Disruption of a ~23–24 nucleotide small RNA pathway elevates DNA damage responses in *Tetrahymena thermophila*

Suzanne R. Lee[a],*, Daniel A. Pollard[a], Domenico F. Galati[a], Megan L. Kelly[a], Brian Miller[a,†], Christina Mong[a,‡], Megan N. Morris[a,‡], Kerry Roberts-Nygren[a,‡], Geoffrey M. Kapler[b], Matthew Zinkgraf[a], Hung Q. Dang[b], Erica Branham[b], Jason Sasser[a,§], Erin Tessier[a,§], Courtney Yoshiyama[a,§], Maya Matsumoto[a], and Gaea Turman[a]
[a]Biology Department, Western Washington University, Bellingham, WA 98225; [b]Molecular and Cellular Medicine, Texas A&M University, College Station, TX 77843

ABSTRACT Endogenous RNA interference (RNAi) pathways regulate a wide range of cellular processes in diverse eukaryotes, yet in the ciliated eukaryote, *Tetrahymena thermophila*, the cellular purpose of RNAi pathways that generate ~23–24 nucleotide (nt) small (s)RNAs has remained unknown. Here, we investigated the phenotypic and gene expression impacts on vegetatively growing cells when genes involved in ~23–24 nt sRNA biogenesis are disrupted. We observed slower proliferation and increased expression of genes involved in DNA metabolism and chromosome organization and maintenance in sRNA biogenesis mutants *RSP1Δ*, *RDN2Δ*, and *RDF2Δ*. In addition, *RSP1Δ* and *RDN2Δ* cells frequently exhibited enlarged chromatin extrusion bodies, which are nonnuclear, DNA-containing structures that may be akin to mammalian micronuclei. Expression of homologous recombination factor Rad51 was specifically elevated in *RSP1Δ* and *RDN2Δ* strains, with Rad51 and double-stranded DNA break marker γ-H2A.X localized to discrete macronuclear foci. In addition, an increase in Rad51 and γ-H2A.X foci was also found in knockouts of TWI8, a macronucleus-localized PIWI protein. Together, our findings suggest that an evolutionarily conserved role for RNAi pathways in maintaining genome integrity may be extended even to the early branching eukaryotic lineage that gave rise to *Tetrahymena thermophila*.

INTRODUCTION

Critically important biological processes like defense against viruses and transposons, cell differentiation, and organismal development depend on endogenous RNA interference (RNAi) pathways. These processes are primarily mediated by small RNAs (sRNAs) that associate with Argonaute/PIWI proteins and base pair to complementary regions in target RNAs for posttranscriptional or transcriptional gene silencing (Czech and Hannon, 2016; Daugaard and Hansen, 2017; Gorski *et al.*, 2017; Ozata *et al.*, 2019). In addition to gene silencing, RNAi contributes to a variety of other DNA-directed processes, including centromere and pericentromere organization, copy number control, release of RNA polymerase from DNA, and DNA repair (Gutbrod and Martienssen, 2020). As RNAi has been primarily studied in *Arabidopsis thaliana* and other model systems of the opisthokont lineage, how broadly conserved these RNAi activities are remains an important open question.

Falling on a less studied branch of the evolutionary tree than more commonly studied eukaryotes, ciliates are alveolates of the SAR (stramenopiles, alveolates, and Rhizaria) supergroup (Lynch *et al.*, 2014). Studies of endogenous RNAi pathways in model ciliates have primarily focused on their developmental roles in somatic nuclear development. Ciliates exhibit nuclear dualism, harboring—

within a single cell—a polyploid (45C), transcriptionally active somatic nucleus (macronucleus; MAC) and a diploid, transcriptionally inactive germline nucleus (micronucleus; MIC). During sexual reproduction, the new MAC develops from a mitotic product of the zygotic MIC through elimination of MIC-limited sequences, chromosome breakage, and telomere addition, and rearrangement of DNA segments (Rzeszutek et al., 2020). Developmentally induced sRNAs associated with PIWI homologues specify many of the DNA segments destined for elimination or retention through a parentally guided inheritance mechanism in which the parental MAC genome specifies the content of the zygotic MAC (Rzeszutek et al., 2020). In Tetrahymena thermophila, the DNA segments that are removed from the developing MAC include 95% of transposable elements and their remnants (Hamilton et al., 2016) and putative centromeres (Cervantes et al., 2006b; Eisen et al., 2006; Hamilton et al., 2016). DNA elimination thus leaves a sequence-streamlined MAC devoid of invasive elements and centromeres that divides through an amitotic process during vegetative growth while the MIC divides through conventional mitosis (Orias and Flacks, 1975; Wong et al., 2000).

We previously identified a set of RNAi pathways in T. thermophila distinct from that driving DNA elimination. These pathways generate ~23–24 nucleotide (nt) sRNAs that accumulate throughout the T. thermophila life cycle (Lee and Collins, 2006). Sequencing of these sRNAs revealed alignment to a number of discrete sites in the MAC genome. The most abundant sRNAs mapped in clusters to processed pseudogene-like loci, widely dispersed high-copy repeats, and intergenic loci yielding long ncRNAs with predicted 3' hairpins (Lee and Collins, 2006; Couvillion et al., 2009; Farley and Collins, 2017). In addition, a subset of sRNAs mapped to a limited number of putative protein-coding loci predicted to generate transcripts that form extensive secondary structure, correspond to paralogous gene clusters, or undergo convergent transcription (Couvillion et al., 2009).

With the exception of sRNAs derived from putative protein-coding loci, accumulation of the sRNA classes described above depends on one or more of three distinct RNA-dependent RNA polymerase complexes (RdRP complexes, or RDRCs; Supplemental Figure S1A). Each RDRC contains the RdRP Rdr1 and one of two nucleotidyltransferases, Rdn2 or Rdn1 (Lee et al., 2009). Rdn1-RDRCs additionally contain one of two novel proteins, Rdf2 or Rdf1 (Lee et al., 2009; Talsky and Collins, 2012). Regardless of composition, the RDRCs generate double-stranded (ds)RNAs that are then processed by Dcr2, a dsRNA nuclease Dicer (Lee and Collins, 2007). Accumulation of all RDRC-dependent sRNAs is also dependent on RNA silencing protein 1 (Rsp1; Talsky and Collins, 2012). Although the pathway yielding RDRC-independent sRNAs remains unknown, accumulation of these sRNAs is repressed in an RSP1- and RDN2-dependent manner (Couvillion et al., 2009; Talsky and Collins, 2012). Ultimately, the ~23–24 nt sRNAs associate with at least two distinct Piwi family proteins that are highly expressed in growing cells: cytoplasmic Twi2 and MAC-localized Twi8 (Couvillion et al., 2009).

Importantly, despite gains in elucidating the biogenesis pathways for ~23–24 nt sRNAs in T. thermophila, insight into the biological functions of these pathways has remained elusive. DCR2, RDR1, and RDN1 are genetically essential for cell viability for yet unknown reasons; however, RSP1, RDN2, RDF1, RDF2, and individual TWI genes are not. To gain deeper insight into the biological contributions of this RNAi pathway, we undertook phenotypic analysis of strains lacking these nonessential RNAi genes during vegetative cell growth. Using a variety of assays, we demonstrate that the absence of RSP1 and RDN2, as well as TWI8, leads to an elevation in markers of DNA damage in MAC nuclei, implicating these RNAi factors in protecting the somatic MAC genome in T. thermophila.

RESULTS AND DISCUSSION

sRNA biogenesis mutants exhibit slower growth and DNA replication phenotypes

Because previous phenotypic characterization of RSP1Δ, RDN2Δ, RDF1Δ, and RDF2Δ strains during growth was limited in scope (Lee et al., 2009; Talsky and Collins, 2012), we initially reexamined these sRNA biogenesis gene knockouts by measuring their growth rates and quantifying cell cycle stage distribution in asynchronous cultures. We found that RSP1Δ, RDN2Δ, and RDF2Δ, but not RDF1Δ, grow approximately two to three times slower than the wild-type parental strain SB210 (Figure 1A). Examination of stained nuclei revealed no statistically significant differences in the distribution of cells in macronuclear G1/S, G2, and amitosis for any knockout strain compared to SB210 (Supplemental Figure S1, B–D; Wei et al., 1998; Cole and Sugai, 2012; Upton et al., 2014). However, because macronuclear G1 and S phase are not easily distinguished by nuclear staining alone, we used pulse-labeling with the modified deoxynucleotide EdU to determine the relative fraction of cells engaged in DNA replication for each knockout strain. Here we found RSP1Δ, RDN2Δ, and RDF2Δ cultures exhibit a statistically significant reduction in EdU-positive cells, while RDF1Δ exhibits a slight elevation, compared to SB210 (Figure 1, B and C). Together, these data suggest that RSP1Δ, RDN2Δ, and RDF2Δ cells are stalled in G1 or at the G1/S transition, and/or the rate of DNA replication in S phase in a subset of cells is reduced to a degree that any incorporated EdU during the pulse is undetectable.

sRNA biogenesis mutants generate extranuclear DNA bodies that are enlarged and at an elevated frequency

Strikingly, among the slower growing strains, we also noticed extranuclear DNA bodies that are enlarged in RSP1Δ and RDN2Δ cells compared to those in SB210 and RDF2Δ (Figure 2, A–C, and Supplemental Figure S1E). Enlarged extranuclear DNA bodies relative to SB210 were also not observed of RDF1Δ cells (unpublished data). These DAPI-staining bodies were typically well separated from the MAC and did not stain with a mitotic MIC marker, which distinguishes them from MIC nuclei as MICs closely associate with MACs when not in mitosis. These are likely chromatin extrusion bodies (CEBs), which are thought to arise from MAC DNA failing to segregate during amitosis and may ultimately fragment in a TUNEL-positive process (Bodenbender et al., 1992; Gao et al., 2013; Kaczanowski and Kiersnowska, 2018). Indeed, although only ~2–3% of cells in asynchronous cultures are in amitosis (Supplemental Figure S1, B–D), we found MAC DNA distribution to daughter cells often lags behind the formation of cytokinetic furrows in RSP1Δ and RDN2Δ compared to SB210, as illustrated by the representative dividing cells outlined in Figure 2A. In contrast, no micronuclear segregation defect was observed during MIC mitosis.

Although putative CEBs are present in cells at all cell cycle stages regardless of genotype, we also found that RSP1Δ, RDN2Δ, and RDF2Δ strain cultures tend to have more CEB-positive cells compared to SB210 (RSP1Δ: 16–55% of total cells, RDN2Δ: 42–57%, and RDF2Δ: 28–37% vs. SB210: 17–33%, N = 3). This bias was found to be significant for all three knockout strains (logistic regression with genotype and replicate factors, z-test $p < 10^{-4}$ for all three genotype contrasts), while no elevation in CEB production was observed of RDF1Δ cells. The number of CEBs in CEB-positive cells ranged from one to five; for cells with at least one CEB, RSP1Δ and
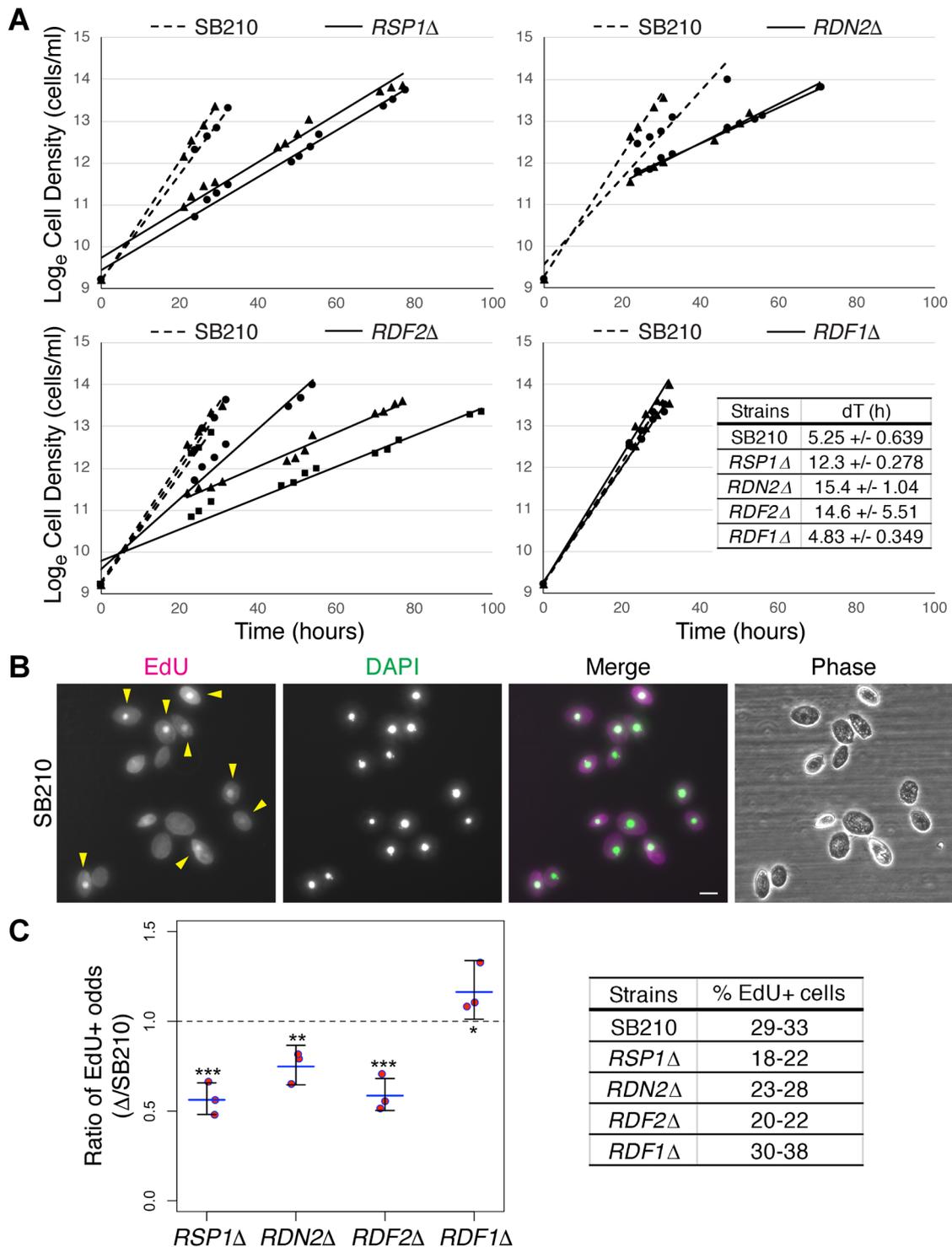
FIGURE 1: *RSP1Δ*, *RDN2Δ*, and *RDF2Δ* cultures exhibit slower proliferation and a reduced proportion of cells engaged in DNA replication. (A) Growth curves for cultures seeded in triplicate at 1×10^4 cells/ml, represented as $\log_e$-transformed averages of triplicate culture densities over time. Linear regression line was fitted to timepoints during exponential growth to compare sRNA biogenesis knockouts (solid lines) to the wild-type SB210 parental strain (dotted lines), with paired biological replicates denoted by data point markers ($N = 2$: circle, triangle; $N = 3$: circle, triangle, square). Inset lists doubling times (dT) $\pm$ SD. (B) Representative cells in asynchronous, log phase cultures undergoing DNA synthesis labeled with EdU (magenta) and stained with DAPI (green). Arrowheads denote cells with a macronucleus scored as EdU-positive. Scale bar = 30 μm. (C) Left, estimated ratio of EdU-positive (EdU+) odds for knockout strains compared to that for SB210 as analyzed by logistic regression with genotype and replicate ($N = 3$) factors. Estimated odds ratios are shown as blue horizontal bars; error bars are 95% confidence intervals about an estimated odds ratio; red data points are replicate measured odds ratios; dotted line is the null expectation of equal odds. ***, $p < 5 \times 10^{-12}$; **, $p < 2 \times 10^{-4}$; *, $p < 0.04$. Right, range of percent EdU+ cells observed across replicates.
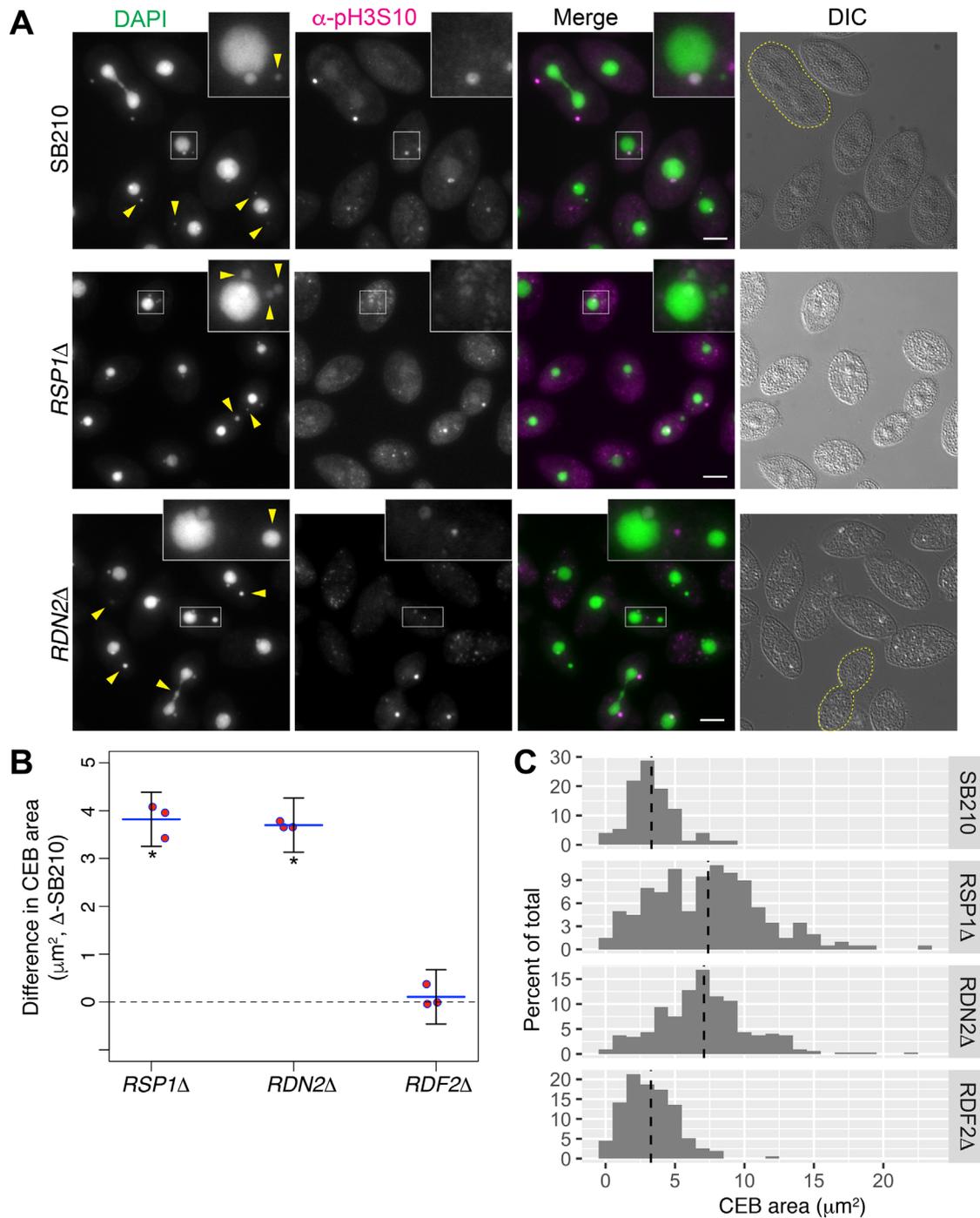
FIGURE 2: Extranuclear chromatin extrusion bodies (CEBs) are generated more frequently and are larger in *RSP1Δ* and *RDN2Δ*. (A) Representative DAPI- (green) and α-phospho-histone H3 serine 10 (pH3S10; magenta) –stained SB210, *RSP1Δ*, and *RDN2Δ* cells. Yellow arrowheads denote CEBs; amitotic cells are outlined with dotted yellow lines. Scale bars = 15 μm. (B) Two-way ANOVA with genotype and replicate factors was used to compare mean CEB areas while controlling for variation between biological replicates ($N = 3$) and was found to be significant ($p < 5 \times 10^{-7}$). Differences in CEB area between knockouts and SB210 were evaluated for significance by Tukey post-hoc analysis. Estimated CEB area differences between knockouts and SB210 are shown as blue horizontal bars; error bars are 95% confidence intervals about an estimated difference in area; red data points represent individual replicate differences; the dotted line is the null expectation of zero difference. *, $p < 1 \times 10^{-5}$. (C) Representative replicate histogram of the CEB area frequency for ≥73 individual CEBs from each strain. Vertical dashed lines represent mean CEB area for replicate.

RDN2Δ were found to have more CEBs per cell than SB210, though the difference was only significant for *RDN2Δ* (logistic regression with genotype and replicate factors, z-test $p = 0.0259$ for *RDN2Δ* vs. SB210).

Ciliate CEBs bear similarity to micronuclei frequently found in cancer cells, and similar structures have also been described in fission yeast and *Arabidopsis* (Sabatinos *et al.*, 2015; Tan *et al.*, 2015; Guo *et al.*, 2019). Micronuclei contain whole chromosomes or

acentric fragments of chromosomes and are formed when DNA is mis-segregated as a result of errors in DNA replication, DNA repair, chromosome segregation, and cell cycle checkpoints (Guo *et al.*, 2019). Similarly, *Tetrahymena* CEB frequency and size increases with disruptions to normal DNA replication (Morrison *et al.*, 2005; Yakisich *et al.*, 2006; Gao *et al.*, 2013; Kaczanowski and Kiersnowska, 2018), depletion of condensin or histone deacetylase Thd1 (Wiley *et al.*, 2005; Cervantes *et al.*, 2006a), or blocks to DNA repair (Marsh *et al.*, 2000; Song *et al.*, 2007). Based on a positive correlation observed between MAC DNA content and CEB production, it has been speculated that production of CEBs may be a mechanism to selectively eliminate extra copies of MAC chromosomes that are aberrantly replicated and/or result from imbalances in amitotic MAC division (Cleffmann, 1980; Bodenbender *et al.*, 1992; Wiley *et al.*, 2005; Kaczanowski and Kiersnowska, 2018). However, hallmarks of DNA damage have also been found in some CEBs, even as they form (Gao *et al.*, 2013), suggesting that at least some CEBs may contain damaged DNA.

DNA replication stress, damage response, and repair genes are elevated in *RSP1Δ* and *RDN2Δ*

The phenotypes described above suggest that sRNA biogenesis factors Rsp1 and Rdn2, and potentially Rdf2, may be required for some aspect of proper macronuclear DNA metabolism and/or chromosome maintenance. To investigate the underlying mechanisms that account for our observations, we performed poly(A)+ RNA sequencing on asynchronous cultures of SB210 and each of the knockout strains *RSP1Δ*, *RDN2Δ*, *RDF2Δ*, and *RDF1Δ*, obtaining an average of 27.3 million mapped reads per sample. Consistent with the degrees to which mutants are affected in their accumulation of different sRNA classes (Couvillion *et al.*, 2009; Talsky and Collins, 2012) as well as the relative extents of observed phenotypes described above, *RSP1Δ* exhibits the greatest number of differentially expressed (DE) genes (8526 genes) relative to SB210, followed by *RDN2Δ* (2470 genes), *RDF2Δ* (1970 genes), and *RDF1Δ* (91 genes; Supplemental Figures S2, A and B, and S3). We noted that the number of down-regulated genes was similar to the number of up-regulated genes in each knockout, suggesting that DE genes in our strains are not limited to genes that may be repressed by sRNAs through canonical RNAi. It is worth noting that unpublished work found that transgene reporters designed to be targeted by endogenous sRNAs in *T. thermophila* are not repressed (K. Collins, personal communication), lending support to endogenous RNAi roles that are distinct from sRNA-guided gene silencing.

To determine what biological processes are enriched among the DE genes for each mutant, we identified biological process gene ontology (GO) terms that were overrepresented relative to their representation in the expressed genome. The GO analysis was facilitated by combining the limited set of publicly available *T. thermophila* gene GO annotations with transferred GO annotations from putative orthologues as determined by BLASTp against the well annotated genomes of four model eukaryotes as well as *Plasmodium falciparum* (an alveolate like *T. thermophila*) and humans (Supplemental Figure S3; see *Materials and Methods*). To focus our GO analysis on the genes with the most consistent differential expression, we limited our analysis to genes exhibiting differential expression with an adjusted p value $\leq 5 \times 10^{-5}$, which uncovered significantly enriched GO terms in *RSP1Δ*, *RDN2Δ*, and *RDF2Δ* DE gene sets, but none in the *RDF1Δ* DE gene set (Supplemental Figure S3). Notably, the majority of enriched GO terms were associated specifically with up-regulated genes and several biological process categories (Supplemental Figure S3).

Given the slow growth, DNA replication, and chromatin extrusion body phenotypes of *RSP1Δ*, *RDN2Δ*, and *RDF2Δ*, we noted with interest enriched GO term categories related to cell cycle and division, DNA metabolism, and chromosome organization and maintenance (Supplemental Figures S2C and S3). Consistent with the contributions of *RDN2* and *RDF2* to the accumulation of distinct, though partially overlapping, subsets of sRNA classes while both *RDN2*- and *RDF2*-dependent sRNAs require *RSP1* for accumulation (Couvillion *et al.*, 2009; Talsky and Collins, 2012), greater overlap in certain enriched GO term categories was seen between *RSP1Δ* and either *RDN2Δ* or *RDF2Δ* than was observed between *RDN2Δ* and *RDF2Δ*. An investigation of the underlying genes in DNA and chromosome-related categories revealed a disproportionate number of up-regulated genes with known or predicted roles in DNA replication stress responses, damage responses, and repair, particularly in *RSP1Δ* and *RDN2Δ* (Figure 3A and Supplemental Figure S3). None of these genes were differentially expressed in *RDF1Δ* (Supplemental Figure S3). These genes include those that are highly conserved with well-characterized functions in DNA repair processes and putative representatives of DNA-directed and/or DNA damage signaling protein families, including DNA helicases, poly (ADP)-ribose polymerases, and kinases. These findings motivated us to examine our knockout strains for additional evidence of DNA damage.

Rad51 protein and γ-H2A.X, markers of double-stranded DNA breaks, are elevated and localized to macronuclear foci in *RSP1Δ* and *RDN2Δ*

Of the up-regulated genes identified above to be involved in DNA replication stress and damage responses, the most well-characterized is RAD51, which encodes a central player in double-stranded DNA break (DSB) repair and homologous recombination (Ciccia and Elledge, 2010). Rad51 up-regulation would be expected of cells responding to DNA damage, so to extend our findings of elevated RAD51 transcript levels in *RSP1Δ* and *RDN2Δ*, we performed Western blots on whole cell lysates to examine Rad51 protein levels. Consistent with our findings at the mRNA level, Rad51 protein levels are elevated in asynchronous cultures of both *RSP1Δ* (by 7.5- to 24-fold) and *RDN2Δ* (by 3.7- to 5-fold) relative to SB210, while they remain unchanged in *RDF2Δ* (Figure 3, B and C) and *RDF1Δ* (unpublished data). Rad51 levels in *RSP1Δ* and *RDN2Δ* approach, but do not exceed, Rad51 levels in SB210 cells treated with the replication elongation inhibitor hydroxyurea (+HU; Figure 3B), which is a particularly strong inducer of Rad51 (Yakisich *et al.*, 2006).

We next examined Rad51 localization in *RSP1Δ* and *RDN2Δ* cells by immunofluorescence, reasoning that if Rad51 is engaged in DNA repair, it would localize to damaged nuclei. Imaging revealed that a subset of cells in asynchronous *RSP1Δ* and *RDN2Δ* cultures exhibit higher α-Rad51 staining in MAC nuclei relative to staining in the cytoplasm (representative images of *RDN2Δ* in Supplemental Figure S4A). In contrast, elevated Rad51 was not observed in either MICs or CEBs. Using a threshold-independent averaging approach to quantify the nuclear-to-cytoplasmic ratio of Rad51 (see *Materials and Methods*; Supplemental Figure S4B), we found that this ratio is significantly elevated in *RSP1Δ* and *RDN2Δ* compared to SB210 (Figure 4, A and B), suggesting greater Rad51 recruitment to MAC nuclei in knockout strains. In some *RSP1Δ* and *RDN2Δ* cells, the Rad51 fluorescence ratios are increased by more than eightfold compared to SB210 (Figure 4B), representing cells with particularly elevated macronuclear Rad51.

To further investigate the mechanism underlying elevated Rad51 levels and nuclear localization, we stained cells for a second marker of DSBs. An early event that occurs near a DSB site is phosphorylation
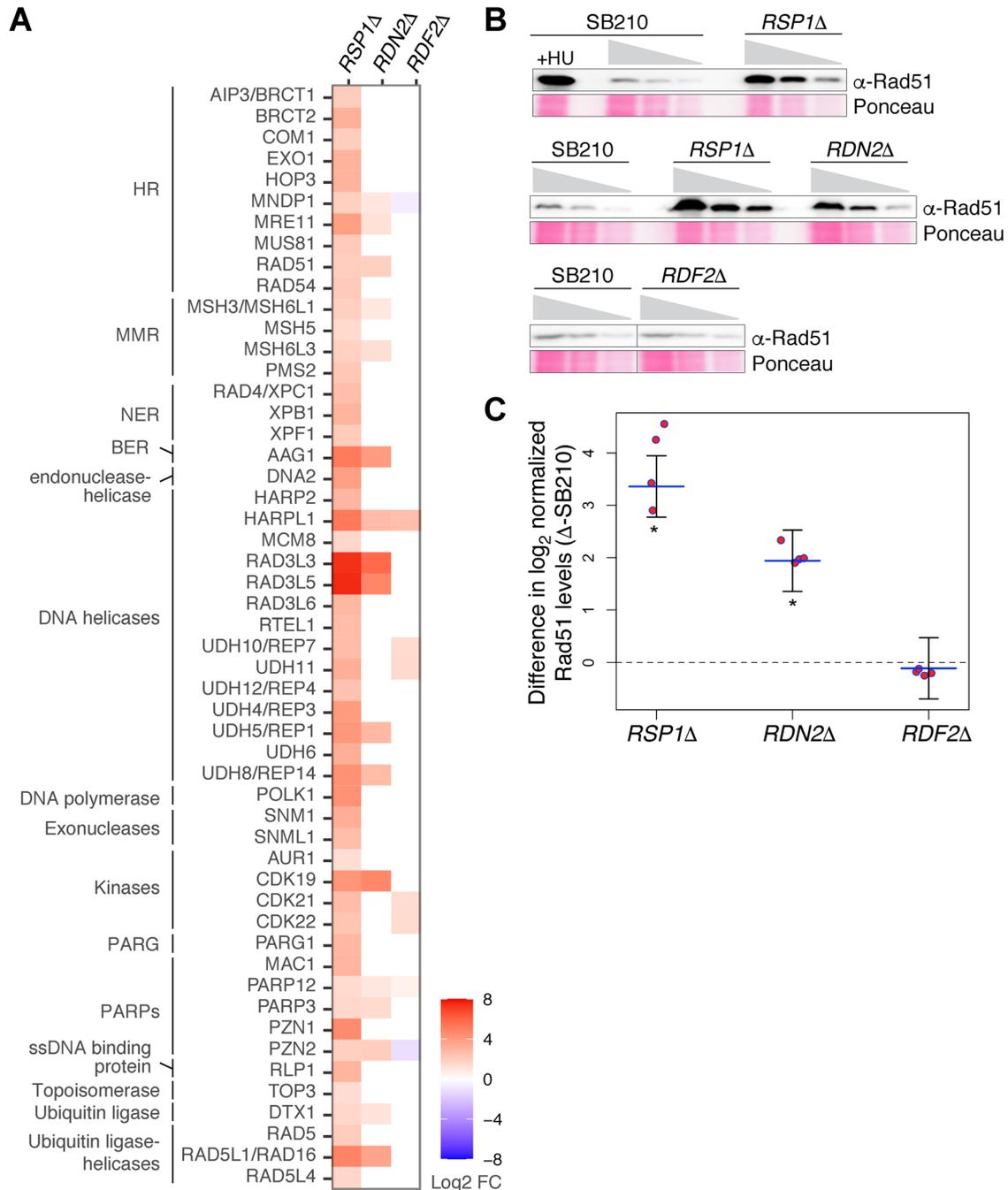
FIGURE 3: *RSP1Δ* and *RDN2Δ* exhibit elevated expression of DNA replication stress, damage response, and repair genes. (A) Differential expression of DNA replication stress, damage response, and repair genes in *RSP1Δ*, *RDN2Δ*, and *RDF2Δ*, relative to SB210, with heat map of significant ($p \leq 0.05$) log$_2$ fold change. Gene names are as listed in the *Tetrahymena* Genome Database. HR, homologous recombination; MMR, mismatch repair; NER, nucleotide excision repair; BER, base excision repair; PARG, poly (ADP-ribose) glycohydrolase; PARP, poly (ADP-ribose) polymerase. (B) Representative Rad51 Western blots of untreated or hydroxyurea-treated ("+HU") whole cell protein samples. Ponceau S–stained regions correspond in size to Rad51 blots. A twofold dilution series for each sample was used to identify signals in a linear range for quantification. SB210 and *RDF2Δ* lanes in the third set of Rad51 and Ponceau S panels are cropped from the same respective images. (C) Two-way ANOVA with genotype and blot-replicate factors was used to compare Ponceau S–normalized Rad51 protein levels while controlling for variation between blots and biological replicates ($N = 4$) and was found to be significant ($p < 6 \times 10^{-9}$). Differences in log$_2$-normalized Rad51 levels between knockouts and SB210 were evaluated for significance by Tukey post-hoc analysis. Estimated differences in log$_2$-normalized levels are shown as blue horizontal bars; error bars, red data points, and dotted line are as in Figure 2B. *, $p < 1 \times 10^{-5}$.
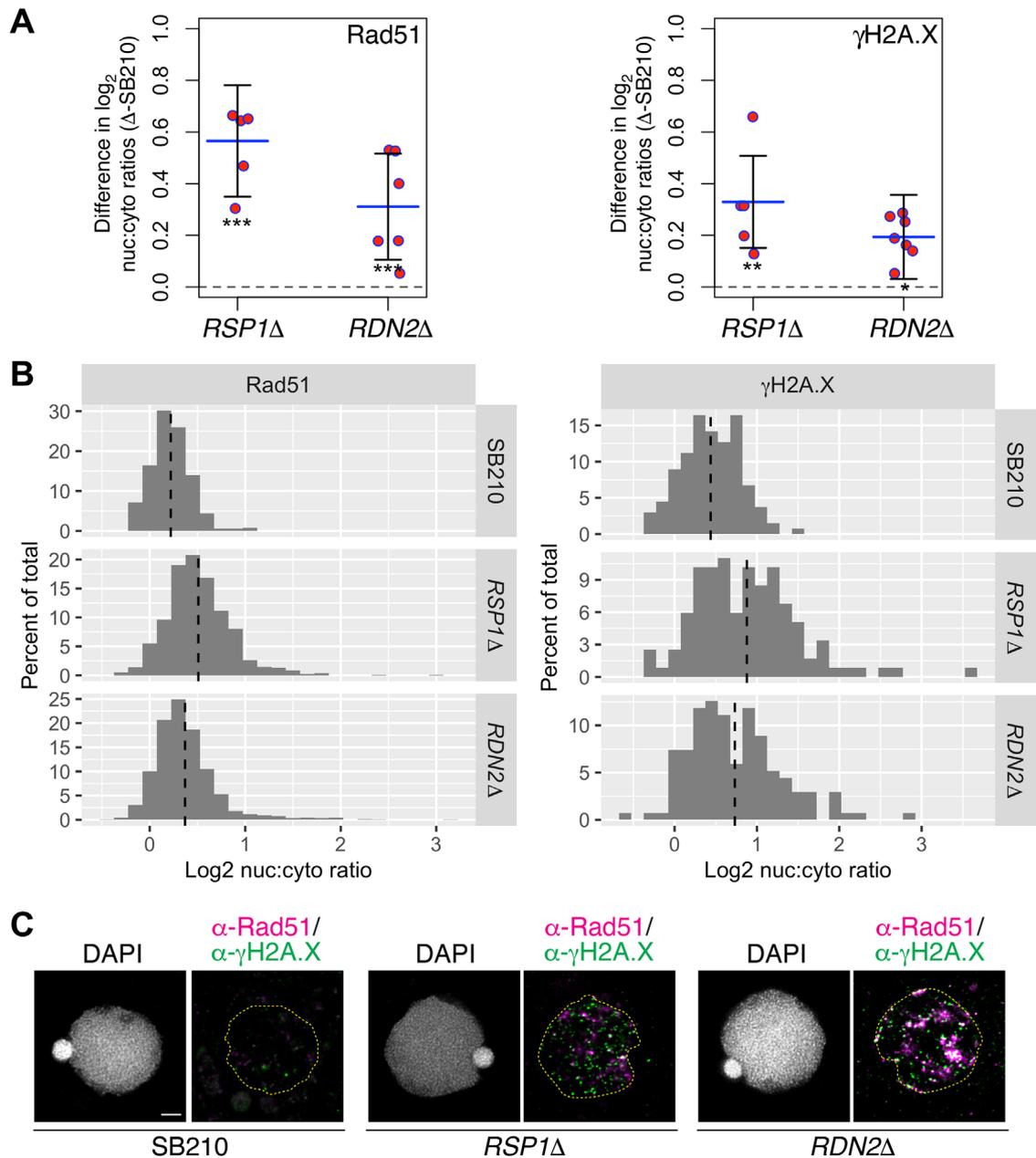
FIGURE 4: Macronuclear Rad51 and γH2A.X are elevated and localize in macronuclear foci in *RSP1Δ* and *RDN2Δ* cells. (A) Two-way ANOVA with genotype and replicate factors was used to compare macronuclear (nuc):cytoplasmic (cyto) ratios while controlling for variation between biological replicates ($N = 7$ for Rad51, $N = 6$ for γH2A.X-*RSP1Δ*, $N = 8$ for γH2A.X-SB210 and γH2A.X-*RDN2Δ*) and was found to be significant (Rad51, $p = 2 \times 10^{-4}$; γH2A.X, $p < 3 \times 10^{-3}$). Differences in $\log_2$ nuc:cyto ratios for Rad51 (left) or γH2A.X (right) between knockouts and SB210 were evaluated for significance by Tukey post-hoc analysis. Estimated differences in $\log_2$ ratios are shown as blue horizontal bars; error bars, red data points, and dotted line are as in Figure 2B. See Supplemental Figure S4 for representative individual and averaged images. ***, $p \leq 6 \times 10^{-3}$; **, $p < 2 \times 10^{-3}$; *, $p < 0.03$. (B) Representative histogram of $\log_2$ nuc:cyto ratios for Rad51 (left) and H2AX (right) for more than 50 individual cells in each strain. Vertical dashed lines represent mean $\log_2$-transformed ratio. (C) Deconvolved laser scanning confocal images showing characteristic punctate staining with α-Rad51 (magenta) and α-phospho-γH2A.X (green) within the macronucleus (dashed line) of SB210, *RSP1Δ*, and *RDN2Δ* cells. Nuclei are larger in these images than those in images elsewhere in this study because these cells were imaged in PBS with a water immersion objective and not subjected to dehydration from mounting in glycerol (see *Materials and Methods* for additional details). Minor punctate signal outside of nuclei is a combination of nonspecific staining and autofluorescence. The α-Rad51 and α-phospho-γH2A.X signals for all images were acquired with the same laser power and detector gain on the same day, and the images are scaled with identical contrast so the relative differences in signal intensity are preserved. Scale bar = 2.5 μm.

of the histone H2A variant H2A.X (γ-H2A.X; Song *et al.*, 2007; Mah *et al.*, 2010). Similar to Rad51, γ-H2A.X is also elevated in MAC nuclei, but not MICs or CEBs, in *RSP1Δ* and *RDN2Δ* cells (Figure 4, A and B, and Supplemental Figure S4, A and B). Importantly, high-resolution confocal microscopy revealed that both Rad51 and γ-H2A.X localize in discrete foci that increase in both size and number in positively staining *RSP1Δ* and *RDN2Δ* MACs compared to SB210 (Figure 4C); an increase in these types of foci is a hallmark of DNA damage through DSBs (Mah *et al.*, 2010). Notably, *RSP1Δ* and *RDN2Δ* cells that costain strongly with the MIC mitotic marker (pH3S10) in asynchronous cultures do not exhibit elevated γ-H2A.X and Rad51, suggesting that an increase in DSBs either occurs or recruits DNA repair machinery in MAC G1 and/or S phase cells, with resolution of DNA damage and/or γ-H2A.X and Rad51 association with DSBs before MAC G2 (Supplemental Figure S1, B and C, and S4A).

Together with our finding that pulse-labeling of MACs with EdU was reduced in *RSP1Δ* and *RDN2Δ* cultures (Figure 1, B and C), the increase in Rad51 and γ-H2A.X MAC foci suggests that DNA damage in *RSP1Δ* and *RDN2Δ* cells delays or slows DNA synthesis. Alternatively, replication stress associated with inhibited replication elongation could lead to DNA damage (Wilhelm *et al.*, 2020). In addition to being a potential consequence of DNA damage, the enlarged micronuclei-like CEBs in *RSP1Δ* and *RDN2Δ* cells (Figure 2, A–C) may also contribute to DNA damage, as micronuclei have recently been linked to chromothripsis, a process in which DNA damage and repair are triggered by the reintegration of fragments from micronucleated DNA into primary nuclei (Sabatinos *et al.*, 2015; Ly and Cleveland, 2017). Identifying the underlying cause of DNA damage in our RNAi gene knockouts will benefit from future studies of the relative timing of DNA replication, DSBs, and CEB formation and fate.

DNA damage markers were not significantly elevated in *TWI2Δ* cells, but were in *TWI8Δ* cells

Taken together, the above data suggest that Rsp1, Rdn2, and/or the ~23–24 nt sRNAs that depend on Rsp1 and Rdn2 for accumulation are important for maintaining the integrity of MAC DNA during growth and limiting the production and size of CEBs. Because ~23–24 nt sRNAs bind to Piwi-homologue Twi proteins, we reasoned that a Twi might also be important for limiting MAC DNA damage in *T. thermophila*. Of the two Twis most highly expressed during growth, a knockout strain lacking cytoplasmically localized Twi2 fully phenocopies *RDN2Δ* in its altered sRNA profiles (Couvillion *et al.*, 2009). In contrast, Twi8, which localizes predominantly to the macronucleus (Couvillion *et al.*, 2009; Farley and Collins, 2017), contributes to the accumulation of a subset of the sRNA classes perturbed *RDN2Δ* cells.

Thus, to determine whether a Twi-sRNA effector complex might have a role in maintaining MAC integrity, we first examined Rad51 protein levels in asynchronous cultures of *TWI2Δ* and *TWI8Δ* cells. To our surprise, although Rad51 protein was modestly elevated in *TWI2Δ* cells compared to SB210, the elevation did not rise to the level of statistical significance (Figure 5, A and B). However, in *TWI8Δ* cells compared to SB210, Rad51 was significantly elevated to levels similar to that observed in *RDN2Δ*, by two- to fivefold. Moreover, similar to *RSP1Δ* and *RDN2Δ*, discrete Rad51 and γ-H2A.X foci were increased in both size and number in *TWI8Δ* cells compared to SB210 (Figure 5C), suggesting that MAC-localized Twi8 is also required to maintain MAC genome integrity.

Perspective
Our results suggest *T. thermophila* has at least two RNAi-dependent mechanisms dedicated to safeguarding the somatic MAC genome.
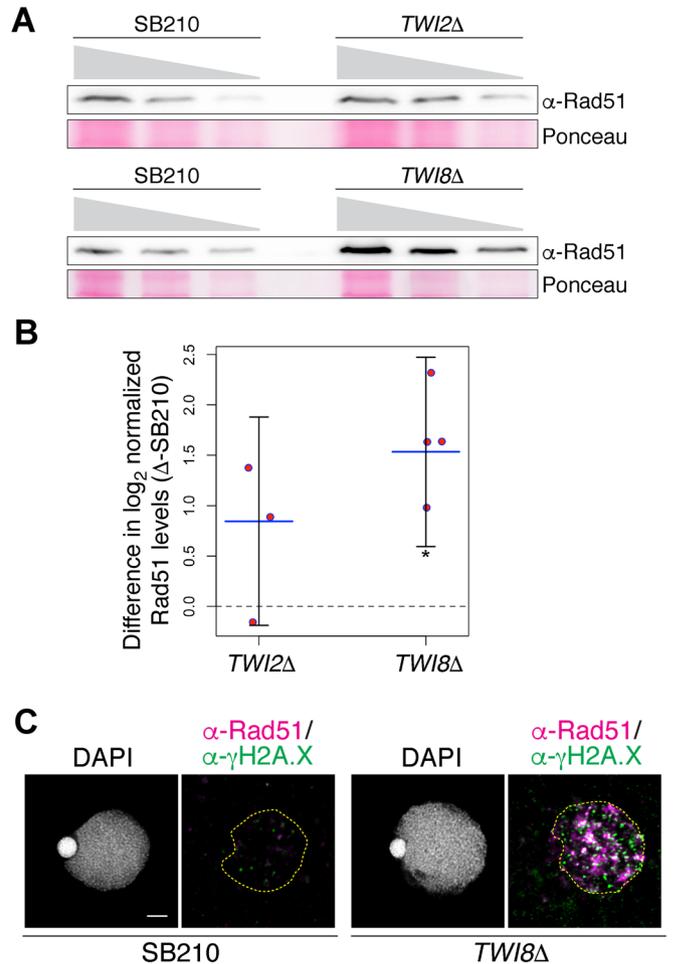

FIGURE 5: *TWI8Δ* cells exhibit elevated Rad51 protein levels and macronuclear Rad51 and γH2A.X foci. (A) Representative Rad51 Western and Ponceau S–stained blot of whole protein samples as in Figure 3B. (B) Two-way ANOVA performed as in Figure 3C was found to be significant ($p < 0.008$). Differences in $\log_2$ normalized Rad51 levels between knockouts and SB210 were evaluated for significance by Tukey post-hoc analysis. Estimated differences in $\log_2$ normalized levels are shown as in Figure 3C. *, $p < 0.008$. (C) Deconvolved laser scanning confocal images of α-Rad51 (magenta) and α-phospho-γH2A.X (green) –stained SB210 and *TWI8Δ* cells, as in Figure 4C.

The *RDN2*-, *RSP1*-, and *TWI8*-dependent pathway we identified here functions to ensure DNA integrity during growth, complementing the developmentally induced RNAi pathway that removes active transposons (Zhao *et al.*, 2019) and other sequences from the MAC. More broadly, our findings indicate that a transposon- and centromere-independent role for RNAi in protecting chromosomes during cell growth may be more widely conserved than previously recognized (Gutbrod and Martienssen, 2020).

Whether the *RDN2*-, *RSP1*-, and *Twi8*-dependent role in maintaining MAC DNA integrity is sRNA-dependent remains an important area for future study. While *RSP1* is needed for accumulation of all RDRC-dependent sRNAs in *T. thermophila* (Talsky and Collins, 2012), *RDN2* specifically contributes to the accumulation of sRNAs derived from clusters of sequence-related processed pseudogenes and high-copy repeats composed of tandem arrays found throughout the genome (Couvillion *et al.*, 2009). One possible explanation for our findings is that Rdn2-RDRCs and/or *RDN2*-dependent sRNAs are important for ensuring proper

chromatin structure at repetitive loci in order to repress HR. This would be similar to a role for Dicer-2 in *Drosophila*, where Dicer-2 mutants exhibit DSBs and nonsegregating extrachromosomal DNA derived from repeats (Peng and Karpen, 2007, 2009). Alternatively, replication or transcription stress in repetitive genomic regions may lead to DNA damage that requires *RDN2* and *TWI8* for proper repair, similar to a role for RNAi in *Neurospora crassa* in promoting DNA repair at repetitive transgenic loci (Zhang *et al.*, 2013; Yang and Qi, 2015) or in human cells at specific transcription termination sites that form RNA–DNA hybrid structures (R-loops; Hatchi *et al.*, 2021). Notably, experimentally induced DSBs, often in repetitive regions, produce sRNAs in *Arabidopsis*, *Drosophila,* and mammalian cells (Francia *et al.*, 2012; Michalik *et al.*, 2012; Wei *et al.*, 2012; Yang and Qi, 2015). Some DNA damage–associated sRNAs appear to nucleate DNA damage response factors and/or DNA repair proteins at DNA breaks (Gao *et al.*, 2014; Francia *et al.*, 2016; Wang and Goldstein, 2016; Liu *et al.*, 2017; Michelini *et al.*, 2017; Hatchi *et al.*, 2021), though RNAi proteins may also have sRNA-independent roles in DSB repair (Lu *et al.*, 2018). We note that while *TWI8* contributes to *RDN2*-dependent sRNA accumulation, this is also true of *TWI2Δ* cells (Couvillion *et al.*, 2009), which do not exhibit significantly elevated Rad51. However, one possibility is that the residual *RDN2*-dependent sRNAs persisting in *TWI2Δ* cells associate with Twi8 and are thus sufficient to support the genome protective role of Twi8-sRNA effectors.

Yet another explanation for our findings could be that aberrant processing or silencing of sRNA precursors in *RDN2Δ*, *RSP1Δ*, and *TWI8Δ* cells contribute to the generation of R-loops that can lead to DNA damage (Sollier and Cimprich, 2015). For example, in cycling fission yeast cells, Dicer Dcr1 promotes RNA pol II termination at probable sites of transcription-replication collision, preventing R-loop accumulation and aberrant HR, though the mechanism is independent of Dcr1 catalytic activity and Ago1 (Zaratiegui *et al.*, 2011; Castel *et al.*, 2014). Moreover, a study in human cells found that overexpression of damage-associated sRNA can increase DNA damage in human cells (Hatchi *et al.*, 2021). Given these examples, it is notable that RDRC-independent sRNAs and/or nonabundant sRNAs from low copy number repeats are elevated in *RDN2Δ*, *RSP1Δ*, and *TWI8Δ* cells (Couvillion *et al.*, 2009).

Finally, given the large evolutionary distance that separates ciliates from the more commonly studied eukaryotes mentioned above, it is also possible that ciliates themselves may have evolved novel roles for RNAi in maintaining genome integrity. Indeed, an entirely novel role for sRNAs in chromosome regulation was recently identified in the ciliate *Oxytricha trifallax* for promoting chromosome copy number (Khurana *et al.*, 2018), with the intriguing finding that, like *T. thermophila RSP1Δ* and *RDN2Δ*, *O. trifallax* Dicer and RdRP mutants exhibit CEB-like structures and differential expression of at least a few genes associated with a DNA damage response. In addition, while *RDF2Δ* and *RDF1Δ* exhibited minimal phenotypes in our study, it is possible that these Rdn1-containing RDRC components may act redundantly to suppress DNA damage, consistent with the relatively stronger phenotypes we observed here for *RSP1Δ* cells compared to *RDN2Δ*. Future investigation into the potential roles that different *T. thermophila* ~23–24 nt sRNA RNAi factors play in different aspects of DNA replication and/or repair will illuminate the extent to which specific mechanisms for DNA-directed RNAi activities in protecting the genome are broadly conserved or adapted through evolution in lineage-specific ways.

MATERIALS AND METHODS
Request a protocol through *Bio-protocol*.

Tetrahymena culturing
Wild-type (SB210; *Tetrahymena* Stock Center) and mutant (Couvillion *et al.*, 2009; Lee *et al.*, 2009; Talsky and Collins, 2012) *T. thermophila* strains were simultaneously revived from frozen or slow-growing soy stocks within a month of experimental analyses to keep passage numbers low. Vegetative cell cultures were grown to mid-log densities (~2–4 × 10^5 cells/ml) in 2% proteose peptone, 0.2% yeast extract, 12 µM FeCl, 0.2% glucose, 250 µg/ml ampicillin and streptomycin, and 1.25 µg/ml amphotericin B (HyClone), shaking at 125–150 rpm at 30°C. For culture growth curves, triplicate cultures were seeded at 1 × 10^4 cells/ml from midlog cultures and samples were taken at regular intervals for counting on a hemocytometer until cultures reached stationary phase. For cells treated with hydroxyurea, midlog cultures were treated at a final concentration of 20 mM for 4 h before harvest.

Poly(A)+ RNA sequencing
Three independent midlog cultures of *RSP1Δ*, *RDN2Δ*, *RDF2Δ*, or *RDF1Δ* were grown in parallel to SB210 for a total of eight replicates of SB210. RNA was prepared from pelleted cells using TRIzol extraction (Invitrogen), DNASe I (New England Biolabs; M0303) treatment, phenol:chloroform:isoamyl alcohol (25:24:1) extraction, and ethanol precipitation, followed by resuspension in 10 mM Tris, pH 7.5. RNA (2 µg) from each sample was submitted for poly(A)+ library construction using the Illumina TruSeq stranded RNA preparation kit and sequencing on a 75 base pair SE NextSeq 500 by the CGB Genomics Service Facility at Indiana University, Bloomington, IN.

RNA-seq analysis: differential expression and GO enrichment
Differential expression analysis. Standard analysis tools and parameters were used for processing, alignment, and analysis of demultiplexed and adapter-trimmed sequences. Briefly, Trimmomatic (v0.36; Bolger *et al.*, 2014) was used to remove all 3′ end nucleotides with a quality score less than 30, and reads less than 36 nt were omitted (TRAILING:30 MINLEN:36). Reads were aligned to the June 2014 genome assembly (www.ciliate.org) using Hisat2 (v2.1.0; Kim *et al.*, 2015) with the parameters -k 1 –rna-strandness "R." For each read, the alignment with the best score was kept (-k 1) and the reads were reversely stranded (–rna-strandness "R"). Counts files were generated using featureCounts (v1.5.3; Liao *et al.*, 2014) with the parameters -s 2 -t exon -g gene_name. The alignment was reversely stranded (-s 2), only reads aligning to coding sequence portions of exons were counted (-t exon), and the output counts file used the TTHERM identifier (-g gene name). Using the counts files, differential expression was calculated using DESeq2 (v1.20.0; Love *et al.*, 2014). The DESeq2 design formula was adjusted to account for an observed batch effect likely due to culturing differences. All differentially expressed genes listed in Supplementary Figure S2 had a Benjamini-Hochberg adjusted *p* value ≤ 0.05.

GO enrichment analysis. Because standard biological process GO annotations from geneontology.org only represented 4599 of the ~27,000 genes in the *T. thermophila* genome (accessed 9/20/2018), we collected additional GO annotations added by the ciliate research community to Tetramine (accessed 07/17/2018; Stover *et al.*, 2012) and the *Tetrahymena* Genome Database (ciliate. org, accessed 8/17/2018; Stover *et al.*, 2012). These *T. thermophila* annotations together represented 7321 genes. We also identified novel annotations for *T. thermophila* genes by using a similar

approach to that used previously (Zinkgraf *et al.*, 2018), transferring the GO annotations of the top BLASTp (v2.2.31+) hits with *E* values < 1 × 10^{-5} in the proteomes of *Saccharomyces cerevisiae* (yeastgenome.org; 1/13/15 release), *Schizosaccharomyces pombe* (pombase.org; annotation version 62), *Arabadopis thaliana* (Arabidopsis.org; TAIR10 release), *Drosophila melanogaster* (flybase.org; release 6.23), *P. falciparum* (geneontology.org and uniprot.org; accessed 8/15/2018), and *Homo sapiens* (geneontology.org and uniprot.org; accessed 8/21/2019; see Supplemental Figure S3 for a list of the top BLASTp hits for each TTHERM in each proteome, if found). This latter analysis expanded the number of genes represented by GO annotations to 12,286 genes (see Supplemental Figure S3). GOStats (v2.46.0; Falcon and Gentleman, 2007) hypergeometric test was then used to identify overenriched GO terms in genes differentially expressed with adjusted *p* values ≤ 5 × 10^{-5} between SB210 and *RSP1Δ*, *RDN2Δ*, *RDF2Δ*, or *RDF1Δ*, omitting the GO terms for genes that are not expressed in both strains for each pairwise comparison as suggested by Falcon 2007. Benjamini-Hochberg false discovery rate of 0.1 was used to correct for multiple testing and define which terms were significantly overrepresented compared to their presence in the expressed genome. Cytoscape (v3.5.1; Shannon *et al.*, 2003) was used to visualize the enriched GO terms for each contrast and identify functional groupings ("categories" and "subcategories") of related GO terms, requiring groupings to consist of at least two or more terms representing more than one gene. GO terms for explicitly non-*Tetrahymena*, organism-specific processes were also omitted from groupings. Genes included in Figure 4A and Supplemental Figure S3 were those that 1) contributed to the enrichment of GO terms associated with DNA replication stress, damage response, and/or repair in at least one knockout DE gene set and 2) have been previously found to be involved in these DNA processes or are conserved in the relevant domain structures of the gene's top BlastP hits in at least two other organisms where the BlastP hits have known roles in DNA replication stress, damage response, and/or repair.

Western blot analysis

Cell lysates for SDS–PAGE gels and blot transfers were prepared from midlog cultures by boiling harvested cells in ice-cold lysis buffer (20 mM Tris, pH 8.0, 10% glycerol, 50 mM NaCl, 1 mM MgCl$_2$, 1× protease inhibitor cocktail [Sigma; P8340], and 0.1 mM phenylmethylsulfonyl fluoride). Blots were stained for total protein with 0.1% Ponceau S in 5% acetic acid and probed with 1:500 dilution of α-Rad51 (3C10) in 5% nonfat milk in Tris-buffered saline. ImageStudioLite was used for quantification of band intensities and only signals in the linear range as determined by sample dilution series were used for Ponceau S normalization of Rad51 bands. ANOVA type II using the aov function in the car package in R, which accounts for unbalanced data, was used to compare differences in Rad51 Western blot levels.

EdU labeling, image acquisition, and image analysis

For EdU labeling, log phase cultures (1.5 × 10^5 cells/ml) were pulsed with 200 μM EdU for 30 min, followed by a wash in 10 mM Tris, pH 7.5, and fixation in 1% formaldehyde in 0.5× phosphate-buffered saline (PBS) for 10–20 min. Cells were dried on coverslips and permeabilized in 0.5% NP-40 in 0.5× PBS for 5 min in preparation for click chemistry where cells on coverslips were incubated in 0.2 M triethylammonium acetate, pH 7.0, 70% dimethyl sulfoxide, 1 mM CuSO$_4$, 2 mM BTTAA, 3 μM AlexaFluor 594 Picolyl-Azide, 20 mM sodium ascorbate for 2 h at 45°C in the dark. Coverslips were then washed with 0.1% NP-40 in PBS and DAPI-containing PBS. Images

were collected using a Nikon Eclipse TE2000-U, taking care to image in nonoverlapping sections distributed evenly across coverslips. For quantification, more than 500 cells were counted for each replicate of each strain. Logistic regression using the standard glm function in R with binomial error was used to compare the log odds of EdU-positive cells.

Immunofluorescence, image acquisition, and image analysis

For immunofluorescence, midlog cells were fixed in 2% paraformaldehyde in PBS and stored 6–24 h in PBS at 4°C before drying on coverslips. For Rad51 and γH2A.X staining, cells were permeabilized in 0.4% Triton X-100 before PBS storage (Gao *et al.*, 2013) and dried coverslips rehydrated in PBST (0.5% Tween-20) before blocking in 3% bovine serum albumin (BSA) and 10% normal goat serum in PBST. For phospho-histone H3 (Ser10; pH3S10) staining, cells were permeabilized in 0.1% Triton X-100 in PBS after drying on coverslips and blocked in 5% BSA in PBST. Cells were then probed with appropriate dilutions of primary antibodies (α-γH2A.X [JBW301; 1:5000], α-Rad51 [3C10; 1:100-1:1000], or α-pH3S10 [Sigma; 06-570, 1:5000]), washed in PBST, and stained with DAPI during incubation with appropriate secondary antibodies in 0.5% BSA in PBST. Coverslips were mounted with 90% glycerol, 0.5% N-propyl gallate in 20 mM Tris, pH 8.0. To generate high-resolution images of Rad51 and γH2A.X foci, cells were fixed, permeabilized, blocked, and stained in solution without drying on coverslips with a directly conjugated version of α-γH2A.X–AlexaFluor 488 (JBW301; 1:1000), and imaged in PBS using a water immersion objective.

For Rad51 and α-γH2A.X foci visualization, 63× images were acquired with a Leica Stellaris8 Falcon FLIM laser scanning confocal microscope equipped with a 63× Plan Apo 1.2 NA water immersion objective. Identical scan settings, laser power, and detector gain were used for all images. Raw images were deconvolved with LIGHTNING deconvolution within the LasX software using the default parameters for the 63× 1.2 NA water immersion objective and aqueous mounting media. Spectral scans and secondary antibody controls revealed that minor punctate signals observed outside the nucleus are caused by a combination of autofluorescence and nonspecific secondary antibody binding. For quantification of CEBs or Rad51/γH2A.X staining, 40× images were acquired with LasX software on a Leica DMi6000 microscope with a Leica HCX 40× Plan Fluor 0.6 NA objective and Leica DFC3000G CCD camera, taking care to image in nonoverlapping sections distributed evenly across coverslips. For CEB analysis, autofluorescent structures were distinguished from CEBs as those that fluoresced outside of the wavelengths of DAPI staining and 2° antibodies used (Supplemental Figure S1D), ≥245 cells per strain per replicate were manually counted for CEB enumeration, and ≥50 CEBs per strain per replicate were analyzed for area using a custom, semiautomated ImageJ macro. Briefly, the locations of manually identified CEBs were stored as regions of interest. To quantify the area of each EB, a local maximum entropy threshold was applied to a 50-pixel box centered over each individual EB. The area of the EB was calculated from the thresholded object. In cases where two EBs were adjacent to one another, the area for the EB furthest from the central point was excluded. This process was repeated for each image and strain. ANOVA type II using the aov function in the car package in R, which accounts for unbalanced data, was used to compare differences in CEB areas. To additionally minimize bias in image acquisition of Rad51-/γH2A.X-stained cells, fields of view were initially selected in the DAPI channel before focusing for corresponding Rad51/γH2A.X stains and image acquisition.

For determining the nuclear:cytoplasmic ratios of Rad51 or γH2A.X staining, custom macros were also used. Briefly, Leica.lif files were separated into individual channels (DAPI and Rad51 or DAPI and γH2A.X) using the LOCI Bio-Formats Importer in Fiji. All macronuclei within a single, randomly selected field of view were identified within the DAPI channel using a three-step filter and threshold routine: 1) a maximum intensity Z projection was created from the DAPI channel, 2) the maximum intensity Z projection was filtered with a 5-pixel radius mean filter, and 3) the RenyEntropy threshold was applied to segment the image. For some images, manual adjustment of the threshold was applied to separate adjacent macronuclei. After thresholding, a 100-pixel square box was centered over the centroid of each macronuclear object that was both less than 45 um^2 in area and had a roundness greater than 0.5. This box was used to duplicate both the DAPI channel and the Rad51 or γH2A.X channel to produce a new image stack that contained centered images of each macronucleus from the larger field of view. The new stack was subjected to an average intensity projection to create the average signal within the macronuclei of more than 50 macronuclear images for a single strain replicate, and the camera background was subtracted from this image. The average nuclear signal was calculated by taking the average signal within a 12-pixel box placed at the center of the average image. The average cytoplasmic signal was calculated by taking the average from four 12-pixel boxes placed at each corner of the average image. Importantly, the above approach does not rely on setting a threshold for DNA damage and scoring cells above the threshold. ANOVA type II using the aov function in the car package in R, which accounts for unbalanced data, was used to compare differences in Rad51 and γH2A.X nuclear-to-cytoplasmic ratios.

Data deposition and code availability

The RNA-seq data were deposited into the Sequence Read Archive database with accession number BioProject: PRJNA669066. RNA-seq analysis was performed using a combination of bash-shell and R scripts and image analysis used ImageJ macro scripts; all scripts are available upon request.

REFERENCES

Bodenbender J, Prohaska A, Jauker F, Hipke H, Cleffmann G (1992). DNA elimination and its relation to quantities in the macronucleus of Tetrahymena. Dev Genet 13, 103–110.

Bolger AM, Lohse M, Usadel B (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120.

Castel SE, Ren J, Bhattacharjee S, Chang AY, Sánchez M, Valbuena A, Antequera F, Martienssen RA (2014). Dicer promotes transcription termination at sites of replication stress to maintain genome stability. Cell 159, 572–583.

Cervantes MD, Coyne RS, Xi X, Yao M-C (2006a). The condensin complex is essential for amitotic segregation of bulk chromosomes, but not nucleoli, in the ciliate Tetrahymena thermophila. Mol Cell Biol 26, 4690–4700.

Cervantes MD, Xi X, Vermaak D, Yao M-C, Malik HS (2006b). The CNA1 histone of the ciliate Tetrahymena thermophila is essential for chromosome segregation in the germline micronucleus. Mol Biol Cell 17, 485–497.

Ciccia A, Elledge SJ (2010). The DNA damage response: making it safe to play with knives. Mol Cell 40, 179–204.

Cleffmann G (1980). Chromatin elimination and the genetic organisation of the macronucleus in Tetrahymena thermophila. Chromosoma 78, 313–325.

Cole E, Sugai T (2012). Developmental progression of Tetrahymena through the cell cycle and conjugation. Methods Cell Biol 109, 177–236.

Couvillion MT, Lee SR, Hogstad B, Malone CD, Tonkin LA, Sachidanandam R, Hannon GJ, Collins K (2009). Sequence, biogenesis, and function of diverse small RNA classes bound to the Piwi family proteins of Tetrahymena thermophila. Genes Dev 23, 2016–2032.

Czech B, Hannon GJ (2016). One loop to rule them all: the ping-pong cycle and piRNA-guided silencing. Trends Biochem Sci 41, 324–337.

Daugaard I, Hansen TB (2017). Biogenesis and function of ago-associated RNAs. Trends Genet 33, 208–219.

Eisen JA, Coyne RS, Wu M, Wu D, Thiagarajan M, Wortman JR, Badger JH, Ren Q, Amedeo P, Jones KM, et al. (2006). Macronuclear genome sequence of the ciliate Tetrahymena thermophila, a model eukaryote. PLoS Biol 4, 1620–1642.

Falcon S, Gentleman R (2007). Using GOstats to test gene lists for GO term association. Bioinformatics 23, 257–258.

Farley BM, Collins K (2017). Trans-generational function of Tetrahymena Piwi protein Twi8p at distinctive non-coding RNA loci. RNA 23, 530–545.

Francia S, Cabrini M, Matti V, Oldani A, D'Adda Di Fagagna F (2016). DICER, DROSHA and DNA damage response RNAs are necessary for the secondary recruitment of DNA damage response factors. J Cell Sci 129, 1468–1476.

Francia S, Michelini F, Saxena A, Tang D, De Hoon M, Anelli V, Mione M, Carninci P, D'Adda Di Fagagna F (2012). Site-specific DICER and DROSHA RNA products control the DNA-damage response. Nature 488, 231–235.

Gao M, Wei W, Li MM, Wu YS, Ba Z, Jin KX, Li MM, Liao YQ, Adhikari S, Chong Z, et al. (2014). Ago2 facilitates Rad51 recruitment and DNA double-strand break repair by homologous recombination. Cell Res 24, 532–541.

Gao J, Xiong J, Zhang C, Berquist BR, Yang R, Zhao M, Molascon AJ, Kwiatkowski SY, Yuan D, Qin Z, et al. (2013). Impaired replication elongation in Tetrahymena mutants deficient in histone H3 Lys 27 monomethylation. Genes Dev 27, 1662–1679.

Gorski SA, Vogel J, Doudna JA (2017). RNA-based recognition and targeting: sowing the seeds of specificity. Nat Rev Mol Cell Biol 18, 215–228.

Guo X, Ni J, Liang Z, Xue J, Fenech MF, Wang X (2019). The molecular origins and pathophysiological consequences of micronuclei: new insights into an age-old problem. Mutat Res - Rev Mutat Res 779, 1–35.

Gutbrod MJ, Martienssen RA (2020). Conserved chromosomal functions of RNA interference. Nat Rev Genet 21, 311–331.

Hamilton EP, Kapusta A, Huvos PE, Bidwell SL, Zafar N, Tang H, Hadjithomas M, Krishnakumar V, Badger JH, Caler EV, et al. (2016). Structure of the germline genome of Tetrahymena thermophila and relationship to the massively rearranged somatic genome. Elife 5, 1–46.

Hatchi E, Goehring L, Landini S, Skourti-Stathaki K, DeConti DK, Abderazzaq FO, Banerjee P, Demers TM, Wang YE, Quackenbush J, et al. (2021). BRCA1 and RNAi factors promote repair mediated by small RNAs and PALB2–RAD52. Nature 591.

Kaczanowski A, Kiersnowska M (2018). Formation and degradation of large extrusion bodies in Tetrahymena thermophila: the role of intramacronuclear microtubules in chromatin segregation. Eur J Protistol 66, 177–188.

Khurana JS, Clay DM, Moreira S, Wang X, Landweber LF (2018). Small RNA-mediated regulation of DNA dosage in the ciliate Oxytricha. RNA 24, 18–29.

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

Lee SR, Collins K (2006). Two classes of endogenous small RNAs in Tetrahymena thermophila. Genes Dev 20, 28–33.

Lee SR, Collins K (2007). Physical and functional coupling of RNA-dependent RNA polymerase and Dicer in the biogenesis of endogenous siRNAs. Nat Struct Mol Biol 14, 604–610.

Lee SR, Talsky KB, Collins K (2009). A single RNA-dependent RNA polymerase assembles with mutually exclusive nucleotidyl transferase subunits to direct different pathways of small RNA biogenesis. RNA 15, 1363–1374.

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

Liu M, Ba Z, Costa-Nunes P, Wei W, Li L, Kong F, Li Y, Chai J, Pontes O, Qi Y (2017). IDN2 interacts with RPA and facilitates DNA double-strand break repair by homologous recombination in Arabidopsis. Plant Cell 29, 589–599.

Love MI, Huber W, Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15, 1–21.

Lu WT, Hawley BR, Skalka GL, Baldock RA, Smith EM, Bader AS, Malewicz M, Watts FZ, Wilczynska A, Bushell M (2018). Drosha drives the formation of DNA:RNA hybrids around DNA break sites to facilitate DNA repair. Nat Commun 9, 1–13.

Ly P, Cleveland DW (2017). Rebuilding chromosomes after catastrophe: emerging mechanisms of Chromothripsis. Trends Cell Biol 27, 917–930.

Lynch M, Field MC, Goodson HV, Malik HS, Pereira-Leal JB, Roos DS, Turkewitz AP, Sazer S (2014). Evolutionary cell biology: two origins, one objective. Proc Natl Acad Sci USA 111, 16990–16994.

Mah LJ, El-Osta A, Karagiannis TC (2010). γh2AX: a sensitive molecular marker of DNA damage and repair. Leukemia 24, 679–686.

Marsh TC, Cole ES, Stuart KR, Campbell C, Romero DP (2000). RAD51 is required for propagation of the germinal nucleus in Tetrahymena thermophila. Genetics 154, 1587–1596.

Michalik KM, Böttcher R, Förstemann K (2012). A small RNA response at DNA ends in Drosophila. Nucleic Acids Res 40, 9596–9603.

Michelini F, Pitchiaya S, Vitelli V, Sharma S, Gioia U, Pessina F, Cabrini M, Wang Y, Capozzo I, Iannelli F, et al. (2017). Damage-induced lncRNAs control the DNA damage response through interaction with DDRNAs at individual double-strand breaks. Nat Cell Biol 19, 1400–1411.

Morrison TL, Yakisich JS, Cassidy-Hanley DM, Kapler GM (2005). TIF1 represses rDNA replication initiation, but promotes normal S phase progression and chromosome transmission in Tetrahymena. Mol Biol Cell 16, 2624–2635.

Orias E, Flacks M (1975). Macronuclear genetics of Tetrahymena. I. Random distribution of macronuclear gene copies in T. pyriformis, syngen 1. Genetics 79, 187–206.

Ozata DM, Gainetdinov I, Zoch A, O'Carroll D, Zamore PD (2019). PIWI-interacting RNAs: small RNAs with big functions. Nat Rev Genet 20, 89–108.

Peng JC, Karpen GH (2007). H3K9 methylation and RNA interference regulate nucleolar organization and repeated DNA stability. Nat Cell Biol 9, 25–35.

Peng JC, Karpen GH (2009). Heterochromatic genome stability requires regulators of histone H3 K9 methylation. PLoS Genet 5, 1–14.

Rzeszutek I, Maurer-Alcalá XX, Nowacki M (2020). Programmed genome rearrangements in ciliates. Cell Mol Life Sci 77, 4615–4629.

Sabatinos SA, Ranatunga NS, Yuan JP, Green MD, Forsburg SL (2015). Replication stress in early S phase generates apparent micronuclei and chromosome rearrangement in fission yeast. Mol Biol Cell 26, 3439–3450.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13, 2498–2504.

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

Song X, Gjoneska E, Ren Q, Taverna SD, Allis CD, Gorovsky MA (2007). Phosphorylation of the SQ H2A.X motif is required for proper meiosis and mitosis in Tetrahymena thermophila. Mol Cell Biol 27, 2648–2660.

Stover NA, Punia RS, Bowen MS, Dolins SB, Clark TG (2012). Tetrahymena genome database Wiki: a community-maintained model organism database. Database 2012, 1–9.

Talsky K, Collins K (2012). Strand-asymmetric endogenous Tetrahymena small RNA production requires a previously uncharacterized uridylyl-transferase protein partner. RNA 18, 1553–1562.

Tan EH, Henry IM, Ravi M, Bradnam KR, Mandakova T, Marimuthu MPA, Korf I, Lysak MA, Comai L, Chan SWL (2015). Catastrophic chromosomal restructuring during genome elimination in plants. Elife 4, 1–16.

Upton HE, Hong K, Collins K (2014). Direct single-stranded DNA binding by Teb1 mediates the recruitment of Tetrahymena thermophila telomerase to telomeres. Mol Cell Biol 34, 4200–4212.

Wang Q, Goldstein M (2016). Small RNAs recruit chromatin-modifying enzymes MMSET and Tip60 to reconfigure damaged DNA upon double-strand break and facilitate repair. Cancer Res 76, 1904–1915.

Wei W, Ba Z, Gao M, Wu Y, Ma Y, Amiard S, White CI, Danielsen JMR, Yang YG, Qi Y (2012). A role for small RNAs in DNA double-strand break repair. Cell 149, 101–112.

Wei Y, Mizzen CA, Cook RG, Gorovsky MA, Allis CD (1998). Phosphorylation of histone H3 at serine 10 is correlated with chromosome condensation during mitosis and meiosis in Tetrahymena. Proc Natl Acad Sci USA 95, 7480–7484.

Wiley EA, Myers T, Parker K, Braun T, Yao MC (2005). Class I histone deacetylase Thd1p affects nuclear integrity in Tetrahymena thermophila. Eukaryot Cell 4, 981–990.

Wilhelm T, Said M, Naim V (2020). DNA replication stress and chromosomal instability: dangerous liaisons. Genes (Basel) 11, 1–34.

Wong L, Klionsky L, Wickert S, Merriam V, Orias E, Hamilton EP (2000). Autonomously replicating macronuclear DNA pieces are the physical basis of genetic coassortment groups in Tetrahymena thermophila. Genetics 155, 1119–1125.

Yakisich JS, Sandoval PY, Morrison TL, Kapler GM (2006). TIF1 activates the intra-S-phase checkpoint response in the diploid micronucleus and amitotic polyploid macronucleus of Tetrahymena. Mol Biol Cell 17, 5185–5197.

Yang Y, Qi Y (2015). RNA-directed repair of DNA double-strand breaks. DNA Repair (Amst) 32, 82–85.

Zaratiegui M, Castel SE, Irvine DV, Kloc A, Ren J, Li F, De Castro E, Marín L, Chang AY, Goto D, et al. (2011). RNAi promotes heterochromatic silencing through replication-coupled release of RNA Pol II. Nature 479, 135–138.

Zhang Z, Chang S, Xue Z, Zhang H, Li S, Liu Y (2013). Homologous recombination as a mechanism to recognize repetitive DNA sequences in an RNAi pathway. Genes Dev 27, 145–150.

Zhao X, Xiong J, Mao F, Sheng Y, Chen X, Feng L, Dui W, Yang W, Kapusta A, Feschotte C, et al. (2019). RNAi-dependent Polycomb repression controls transposable elements in Tetrahymena. Genes Dev 33, 348–364.

Zinkgraf M, Gerttula S, Zhao S, Filkov V, Groover A (2018). Transcriptional and temporal response of Populus stems to gravi-stimulation. J Integr Plant Biol 60, 578–590.