RNA-DNA hybrids on protein coding genes are stabilized by loss of RNase H and are associated with DNA damages during S-phase in fission yeast.

Tomoko Sagi[1], Daichi Sadato[2], Kazuto Takayasu[1,3], Hiroyuki Sasanuma[1], Yutaka Kanoh[1], Hisao Masai[1]

[1]Genome Dynamics Project, Department of Basic Medical Sciences, Tokyo Metropolitan Institute of Medical Science, Setagaya-ku, Tokyo 156-8506, Japan

[2] Clinical Research and Trials Center, Tokyo Metropolitan Cancer and Infectious Diseases Center, Komagome Hospital, Tokyo, Japan

[3] Department of Biosciences, School of Science, Kitasato University, Sagamihara, Kanagawa, Japan

Running title: RNase H maintains the genomic integrity.

Key words: RNA-DNA hybrid, R-loop, RNase H, Rad52, DNA damages

Correspondence should be addressed to

Yutaka Kanoh

E-mail: kanou-yt@igakuken.or.jp

Tel: +81-3-5316-3117, Extension 1692

Abstract

RNA-DNA hybrid is a part of the R-loop which is an important non-standard nucleic acid structure. RNA-DNA hybrid/R-loop causes genomic instability by inducing DNA damages or inhibiting DNA replication. It also plays biologically important roles in regulation of transcription, replication, recombination and repair. Here, we have employed catalytically inactive human RNase H1 mutant (D145N) to visualize RNA-DNA hybrids and map their genomic locations in fission yeast cells. The RNA-DNA hybrids appear as multiple nuclear foci in *rnh1*Δ*rnh201*Δ cells lacking cellular RNase H activity, but not in the wild-type. The majority of RNA-DNA hybrid loci are detected at the protein coding regions and tRNA. In *rnh1*Δ*rnh201*Δ cells, cells with multiple Rad52 foci increase during S-phase and about 20% of the RNA-DNA hybrids overlap with Rad52 loci. During S-phase, more robust association of Rad52 with RNA-DNA hybrids was observed in the protein coding region than in M-phase. These results suggest that persistent RNA-DNA hybrids in the protein coding region in *rnh1*Δ*rnh201*Δ cells generate DNA damages during S-phase, potentially through collision with DNA replication forks.

Introduction

The genomic stability is tightly regulated through multiple layers of mechanisms, including rapid detection of stalled replication forks and DNA damages, and repair of any lesions on the genome, to prevent mutations and detrimental gene rearrangement. Genomic instability is induced by DNA damage caused through exogenous incidents, UV irradiation and DNA damaging drugs, or through endogenous cellular events during the course of DNA replication and transcription. Double-strand breaks (DSB) are repaired though two major pathways, homologous recombination (HR) and non-homologous end joining (NHEJ). These pathways are well conserved from yeast to human. The HR pathway accurately repairs DSBs and is essential for maintaining genomic stability. In this pathway, DNA double-stand breaks are recognized and resected by the Mre11-Rad50-Nbs1 complex, and the single-strand DNA overhangs are generated. The exposed single strand DNA is coated with single strand binding protein RPA, and the Rad52 heptamer replaces RPA with Rad51 recombinase. Rad51-DNA filaments invade into the homologous sister chromatid and form D-loops. The invading DNA strand is extended on the sister chromatid strands by a DNA polymerase for repair of DSB (Scully, Panday, Elango, & Willis, 2019). Failure in this process increases genomic instability, ultimately leading to cancer and neurodegenerative disorders.

RNA-DNA hybrid is a part of R-loop and is associated with various important cellular functions including genomic DNA replication, mitochondrial DNA replication, transcription, telomere maintenance and recombination such as immunoglobulin class switching in B cells (Yu, Chedin, Hsieh, Wilson, & Lieber, 2003). On the other hand, unscheduled occurrence of RNA-DNA hybrids on the genome is known to cause genomic instability (Rondón & Aguilera, 2019). Therefore, the regulation of RNA-DNA hybrid formation and resolution must be tightly controlled to maintain genome stability. RNA-DNA hybrids are eliminated by ribonuclease H (RNase H). Other helicases can also resolve RNA-DNA hybrids. These include senataxin (SETX) (Mischo et al., 2011; Skourti-Stathaki, Proudfoot, & Gromak, 2011), Aquarius (AQR) (Sollier et al., 2014), WRN RecQ like helicase (WRN), BLM RecQ like helicase (BLM), regulator of telomere elongation helicase 1 (RTEL1) (Kotsantis et al., 2020), PIF1 (Chib, Byrd, & Raney, 2016; Osmundson, Kumar, Yeung, & Smith, 2017), ATRX, and FA complementation group M (FANCM). Mutations in these genes cause aberrant accumulation of RNA-DNA hybrids in cells, increasing genomic instability.

In mammalian cells, RNase H is encoded by RNase H1 and RNase H2 (Cerritelli & Crouch, 2009). RNase H1 degrades RNA-DNA hybrids by hydrolyzing the RNA strands (Cerritelli & Crouch, 2009; Nowotny, Gaidamakov, Crouch, & Yang, 2005; Nowotny et al., 2007). RNase H2 consists of three subunits (Crow et al., 2006; Jeong, Backlund, Chen, Karavanov, & Crouch, 2004) and can remove

ribonucleoside monophosphates (rNMPs) misincorporated into DNA in addition to degrading RNA strand of RNA-DNA hybrid (Cerritelli & Crouch, 2009; Hiller et al., 2012; Reijns et al., 2012; Sparks et al., 2012). RNase H1 is constitutively expressed throughout cell cycle, while RNase H2 is expressed at a higher level from late S to G2-phase in budding yeast (Lockhart et al., 2019). Thus, cellular RNA-DNA hybrid levels are tightly regulated.

Defects in RNase H activity cause accumulation of pathogenic R-loop or RNA-DNA hybrids that can block replication fork progression, resulting in Replication-Transcription Conflicts (RTC) (Helmrich, Ballarino, & Tora, 2011). RNase H1 knockout mice are embryonic lethal due to the effect on mitochondrial DNA replication (Cerritelli et al., 2003; Lima et al., 2016). Mutations in RNase H2 are known to cause Aicardi-Goutières syndrome (AGS) (Crow et al., 2006).

In this study, by using *Schizosaccharomyces pombe* (*S. pombe*) as a model, we determined the genomic profiles of RNA-DNA hybrids/ R-loops and examined if they are associated with DNA damages. We show that RNA-DNA hybrids form nuclear foci in the RNase H mutant. Approximately 80% of the RNA-DNA hybrids by ChIP-seq was detected at the Protein coding region and tRNA. 37 or 20% of the RNA-DNA hybrid peaks are associated with DNA damages detected by Rad52 in wild-type or RNase H mutant cells, respectively. Cells with multiple Rad52 foci increase during S-phase in the RNase H mutant and Rad52 association increase around the RNA-DNA hybrids in the protein coding region during S-phase, suggesting that RTC may cause DNA damages.

Results

Identification of RNA-DNA hybrids in *S. pombe* using human catalytic inactive RNASE H1 probe.

To identify RNA-DNA hybrids in *S. pombe*, we utilized human catalytically inactive RNASE H1. The active center in the catalytic domain of RNASE H1 has been identified at the residues D145, E186 or D210 (Nowotny et al., 2008; Nowotny et al., 2007; Wu, Lima, & Crooke, 2001). The D145N mutant of human RNASE H1 does not degrade RNA-DNA hybrids, but can still recognize and bind to RNA-DNA hybrids. We used the D145N mutant of RNASE H1 as the probe for RNA-DNA hybrids. Residues 1-26 amino acids containing the mitochondrial targeting sequence (MTS) were eliminated, replaced with the nuclear localization signal (NLS), and 3×FLAG were fused to the C-terminus. This probe was expressed under the inducible nmt41 (No Message in Thiamine 1) promoter on pREP41 (**Figure S1a**). The plasmid was transformed into wild-type, *rnh1*Δ, *rnh201*Δ, and *rnh1*Δ*rnh201*Δ strains. *rnh1* encodes RNase H1 that degrades RNA strand on RNA-DNA hybrids. *rnh201* encodes RNase H2A, one of the subunits in the heterotrimeric complex of the RNase H2. RNase H2 removes single rNMPs in RNA-DNA hybrids for ribonucleotide excision repair (RER) activity and can degrade the RNA strand of RNA-DNA hybrid. Expression of the RNASE H1(D145N) probe was induced for 18 h in medium without thiamine and its expression was confirmed by western blotting analysis (**Figure S1b**).

Next, we fused EGFP to NLS-RNASE H1(D145N) at the C-terminus to visualize the RNA-DNA hybrids formed in the cells. The expression of RNASE H1(D145N)-EGFP was induced for 18 h in medium without thiamine. RNASE H1(D145N)-EGFP signals were detected in the nuclei, and the stronger signals were observed in the nucleoli, which was weakly stained with Hoechst 33342. Nucleoplasmic foci were detected in *rnh1*Δ cells, but not in the wild-type or barely in *rnh201*Δ cells. The numbers of the cells with multiple foci dramatically increased in the *rnh1*Δ*rnh201*Δ cells compared to *rnh1*Δ or *rnh201*Δ cells (**Figure 1a,b**). The results suggested that this probe could detect the RNA-DNA hybrids that were not resolved due to the lack of RNase H1 and RNase H2 functions. Cell shapes were elongated in *rnh1*Δ*rnh201*Δ cells accumulating DNA damage as reported (Zhao, Zhu, Limbo, & Russell, 2018). Furthermore, nuclei were also stretched in the elongated cells. We validated the detected foci by examining whether exogeneous expression of a RNase HI would diminish the signals. Indeed, the expression of *E. coli* RNase HI decreased the number of observed RNA-DNA hybrid foci in *rnh1*Δ*rnh201*Δ cells (**Figure 1c,d**). These results show that the RNASE H1(D145N)-EGFP probe can recognize and visualize RNA-DNA hybrid in fission yeast cells.

RNA-DNA hybrid signals do not increase in RNA/DNA helicase mutant strains.

RNA-DNA hybrids are known to be resolved by helicases, including Sen1, a homologue of senataxin (SETX) encoding an RNA/DNA helicase (Kim, Choe, & Seo, 1999), Rqh1, a RecQ type DNA helicase (Ahmad, Kaplan, & Stewart, 2002), and Srs2 (Rhind et al., 2011), a DNA helicase in the UvrD subfamily. We tested whether RNA-DNA hybrid foci would be affected in these helicase defective mutants. RNASE H1(D145N)-EGFP was expressed in these mutants and the cells were observed under fluorescence microscope (**Figure S2**). The nucleoplasmic RNA-DNA hybrid foci, observed in RNase H mutants, were not observed. The results suggest that these helicases would unwind RNA-DNA hybrids rather locally at specific loci, but not extensively, in comparison to the RNase H that would degrade RNA-DNA hybrids more extensively along the genome.

RNASE H1(D145N) probe recognizes RNA-DNA hybrids in *S. pombe*.

We observed that the number of RNA-DNA hybrid foci increased in living *rnh1*Δ*rnh201*Δ cells compared to wild-type cells. Next, we mapped the RNA-DNA hybrid forming loci by chromatin immunoprecipitation (ChIP) seq using the RNASE H1(D145N) in wild-type and *rnh1*Δ*rnh201*Δ cells to compare the RNA-DNA hybrid peaks in each strain. The wild-type or *rnh1*Δ*rnh201*Δ cells harboring pREP41-NLS-RNASE H1(D145N)-3×FLAG was used for ChIP-seq analyses. ChIP-seq data show that the peaks on some gene loci detected in *rnh1*Δ*rnh201*Δ cells exhibited higher peak intensity than that in wild-type cells (**Figure 2a**) The total numbers of RNase H1(D145N)-3×FLAG peaks detected was 322 and 496 in wild-type and *rnh1*Δ*rnh201*Δ cells, respectively.

We conducted by ChIP-qPCR to validate whether detected peaks were RNA-DNA hybrid and confirm the peak intensity fluctuation. The result shows that the IP efficiency increased in *rnh1*Δ*rnh201*Δ cells compared to the wild-type at most of loci observed in ChIP-seq result (**Figure 2b**). It was reported that RNA-DNA hybrid peaks detected by DNA-RNA immunoprecipitation (DRIP)-seq at the tRNA region were higher in *rnh1*Δ*rnh201*Δ cells compared to the wild-type cells (El Hage, Webb, Kerr, & Tollervey, 2014), consistent with the results shown in this study. To test whether the detected RNA-DNA hybrids were resolved by co-expression of exogeneous RNase H, we expressed RNase HI derived from *E. coli* in both wild-type and the *rnh1*Δ*rnh201*Δ cells expressing the RNASE H1(D145N) probe and then performed ChIP-qPCR (**Figure 2c**). IP efficiencies at almost all the sites in the *rnh1*Δ*rnh201*Δ cells decreased significantly, to the levels exhibited in the wild-type cells. These results led us to conclude that the peaks detected by the RNASE H1(D145N) probe represent RNA-DNA hybrids.

RNA-DNA hybrids accumulate at protein coding region in RNase H mutant.

Next, we conducted detailed analyses of RNA-DNA hybrids on the genome. The peaks of 3×FLAG as a negative control were detected at 22 and 46 loci in wild-type and *rnh1*Δ*rnh201*Δ cells, respectively. These peaks, except for one locus in *rnh1*Δ*rnh201*Δ cells, overlapped with the peaks of RNA-DNA

hybrid. The overlapping peaks were eliminated as non-specific peaks. As a result, we identified 300 and 451 peaks of RNA-DNA hybrid in wild-type and *rnh1*Δ*rnh201*Δ cells, respectively (**Figure S3a**). Among all the peaks detected in the wild-type and *rnh1*Δ*rnh201*Δ cells combined, 43.0% (226/526), 14.2% (75/526) and 42.8% (225/526) were detected only in *rnh1*Δ*rnh201*Δ, only in the wild-type, and in common, respectively. We analyzed in detail the functional annotations and gene types where RNA-DNA hybrids were located on the genome in the wild-type and *rnh1*Δ*rnh201*Δ cells. Approximately 63% of the detected RNA-DNA hybrid peaks were present at promoter-TSS in both wild-type and *rnh1*Δ*rnh201*Δ cells (190/300 and 282/451, respectively; **Figure S3b**). This result is consistent with previously reported characteristics of this probe (Castillo-Guzman & Chédin, 2021; Chen et al., 2017). Among all the RNA-DNA hybrids detected, 43.7% (131/300), 33.7% (101/300), and 22.6% (68/300) were at protein coding region, tRNA and other genes, respectively, in the wild-type cells, whereas the distribution was 55.4% (250/451), 26.0% (117/451) and 18.6% (84/451), respectively, in *rnh1*Δ*rnh201*Δ cells (**Figure 2d, Table S1, Table S2**). Notably, the fraction of RNA-DNA hybrid peaks in the protein coding region increased in the *rnh1*Δ*rnh201*Δ cells compared to the wild-type cells.

Furthermore, we analyzed peak distributions in "wild-type unique", "common", and "*rnh1*Δ*rnh201*Δ unique" groups. The promoter-TSS region occupied the largest fraction in all the groups. The fraction of TSS was largest in "common", whereas that of TTS was smallest (**Figure S3c**). In the peak distribution of gene types, the fraction of tRNA was the largest in "common", whereas it occupied small fractions in "wild-type unique" and "*rnh1*Δ*rnh201*Δ unique". The fractions of protein coding regions were 61.3% (46/75) and 73.0% (165/226) in "wild-type unique" and "*rnh1*Δ*rnh201*Δ unique", respectively, whereas it was 38.1% (86/225) in "common" (**Figure S3d**). tRNA and rRNA were predominant in "common" and the peak score of these loci was higher in *rnh1*Δ*rnh201*Δ cells than these in wild-type cells (**Figure S4a**).

To evaluate the peak intensity variation between wild-type and *rnh1*Δ*rnh201*Δ cells, we analyzed the peak distributions by Z-scoring normalization in each strain and compared these distributions. In the *rnh1*Δ*rnh201*Δ cells, the number of high-value peaks ($0 < Z$) slightly increased and that of low-value peaks ($-1.0 < Z < 0$) tended to increase compared to the wild-type cells (**Figure 2e,** Whole). We then broke down the peak distributions according to gene types. The peak numbers of protein coding regions particularly increased within $-1.0 < Z < 0$ scores, whereas those of tRNA was shifted from lower Z-score to higher Z-score ($1.0 < Z$) in the *rnh1*Δ*rnh201*Δ cells compared to the wild-type cells (**Figure 2e, Figure S4b, Table S1, Table S2**). The results indicated that the RNA-DNA hybrids escaped from degradation and accumulated at the protein coding regions in the *rnh1*Δ*rnh201*Δ cells, whereas those at the highly transcribed tRNA and rRNA genes could not be completely removed by

211 endogenous RNase H1 and H2 in the wild-type cells, albeit with the reduction of peak intensities.
212 These results support that RNA-DNA hybrids are efficiently formed at highly transcribed tRNA and
213 rRNA regions in *S. pombe*, and they are not completely eliminated by cellular RNase H. On the other
214 hand, the RNA-DNA hybrids at the protein coding region may be more efficiently degraded by
215 endogenous RNase H in the wild-type cells.
216
217 **A fraction of the RNA-DNA hybrids overlaps with DNA damages.**
218 It has been reported that DNA damage Rad52 foci increased in *rnh1*Δ*rnh201*Δ cells in *S. pombe* (Zhao
219 et al., 2018). Therefore, we first examined whether Rad52 damage foci accumulate in *rnh1*Δ*rnh201*Δ
220 cells. The wild-type or *rnh1*Δ*rnh201*Δ cells, expressing Rad52 fused with mCherry at its C-terminus,
221 were grown to log phase. Rad52 foci were observed more frequently in *rnh1*Δ*rnh201*Δ cells than in
222 wild-type (**Figure 3a,b**). This result prompted us to investigate whether unresolved RNA-DNA
223 hybrids are involved in the DNA damage detected by Rad52. We performed Rad52 ChIP-seq to
224 compare the genome profiles of Rad52 binding and RNA-DNA hybrid formation (**Figure 3c**).
225 Unexpectedly, the distributions of Rad52 binding regions did not change between the wild-type and
226 the *rnh1*Δ*rnh201*Δ cells (**Figure S5a, Table S3, Table S4**). We further examined the distributions of
227 the peaks in each of the "RNA-DNA hybrid unique", "overlap", and "Rad52 unique" groups (**Figure
228 3d, Figure S5b, Table S5, Table S6**). 36.7% (110/300) and 20.0% (90/451) of the RNA-DNA hybrid
229 peaks overlapped with Rad52 peaks in the wild-type and *rnh1*Δ*rnh201*Δ cells, respectively (**Figure
230 S5b, Table S5, Table S6**). The fraction of protein coding regions significantly increased in
231 *rnh1*Δ*rnh201*Δ cells among the overlapping peaks (**Figure 3d,** overlap). This result is consistent with
232 the increased RNA-DNA hybrid peaks at the protein coding regions in *rnh1*Δ*rnh201*Δ cells in **Figure
233 2d**.
234
235 **Rad52 foci accumulate during S-phase in *rnh1*Δ*rnh201*Δ cells.**
236 When DNA replication forks collide with unresolved RNA-DNA hybrids, DNA lesions may be
237 generated, leading to increased genome instability (Helmrich et al., 2011). To investigate whether the
238 appearance of the Rad52 foci depends on cell cycle, we observed them in the synchronized cells. Cells
239 were arrested at M-phase by *nda3-KM311* and then released into cell cycle. Cell cycle progression
240 was monitored by flow cytometry and the late S-phase population was enriched at 60 min after release
241 (**Figure S6**). The cells were observed under fluorescence microscope every 30 min for 2 h after release
242 (**Figure 4a,b**). In the wild-type cells, the number of cells with a single Rad52 focus slightly increased
243 from 60 min to 90 min after release. In *rnh1*Δ*rnh201*Δ cells, in contract, ~35% of cells showed a single
244 Rad52 focus at the time of release, and the cells with multiple Rad52 foci strikingly increased, as cells
245 progressed through S-phase, although intensity of each focus was lower than that of a single focus
246 observed in a wild-type cell (**Figure 4b,c, Figure S7**). These results show that DNA damages

spontaneously accumulate in *rnh1*Δ*rnh201*Δ cells during S-phase.

Unresolved RNA-DNA hybrids may collide with replication forks and increase genomic instability.

In *rnh1*Δ*rnh201*Δ cells, RNA-DNA hybrids were not resolved in the protein coding region, and DNA damages accumulated there, leading to the dramatic increase of cells with multiple Rad52 foci during S-phase in *rnh1*Δ*rnh201*Δ cells. Rad52 binding regions were analyzed by ChIP-seq during cell cycle progression. The Rad52 peaks were assigned to the gene types. The number and gene type distribution of the peaks at 0 min (M-phase) were similar to those at 60 min (S-phase) in the wild-type. In *rnh1*Δ*rnh201*Δ cells, in contrast, the Rad52 peaks on the protein coding region increased two times at 60 min compared to those at 0 min (**Figure 5a**). The Rad52 binding loci at 0 min or 60 min were compared with RNA-DNA hybrid peaks in wild-type and *rnh1*Δ*rnh201*Δ cells. Rad52 was recruited near RNA-DNA hybrids of highly transcribing tRNA throughout the cell cycle in both strains. Rad52 binding was more robustly associated with the RNA-DNA hybrid signals within 3 kb of the RNA-DNA hybrid peaks in the protein coding region at 60 min (S-phase) than at 0 min (M-phase; **Figure 5b**).

Furthermore, we broke down the Rad52 peaks in close association with RNA-DNA hybrid according to gene types (**Figure 5c**). Peaks of the protein coding region at 60 min shifted down to a lower Z-score in wild-type as well as in *rnh1*Δ*rnh201*Δ cells compared to the peaks at 0 min. Some peaks did not change and were still at higher score ($0 < Z < 2.0$). These peaks may contain tRNA because the distribution of tRNA peaks around RNA-DNA hybrids did not change at 0 min and 60 min (**Figure 5b**). In fact, Rad52 binding was equally observed at strong RNA-DNA hybrid peaks at both 0 and 60 min (**Figure S8a,b**). The increased number of peaks may be due to collision of DNA replication forks with unresolved RNA-DNA hybrids during S-phase. It was reported that unprocessed R-loops collapse replication forks (Zhao et al., 2018). These results suggest that DNA damage increased on the entire genome during S-phase in *rnh1*Δ*rnh201*Δ cells. Rad52 appears to be recruited to RNA-DNA hybrids in more selected regions such as tRNA (**Figure 3d**, overlap; ex. tRNA), which is highly transcribed and tends to cause topological problems. In fact, the mutation rate of tRNA loci is reported to be 7 to 10 times higher than that of the entire genome (Milano, Gautam, & Caldecott, 2024). At 60 min in wild-type, DNA damage can be caused more globally through conflicts between RNA-DNA hybrids or transcription with replication. As a result, Rad52 recruited sites were dispersed and the score decreased, although the higher score peaks at 0 min were still maintained at 60 min (**Figure 5c**, wild-type). In *rnh1*Δ*rn201*Δ, the high score peaks also remained and the protein coding peaks increased. Due to collision with forks, low score peaks significantly increased compared to those at 0 min (**Figure 5c**, *rnh1*Δ*rn201*Δ). The fraction of Rad52 overlapping with RNA-DNA hybrids on the protein coding

region remarkably increased in *rnh1*Δ*rnh201*Δ compared to that in the mutant at 0 min or to that in the wild-type at 60 min.

Taken together, these results indicate that the unresolved RNA-DNA hybrids on the protein coding region accumulate in *rnh1*Δ*rnh201*Δ cells and that Rad52 is recruited to these RNA-DNA hybrid loci, especially during S-phase.

Discussion

Probes have been developed that can detect RNA-DNA hybrids. These include S9.6 antibody and a hybrid binding domain of RNase H1 (Boguslawski et al., 1986; Chan et al., 2014; Ginno, Lott, Christensen, Korf, & Chédin, 2012; Nowotny et al., 2008; Wang et al., 2021). These probes have been shown to have some specificity in terms of target recognition (Castillo-Guzman & Chédin, 2021). Catalytically inactive RNASE H1 proteins have been shown to recognize RNA-DNA hybrid in the promoter region, while signals detected by S9.6 antibody spreads more broadly on genes. S9.6 antibody was developed against the transcription reaction mix generated on a single-stranded DNA, and thus, it may recognize not only RNA-DNA hybrid but also template DNA as well as free RNA including dsRNA (Boguslawski et al., 1986). Indeed, analyses of RNA-DNA hybrids by S9.6 antibody showed that it recognizes dsRNA which is degraded by RNase III (Hartono et al., 2018). In contrast, catalytically inactive RNASE H1 has very low affinity with dsRNA (Crossley et al., 2021). Thus, RNASE H1(D145N) probe is expected to have higher specificity toward RNA-DNA hybrids.

In this work, we have used RNASE H1(D145N), a catalytically inactive derivative of human RNase H1, to detect RNA-DNA hybrids in fission yeast cells. The fluorescent probe was able to detect signals that were enhanced by loss of RNase H genes and were reduced by expression of exogenous RNase H derived from *E. coli*, indicating that the observed signals represent RNA-DNA hybrids.

We observed strong RNase H1(D145N)-EGFP signals in **nucleoli. *S. pombe* ribosomal DNA (rDNA) tandem repeats are compartmentalized within the nucleolus. rDNA consists of 100-120 tandem repeats of a 10.4-kb fragment containing the 5.8S, 18S and 25S ribosomal RNA (rRNA) genes and account for around 1.1 Mb in the two telomeric regions of chromosome III** (Wood et al., 2002). **rRNA is actively transcribed and large amounts of rRNA are generated from the rDNA repeats.** RNA-DNA hybrids may be formed on the rDNA tandem repeat region during rRNA transcription. However, because the rDNA repeat copy numbers were variable in each cell, it is difficult to quantify the efficiency of RNA-DNA hybrid formation on rDNA repeats by ChIP analyses. Alternatively, it has also been reported that S9.6 immunofluorescence signals were observed in the **nucleoli** of human cells, but the signals were not derived from RNA-DNA hybrids because RNase H1 treatment did not affect the signals in the **nucleoli** (Smolka, Sanz, Hartono, & Chédin, 2021). It is possible that the probe recognized rRNA and that the signals are accumulated in the **nucleoli** because the catalytically inactive RNase H1 protein has a low affinity for dsRNA (Nowotny et al., 2008).

Previous study of DRIPc-seq using the S9.6 antibody on *S. pombe* indicated the presence of two types of RNA-DNA hybrids/R-loops on the basis of the sensitivity to RNase H, the peaks observed in the wild-type (class A) and the other peaks observed in the RNase H mutant (class B). Class B genes are

transcribed less robustly than class A (Hartono et al., 2018). In our analyses, the highly transcribed tRNA was detected relatively more abundantly by RNase H1(D145N) in wild-type than in *rnh1*Δ*rnh201*Δ cells. And more than half of the peaks observed in *rnh1*Δ*rnh201*Δ cells represent low intensity peaks such as those found in protein coding genes. This is consistent with the previous report that tRNA loci belong to the class A, while protein coding regions to the class B which are generally more readily destabilized by RNase H.

Based on the genome-wide mapping data from various species, RNA-DNA hybrids/R-loops were categorized into two classes, Class I (promoter-paused R-loops) and Class II (Elongation-associated R-loops) (Castillo-Guzman & Chédin, 2021). The genomic RNA-DNA hybrids mapped by RNASE H1(D145N) probe in human cells were predominantly Class I R-loops that form at the promoter-proximal regions of paused promoters. This is consistent with the results of the RNASE H1(D145N) probe in *S. pombe*, where approximately 63% of the RNA-DNA hybrids were detected at the promoter-TSS (**Figure S3c**).

Furthermore, detailed analyses show that RNA-DNA hybrids accumulate at the promoter-TSS of the protein coding region in *rnh1*Δ*rnh201*Δ cells compared to the wild-type cells. This may suggest that in *S. pombe*, stable RNA-DNA hybrids are rarely formed in the gene bodies. This could be due to relatively short gene sizes and AT-rich nature of the *S. pombe* genome, which makes the chance of RNA-DNA hybrid formation low. The reason for the detection of tRNA and 5S RNA as strong RNA-DNA hybrid loci is not clear (**Figure S4a**, **Table S1, 2**). tRNA and 5S rRNA are strongly transcribed by RNA polymerase III and have a high propensity to form secondary structures, which may facilitate the formation of RNA-DNA hybrids.

We adapted Z-scoring to evaluate the peak intensity of RNA-DNA hybrids and Rad52 bindings in this study. Although Z-scoring can evaluate the ChIP peaks intensity comparatively within each strain, precise quantitative comparison is difficult between different strains since they possibly have different peak distribution. It is necessary to use spike in DNA for ChIP-seq analysis to quantitatively compare the intensities of RNA-DNA hybrid peaks between different strains. Thus, our Z-score analyses (**Figure 2e**, **Figure 5c**) should be treated only as “semi-quantitative” data.

The presence of RNA-DNA hybrids/ R-loops and their potential to induce genome instability has been reported in several studies (Castillo-Guzman & Chédin, 2021), and it has also been reported that *rnh1*Δ*rnh201*Δ cells accumulate DNA damages from collapsed replication forks (Zhao et al., 2018). In this report, we have shown that Rad52 foci increase during S-phase (**Figure 4b,c**), and that RNA-DNA hybrids colocalize with Rad52 binding sites (**Figure 5b**). We speculate that this is due to collision

364 of the replication fork with the unresolved RNA-DNA hybrids, leading to DSB. Chk1 is constitutively
365 activated in RNase H mutants (Zhao et al., 2018), suggesting that RNA-DNA hybrids and associated
366 DSBs are present outside S-phase as well.
367 The RNA-DNA hybrid and Rad52 loci overlap, and they increase during S-phase (**Figure S5, Figure**
368 **S8**), but we do not know what distinguishes damage-inducing RNA-DNA hybrid and more latent
369 RNA-DNA hybrids. RNA-DNA hybrids accumulate at the protein coding region in *rnh1*Δ*rnh201*Δ
370 cells, and Rad52 is also recruited to the promoter-TSS of the protein coding region close to RNA-
371 DNA hybrids. Rad52 binding scores are elevated during S-phase. This indicates that DNA at the
372 promoter-TSS of the protein coding region is damaged in RNase H mutant cells, and is consistent with
373 the speculation that DNA replication forks are collapsed at the sites of unresolved RNA-DNA hybrids.
374 DNA damage in the protein coding region potentially induces the mutation or deletion in the gene,
375 resulting in critical events such as the cell death, cancer and diseases. Thus, our results reinforce the
376 importance of the cellular RNase H activity in maintaining the genome stability by reducing generation
377 of RNA-DNA hybrids in the protein coding regions.
378
379
380

Experimental procedures

Strains and Medium

All the strains and plasmids used in this study are listed in **Table 1**. Epitope-tagged strains were made by integrating a 3×FLAG-, 6×PK- or mCherry-tagged gene fragment into the endogenous genes. All the tags were located at the C-terminus. Yeast extract with supplements (YES) medium containing 0.5% yeast extract, 3% glucose and 0.1 mg/mL each of adenine, uracil, leucine, lysine and histidine was used for cell culture. YES plates were made by adding 2% agar to YES medium. Synthetic dextrose minimal (SD) medium contains 6.3g/L Yeast Nitrogen Base w/o Amino Acids (DB DIFCO™, 233520), 2% glucose and 0.1 mg/mL each of the required amino acids. Pombe Minimal Glutamate (PMG) medium contains 27.3 g/L EMM Broth without Nitrogen (FORMEDIM, PMD1302), 5 g/L L-glutamic acid, and 0.1 mg/mL of each required amino acid. Edinburgh Minimal Medium (EMM) contains 12.3 g/L EMM Broth without Dextrose (FORMEIUM, PMD0402), 2% glucose and 0.1 mg/mL of each required amino acid. 15 μM thiamine was added to EMM or PMG medium to suppress transcription from the nmt1 promoter on pREP expression plasmids. To induce the expression of the plasmid, cells were cultured in medium without thiamine for 18 h.

Plasmid construction

To generate pREP41-NLS-RNase H1(D145N)-3×FLAG construct, RNase H1(D145N)-3×FLAG fragment was amplified by PCR using pUC18-RNase H1(D145N)-3×FLAG as template. NLS sequence was amplified by PCR using the plasmid containing NLS as template. pREP41 was used as a vector for expression in *S. pombe*. The vector was digested with NdeI and BamHI. The two fragments were cloned into the NdeI-BamHI site of pREP41 using the In-Fusion® HD cloning kit (Takara Bio, 639648). To generate pREP42-NLS-RNase HI-6×PK construct, *E. coli rnhA*$^+$ fragment was amplified by PCR using *E. coli* genomic DNA as a template. NLS sequence was included into the primer to amplify the *rnhA*$^+$ fragment. pREP42-6×PK was digested with NdeI and XhoI. The fragment was cloned into NdeI-XhoI site of pREP42-6×PK using the In-Fusion® HD cloning kit.

Cell cycle synchronization and analysis by flow cytometry

The strains containing *nda3-KM311* mutation were arrested at 20 ºC for 6 h to synchronize cell cycle at M-phase and then released into cell cycle at 30 ºC. To induce RNASE H1(D145N) expression, the cells were grown in PMG medium containing 15 μM thiamine and then transferred to PMG medium without thiamine for 12 h before cell cycle synchronization. Cells in 5 mL culture were collected and resuspended in 200 μL water and were fixed with 600 μL ethanol. The cells were washed with 50 mM sodium citrate (pH7.5) and were treated in 300 μL of 50 mM sodium citrate containing 0.1 mg/mL RNase A at 37 ºC for 2 h. The cells were stained with 4 μg/mL propidium iodide (PI) for 1 h at room temperature. After sonication, the fluorescence intensity of

418 the intercalated PI was measured by BD LSR Fortessa™ X-20 to analyze the DNA content in each
419 cell population.
420
421 **Protein extraction from *S. pombe* and immunoblotting**
422 0.5×10^8 cells were harvested and washed with water. The cells were resuspended in 90 μL water and
423 boiled at 100 ºC for 5 min. 100 μL of 2×SDS Sample Buffer (Tris-HCl [pH6.8], 4% sodium dodecyl
424 sulfate [SDS], 8 M urea, 125 mM 20% glycerol, 1.43 M β-mercaptoethanol, 0.2% bromophenol blue)
425 was added to the cell suspension and the cells were crushed with glass beads by Multi-beads Shocker®
426 (Yasui Kikai Co). The lysates were boiled at 96 ºC for 5 min and the debris were removed by
427 centrifugation at max-speed. The extracted protein samples and ExcelBand 3-color High Range
428 Protein Marker (SMOBIO TECHNOLOGY, PM2600) were loaded and run on 4~20% gradient precast
429 gel (WSHTBIO, GSH2001-420F) and transferred to PVDF membranes (Millipore, IPVH00010). The
430 membranes were blocked with 5% skim milk in TBS-T. The target proteins were detected with ANTI-
431 FLAG® M2-Peroxidase (HRP) antibody (Sigma-Aldrich, A8592). To verify the amounts of loaded
432 proteins, the transferred proteins on the membrane were stained with Ponceau BS.
433
434 **Chromatin immunoprecipitation (ChIP)**
435 1.0×10^9 cells in the culture were cross-linked with 1% formaldehyde for 15 min at 30 ºC and prepared
436 for ChIP as previously described (Kanoh et al., 2015). Briefly, the cross-linked cells were suspended
437 in lysis buffer (50 mM Hepes-KOH [pH 7.5], 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1%
438 sodium deoxycholate, 1 mM PMSF, 1×Protease Inhibitor Cocktail [Sigma-Aldrich, P8215],
439 1×cOmplete™ Protease Inhibitor Cocktail [Roche Diagnostics, 11697498001], and 0.1 mM MG-132
440 [PEPTIDE INSTITUTE INC, 3175-v]), and the cell lysates were prepared by extracting DNA-
441 protein complexes using Multi-beads shocker® (Yasui Kikai Co.) and shearing genomic DNA using
442 S220 Focused-ultrasonicator (Covaris). The lysates were incubated with ANTI-FLAG® M2 antibody
443 (Sigma-Aldrich, F3165) or anti-V5-Tag antibody (BIO-RAD, MCA1360) conjugated to Protein G
444 Dynabeads (ThermoFisher, 10003D) for 6 h at 4 ºC. The beads were washed several times and the
445 precipitated materials were eluted with Elution buffer (50 mM Tris-HCl [pH 8.0], 10 mM EDTA, and
446 1% SDS) for 15 min at 68 ºC. DNA-protein complexes in the eluates were reverse cross-linked
447 overnight at 68 ºC and then treated with RNase A and Proteinase K. DNA was precipitated with ethanol
448 and further purified using QIAquick PCR purification kit (QIAGEN, 28104).
449
450 **ChIP-qPCR**
451 Quantitative PCR was performed using TB Green® Premix Ex Taq™ II (TAKARA Bio, RR820) and
452 a Lightcycler® 480 System II (Roche Diagnostics). The immunoprecipitation efficiency was estimated
453 by calculation as a percentage of input chromatin. The primer sets in this study used are listed in **Table**

454 **2**.

455

456 **Fluorescence microscope observation**

457 RNA-DNA hybrids were visualized by expressing RNASE H1(D145N)-EGFP. DNA damage was
458 detected by observation of fluorescent Rad52-mCherry foci. The cells were stained with 1 μg/mL
459 Hoechst® 33342 to visualize nuclear DNA. RNASE H1(D145N)-EGFP, Rad52-mCherry and
460 Hoechst® 33342 were observed on BZ-X700 (KEYENCE) equipped with Nicon PlanApoλ 100× (NA-
461 1.45) using IMMERSION OIL TYPE NF (Nicon).

462

463 **Next-generation sequencing (NGS) and ChIP-seq**

464 Next Generation Sequencing was performed as previously described (Kanoh et al., 2015). The input
465 and immunoprecipitated DNA were fragmented to an average size of ~ 150 bp using S220 Focused-
466 ultrasonicator (Covaris). The fragmented DNAs were end-repaired, ligated to sequence adapters, and
467 amplified using NEBNext® Ultra™ II DNA Library Prep Kit for Illumina® (New England Biolabs,
468 E7645) and NEBNext® Multiplex Oligos for Illumina® (New England Biolabs, E6440) according to
469 the manual. The amplified DNA (approximately 275 bp in size) was sequenced on Illumina NextSeq
470 to generate 50 bp pair ends reads.

471 The obtained reads from ChIP and input sample sequencing were aligned to the *S. pombe* genomic
472 reference sequence (ASM294v2) provided by PomBase using bowtie2-2.4.4 with the default settings.
473 The sam files generated by bowtie2 were converted to bam files, then sorted, and indexed using
474 samtools. To quantify and visualize ChIP peaks, converted bam files were further converted to bigwig
475 file using deeptools. Peaks were called using findPeaks program included in Homer (Heinz et al.,
476 2010) through comparing immunoprecipitated sample data with that of input. In Homer peak detection
477 process, the filtering Peaks options "-F 2.0" and "-L 2.0" were applied to detect peaks with low
478 intensity. Peak-called genomic loci were annotated by annotatePeaks function in Homor to aggregate
479 the numbers of functionally assigned loci. To convert gene symbol to gene name,
480 "Schizosaccharomyces_pombe.ASM294v2.58.gtf" published in PomBase was used. The peak
481 intensities in each data were semi-quantitatively compared thorough calculating Z-score and
482 visualized using R. For visualizing peaks, Integrative Genomics Viewer (IGV) was used. Gene
483 distance between the RNA/DNA hybrid and these of Rad52 peaks were calculated using
484 computeMatrix function in deepTool (Ramírez et al., 2016) with the positional data of RNA-DNA
485 hybrid loci in BED file and Rad52 peak reads in bigwig files. The results were visualized as the
486 heatmap with the summary plot showing the density of peaks on each distance.

487

488 **Acknowledgements**

489 We are grateful to Dr.Yota Murakami, Dr. Yasukazu Daigaku and National Bio-Resource Project

(NBRP) for providing the mutant strains. This work was funded in part by Grant-in-Aid for Scientific Research (A) (20H00463) to HM.

Author Contributions

T.S., H.S., and Y.K. designed and performed experiments. D.S. analyzed all the NGS data. K.T. assisted the analysis of NGS data. T.S., Y.K, D.S, and M.S designed the study and wrote the manuscript.

References

Ahmad, F., Kaplan, C. D., & Stewart, E. (2002). Helicase activity is only partially required for Schizosaccharomyces pombe Rqh1p function. *Yeast, 19*(16), 1381-1398. doi:10.1002/yea.917

Boguslawski, S. J., Smith, D. E., Michalak, M. A., Mickelson, K. E., Yehle, C. O., Patterson, W. L., & Carrico, R. J. (1986). Characterization of monoclonal antibody to DNA.RNA and its application to immunodetection of hybrids. *J Immunol Methods, 89*(1), 123-130. doi:10.1016/0022-1759(86)90040-2

Castillo-Guzman, D., & Chédin, F. (2021). Defining R-loop classes and their contributions to genome instability. *DNA Repair (Amst), 106*, 103182. doi:10.1016/j.dnarep.2021.103182

Cerritelli, S. M., & Crouch, R. J. (2009). Ribonuclease H: the enzymes in eukaryotes. *Febs j, 276*(6), 1494-1505. doi:10.1111/j.1742-4658.2009.06908.x

Cerritelli, S. M., Frolova, E. G., Feng, C., Grinberg, A., Love, P. E., & Crouch, R. J. (2003). Failure to produce mitochondrial DNA results in embryonic lethality in Rnaseh1 null mice. *Mol Cell, 11*(3), 807-815. doi:10.1016/s1097-2765(03)00088-1

Chan, Y. A., Aristizabal, M. J., Lu, P. Y., Luo, Z., Hamza, A., Kobor, M. S., . . . Hieter, P. (2014). Genome-wide profiling of yeast DNA:RNA hybrid prone sites with DRIP-chip. *PLoS Genet, 10*(4), e1004288. doi:10.1371/journal.pgen.1004288

Chen, L., Chen, J. Y., Zhang, X., Gu, Y., Xiao, R., Shao, C., . . . Fu, X. D. (2017). R-ChIP Using Inactive RNase H Reveals Dynamic Coupling of R-loops with Transcriptional Pausing at Gene Promoters. *Mol Cell, 68*(4), 745-757.e745. doi:10.1016/j.molcel.2017.10.008

Chib, S., Byrd, A. K., & Raney, K. D. (2016). Yeast Helicase Pif1 Unwinds RNA:DNA Hybrids with Higher Processivity than DNA:DNA Duplexes. *J Biol Chem, 291*(11), 5889-5901. doi:10.1074/jbc.M115.688648

Crossley, M. P., Brickner, J. R., Song, C., Zar, S. M. T., Maw, S. S., Chédin, F., . . . Cimprich, K. A. (2021). Catalytically inactive, purified RNase H1: A specific and sensitive probe for RNA-DNA hybrid imaging. *J Cell Biol, 220*(9). doi:10.1083/jcb.202101092

Crow, Y. J., Leitch, A., Hayward, B. E., Garner, A., Parmar, R., Griffith, E., . . . Jackson, A. P. (2006). Mutations in genes encoding ribonuclease H2 subunits cause Aicardi-Goutières syndrome and mimic congenital viral brain infection. *Nat Genet, 38*(8), 910-916. doi:10.1038/ng1842

El Hage, A., Webb, S., Kerr, A., & Tollervey, D. (2014). Genome-wide distribution of RNA-DNA hybrids identifies RNase H targets in tRNA genes, retrotransposons and mitochondria. *PLoS Genet, 10*(10), e1004716. doi:10.1371/journal.pgen.1004716

Ginno, P. A., Lott, P. L., Christensen, H. C., Korf, I., & Chédin, F. (2012). R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. *Mol Cell, 45*(6), 814-825. doi:10.1016/j.molcel.2012.01.017

Hartono, S. R., Malapert, A., Legros, P., Bernard, P., Chédin, F., & Vanoosthuyse, V. (2018). The Affinity

of the S9.6 Antibody for Double-Stranded RNAs Impacts the Accurate Mapping of R-Loops in Fission Yeast. *J Mol Biol, 430*(3), 272-284. doi:10.1016/j.jmb.2017.12.016

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., . . . Glass, C. K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell, 38*(4), 576-589. doi:10.1016/j.molcel.2010.05.004

Helmrich, A., Ballarino, M., & Tora, L. (2011). Collisions between replication and transcription complexes cause common fragile site instability at the longest human genes. *Mol Cell, 44*(6), 966-977. doi:10.1016/j.molcel.2011.10.013

Hiller, B., Achleitner, M., Glage, S., Naumann, R., Behrendt, R., & Roers, A. (2012). Mammalian RNase H2 removes ribonucleotides from DNA to maintain genome integrity. *J Exp Med, 209*(8), 1419-1426. doi:10.1084/jem.20120876

Jeong, H. S., Backlund, P. S., Chen, H. C., Karavanov, A. A., & Crouch, R. J. (2004). RNase H2 of Saccharomyces cerevisiae is a complex of three proteins. *Nucleic Acids Res, 32*(2), 407-414. doi:10.1093/nar/gkh209

Kanoh, Y., Matsumoto, S., Fukatsu, R., Kakusho, N., Kono, N., Renard-Guillet, C., . . . Masai, H. (2015). Rif1 binds to G quadruplexes and suppresses replication over long distances. *Nat Struct Mol Biol, 22*(11), 889-897. doi:10.1038/nsmb.3102

Kim, H. D., Choe, J., & Seo, Y. S. (1999). The sen1(+) gene of Schizosaccharomyces pombe, a homologue of budding yeast SEN1, encodes an RNA and DNA helicase. *Biochemistry, 38*(44), 14697-14710. doi:10.1021/bi991470c

Kotsantis, P., Segura-Bayona, S., Margalef, P., Marzec, P., Ruis, P., Hewitt, G., . . . Boulton, S. J. (2020). RTEL1 Regulates G4/R-Loops to Avert Replication-Transcription Collisions. *Cell reports, 33*(12), 108546. doi:10.1016/j.celrep.2020.108546

Lima, W. F., Murray, H. M., Damle, S. S., Hart, C. E., Hung, G., De Hoyos, C. L., . . . Crooke, S. T. (2016). Viable RNaseH1 knockout mice show RNaseH1 is essential for R loop processing, mitochondrial and liver function. *Nucleic Acids Res, 44*(11), 5299-5312. doi:10.1093/nar/gkw350

Lockhart, A., Pires, V. B., Bento, F., Kellner, V., Luke-Glaser, S., Yakoub, G., . . . Luke, B. (2019). RNase H1 and H2 Are Differentially Regulated to Process RNA-DNA Hybrids. *Cell reports, 29*(9), 2890-2900.e2895. doi:10.1016/j.celrep.2019.10.108

Milano, L., Gautam, A., & Caldecott, K. W. (2024). DNA damage and transcription stress. *Mol Cell, 84*(1), 70-79. doi:10.1016/j.molcel.2023.11.014

Mischo, H. E., Gómez-González, B., Grzechnik, P., Rondón, A. G., Wei, W., Steinmetz, L., . . . Proudfoot, N. J. (2011). Yeast Sen1 helicase protects the genome from transcription-associated instability. *Mol Cell, 41*(1), 21-32. doi:10.1016/j.molcel.2010.12.007

Nowotny, M., Cerritelli, S. M., Ghirlando, R., Gaidamakov, S. A., Crouch, R. J., & Yang, W. (2008). Specific recognition of RNA/DNA hybrid and enhancement of human RNase H1 activity by HBD.

Embo j, 27(7), 1172-1181. doi:10.1038/emboj.2008.44

Nowotny, M., Gaidamakov, S. A., Crouch, R. J., & Yang, W. (2005). Crystal structures of RNase H bound to an RNA/DNA hybrid: substrate specificity and metal-dependent catalysis. *Cell, 121*(7), 1005-1016. doi:10.1016/j.cell.2005.04.024

Nowotny, M., Gaidamakov, S. A., Ghirlando, R., Cerritelli, S. M., Crouch, R. J., & Yang, W. (2007). Structure of human RNase H1 complexed with an RNA/DNA hybrid: insight into HIV reverse transcription. *Mol Cell, 28*(2), 264-276. doi:10.1016/j.molcel.2007.08.015

Osmundson, J. S., Kumar, J., Yeung, R., & Smith, D. J. (2017). Pif1-family helicases cooperatively suppress widespread replication-fork arrest at tRNA genes. *Nat Struct Mol Biol, 24*(2), 162-170. doi:10.1038/nsmb.3342

Ramírez, F., Ryan, D. P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., . . . Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res, 44*(W1), W160-165. doi:10.1093/nar/gkw257

Reijns, M. A., Rabe, B., Rigby, R. E., Mill, P., Astell, K. R., Lettice, L. A., . . . Jackson, A. P. (2012). Enzymatic removal of ribonucleotides from DNA is essential for mammalian genome integrity and development. *Cell, 149*(5), 1008-1022. doi:10.1016/j.cell.2012.04.011

Rhind, N., Chen, Z., Yassour, M., Thompson, D. A., Haas, B. J., Habib, N., . . . Nusbaum, C. (2011). Comparative functional genomics of the fission yeasts. *Science, 332*(6032), 930-936. doi:10.1126/science.1203357

Rondón, A. G., & Aguilera, A. (2019). What causes an RNA-DNA hybrid to compromise genome integrity? *DNA Repair (Amst), 81*, 102660. doi:10.1016/j.dnarep.2019.102660

Scully, R., Panday, A., Elango, R., & Willis, N. A. (2019). DNA double-strand break repair-pathway choice in somatic mammalian cells. *Nat Rev Mol Cell Biol, 20*(11), 698-714. doi:10.1038/s41580-019-0152-0

Skourti-Stathaki, K., Proudfoot, N. J., & Gromak, N. (2011). Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. *Mol Cell, 42*(6), 794-805. doi:10.1016/j.molcel.2011.04.026

Smolka, J. A., Sanz, L. A., Hartono, S. R., & Chédin, F. (2021). Recognition of RNA by the S9.6 antibody creates pervasive artifacts when imaging RNA:DNA hybrids. *J Cell Biol, 220*(6). doi:10.1083/jcb.202004079

Sollier, J., Stork, C. T., García-Rubio, M. L., Paulsen, R. D., Aguilera, A., & Cimprich, K. A. (2014). Transcription-coupled nucleotide excision repair factors promote R-loop-induced genome instability. *Mol Cell, 56*(6), 777-785. doi:10.1016/j.molcel.2014.10.020

Sparks, J. L., Chon, H., Cerritelli, S. M., Kunkel, T. A., Johansson, E., Crouch, R. J., & Burgers, P. M. (2012). RNase H2-initiated ribonucleotide excision repair. *Mol Cell, 47*(6), 980-986. doi:10.1016/j.molcel.2012.06.035

605 Vachez, L., Teste, C., & Vanoosthuyse, V. (2022). DNA:RNA Immunoprecipitation from S. pombe Cells
606 for qPCR and Genome-Wide Sequencing. *Methods Mol Biol, 2528*, 411-428. doi:10.1007/978-1-
607 0716-2477-7_27
608 Wang, K., Wang, H., Li, C., Yin, Z., Xiao, R., Li, Q., . . . Liang, K. (2021). Genomic profiling of native R
609 loops with a DNA-RNA hybrid recognition sensor. *Sci Adv, 7*(8). doi:10.1126/sciadv.abe3516
610 Wood, V., Gwilliam, R., Rajandream, M. A., Lyne, M., Lyne, R., Stewart, A., . . . Nurse, P. (2002). The
611 genome sequence of Schizosaccharomyces pombe. *Nature, 415*(6874), 871-880.
612 doi:10.1038/nature724
613 Wu, H., Lima, W. F., & Crooke, S. T. (2001). Investigating the structure of human RNase H1 by site-directed
614 mutagenesis. *J Biol Chem, 276*(26), 23547-23553. doi:10.1074/jbc.M009676200
615 Yu, K., Chedin, F., Hsieh, C. L., Wilson, T. E., & Lieber, M. R. (2003). R-loops at immunoglobulin class
616 switch regions in the chromosomes of stimulated B cells. *Nat Immunol, 4*(5), 442-451.
617 doi:10.1038/ni919
618 Zhao, H., Zhu, M., Limbo, O., & Russell, P. (2018). RNase H eliminates R-loops that disrupt DNA
619 replication but is nonessential for efficient DSB repair. *EMBO Rep, 19*(5).
620 doi:10.15252/embr.201745335
621
622

623 **Legends to Figures**

624

625 **Figure 1. RNA-DNA hybrids in living cells can be visualized using a fluorescent RNASE**
626 **H1(D145N) probe.**

627 (a) pREP41-NLS-*RNASE H1(D145N)*-3×FLAG-EGFP was transformed into the wild- type, *rnh1*Δ,
628 *rnh201*Δ, and *rnh1*Δ*rnh201*Δ cells to visualize RNA-DNA hybrids. RNASE H1(D145N)-EGFP was
629 expressed by thiamine depletion in PMG medium for 18 h. *RNASE H1(D145N)*-EGFP foci (green)
630 were observed under fluorescence microscope KEYENCE BZ-X700 after staining with Hoechst
631 33342 (blue).

632 (b) The numbers of the nuclear *RNASE H1(D145N)*-EGFP foci from (a) were counted, and the
633 fractions (%) of the cells containing 0, 1, 2, or 3 ≦ foci were quantified.

634 (c) pREP42 vector (upper panel) or pREP42-NLS-*rnhA*-6×PK expressing the *E. coli rnhA* gene (lower
635 panel) was introduced into *rnh1*Δ*rnh201*Δ cells harboring pREP41-NLS-*RNASE H1(D145N)*-
636 3×FLAG-EGFP, and cells were grown in PMG medium without thiamine for 18 h. *RNASE*
637 *H1(D145N)*-EGFP foci were observed under fluorescence microscope KEYENCE BZ-X700 after
638 staining with Hoechst 33342.

639 (d) The numbers of the nuclear *RNASE H1(D145N)*-EGFP foci from (c) were counted, and the
640 fractions (%) of the cells containing 0, 1, 2, or 3 ≦ foci are shown.

641

642 **Figure 2. RNA-DNA hybrids are more enriched in the protein coding region in the**
643 ***rnh1*Δ*rnh201*Δ than those in the wild-type cells.**

644 (a) Representative RNA-DNA hybrids peaks from ChIP-seq analysis. Normalized coverage values are
645 shown in vertical axis. The wild-type or *rnh1*Δ*rnh201*Δ cells harboring pREP41-NLS-*RNASE*
646 *H1(D145N)*-3×FLAG or pREP41-NLS-3×FLAG were grown in PMG medium containing 15 μM
647 thiamine, and then transferred to fresh PMG medium without thiamine at 30 ºC for 18 h to induce the
648 expression of RNASE H1(D145N). The binding of RNASE H1(D145N) on the genome was analyzed
649 by ChIP-seq. The ChIP-seq peaks at selected loci [protein coding (rna15, fba1and srp7), tRNA (Arg,
650 Gln and His) and S5-RNA] are shown. II:4282kb region is negative control loci. fba1 and srp7 were
651 previously reported to be the sites of RNA-DNA hybrid by DRIP-seq (Vachez, Teste, & Vanoosthuyse,
652 2022).

653 (b) ChIP-qPCR analysis of RNASE H1(D145N) precipitated RNA-DNA hybrids. The binding of
654 RNASE H1(D145N) were measured by ChIP-qPCR at the sites of the peaks indicated in (a) in the
655 wild-type or *rnh1*Δ*rnh201*Δ cells harboring pREP41-NLS-*RNASE H1(D145N)*-3×FLAG plasmid
656 (mean ± SD (standard deviation), n = 3 biological replicates).

657 (c) ChIP-qPCR analysis of RNASE H1(D145N) precipitant with RNase HI co-expression condition.
658 ChIP-qPCR were performed under the same condition with (b) except co-express RNase HI. The

bindings of RNASE H1(D145N) were measured by ChIP-qPCR at the sites of the peaks indicated in (a) in the wild-type or *rnh1*Δ*rnh201*Δ cells harboring pREP41-NLS-*RNASE H1(D145N)*-3×FLAG and pREP42-NLS-*rnhA*-6×PK plasmids. *rnhA* encodes RNase HI in *E. coli* (mean ± SD, $n = 3$ biological replicates). In (b) and (c), black bars; wild-type; gray bars, *rnh1*Δ*rnh201*Δ. *$p < 0.05$, **$p < 0.01$ (unpaired two-tailed Student's t test).

(d) Distributions of gene type (%) among the detected RNA-DNA hybrids in wild-type and *rnh1*Δ*rnh201*Δ cells. RNA-DNA hybrids peaks assigned to gene type in wild-type and *rnh1*Δ*rnh201*Δ cells

(e) Distribution histogram of the RNA-DNA hybrid peak score showing that the peak intensity of whole genome, the protein coding region and tRNA region in wild-type and *rnh1*Δ*rnh201*Δ cells. X-axis shows Z-score and Y-axis shows the counts. Intensities were compared between the wild-type and *rnh1*Δ*rnh201*Δ cells. Peaks with Z-score in the range of $-2.0 < Z < 2.0$ were plotted. All the data are presented in Tables S1 and S2.

Figure 3. Rad52 peaks overlapping with RNA-DNA hybrids in the protein coding region increase in *rnh1*Δ*rnh201*Δ cells.

(a) DNA damage foci visualized by Rad52-mCherry (magenta) were observed under fluorescence microscope in the wild-type and *rnh1*Δ*rnh201*Δ cells.

(b) The numbers of Rad52 foci were counted in (a), and the fractions of cells containing 0, 1, 2 or 3 ≦ foci were calculated.

(c) The wild-type (*rad52-6×PK*, *nda3-KM311*) or *rnh1*Δ*rnh201*Δ (*rad52-6×PK*, *nda3-KM311*) cells harboring pREP41-NLS-*RNASE H1(D145N)*-3×FLAG was grown in PMG medium containing 15 μM thiamine. RNASE H1(D145N) expression was induced in the fresh PMG medium without thiamine at 30 ºC for 12 h. Cells were then synchronized at M-phase by shift to 20ºC for 6 h in PMG medium, and were released by shift to 30 ºC. Cells were collected at 60 min after release. The bindings of RNASE H1(D145N) and Rad52 on the genome were analyzed by ChIP-seq. Profiles of RNA-DNA hybrids and Rad52 bindings at the selected loci are shown for the wild-type and *rnh1*Δ*rnh201*Δ cells. Normalized coverage values are shown in vertical axis.

(d) Distributions of gene type in "RNA-DNA hybrid unique" (left), "overlap" (center), and "Rad52 unique" (right) peaks in the wild-type and *rnh1*Δ*rnh201*Δ cells.

Figure 4. Rad52 damage foci increase during S-phase in the *rnh1*Δ*rnh201*Δ cells.

(a) Outlines of the cell cycle time-course experiment. The *nda3-KM311* cells, grown in the absence of thiamine for 12 h, were synchronized at 20 ºC for 6 h and then released into cell cycle at 30 ºC, and Rad52 foci were analyzed at the times indicated after release.

(b) The wild-type (*rad52-mCherry, nda3-KM311*) or *rnh1*Δ*rnh201*Δ (*rad52-mCherry, nda3-KM311*)

cells harboring pREP41-NLS-*RNASE H1(D145N)*-3×FLAG-EGFP was grown in PMG medium containing 15 μM thiamine, transferred to PMG medium without thiamine and cultured at 30 ºC for 12 h. Cells were then synchronized at the M-phase by shift to 20 ºC for 6 h in PMG medium, and then were released from M-phase by shift to 30 ºC. The mCherry (magenta) and EGFP (green) signals were observed under fluorescence microscope every 30 min after release.

(c) The numbers of Rad52 foci in (b) were counted, and cells containing 0, 1, 2 or 3 ≦ foci were quantified.

Figure 5. Rad52 peaks overlapping with RNA-DNA hybrids in the protein coding region increase during S-phase.

(a) A Rad52 binding peaks at 0 min (M-phase) or 60 min (S-phase) after release from *nda3-KM311* arrest was assigned to gene types in wild-type and *rnh1*Δ*rnh201*Δ cells.

(b) Metagene plot and heatmap showing the enrichment of Rad52 binding sites at the RNA-DNA hybrids in wild-type and *rnh1*Δ*rnh201*Δ cells. Enrichment is shown for whole genome, protein coding segment and tRNA segment at 0 min (M-phase) or 60 min (S-phase). Heatmap shows whole genome results. TSS, transcription start site; TES, transcription end site.

(c) Distribution histograms of Rad52 peaks overlapping with RNA-DNA hybrids in wild-type and *rnh1*Δ*rnh201*Δ cells. X-axis, Z-score; Y-axis, counts. Peak distributions are color-coded according to the indicated gene types in the peaks with Z-score within $-1.0 < Z < 1.0$. The black dotted line indicates the center of the detected peaks on the protein coding region at 0 min in wild-type cells.

Supplementary Figure S1. Expression of human catalytically inactive RNASE H1 in fission yeast cells.

(a) Schematic drawing of the probe capable of recognizing RNA-DNA hybrids in *S. pombe*. Human ribonuclease H1 functional domain consists of a mitochondrial targeting sequence (MTS: 1-26 a.a), a hybrid binding domain (HBD: 27-70 a.a), and a catalytic domain (CD: 136-286 a.a). The SV40 nuclear localization signal (NLS) and 3×FLAG were fused to catalytically inactive human RNASE H1(D145N) (28-286 a.a) at the N- and C-terminus respectively. The resulting NLS-*RNASE H1(D145N)*-3xFLAG was cloned under the inducible nmt1 (No Message in Thiamine1) promoter on the vector pREP41.

(b) Expression level of NLS-*RNASE H1(D145N)*-3×FLAG in wild-type, *rnh1*Δ, *rnh201*Δ, and *rnh1*Δ*rnh201*Δ cells. The cells were grown in PMG medium with 15 μM thiamine until log phase, then transferred to the PMG medium without thiamine and cultured for 18 h. The expression of NLS-*RNASE H1(D145N)*-3×FLAG was analyzed by western blotting (upper). Ponceau staining serves as a loading control (lower).

Supplementary Figure S2. Effects of various helicase mutations on RNASE H1(D145N) foci.
*rqh1*Δ, *sen1*Δ and *srs2*Δ cells harboring pREP41-NLS-*RNASE H1(D145N)*-3×FLAG-EGFP were grown in PMG medium with 15 μM thiamine, transferred to the PMG medium without thiamine and grown for 18 h. NLS-*RNASE H1(D145N)*-3×FLAG-EGFP signals (green) were observed using fluorescence microscope after staining with Hoechst 33342 (blue).

Supplemental Figure S3. Genome-wide profiles of RNA-DNA hybrids in the wild-type and *rnh1*Δ*rnh201*Δ cells.
(a) The wild-type or *rnh1*Δ*rnh201*Δ cells harboring pREP41-NLS-3×FLAG or pREP41-NLS-*RNASE H1(D145N)*-3×FLAG was cultured in PMG medium containing 15 μM thiamine. Cells were then transferred to PMG medium without thiamine and grown at 30 ºC for 18 h. Cells were harvested and were fixed with 1% formalin, and ChIP-seq was conducted. The wild-type or *rnh1*Δ*rnh201*Δ cells harboring pREP41-NLS-*RNASE H1(D145N)*-3×FLAG or pREP41-NLS-3×FLAG was grown in PMG medium containing 15 μM thiamine, and were transferred to fresh PMG medium without thiamine and were grown at 30 ºC for 18 h. The binding of RNASE H1(D145N) on the genome was analyzed by ChIP-seq. ChIP-seq peaks of RNASE H1 (D145N)-3×FLAG or 3×FLAG as a mock control is presented on the whole genome of *S. pombe*.
(b) Venn diagram showing the overlapping and unique peaks of RNA-DNA hybrids between wild-type and *rnh1*Δ*rnh201*Δ cells.
(c) Distributions of functional annotations of RNA-DNA hybrids peaks in wild-type or *rnh1*Δ*rnh201*Δ cells.
(d) Distributions of functional annotations in the "wild-type unique", "*rnh1*Δ*rnh201*Δ unique" and "common" peaks of RNA-DNA hybrids.
(e) Distributions of gene types in the "wild-type unique", "*rnh1*Δ*rnh201*Δ unique" and "common" peaks of RNA-DNA hybrids.

Supplementary Figure S4. RNA-DNA hybrid peaks detected by RNase H1(D145N) probe: association with gene types.
(a) The peak scores at the RNA-DNA hybrid peaks overlapping with Rad52 peaks in the wild-type and *rnh1*Δ*rnh201*Δ cells.
(b) Distribution histograms of RNA-DNA hybrids among different gene types in wild-type and *rnh1*Δ*rnh201*Δ cells. X-axis, Z-score; Y-axis, counts. Peak distributions are color-coded according to the indicated gene types in the peaks with Z-score within $-2.0 < Z < 2.0$. The peak score of RNA-DNA hybrids on tRNA and rRNA are higher in *rnh1*Δ*rnh201*Δ cells than in the wild-type cells. RNA-DNA hybridscounts in protein coding regions increased within the lower peak score range.

Supplementary Figure S5. DNA damage peaks detected by Rad52.

(a) The distributions of gene types among the detected Rad52 binding sites in wild-type and *rnh1*Δ*rnh201*Δ cells.

(b) Venn diagram showing the overlapping and unique peaks of RNA-DNA hybrids and Rad52 binding in the wild-type and *rnh1*Δ*rnh201*Δ cells.

Supplementary Figure S6. Analysis of cell cycle progression in the wild-type and *rnh1*Δ*rnh201*Δ cells after release from M-phase arrest.

The wild-type (*nda3-KM311*) or *rnh1*Δ*rnh201*Δ (*nda3-KM311*) were arrested at the M-phase by incubation at 20 ºC for 6 h with and without thiamine. The cells were released into the cell cycle at 30 ºC and harvested every 15 min for 2 h. The cell cycle progression was monitored by flow cytometry.

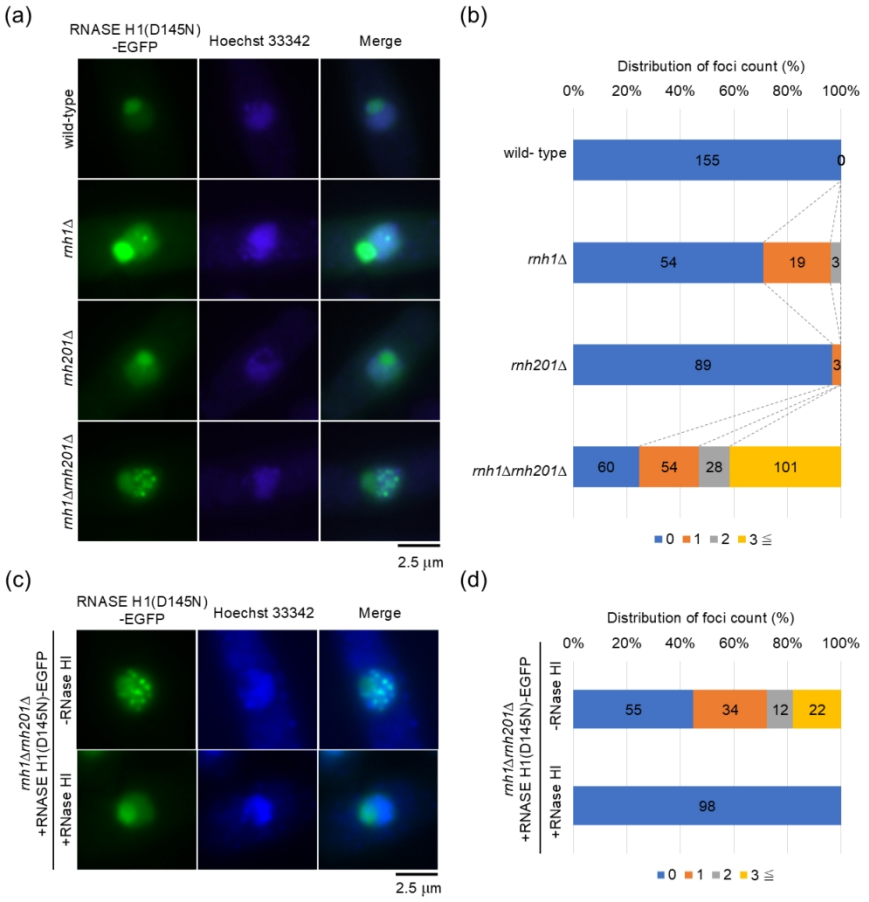
Supplementary Figure S7. Rad52 damage foci increase during S-phase in *rnh1*Δ*rnh201*Δ cells.

The wild-type (*rad52-mCherry*, *nda3-KM311*) or *rnh1*Δ*rnh201*Δ (*rad52-mCherry*, *nda3-KM311*) cells harboring pREP41-NLS-*RNASE H1(D145N)*-3×FLAG-EGFP were grown in PMG medium containing 15 μM thiamine, transferred to PMG medium without thiamine and cultured at 30 ºC for 12 h. Cells were synchronized at the M-phase by shift to 20 ºC for 6 h in PMG medium, and then were released from M-phase by shift to 30 ºC. The mCherry (magenta) and EGFP signals (green) were observed under fluorescence microscope every 30 min after release.

Supplementary Figure S8. Rad52 damage foci and RNA-DNA hybrids: overlapping loci.

(a) The wild-type (*rad52-mCherry*, *nda3-KM311*) or *rnh1*Δ*rnh201*Δ (*rad52-mCherry*, *nda3-KM311*) cells were grown in PMG medium and synchronized at M-phase by incubation for 6 h at 20 ºC. After release from M-phase arrest at 30 ºC, cells were harvested at 0 min or 60 min. ChIP-seq patterns at the selected loci of Rad52 binding closed to the RNA-DNA hybrid are shown at 0 min (M-phase) and 60 min (late S-phase) after release. I:1027kb and II:4282kb serve as negative controls.

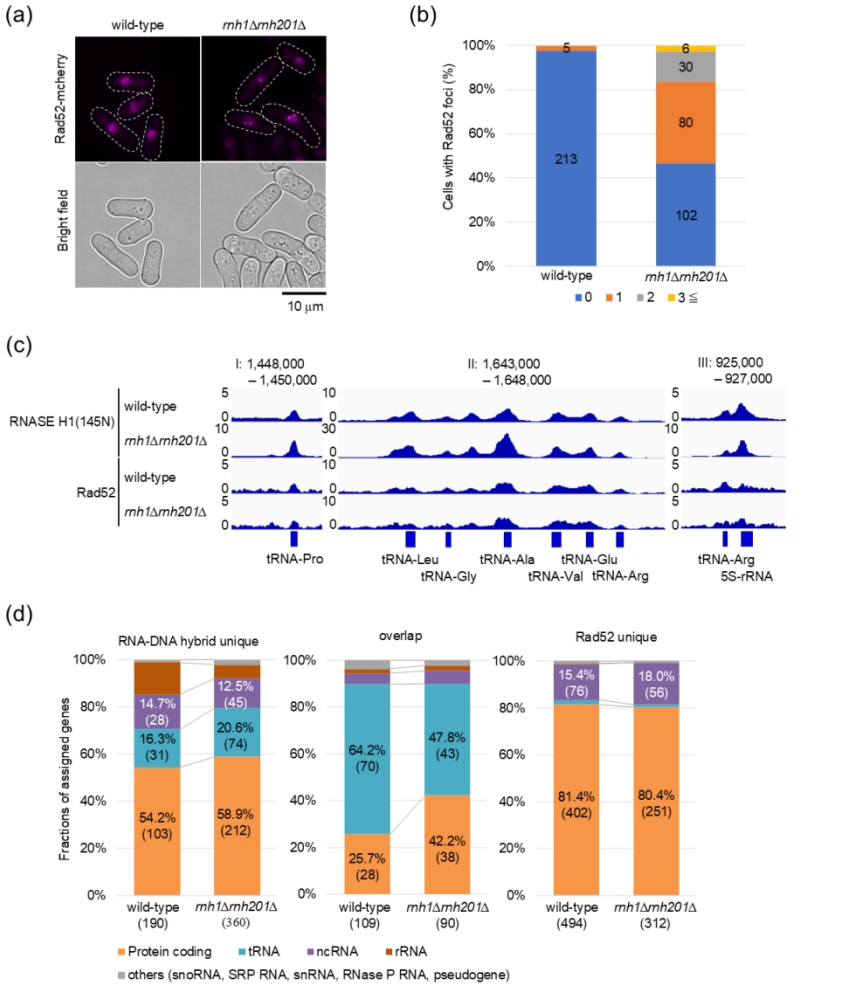
(b) Rad52 binding was validated by ChIP-qPCR at the peaks overlapping with RNA-DNAs (mean ± SD, $n = 3$ biological replicates) indicated in (a). $^{*}p < 0.05$, $^{**}p < 0.01$ (unpaired two-tailed Student's t test).



Sagi *et al.* Figure 1

Figure 1. RNA-DNA hybrids in living cells can be visualized using a fluorescent RNASE H1(D145N) probe.

338x451mm (96 x 96 DPI)



Sagi *et al.* Figure 3

Figure 3. Rad52 peaks overlapping with RNA-DNA hybrids in the protein coding region increase in *rnh1Δrnh201Δ* cells.

338x451mm (96 x 96 DPI)

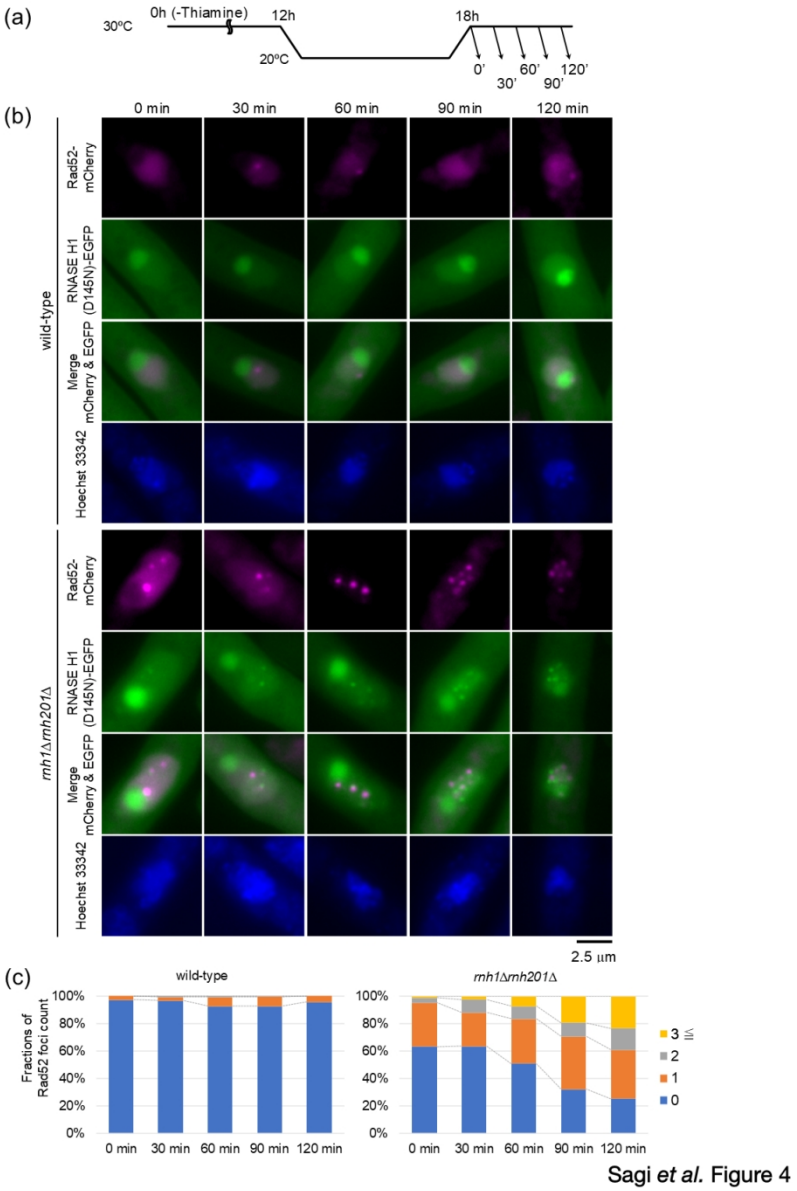
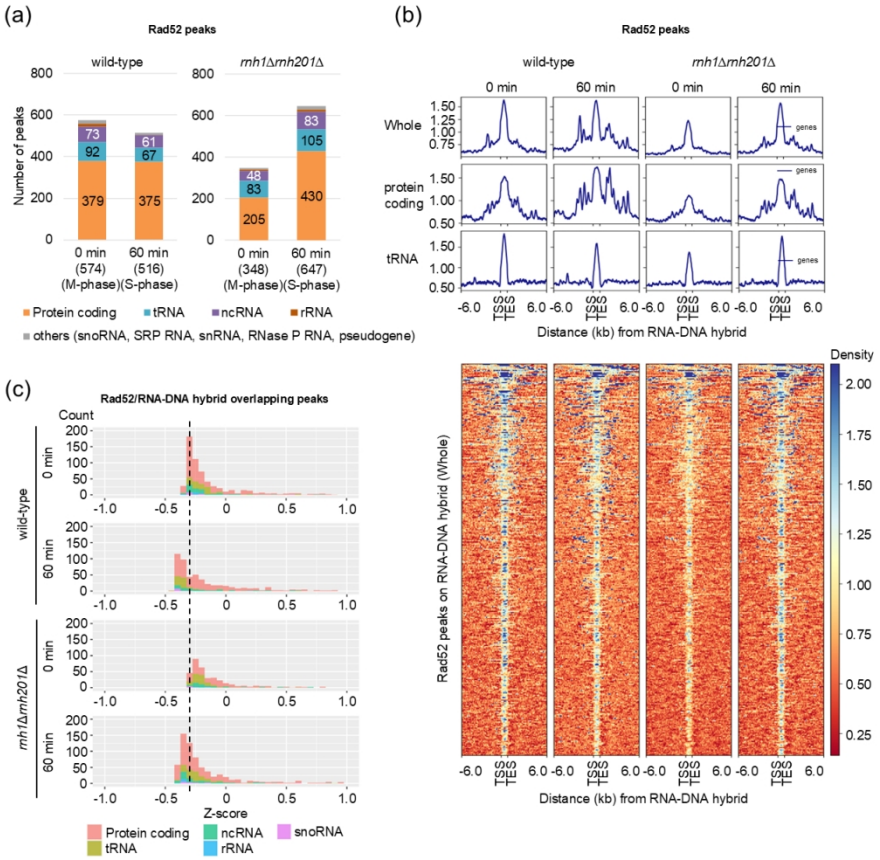


Figure 4. Rad52 damage foci increase during S-phase in the *rnh1Δrnh201Δ* cells.

338x451mm (96 x 96 DPI)



Sagi *et al.* Figure 5

Figure 5. Rad52 peaks overlapping with RNA-DNA hybrids in the protein coding region increase during S-phase.

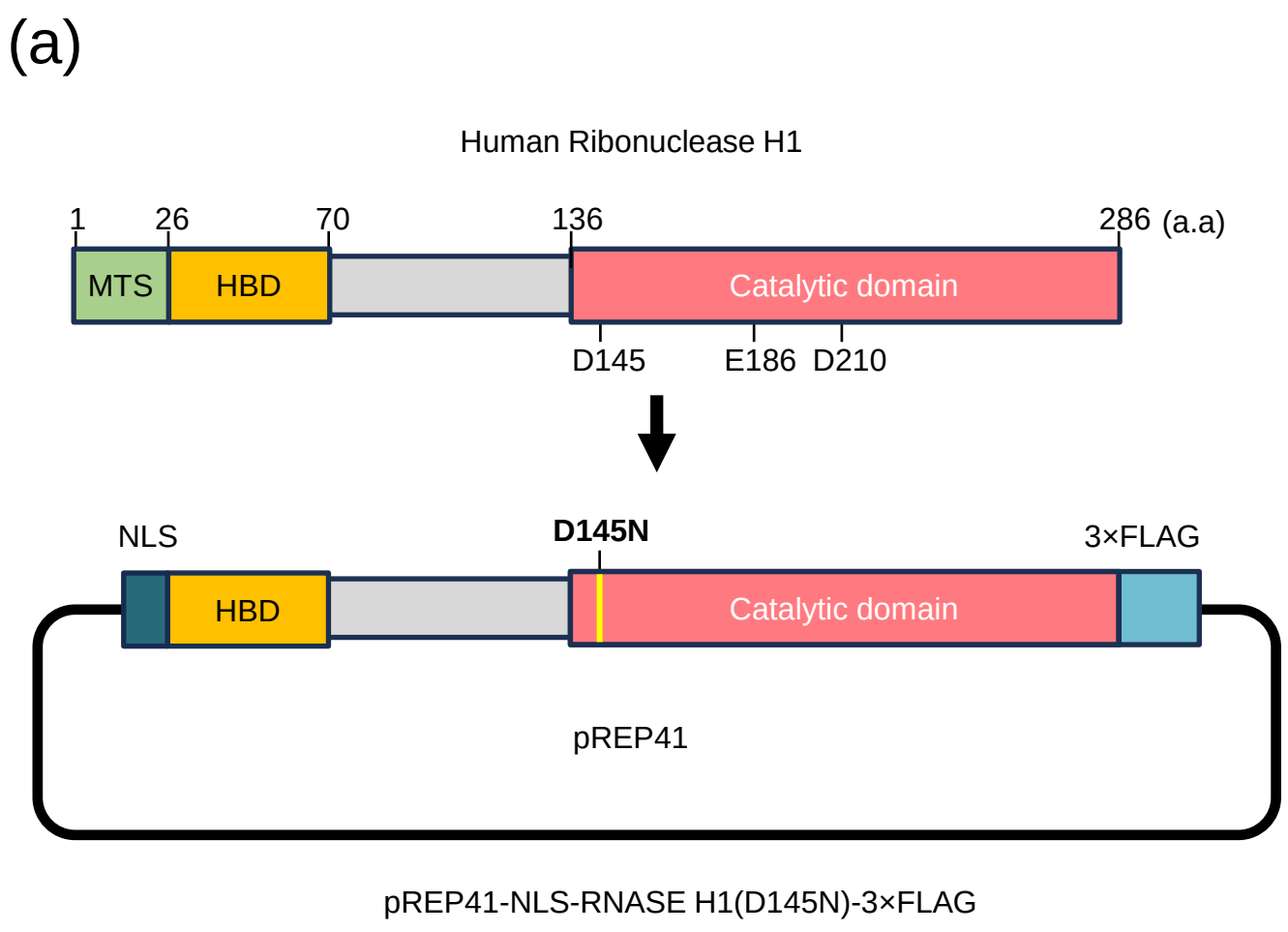
338x451mm (96 x 96 DPI)

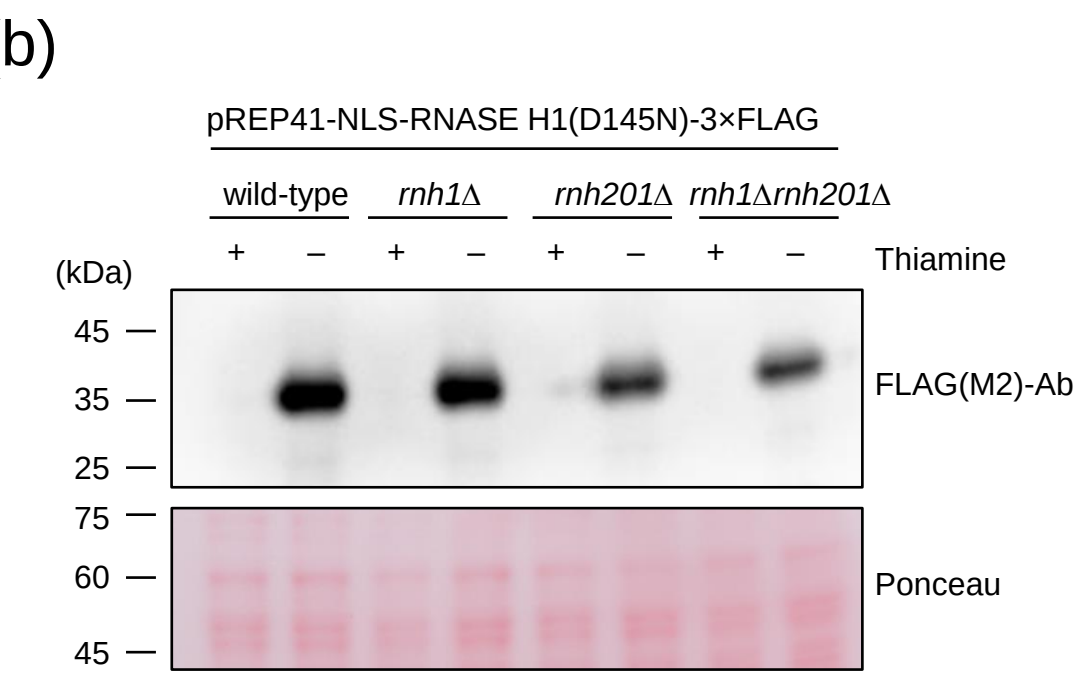
Table 1. Strain list used in this study

Name	Genotype	source
MN-57	*h+ ade6-DN/N leu1-32 ura4-DS/E imrIL::ura4+ OtrIR::ade6+ rnh1ΔhphMX6*	gifted from Dr. Murakami
IM1048	*h– ade6-704 leu1-32 ura4-D18 rnh201::KanMX6*	gifted from Dr. Daigaku
FY11019(MY3096)	*h– leu1 ura4 Δrqh1::ura4+*	gifted from NBRP
FY23524 (SP1347)	*h90 leu1-32 ura4-D18 ade6-M210 sen1D::kan*	gifted from NBRP
SG10	*h– leu1-32 ura4-D18 ura4-DS/E imrIL::ura4+? OtrIR::ade6+? rnh1hphMX6 rnh201::KanMX6*	this study
SG16	*h– leu1-32 ura4-D18 pREP41-NLS-RNaseH1(D145N)*-3×FLAG	this study
SG19	*h– leu1-32 ura4-D18 ura4-DS/E imrIL::ura4+? OtrIR::ade6+? rnh1ΔhphMX6 rnh201::KanMX6 pREP41-NLS-RNaseH1(D145N)-3×FLAG*	this study
SG34	*h– leu1-32 ura4-D18 pREP41-NLS-RNaseH1(D145N)-3×FLAG-EGFP*	this study
SG37	*h– leu1-32 ura4-D18 ura4-DS/E imrIL::ura4+? OtrIR::ade6+? rnh1ΔhphMX6 rnh201::KanMX6 pREP41-NLS-RNaseH1(D145N)-3×FLAG-EGFP*	this study
SG40	*h– ade6-704 leu1-32 ura4-D18 rnh201::KanMX6 pREP41-NLS-RNaseH1(D145N)-3×FLAG-EGFP*	this study
SG46	*h+ ade6-DN/N leu1-32 ura4-DS/E imrIL::ura4+? OtrIR::ade6+? rnh1ΔhphMX6 pREP41-NLS-RNaseH1(D145N)-3×FLAG-EGFP*	this study
SG97	*h– leu1-32 ura4-D18 rad52-mcherry::ura4 nda3-KM311* pREP41-NLS-RNaseH1(D145N)-3×FLAG-EGFP	this study
SG98	*h+ ade6-DN/N leu1-32 ura4-DS/E imrIL::ura4+? OtrIR::ade6+? rnh1ΔhphMX6 nda3-KM311 rad52-mcherry::ura4 pREP41-NLS-RNaseH1(D145N)-3×FLAG-EGFP*	this study
SG99	*h– ade6-704 leu1-32 ura4-D18 rnh201::KanMX6 nda3-KM311 rad52-mcherry::ura4 pREP41-NLS-RNaseH1(D145N)-3×FLAG-EGFP*	this study
SG100	*h– leu1-32 ura4-D18 ura4-DS/E imrIL::ura4+? OtrIR::ade6+? rnh1ΔhphMX6 rnh201::KanMX6 pREP41-NLS-RNaseH1(D145N)-3×FLAG-EGFP*	this study
SG128	*h– leu1 ura4 Δrqh1::ura4+ pREP41-NLS-RNaseH1(D145N)-3×FLAG-EGFP*	this study
SG162	*h– leu1-32 ura4-D18 pREP41-NLS-RNaseH1(D145N)-3×FLAG-EGFP pREP42-NLS-rnhA-6xPK*	this study
SG163	*h– leu1-32 ura4-D18 ura4-DS/E imrIL::ura4+? OtrIR::ade6+? rnh1ΔhphMX6 rnh201::KanMX6 pREP41-NLS-RNaseH1(D145N)-3×FLAG-EGFP pREP42-NLS-rnhA-6xPK*	this study
SG193	*h– leu1 ura4 Δrqh1::ura4+ pREP41-NLS-RNaseH1(D145N)-3×FLAG_EGFP*	this study
SG195	*h90 leu1-32 ura4-D18 ade6-M210 sen1D::kan pREP41-NLS-RNaseH1(D145N)-3×FLAG-EGFP*	this study
SG197	*h– leu1-32 srs2::ura4 pREP41-NLS-RNaseH1(D145N)-3×FLAG-EGFP*	this study
SG203	*h– leu1-32 ura4-D18 pREP41-NLS-RNaseH1(D145N)-3×FLAG pREP42-NLS-rnhA-6xPK*	this study
SG205	*h– leu1-32 ura4-D18 ura4-DS/E imrIL::ura4+? OtrIR::ade6+? rnh1ΔhphMX6 rnh201::KanMX6 pREP41-NLS-RNaseH1(D145N)-3×FLAG pREP42-NLS-rnhA-6xPK*	this study
SG207	*h– leu1-32 ura4-D18 pREP41-NLS-3×FLAG*	this study
SG209	*h– leu1-32 ura4-D18 ura4-DS/E imrIL::ura4+? OtrIR::ade6+? rnh1ΔhphMX6 rnh201::KanMX6 pREP41-NLS-3×FLAG*	this study
SG262	*h– leu1-32 ura4-D18 rad52-mcherry::ura4 nda3-KM311 pREP41-NLS-RNaseH1(D145N)-3×FLAG-EGFP*	this study
SG263	*h+ ade6-DN/N leu1-32 ura4-DS/E imrIL::ura4+? OtrIR::ade6+? rnh1ΔhphMX6 rad52-mcherry::ura4 nda3-KM311 pREP41-NLS-RNaseH1(D145N)-3×FLAG-EGFP*	this study
SG264	*h– ade6-704 leu1-32 ura4-D18 rnh201::KanMX6 rad52-mcherry::ura4 nda3-KM311 pREP41-NLS-RNaseH1(D145N)-3×FLAG-EGFP*	this study
SG265	*h– leu1-32 ura4-D18 ura4-DS/E imrIL::ura4+? OtrIR::ade6+? rnh1ΔhphMX6 rnh201::KanMX6 pREP41-NLS-RNaseH1(D145N)-3×FLAG-EGFP*	this study

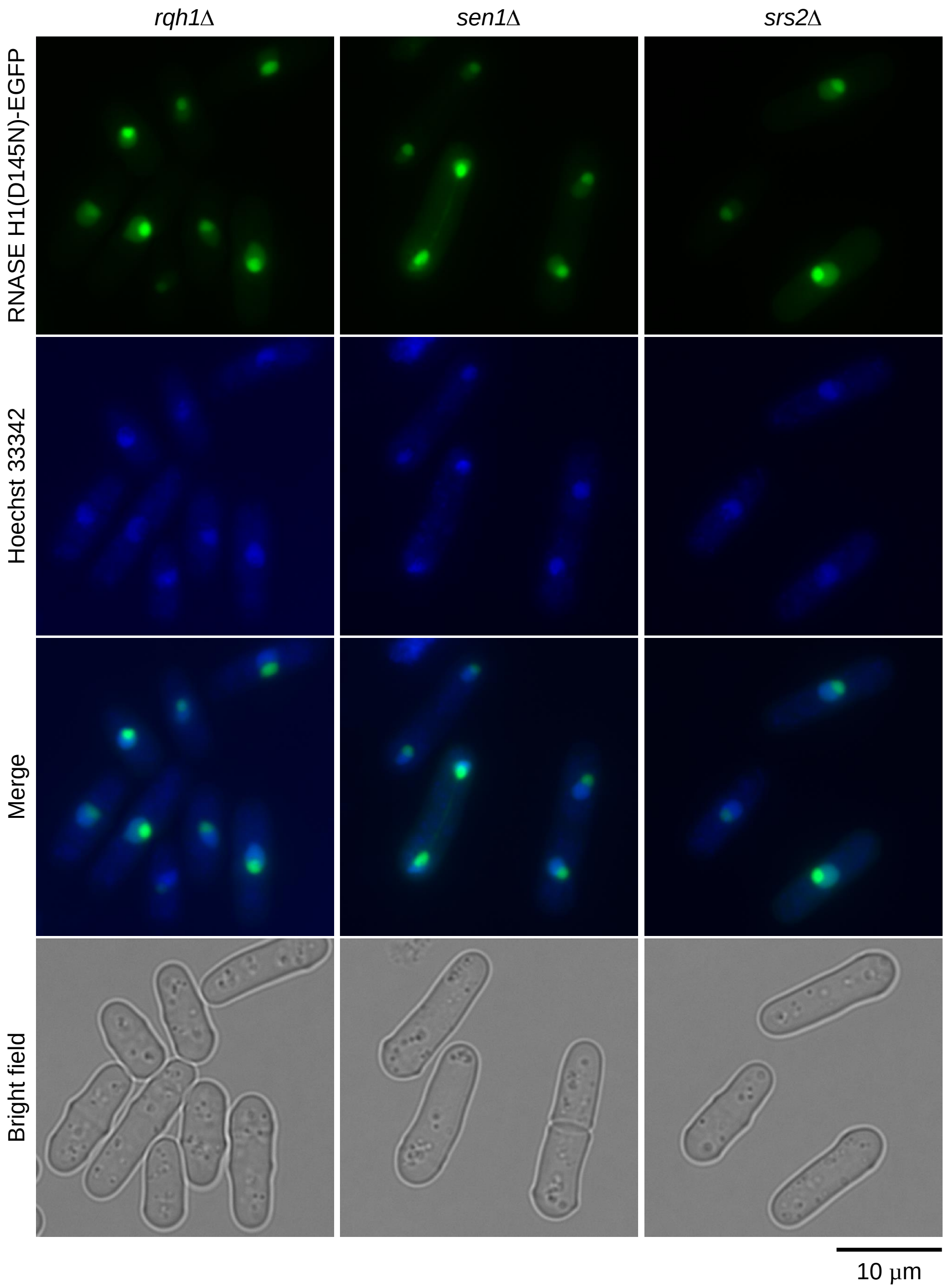
Table 2. Primer sequence for qPCR

Primer name	Sequence	locus
I:2701kb_Fwd	ATGGCTGAGGAACAGGTCATTGACA	I:2,701,712-2,701,739
I:2701kb_Rev	CGAACTGCAGACGCTGCAGTAG	I:2,701,836-2,701,860
I:3443kb_Fwd	TCGTGGCCTAGAAGTAAGGCGT	I:3,443,057-3,443,081
I:3443kb_Rev	GTTGACAGCAAATTTGCTTAGATGGAA	I:3,443,201-3,443,230
II:356kb_Fwd	AGCTGAGAAATCGGTTCTCAACCA	II:356,093-356,119
II:356kb_Rev	TCGTGATAGCGTAACGTTCGTCCA	II:356,185-356,211
II:464kb_Fwd	GTCTACGGCCATACCTAGGCGAA	II:464,525-464,550
II:464kb_Rev	AAGCCTACAGCACCCCGGATTC	II:464,623-464,647
fba1_qL1	TCAAGACCACCAACGACAAG	II:1,689,117-1,689,137
fba1_qR1	AGGCGAATTGGGTATCAGTG	II:1,689,229-1,689,249
srp7_qL2	AAGAGTAGTCTTCGTGGCAACTG	I:4,268,764-4,268,789
srp7 qR2	GATGTGCATTGTTTCCAACC	I:4,268,815-4,268,835
II:4282kb_Fwd	ATGCTTCACATATGGTAGGCTCAGA	II:4,281,887-4,281,914
II:4282kb_Rev	GACCAAGGCAAGAAGTTAGTTTACGA	II:4,281,973-4,282,001
I:3991kb_Fwd	CAGTCGAATGATCGCAGCATAGC	I:3,991,540-3,991,565
I:3991kb_Rev	CTTCATGCGCTTCTCGTTCCA	I:3,991,634-3,991,656
I:4139kb_Fwd	CACTTCCATTATGTGCCTGCCGTT	I:4,139,723-4,139,749
I:4139kb_Rev	AGTGCATCTAGTCTGACAGCGA	I:4,139,819-4,139,843
II:2807kb_Fwd	ACCAACGAAGCAACGACCATCTC	II:2,807,351-2,807,376
II:2807kb_Rev	GACTATCTTTACCACAATTGCCTCGCACT	II:2,807,467-2,807,498
III:278kb_Fwd	CCTTCTCGCCACCGATTTTCCTA	III:277,458-277,483
III:278kb_Rev	CTTGACAATCGATTTTGCATTGGCAT	III:277,599-277,627
I:1027kb_Fwd	GTGCCCATATCAATCGGACATGCT	I:1,027,159-1,027,185
I:1027kb_Rev	CATGGATCTTATACAGTGCATTCACCGT	I:1,027,281-1,027,311

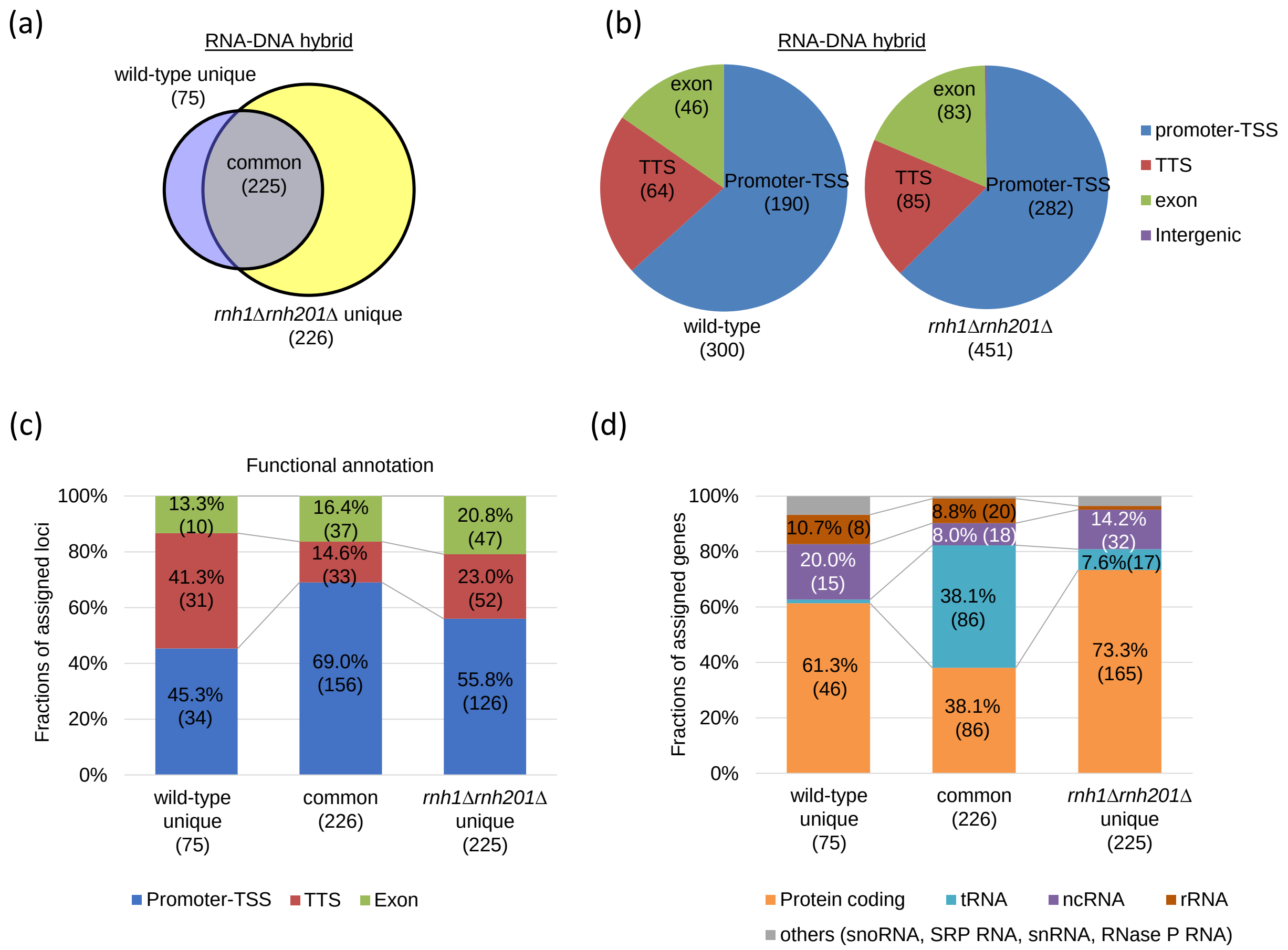




Sagi *et al.* Supplementary Figure S1

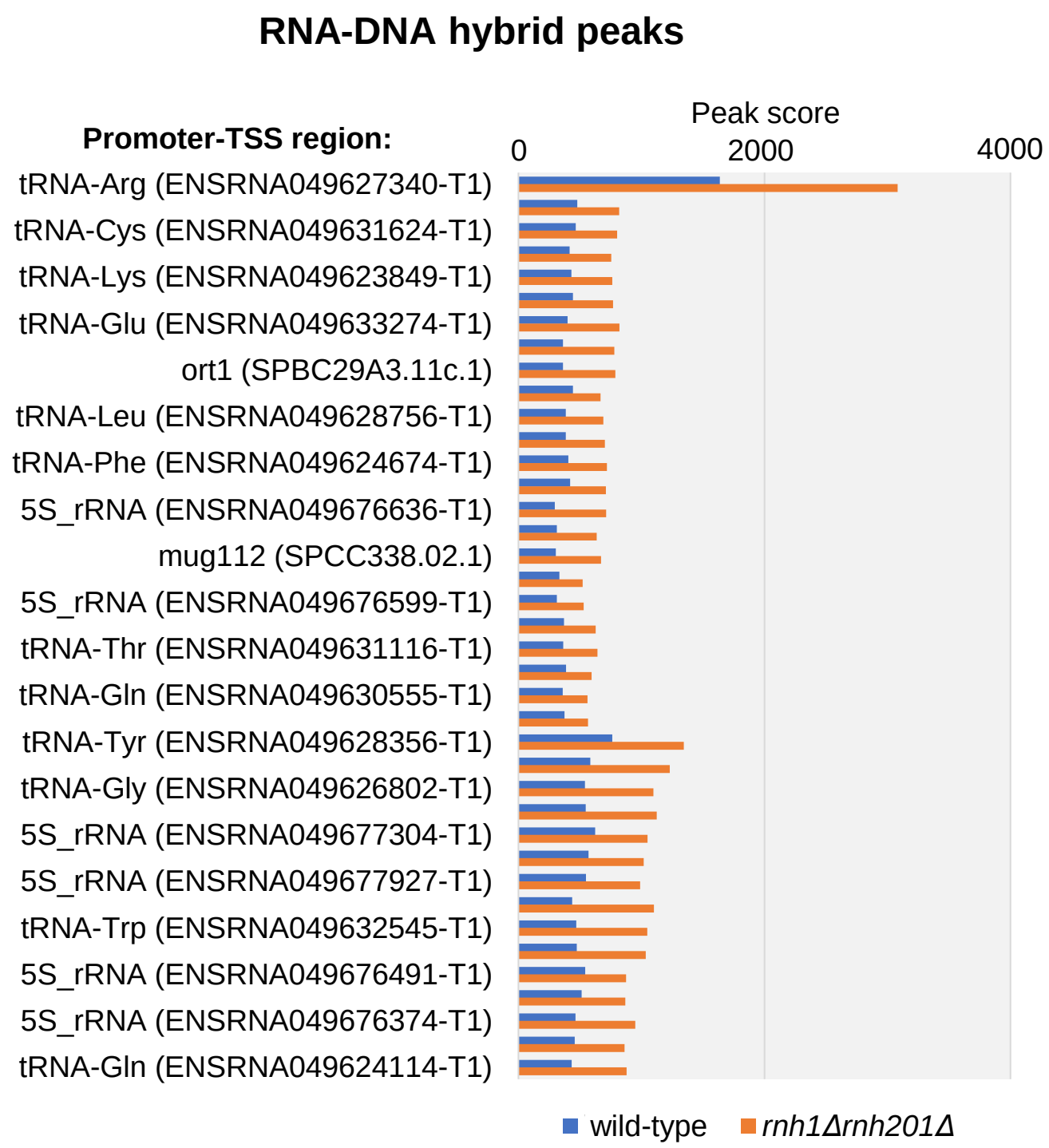


Sagi *et al.* Supplementary Figure S2

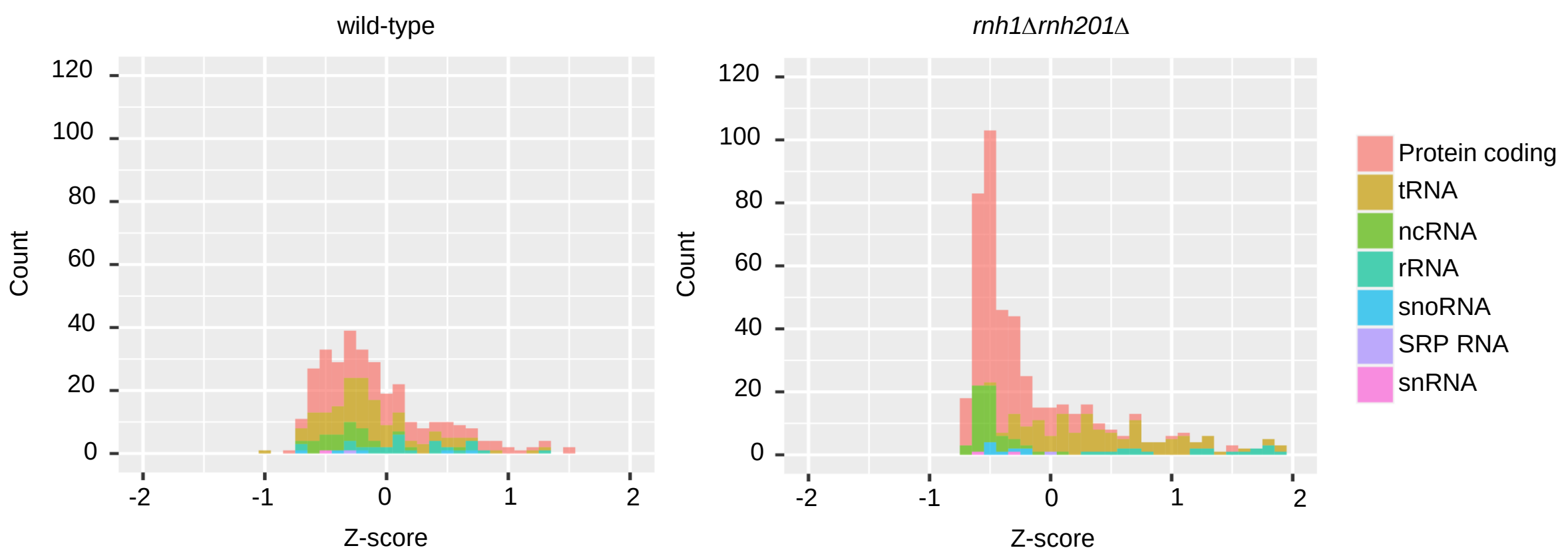


Sagi *et al.* Supplementary Figure S3

(a)

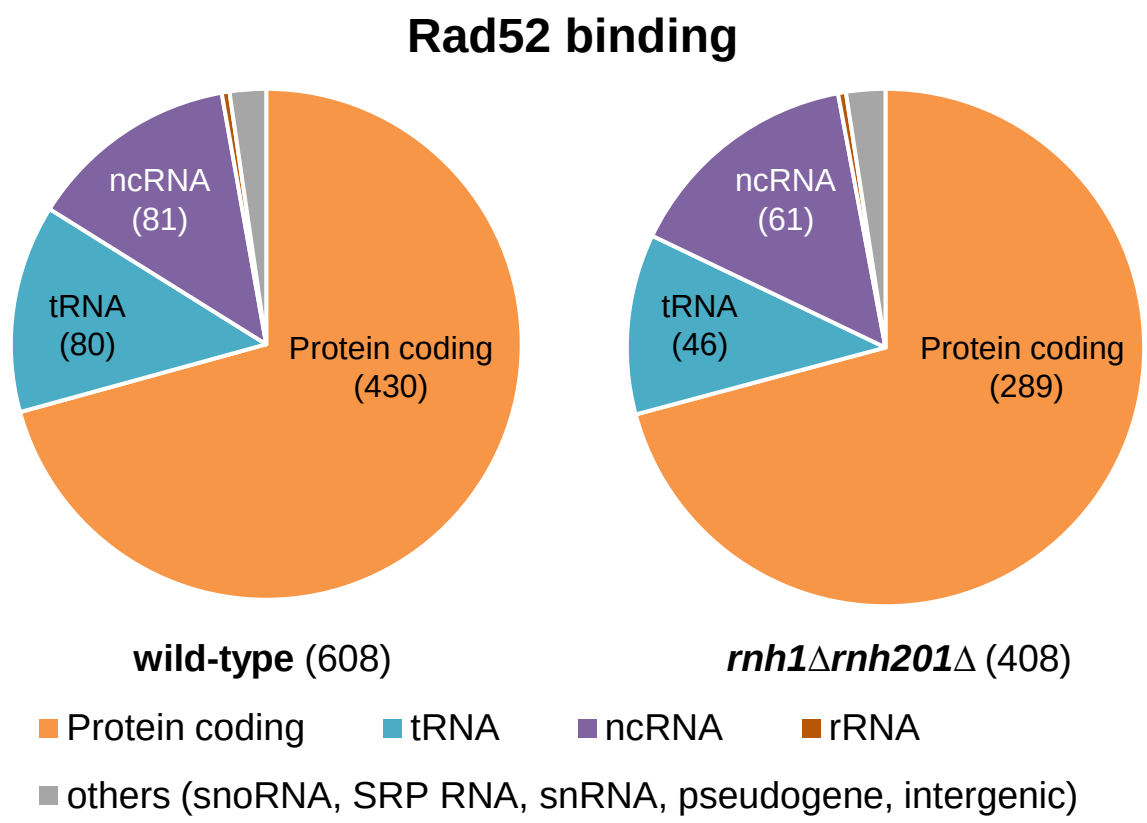


(b)

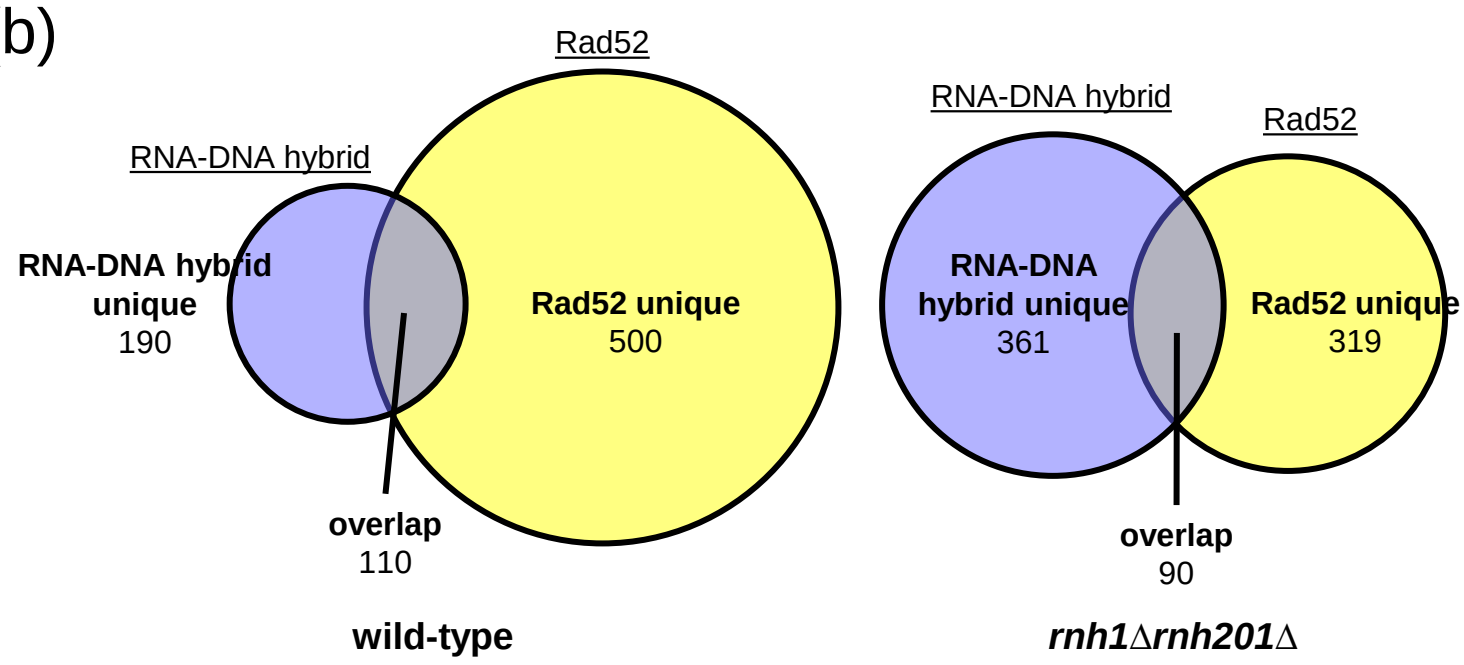


Sagi *et al.* Supplementary Figure S4

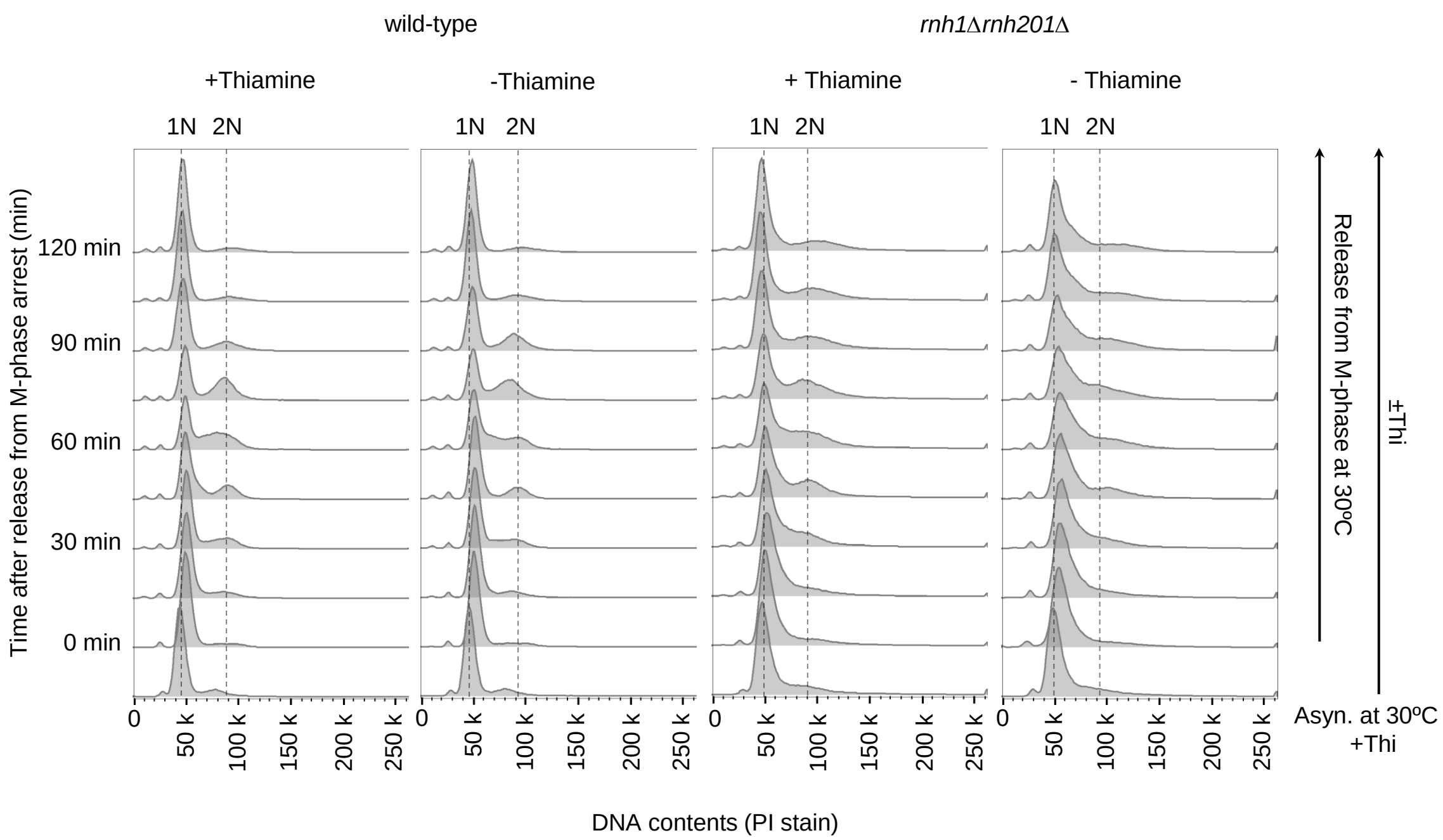
(a)



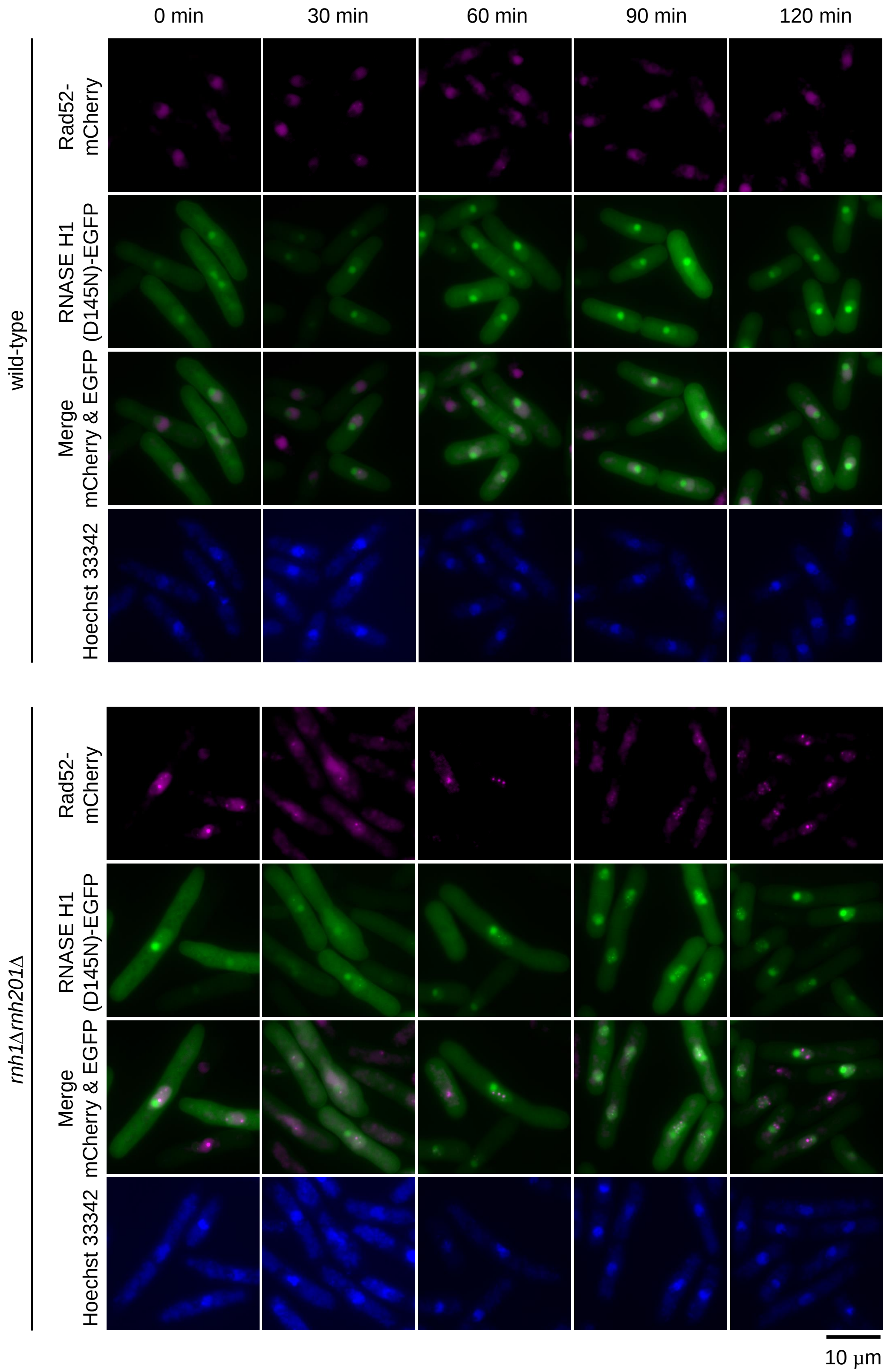
(b)



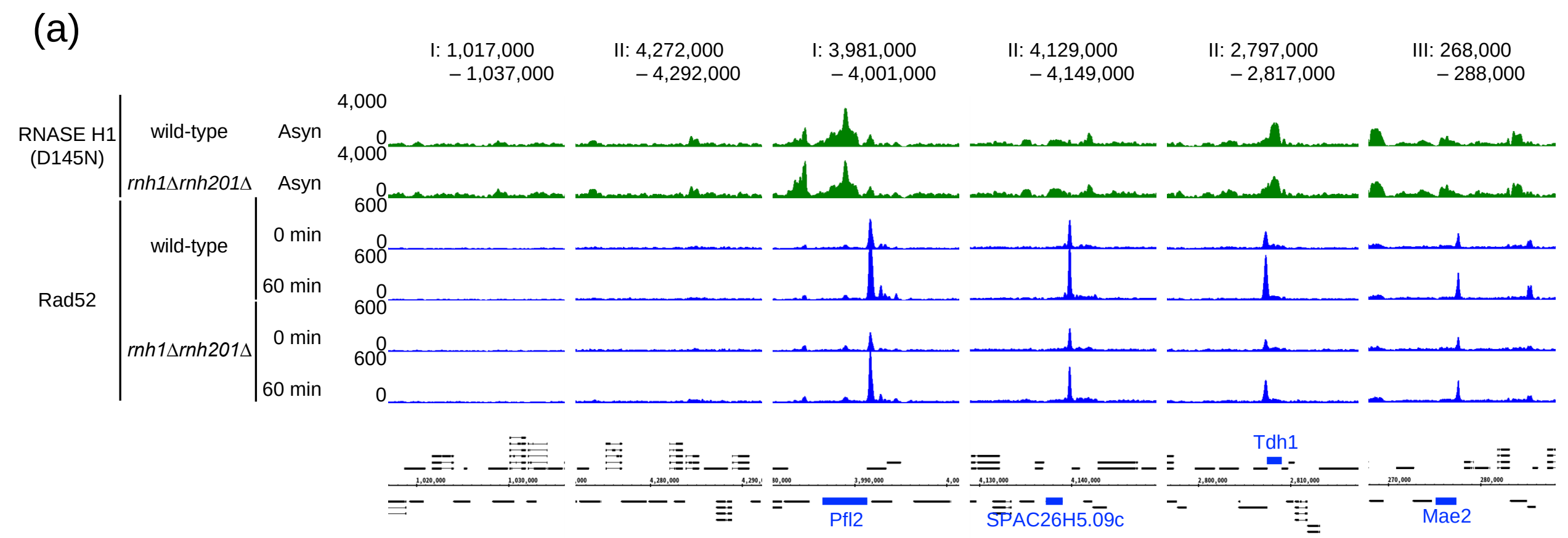
Sagi *et al.* Supplementary Figure S5

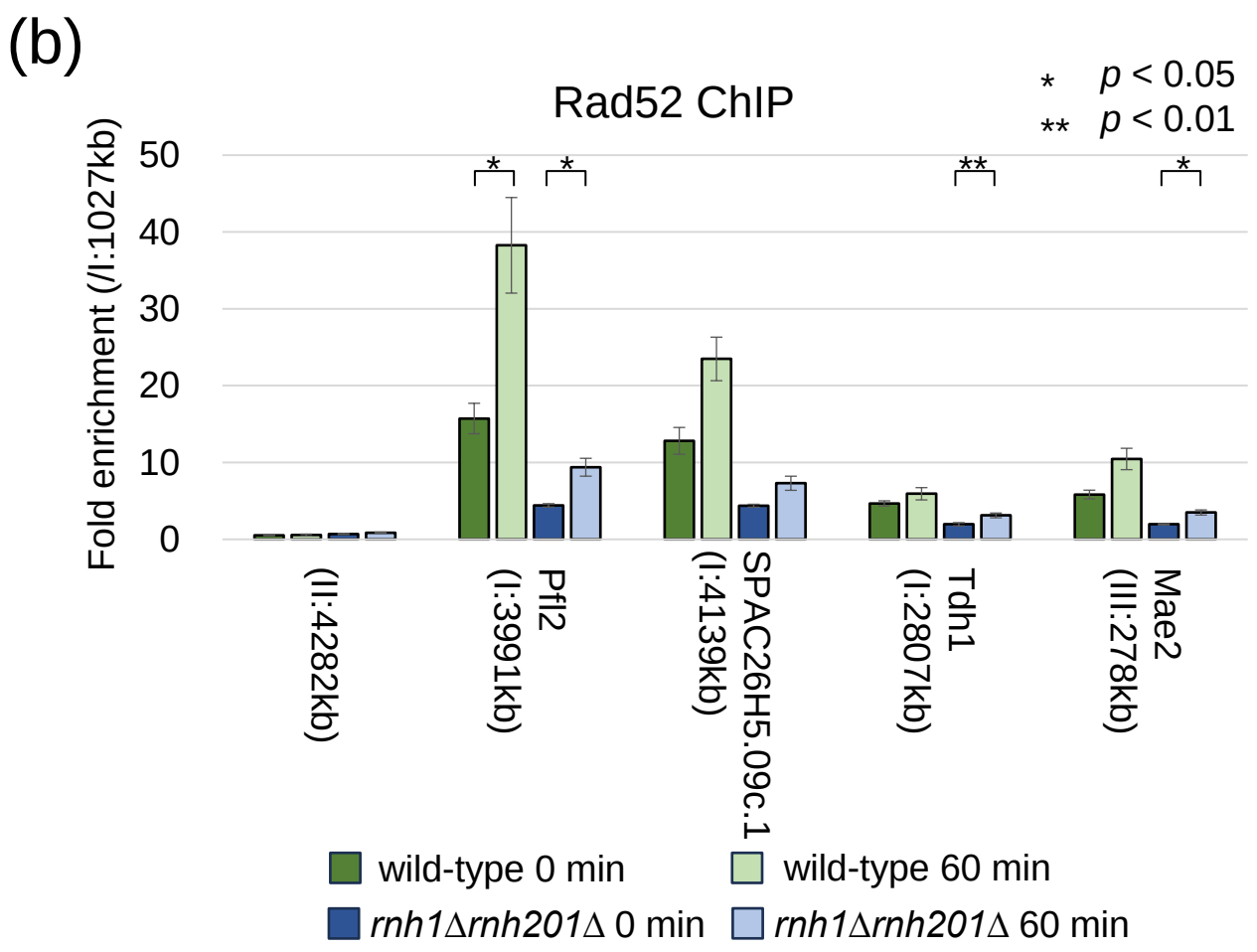


Sagi *et al.* Supplementary Figure S6



Sagi *et al.* Supplementary Figure S7





Sagi *et al.* Supplementary Figure S8