was required. Serial dilution assays and growth curve analysis were performed as described (McLellan et al. 2009; Stirling et al. 2011). Supplemental Table S6 contains a list of yeast strains and plasmids.

Rad52-YFP foci screening and CIN assays

Rad52-YFP fusions were made by direct transformation and introduced into the relevant strains by SGA (Tong et al. 2004). Log phase cultures expressing Rad52-YFP were shifted for 3.5 h from 25°C to 37°C and then prepared for microscopy as described (Carroll et al. 2009). Both DIC and fluorescence images were collected with Metamorph (Molecular Devices) and analyzed using ImageJ (http://rsbweb.nih.gov/ij/index.html). In Figure 3, cells growing at 25°C were pretreated with 100 μg/mL 6AU for 20 min before shifting for 3.5 h to 37°C and were imaged as for untreated cells. For each putative hit, the experiment was triplicated and ≥100 cells were counted per replicate. Budding indices were derived from these images. CIN assays were performed exactly as described (Spencer et al. 1990; Yuen et al. 2007).

ChIP–chip

ChIP–chip experiments were performed as described (Schulze et al. 2009) using a ChIP-grade anti-phospho-Ser129 H2A (ab15083, Abcam), except ts alleles were inactivated by shifting for 3 h to

37°C before cross-linking. Data analysis and plotting of enriched features were performed using custom R, MatLab, and python scripts essentially as described (Schulze et al. 2009; Takahashi et al. 2011). All profiles were generated in duplicate, except the *h2a*-S129A control. The enrichment threshold was 2.5-fold for all analyses. Features were considered enriched if 25% of the feature length (base pair) met the enrichment threshold when normalized to *h2a*-s129A and if 10% of the feature met the enrichment criteria for mutants normalized to wild type.

ChIP–chip statistical analysis

To determine whether the data recovered more of a particular feature than would be predicted at random, a Monte Carlo simulation on the relevant data set was run using randomly generated start positions for a particular set of features (e.g., ARSs), while maintaining the total number of features within a chromosome (Schulze et al. 2009). Five-hundred simulations were run for each feature to generate a mean and standard deviation. These values were compared with the observed score using the cumulative normal distribution, which calculated the probability of seeing a lower score if a value was selected at random (http://www.stattrek.com). Transcriptional frequency categories and Rad52 foci counts were compared with controls using a Student's *t*-test (http://www.graphpad.com). Overall transcriptional frequencies for smaller samples were compared using the Mann-Whitney test. Significance levels are indicated in the text or figure legends.

Yeast chromosome spreads

Chromosome spreads were done essentially as described, except that cells were shifted for 4–6 h to 37°C before lysis (Hartman et al. 2000; Wahba et al. 2011). The S9.6 antibody (Hu et al. 2006) was used at 1 µg/mL in blocking buffer (1× PBS with 5% BSA, 0.2% skim milk powder) and detected with Cy3-conjugated goat anti-mouse antibody (Jackson Laboratories, #115-165-003). For RNaseH treatments, each spread was incubated with 2 U of RNaseH (Roche) in blocking buffer for 15 min prior to antibody treatment.

Mammalian tissue culture

The HCT116 cell line was cultured under standard conditions as previously described (Barber et al. 2008). siRNAs were obtained from Dharmacon and applied according to the manufacturer's instructions. Mitotic chromosome spreads were performed essentially as described (Barber et al. 2008). Briefly, actively growing cells were treated with 0.1 µg/mL colcemid for at least 90 min before washing, trypsinization, and addition of hypotonic (0.075 M KCl) for 5 min at room temperature. Cells were fixed in 3:1 methanol:glacial acetic acid, centrifuged (800 rpm for 5 min) and fixed again before mounting on clean slides, DAPI staining, and visualizing. 53BP1 foci were detected in paraformaldehyde-fixed HCT116 cells constitutively expressing 53BP1-mCherry. Foci were scored in Metamorph from three sets of images derived from two independent siRNA transfections.

References

Alvaro D, Lisby M, Rothstein R. 2007. Genome-wide analysis of Rad52 foci reveals diverse mechanisms impacting recombination. *PLoS Genet* **3:** e228. doi: 10.1371/journal.pgen.0030228.

Andersen MP, Nelson ZW, Hetrick ED, Gottschling DE. 2008. A genetic screen for increased loss of heterozygosity in *Saccharomyces cerevisiae*. *Genetics* **179:** 1179–1195.

Barber TD, McManus K, Yuen KW, Reis M, Parmigiani G, Shen D, Barrett I, Nouhi Y, Spencer F, Markowitz S, et al. 2008. Chromatid cohesion defects may underlie chromosome instability in human colorectal cancers. *Proc Natl Acad Sci* **105:** 3443–3448.

Bell DW, Sikdar N, Lee KY, Price JC, Chatterjee R, Park HD, Fox J, Ishiai M, Rudd ML, Pollock LM, et al. 2011. Predisposition to cancer caused by genetic and functional defects of Mammalian atad5. *PLoS Genet* **7:** e1002245. doi: 10.1371/journal.pgen.1002245.

Birse CE, Minvielle-Sebastia L, Lee BA, Keller W, Proudfoot NJ. 1998. Coupling termination of transcription to messenger RNA maturation in yeast. *Science* **280:** 298–301.

Brodsky AS, Silver PA. 2000. Pre-mRNA processing factors are required for nuclear export. *RNA* **6:** 1737–1749.

Buitenhuis M, Verhagen LP, Cools J, Coffer PJ. 2007. Molecular mechanisms underlying FIP1L1-PDGFRA-mediated myeloproliferation. *Cancer Res* **67:** 3759–3766.

Cahill DP, Kinzler KW, Vogelstein B, Lengauer C. 1999. Genetic instability and darwinian selection in tumours. *Trends Cell Biol* **9:** M57–M60. doi: 10.1016/S0962-8924(99)01661-X.

Carroll SY, Stirling PC, Stimpson HE, Giesselmann E, Schmitt MJ, Drubin DG. 2009. A yeast killer toxin screen provides insights into a/b toxin entry, trafficking, and killing mechanisms. *Dev Cell* **17:** 552–560.

Chavez S, Beilharz T, Rondon AG, Erdjument-Bromage H, Tempst P, Svejstrup JQ, Lithgow T, Aguilera A. 2000. A protein complex containing Tho2, Hpr1, Mft1 and a novel protein, Thp2, connects transcription elongation with mitotic recombination in *Saccharomyces cerevisiae*. *EMBO J* **19:** 5824–5834.

Cools J, DeAngelo DJ, Gotlib J, Stover EH, Legare RD, Cortes J, Kutok J, Clark J, Galinsky I, Griffin JD, et al. 2003. A tyrosine kinase created by fusion of the PDGFRA and FIP1L1 genes as a therapeutic target of imatinib in idiopathic hypereosinophilic syndrome. *N Engl J Med* **348:** 1201–1214.

Deshpande AM, Newlon CS. 1996. DNA replication fork pause sites dependent on transcription. *Science* **272:** 1030–1033.

Doheny KF, Sorger PK, Hyman AA, Tugendreich S, Spencer F, Hieter P. 1993. Identification of essential components of the *S. cerevisiae* kinetochore. *Cell* **73:** 761–774.

Dominguez-Sanchez MS, Barroso S, Gomez-Gonzalez B, Luna R, Aguilera A. 2011. Genome instability and transcription elongation impairment in human cells depleted of THO/TREX. *PLoS Genet* **7:** e1002386. doi: 10.1371/journal.pgen.1002386.

Downs JA, Lowndes NF, Jackson SP. 2000. A role for *Saccharomyces cerevisiae* histone H2A in DNA repair. *Nature* **408:** 1001–1004.

El Hage A, French SL, Beyer AL, Tollervey D. 2010. Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis. *Genes Dev* **24:** 1546–1558.

Fujiwara T, Bandi M, Nitta M, Ivanova EV, Bronson RT, Pellman D. 2005. Cytokinesis failure generating tetraploids promotes tumorigenesis in p53-null cells. *Nature* **437:** 1043–1047.

Gan W, Guan Z, Liu J, Gui T, Shen K, Manley JL, Li X. 2011. R-loop-mediated genomic instability is caused by impairment of replication fork progression. *Genes Dev* **25:** 2041–2056.

Gomez-Gonzalez B, Felipe-Abrio I, Aguilera A. 2009. The S-phase checkpoint is required to respond to R-loops accumulated in THO mutants. *Mol Cell Biol* **29:** 5203–5213.

Gotlib J, Cools J. 2008. Five years since the discovery of FIP1L1-PDGFRA: What we have learned about the fusion and other molecularly defined eosinophilias. *Leukemia* **22:** 1999–2010.

Gottipati P, Cassel TN, Savolainen L, Helleday T. 2008. Transcription-associated recombination is dependent on replication in Mammalian cells. *Mol Cell Biol* **28:** 154–164.

Gross S, Moore C. 2001. Five subunits are required for reconstitution of the cleavage and polyadenylation activities of *Saccharomyces cerevisiae* cleavage factor I. *Proc Natl Acad Sci* **98:** 6080–6085.

Hartman T, Stead K, Koshland D, Guacci V. 2000. Pds5p is an essential chromosomal protein required for both sister chromatid cohesion and condensation in *Saccharomyces cerevisiae*. *J Cell Biol* **151:** 613–626.

Holstege FC, Jennings EG, Wyrick JJ, Lee TI, Hengartner CJ, Green MR, Golub TR, Lander ES, Young RA. 1998. Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* **95:** 717–728.

Hu Z, Zhang A, Storz G, Gottesman S, Leppla SH. 2006. An antibody-based microarray assay for small RNA detection. *Nucleic Acids Res* **34:** e52. doi: 10.1093/nar/gkl142.

Huvet M, Nicolay S, Touchon M, Audit B, d'Aubenton-Carafa Y, Arneodo A, Thermes C. 2007. Human gene organization driven by the coordination of replication and transcription. *Genome Res* **17:** 1278–1285.

Iwase S, Akiyama N, Sekikawa T, Saito S, Arakawa Y, Horiguchi-Yamada J, Yamada H. 2003. Both NUP98/TOP1 and TOP1/NUP98 transcripts are detected in a de novo AML with t(11;20)(p15;q11). *Genes Chromosomes Cancer* **38:** 102–105.

Jackson SP, Bartek J. 2009. The DNA-damage response in human biology and disease. *Nature* **461:** 1071–1078.

Kanellis P, Gagliardi M, Banath JP, Szilard RK, Nakada S, Galicia S, Sweeney FD, Cabelof DC, Olive PL, Durocher D. 2007. A screen for suppressors of gross chromosomal rearrangements identifies a conserved role for PLP in preventing DNA lesions. *PLoS Genet* **3:** e134. doi: 10.1371/journal.pgen.0030134.

Koerber RT, Rhee HS, Jiang C, Pugh BF. 2009. Interaction of transcriptional regulators with specific nucleosomes across the *Saccharomyces* genome. *Mol Cell* **35:** 889–902.

Li X, Manley JL. 2005. Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. *Cell* **122:** 365–378.

Li Z, Vizeacoumar FJ, Bahr S, Li J, Warringer J, Vizeacoumar FS, Min R, Vandersluis B, Bellay J, Devit M, et al. 2011. Systematic exploration of essential yeast gene function with temperature-sensitive mutants. *Nat Biotechnol* **29:** 361–367.

Lisby M, Rothstein R, Mortensen UH. 2001. Rad52 forms DNA repair and recombination centers during S phase. *Proc Natl Acad Sci* **98:** 8276–8282.

Loeb LA. 2011. Human cancers express mutator phenotypes: Origin, consequences and targeting. *Nat Rev Cancer* **11:** 450–457.

Luna R, Jimeno S, Marin M, Huertas P, Garcia-Rubio M, Aguilera A. 2005. Interdependence between transcription and mRNP processing and export, and its impact on genetic stability. *Mol Cell* **18:** 711–722.

Mapendano CK, Lykke-Andersen S, Kjems J, Bertrand E, Jensen TH. 2010. Crosstalk between mRNA 3′ end processing and transcription initiation. *Mol Cell* **40:** 410–422.

McLellan J, O'Neil N, Tarailo S, Stoepel J, Bryan J, Rose A, Hieter P. 2009. Synthetic lethal genetic interactions that decrease somatic cell proliferation in *Caenorhabditis elegans* identify the alternative RFC CTF18 as a candidate cancer drug target. *Mol Biol Cell* **20:** 5306–5313.

Menezes J, Acquadro F, Perez-Pons de la Villa C, Garcia-Sanchez F, Alvarez S, Cigudosa JC. 2011. FIP1L1/RARA with breakpoint at FIP1L1 intron13: A variant translocation in acute promyelocytic leukemia. *Haematologica* **96:** 1565–1566.

Metzgeroth G, Walz C, Score J, Siebert R, Schnittger S, Haferlach C, Popp H, Haferlach T, Erben P, Mix J, et al. 2007. Recurrent finding of the FIP1L1-PDGFRA fusion gene in eosinophilia-associated acute myeloid leukemia and lymphoblastic T-cell lymphoma. *Leukemia* **21:** 1183–1188.

Michaelis C, Ciosk R, Nasmyth K. 1997. Cohesins: Chromosomal proteins that prevent premature separation of sister chromatids. *Cell* **91:** 35–45.

Mischo HE, Gomez-Gonzalez B, Grzechnik P, Rondon AG, Wei W, Steinmetz L, Aguilera A, Proudfoot NJ. 2011. Yeast Sen1 helicase protects the genome from transcription-associated instability. *Mol Cell* **41:** 21–32.

Mortensen UH, Lisby M, Rothstein R. 2009. Rad52. *Curr Biol* **19:** R676–R677. doi: 10.1016/j.cub.2009.06.001.

Paulsen RD, Soni DV, Wollman R, Hahn AT, Yee MC, Guan A, Hesley JA, Miller SC, Cromwell EF, Solow-Cordero DE, et al. 2009. A genome-wide siRNA screen reveals diverse cellular processes and pathways that mediate genome stability. *Mol Cell* **35:** 228–239.

Prado F, Aguilera A. 2005. Impairment of replication fork progression mediates RNA polII transcription-associated recombination. *EMBO J* **24:** 1267–1276.

Prado F, Piruat JI, Aguilera A. 1997. Recombination between DNA repeats in yeast hpr1δ cells is linked to transcription elongation. *EMBO J* **16:** 2826–2835.

Rougemaille M, Dieppois G, Kisseleva-Romanova E, Gudipati RK, Lemoine S, Blugeon C, Boulay J, Jensen TH, Stutz F, Devaux P, et al. 2008. THO/Sub2p functions to coordinate 3′-end processing with gene-nuclear pore association. *Cell* **135:** 308–321.

Schulze JM, Jackson J, Nakanishi S, Gardner JM, Hentrich T, Haug J, Johnston M, Jaspersen SL, Kobor MS, Shilatifard A. 2009. Linking cell cycle to histone modifications: SBF and H2B monoubiquitination machinery and cell-cycle regulation of H3K79 dimethylation. *Mol Cell* **35:** 626–641.

Schvartzman JM, Sotillo R, Benezra R. 2010. Mitotic chromosomal instability and cancer: Mouse modelling of the human disease. *Nat Rev Cancer* **10:** 102–115.

Sikdar N, Banerjee S, Zhang H, Smith S, Myung K. 2008. Spt2p defines a new transcription-dependent gross chromosomal rearrangement pathway. *PLoS Genet* **4:** e1000290. doi: 10.1371/journal.pgen.1000290.

Smith S, Hwang JY, Banerjee S, Majeed A, Gupta A, Myung K. 2004. Mutator genes for suppression of gross chromosomal rearrangements identified by a genome-wide screening in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci* **101:** 9039–9044.

Spencer F, Gerring SL, Connelly C, Hieter P. 1990. Mitotic chromosome transmission fidelity mutants in *Saccharomyces cerevisiae*. *Genetics* **124:** 237–249.

Stirling PC, Bloom MS, Solanki-Patil T, Smith S, Sipahimalani P, Li Z, Kofoed M, Ben-Aroya S, Myung K, Hieter P. 2011. The

complete spectrum of yeast chromosome instability genes identifies candidate CIN cancer genes and functional roles for ASTRA complex components. *PLoS Genet* **7:** e1002057. doi: 10.371/journal.pgen.1002057.

Stover EH, Chen J, Folens C, Lee BH, Mentens N, Marynen P, Williams IR, Gilliland DG, Cools J. 2006. Activation of FIP1L1-PDGFRα requires disruption of the juxtamembrane domain of PDGFRα and is FIP1L1-independent. *Proc Natl Acad Sci* **103:** 8078–8083.

Szilard RK, Jacques PE, Laramee L, Cheng B, Galicia S, Bataille AR, Yeung M, Mendez M, Bergeron M, Robert F, et al. 2010. Systematic identification of fragile sites via genome-wide location analysis of γ-H2AX. *Nat Struct Mol Biol* **17:** 299–305.

Takahashi YH, Schulze JM, Jackson J, Hentrich T, Seidel C, Jaspersen SL, Kobor MS, Shilatifard A. 2011. Dot1 and histone H3K79 methylation in natural telomeric and HM silencing. *Mol Cell* **42:** 118–126.

Takeuchi Y, Horiuchi T, Kobayashi T. 2003. Transcription-dependent recombination and the role of fork collision in yeast rDNA. *Genes Dev* **17:** 1497–1506.

Tanabe S, Bohlander SK, Vignon CV, Espinosa R 3rd, Zhao N, Strissel PL, Zeleznik-Le NJ, Rowley JD. 1996. AF10 is split by MLL and HEAB, a human homolog to a putative *Caenorhabditis elegans* ATP/GTP-binding protein in an invins(10;11)(p12;q23q12). *Blood* **88:** 3535–3545.

Tong AH, Lesage G, Bader GD, Ding H, Xu H, Xin X, Young J, Berriz GF, Brost RL, Chang M, et al. 2004. Global mapping of the yeast genetic interaction network. *Science* **303:** 808–813.

Tous C, Rondon AG, Garcia-Rubio M, Gonzalez-Aguilera C, Luna R, Aguilera A. 2011. A novel assay identifies transcript elongation roles for the Nup84 complex and RNA processing factors. *EMBO J* **30:** 1953–1964.

Tuduri S, Crabbe L, Conti C, Tourriere 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.

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

Yuen KW, Warren CD, Chen O, Kwok T, Hieter P, Spencer FA. 2007. Systematic genome instability screens in yeast and their potential relevance to cancer. *Proc Natl Acad Sci* **104:** 3925–3930.