

# Immunoprecipitation of RNA–DNA hybrid interacting proteins in *Trypanosoma brucei* reveals conserved and novel activities, including in the control of surface antigen expression needed for immune evasion by antigenic variation

Mark J. Girasol<sup>1,2</sup>, Emma M. Briggs<sup>1,3,\*</sup>, Catarina A. Marques<sup>1</sup>, José M. Batista<sup>1</sup>, Dario Beraldi<sup>1</sup>, Richard Burchmore<sup>1</sup>, Leandro Lemgruber<sup>1</sup> and Richard McCulloch<sup>1</sup> <sup>1,\*</sup>

<sup>1</sup>University of Glasgow, College of Medical, Veterinary and Life Sciences, School of Infection and Immunity, Wellcome Centre for Integrative Parasitology, Glasgow, UK

<sup>2</sup>University of the Philippines Manila, College of Medicine, Manila, Philippines

<sup>3</sup>University of Edinburgh, Institute for Immunology and Infection Research, School of Biological Sciences, Edinburgh, UK

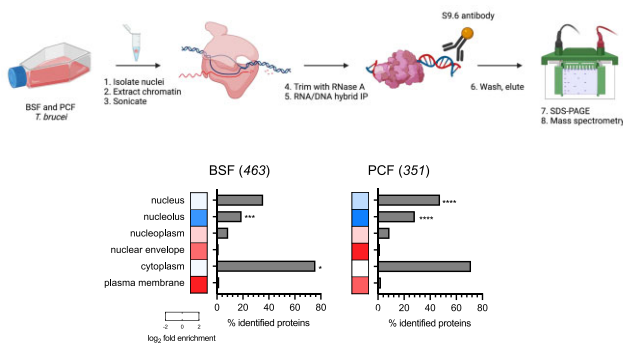
\*To whom correspondence should be addressed. Tel: +44 1413305946; Email: Richard.mcculloch@glasgow.ac.uk  
Correspondence may also be addressed to Emma M. Briggs. Email: Emma.Briggs@ed.ac.uk

## Abstract

RNA–DNA hybrids are epigenetic features of genomes that provide a diverse and growing range of activities. Understanding of these functions has been informed by characterising the proteins that interact with the hybrids, but all such analyses have so far focused on mammals, meaning it is unclear if a similar spectrum of RNA–DNA hybrid interactors is found in other eukaryotes. The African trypanosome is a single-cell eukaryotic parasite of the Discoba grouping and displays substantial divergence in several aspects of core biology from its mammalian host. Here, we show that DNA–RNA hybrid immunoprecipitation coupled with mass spectrometry recovers 602 putative interactors in *T. brucei* mammal- and insect-infective cells, some providing activities also found in mammals and some lineage-specific. We demonstrate that loss of three factors, two putative helicases and a RAD51 paralogue, alters *T. brucei* nuclear RNA–DNA hybrid and DNA damage levels. Moreover, loss of each factor affects the operation of the parasite immune survival mechanism of antigenic variation. Thus, our work reveals the broad range of activities contributed by RNA–DNA hybrids to *T. brucei* biology, including new functions in host immune evasion as well as activities likely fundamental to eukaryotic genome function.

## Graphical abstract

Identification of RNA–DNA hybrid interactors in *Trypanosoma brucei*



## Introduction

RNA–DNA hybrids are ubiquitous features of genomes in all domains of life. R-loops are a form of RNA–DNA hybrid in which an RNA molecule base-pairs with one strand of double-stranded DNA, causing displacement of a DNA single-strand.

Though initially found to form during transcription (1), R-loops are increasingly known to be widespread in genomes and to have wide impacts on genome function, both positive and negative (2–4). Such activities include initiation and arrest of DNA replication (5–7), transcription activation and

Received: May 26, 2023. Revised: September 14, 2023. Editorial Decision: September 15, 2023. Accepted: September 28, 2023

© The Author(s) 2023. Published by Oxford University Press on behalf of Nucleic Acids Research.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

termination (8), chromatin formation (9,10), and telomere function (11,12). R-loops can result in genome instability (1,13,14), at least in part by generating DNA breaks, which can be harmful (15), but can also be used functionally, such as during class switch recombination in mammalian B cells (16–18). RNA–DNA hybrids and R-loops can also contribute to DNA damage repair, including of DNA double-strand breaks (3,19,20).

A number of studies have recently begun to dissect the various activities provided by RNA–DNA hybrids and R-loops by characterising the proteins with which they interact. The earliest of these studies searched for RNA–DNA hybrid interactors in HeLa cells using immunoprecipitation with the S9.6 monoclonal antibody (21) and mass spectrometry (22), revealing hundreds of potential activities. A similar approach in mouse embryonic stem cells also revealed hundreds of putative RNA–DNA hybrid interactors (23). A distinct approach used two large, synthetic RNA–DNA hybrids and recovered >1000 proteins each from lysates of human B cells (24). Two more recent studies relied on proximity-labelling based on the DNA binding domain of RNase H1, identifying ~300–400 proteins in immortalised human cells (25,26). Finally, based on all these datasets, Kumer *et al.* (27) searched for common features of the recovered proteins and used machine learning to predict RNA–DNA hybrid interacting proteins across the human proteome. Together, these studies have revealed a wealth of potential RNA–DNA hybrid and R-loop associated activities. However, all these studies are limited to mammals, and no study has asked if similar or distinct activities are found in other eukaryotes. Here, we have adapted the DNA–RNA immunoprecipitation-mass spectrometry (DRIP-MS) approach of Cristini *et al.* (22) to explore the RNA–DNA hybrid interactome of the protozoan parasite, *Trypanosoma brucei*, where mapping of R-loops has predicted both conserved and diverged genomic activities (28).

The genome of *T. brucei* is unusual for a eukaryote in several respects (29,30). All but two of the ~8000 protein-coding genes in *T. brucei*'s genome 'core' (see below) are transcribed by RNA Polymerase (Pol) II from multigene clusters, each of which contains potentially hundreds of genes and is transcribed from a single, still only partly understood transcription start site (31–35). Unlike in bacteria, genes in such operon-like transcription units appear not be functionally related (36). This arrangement means that the genome core contains relatively limited content that is not traversed by RNA Pol-II, and mature mRNAs are generated from pre-mRNA transcripts by extensive, coupled trans-splicing and polyadenylation (37). In addition, each multigene transcription unit has a single transcription termination site, which contains a novel base, termed J, that acts to recruit a number of termination factors (38). This unusual organisation of gene expression appears to also reflect DNA replication organisation, since mapping sites of replication initiation, termed origins, reveals close overlap with transcription start and termination sites (39,40). Furthermore, RNAi against a subunit of the origin recognition complex, which defines origins (41), suggests functional interaction between the replication and transcription machineries (39). All the above aspects of gene expression appear conserved with the wider grouping of kinetoplastids, while other aspects of the *T. brucei* genome may be specific. Survival of African trypanosomes in the mammal relies on a process termed antigenic variation, which involves continuous changes in the identity of a surface expressed 'coat', which is

composed of a single Variant Surface Glycoprotein (VSG) in a single cell at a given time (42,43). Switching from one VSG coat to another in the mammal relies on both transcriptional changes between approximately 15 telomeric VSG transcription sites, termed bloodstream expression sites (BES), and recombination reactions that move silent VSG genes into the actively transcribed BES (43,44). Each BES is also a multigene transcription unit but is, remarkably, transcribed by RNA Pol-I from a promoter that shares some homology with those at rRNA gene clusters (45). Recombination relies on a huge archive of 1000s of silent VSGs, which are mainly found in arrays that occupy the chromosome subtelomeres (29,46,47). Each chromosome in *T. brucei* thus comprises a highly transcribed core and predominantly untranscribed subtelomeres, with chromosome conformation capture and ATAC-seq analyses indicating that the two genome compartments rarely interact and display differing levels of chromatin-mediated compaction (30).

Previous work has mapped R-loops in *T. brucei*, revealing that their localisation and potential functions span all the above aspects of the genome. DNA–RNA immunoprecipitation and sequencing (DRIP-seq) indicates that R-loops localise to the start and, to a lesser extent, end of the RNA Pol-II multigene clusters, as well as to intra-cluster regions of splicing and polyadenylation (28). The same approach showed that R-loops localise to the single centromere found in each chromosome. Mutation or RNAi has also been used to examine the impact of loss of either *T. brucei* RNase H1 or the A subunit of RNase H2, both of which are RNA–DNA endonucleases that remove RNA from RNA–DNA hybrids (48). Mutation of RNase H1 has no effect on growth, while RNAi of RNase H2A stops growth and results in DNA damage associated with RNA Pol-II transcription initiation; loss of either RNase H results in R-loop and DNA damage accumulation within the VSG BESs, as well as loss of VSG expression control and increased VSG switching (49–51). Loss or overexpression of RNase H1 or RNase H2A also alters levels of telomeric RNA–DNA hybrids (49,50,52,53), and such telomere-focused changes may also influence VSG expression and switching. To begin to explore this diverse range of RNA–DNA hybrid functions in *T. brucei*, we show here that DRIP-MS recovers a similarly large number of interacting proteins as is observed in mammalian cells. Amongst these putative interactors we can identify functions conserved in mammals, including ribosome- and mRNA-associated factors and helicases, as well as activities that may reflect the unusual *T. brucei* genome, including histone variants and centromere-binding kinetochore proteins. We provide functional analysis of three DRIP-MS factors and show that loss of any of them increases nuclear damage and RNA–DNA hybrid levels, as well as altering VSG expression. One of these factors is a RAD51-related protein previously described to act in VSG recombination (54,55), while the two others are putative helicases that have never been functionally examined in *T. brucei*.

## Materials and methods

### Trypanosome culture and genetic editing

*Trypanosoma brucei* MiTat1.2 (Lister 427) bloodstream forms were cultured in HMI-9 medium (Life Technologies) supplemented with 10% foetal calf serum (FCS) at 37°C,

5% CO<sub>2</sub>, and Lister 427 procyclic forms were cultured in SDM-79 medium with 10% FCS and 0.2% hemin at 27°C. Bloodstream forms capable of CRSIPR-Cas9 modification were generated by transfection with plasmid pJ1399 (56), containing T7 polymerase and Cas9. Expression of Cas9 was confirmed by RT-PCR (Supplementary Figure S4). For epitope tagging and allele knockout, primers were designed with LeishGedit (<http://www.leishgedit.net/>) and donor sequences were PCR amplified from the pPOT plasmid (containing mNeonGreen and blasticidin resistance gene) as previously described (57); PCR products were transfected along with a distinct PCR product to generate guide RNA, which was amplified using self-annealing primers derived also from LeishGedit. Mutant lines were generated by Amaxa transfection with ethanol-purified PCR products and drug-selection to generate clonal line, which were confirmed by PCR. Tagged proteins were confirmed by western blot using anti-mNeonGreen antibody (1:1000) and anti-EF1alpha (1:25000) as a loading control. For inducible RNAi, the BSF Lister 427 derivative 2T1 (58) was used. Target sequences and primers were designed with RNAi (59). PCR amplified fragments were cloned into the pRPaiSL vector plasmid (60), which was linearised and transfected to generate RNAi clones, confirmed by RT-PCR.

### RT-qPCR

RNA was extracted from  $8 \times 10^6$  cells using the RNeasy Mini Kit (Qiagen) protocol. Genomic DNA was digested on-column for 15 min at room temperature using the RNase-free DNase I Set (Qiagen). First-strand complementary DNA (cDNA) synthesis cDNA was generated from 500 ng total RNA using SuperScript™ IV First-Strand Synthesis System (Invitrogen), following the manufacturer's protocol using random hexamer primers. Previously used primer sequences were used for VSG RT-qPCR (39). Each primer pair target was run in two biological replicates and three technical replicates for each cell line. A 20 µl reaction contains 1X SYBR. Green PCR Master Mix (Applied Biosystems), 250 nM of forward and reverse primers and 1 µl cDNA. All qPCR experiments were run in 7500 Real-Time PCR system (Applied Biosystems) using the following cycling conditions: 1 cycle at 95°C for 10 min, followed by 40 cycles of 95°C for 15 s and 60°C for 1 min. Fluorescence intensity was measured at the end of each extension step (60°C for 1 min). For normalization across different samples, actin amplification was used as endogenous control. For calculation of relative mRNA levels, the 2- $\Delta\Delta C_t$  method was used (61).

### Fluorescence microscopy

For imaging mNeonGreen fluorescent proteins,  $\sim 8 \times 10^6$  parasites were pelleted by centrifugation at 800 rcf for 7 min and resuspended in FCS-free HMI-9 media with 1 µg/ml Hoechst 33342 (Sigma-Aldrich). Cells were pelleted again, resuspended in 0.05% (v/v) formaldehyde in FCS-free HMI-9 to immobilise parasite flagella and adhered to a Poly-L-Lysine coated slide for immediate imaging. For immunofluorescence analysis of DNA damage, washed parasites were fixed in 4% formaldehyde in vPBS for 10 min and permeabilised with 0.1% IGEPAL CA-630 for 10 min. Cells were washed in vPBS and adhered to a slide, before incubating in PSB + 1% (w/v) glycine 5 min and blocked with 1% bovine serum albumin (BSA) for 1 h. Staining was performed with  $\alpha$ - $\gamma$ H2A primary (1:1000) and  $\alpha$ -rabbit Alexa Fluor®

Plus 488 (1:1000) secondary antibodies diluted in 1% BSA. Slides were washed 1× PBS before mounting with 5 µl DAPI Fluoromount-G® (Southern Biotech) and imaging. For imaging DNA–RNA hybrids, parasites were instead fixed in 70% ice-cold methanol for 1 h, permeabilised with 0.5% v/v Triton X-100 for 10 min and blocked with 1× PBS, 0.01% v/v Tween-20, 0.1% w/v BSA for 1 h at 37°C. S9.6 (Kerafast) primary (1:1000) and Alexa Fluor Plus 488 Goat anti-Mouse IgG (H + L) (ThermoFisher) secondary (1:3000) antibodies were diluted in blocking solution for staining in suspension while shaking before adhering to slides. VSG immunofluorescence analysis was performed following the protocol of Glover *et al.* (62). Briefly, formaldehyde fixed parasites were adhered to glass slides and blocked with 50% FCS in PBS for 45 min before staining with primary anti-VSG (1:10000) and secondary Alexa Fluor (1:1000) antibodies and mounting with DAPI Fluoromount-G® (Southern Biotech). Imaging was performed with an Axioscope 2 widefield fluorescence microscope (Zeiss) using a 63×/1.40 oil objective, or a Leica DiM8 widefield fluorescence microscope to acquire Z-stacks of 5 µm thickness in 25 sections. The images were later processed on Fiji/ImageJ (63) using the same parameters ([http://imagej.net/Rolling\\_Ball\\_Background\\_Subtraction](http://imagej.net/Rolling_Ball_Background_Subtraction)). For fluorescence intensity quantification, all images were obtained using the same exposure times and were later processed on ImageJ using the same parameters. After image processing, a circular 21 × 21 pixel region of interest (ROI) was drawn around each nucleus, and the mean pixel intensity per nucleus was plotted into a vertical scatter plot using Prism 9 (GraphPad).

### DRIP-MS

DRIP-MS approach was adapted from Cristini *et al.* (22).  $2.5 \times 10^8$  mid-log growth phase parasites were pelleted and washed in 1 ml of 1× PBS, before incubation in 1 ml of lysis buffer (80 mM KCl, 5 mM PIPES pH 8.0, 0.5% Nonidet P-40 substitute) on ice for 20 min. The pellet was homogenized using a tight-fitting dounce homogenizer until the nuclei were released, as checked with microscopy. Nuclei were enriched by centrifugation (2400 rcf for 10 min, 4°C) and resuspended in 200 µl of RSB sonication buffer [1× RSB buffer (10 mM Tris–HCl pH 7.5, 200 mM NaCl, 2.5 mM MgCl<sub>2</sub>), 0.2% sodium deoxycholate, 0.1% SDS, 0.05% sodium lauryl sarcosinate, 0.5% Triton X-100] before sonicating for 10 min in the Diagenode Bioruptor using cycles of 30 s ON and 90 s OFF. The sonicated samples were then diluted with 600 µl RSB + T buffer (1× RSB buffer, 0.5% Triton X-100) plus 0.8 ng RNase A to degrade single-stranded RNA. For benzonase control IPs, 1 U/µl benzonase was added and incubated for 30 min at 37°C. 20 µl was reserved as input control samples. Pre-prepared Protein A Dynabeads (incubated with 5 µg of S9.6 antibody (Millipore) in 1× PBS plus 0.5% BSA overnight at 4°C), were washed in 1× PBS + 0.5% BSA and suspended in the sonicated sample. Samples were incubated on a rotor at 4°C for 2 h to bind DNA–RNA hybrids and interacting proteins. Beads were then wash in 1 ml cold RSB + T buffer (4 times) and 1 ml cold RSB buffer (twice), before collecting by centrifugation at 1000 rcf for 3 min at 4°C. Beads were retained using a magnetic rack and supernatant was discarded. Beads were suspended in 15 µl loading buffer (1× NuPAGE LDS loading buffer, 1× PBS, 0.1%  $\beta$ -mercaptoethanol) for 10 min at 70°C. Beads were pelleted and supernatant, containing eluted proteins collected.

Proteins were resolved using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and stained using InstantBlue® Coomassie Protein Stain (Abcam). Gel segments only found in S9.6 IP samples and not benzonase controls were excised for Nanoflow HPLC Electrospray Tandem Mass Spectrometry (nLC-ESI-MS/MS), performed by Proteomics Facility of Glasgow Polyomics. Briefly, the gel pieces were de-stained by incubation in 500 µl of 50% acetonitrile in 100 mM ammonium bicarbonate for 30 min on a shaker, dehydrated by incubation in acetonitrile for 10 min then dried a vacuum centrifuge. Trypsin was added to rehydrate the gel pieces and digested overnight at 37°C, before gel pieces were pelleted. The supernatant containing peptides was dried in a vacuum centrifuge and solubilized in 20 µl 5% acetonitrile with 0.5% formic acid.

Trypsinized peptide samples were analyzed using nanoscale liquid chromatography coupled to electrospray ionization tandem mass spectrometry (nLC-ESI-MS/MS). Online detection of peptide ion was by electrospray ionization mass spectrometry using an Orbitrap Elite MS (Thermo Scientific). Peptides were separated on a Pepmap™ C18 reversed phase column (3 µm, 100 Å, 75 µm × 50 cm) (Thermo Scientific). Samples were fractionated with mobile phase A consisting of 0.1% (v/v) formic acid in water and mobile phase B consisting of acetonitrile (80% v/v) and water (20% v/v). The peptide separation was performed at a fixed solvent flow rate of 0.3 µl/min, using a gradient of 4–100% mobile phase B over 120 min. The Orbitrap Elite™ MS acquired full-scan spectra in the mass range of  $m/z$  300–2000 Da for a high-resolution precursor scan at a set mass resolving power of 60 000 (at 400  $m/z$ ). Collision-induced dissociation was performed in the linear ion trap with the 20 most abundant precursors using rapid scan mode.

Proteins were identified using the Mascot search engine (v2.6.2, Matrix Science) by interrogating MS data against protein databases of *Trypanosoma brucei* Lister 427. A mass tolerance of 10 ppm was allowed for the precursor and 0.3 Da for MS/MS matching, with the following search parameters: trypsin enzyme specificity, allowing one missed cleavage; cysteine carbamidomethylation was selected as fixed modification, while N-terminal carbamidomethylation, asparagine-glutamine deamidation, tyrosine iodination, and methionine oxidation were selected as variable modifications. Proteins within a significance score of  $P < 0.05$  and with at least one unique peptide were considered. Four and two independent biological replicates were generated for BSF and PCF *T. brucei*, respectively. Corresponding emPAI values were compared between IP and benzonase controls; proteins with mean  $\log_2$  fold-change  $> 0$  across all biological replicates constituted the RNA/DNA hybrid interactome for each form. BSF-specific interactome was defined as proteins with  $> 0 \log_2$  fold-change in BSF versus PCF.

### Homology searches, domain analysis and phylogenetics

To look for putative homologs of Tb927.3.2600 and Tb927.3.5440 in *Discoba* and elsewhere, sequence- and HMM-based similarity searches were done in HMMER (64) and HHpred (65,66). KEGG Orthology (KO) scores, which guided our searches for homologs and phylogenetic analyses, were obtained with KofamKOALA (67). To determine the Pfam domain organisation of every analysed protein, se-

quences were run through InterProScan (68). Helicase ‘core’ regions were defined according to Fairman-Williams *et al.* (69). The regions of sequence similarity shown in Figure 2, were obtained from HHpred searches using Tb927.3.2600 and Tb927.3.5440 as queries against HMM databases representing the proteomes of *Drosophila melanogaster* and/or *Homo sapiens*; for every protein pair that we considered, we defined their ‘region of similarity’ as that corresponding to their whole alignment as retrieved by HHpred. For the phylogenetic analysis of Tb927.3.2600, we first collected putative Tb927.3.2600 homologs from across *Discoba* (searches were done in publicly available genomic and transcriptomic datasets), as well as representative orthologs of Tb927.3.2600’s top 4 hits from KofamKOALA, namely DDX60 (KO entry: K20103), MTR4 (K12598), ISE2 (K26077), and Hely (K03727), all of which are members of the Ski2-like family of SF2 helicases (69–71). We then used MAFFT-DASH L-INS-i (72,73) to align each of the 5 groups of protein sequences (Tb927.3.2600 and its putative homologs in *Discoba*; DDX60; MTR4; ISE2; and Hely) separately; in each case, alignment was followed by manual trimming in Jalview (74), down to the helicase ‘core’ plus short flanking regions (approximately 50–100 aa, depending on sequence conservation within each group). All resulting subsequences were then aligned with MAFFT-DASH L-INS-i, and the obtained alignment trimmed with trimAl (75) using a column gap threshold of 10% (‘gt -0.1’ option). This yielded a ‘final alignment’, which was used as input in an IQ-TREE (76) run to infer a maximum-likelihood phylogenetic tree; the best-fit evolutionary model was selected by ModelFinder (in addition to default ‘standard’ models, several mixture models, e.g. EX\_EHO, were included in ModelFinder’s search), and both SH-aLRT (77) and UFBoot2 (78) were performed with 1000 replicates each. For the phylogenetic analysis of Tb927.3.5440, a similar protocol was followed. Tb927.3.5440’s top 4 hits from KofamKOALA were ATRX (K10779), ARIP4 (K10876), RAD54 (K10875) and RAD54B (K10877), all of which are members of the Rad54-like grouping within the Snf2 family of SF2 helicases (79). KO’s RAD54 entry confusingly includes a mixture of RAD54 and DRD1 orthologs, which we were able to identify in preliminary trees; because DRD1 is also part of the Rad54-like grouping, we included it in our analysis. Phylogenetic trees were rooted with MAD (80), and visualized and edited in iTOL (81).

## Results and discussion

### Identifying the *T. brucei* RNA–DNA hybrid interactome

RNA–DNA hybrids and associated proteins were enriched through immunoprecipitation (DRIP) using the S9.6 antibody, which recognizes RNA–DNA hybrids at an affinity as low as 0.6 nM (82). Since S9.6 can also recognise double-stranded RNA (83), nuclei were first enriched using homogenization after mild lysis and centrifugation. The released nuclei were then sonicated to fragment the genomic DNA and limit co-precipitation of non-specific proteins during RNA–DNA hybrid recovery. DRIP was then performed on native chromatin, without cross-linking, in the presence of RNase A to deplete nuclear RNA not associated with DNA, and thereby limit DRIP of RNA binding proteins. In addition, half of the

nuclear extracts were treated with benzonase prior to DRIP. Benzonase digests all forms of nucleic acids (84), including RNA/DNA hybrids, and so these DRIP samples provided controls for recovery of non-specific proteins by S9.6. DRIPs, plus benzonase controls, were performed for both *T. brucei* procyclic forms (infective to the insect vector; two biological replicates) and bloodstream forms (mammal-infective; four replicates) grown in culture. Figure S1A shows the results of such DRIPs after gel separation, revealing that treatment with benzonase caused loss of many detectable DRIP proteins from both PCF and BSF cells. To characterize the recovered proteins, matched sections of gels for both benzonase-treated and non-treated DRIP samples from all independent replicates were analysed by mass spectrometry (MS).

To identify putative RNA–DNA hybrid interacting proteins from the label-free DRIP-MS data, empAI values of all proteins were compared, in each DRIP sample, to protein empAI values in their cognate benzonase-treated control DRIP samples and log<sub>2</sub> enrichment determined. In total, 602 proteins were enriched in at least one DRIP-MS replicate relative to the benzonase controls: 463 and 351 putative RNA–DNA hybrid-interacting proteins in BSF and PCF cells, respectively (Figure 1, Table S1). To ask if the DRIP was selective, fold BSF/PCF enrichment of the 602 DRIP-MS enriched proteins was compared with whole-cell proteomic data from Butter *et al.* (85) (Supplementary Figure S1B). No correlation in BSF/PCF enrichment ratios was seen between these two studies, indicating DRIP enriched for a non-random set of *T. brucei* proteins. Cellular compartment gene ontology (GO) term enrichment analysis, which has been greatly aided by epitope tag-mediated evaluation of the subcellular localisation of most *T. brucei* proteins (86), was next performed on the putative RNA–DNA hybrid interactors. Consistent with nuclear enrichment prior to DRIP, and with the likelihood that most RNA–DNA hybrids will form on the nuclear genome, predicted nuclear and nucleolar proteins showed significant GO term enrichment in the PCF data, and nucleolar proteins were significantly enriched in the BSF data (Figure 1A). In contrast, nuclear envelope and plasma membrane proteins were under-represented. Though cytoplasmic proteins were not enriched (Figure 1A), they represented the majority of proteins recovered from each life cycle stage, perhaps suggesting recovery of proteins that are not RNA–DNA hybrid interactors; for instance, ribosomal proteins were prominent in the DRIP-MS dataset (Supplementary Table S1), possibly due to S9.6 recognition of double-stranded RNA. However, the enrichment of ribosomal proteins may also reflect growing evidence for roles of R-loops in rRNA transcription and processing, leading to ribosome biogenesis (87,88), and is consistent with machine-learning predictions of human RNA–DNA hybrid interacting protein types (27). Moreover, recent work has shown R-loops formed in the nucleus can be found in the cytoplasm of human cells (89), where it is possible that they interact with cytoplasmic proteins. Indeed, below we provide evidence that at least one *T. brucei* RNA–DNA hybrid interactor predominantly localises to the cytoplasm but can provide nuclear activities.

### Conserved eukaryotic RNA–DNA hybrid-associated functions predicted by DRIP-MS

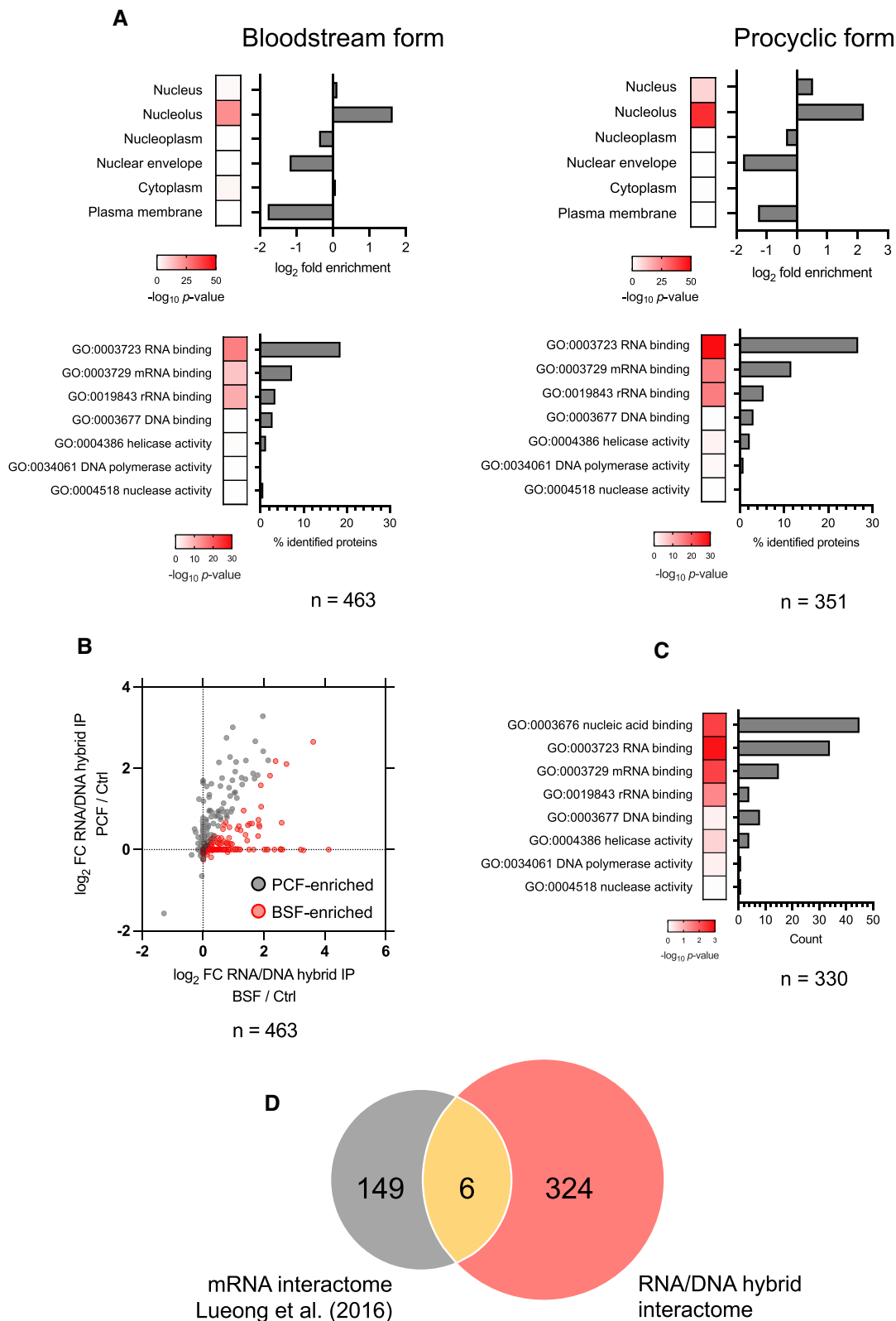
Next, molecular function GO term analysis was used to predict activities of the proteins enriched by DRIP (Figure 1A). Multiple proteins with predicted nucleic acid-binding activi-

ties were enriched, mirroring previous interactome data from mammalian cells (22) and consistent with selective S9.6 immunoprecipitation. RNA-binding proteins showed the most significant enrichment in both the BSF and PCF DRIP interactomes. Pronounced enrichment of mRNA-binding proteins is a common feature of our data and that of mammalian HeLa cells (22), and here might reflect the ubiquitous localisation of R-loops with pre-mRNA processing regions in RNA Pol-II polycistrons (28). Enrichment of mRNA binding proteins was non-random (see below) and may suggest the hybrids provide activities linked to mRNA maturation, such as provided by the numerous predicted helicases recovered (Supplementary Table S1). rRNA-binding proteins were also enriched, as also seen by Wu *et al.* in mouse cells (23), providing further evidence of ribosome-associated R-loop functions (Figure. 1A).

A number of classes of enriched factors provide evidence for RNA–DNA hybrids acting during gene expression in *T. brucei*. All four known *T. brucei* histone variants were enriched in the DRIP-MS data (Supplementary Table S1): H2A.Z and H2B.V, which localise at transcription initiation regions, and H4.V and H3.V, which localise to termination regions (33). Thus, the DRIP-MS data reinforces previous *T. brucei* DRIP-seq mapping that showed enrichment of RNA–DNA hybrids at transcription start sites and, to a lesser extent, termination sites, indicating R-loop roles in gene expression organisation (28). In mammalian cells, RNA–DNA hybrids were found to interact with histone H3 (22), and the hybrids are known to recruit histone H3 modifications at promotor regions (90). R-loops have not been mapped in further trypanosomatids but such roles may be conserved, as the histone variants show the same localisations (91,92). How the deposition and functions of trypanosome R-loops and histone variants intersect is unknown, and the potential contribution of the hybrids to transcription remains unclear. However, although the DRIP did not enrich for RNase H1 or RNase H2 (in common with Crisitini *et al.* (22)), loss of RNase H2A causes pronounced DNA damage accumulation at transcription start sites (49). In addition, immunoprecipitation of the histone methyltransferases DOT1A or DOT1B reveals interaction with RNase H2 (32,51), with DOT1A-RNase H2 activity potentially resolving R-loops at transcription termination sites containing RNA Pol-III genes (32). We did not detect either DOT1A or DOT1B in the DRIP-MS data, perhaps suggesting the methyltransferases do not directly interact with RNA–DNA hybrids.

Adding to the above gene expression-associated activities, two of four *T. brucei* (93) Alba proteins were enriched: ALBA3 and ALBA4 (Supplementary Table S1). Alba proteins are found in both archaea and eukaryotes and bind DNA and/or RNA (94). Studies to date have shown *T. brucei* ALBAs to be cytoplasmic and possess RNA binding roles in translation (95–97), and so interaction with RNA–DNA hybrids may be surprising. However, in *Arabidopsis thaliana* ALBA1 binds RNA–DNA hybrids and ALBA2 interacts with the displaced single-stranded DNA, and together acts as an R-loop reader complex (98). Additionally, the four ALBA proteins found in *Plasmodium falciparum* bind both DNA and RNA (99), and so the prediction of RNA–DNA hybrid interaction for *T. brucei* ALBA proteins may indicate previously unknown activities.

DNA-binding proteins showed only weak evidence for enrichment (Figure 1A), unlike the pronounced enrichment seen in mammal DRIP-MS data (22). However, this grouping included some notable factors when specific functions were



**Figure 1.** Characterisation of the *T. brucei* RNA–DNA hybrid interactome. **(A)** The top two panels show cellular compartment GO term analysis of proteins recovered by RNA–DNA hybrid immunoprecipitation and mass spectrometry (DRIP–MS) from bloodstream (BSF) and procyclic form cells (PCF), while the lower panels show molecular function GO term analysis of the same datasets. For cellular compartment analysis, fold enrichment relative to the proteome is indicated, while for molecular function analysis the different categories are shown as percentage of total recovered proteins (n); for both, Bonferroni-adjusted p-values are shown as heatmaps. **(B)** Scatter plot of  $\log_2$ -transformed mean emPAI values of BSF and PCF DRIP–MS proteins relative to benzonase controls, with those proteins enriched in BSF cells shown in red and those enriched in PCF cells shown in grey. **(C)** Molecular function GO term analysis of BSF-enriched proteins (details as in A). **(D)** Comparison of the BSF RNA/DNA hybrid interactome and mRNA interactome obtained from a study by Lueong *et al.* (163).

examined, suggesting roles for *T. brucei* RNA–DNA hybrid interactors in chromosome functions. Previously, we described localisation of R-loops at *T. brucei* centromeres using DRIP-seq (28). Consistent with such localisation, three kinetochore proteins (100) were enriched by DRIP: KKT1, KKT3 and KKT19 (Supplementary Table S1). Though R-loops have also been shown to localise to centromeres in yeast (101), mammals (102) and plants (103,104), it is unclear how they might contribute to, or indeed impede, centromere function. Indeed, most trypanosomatid kinetochore subunits show no evidence of orthology with kinetochores in other eukaryotes (100,105). Intriguingly, *T. brucei* KKT3 is thought to be one of two kinetochore proteins that are positioned most proximal to the centromere (105), where they localise throughout the cell cycle (100). In addition, they contain novel domains for centromere association (106) and recruit KKT1 to assemble the kinetochore (107). DRIP recovery of KKT3 and KKT1 may then suggest that *T. brucei* centromeric R-loops help maintain and guide interaction of the kinetochore to the centromere throughout the cell cycle. Alternatively, R-loops may provide epigenetic definition of the *T. brucei* centromere in the absence of the histone H3 variant, CENP-A (108). SMC (structural maintenance of chromosome) proteins are ATPases found in all domains of life and are core subunits of larger protein complexes needed to organise the genome through conformational change (109). Two of these complexes in eukaryotes are cohesin, which contains Smc1 and Smc3, and condensin, containing Smc2 and Smc4. Here, DRIP enriched for both *T. brucei* SMC2 and SMC4 (Supplementary Table S1). Unlike for cohesin (110,111), no work to date has described condensin functions in any kinetoplastid, and so the prediction of R-loop interaction may provide a route to examine where and how the complex acts in *T. brucei*, where chromosome condensation during mitosis appears to be minimal (112).

Beyond the above protein cohorts, DRIP-MS implicated a number of further, less easy to predict RNA–DNA hybrid-associated activities. For instance, several protein kinases were enriched (Supplementary Table S1), including three NEK family kinases (113), which have diverse roles including in cell cycle control and DNA damage repair (114). Though mitochondrial proteins are likely to be under-represented in our approach, several kinetoplast proteins were enriched (Supplementary Table S1), including two DNA Pols (IC and Beta-PAK) (115,116) and a putative RNA-editing nuclease. Finally, nearly a third of the proteins (180) enriched by DRIP are annotated as hypothetical or hypothetical conserved (Supplementary Table S1), and so no functions can be predicted currently.

### Searching for RNA–DNA hybrid interactor roles in *T. brucei* antigenic variation

Amongst the wealth of potential RNA–DNA hybrid interactors revealed by DRIP-MS, we decided to ask if activities associated with antigenic variation could be identified, since R-loops are involved in the pathway in ways that are not yet clear (117,118). In addition, we reasoned that some R-loop activities that act in antigenic variation could be unique to African trypanosomes and might therefore yield a means to impede this crucial survival mechanism (43).

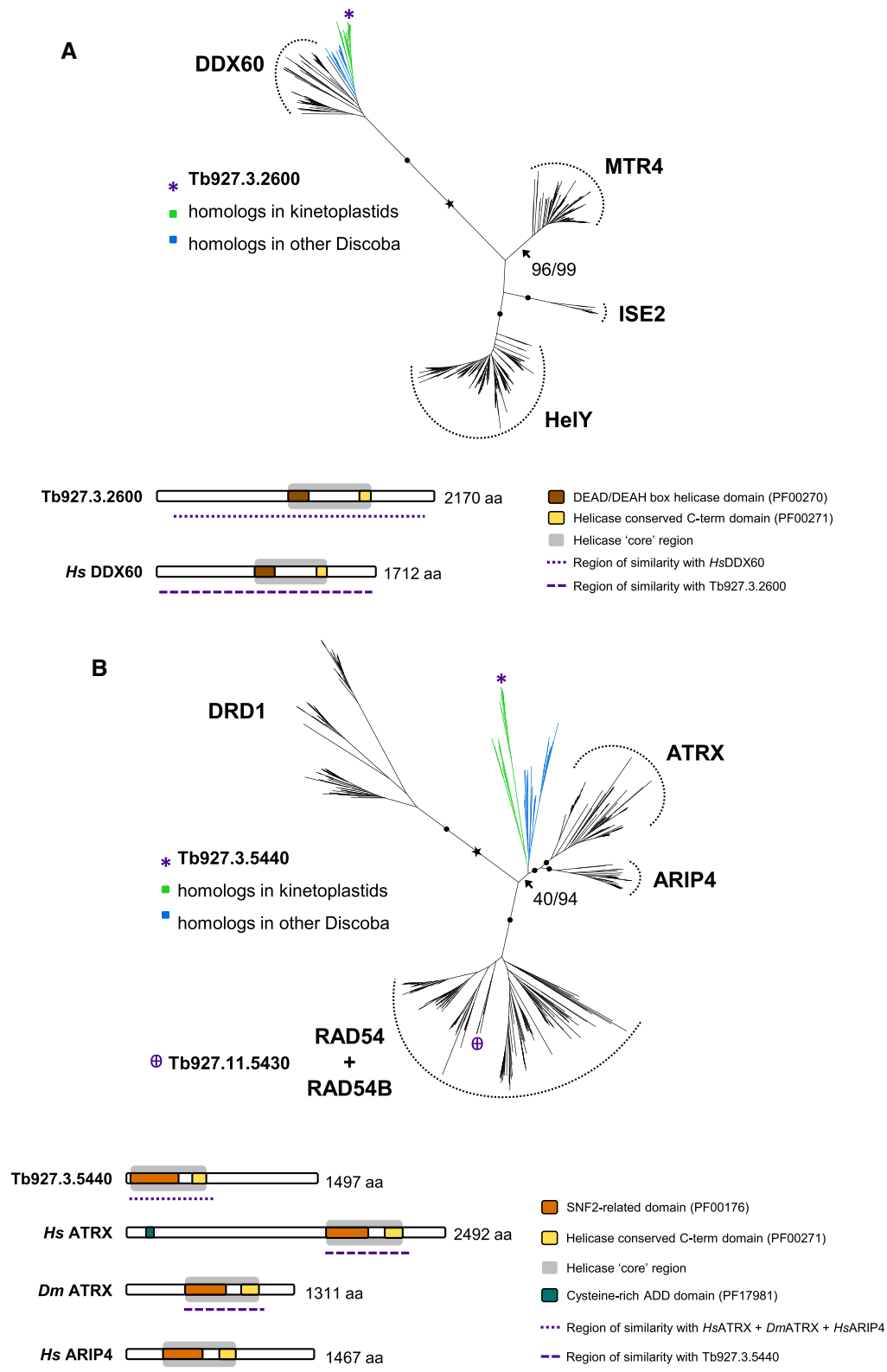
Despite DRIP-MS not detecting direct RNase H2A or DOT1B interaction with RNA–DNA hybrids, there is clear evidence of functional interaction between these proteins and

their involvement in antigenic variation: loss of DOT1B or RNase H2A results in similar changes in R-loop levels, DNA damage levels and Variant Surface Glycoprotein (VSG) expression alterations (49,51). In fact, a connection between R-loops and VSG expression may also extend to the histone variants recovered by DRIP-MS, given similarities in changes to VSG transcription control after loss of either RNase H (49,50) or H3.V and/or H4.V (30,119,120). Intriguingly, DRIP-MS data predicted VEX1 as a hybrid interactor, perhaps suggesting even more widespread roles for R-loops in VSG expression (Supplementary Table S1), consistent with the similarity in de-repression of silent VSG expression sites after loss or overexpression of VEX1 (62), in RNase H1 mutants (50), and after RNase H2A RNAi (49,51). One explanation for DRIP enrichment of VEX1 may be that the protein localises within or proximal to the telomeres of VSG expression sites (121), where RNA–DNA hybrids are present (49,50,52). We did not detect the other components of the VEX complex (121), such as VEX2, however, and so these data may suggest a specific R-loop role of VEX1 (122).

To ask if other RNA–DNA interacting proteins functionally link R-loops and antigenic variation, including by previously unknown mechanisms, we first attempted to compare the abundance of DRIP enriched proteins in BSF and PCF cells, asking which are more prevalent in the former, since VSG is not expressed in the latter (Figure 1B). A total of 330 putative interactors were found to have a greater level of enrichment in BSF cells than PCF relative to benzonase controls. Molecular function GO term analysis of this subset did not reveal any clear difference to all 602 proteins recovered (Figure 1C). Nonetheless, we were now more able to compare the RNA/DNA hybrid interactome with the BSF mRNA proteome generated by Lueong et al (123) (Figure 1D), since the GO term ‘mRNA binding’ was consistently enriched in all DRIP-MS analyses (Figure 1A, C). Overlap between the two proteomic datasets was very limited (Figure 1D), further showing that S9.6 immunoprecipitation recovers a non-random selection of *T. brucei* proteins.

Given that comparing predicted RNA–DNA hybrid interactomes between two life cycle stages did not yield any obvious difference in enrichment patterns, we decided to narrow the search based on two criteria: looking for proteins with annotations of relevant predicted activity; and DRIP-MS indication of recovery only from BSF cells. Amongst the proteins that fulfilled these criteria, four were selected for further analysis. The first two proteins were chosen because they are known to act in *T. brucei* VSG switching: RAD51 (Tb927.11.8190) (124–126) and the RAD51 paralogue, RAD51-3 (Tb927.11.2550) (54,55). RAD51 from yeast and mammals has previously been shown to bind RNA–DNA hybrids (127,128), and we describe the *T. brucei* RAD51-directed connection between R-loops and VSG switching elsewhere (Girasol *et al.*, BioRxiv). RAD51 paralogues are related to RAD51 and provide a range of activities in DNA damage repair and replication (129,130), but no work to date has suggested interaction with RNA–DNA hybrids. The two other proteins, encoded by Tb927.3.2600 and Tb927.3.5440 (Figure 2), were chosen as they provide relevant predicted functions (below) that have not been experimentally examined in *T. brucei* to date.

Tb927.3.2600 encodes one of a number of putative ATP-dependent DExD-box RNA helicases (DDXs) recovered by DRIP (Figure 2A, Table S1). DDXs are one grouping of helicases within a larger superfamily (SF2) (69) and, in other





eukaryotes, a number of DDXs have been shown to act on R-loops: DDX21, DDX23, DDX38B and DDX41 each limit DNA damage during transcription (25,131–133); DDX19 is a nucleopore-associated factor that can translocate to the nucleus and act with the kinase ATR during transcription-replication clashes (134,135); DDX1 contributes to immunoglobulin class switch recombination (136); and DDX1, DDX5 and DDX18 act on R-loops associated with DNA damage, including through interaction with DNA repair factors (137–140). Many of these helicases were enriched by DRIP-MS from mouse cells (23). Homology searches and phylogenetic analyses strongly suggest the *T. brucei* Tb927.3.2600-encoded DDX to be a homologue of DDX60 (Figure 2A) and a member of the Ski2 helicase family (69), which has not been implicated in R-loop functions in any eukaryote and has only been functionally characterised in *T. brucei* through its putative interaction with the stress response mRNA binding factor, MKT1 (141,142). The syntenically encoded protein in *T. cruzi* has recently been shown to be part of 43S preinitiation complex of the assembling ribosome (143), an association not described in other eukaryotes, including mammals, where DDX60 is non-essential and has been instead implicated in antiviral activities and cancer (144,145). The novelty of *T. cruzi* DDX60 interaction with the ribosome appears to be reflected in structural features not found in its mammalian orthologues (143).

Tb927.3.5440 has been annotated (tritypdb.org) as encoding a putative SNF2 DNA repair protein, merely suggesting that it belongs to the large Snf2 family of helicases (69) whose members provide a wide range of activities (146,147), including chromatin remodelling, transcription and DNA repair. Homology searches and phylogenetic analyses suggest the protein encoded by Tb927.3.5440 belongs to a somewhat distinct Discoba grouping that is most closely related to ATRX, which is widely distributed in eukaryotes, and ARIP4, which appears limited to animals (Figure 2B). No function has been ascribed to *T. brucei* ATRX or its relative, RAD54/B (encoded by Tb927.11.5430; Figure 2B), which was not detected in the DRIP-MS data (Supplementary Table S1). However, ATRX in other eukaryotes has been shown to have roles in alternative lengthening of telomeres (148–150), in homologous recombination pathway selection (151,152), and in suppression of R-loops in transcribed telomeres (26,153,154). Many of such roles could be consistent with RNA–DNA hybrid functions at the intersection of transcription and recombination in telomeric BESs during VSG switching in *T. brucei*.

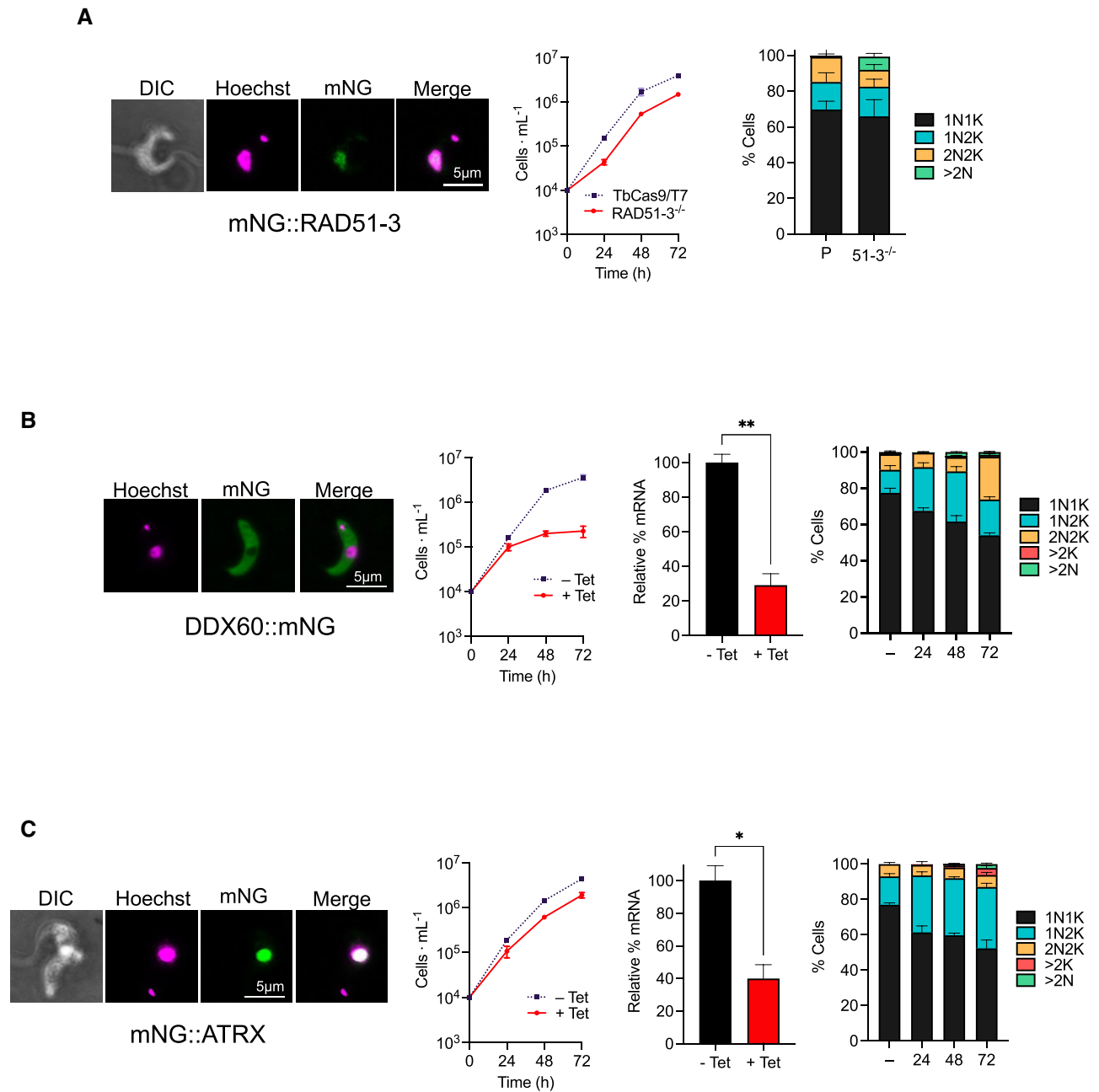
### Loss of RAD51-3, DDX60 or ATRX leads to increased nuclear DNA damage

To begin to test the functions of the three predicted RNA–DNA interactors, we engineered MiTat1.2 BSF *T. brucei* cells to permit genetic modification via CRISPR–Cas9 (56). Using these cells, we generated variant parasites expressing each protein as a translational fusion with mNeonGreen (mNG) and, in addition, attempted to make null mutants by replacing each allelic ORF with antibiotic resistance markers.

Both alleles of all genes were successfully tagged with mNG: RAD51-3 and ATRX at the N-terminus, and DDX60 at the C-terminus (Figures 3 and S2). In each case PCR showed integration of *BSD* and *NEO* constructs, with concomitant loss of the wild type untagged allele, and western blotting using anti-mNG antibody showed expression of

fusion proteins of the expected size (Supplementary Figure S2). Tagging of RAD51-3 or DDX60 did not impair parasite growth in culture (Supplementary Figure S2A,B), whereas mNG appeared to at least partially impede ATRX protein function, since the tagged cells exhibited a growth defect compared with parental TbCas9/T7 cells (Supplementary Figure S2C). Live fluorescence microscopy revealed nuclear localization of mNG::RAD51-3 and mNG::ATRX (Figure 3A,C) in all cell cycle stages (Supplementary Figure S2A,C). Fluorescence signal for mNG::RAD51-3 appeared more focal in cells with 1N1K and 1N2K compositions of nucleus (N) and kinetoplast (K; elongated, eK) staining (see Figures 3 and S2 for explanation), perhaps indicating recruitment to subnuclear loci during DNA replication (Supplementary Figure S2A). In all cell cycle stages DDX60::mNG signal was cytoplasmic (Figures 3B and S2B), perhaps consistent with a ribosomal function (143). Localisation of each protein in BSF cells essentially matches what is seen in PCF cells (86), suggesting conserved roles in at least these two life cycle stages.

A *RAD51-3* null mutant (*RAD51-3*–/–) was generated in a single round of transfection, with PCR demonstrating replacement of both WT alleles with *BSD* and *NEO*, and RT-PCR showing loss of *RAD51-3* transcript (Supplementary Figure S3A). CRISPR-mediated deletion of *RAD51-3* confirms previous observations that the paralogue is not essential in *T. brucei* (54), which may differ from *Leishmania* (155,156). The absence of *RAD51-3* did result in a growth defect, however (Figure 3A), which may be explained by the *RAD51-3*–/– mutants showing an accumulation of cells with more than two nuclei (>2N;  $7.4 \pm 1.0\%$  in –/–,  $0.8 \pm 0.5\%$  in parental) and a reduction in the proportion of 2N2K cells ( $9.4 \pm 1.7\%$  in –/–,  $14 \pm 2.9\%$  in parental), suggesting a mitotic defect. Attempts to make null mutants of *ATRX* or *DDX60* failed. For *DDX60*, double antibiotic-resistance transformant clones were recovered, but all retained an intact ORF and displayed improper integration of *BSD* (Supplementary Figure S3B). Attempts to remove even just a single *ATRX* allele failed to yield viable antibiotic-resistant clones. To examine functions, we instead used tetracycline-inducible RNAi (157). For both genes a growth defect emerged from 24 h after RNAi induction, though this was more pronounced for *DDX60* than *ATRX* (Figure 3B, C). Concomitant with the growth defects emerging, DAPI staining revealed perturbation in the DNA content of cells within the populations. For *ATRX*, the most pronounced change was an increase in 1N2K cells (32% in induced, 16% in uninduced at 24 h; Figure 3C), with an associated reduction in 1N1K cells, suggesting a stall in S/G2 phase. After 72 h some cells harbouring more than two kinetoplasts could be seen (4% in induced, 0% in uninduced), indicating kDNA replication and division can occur to at least some extent. The cell cycle defect after *DDX60* RNAi was distinct (Figure 3B), with loss of 1N1K cells associated initially with accumulation of 1N2K cells (24–48 h post-induction) and later by an increase in the proportion of 2N2K cells (72 h). These effects may be explained by death of S/G2-stalled cells from 24–48 h, as there was little increase in cell density at these time points (Figure 3B), or by an S/G2 stall caused by *DDX60* loss that is not absolute, with some cells progressing into but not through mitosis. The cell cycle effects described by DAPI staining after RNAi against *DDX60* and *ATRX* appear consistent with recent analysis of cell cycle perturbations using genome-wide RNAi and characterisation of gene-specific

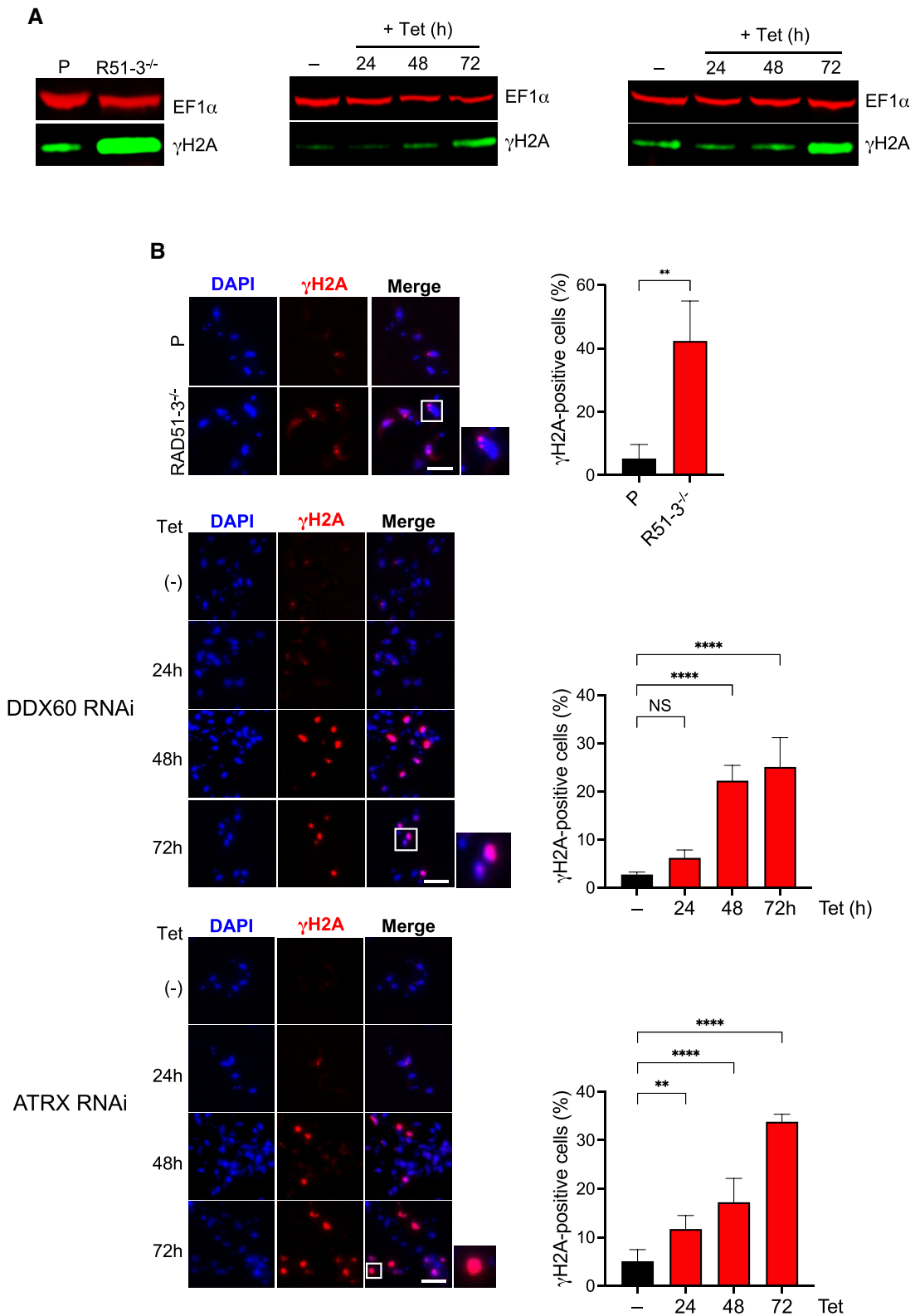


**Figure 3.** Functional characterisation of *T. brucei* RAD51-3, ATRX and DDX60. For each of RAD51-3 (**A**), DDX60 (**B**) and ATRX (**C**), the following are shown (left to right): representative microscopy images of live fluorescence imaging of a cell expressing the proteins as fusions with mNeonGreen (mNG; scale bar = 5  $\mu\text{m}$ ); growth after loss of expression (for RAD51-3 this is a comparison of parental TbCas9T7 cells relative to null mutants (-/-), while for DDX60 and ATRX RNAi induced (+ Tet) and uninduced (- Tet) cells are shown, including relative RNA levels after 24 h of RNAi induction (uninduced RNA level was set at 100%); cell cycle profile of parental and -/- cells (RAD51-3), or before and after RNAi induction for 24, 48 and 72 h (DDX60 and ATRX), as determined through DAPI staining of nucleus (N) and kinetoplast (K). For growth analysis, error bars represent SEM from three independent experiments. For RT-qPCR to determine RNA levels, error bars show SEM from two independent experiments and statistical significance was determined using *t*-test ( $*P < 0.05$ ). For cell cycle analysis, values are shown as a proportion of cells with specific N-K configurations (1N1K, 1N2K, 2N2K, >2K, >2N) in a cell population (>300 cells); error bars represent SEM from three independent experiments.

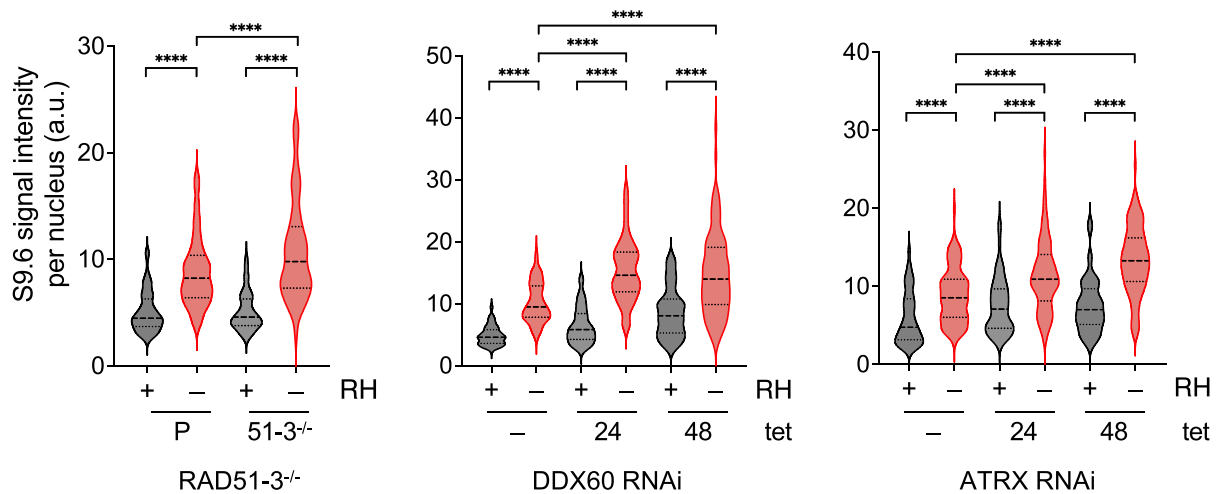
defects by fluorescent cell sorting and next generation sequencing (158).

To ask if the loss of the three putative RNA-DNA hybrid interactors affects nuclear genome integrity, we tested for levels of Thr130-phosphorylated histone H2A ( $\gamma\text{H2A}$ ), which is a marker for nuclear DNA damage (159). Western blots indicated an increased level of  $\gamma\text{H2A}$  in RAD51-3-/- cells compared with parental, while  $\gamma\text{H2A}$  levels increased after 72 h

of RNAi against DDX60 or ATRX (Figure 4A). To explore these effects further,  $\gamma\text{H2A}$  was localised and quantified by immunofluorescence (Figure 4B). An increase in the proportion of cells with  $\gamma\text{H2A}$ -positive nuclei was detected in RAD51-3-/- cells (~42% compared with ~5% in parental). Moreover,  $\gamma\text{H2A}$  nuclear signal in the mutants was notably focal (Figure 4B), suggesting discrete DNA damage accumulation and perhaps reflecting the localisation of mNG::RAD51-3 protein



**Figure 4.** Loss of *T. brucei* RAD51-3, ATRX or DDX60 leads to increased nuclear DNA damage. **(A)** Anti- $\gamma$ H2A western blot comparing *RAD51-3*<sup>-/-</sup> mutants with parental (P) cells, and before (Tet-) and after *DDX60* or *ATRX* RNAi induction (Tet+) for 24, 48 and 72 h (anti-EF1  $\alpha$  was used as a loading control). **(B)** Representative microscopy images of  $\gamma$ H2A immunofluorescence in *RAD51-3*<sup>-/-</sup> mutants and parental (P) cells and before (Tet-) and after *DDX60* or *ATRX* RNAi induction (Tet+) for 24, 48 and 72 h.  $\gamma$ H2A localisation is quantified in the adjacent graphs, which show the percent of cells with nuclear anti- $\gamma$ H2A signal. Error bars signify SEM from three independent experiments, counting at least 50 cells; statistical significance was determined for the *RAD51-3*<sup>-/-</sup> mutants using a *t*-test (\*\* $P < 0.01$ ), while statistical significance following RNAi was determined through one-way ANOVA followed by Šidák's multiple comparisons tests (ns, not significant, \*\* $P < 0.01$ , \*\*\*\* $P < 0.0001$ ).



**Figure 5.** Loss of *T. brucei* RAD51-3, ATRX or DDX60 leads to increased levels of nuclear RNA–DNA hybrids. Violin plots show the intensity of nuclear S9.6 immunofluorescence signal in parental (P) and RAD51-3<sup>-/-</sup> cells, or before (Tet-) and after DDX60 or ATRX RNAi induction (Tet+) for 24 or 48 h; in all cases the intensity was measured with (+) and without (-) treatment with *E. coli* RNase H1 (RH). Data is shown in each case for > 100 cells, the median is shown by a heavily dotted line, and the interquartile range by surrounding lightly dotted lines; statistical significance was determined through one-way ANOVA followed by Šidák's multiple comparisons test (\*\*\*\* $P < 0.0001$ ).

(Supplementary Figure S2A). Whether or not these effects are related to replication-associated DNA damage observed after loss of RAD51-3 in *L. major* is unclear (156). Accumulation of nuclear  $\gamma$ H2A signal followed the growth defects seen after RNAi of DDX60 or ATRX (Figure 4B, Figure 3B,C): for the former, no change in the proportion of cells harbouring  $\gamma$ H2A signal was seen 24 h after RNAi induction, whereas the signal increased significantly by 48 h and remained essentially the same at 72 h (22% and 25%, respectively); for ATRX, the proportion of cells expressing  $\gamma$ H2A increased significantly 24 h post-induction (~12% in induced, ~5% in uninduced) and continued to increase from 48–72 h (~17% to ~34%, respectively). In both cases,  $\gamma$ H2A signal was distinct from that seen in RAD51-3<sup>-/-</sup> cells, in that it appeared throughout the nucleus (Figure 4B). Nonetheless, loss of either of these factors also resulted in nuclear DNA damage, which is perhaps most surprising for DDX60, as localisation of DDX60::mNG suggested it is cytoplasmic (Figures 3 and S2). Whether these data indicate an undetected population of nuclear DDX60, or if the protein can dynamically move between the nucleus and cytoplasm, is unclear.

### Loss of RAD51-3, DDX60 or ATRX alters RNA–DNA hybrid homeostasis

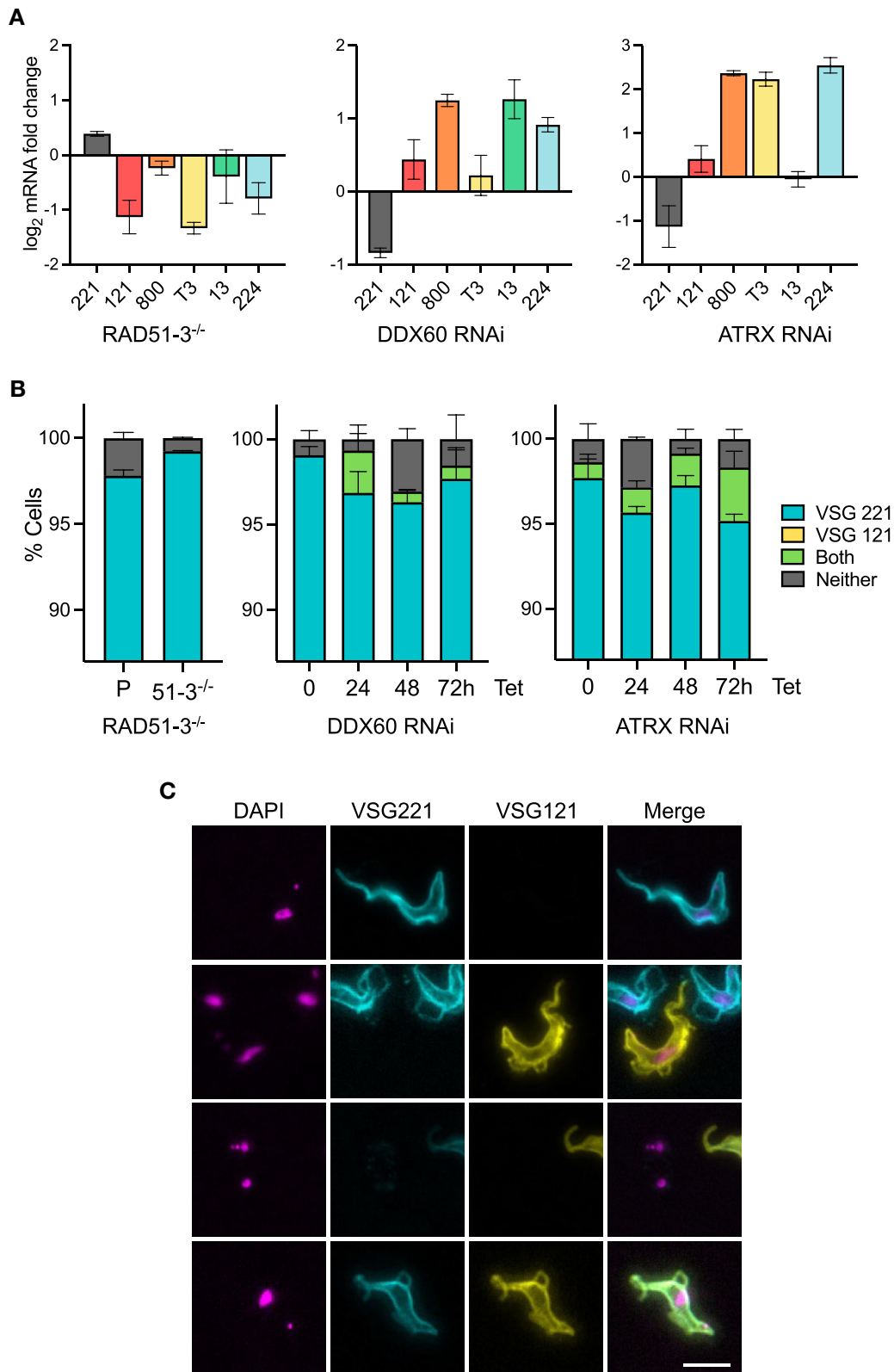
To ask if loss of the putative interactors affects RNA–DNA hybrid dynamics, we performed immunofluorescence with the S9.6 antibody (Figures 5 and S4). Unlike in mammalian cells (160,161), the majority of anti-S9.6 signal detected in parental or uninduced RNAi *T. brucei* BSF cells was nuclear (Supplementary Figure S4). In addition, and notwithstanding concerns about its effectiveness (161), treatment with *E. coli* RNase H1 significantly reduced S9.6 nuclear fluorescence intensity in the same cells (Figures 5 and S4), indicating much of the signal represents RNA–DNA hybrids, including R-loops.

RAD51-3<sup>-/-</sup> cells displayed significantly increased S9.6 fluorescence compared with the parental cells (Figure 5). In addition, tetracycline induction of the ATRX or DDX60 RNAi cells led to an increase in S9.6 nuclear signal compared to the uninduced (Figure 5). In fact, the temporal changes in

S9.6 signal appeared to have parallels with the growth curves (Figure 3B, C) and  $\gamma$ H2A immunofluorescence (Figure 4B: for ATRX, median fluorescence increased from 24–48 h after RNAi, whereas for DDX60 median fluorescence increased by 24 h and was unchanged 48 h post-induction (Figure 5). Hence, growth impairment and nuclear DNA damage may follow from increased levels of RNA–DNA hybrids due to the loss of the factors. In addition, the findings reiterate a nuclear function for DDX60. Taken together, these data indicate each of these factors acts in RNA–DNA hybrid homeostasis, consistent with their recovery and identification by DRIP-MS. Whether the altered levels of S9.6 reflect roles for the factors in depositing, resolving or acting upon RNA–DNA hybrids is unclear.

### Loss of RAD51-3, DDX60 or ATRX alters VSG expression

Given the above evidence linking RAD51-3, DDX60 and ATRX with homeostasis of nuclear RNA–DNA hybrids and with nuclear DNA damage, we next asked if their loss has an impact on VSG switching. MiTat1.2 BSF cells (the strain used for all experiments here) predominantly express VSG221 (also named VSG2) from BES1 [39,40,42]. When wild type MIT1.2 BSF cells are grown in culture, a small proportion (~1–3%) of cells switch off expression of VSG221 and activate a distinct VSG [39–41,66,67]. To ask if this stochastic switching frequency is altered by loss of the RNA–DNA hybrid interactors, RT-qPCR was first used to assess RNA levels of VSG221 and four VSGs in normally silent BESs (Figure 6A). In the RAD51-3<sup>-/-</sup> parasites RT-qPCR indicated increased levels of VSG221 transcript relative to parental cells, and an associated reduction in all VSG transcripts from the mainly silent BESs. RNAi of DDX60 or ATRX had the opposite effect: in both cases less VSG221 transcript was expressed in the 24 h induced populations relative to uninduced, and four or five of the silent BES VSG transcripts increased in abundance. These data suggest that loss of RAD51-3 reduces switching away from VSG221 towards any of the silent BES-resident VSGs tested, while loss of DDX60 or ATRX increases switch-



**Figure 6.** Loss of *T. brucei* RAD51-3, ATRX or DDX60 leads to altered VSG expression. **(A)** Quantification of the relative levels of VSG RNAs, comparing parental and *RAD51-3*<sup>-/-</sup> mutants, or in cells 24 h post-induction of RNAi against *DDX60* or *ATRX* compared with uninduced cells; error bars indicate SEM of two independent experiments. **(B)** Graphical representation of VSG immunofluorescence analysis that shows the proportion of cells staining positive for VSG221 (teal), VSG121 (not detected), both VSGs (green), or neither VSG (grey) in parental (P) and *RAD51-3*<sup>-/-</sup> cells, or 0, 24, 48 and 72 h after RNAi induction (Tet) against *DDX60* or *ATRX*; error bars represent SEM from three independent experiments, counting >300 cells. **(C)** Sample VSG immunofluorescence images are shown 24 h after *ATRX* RNAi; scale bar = 10 μm.

ing away from VSG221 and increases activation of the silent BES VSGs. Notably, in the latter cases, switching alteration was detected prior to the significant accumulation of DNA damage or pronounced growth defects beyond 24 h of RNAi induction.

To test the effects on VSG expression further, we performed immunofluorescence on live cells with antiserum against VSG221 and VSG121, which is expressed from the predominantly silent BES3 (Figure 6B). Consistent with the RT-qPCR, more *RAD51-3-/-* cells that express VSG221 were detected in the population compared to the parental line. In addition, while VSG121 was not detected in parental or mutant populations, a reduced number of *RAD51-3-/-* cells were found that did not express either coat protein. These findings are consistent with previous reports that measured VSG switching using *in vivo* immune selection against VSG221 (54,55), and confirm that *RAD51-3* loss reduces the efficiency of *T. brucei* VSG switching. Loss of *RAD51* also leads to decreased VSG switching but, distinct from *RAD51-3* mutants, levels of RNA–DNA hybrids are reduced (Girasol et al, BioRxiv). Hence, more work is needed to understand how the patterns of global and local R-loops are influenced by HR factors, including in the BESs. VSG221 and VSG121 immunofluorescence provided a fuller explanation of the RT-qPCR analysis after *DDX60* or *ATRX* RNAi (Figure 6B). Here, the analysis was conducted over 72 h after RNAi induction and, in both cases, cells were detected that expressed both VSG221 and VSG121 as early as 24 h (an example is shown in Figure 6C). Before induction of *DDX60* RNAi no such cells were detected, whereas a small number were seen prior to induction of *ATRX* RNAi and their numbers increased after induction. Taken together, these findings indicate that loss of either factor impairs the gene expression controls that normally operate to ensure only a single VSG BES is transcribed, or their RNAi delays the process of transition from expressing VSG221 to VSG121 (and potentially to any new VSG), as seen after loss of *DOT1B* (162). These same effects - loss of VSG expression control and altered or increased VSG switching - are also seen after loss of RNase H1 or RNase H2A (49,50), where R-loop levels increase in the VSG BESs, indicating a potential link in terms of R-loop homeostasis.

## Conclusions

RNA–DNA hybrids are ubiquitous epigenetic features of all DNA genomes, where their list of functions continues to expand. Understanding this range of functions can be aided by describing the proteins that interact with RNA–DNA hybrids, though such studies have to date only been conducted in mammals. Here, we describe a large cohort of putative RNA–DNA hybrid interactors in the protozoan parasite, *Trypanosoma brucei*, where R-loops have been mapped genome-wide and implicated in both conserved and lineage-specific activities (28). Consistent with predictions of conserved activities, we find overlap between our data and mammalian studies (22,27), including ribosome-associated proteins and predicted RNA helicases. In fact, our data provide a potentially novel link between these activities, in that we describe a predominantly cytoplasmic DDX (*DDX60*) that has been found to associate with the ribosome in *T. brucei*'s relative, *T. cruzi* (143) and influences RNA–DNA hybrid levels in the parasite and, moreover, appears to have moonlighting activities in nuclear DNA repair. Our data also provides a num-

ber of predicted RNA–DNA hybrid interactors with activities that may reflect particular features of trypanosome gene expression and chromosome biology, including histone variants (33) and kinetochore components (100), though it should be noted that tests are needed to determine that these factors do indeed bind RNA–DNA hybrids.

A crucial activity used by *T. brucei* to survive in its mammalian host is antigenic variation, which is driven by transcriptional controls and recombination of VSG genes (43). Amongst the proteins recovered by DRIP-MS, we have found four that act in VSG switching. Two are key determinants of DNA repair by homologous recombination: *RAD51* (Girasol et al., BioRxiv) and *RAD51-3*. Though both proteins have previously been shown to act in VSG switching by recombination (44), their interaction with RNA–DNA hybrids provides new mechanistic understanding and builds upon recent work that has implicated R-loops in the reaction (49,50,53,118). Two other factors, *DDX60* and *ATRX*, have never before been implicated in VSG switching. How these factors act in antigenic variation is not yet clear, but the observation that their loss, like that of *RAD51-3*, leads to increased RNA–DNA hybrids levels, while they have distinct effects on VSG switching to the *RAD51* paralogue, illustrates that emerging links between R-loops and VSG transcription or recombination deserve further analysis. For instance, the increased levels of RNA–DNA hybrids seen in *RAD51-3* mutants perhaps indicates DNA repair functions that operate more widely than localised activities dedicated to VSG recombination.

## Data availability

DRIP-mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository, with the dataset identifier PXD042146.

## Supplementary data

Supplementary Data are available at NAR Online.

## Acknowledgements

We thank Suzanne McGill and all staff at Glasgow Polyomics for support in mass spectrometry analysis, and all lab members for helpful discussions. The graphical abstract was created with Biorender.com.

## Funding

Wellcome Trust Investigator Award [224501/Z/21/Z to R.M.]; Philippine Council for Health Research and Development (DOST-PCHRD) (to M.G.); Wellcome Trust Sir Henry Wellcome fellowship [218648/Z/19/Z to E.M.B.]; Wellcome Institutional Strategic Support Fund (ISSF3) award held at the University of Glasgow [204820/Z/16/Z to E.M.B., R.M.]; BBSRC [BB/N016165/1, BB/R017166/1, BB/W001101/1 to R.M., C.A.M.]; Wellcome Centre for Integrative Parasitology is funded by a Wellcome Trust Strategic Award [104111/Z/14/Z/A]. Funding for open access charge: UKRI/Wellcome Trust block grant.

## Conflict of interest statement

None declared.

## References

- Aguilera, A. and Garcia-Muse, T. (2012) R loops: from transcription byproducts to threats to genome stability. *Mol. Cell*, **46**, 115–124.
- Petermann, E., Lan, L. and Zou, L. (2022) Sources, resolution and physiological relevance of R-loops and RNA–DNA hybrids. *Nat. Rev. Mol. Cell Biol.*, **23**, 521–540.
- Brickner, J.R., Garzon, J.L. and Cimprich, K.A. (2022) Walking a tightrope: the complex balancing act of R-loops in genome stability. *Mol. Cell*, **82**, 2267–2297.
- Niehrs, C. and Luke, B. (2020) Regulatory R-loops as facilitators of gene expression and genome stability. *Nat. Rev. Mol. Cell Biol.*, **21**, 167–178.
- Stuckey, R., Garcia-Rodriguez, N., Aguilera, A. and Wellinger, R.E. (2015) Role for RNA:DNA hybrids in origin-independent replication priming in a eukaryotic system. *Proc. Nat. Acad. Sci. U.S.A.*, **112**, 5779–5784.
- Lombrana, R., Almeida, R., Alvarez, A. and Gomez, M. (2015) R-loops and initiation of DNA replication in human cells: a missing link? *Front. Genet.*, **6**, 158.
- Tran, P.L.T., Pohl, T.J., Chen, C.F., Chan, A., Pott, S. and Zakian, V.A. (2017) PIF1 family DNA helicases suppress R-loop mediated genome instability at tRNA genes. *Nat. Commun.*, **8**, 15025.
- Santos-Pereira, J.M. and Aguilera, A. (2015) R loops: new modulators of genome dynamics and function. *Nat. Rev. Genet.*, **16**, 583–597.
- Chedin, F. (2016) Nascent connections: r-loops and chromatin patterning. *Trends Genet.*, **32**, 828–838.
- Zhang, H., Gan, H., Wang, Z., Lee, J.H., Zhou, H., Ordog, T., Wold, M.S., Ljungman, M. and Zhang, Z. (2017) RPA interacts with HIRA and regulates H3.3 deposition at gene regulatory elements in mammalian cells. *Mol. Cell*, **65**, 272–284.
- Nanavaty, V., Sandhu, R., Jehi, S.E., Pandya, U.M. and Li, B. (2017) *Trypanosoma brucei* RAP1 maintains telomere and subtelomere integrity by suppressing TERRA and telomeric RNA:DNA hybrids. *Nucleic Acids Res.*, **45**, 5785–5796.
- Rippe, K. and Luke, B. (2015) TERRA and the state of the telomere. *Nat. Struct. Mol. Biol.*, **22**, 853–858.
- Stirling, P.C. and Hieter, P. (2016) Canonical DNA repair pathways influence R-loop-driven genome instability. *J. Mol. Biol.*, **429**, 3132–3138.
- Sollier, J. and Cimprich, K.A. (2015) Breaking bad: r-loops and genome integrity. *Trends Cell Biol.*, **25**, 514–522.
- Garcia-Muse, T. and Aguilera, A. (2019) R Loops: from Physiological to Pathological Roles. *Cell*, **179**, 604–618.
- Daniels, G.A. and Lieber, M.R. (1995) RNA:DNA complex formation upon transcription of immunoglobulin switch regions: implications for the mechanism and regulation of class switch recombination. *Nucleic Acids Res.*, **23**, 5006–5011.
- Yu, K., Chedin, F., Hsieh, C.L., Wilson, T.E. and Lieber, M.R. (2003) R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nat. Immunol.*, **4**, 442–451.
- Matthews, A.J., Zheng, S., DiMenna, L.J. and Chaudhuri, J. (2014) Regulation of immunoglobulin class-switch recombination: choreography of noncoding transcription, targeted DNA deamination, and long-range DNA repair. *Adv. Immunol.*, **122**, 1–57.
- Ketley, R.F. and Gullerova, M. (2020) Jack of all trades? The versatility of RNA in DNA double-strand break repair. *Essays Biochem.*, **64**, 721–735.
- Marnef, A. and Legube, G. (2021) R-loops as Janus-faced modulators of DNA repair. *Nat. Cell Biol.*, **23**, 305–313.
- Boguslawski, S.J., Smith, D.E., Michalak, M.A., Mickelson, K.E., Yehle, C.O., Patterson, W.L. and Carrico, R.J. (1986) Characterization of monoclonal antibody to DNA:RNA and its application to immunodetection of hybrids. *J. Immunol. Methods*, **89**, 123–130.
- Cristini, A., Groh, M., Kristiansen, M.S. and Gromak, N. (2018) RNA/DNA hybrid interactome identifies DXH9 as a molecular player in transcriptional termination and R-loop-associated DNA damage. *Cell Rep.*, **23**, 1891–1905.
- Wu, T., Nance, J., Chu, F. and Fazio, T.G. (2021) Characterization of R-loop-interacting proteins in embryonic stem cells reveals roles in rRNA processing and gene expression. *Mol. Cell. Proteomics*, **20**, 100142.
- Wang, I.X., Grunseich, C., Fox, J., Burdick, J., Zhu, Z., Ravazian, N., Hafner, M. and Cheung, V.G. (2018) Human proteins that interact with RNA/DNA hybrids. *Genome Res.*, **28**, 1405–1414.
- Mosler, T., Conte, F., Longo, G.M.C., Mikicic, J., Kreim, N., Mockel, M.M., Petrosino, G., Flach, J., Barau, J., Luke, B., et al. (2021) R-loop proximity proteomics identifies a role of DDX41 in transcription-associated genomic instability. *Nat. Commun.*, **12**, 7314.
- Yan, Q., Wulfridge, P., Doherty, J., Fernandez-Luna, J.L., Real, P.J., Tang, H.Y. and Sarma, K. (2022) Proximity labeling identifies a repertoire of site-specific R-loop modulators. *Nat. Commun.*, **13**, 53.
- Kumar, A., Fournier, L.A. and Stirling, P.C. (2022) Integrative analysis and prediction of human R-loop binding proteins. *G3*, **12**, jkac142.
- Briggs, E., Hamilton, G., Crouch, K., Lapsley, C. and McCulloch, R. (2018) Genome-wide mapping reveals conserved and diverged R-loop activities in the unusual genetic landscape of the African trypanosome genome. *Nucleic Acids Res.*, **46**, 11789–11805.
- Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D.C., Lennard, N.J., Caler, E., Hamlin, N.E., Haas, B., et al. (2005) The genome of the African trypanosome *Trypanosoma brucei*. *Science*, **309**, 416–422.
- Muller, L.S.M., Cosentino, R.O., Forstner, K.U., Guizetti, J., Wedel, C., Kaplan, N., Janzen, C.J., Arampatzi, P., Vogel, J., Steinbiss, S., et al. (2018) Genome organization and DNA accessibility control antigenic variation in trypanosomes. *Nature*, **563**, 121–125.
- Staneva, D.P., Bresson, S., Auchynnika, T., Spanos, C., Rappsilber, J., Jayaprakash, A.A., Tollervey, D., Matthews, K.R. and Allshire, R.C. (2022) The SPARC complex defines RNAPII promoters in *Trypanosoma brucei*. *eLife*, **11**, e83135.
- Staneva, D.P., Carloni, R., Auchynnika, T., Tong, P., Rappsilber, J., Jayaprakash, A.A., Matthews, K.R. and Allshire, R.C. (2021) A systematic analysis of *Trypanosoma brucei* chromatin factors identifies novel protein interaction networks associated with sites of transcription initiation and termination. *Genome Res.*, **31**, 2138–2154.
- Siegel, T.N., Hekstra, D.R., Kemp, L.E., Figueiredo, L.M., Lowell, J.E., Fenyo, D., Wang, X., Dewell, S. and Cross, G.A. (2009) Four histone variants mark the boundaries of polycistronic transcription units in *Trypanosoma brucei*. *Genes Dev.*, **23**, 1063–1076.
- Wedel, C., Forstner, K.U., Derr, R. and Siegel, T.N. (2017) GT-rich promoters can drive RNA pol II transcription and deposition of H2A.Z in African trypanosomes. *EMBO J.*, **36**, 2581–2594.
- Cordon-Obras, C., Gomez-Linan, C., Torres-Rusillo, S., Vidal-Cobo, J., Lopez-Farfan, D., Barroso-Del Jesus, A., Rojas-Barros, D., Carrington, M. and Navarro, M. (2022) Identification of sequence-specific promoters driving polycistronic transcription initiation by RNA polymerase II in trypanosomes. *Cell Rep.*, **38**, 110221.
- Daniels, J.P., Gull, K. and Wickstead, B. (2010) Cell biology of the trypanosome genome. *Microbiol. Mol. Biol. Rev.*, **74**, 552–569.
- Siegel, T.N., Gunasekera, K., Cross, G.A. and Ochsenreiter, T. (2011) Gene expression in *Trypanosoma brucei*: lessons from high-throughput RNA sequencing. *Trends Parasitol.*, **27**, 434–441.
- Kieft, R., Zhang, Y., Marand, A.P., Moran, J.D., Bridger, R., Wells, L., Schmitz, R.J. and Sabatini, R. (2020) Identification of a novel base J binding protein complex involved in RNA polymerase II transcription termination in trypanosomes. *PLoS Genet.*, **16**, e1008390.

39. Tiengwe,C., Marcello,L., Farr,H., Dickens,N., Kelly,S., Swiderski,M., Vaughan,D., Gull,K., Barry,J.D., Bell,S.D., *et al.* (2012) Genome-wide analysis reveals extensive functional interaction between DNA replication initiation and transcription in the genome of *Trypanosoma brucei*. *Cell Rep.*, **2**, 185–197.
40. Devlin,R., Marques,C.A., Paape,D., Prorocic,M., Zurita-Leal,A.C., Campbell,S.J., Lapsley,C., Dickens,N. and McCulloch,R. (2016) Mapping replication dynamics in *Trypanosoma brucei* reveals a link with telomere transcription and antigenic variation. *eLife*, **5**, e12765.
41. Marques,C.A., Tiengwe,C., Lemgruber,L., Damasceno,J.D., Scott,A., Paape,D., Marcello,L. and McCulloch,R. (2016) Diverged composition and regulation of the *Trypanosoma brucei* origin recognition complex that mediates DNA replication initiation. *Nucleic Acids Res.*, **44**, 4763–4784.
42. McCulloch,R., Cobbold,C.A., Figueiredo,L., Jackson,A., Morrison,L.J., Mugnier,M.R., Papavasiliou,N., Schnauffer,A. and Matthews,K. (2017) Emerging challenges in understanding trypanosome antigenic variation. *Emerg. Top. Life Sci.*, **1**, 585–592.
43. Faria,J., Briggs,E.M., Black,J.A. and McCulloch,R. (2022) Emergence and adaptation of the cellular machinery directing antigenic variation in the African trypanosome. *Curr. Opin. Microbiol.*, **70**, 102209.
44. McCulloch,R., Morrison,L.J. and Hall,J.P. (2015) DNA recombination strategies during antigenic variation in the African trypanosome. *Microbiol. Spectr.*, **3**, MDNA3–A0016–2014.
45. Gunzl,A., Bruderer,T., Laufer,G., Schimanski,B., Tu,L.C., Chung,H.M., Lee,P.T. and Lee,M.G. (2003) RNA polymerase I transcribes procyclin genes and variant surface glycoprotein gene expression sites in *Trypanosoma brucei*. *Eukaryot. Cell*, **2**, 542–551.
46. Cross,G.A., Kim,H.S. and Wickstead,B. (2014) Capturing the variant surface glycoprotein repertoire (the VSGnome) of *Trypanosoma brucei* Lister 427. *Mol. Biochem. Parasitol.*, **195**, 59–73.
47. Marcello,L. and Barry,J.D. (2007) Analysis of the VSG gene silent archive in *Trypanosoma brucei* reveals that mosaic gene expression is prominent in antigenic variation and is favored by archive substructure. *Genome Res.*, **17**, 1344–1352.
48. Cerritelli,S.M. and Crouch,R.J. (2009) Ribonuclease H: the enzymes in eukaryotes. *FEBS J.*, **276**, 1494–1505.
49. Briggs,E., Crouch,K., Lemgruber,L., Hamilton,G., Lapsley,C. and McCulloch,R. (2019) *Trypanosoma brucei* ribonuclease H2A is an essential R-loop processing enzyme whose loss causes DNA damage during transcription initiation and antigenic variation. *Nucleic Acids Res.*, **47**, 9180–9197.
50. Briggs,E., Crouch,K., Lemgruber,L., Lapsley,C. and McCulloch,R. (2018) Ribonuclease H1-targeted R-loops in surface antigen gene expression sites can direct trypanosome immune evasion. *PLoS Genet.*, **14**, e1007729.
51. Eisenhuth,N., Vellmer,T., Rauh,E.T., Butter,F. and Janzen,C.J. (2021) A DOT1B/ribonuclease H2 protein complex is involved in R-loop processing, genomic integrity, and antigenic variation in *Trypanosoma brucei*. *mBio*, **12**, e0135221.
52. Nanavaty,V., Sandhu,R., Jehi,S.E., Pandya,U.M. and Li,B. (2017) *Trypanosoma brucei* RAP1 maintains telomere and subtelomere integrity by suppressing TERRA and telomeric RNA:DNA hybrids. *Nucleic Acids Res.*, **45**, 5785–5796.
53. Saha,A., Gaurav,A.K., Pandya,U.M., Afrin,M., Sandhu,R., Nanavaty,V., Schnur,B. and Li,B. (2021) TbTRF suppresses the TERRA level and regulates the cell cycle-dependent TERRA foci number with a TERRA binding activity in its C-terminal Myb domain. *Nucleic Acids Res.*, **49**, 5637–5653.
54. Proudfoot,C. and McCulloch,R. (2005) Distinct roles for two RAD51-related genes in *Trypanosoma brucei* antigenic variation. *Nucleic Acids Res.*, **33**, 6906–6919.
55. Dobson,R., Stockdale,C., Lapsley,C., Wilkes,J. and McCulloch,R. (2011) Interactions among *Trypanosoma brucei* RAD51 paralogues in DNA repair and antigenic variation. *Mol. Microbiol.*, **81**, 434–456.
56. Rojas,F., Silvester,E., Young,J., Milne,R., Tettey,M., Houston,D.R., Walkinshaw,M.D., Perez-Pi,J., Auer,M., Denton,H., *et al.* (2019) Oligopeptide signaling through TbGPR89 drives trypanosome quorum sensing. *Cell*, **176**, 306–317.
57. Beneke,T., Madden,R., Makin,L., Valli,J., Sunter,J. and Gluenz,E. (2017) A CRISPR Cas9 high-throughput genome editing toolkit for kinetoplastids. *R. Soc. Open Sci.*, **4**, 170095.
58. Alsford,S., Glover,L. and Horn,D. (2005) Multiplex analysis of RNA interference defects in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.*, **139**, 129–132.
59. Redmond,S., Vadivelu,J. and Field,M.C. (2003) RNAit: an automated web-based tool for the selection of RNAi targets in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.*, **128**, 115–118.
60. Alsford,S. and Horn,D. (2007) RNA polymerase I transcription stimulates homologous recombination in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.*, **153**, 77–79.
61. Livak,K.J. and Schmittgen,T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-</sup>(Delta Delta C(T)) Method. *Methods*, **25**, 402–408.
62. Glover,L., Hutchinson,S., Alsford,S. and Horn,D. (2016) VEX1 controls the allelic exclusion required for antigenic variation in trypanosomes. *Proc. Nat. Acad. Sci. U.S.A.*, **113**, 7225–7230.
63. Schindelin,J., Arganda-Carreras,I., Frise,E., Kaynig,V., Longair,M., Pietzsch,T., Preibisch,S., Rueden,C., Saalfeld,S., Schmid,B., *et al.* (2012) Fiji: an open-source platform for biological-image analysis. *Nat. Methods*, **9**, 676–682.
64. Potter,S.C., Luciani,A., Eddy,S.R., Park,Y., Lopez,R. and Finn,R.D. (2018) HMMER web server: 2018 update. *Nucleic Acids Res.*, **46**, W200–W204.
65. Zimmermann,L., Stephens,A., Nam,S.Z., Rau,D., Kubler,J., Lozajic,M., Gabler,F., Soding,J., Lupas,A.N. and Alva,V. (2018) A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J. Mol. Biol.*, **430**, 2237–2243.
66. Steinegger,M., Meier,M., Mirdita,M., Vohringer,H., Haunsberger,S.J. and Soding,J. (2019) HH-suite3 for fast remote homology detection and deep protein annotation. *BMC Bioinf.*, **20**, 473.
67. Aramaki,T., Blanc-Mathieu,R., Endo,H., Ohkubo,K., Kanehisa,M., Goto,S. and Ogata,H. (2020) KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics*, **36**, 2251–2252.
68. Paysan-Lafosse,T., Blum,M., Chuguransky,S., Grego,T., Pinto,B.L., Salazar,G.A., Bileschi,M.L., Bork,P., Bridge,A., Colwell,L., *et al.* (2023) InterPro in 2022. *Nucleic Acids Res.*, **51**, D418–D427.
69. Fairman-Williams,M.E., Guenther,U.P. and Jankowsky,E. (2010) SF1 and SF2 helicases: family matters. *Curr. Opin. Struct. Biol.*, **20**, 313–324.
70. Bobik,K., McCray,T.N., Ernest,B., Fernandez,J.C., Howell,K.A., Lane,T., Staton,M. and Burch-Smith,T.M. (2017) The chloroplast RNA helicase ISE2 is required for multiple chloroplast RNA processing steps in *Arabidopsis thaliana*. *Plant J.*, **91**, 114–131.
71. Uson,M.L., Ordonez,H. and Shuman,S. (2015) *Mycobacterium smegmatis* HelY Is an RNA-activated ATPase/dATPase and 3'-to-5' helicase that unwinds 3'-tailed RNA duplexes and RNA:DNA hybrids. *J. Bacteriol.*, **197**, 3057–3065.
72. Katoh,K. and Standley,D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.*, **30**, 772–780.
73. Rozewicki,J., Li,S., Amada,K.M., Standley,D.M. and Katoh,K. (2019) MAFFT-DASH: integrated protein sequence and structural alignment. *Nucleic Acids Res.*, **47**, W5–W10.
74. Waterhouse,A.M., Procter,J.B., Martin,D.M., Clamp,M. and Barton,G.J. (2009) Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, **25**, 1189–1191.



75. Capella-Gutierrez,S., Silla-Martinez,J.M. and Gabaldon,T. (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, **25**, 1972–1973.
76. Minh,B.Q., Schmidt,H.A., Chernomor,O., Schrempf,D., Woodhams,M.D., Haeseler,A. and Lanfear,R. (2020) IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.*, **37**, 1530–1534.
77. Guindon,S., Dufayard,J.F., Lefort,V., Anisimova,M., Hordijk,W. and Gascuel,O. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.*, **59**, 307–321.
78. Hoang,D.T., Chernomor,O., von Haeseler,A., Minh,B.Q. and Vinh,L.S. (2018) UFBoot2: improving the Ultrafast Bootstrap Approximation. *Mol. Biol. Evol.*, **35**, 518–522.
79. Flaus,A., Martin,D.M., Barton,G.J. and Owen-Hughes,T. (2006) Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. *Nucleic Acids Res.*, **34**, 2887–2905.
80. Tria,F.D.K., Landan,G. and Dagan,T. (2017) Phylogenetic rooting using minimal ancestor deviation. *Nat. Ecol. Evol.*, **1**, 193.
81. Letunic,I. and Bork,P. (2021) Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.*, **49**, W293–W296.
82. Phillips,D.D., Garboczi,D.N., Singh,K., Hu,Z., Leppla,S.H. and Leysath,C.E. (2013) The sub-nanomolar binding of DNA–RNA hybrids by the single-chain Fv fragment of antibody S9.6. *J. Mol. Recognit.*, **26**, 376–381.
83. Hartono,S.R., Malapert,A., Legros,P., Bernard,P., Chedin,F. and Vanoosthuysen,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**, 272–284.
84. Moreno,J.M., Sanchez-Montero,J.M., Ballesteros,A. and Sinisterra,J.V. (1991) Hydrolysis of nucleic acids in single-cell protein concentrates using immobilized benzonase. *Appl. Biochem. Biotechnol.*, **31**, 43–51.
85. Butter,F., Bucerius,F., Michel,M., Cicova,Z., Mann,M. and Janzen,C.J. (2013) Comparative proteomics of two life cycle stages of stable isotope-labeled *Trypanosoma brucei* reveals novel components of the parasite's host adaptation machinery. *Mol. Cell. Proteomics*, **12**, 172–179.
86. Billington,K., Halliday,C., Madden,R., Dyer,P., Barker,A.R., Moreira-Leite,F.F., Carrington,M., Vaughan,S., Hertz-Fowler,C., Dean,S., *et al.* (2023) Genome-wide subcellular protein map for the flagellate parasite *Trypanosoma brucei*. *Nat. Microbiol.*, **8**, 533–547.
87. Lawrimore,J., Kolbin,D., Stanton,J., Khan,M., de Larminat,S.C., Lawrimore,C., Yeh,E. and Bloom,K. (2021) The rDNA is biomolecular condensate formed by polymer-polymer phase separation and is sequestered in the nucleolus by transcription and R-loops. *Nucleic Acids Res.*, **49**, 4586–4598.
88. Abraham,K.J., Khosraviani,N., Chan,J.N.Y., Gorthi,A., Samman,A., Zhao,D.Y., Wang,M., Bokros,M., Vidya,E., Ostrowski,L.A., *et al.* (2020) Nucleolar RNA polymerase II drives ribosome biogenesis. *Nature*, **585**, 298–302.
89. Crossley,M.P., Song,C., Bocek,M.J., Choi,J.H., Kousorous,J., Sathirachinda,A., Lin,C., Brickner,J.R., Bai,G., Lans,H., *et al.* (2023) R-loop-derived cytoplasmic RNA–DNA hybrids activate an immune response. *Nature*, **613**, 187–194.
90. Sanz,L.A., Hartono,S.R., Lim,Y.W., Steyaert,S., Rajpurkar,A., Ginno,P.A., Xu,X. and Chedin,F. (2016) Prevalent, dynamic, and conserved R-loop structures associate with specific epigenomic signatures in mammals. *Mol. Cell*, **63**, 167–178.
91. Anderson,B.A., Wong,I.L., Baugh,L., Ramasamy,G., Myler,P.J. and Beverley,S.M. (2013) Kinetoplastid-specific histone variant functions are conserved in *Leishmania major*. *Mol. Biochem. Parasitol.*, **191**, 53–57.
92. Roson,J.N., Vitarelli,M.O., Costa-Silva,H.M., Pereira,K.S., Pires,D.D.S., Lopes,L.S., Cordeiro,B., Kraus,A.J., Cruz,K.N.T., Calderano,S.G., *et al.* (2022) H2B.V demarcates divergent strand-switch regions, some tDNA loci, and genome compartments in *Trypanosoma cruzi* and affects parasite differentiation and host cell invasion. *PLoS Pathog.*, **18**, e1009694.
93. Mani,J., Guttinger,A., Schimanski,B., Heller,M., Acosta-Serrano,A., Pescher,P., Spath,G. and Roditi,I. (2011) Alba-domain proteins of *Trypanosoma brucei* are cytoplasmic RNA-binding proteins that interact with the translation machinery. *PLoS One*, **6**, e22463.
94. Goyal,M., Banerjee,C., Nag,S. and Bandyopadhyay,U. (2016) The Alba protein family: structure and function. *Biochim. Biophys. Acta*, **1864**, 570–583.
95. Subota,I., Rotureau,B., Blisnick,T., Ngwabyt,S., Durand-Dubief,M., Engstler,M. and Bastin,P. (2011) ALBA proteins are stage regulated during trypanosome development in the tsetse fly and participate in differentiation. *Mol. Biol. Cell*, **22**, 4205–4219.
96. Bevkals,S., Naguleswaran,A., Rehmann,R., Kaiser,M., Heller,M. and Roditi,I. (2021) An Alba-domain protein required for proteome remodelling during trypanosome differentiation and host transition. *PLoS Pathog.*, **17**, e1009239.
97. Melo do Nascimento,L., Egler,F., Arnold,K., Papavasiliou,N., Clayton,C. and Erben,E. (2021) Functional insights from a surface antigen mRNA-bound proteome. *eLife*, **10**, e68136.
98. Yuan,W., Zhou,J., Tong,J., Zhuo,W., Wang,L., Li,Y., Sun,Q. and Qian,W. (2019) ALBA protein complex reads genetic R-loops to maintain genome stability in Arabidopsis. *Sci. Adv.*, **5**, eaav9040.
99. Chene,A., Vembar,S.S., Riviere,L., Lopez-Rubio,J.J., Claes,A., Siegel,T.N., Sakamoto,H., Scheidig-Benatar,C., Hernandez-Rivas,R. and Scherf,A. (2012) PfAlbas constitute a new eukaryotic DNA/RNA-binding protein family in malaria parasites. *Nucleic Acids Res.*, **40**, 3066–3077.
100. Akiyoshi,B. and Gull,K. (2014) Discovery of unconventional kinetochores in kinetoplastids. *Cell*, **156**, 1247–1258.
101. Mishra,P.K., Chakraborty,A., Yeh,E., Feng,W., Bloom,K.S. and Basrai,M.A. (2021) R-loops at centromeric chromatin contribute to defects in kinetochore integrity and chromosomal instability in budding yeast. *Mol. Biol. Cell*, **32**, 74–89.
102. Kabeche,L., Nguyen,H.D., Buisson,R. and Zou,L. (2018) A mitosis-specific and R loop-driven ATR pathway promotes faithful chromosome segregation. *Science*, **359**, 108–114.
103. Liu,Q., Yi,C., Zhang,Z., Su,H., Liu,C., Huang,Y., Li,W., Hu,X., Liu,C., Birchler,J.A., *et al.* (2023) Non-B-form DNA tends to form in centromeric regions and has undergone changes in polyploid oat subgenomes. *Proc. Nat. Acad. Sci. U.S.A.*, **120**, e2211683120.
104. Liu,Y., Liu,Q., Su,H., Liu,K., Xiao,X., Li,W., Sun,Q., Birchler,J.A. and Han,F. (2021) Genome-wide mapping reveals R-loops associated with centromeric repeats in maize. *Genome Res.*, **31**, 1409–1418.
105. D'Archivio,S. and Wickstead,B. (2016) Trypanosome outer kinetochore proteins suggest conservation of chromosome segregation machinery across eukaryotes. *J. Cell Biol.*, **216**, 379–391.
106. Marciano,G., Ishii,M., Nerusheva,O.O. and Akiyoshi,B. (2021) Kinetoplastid kinetochore proteins KKT2 and KKT3 have unique centromere localization domains. *J. Cell Biol.*, **220**, e202101022.
107. Ishii,M., Ludzia,P., Marciano,G., Allen,W., Nerusheva,O.O. and Akiyoshi,B. (2022) Divergent polo boxes in KKT2 bind KKT1 to initiate the kinetochore assembly cascade in *Trypanosoma brucei*. *Mol. Biol. Cell*, **33**, ar143.
108. Akiyoshi,B. and Gull,K. (2013) Evolutionary cell biology of chromosome segregation: insights from trypanosomes. *Open Biology*, **3**, 130023.
109. Hassler,M., Shaltiel,I.A. and Haering,C.H. (2018) Towards a unified model of SMC complex function. *Curr. Biol.*, **28**, R1266–R1281.
110. Gluenz,E., Sharma,R., Carrington,M. and Gull,K. (2008) Functional characterisation of cohesin subunit SCC1 in

- Trypanosoma brucei and dissection of mutant phenotypes in two life cycle stages. *Mol. Microbiol.*, **69**, 666–680.
111. Landeira,D., Bart,J.M., Van Tyne,D. and Navarro,M. (2009) Cohesin regulates VSG monoallelic expression in trypanosomes. *J. Cell Biol.*, **186**, 243–254.
  112. Yamin,K., Bijlani,S., Berman,J., Soni,A., Shlomai,J., Buragohain,B.M., Werbner,M., Gal-Tanamy,M., Matityahu,A. and Onn,I. (2022) Fold-change of chromatin condensation in yeast is a conserved property. *Sci. Rep.*, **12**, 17393.
  113. Jones,N.G., Thomas,E.B., Brown,E., Dickens,N.J., Hammarton,T.C. and Mottram,J.C. (2014) Regulators of Trypanosoma brucei cell cycle progression and differentiation identified using a kinome-wide RNAi screen. *PLoS Pathog.*, **10**, e1003886.
  114. Pavan,I.C.B., Peres de Oliveira,A., Dias,P.R.F., Basei,F.L., Issayama,L.K., Ferezin,C.C., Silva,F.R., Rodrigues de Oliveira,A.L., Alves Dos Reis Moura,L., Martins,M.B., et al. (2021) On broken ne(c)ks and broken DNA: the role of human NEKs in the DNA damage response. *Cells*, **10**, 507.
  115. Miller,J.C., Delzell,S.B., Concepcion-Acevedo,J., Boucher,M.J. and Klingbeil,M.M. (2020) A DNA polymerization-independent role for mitochondrial DNA polymerase I-like protein C in African trypanosomes. *J. Cell Sci.*, **133**, jcs.233072.
  116. Maldonado,E., Morales-Pison,S., Urbina,F. and Solari,A. (2021) Molecular and functional characteristics of DNA polymerase beta-like enzymes from Trypanosomatids. *Front. Cell. Infect. Microbiol.*, **11**, 670564.
  117. Damasceno,J.D., Marques,C.A., Black,J., Briggs,E. and McCulloch,R. (2021) Read, write, adapt: challenges and opportunities during kinetoplast genome replication. *Trends Genet.*, **37**, 21–34.
  118. Saha,A., Nanavaty,V.P. and Li,B. (2020) Telomere and subtelomere R-loops and antigenic variation in Trypanosomes. *J. Mol. Biol.*, **432**, 4167–4185.
  119. Kim,H.S. (2021) Genetic interaction between site-specific epigenetic marks and roles of H4v in transcription termination in *Trypanosoma brucei*. *Front. Cell Dev. Biol.*, **9**, 744878.
  120. Reynolds,D., Hofmeister,B.T., Cliffe,L., Alabady,M., Siegel,T.N., Schmitz,R.J. and Sabatini,R. (2016) Histone H3 variant regulates RNA polymerase II transcription termination and dual strand transcription of siRNA loci in *Trypanosoma brucei*. *PLoS Genet.*, **12**, e1005758.
  121. Faria,J., Glover,L., Hutchinson,S., Boehm,C., Field,M.C. and Horn,D. (2019) Monoallelic expression and epigenetic inheritance sustained by a *Trypanosoma brucei* variant surface glycoprotein exclusion complex. *Nat. Commun.*, **10**, 3023.
  122. Faria,J., Luzak,V., Muller,L.S.M., Brink,B.G., Hutchinson,S., Glover,L., Horn,D. and Siegel,T.N. (2021) Spatial integration of transcription and splicing in a dedicated compartment sustains monogenic antigen expression in African trypanosomes. *Nat. Microbiol.*, **6**, 289–300.
  123. Lueong,S., Merce,C., Fischer,B., Hoheisel,J.D. and Erben,E.D. (2016) Gene expression regulatory networks in *Trypanosoma brucei*: insights into the role of the mRNA-binding proteome. *Mol. Microbiol.*, **100**, 457–471.
  124. McCulloch,R. and Barry,J.D. (1999) A role for RAD51 and homologous recombination in *Trypanosoma brucei* antigenic variation. *Genes Dev.*, **13**, 2875–2888.
  125. Glover,L., Alsford,S. and Horn,D. (2013) DNA break site at fragile subtelomeres determines probability and mechanism of antigenic variation in African trypanosomes. *PLoS Pathog.*, **9**, e1003260.
  126. Jehi,S.E., Wu,F. and Li,B. (2014) *Trypanosoma brucei* TIF2 suppresses VSG switching by maintaining subtelomere integrity. *Cell Res.*, **24**, 870–885.
  127. Feretaki,M., Pospisilova,M., Valador Fernandes,R., Lunardi,T., Krejci,L. and Lingner,J. (2020) RAD51-dependent recruitment of TERRA lncRNA to telomeres through R-loops. *Nature*, **587**, 303–308.
  128. Wahba,L., Gore,S.K. and Koshland,D. (2013) The homologous recombination machinery modulates the formation of RNA–DNA hybrids and associated chromosome instability. *eLife*, **2**, e00505.
  129. Bhattacharya,D., Sahoo,S., Nagraj,T., Dixit,S., Dwivedi,H.K. and Nagaraju,G. (2022) RAD51 paralogs: expanding roles in replication stress responses and repair. *Curr. Opin. Pharmacol.*, **67**, 102313.
  130. Bonilla,B., Hengel,S.R., Grundy,M.K. and Bernstein,K.A. (2020) RAD51 gene family structure and function. *Annu. Rev. Genet.*, **54**, 25–46.
  131. Sridhara,S.C., Carvalho,S., Grosso,A.R., Gallego-Paez,L.M., Carmo-Fonseca,M. and de Almeida,S.F. (2017) Transcription dynamics prevent RNA-mediated genomic instability through SRPK2-dependent DDX23 phosphorylation. *Cell Rep.*, **18**, 334–343.
  132. Perez-Calero,C., Bayona-Feliu,A., Xue,X., Barroso,S.I., Munoz,S., Gonzalez-Basallote,V.M., Sung,P. and Aguilera,A. (2020) UAP56/DDX39B is a major cotranscriptional RNA–DNA helicase that unwinds harmful R loops genome-wide. *Genes Dev.*, **34**, 898–912.
  133. Song,C., Hotz-Wagenblatt,A., Voit,R. and Grummt,I. (2017) SIRT7 and the DEAD-box helicase DDX21 cooperate to resolve genomic R loops and safeguard genome stability. *Genes Dev.*, **31**, 1370–1381.
  134. Hodroj,D., Recolin,B., Serhal,K., Martinez,S., Tsanov,N., Merhi,A. and Maiorano,D. (2017) An ATR-dependent function for the Ddx19 RNA helicase in nuclear R-loop metabolism. *EMBO J.*, **36**, 1182–1198.
  135. Hodroj,D., Serhal,K. and Maiorano,D. (2017) Ddx19 links mRNA nuclear export with progression of transcription and replication and suppresses genomic instability upon DNA damage in proliferating cells. *Nucleus*, **8**, 489–495.
  136. Ribeiro de Almeida,C., Dhir,S., Dhir,A., Moghaddam,A.E., Sattentau,Q., Meinhart,A. and Proudfoot,N.J. (2018) RNA helicase DDX1 converts RNA G-quadruplex structures into R-loops to promote IgH class switch recombination. *Mol. Cell*, **70**, 650–662.
  137. Li,L., Germain,D.R., Poon,H.Y., Hildebrandt,M.R., Monckton,E.A., McDonald,D., Hendzel,M.J. and Godbout,R. (2016) DEAD Box 1 facilitates removal of RNA and homologous recombination at DNA double-strand breaks. *Mol. Cell. Biol.*, **36**, 2794–2810.
  138. Lin,W.L., Chen,J.K., Wen,X., He,W., Zarceno,G.A., Chen,Y., Chen,S., Paull,T.T. and Liu,H.W. (2022) DDX18 prevents R-loop-induced DNA damage and genome instability via PARP-1. *Cell Rep.*, **40**, 111089.
  139. Saha,S., Yang,X., Huang,S.N., Agama,K., Baechler,S.A., Sun,Y., Zhang,H., Saha,L.K., Su,S., Jenkins,L.M., et al. (2022) Resolution of R-loops by topoisomerase III-beta (TOP3B) in coordination with the DEAD-box helicase DDX5. *Cell Rep.*, **40**, 111067.
  140. Yu,Z., Mersaoui,S.Y., Guitton-Sert,L., Coulombe,Y., Song,J., Masson,J.Y. and Richard,S. (2020) DDX5 resolves R-loops at DNA double-strand breaks to promote DNA repair and avoid chromosomal deletions. *NAR Cancer*, **2**, zcaa028.
  141. Singh,A., Minia,I., Droll,D., Fadda,A., Clayton,C. and Erben,E. (2014) Trypanosome MKT1 and the RNA-binding protein ZC3H11: interactions and potential roles in post-transcriptional regulatory networks. *Nucleic Acids Res.*, **42**, 4652–4668.
  142. Ooi,C.P., Benz,C. and Urbaniak,M.D. (2020) Phosphoproteomic analysis of mammalian infective *Trypanosoma brucei* subjected to heat shock suggests atypical mechanisms for thermotolerance. *J. Proteomics*, **219**, 103735.
  143. Boehler,A., Querido,J.B., Prilepskaja,T., Soufari,H., Simonetti,A., Del Cistia,M.L., Kuhn,L., Ribeiro,A.R., Valasek,L.S. and Hashem,Y. (2020) Structural differences in translation initiation between pathogenic Trypanosomatids and their mammalian hosts. *Cell Rep.*, **33**, 108534.

144. Sadic,M., Schneider,W.M., Katsara,O., Medina,G.N., Fisher,A., Mogulothu,A., Yu,Y., Gu,M., de Los Santos,T., Schneider,R.J., et al. (2022) DDX60 selectively reduces translation off viral type II internal ribosome entry sites. *EMBO Rep.*, **23**, e55218.
145. Geng,N., Hu,T. and He,C. (2022) Identification of DDX60 as a regulator of MHC-I class molecules in colorectal cancer. *Biomedicines*, **10**, 3092.
146. Hopfner,K.P., Gerhold,C.B., Lakomek,K. and Wollmann,P. (2012) Swi2/Snf2 remodelers: hybrid views on hybrid molecular machines. *Curr. Opin. Struct. Biol.*, **22**, 225–233.
147. Joseph,S.A., Tagliatalata,A., Leuzzi,G., Huang,J.W., Cuella-Martin,R. and Ciccia,A. (2020) Time for remodeling: sNF2-family DNA translocases in replication fork metabolism and human disease. *DNA Repair (Amst.)*, **95**, 102943.
148. Geiller,H.E.B., Harvey,A., Jones,R.E., Grimstead,J.W., Cleal,K., Hendrickson,E.A. and Baird,D.M. (2022) ATRX modulates the escape from a telomere crisis. *PLoS Genet.*, **18**, e1010485.
149. Kim,J., Sun,C., Tran,A.D., Chin,P.J., Ruiz,P.D., Wang,K., Gibbons,R.J., Gamble,M.J., Liu,Y. and Oberdoerffer,P. (2019) The macroH2A1.2 histone variant links ATRX loss to alternative telomere lengthening. *Nat. Struct. Mol. Biol.*, **26**, 213–219.
150. Bhargava,R., Lynskey,M.L. and O’Sullivan,R.J. (2022) New twists to the ALternative endings at telomeres. *DNA Repair (Amst.)*, **115**, 103342.
151. Elbakry,A., Juhasz,S., Chan,K.C. and Lohrich,M. (2021) ATRX and RECQ5 define distinct homologous recombination subpathways. *Proc. Nat. Acad. Sci. U.S.A.*, **118**, e2010370118.
152. Juhasz,S., Elbakry,A., Mathes,A. and Lohrich,M. (2018) ATRX Promotes DNA Repair Synthesis and Sister Chromatid Exchange during Homologous Recombination. *Mol. Cell*, **71**, 11–24.
153. Nguyen,D.T., Voon,H.P.J., Xella,B., Scott,C., Clynes,D., Babbs,C., Ayyub,H., Kerry,J., Sharpe,J.A., Sloane-Stanley,J.A., et al. (2017) The chromatin remodelling factor ATRX suppresses R-loops in transcribed telomeric repeats. *EMBO Rep.*, **18**, 914–928.
154. Toubiana,S. and Selig,S. (2018) DNA:RNA hybrids at telomeres - when it is better to be out of the (R) loop. *FEBS J.*, **285**, 2552–2566.
155. Genois,M.M., Plourde,M., Ethier,C., Roy,G., Poirier,G.G., Ouellette,M. and Masson,J.Y. (2015) Roles of Rad51 paralogs for promoting homologous recombination in *Leishmania infantum*. *Nucleic Acids Res.*, **43**, 2701–2715.
156. Damasceno,J.D., Reis-Cunha,J., Crouch,K., Beraldi,D., Lapsley,C., Tosi,L.R.O., Bartholomeu,D. and McCulloch,R. (2020) Conditional knockout of RAD51-related genes in *Leishmania major* reveals a critical role for homologous recombination during genome replication. *PLoS Genet.*, **16**, e1008828.
157. Alsford,S. and Horn,D. (2008) Single-locus targeting constructs for reliable regulated RNAi and transgene expression in *Trypanosoma brucei*. *Mol. Biochem. Parasitol*, **161**, 76–79.
158. Marques,C.A., Ridgway,M., Tinti,M., Cassidy,A. and Horn,D. (2022) Genome-scale RNA interference profiling of *Trypanosoma brucei* cell cycle progression defects. *Nat. Commun.*, **13**, 5326.
159. Glover,L. and Horn,D. (2012) Trypanosomal histone gammaH2A and the DNA damage response. *Mol. Biochem. Parasitol*, **183**, 78–83.
160. Skourti-Stathaki,K. (2022) Detection of R-loop structures by immunofluorescence using the S9.6 monoclonal antibody. *Methods Mol. Biol.*, **2528**, 21–29.
161. Smolka,J.A., Sanz,L.A., Hartono,S.R. and Chedin,F. (2021) Recognition of RNA by the S9.6 antibody creates pervasive artifacts when imaging RNA:DNA hybrids. *J. Cell Biol.*, **220**, e202004079.
162. Figueiredo,L.M., Janzen,C.J. and Cross,G.A. (2008) A histone methyltransferase modulates antigenic variation in African trypanosomes. *PLoS Biol.*, **6**, e161.
163. Lueong,S., Merce,C., Fischer,B., Hoheisel,J.D. and Erben,E.D. (2016) Gene expression regulatory networks in *Trypanosoma brucei*: insights into the role of the mRNA-binding proteome. *Mol. Microbiol.*, **100**, 457–471.