The Affinity of the S9.6 Antibody for Double-Stranded RNAs Impacts the Accurate Mapping of R-Loops in Fission Yeast

Stella R. Hartono[1,2,†], **Amélie Malapert**[3,†], **Pénélope Legros**[3], **Pascal Bernard**[3], **Frédéric Chédin**[1], and **Vincent Vanoosthuyse**[3]

[1]Department of Molecular and Cellular Biology and Genome Center, University of California, Davis, CA, 95616, United States

[2]Integrative Genetics and Genomics Graduate Group, University of California, Davis, CA, 95616, United States

[3]UMR5239 CNRS/Université de Lyon/ENS-Lyon, 46 Allée d'Italie, 69007 Lyon, France

Abstract

R-loops, which result from the formation of stable DNA:RNA hybrids, can both threaten genome integrity and act as physiological regulators of gene expression and chromatin patterning. To characterize R-loops in fission yeast, we used the S9.6 antibody-based DRIPc-seq method to sequence the RNA strand of R-loops and obtain strand-specific R-loop maps at near nucleotide resolution. Surprisingly, preliminary DRIPc-seq experiments identified mostly RNase H-resistant but exosome-sensitive RNAs that mapped to both DNA strands and resembled RNA:RNA hybrids (dsRNAs), suggesting that dsRNAs form widely in fission yeast. We confirmed *in vitro* that S9.6 can immuno-precipitate dsRNAs and provide evidence that dsRNAs can interfere with its binding to R-loops. dsRNA elimination by RNase III treatment prior to DRIPc-seq allowed the genome-wide and strand-specific identification of genuine R-loops that responded *in vivo* to RNase H levels and displayed classical features associated with R-loop formation. We also found that most transcripts whose levels were altered by *in vivo* manipulation of RNase H levels did not form detectable R-loops, suggesting that prolonged manipulation of R-loop levels could indirectly alter the transcriptome. We discuss the implications of our work in the design of experimental strategies to probe R-loop functions.

Keywords

R-loops; S9.6; dsRNA; DRIPc-seq; RNAse H1

Introduction

R-loops are three-stranded nucleic acid structures composed of a stable DNA:RNA hybrid and a displaced single-stranded DNA [1]. R-loops have been found from bacteria to higher eukaryotes and usually form upon re-annealing of the nascent transcript to its DNA template in the wake of the RNA polymerase. R-loop formation is antagonized by a number of enzymes, including RNase H enzymes or DNA&RNA helicases such as the highly conserved Senataxin (reviewed in Ref. [2]). In addition, nucleosome turnover [3,4], chromatin structure and topology or the efficient packaging of nascent RNA into ribonucleotide particles (mRNPs) were also proposed to limit R-loop formation because they restrict the possibility of hybridization between the nascent RNA and its DNA template (reviewed in Ref. [1]).

R-loops are widely mapped genome-wide by exploiting the strong affinity of the S9.6 antibody for DNA:RNA hybrids [5,6]. Recent maps have shown that R-loops can occupy between 5% and 10% of the genome in yeast, vertebrates and *Arabidopsis thaliana* [7–9]. R-loop formation has been associated with replication stress and genome instability, and it was recently proposed that defective control of R-loop formation could contribute to a number of human pathologies (reviewed in Refs. [1,2,10]). However, a mounting body of evidence also suggests that R-loops play physiological roles in transcriptional control and chromatin patterning (reviewed in Refs. [1,2,5]). What distinguishes toxic and beneficial R-loops is still unclear. To answer this question, accurate maps of R-loops must be obtained and assays must be developed to sort R-loops into functional categories.

Whether the S9.6 antibody is the best tool to map R-loops has recently been debated. In-depth characterization of the S9.6 antibody *in vitro* has shown that it recognizes preferentially DNA:RNA hybrids but that it can also recognize AT-rich RNA:RNA hybrids (double-stranded RNAs, or dsRNAs), albeit with a fivefold lower affinity [11]. In addition, it was argued that the affinity of S9.6 for R-loops is influenced by R-loop sequences [12]. The direct consequences of these observations on the accuracy of R-loop mapping by DNA:RNA immuno-precipitation (DRIP)-like methods have not yet been rigorously assessed. Preliminary RNase A treatment has been proposed to increase the specificity of R-loop mapping using DRIP approaches [13], but this was later disputed [14]. However, whether or not RNase A treatment does modulate the efficacy of DRIP, it does not help to evaluate whether dsRNAs are able to interfere with the quantitative recovery of R-loops using S9.6, as RNase A will degrade single-stranded RNA but not dsRNA. To date, it is still not clear whether S9.6 can immuno-precipitate dsRNAs *in vivo* or whether dsRNA can interfere with the interaction between S9.6 and DNA:RNA hybrids.

Here we sought to identify R-loops in the fission yeast *Schizosaccharomyces pombe* using the recently described DRIPc-seq approach, where RNA:DNA hybrids are immuno-precipitated with the S9.6 antibody and their RNA strands are sequenced in a strand-specific manner to map R-loops at near base-pair resolution [8]. This strategy is well suited to mapping R-loops in compact genomes such as fission yeast where genes frequently overlap. Surprisingly, we found that the presence of exosome-sensitive dsRNAs in fission yeast interfered with the mapping of R-loops using this method. Robust R-loop identification by

DRIPc-seq required dsRNA elimination using RNase III treatment. We show that these newly mapped R-loops respond to RNase H levels and exhibit the expected characteristics of R-loop-forming genes. Surprisingly, our results did not confirm that R-loops form over centromeric heterochromatin repeats, as suggested previously [15,16]. Our data indicate instead that S9.6 was able to immuno-precipitate RNase III-sensitive dsRNAs derived from centromeric hetrochromatic repeats. Taken together, our results demonstrate that the affinity of the S9.6 antibody for dsRNAs has the potential to interfere with the accurate and quantitative identification of R-loops by DRIP-like approaches. In addition, we provide evidence that long-term manipulation of RNase H levels causes a consistent transcriptional response affecting mostly non-R-loop-forming genes.

Results

The S9.6 antibody immuno-precipitates dsRNAs *in vitro* and *in vivo*

To map R-loops in fission yeast, we used the recently described DRIPc-seq method, where RNA:DNA hybrids are immuno-precipitated with the S9.6 antibody and their associated RNA strands are sequenced in a strand-specific manner to map R-loops at near base-pair resolution [8] (Fig. S1). Unexpectedly, our first attempts at DRIPc-seq failed to deliver strand-specific signals. Instead, the vast majority of RNAs sequenced mapped to both DNA strands over genes in a nearly symmetrical pattern. This pattern was clearly distinct from the one obtained when sequencing total RNA, suggesting that DRIPc-seq enriched for a specific class of low-abundance RNAs (Fig. 1a). Such DRIPc-seq RNAs were largely resistant to exogenous RNase H treatment and to *in vivo* expression of exogenous RNase H (*Escherichia coli* RnhA), strongly suggesting that they did not correspond to R-loops (Fig. 1a and data not shown). For about half the genes, those RNase H-resistant DRIPc-seq RNAs were particularly enriched at gene ends (Fig. 1a and see gene clusters 1–3, Fig. S2), consistent with the idea that they only represent a specific subset of total RNAs. Interestingly, RNase H-resistant DRIPc-seq RNAs were more abundant when the activity of the RNA exosome was impaired by the mutation of either of its two catalytic sub-units (deletion of the non-essential Rrp6 sub-unit (*rrp6Δ*) or mutation of the essential Dis3 sub-unit (*dis3–54*) (Fig. 1b, c, d). In addition, the proportion of each gene covered by such RNase H-resistant DRIPc-seq RNAs was significantly greater when the activity of the RNA exosome was impaired (Fig. S3b). Because such RNase H-resistant and exosome-sensitive RNAs mapped to both DNA strands, we speculated that they were dsRNAs. Dot blot analysis with the dsRNA-specific J2 antibody confirmed that both exosome mutants accumulate dsRNAs (Fig. 1e). Taken together, these observations suggested that our first attempts at DRIPc-seq in fission yeast identified mostly RNase H-resistant and exosome-sensitive dsRNAs but not R-loops.

It was reported previously that the S9.6 antibody can recognize AT-rich dsRNAs *in vitro*, albeit with a lower affinity than DNA:RNA hybrids (K_D of 0.6 nM for DNA:RNA hybrids against a K_D of 2.7 nM for RNA:RNA hybrids [11]). Here we show that the S9.6 antibody is able to recognize and immuno-precipitate long GC-rich dsRNAs produced *in vitro* from the mouse *Airn* gene [17] (Fig. 2ab). The R-loop-forming *Airn* gene encoded on a plasmid was transcribed *in vitro* either in the sense orientation (using T3 RNA polymerase) or in both sense and antisense orientation (using both T3 and T7 RNA polymerases). As described

previously [17], transcription in the sense orientation was associated with an R-loop-dependent topological shift of the plasmid on agarose gels, which was sensitive to RNase H but resistant to RNase III treatment (Fig. 2a, lanes 2–5). Sense transcription also produced strong RNase H-sensitive and RNase III-resistant signals on S9.6 dot blots (Fig. 2b, dots 2–5). Transcription in both orientations was associated with the production of RNase H- and RNase A-resistant but RNase III-sensitive dsRNAs that ran on agarose gels as a smear (Fig. 2a, lanes 6–9). Strikingly, the associated S9.6 dot blot signals were partially resistant to RNase H treatment (dot 7, Fig. 2b) and only disappeared fully when treated with both RNase H and RNase III (dot 9, Fig. 2b). This strongly suggested that S9.6 was able to recognize GC-rich dsRNAs produced from *Airn* as well as R-loops. This was confirmed by spotting only dsRNAs on S9.6 dot blots (dot 10, Fig. 2c). To build on these observations, we tested whether S9.6 could immuno-precipitate dsRNAs produced from *Airn*. dsRNAs were incubated either alone or in combination with T3-transcribed *Airn* as a source of R-loops and immuno-precipitated with S9.6. We analyzed the input (IN), the unbound fraction (FT), the last wash (W4) and the bound fraction (IP) using agarose gel (Fig. 2d), dot blots (Fig. 2e) and qPCR (Fig. 2f; see Materials and Methods). dsRNAs alone were efficiently captured by S9.6, whether or not R-loops were present. In the absence of dsRNAs, the vast majority of R-loops were captured by the S9.6 antibody since no R-loop was found in the unbound fraction (Fig. 2e). Strikingly, however, when dsRNAs were incubated with R-loops, some R-loops did not bind to S9.6 and remained in the unbound fraction (Fig. 2e, f). As a result, fewer R-loops were found in the bound fraction (Fig. 2f). This experiment showed conclusively that S9.6 could immuno-precipitate dsRNAs and that dsRNAs were able to interfere with the association of S9.6 with R-loops, at least *in vitro* (Fig. 2d–f).

To test whether S9.6 can also recognize dsRNA produced *in vivo*, we studied the centromeric *dg* and *dh* repeats. These repeats generate dsRNAs from bi-directional promoters that are then processed by the RNAi machinery to promote heterochromatin formation (reviewed in Ref. [19]). DRIPc-seq data from *rnh1 rnh201* cells indicated that RNAs originating from centromeric repeats were indeed immuno-precipitated by S9.6 (Fig. 2g). These centromeric DRIPc-seq signals were resistant to exogenous RNase H treatment (Fig. 2g) or *in vivo* expression of exogenous RnhA (data not shown) but sensitive to RNase III treatment, strongly suggesting that they corresponded to dsRNAs and not to RNA:DNA hybrids. We confirmed that RNase III did not affect R-loop formation and detection using *in vitro* substrates (Fig. 2a). Taken together, these observations establish that S9.6 can immuno-precipitate *in vivo* produced and biologically relevant dsRNAs. Furthermore, these data strongly suggest that the affinity of the S9.6 antibody for dsRNAs can confound results of R-loop mapping approaches through DRIPc-seq type experiments, at least in fission yeast where significant loads of exosome-sensitive dsRNAs are produced.

Improved R-loop mapping identifies two classes of R-loops

To map R-loops in the absence of contaminating dsRNAs, we treated genomic DNA with RNase III prior to S9.6 immuno-precipitation. Strikingly, DRIPc-seq signals after RNase III treatment exhibited all the hallmarks of genuine R-loops: they mapped to transcribed gene units, were strand-specific and co-directional with transcription, were sensitive to the *in vivo* expression of exogenous RNase H (*E. coli* RnhA) and were stabilized by loss of endogenous

RNase H activity (*rnh1 rnh201*) (Fig. 3a, b). Upon expression of RnhA, stranded DRIPc-seq signals were still detectable at ~25% of loci forming R-loops in wild-type cells, albeit with significantly reduced intensity and length (Fig. 3b, c). On the contrary, in the absence of endogenous RNase H activity (*rnh1 rnh201*), both the number of R-loop-forming loci and the percentage of the genome covered by R-loops increased significantly (Fig. 3b, c). Taken together, these data show that the mapping of R-loops by DRIPc-seq in fission yeast requires the elimination of dsRNAs by RNase III treatment. Unless further specified, all references to DRIPc from this point on refer to DRIPc-seq signal after RNase III treatment.

Our improved DRIPc-seq data identified two classes of R-loops, here referred to as classes A and B, which differ mainly by their sensitivity to endogenous RNase H activity. Class A R-loops were reproducibly detected at ~500 loci in wild-type cells, representing roughly 5% of the 12.57-Mbp-long genome (Fig. 3c). In cells lacking endogenous RNase H activity (*rnh1 rnh201*), we detected an additional ~1000 R-loop-forming loci, which we refer to as class B R-loops. Overall, class B R-loops cover an additional ~10% of the genome. The average size of individual class A R-loops was longer than class B R-loops (Fig. S4) and formed predominantly over highly transcribed genes and Tf2 retrotransposons, in a similar pattern to R-loop formation in *Saccharomyces cerevisiae* [7,20,21] (Fig. 3d and data not shown). Upon expression of bacterial RnhA, the few remaining detectable R-loops also formed at highly expressed genes, presumably from incomplete degradation of class A R-loops. On the contrary, class B R-loops formed at loci with weaker transcript levels (Fig. 3d) and likely correspond to transcription units where transient R-loop formation is efficiently antagonized by endogenous RNase H enzymes. The heavily transcribed tDNAs, which accumulated class B R-loops in the absence of endogenous RNase H activity (Fig. 2c and see [18]), were an exception to this low-expression rule for class B R-loops. The presence of R-loops at tDNAs confirms our previous observations [18] and observations made in budding yeast [7,20,21] and *A. thaliana* [9]. The low-level R-loop formation at tDNAs in wild-type cells is also consistent with the high recruitment of RNase H1 and Senataxin (Sen1) at these loci in fission yeast [18].

We did not detect a significant bias for R-loop formation toward either the promoter or the terminator regions of genes (Fig. 3e), but we detected a significant strand bias for G over C (GC skew) in the terminal regions of class A R-loops (Fig. 3f). By contrast, AT skew, which was identified as a feature of *S. cerevisiae* R-loops [7], was not strongly different in R-loop-forming *versus* expression-matched non-R-loop genes (Fig. 3f). Finally, we analyzed whether R-loop-forming genes were associated with certain chromatin marks or factors compared to expression- and length-matched non-R-loop genes. R-loop-forming genes were significantly associated with a number of chromatin factors and marks compared to matched genes. The SMARCAD1 family ATP-dependent DNA helicase Fft3, the H3K4 methyltransferase Swd1, the H3K36 methyltransferase Set2, the histone chaperone Spt6 and the Paf1 complex subunit Ctr9[Tpr1] were significantly enriched over R-loop-forming genes (Fig. S5). In conclusion, R-loop formation in fission yeast is widespread and associated with specific chromatin features characteristic of transcription elongation.

RNase H level manipulations in fission yeast alters the transcriptome

Because we showed that manipulating RNase H levels *in vivo* affected R-loop formation, we assessed the impact of R-loop levels on the transcriptome of exponentially growing vegetative cells. For this, R-loops were either resolved by expression of bacterial RnhA in fission yeast cells or stabilized by deletion of both endogenous RNase H genes, and the consequences on the transcriptome determined by strand-specific RNA-seq on ribo-depleted RNAs. RnhA expression impacted the steady-state levels of 1218 transcripts out of 7016, including 1012 protein-coding mRNAs (Fig. 4a). Of those protein-coding mRNAs, the abundance of 422 transcripts increased. Gene ontology (GO) analysis indicated that many up-regulated genes took part in ribosome biogenesis and nucleolus function (Fig. 4a). Deletion of both endogenous RNase H genes was accompanied by significant changes in expression of 660 transcripts, including 584 protein-coding genes (Fig. 4a). Many up-regulated genes again showed strong enrichment for GO terms related to ribosome assembly and nucleolus biogenesis. Further analysis revealed a significant overlap between the transcripts impacted by RnhA expression (loss of R-loops) and those affected by RNase H deletions (R-loop stabilization) (Fig. 4b). For instance, genes involved in ribosome biogenesis and assembly tended to be up-regulated in both conditions, while genes involved in iron assimilation and glycan metabolism tended to be down-regulated. This unexpected observation suggests that long-term R-loop level manipulation through RNase H deletion or over-expression strategies impart significant and specific transcriptional responses in fission yeast.

RNase H level manipulations in fission yeast alters predominantly the expression of genes that do not form detectable R-loops

To further understand the relationship between R-loop level perturbations and the associated transcriptional response, we looked for R-loop-forming transcripts whose expression was impacted by RNase H over-expression or deletion. Surprisingly, the expression of less than a quarter of R-loop-forming genes in wild-type cells (class A) was affected by RnhA expression (pRnhA, Fig. 4c). Conversely, only about 10% of genes showing significant gene expression changes upon RnhA expression were R-loop forming in wild-type cells (Fig. 4c). Similar results were obtained when focusing on class B R-loops (Fig. 4d): transcript abundance at genes that gain R-loops upon RNase H deletion (class B) was for the most part not affected by loss of RNase H (Fig. 4d, left panel) or gain of RNase H (Fig. 4d, right panel). This shows that while R-loop level manipulation caused significant alterations to the transcriptome, the vast majority of responsive genes did not form R-loops that we could detect using DRIPc-seq. Taken together, our results suggest that the manipulation of RNase H levels in fission yeast triggers a global transcriptional response that impacts predominantly transcripts involved in ribosome biogenesis, irrespective of whether or not they can form detectable R-loops.

Discussion

Caveats associated with the use of the S9.6 antibody to map R-loop-forming regions

We provide evidence that the S9.6 antibody used in all variants of the DRIP method to map R-loops can also recognize and immuno-precipitate RNase III-sensitive dsRNAs both *in*

vitro and *in vivo* in fission yeast (Figs. 1 and 2). As a result, it is likely that S9.6 can also recognize structured single-stranded RNA molecules that form hairpins. Our work shows that the affinity of S9.6 for dsRNAs becomes a problem when DRIP is followed by RNA sequencing for increased resolution, such as in DRIPc-seq, because it can lead to mixed signals from R-loops and dsRNAs. This is true in organisms or genotypes where significant dsRNA loads are produced. Recent DRIPc-seq maps in mammalian cells were strictly directional even in the absence of RNase III treatment [8], suggesting that dsRNA formation was not an issue, at least in the conditions tested. Based on our *in vitro* results, it is also conceivable that even when DRIP is followed by direct DNA sequencing or qPCR, such as in classic DRIP-seq or DRIP-qPCR, competition from dsRNAs could affect the efficiency of the S9.6 immunoprecipitation (Fig. 2). To circumvent this issue and ensure the quantitative enrichment of R-loops with the S9.6 antibody, we would recommend the use of RNase III treatments prior to immuno-precipitation. Finally, it is noteworthy that in immuno-fluorescence experiments, the strongest nuclear S9.6 signals consistently map to nucleoli and is at least partially resistant to RNase H treatment. This might also reflect S9.6 binding to a mixture of R-loops and to highly structured and abundant rRNAs.

DRIPc-seq data prior to RNase III treatment suggest that exosome-sensitive dsRNA species form widely in fission yeast, arguing for widespread antisense transcription. Such dsRNAs were not observed by regular RNA-seq experiments, suggesting that they most likely correspond to low-abundance RNA species that are detected after enrichment by the S9.6 antibody. dsRNAs were also recently detected in the distant budding yeast [22], suggesting that widespread dsRNA formation is a conserved phenomenon, at least in yeasts. dsRNA formation in fission yeast is associated with gene silencing through heterochromatin formation *via* the RNAi pathway [19]. Our data indicate that the percentage of the genome to which these dsRNAs map (4.6 Mb corresponding to 36.5% of the genome) is a lot greater than the percentage of the genome covered by heterochromatin (roughly 500 kb, corresponding to ~4% of the genome). Further investigation will be required to establish whether or not such widespread low-level dsRNA formation plays a biological role.

Improved R-loop mapping in fission yeast

The R-loop maps provided here represent a significant improvement over the two maps that were published previously in fission yeast. The first map was generated by DRIP-seq and as a result was not stranded and had low resolution [16]. The second map was made using a ChIP-exo approach to increase resolution, but it only provided a map of R-loops in cells lacking RNase H activity (*rnh1 rnh201* cells) and the sensitivity of the signals to RNase H was not controlled [23]. The maps produced here are stranded and have high resolution, and the sensitivity of the signals to *in vivo* RNase H level manipulations has been validated.

While our maps recapitulate many observations made previously in yeast (R-loop formation at highly expressed G-rich genes, at Tf2 retrotransposons and at tDNAs), they differ from previous maps in several respects. For instance, contrary to what was proposed before [16,23], we could not detect significant R-loop formation either over heterochromatin repeats at centromeres or over gene terminator regions in fission yeast. Although we confirmed that the S9.6 antibody immuno-precipitates RNA species associated with

centromeric repeats and terminator regions (Figs. 1, 2 and S2), our data suggest that these RNAs are resistant to both *in vivo* RNase H over-expression and *in vitro* RNase H treatment while being sensitive to RNase III treatment. As such, these RNAs are more likely to correspond to dsRNAs rather than R-loops. We speculate that at least a fraction of these dsRNAs remain associated with chromatin as was suggested previously for heterochromatin transcripts [24], which might explain why they have been previously detected using DRIP. Interestingly, lack of a significant enrichment for R-loops at terminator regions was also recently observed in *A. thaliana* [9]. The situation in *S. pombe* and *A. thaliana* is therefore different from human and mouse where genuine R-loops form at a subset of terminator regions [8,17,25–28].

Altogether, our improved R-loop maps show that 5% of the fission yeast genome can exist under an R-loop form in unperturbed cells, an observation consistent with findings in mammals and plants [8,9]. Under these conditions, R-loop formation is principally a property of highly transcribed genes (class A). When cells lack endogenous RNase H activity, a second class of R-loops becomes detectable, covering an additional 8% of the genome. This shows that one of the roles of endogenous RNase H enzymes is to suppress low-level R-loop formation over a significant portion of the genome. Our data, combined with observations from mammalian cells and *A. thaliana* [8,9,17,25], reinforce the idea that genuine R-loops share universal hallmarks, including strand specificity, co-directionality with transcription, co-dependence with transcription levels and a strong relationship with favorable DNA sequence features such as GC skew (Fig. 3). In addition, we show that R-loop-forming loci in fission yeast significantly associate with chromatin marks of active transcription and transcription elongation, consistent with properties of mammalian and plant R-loops [8,9] These associations were significant even after stringent matching to non-R-loop-forming genes with similar lengths and expression levels.

Manipulation of R-loop levels over many cell generations imparts significant transcriptome alterations

We report here that *in vivo* manipulations of R-loop levels trigger significant transcriptome alterations and that the responses to R-loop loss or stabilization are unexpectedly similar (Fig. 4). This could be consistent with the idea proposed previously that the formation and the resolution of R-loops are equally important for gene expression [28]. How opposite changes in R-loop turnover rates are translated to similar gene expression responses remains unclear, however. Alternatively, we propose that the manipulation of R-loop levels over many generations might trigger a global cellular stress response. Our observation that up-regulated genes are enriched for functions associated with ribosome biogenesis and protein synthesis suggests that cells are coping with perturbations in the process of protein translation. An attractive, but as yet untested, hypothesis to account for this transcriptional response is that perturbations of R-loop levels negatively impact the transcription, processing or export of R-loop prone rRNAs, ultimately interfering with ribosome assembly. Interestingly, perturbations of ribosome assembly were shown recently to result in the Spt6-dependent activation of ribosome-related genes in *S. cerevisiae* [29]. If this feedback response also exists in *S. pombe,* it could explain why R-loop level manipulations result in the increased abundance of transcripts involved in ribosome assembly (Fig. 4).

Regardless of its origin, this cellular response may mask the immediate impact of RNase H level manipulations on the expression of R-loop-forming genes. Indeed, consistent with previous observations in *S. cerevisiae* [20], we observed that transcript levels of most R-loop-forming genes were not significantly altered by RNase H level manipulations over long periods of time (Fig. 4), A corollary to these observations is that collateral alterations to the transcriptome and the proteome could contribute to the phenotypes associated with prolonged RNase H1 over-expression, complicating the interpretation of such experiments. Although technically challenging, a suitable alternative to this approach could be to expose cells to high levels of RNase H1 for a very short period of time only.

To conclude, our work shows that the cross-reactivity of the S9.6 for dsRNA and possible indirect perturbations to the transcriptome associated with prolonged manipulation of R-loop levels are two confounding parameters that must be taken into account in the design of experimental strategies to understand R-loop functions.

Material and Methods

Fission yeast strains

A complete list of all of the strains used in this study is given in Table S1. Standard genetic crosses were employed to construct all strains. To express bacterial RNase H1 in fission yeast, the *RnhA* gene from *E. coli* was placed under the control of the thiamine-repressible *nmt1* promoter and integrated stably into the genome as described previously [18].

R-loop mapping *via* DRIPc-seq

DRIP was performed as before [18] with the following modifications to the way the genomic DNA was extracted: 6×10^8 cells were resuspended in 7.5 mL of pre-incubation buffer [20 mM citric acid, 20 mM Na_2HPO_4, 40 mM EDTA (pH 8.0) to which 30 mM of β-mercaptoethanol was added just before use] and incubated at 30 °C for 10 min under vigorous agitation. Cells were then pelleted and resuspended in 3.5 mL of Sorb/Tris buffer [1 M Sorbitol, 50 mM Tris/HCl (pH 7.4), 10 mM β-mercaptoethanol containing 10 mg of Zymolyase 100 T] and incubated at 30 °C for 20 min under vigorous agitation then immediately transferred on ice. The resulting spheroplasts were washed in 15 mL of Sorb/Tris buffer and resuspended in 375 μL of NP buffer [1 M Sorbitol, 50 mM NaCl, 10 mM Tris/HCl (pH 7.4), 5 mM $MgCl_2$, 1 mM $CaCl_2$, 0.75% IGEPAL, 1 mM β-mercaptoethanol] to which 3 μg/mL of RNase A and 700 μL of 10% SDS were added. After 1 h at 37 °C, phenol/chloroform and ethanol precipitation was used to purify the genomic DNA. The DNA was digested with restriction enzymes (EcoRI, SspI, BsrGI, XbaI, HindIII) as described before [18], in the presence of 2 U of RNase III (Ambion, AM2290) or 4 U of RNase H (Roche 10786357001). The digested DNA was then separated into two, and each half was incubated overnight at 4 °C with 40 μL of Protein A-coupled Dynabeads (Invitrogen 10002D) previously incubated with 10 μg of S9.6 antibody according to the manufacturer's instructions. After extensive washes of the beads, the immuno-precipitated nucleic acids were pooled, purified as previously described [18] and heated for 1 min at 94 °C to denature DNA:RNA hybrids. DNA molecules were then digested with DNase I (Invitrogen AM1907) for 1 h at 37 °C and the RNA molecules were recovered after ethanol

precipitation. The RNA molecules were analyzed and quantified using both Bioanalyzer (Agilent) and Quantus (Promega). Sequencing libraries were prepared using NEXTflex ™ Rapid Directional mRNA-Seq Kit (dUTP-Based; Bioo Scientific NOVA-5138-07) from 10 ng of RNA, without Poly-A enrichment or RNA fragmentation. Paired-End sequencing (75 bp) was performed on a NextSeq500 (Illumina).

Bioinformatics analysis

Both DRIPc-seq and RNA-seq reads were trimmed with fastq-mcf. DRIPc-seq reads were mapped to *S. pombe* genome version ASM294v2 (accession GCA_000002945.2) using Bowtie v2.1.0, and RNA-seq reads were mapped with Tophat v2.0.9. Gene sequence and annotation (version 4/2015) were downloaded from pombase. DRIPc-seq peaks were called using StochHMM as previously described in Ref. [8]. Reads were counted using htseq-count v0.6.0; RNA-seq reads were counted per annotated gene, while DRIPc-seq reads were counted per DRIPc-seq peaks. For both RNA-seq and DRIPc-seq, significantly up- and down-regulated genes and peaks were determined using DESeq2 v1.10.1, with threshold of at least $1.25\times$ fold change compared to wild-type (higher or lower) and an adjusted *P*-value of 0.1 or less. GO analysis was performed using DAVID v6.8 with *S. pombe* as reference. All programs were run with default parameters unless mentioned. External ChIP-seq data set was downloaded from GEO accession number GSE49575. To test if R-loop-forming genes are enriched or depleted over a specific factor, we performed shuffling test as described in Ref. [8]. Briefly, R-loop-forming genes were shuffled 125 times on a set of randomly chosen non-R-loop-forming genes and matched for gene expression and length, and the results were used to calculate fold enrichment and empirical *P*-values.

Dot blot analysis on genomic DNA

The genomic DNA was purified and digested with restriction enzymes as for the DRIPc (see above). After digest, the genomic DNA was quantified by qPCR using three independent primer pairs. DNA concentrations were then normalized between samples, and a new round of qPCR with the same three independent primer pairs was used to confirm the correct normalization. Two microliters of DNA was spotted on a Hybond ™-N+ membrane (Amersham RPN203B) and cross-linked twice with UV (0.12 J). To quantify R-loop formation, the membrane was blocked in Mes/BSA buffer [100 mM Mes (pH 6.6), NaCl 500 mM, 0.05% Triton, 2 mg/mL BSA] for 30 min hour at RT, and 0.02 μg/mL of S9.6 antibody was added and incubated overnight. After extensive washes in Mes/BSA buffer and exposure to a secondary antibody, ECL revelation was performed using SuperSignal West Femto Maximum Sensitivity Substrate (ThermoScientific 34,096) according to standard protocols. To quantify dsRNA or dsDNA, the membrane was first blocked in PBS $1\times$ 0.02% Tween-20 containing 5% milk and incubated overnight with the dsRNA-specific J2 antibody (SCICONS J2 tech, dilution 1/1000) or the dsDNA-specific (ab27156 Abcam, dilution 1/1000).

In vitro transcription of R-loops

The pFC53 plasmid contains the R-loop-forming part of the *Airn* transcription unit from mouse [17]. 750 nanograms of pFC53 were transcribed for 30 min at 37 °C in the sense orientation using the T3 RNA polymerase according to the manufacturer's instructions.

After inactivation of the polymerase at 65 °C for 10 min, the reaction was incubated for 2 h at 37 °C with 0.25 μg RNase A and 20 units of ApaLI to cleave the plasmid in two when needed (Fig. 2b). When required, 2 units of RNase H was added to the reaction. The reaction buffer was as follows: 50 mM potassium acetate, 20 mM Tris–acetate, 10 mM magnesium acetate and 100 μg/mL BSA (pH 7.9). Ten micrograms of Proteinase K was then added for 30 min at 37 °C before purifying the DNA using standard phenol/chloroform and ethanol precipitation techniques.

Production of dsRNAs

4.5 micrograms of pFC53 (see above) were transcribed for 30 min at 37 °C in both orientations using both T3 and T7 RNA polymerases. After inactivation of the polymerase at 65 °C for 10 min, the reaction was incubated with 30 μg of RNase A for 20 min at 37 °C. The plasmid DNA was then digested with DNase I (Invitrogen AM1907) according to the manufacturer's instructions. dsRNAs were then purified using standard phenol/chloroform and ethanol precipitation techniques. RNase III was purchased from Ambion (reference AM2290), and digests were performed according to the manufacturer's instructions.

Immuno-precipitation of R-loops and dsRNA *in vitro*

The purified R-loops produced from 750 ng of pFC53 and the dsRNAs purified from 4.5 μg of pFC53 were incubated alone or in combination in 300 μL of interaction buffer [Tris 10 mM (pH 8.0), NaCl 140 mM, 0.5% Triton]. One hundred microliters was kept as input (fraction IN, Fig. 2b). The remaining 200 μL was incubated with 4 μg of S9.6 antibody coupled to 15 μL of Protein A-coupled Dynabeads at 4 °C with agitation. After an hour of incubation, the beads were collected and the FT was kept. The beads were washed three times with 1 mL of interaction buffer (5-min incubation each time) and then resuspended in 200 μL of incubation buffer. After 5-min incubation, the 200 μL was collected and kept as fraction W4. The beads were resuspended in 200 μL of elution buffer [50 mM Tris (pH 8.0), 10 mM EDTA, 0.5% SDS] containing 50 μg of Proteinase K (ThermoFisher EO0491) and incubated at 55 °C. After 30 min, the beads were eliminated and the supernatant was recovered (IP fraction). The IN, FT, W4 and IP fractions were then purified using standard phenol/chloroform and ethanol precipitation techniques and resuspended in 22 μL of Tris 10 mM (pH 8.0). Six microliters was used for dot blot analysis with the S9.6 (DNA:RNA hybrids), J2 (RNA:RNA hybrids) or ab27156 (dsDNA) antibodies. Twelve microliters was run on a 0.8% agarose TBE 1× gel. The different fractions were then diluted 1/10,000 for qPCR analysis using the primers AIRN qL1 (5′-TCAG CACAACCAAGGATCAC-3′) and AIRN qR1 (5′-AGG GTGAAAAGCTGCACAAG-3′) to monitor the enrichment of the R-loop-forming fragment and using the primers Amp 216 FW (5′-AACTTTATCCGCCTCCAT CC-3′) and Amp 353 RV (5′-AAGCCATACCAAAC GACGAG-3′) to monitor the enrichment of the control fragment.

RNA-seq

RNAs were extracted as previously described [30] from biological triplicates (2×10^8 cells per sample) and analyzed and quantified using both Bioanalyzer (Agilent) and Quantus (Promega). Ribosomal RNAs were depleted from 2 μg of total RNAs. Strand-specific sequencing libraries were prepared from 2 ng of rRNA-depleted RNAs using the NEXT-

flex™ Rapid Directional mRNA-seq (Bioo Scientific) and sequenced on a HiSeq2500 Illumina (single reads 50 bp).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Abbreviations used

|	
dsRNAs	double-stranded RNAs
IN	input
FT	unbound fraction
W4	last wash
IP	bound fraction
GO	Gene ontology
DRIP, DNA	RNA immuno-precipitation

References

1. Santos-Pereira JM, Aguilera A. R loops: new modulators of genome dynamics and function. Nat Rev Genet. 2015; 16:583–597. [PubMed: 26370899]

2. Skourti-Stathaki K, Proudfoot NJ. A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. Genes Dev. 2014; 28:1384–1396. [PubMed: 24990962]

3. Herrera-Moyano E, Mergui X, Garcia-Rubio ML, Barroso S, Aguilera A. The yeast and human FACT chromatin-reorganizing complexes solve R-loop-mediated transcription-replication conflicts. Genes Dev. 2014; 28:735–748. [PubMed: 24636987]

4. Taneja N, Zofall M, Balachandran V, Thillainadesan G, Sugiyama T, Wheeler D, et al. SNF2 family protein Fft3 suppresses nucleosome turnover to promote epigenetic inheritance and proper replication. Mol Cell. 2017; 66:50–62(e6). [PubMed: 28318821]

5. Chedin F. Nascent connections: R-loops and chromatin patterning. Trends Genet. 2016; 32:828–838. [PubMed: 27793359]

6. Boguslawski SJ, Smith DE, Michalak MA, Mickelson KE, Yehle CO, Patterson WL, et al. Characterization of monoclonal antibody to DNA.RNA and its application to immunodetection of hybrids. J Immunol Methods. 1986; 89:123–130. [PubMed: 2422282]

7. Wahba L, Costantino L, Tan FJ, Zimmer A, Koshland D. S1-DRIP-seq identifies high expression and polyA tracts as major contributors to R-loop formation. Genes Dev. 2016; 30:1327–1338. [PubMed: 27298336]

8. Sanz LA, Hartono SR, Lim YW, Steyaert S, Rajpurkar A, Ginno PA, et al. Prevalent, dynamic, and conserved R-loop structures associate with specific epigenomic signatures in mammals. Mol Cell. 2016; 63:167–178. [PubMed: 27373332]

9. Xu W, Xu H, Li K, Fan Y, Liu Y, Yang X, et al. The R-loop is a common chromatin feature of the *Arabidopsis* genome. Nat Plants. 2017; 3(9):704–714. [PubMed: 28848233]

10. Richard P, Manley JL. R loops and links to human disease. J Mol Biol. 2016; 429:3168–3180. [PubMed: 27600412]

11. Phillips DD, Garboczi DN, Singh K, Hu Z, Leppla SH, Leysath CE. The sub-nanomolar binding of DNA–RNA hybrids by the single-chain Fv fragment of antibody S9.6. J Mol Recognit. 2013; 26:376–381. [PubMed: 23784994]

12. Konig F, Schubert T, Langst G. The monoclonal S9.6 antibody exhibits highly variable binding affinities towards different R-loop sequences. PLoS One. 2017; 12:e0178875. [PubMed: 28594954]

13. Zhang ZZ, Pannunzio NR, Hsieh CL, Yu K, Lieber MR. Complexities due to single-stranded RNA during antibody detection of genomic rna:dna hybrids. BMC Res Notes. 2015; 8:127. [PubMed: 25890199]

14. Halasz L, Karanyi Z, Boros-Olah B, Kuik-Rozsa T, Sipos E, Nagy E, et al. RNA–DNA hybrid (R-loop) immuno-precipitation mapping: an analytical workflow to evaluate inherent biases. Genome Res. 2017; 27:1063–1073. [PubMed: 28341774]

15. Nakama M, Kawakami K, Kajitani T, Urano T, Murakami Y. DNA–RNA hybrid formation mediates RNAi-directed heterochromatin formation. Genes Cells. 2012; 17:218–233. [PubMed: 22280061]

16. Castel SE, Ren J, Bhattacharjee S, Chang AY, Sanchez M, Valbuena A, et al. Dicer promotes transcription termination at sites of replication stress to maintain genome stability. Cell. 2014; 159:572–583. [PubMed: 25417108]

17. Ginno PA, Lott PL, Christensen HC, Korf I, Chedin F. R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. Mol Cell. 2012; 45:814–825. [PubMed: 22387027]

18. Legros P, Malapert A, Niinuma S, Bernard P, Vanoosthuyse V. RNA processing factors Swd2.2 and Sen1 antagonize RNA Pol III-dependent transcription and the localization of condensin at Pol III genes. PLoS Genet. 2014; 10:e1004794. [PubMed: 25392932]

19. Martienssen R, Moazed D. RNAi and heterochromatin assembly. Cold Spring Harb Perspect Biol. 2015; 7:a019323. [PubMed: 26238358]

20. Chan YA, Aristizabal MJ, Lu PY, Luo Z, Hamza A, Kobor MS, et al. Genome-wide profiling of yeast DNA:RNA hybrid prone sites with DRIP-chip. PLoS Genet. 2014; 10:e1004288. [PubMed: 24743342]

21. El Hage A, Webb S, Kerr A, Tollervey D. Genome-wide distribution of RNA–DNA hybrids identifies RNase H targets in tRNA genes, retrotransposons and mitochondria. PLoS Genet. 2014; 10:e1004716. [PubMed: 25357144]

22. Wery M, Descrimes M, Vogt N, Dallongeville AS, Gautheret D, Morillon A. Nonsense-mediated decay restricts LncRNA levels in yeast unless blocked by double-stranded RNA structure. Mol Cell. 2016; 61:379–392. [PubMed: 26805575]

23. Ohle C, Tesorero R, Schermann G, Dobrev N, Sinning I, Fischer T. Transient RNA–DNA hybrids are required for efficient double-strand break repair. Cell. 2016; 167:1001–1013(e7). [PubMed: 27881299]

24. Bronner C, Salvi L, Zocco M, Ugolini I, Halic M. Accumulation of RNA on chromatin disrupts heterochromatic silencing. Genome Res. 2017; 27:1174–1183. [PubMed: 28404620]

25. Ginno PA, Lim YW, Lott PL, Korf I, Chedin F. GC skew at the 5′ and 3′ ends of human genes links R-loop formation to epigenetic regulation and transcription termination. Genome Res. 2013; 23:1590–1600. [PubMed: 23868195]

26. Hatchi E, Skourti-Stathaki K, Ventz S, Pinello L, Yen A, Kamieniarz-Gdula K, et al. BRCA1 recruitment to transcriptional pause sites is required for R-loop-driven DNA damage repair. Mol Cell. 2015; 57:636–647. [PubMed: 25699710]

27. Skourti-Stathaki K, Kamieniarz-Gdula K, Proudfoot NJ. R-loops induce repressive chromatin marks over mammalian gene terminators. Nature. 2014; 516:436–439. [PubMed: 25296254]

28. Skourti-Stathaki K, Proudfoot NJ, Gromak N. Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. Mol Cell. 2011; 42:794–805. [PubMed: 21700224]

29. Gomez-Herreros F, Margaritis T, Rodriguez-Galan O, Pelechano V, Begley V, Millan-Zambrano G, et al. The ribosome assembly gene network is controlled by the feedback regulation of transcription elongation. Nucleic Acids Res. 2017; 45:9302–9318. [PubMed: 28637236]

30. Vanoosthuyse V, Legros P, van der Sar SJ, Yvert G, Toda K, Le Bihan T, et al. CPF-associated phosphatase activity opposes condensin-mediated chromosome condensation. PLoS Genet. 2014; 10:e1004415. [PubMed: 24945319]

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2017.12.016.
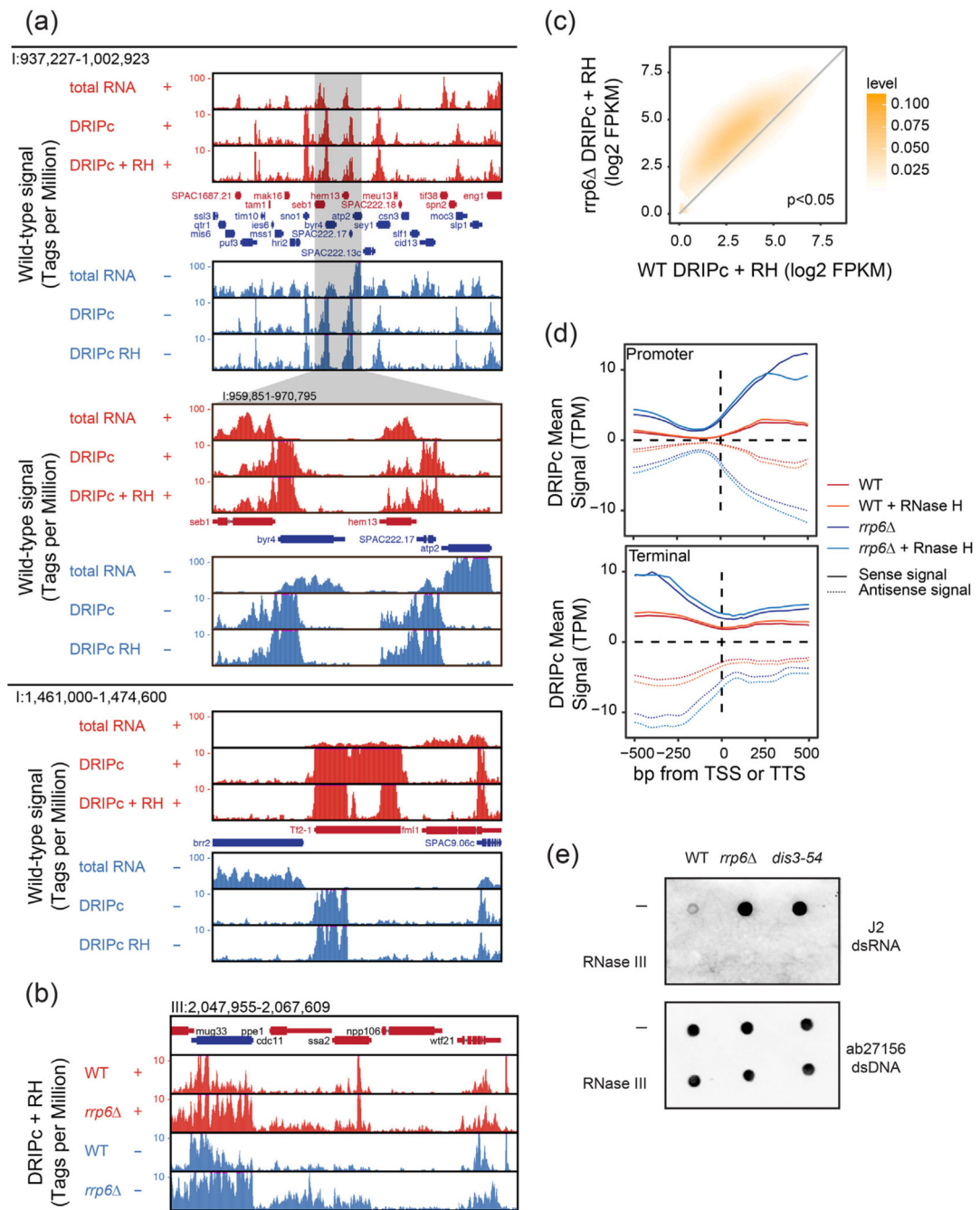
Fig. 1.
DRIPc-seq identifies RNase H-resistant but exosome-sensitive transcripts in fission yeast.
(a) Representative genomic regions showing RNase H-resistant DRIPc-seq signals in wild-type cells. Signals are broken down by strand (red, + strand; blue, − strand) and represent total RNA, DRIPc-seq and RNase H-pre-treated DRIPc-seq in each sub-panel. Annotated genes are shown. Note that RNase H-resistant DRIPc-seq RNAs are often enriched at regions of gene convergence and that DRIPc-seq RNAs at Tf2 transposons are partially sensitive to exogenous RNase H pre-treatment. (b) Representative genomic region showing

RNase H-resistant DRIPc-seq signals in rrp6+ and *rrp6* cells. (c) *X–Y* plot of RNase H-resistant DRIPc-seq signals at each locus in *rrp6* (*y*-axis) *versus* rrp6+ (*x*-axis) cells. (d) Metaplot of DRIPc-seq signals in rrp6+ and *rrp6* cells at promoter and terminator regions. (e) Nucleic acids extracted from wild-type or exosome-deficient *rrp6* and *dis3–54* cells were pre-treated or not with RNase III, and the amount of dsRNA or dsDNA was analyzed on separate dot blots.
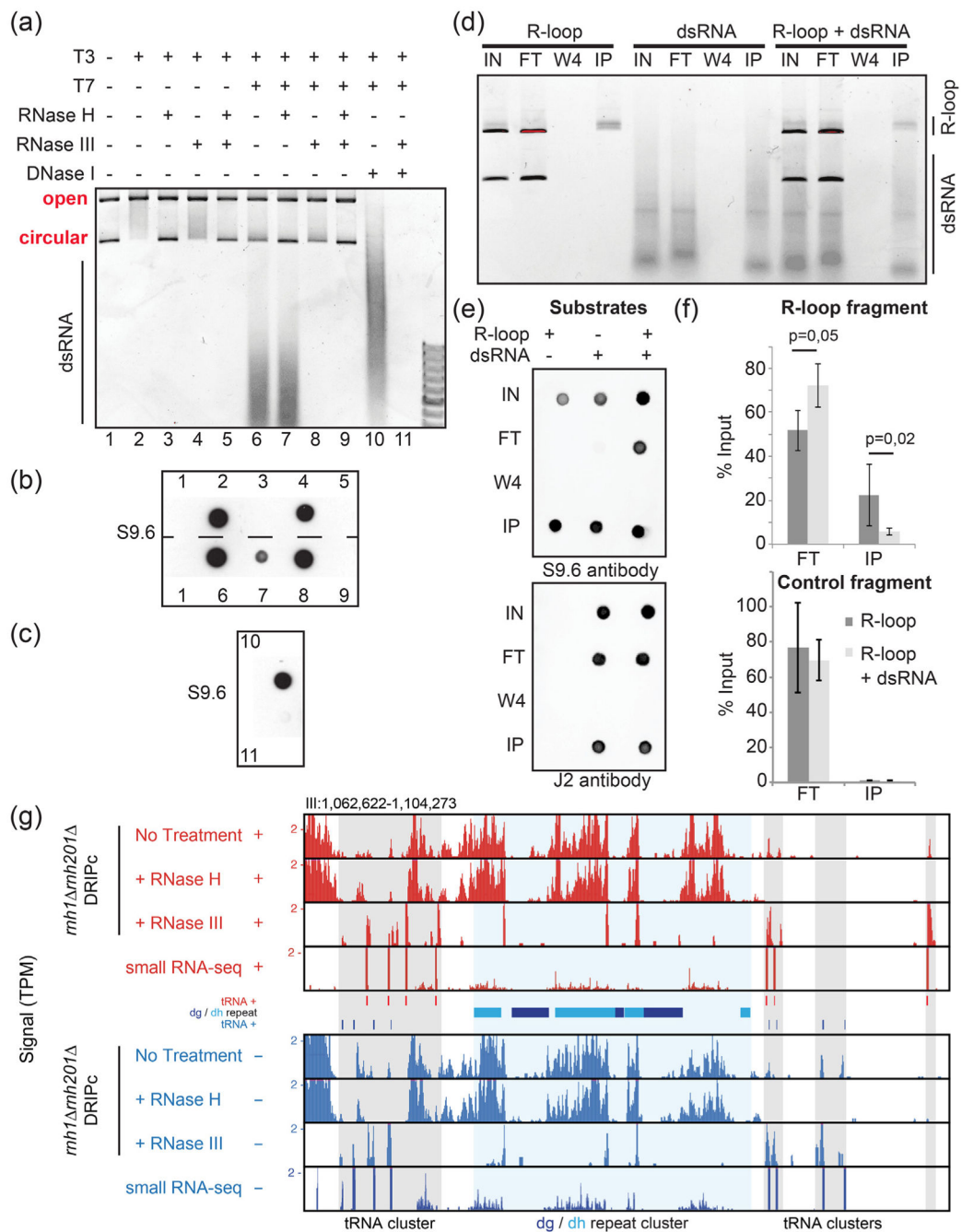
Fig. 2.
The S9.6 antibody recognizes and immune-precipitates dsRNAs both *in vitro* and *in vivo* in fission yeast. (a) A plasmid carrying the R-loop-forming *AIRN* mouse gene [17] was transcribed *in vitro* in the sense (T3) or antisense (T7) orientations as indicated. Soluble RNAs and proteins were then digested with RNase A and Proteinase K, respectively. After concomitant sense and antisense transcription, the reaction was split into two. In one half of the reaction, the plasmid was digested using DNase I treatment and only the dsRNAs were kept. (a) The reactions were run on an agarose gel. R-loops formed upon sense transcription

(T3) induced a topological shift in the plasmid as reported previously [17]. When both T3 and T7 polymerases were used at the same time, RNase III-sensitive dsRNAs appeared in the gel. (b) Dot blot analysis of the above reactions using the S9.6 antibody. Upon sense and antisense transcription, the RNase H treatment is not sufficient to remove the signal. The signal fully disappears only after treatment with both RNase H and RNase III, suggesting that the S9.6 can recognize dsRNA *in vitro.* This was confirmed using the purified dsRNA (c). (d) R-loops and dsRNA were produced from the pFC53 plasmid. After transcription, the R-loop containing plasmids were cut with ApaLI to produce 2 bands, the longest of which contains R-loops. The smallest band acts as negative control because it does not carry R-loops. The cut plasmids and dsRNAs were then incubated alone or in combination with 4 μg of purified S9.6 antibody. The input (IN), corresponding to half of what was incubated with the antibody; the flow-through fraction (FT), corresponding to the material that did not bind to the antibody; the W4; and the immuno-precipitated fraction (IP) were then purified using phenol/chloroform extraction and ethanol precipitation and either run on an agarose gel (d) or analyzed by dot blot with the indicated antibodies (e). (f) The long fragment where R-loops can form (R-loop fragment) and the short fragment where R-loops do not form (control fragment) were also quantified by qPCR ($n = 4$ independent experiments, *P*-values with Wilcoxon–Mann–Whitney). The S9.6 antibody can immuno-precipitate GC-rich dsRNAs, and when dsRNAs are mixed with R-loops, the recovery of R-loops by the S9.6 antibody is reduced (see text). (g) Representative centromeric region of chromosome III showing RNase H-resistant but RNase III-sensitive DRIPc-seq signals at the *dg* and *dh* centromeric repeats. Note that most DRIPc-seq signals over tRNA genes are RNase H-sensitive, confirming that genuine R-loops form at tRNA genes [18].
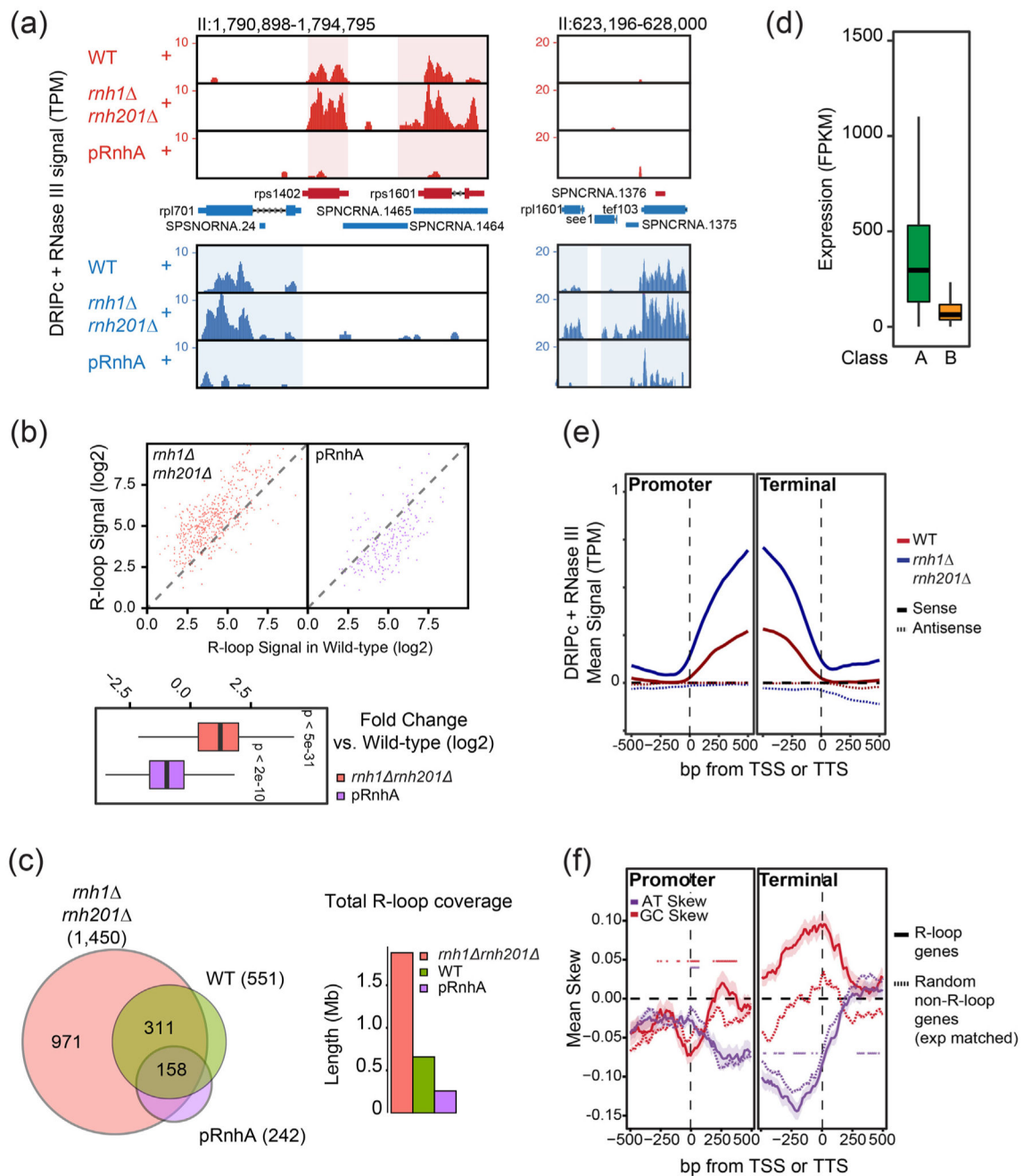
Fig. 3.
Improved R-loop mapping identifies two classes of R-loops in fission yeast. (a) Representative genomic regions showing DRIPc-seq signals after pre-treatment with RNase III. Color schemes are as in Fig. 1A. Data are shown for WT, *rnh1Δ rnh201Δ* and pRnhA cells. (b) Top: *X–Y* plot of RNase III-treated DRIPc-seq signal at each locus in WT (*x*-axis) *versus rnh1Δ rnh201Δ* or RnhA expression (*y*-axis). This confirms that R-loop signals are stronger in *rnh1Δ rnh201Δ* and weaker in pRnhA for the vast majority of R-loop-forming loci. Bottom: A box-plot representation of the same result. DRIPc-seq signals are significantly higher in *rnh1Δ rnh201Δ* and lower in pRnhA cells (*P*-values measured by

Wilcoxon–Mann–Whitney test). (c) Left: Venn diagram showing the number of genes with R-loop peaks and their overlap in WT, *rnh1 rnh201* and *pRnhA* conditions. Right: Bar-plot showing total DRIPc peak length in the indicated samples. (d) The wild-type transcript levels at genes with a class A or class B DRIPc-seq peak are represented as box-plots. Genes forming class B R-loops tend to be less expressed than genes forming class A R-loops. (e) Metaplots of RNase III-treated DRIPc-seq signals at promoter and terminator regions. (f) Metaplots of AT- and GC-skew at promoter and terminator regions of genes with class A R-loops and random non-R-loop-forming genes matched for gene expression and length.
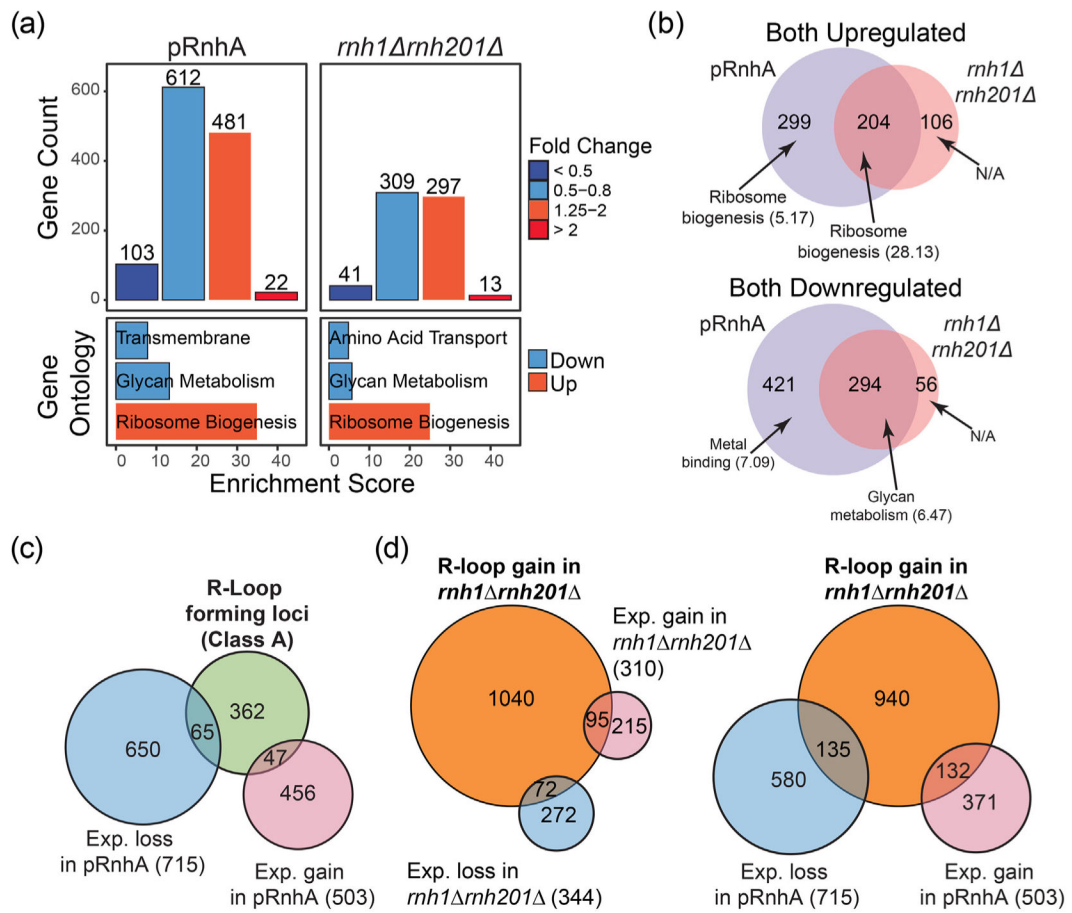
Fig. 4.
Manipulations of R-loop levels impact transcription mostly at genes that do not form detectable R-loops. (a) Top: Histogram showing the number of transcripts that were down-regulated and up-regulated after expression of RnhA (pRnhA) or deletion of both endogenous RNase H enzymes (*rnh1Δrnh201Δ*). Bottom: GO analysis of up- and down-regulated transcripts. (b) Venn diagram showing the overlap between the transcripts that are up- (top) and down-regulated (bottom) in *pRnhA* and *rnh1Δrnh201Δ*. GO term enrichment is indicated. (c) Venn diagram showing the overlap between the transcripts forming class A R-loops and the transcripts whose levels are affected upon RnhA expression. (d) Venn diagram showing the overlap between the transcripts forming R-loops in *rnh1Δrnh201Δ* and the transcripts whose levels are affected in *rnh1Δrnh201Δ* (left) or upon RnhA expression (right).