

**Site-specific epigenetic marks in *Trypanosoma brucei* transcription termination, antigenic variation, and proliferation**

Hee-Sook Kim

Public Health Research Institute, New Jersey Medical School, Rutgers, The State University of New Jersey, Newark, NJ 07103, USA

\* Correspondence: [heesook@njms.rutgers.edu](mailto:heesook@njms.rutgers.edu)

Short title: Chromatin, transcription, and replication in *Trypanosoma brucei*

# **Abstract**

In *Trypanosoma brucei*, genes assemble into polycistronic transcription units (PTUs). Transcription termination sites (TTSs) hold deposition sites for three non-essential chromatin factors, histone variants (H3v and H4v) and a DNA modification (base J, a hydroxyl-glucosyl dT). Here, I found that H4v is a major sign for transcription termination at TTSs and readthrough transcription machineries progress until they encounter the next bidirectional transcription start site. While having a secondary function at TTSs, H3v is important for monoallelic transcription of telomeric antigen genes. The simultaneous absence of both histone variants leads to proliferation and replication defects, which are exacerbated by the J deficiency, accompanied by accumulation of sub-G1 population. Base J likely contributes to DNA replication and cell-cycle control. I propose that the coordinated actions of H3v, H4v and J function in concert for cellular fate determination and provide compensatory mechanisms for each other in chromatin organization, transcription, and replication.

## Introduction

*Trypanosoma brucei* is a parasitic protist that causes African sleeping sickness in humans and related diseases in animals, predominantly in sub-Saharan Africa. Disease transmission to mammals occurs by infected tsetse fly bites. To adapt in two different hosts, *T. brucei* undergoes stages of life-cycle specific differentiation. In insect vector, trypanosomes proliferate as a 'procyclic form (PF)' in the midgut and migrate to the fly's salivary gland and differentiate into a non-proliferative metacyclic form. The infected fly injects the metacyclic form *T. brucei* into the mammalian host's bloodstream. After differentiating into a 'bloodstream form (BF)', trypanosomes proliferate in the host's bloodstream and extracellular spaces. BF trypanosomes are transferred during a blood meal and differentiate into a procyclic form and repeat the life cycle<sup>1</sup>. The surface of BF trypanosome is densely coated with a single type of Variant Surface Glycoprotein (VSG), which can trigger a strong antibody response in the mammalian host. But, trypanosomes can escape the host's immune system with their reservoir of over 2,500 VSG genes by sequentially expressing one VSG gene at any given time, which drives disease persistence. Thus, transcriptional control of VSG genes and their genomic arrangement are key to *T. brucei* immune evasion mechanisms.

In *T. brucei*, groups of 10-100 genes are assembled in hundreds of Polycistronic Transcription Units (PTUs). Gene expression in *T. brucei* is controlled in three ways: 1) control of transcription initiation and termination at boundaries of PTUs, 2) processing of precursor RNAs (*trans*-splicing and polyadenylation) that generates mature mRNAs, and 3) regulating individual mRNA using regulatory information present in 3' UTRs and *trans*-acting elements like RNA-binding proteins<sup>2-4</sup>. PTU boundaries are demarcated by specific epigenetic marks in *T. brucei*. Transcription Start Sites (TSSs) feature acetylated H4 (H4K10ac), methylated H3 (H3K4me3), two of essential histone variants (H2Az and H2Bv), and base J ( $\beta$ -D-glucosyl-hydroxymethyluracil, glucosyl-hmU)<sup>5-7</sup>, a *kinetoplastid*-specific DNA modification. Acetylated H2Az is required for H2Az deposition at TSSs. Depletion of HAT1, an acetyltransferase for H4, reduces total mRNA levels by 50%<sup>8</sup>. Deposition of H2Az at TSSs requires GT-rich sequence element present in TSSs<sup>9</sup>.

Transcription termination regulation in *T. brucei* remains less known. While TSSs are associated with essential histones, histone variants, and histone modifying enzymes, Transcription Termination Sites (TTSs) feature three non-essential factors, two of non-

essential histone variants (H3v and H4v) and base J<sup>5,7</sup>. Base J modification uses a two-step process. J-Binding Protein-1 and 2 (JBP1, JBP2) mediate thymidine hydroxylation to generate a hydroxymethyl-dU (hmU) intermediate. J-associated Glucosyl Transferase (JGT) glucosylates hmU to generate base J<sup>10,11</sup>. Deletion of both JBP1 and JBP2 (J $\Delta$ ) abolished base J modification in *T. brucei* but enable normal growth<sup>10</sup>, unlike in *Leishmania* species with essential JBP1 functions<sup>12</sup>. Thus, base J may function differently than in the closely related *kinetoplastid* parasites. Approximately 50% of base J localizes at PTU boundaries of PTUs in *T. brucei*, but, only 1% of base J is chromosome internally primarily at TTSs in *Leishmania*<sup>13–16</sup>. J depletion significantly increased transcription readthrough at TTSs in *Leishmania*<sup>17,18</sup>. However, J null trypanosomes showed minor transcription termination defects with increased antisense transcripts concentrated near TTSs<sup>19,20</sup>. In both organisms, no significant transcription termination defects were observed in the H3v KO single mutant<sup>19,21</sup>. However, antisense transcript levels were higher in trypanosome cells lacking both H3v and J than J null cells<sup>19,20</sup>, indicating that H3v also contributes to transcription termination in *T. brucei*.

The last PTU in some megabase chromosomes houses a VSG, a trypanosome surface antigen gene. The relationship between H3v and J is mirrored in transcription regulation at these telomeric PTUs<sup>19</sup>. Allelic exclusion of VSG among the 2,500 VSG genes and periodic switching of the expressed allele allows trypanosomes to evade the host immune response, a phenomenon known as the antigenic variation. VSG genes occur at four types of chromosomal loci<sup>22–27</sup>: 1) Bloodstream-form Expression Site (BES), 2) Metacyclic Expression Site (MES), 3) Minichromosome (MC), and 4) subtelomeric region. *T. brucei* genome has about 15-20 BESs. A BES, a special telomeric PTU, contains a promoter associated with RNA pol I transcription (not pol II), several of Expression-Site Associated Genes (ESAGs), 70-bp repeats, and a VSG gene located immediately upstream of telomere repeat. There are 6 MESs, each containing a VSG immediately upstream of telomere repeat. MES VSGs are expressed specifically in the metacyclic stage by RNA pol I. *T. brucei* has about 60 small linear chromosomes called minichromosomes (about 30-150kb in size). MCs are organized with an inverted 177-bp repeat in the middle and telomeres at the ends. Some MCs have a VSG located immediately upstream of telomere repeat, but no promoter is present<sup>24</sup>. The remaining VSGs are found in arrays without promoters at subtelomeric regions. Only one BES promoter is transcriptionally active, while the others are repressed to enable monoallelic expression of VSG. A new VSG can be activated by changes in transcriptional status or

genetic rearrangement that moves a silent VSG to the active VSG location (gene conversion)<sup>28</sup>. Trypanosomes coated with a new VSG are undetected by host antibody defenses tuned to prior VSG. H3v and base J reside in telomere repeats in *T. brucei*<sup>5,7</sup>, but only H3vΔ showed increased derepression of telomeric silent VSGs<sup>19</sup>. Level of silent VSG derepression was higher in H3vΔ JΔ double mutant compared to H3vΔ<sup>19</sup>, indicating that VSG silencing requires H3v with minor contribution by J.

In contrast to TSS chromatin marks, TTS marks are non-essential<sup>5,10,19,29</sup>. I hypothesized that H3v, J and H4v will perform redundant functions at TTSs. If this supposition were true, deleting all three marks will induce severe termination defects and result in cell lethality or sickness. Here, our study indicates that H3v, H4v and base J did show synthetic lethal genetic interaction due to uncontrolled cell division with stalled replication, not transcription termination defects. H4v is a major mark of transcription termination but H4vΔ cells grow normally. While having a secondary function at TTSs, H3v is important for monoallelic transcription of VSG genes. The absence of both histone variants promoted cell growth and replication defects, which increased in the absence of base J. Thus, the coordinated actions of H4v, H3v and J in chromatin determine cell fate.

## Results

### Synthetic lethal genetic interaction among TTS-associated chromatin marks

I predicted that successful transcription termination in *T. brucei* requires three non-essential chromatin marks (H3v, H4v, and base J) that coincide at TTSs, which may functionally "signal" RNA polymerase II to stop. All experiments were performed in BF trypanosomes, as J modification is detected only in BF stage. I generated individual single gene knockout (KO) mutants, H3vΔ<sup>19</sup>, H4vΔ, JΔ<sup>19</sup>, using an established Cre-loxP system<sup>30</sup>. Although JΔ was generated by sequentially deleting both alleles of JBP1 and JBP2 genes, this study will refer to them as a single mutant hereafter, as J is one of three TTS marks. Briefly, all parental cell lines contain a Cre recombinase gene whose expression can be induced by adding tetracycline ('Tet-on'). Gene deletion cassettes have upstream and downstream homology sequences and a floxed selection marker in the middle. Expression of Cre by Tet addition can remove floxed selection markers, so that they can be reused. JΔ strain without markers serve to generate H3vΔ JΔ<sup>19</sup> and H4vΔ JΔ strains. Both double

mutants and the H4vΔ mutant grew normally and showed no morphological abnormalities. To generate a H4vΔ H3vΔ mutant, I transfected a H4v KO vector in the H4vΔ/+ heterozygous H3vΔ strain. The six viable clones grew slower than WT. Four clones showed an unusual growth pattern, as cell growth temporarily arrested before growth resumed (Supplementary Fig. 1, arrows). These double KO cells may have experienced replication challenges but recovered through epigenetic adaptations. To avoid any obscurity cause by unknown events occurring during prolonged culturing and analyze the mutant phenotypes immediately after the knockout, I generated a conditional KO mutant, a H4vΔ strain with a floxed H3v-Ty1, termed conditional double KO (DKO). Cre expressed by tetracycline can remove the floxed H3v-Ty1 allele to produce H3vΔ H4vΔ cells (DKO + Tet) (Fig. 1A). No viable clones emerged after deleting the remaining H4v allele in the H4vΔ/+ heterozygous H3vΔ JΔ strain. So, I generated a conditional triple KO mutant, a H4vΔ JΔ strain with a floxed H3v-Ty1, termed conditional triple KO (TKO). Genotypically, TKO is a quadruple null of H3v, H4v, JBP1 and JBP2 genes.

After adding Tet, both DKO and TKO cells exhibited severe growth defects and abnormal cell-cycle distribution (greater in the Tet-treated TKO) (Fig. 1B & 1C). Cell populations containing less than 2C DNA content (sub-G1) significantly increased after removing H3v in the conditional TKO strain (Fig 1C). Levels of VSG3, a silent VSG protein, increased after Tet treatment in DKO and TKO (Fig. 1D). The DKO double mutant grew more poorly than those generated by transfection, suggesting that cells compensate with the loss of H3v and H4v marks in different mechanisms over time.

I observed increased sub-G1 population in *Tb*MCM-BP depleted cells<sup>31,32</sup>, similar to Tet-treated TKO. *Tb*MCM-BP depletion also caused a DNA replication defect<sup>32</sup>. To determine if replication problems triggered these cell-cycle defects in TKO cells, I measured the efficiency of DNA synthesis using bromo-2'-deoxyuridine (BrdU) incorporation. Tet-untreated WT, DKO and TKO showed a normal pattern of BrdU labeling. BrdU incorporating S phase population was decreased in Tet-treated DKO and this reduction was more pronounced in Tet-treated TKO (Fig. 1E, F). I posit that sub-G1 population in Tet-treated TKO cells arise as mutant cells enter cytokinesis before completing DNA replication.

## **H4v drives transcription termination with H3v-J contributions**

JΔ and H3vΔ JΔ mutants showed mild transcription termination defects and wild type growth rates<sup>19,20</sup>. I speculated that these effects reflected the compensatory presence of H4v,

such that simultaneous deletion of H3v, H4v, and J would cause stronger termination defects. So, I performed stranded RNA-seq experiment with triplicated cultures of WT and 9 KO mutants, including J $\Delta$ , H3v $\Delta$ , H4v $\Delta$ , H3v $\Delta$  J $\Delta$ , H4v $\Delta$  J $\Delta$ , DKO without or with Tet treatment, and TKO without or with Tet treatment. Strains were confirmed by PCR and/or western blot and flow cytometry (Supplementary Fig. 2 & 3). Stranded total RNA was prepared by rRNA removal to avoid the loss of antisense transcripts that are not polyadenylated. Sequence reads were analyzed with Bowtie 2<sup>33</sup> and Seqmonk (Babraham Bioinformatics).

Stranded RNA-seq reads were aligned to the Lister 427 genome. 11 megabase chromosomes were analyzed with sliding windows (5 kb bin, 1kb step). Reads Per Million mapped reads (RPM) values were generated separately from forward and reverse reads and fold changes compared to WT were plotted. Chromosome 10 (Fig. 2) and 7 (Supplementary Fig. 5) are shown as examples. TSSs are shown with H4K10ac peaks<sup>5</sup>. Forward reads mapping to reverse PTUs (blue) (transcription moving reverse direction) represent sense transcription, and forward reads mapping to forward PTUs (red) (transcription moving forward) represent antisense transcription. Reverse reads mapping to forward PTUs (red) represent sense and to reverse PTUs represent antisense transcription. Transcript level was unchanged in J $\Delta$  and H3v $\Delta$ , compared to WT (Fig. 2A). H4v $\Delta$  (pink lines) showed an approximately 2-fold increase in antisense transcription throughout chromosome 10. In the region for PTU assembly in head-to-tail (HT) organization (Fig. 2), all PTUs have the same increase in antisense transcript levels. The data indicate that H4v is an important signal for transcription stop. RNA pol II machinery that missed its stop sign progresses until encountering a bidirectional transcription start site, where directionality of RNA pol II should be determined<sup>9,34,35</sup>.

No further increase in antisense transcription was found in H4v $\Delta$  J $\Delta$  compared to H4v $\Delta$  strain (Fig. 2B). This same pattern of antisense-transcription increase occurred in H3v $\Delta$  J $\Delta$  strain (Fig. 2C), but to a lesser degree. Two possible mechanisms could underlie this. First, H4v and H3v-J independently control transcription termination, which predicts that an antisense transcription increase would be higher in triple KO than in H4v $\Delta$  mutant (additive or synergistic). Second, H4v is the major signal for termination with assistance from H3v-J, so H4v $\Delta$  H3v $\Delta$  and H4v $\Delta$  H3v $\Delta$  J $\Delta$  should show similar levels of antisense transcription as the H4v $\Delta$ . As shown in Fig. 2D and 2E, the increase in antisense transcription was comparable between H3v $\Delta$  H4v $\Delta$  (DKO + Tet) and H4v $\Delta$  and between H3v $\Delta$  H4v $\Delta$  J $\Delta$  (TKO + Tet) and H4v $\Delta$  J $\Delta$ . The same pattern occurred along all chromosomes (Supplementary Fig. 5).



Therefore, H4v is the major chromatin mark for transcription termination with secondary contribution from H3v-J.

RPKM values of all, sense, or antisense reads mapping to 8,428 CDSs (excluding those located in subtelomeric region), are presented in Supplementary Fig. 6A. Mean values of antisense transcription increased about 1.44 (H3vΔ JΔ), 1.64 (H4vΔ), 1.75 (DKO + Tet) and 1.77-fold (TKO + Tet) compared to WT, while sense transcription did not significantly change. Location of CDSs with upregulated antisense transcription confirmed that the antisense transcription increase is global and evenly distributed throughout all PTUs (Supplementary Fig. 6B).

### **TTS chromatin marks affect polyadenylation of antisense transcripts**

We demonstrated that antisense transcription was significantly increased at and near TTSs in JΔ in polyA selected stranded RNA-seq experiments<sup>19</sup>. In the rRNA deletion method here, antisense transcription showed minimal alterations in JΔ mutant cells (Fig. 2). This suggests that TTS chromatin marks may be involved also in polyadenylation of antisense RNA. So, I compared WT stranded RNA-seq data prepared by polyA selection and rRNA depletion. Reads mapped to the Lister 427 genome were analyzed with sliding windows (5 kb bin, 1kb step), as in Fig. 2. RPM values for each method were plotted for chromosome 10 (Fig. 3A). Forward reads mapping represents sense transcription for reverse PTUs (blue) and antisense for forward PTUs (red), and vice versa for reverse reads mapping. For sense transcription, I found no change between samples prepared by polyA selection and rRNA depletion, but a higher level of antisense transcripts (about 10-100-fold) in rRNA depleted samples. This result indicates that only a small fraction of antisense transcripts undergoes polyadenylation.

The increase of polyadenylated antisense transcripts (spiking peaks near TTSs: green lines in Fig. 3B) was observed near TTSs in J Δ, consistent with prior results<sup>19</sup>. This was also observed in H4vΔ JΔ and TKO without Tet (Flag-H4v cKO in H3vΔ JΔ) (Fig. 3D, E). Interestingly, Tet-treated TKO cells undergo an increase in polyadenylated antisense transcripts that were not only localized at TTSs, but also at all PTU regions (Fig. 3F). H4vΔ alone did not significantly change the level of polyadenylated antisense transcripts (Fig. 3C). The data suggest that H4v is required for stopping the RNA pol II at TTSs, but all three TTS marks are required to stop the polyadenylation of non-coding RNA produced by readthrough transcription. Transcriptome analysis of 8,428 CDSs showed that antisense transcription was



1.69-fold higher in the triple KO mutant compared to WT (Supplementary Fig. 6C). Overall, I conclude that TTS chromatin marks control transcription termination and also define the junction of coding and non-coding regions within a long precursor RNA by preventing polyadenylation of antisense RNA. Statistical analyses for all transcriptome data are summarized in Supplementary Table 7.

## **TTS-chromatin marks control VSG transcription in a synergistic manner**

While chromosome internal TTSs have all three marks, telomeres are enriched with H3v and base J, but not H4v<sup>5,7,29</sup>. Transcription is repressed at telomeric regions by heterochromatin structure in *T. brucei* as in other eukaryotes<sup>36–38</sup>, and transcription of telomeric VSG utilizes RNA pol I<sup>39</sup>, not pol II. Therefore, I speculated that the genetic requirement for VSG transcription control would be different from that for RNA pol II-mediated transcription control. Four types of VSG locations are depicted in Fig. 4A. A single VSG is transcribed from one BES (telomeric PTUs with pol I promoter in 10-60kb length)<sup>25,26,39,40</sup>. Additional genomic locations that contain VSGs are also telomeric (MES<sup>26,27</sup> or minichromosomes<sup>24</sup>) or subtelomeric, chromosome-internal VSGs (mostly partial genes or pseudogenes). Only one BES is transcriptionally active at a time, while the remaining BESs are silent. Sequence reads were mapped to the VSGnome<sup>26</sup> and LOG<sub>2</sub>(RPKM) values of VSG CDSs were compared between WT and KO mutants (Fig. 4). H4vΔ or H4vΔ JΔ cells exhibited no significant changes in levels of silent BES and MES VSG transcripts (Fig. 4B, C). Mutants lacking H3v showed increased levels of silent BES and MES VSGs like DKO+Tet, TKO+Tet > H3vΔ JΔ > H3vΔ (high to low). Given that H4vΔ and H4vΔ JΔ showed no change in VSG silencing and H4v does not localize at the telomere, the synergistic upregulation of silent VSGs in Tet-treated DKO and TKO cells was puzzling. H4v could bind telomere but are not detectable, or H4v may bind telomeres only in the absence of H3v. Transcription at the entire BES units were also examined. Interestingly, only those genes adjacent to promoter or telomere (VSG) were significantly affected by H3v, H4v and J (Supplementary Fig. 7). Telomere unprotection in Tet-treated DKO or TKO cells may cause disruption of silent BES heterochromatin structure, leading to the loss of promoter and VSG silencing.

Although MC VSGs do not have a promoter, they experience similar regulation by H3v, H4v, and J (Fig. 4D). The mean values of MC VSG transcript levels increased in mutants lacking H3v, in the order of DKO+Tet, TKO+Tet, H3vΔ JΔ and H3vΔ (high to low), though the

underlying mechanism for the derepression of these promoter-less VSGs remains unclear. They may contain 'promoter-like' elements, which become more accessible by RNA pol I, as the heterochromatin structure is disrupted in H3vΔ H4vΔ cells. Chromosome-internal VSGs were also upregulated similarly (Fig. 4E). Statistical analyses for all VSG expression data are in Supplementary Table 9.

Müller et al observed an increased number of switched cells in H3vΔ H4vΔ cells<sup>41</sup>. To determine if VSG switching also increases in Tet-treated DKO and TKO cells, I examined VSG switching using flow cytometry. Because removing H3v is not reversible and Tet-treated conditional KO cells are very sick, accurately measuring the VSG switching rate posed challenges. Therefore, I looked at switching between two VSGs by staining them with antibodies conjugated with different fluorophores, the active VSG2 with Dylight 488 (green) and the silent VSG3 with Dylight 650 (red). VSG3 was chosen because VSG3 is the most frequently selected silent VSGs in switching assays in the SM strain background<sup>19,42–44</sup>. VSG3 expressing cells were added in to VSG2-expressing cells at 0.01%, 0.1%, 1%, and 10%, and cells were stained with anti-VSG2 and anti-VSG3 antibodies and analyzed by flow cytometry. The percent of VSG3 expressing cells mixed matched well with % VSG3 detected (Supplementary Fig. 8A, B). However, mixing of two strains expressing two different VSG coats produced a population of cells that colocalized both antibodies. This population only occurred when two strains were mixed, as they were not detected in 100% VSG2 expressing cells (VSG2 only). I speculate that there might be some cell-to-cell interaction. Using this method, silent VSG3 expression (derepression or switching) was examined in WT, DKO and TKO treated with or without Tet (Supplementary Fig. 8C). The number of double positive population increased in Tet-treated DKO and TKO cells, but I did not detect any obvious VSG3 switcher population in either sample. Because the abnormal cell population accumulated in mutants, I could not gate other VSG switchers that are not VSG3 expressors. However, at least for VSG3, it is clear that switching did not increase. I conclude that the VSG3 protein detected upon Tet induction likely arise from derepression of VSG3 locus, rather than from switching to VSG3.

## Global DNA replication impairment in TKO cells

The triple KO cells showed inefficient DNA replication and abnormal cell division, producing a large population of cells containing no or less than 2C DNA content. TSSs are also sites of *Tb*ORC1 occupancy in *T. brucei*<sup>45</sup>, so transcription must coordinate with

replication progression to prevent genome instability. To investigate global replication initiation at early replicating origins, I performed MFA-seq experiment (Marker Frequency Analysis followed by HT sequencing) in TKO cells, as previously described<sup>32,45,46</sup>.

PI-stained TKO cells treated with tetracycline for 0, 1, and 2 days were sorted by Fluorescence-Activated Cell Sorting (FACS) (supplementary Fig. 9). Genomic DNA was isolated from cells in G1, early S, late S, or G2 phase, and sequenced. Sequence reads were mapped to the Lister 427 reference genome and analyzed with sliding windows (10kb bin, 2.5kb step) (Fig. 5). I re-analyzed WT MFA-seq reads (which were mapped to the *Tb927* reference genome<sup>32</sup>) for the Lister 427 genome. I found that subtelomeres also contained several of strong early firing origins (indicated as arrows in red in WT). Untreated TKO strain showed the same peak pattern as the WT. At day 1 post Tet induction, the peak pattern was the same but with a reduced intensity compared to day 0. Although several of small peaks were visible at some of strong origin sites at day 2 (indicated as arrows in black), no major peaks were detected, indicating a global impairment of DNA replication initiation.

Cell-cycle sorting (supplementary Fig. 9C, D) indicates that early-S phase cells should yield about 30% more DNA compared to G1 cells. While this increased DNA amount was clearly mapped to regions of early replicating origins in WT and untreated TKO cells, in TKO cells treated with Tet for 2 days, it is impossible to identify regions where this extra DNA was generated from, as no peak was detected. One possibility is that DNA synthesis may be initiated promiscuously (e.g., activating more origins including dormant ones) but replication forks progress inefficiently, producing shorter BrdU-labeled strands. Alternatively, unequal chromosome segregation in a replication-defective parent cell may produce daughter cells with one containing more DNA (early-S DNA content) and the other containing less DNA (sub-G1 DNA content).

### **Emergence of surviving cell populations in cells lacking H3v and H4v**

H3vΔ H4vΔ cells generated by conditional KO showed defects in growth, cell-cycle and DNA replication within two days (Fig. 1). Growth stalling occurred in some of H3vΔ H4vΔ clones obtained by transfection (Supplementary Fig. 1). These data suggest that removing H3v and H4v may cause growth arrest initially, but some cells adapt and manage to proliferate. To determine whether Tet-treated DKO cells are inviable or growth arrested, I measured the viability of DKO and TKO cells 2 days after Tet induction. The viability of Tet-treated DKO and TKO mutants was both low (Fig. 6A), indicating that H3vΔ H4vΔ is lethal

primarily (94.4% and 96.9% lethal). Seven survived clones from each strain (DKO survivor, DS and TKO survivor, TS) were analyzed. They grew at different rates (supplementary Fig. 10A). DS clones showed different levels of VSG3 proteins, while TS clones showed similar levels of VSG3 protein but lower than Tet-treated TKO (Fig. 6C). Three of DS and TS clones were analyzed further in stranded RNA-seq (rRNA depleted). Three DS clones showed different patterns of BES VSG expression (Fig. 6D, left). In DS2 and DS3, silent VSG expression levels were lower than Tet-treated DKO replicates, but majority of silent VSGs in DS5 expressed the same level or more than in Tet-treated DKO cells. The level of MES VSG transcripts in DS clones was lower than Tet-treated DKO and was generally uniform. Compared to DS clones, less clonal variation occurred in TS clones in BES or MES VSG expression (Fig. 6D, right). VSG expression profiles of all replicates are summarized in box plots (Supplementary Fig. 11).

Because H4v is the main factor for transcription termination control, I did not suspect major changes in transcription profiles as H4v was already absent in DKO and TKO strains. Interestingly, transcription termination defects appear partially alleviated in DS5 (Fig. 6E, left). In addition, levels of elevated antisense transcripts in TS clones were slightly lower than Tet-treated TKO (Fig. 6E, right). Collectively, I conclude that the loss of H3v, H4v (and base J) generates toxic effects initially but cells that overcome do emerge.

### **H3v and H4v do not reside in the same nucleosome**

To ensure that phenotypes observed from conditional KO strains were due to the H3v removal and not some artefacts, I conducted complementation experiments. I reintroduced an 'un-floxed' WT-H3v allele in the TKO strain (floxed H3v-HA in H4v $\Delta$  J $\Delta$ ) (Fig. 7A). Cre can remove only the floxed H3v-HA allele and not the complementing WT-H3v, as shown in PCR genotyping (Fig. 7B). H3v fully complemented defects in cell growth, cell-cycle progression and VSG silencing defects (Fig. 7C-F), indicating that H3v removal caused the observed phenotypes. WT-JBP1 or JBP2 was also introduced in the TKO strain using a similar approach. Slight growth improvements were observed in JBP2 complemented Tet-treated TKO cells. JBP2 partially rescued the cell-cycle defect (Fig. 7C-E), as the number of sub-G1 cells was significantly reduced by JBP2 introduction. JBP1 also reduced the number of sub-G1 cells, but fewer than JBP2. Introduction of JBP1 or JBP2 can restore base J modification in J $\Delta$  cells at different genomic loci<sup>7,10</sup>. JBP1 stimulated base J modification at PTU boundaries, including TTSs and TSSs, while JBP2 did not<sup>7</sup>. Instead, JBP2 partially restored J

modifications associated at telomere, 70bp repeats, and 177 bp repeats in minichromosomes<sup>10</sup>. I posit that base J (or JBP1/2 protein) is essential to coordinate DNA replication completion and cell division.

The largest subunit of Replication Protein A complex (RPA), RPA1, forms nuclear foci during replication stress and DNA damage in eukaryotes, including *T. brucei*<sup>47–49</sup>. Tet-treated TKO cells displayed nuclear *TbRPA1* foci (Supplementary Fig. 12). Collectively, the data suggest that the loss of H3v and H4v induces replication stress and activates the DNA damage response (DDR) including *TbRPA1* recruitment to the lesions, and cell-cycle arrest in a J dependent manner.

The genetic interaction thus far strongly suggests that H3v and H4v do not reside in the same nucleosome and this was confirmed by co-immunoprecipitation (co-IP) assays using H3v-HA and/or Flag-H4v expressing strains (Supplementary Fig. 13).

## Discussion

The data here indicate that genome-wide transcription termination readthrough neither interferes with DNA replication nor with cell growth in *T. brucei*. H3v and H4v deletion led to problems in DNA synthesis and growth in DKO and TKO mutant strains. However, Tet-treated DKO and TKO cells produced about 5.6% and 3.1% viable clones. Since the high rate suggested another mechanism than genetic suppression, I hypothesized that epigenetic adaptation may underlie this effect. Initial growth defects in Tet-treated DKO and TKO cells could be caused due to unprotection of TTS DNA. In mice, cells with less nucleosomes accumulated more DNA damage and nucleosome-occupied regions had more protection from irradiation induced DNA damage<sup>50</sup>. Therefore, nucleosome-free TTS DNA may become fragile, accumulate DNA lesions, and activate DNA damage response in Tet-treated DKO and TKO cells. To protect TTS DNA, cells may attempt to randomly deposit core H3-H4, perhaps with specific PTMs at TTSSs, most of which are likely unsuccessful. If a cell deposited a functional pair of H3-H4 PTMs capable of stable nucleosome formation at TTSSs, these cells could emerge and survive (Fig. 8A). Several of H3v and H4v PTM are now identified<sup>8</sup>, although their functions are not yet known, as well as chromatin proteins associated at TTSSs, which include readers and writers of histone PTMs (BDF7, PHD2, PHD4, DOT1A)<sup>51</sup>. H4 (or H3) with a specific PTM may associate with H3v (or H4v) for depositing to TTSSs. I conclude



that PTMs of histones and histone variants marking TTSs are critical for regulating transcription termination.

Sub-G1 population increase was more pronounced in Tet-treated TKO, than DKO, indicating that J has a special function in cell-cycle control. In other eukaryotes, activation of cell-cycle checkpoint upon DNA damage or replication stress leads to cell-cycle arrest at G1 or G2/M, or delays in S-phase progression<sup>52</sup>. Thus, accumulation of sub-G1 population in *T. brucei* is an unusual phenotype for a checkpoint defective mutant. Sub-G1 populations can arise through an unequal segregation of chromosome and/or uncoupling of mitosis/cytokinesis and replication completion. In the presence of base J, nucleosome-free TTSs may signal cells to slow down replication and cell-cycle progression to repair the damage. When base J is absent, H3v $\Delta$  H4v $\Delta$  cells may fail to respond to this signal and proceed with mitosis and cytokinesis with incomplete replication and unrepaired DNA, produce daughter cells lacking DNA and induce cell death.

*T. brucei* JBP proteins prompted the discovery of mammalian TET enzymes (Ten Eleven Translocation) that convert 5-methylcytosine (5mC) to hydroxy-5mC (5hmC), similar to JBP proteins converting methyl-U (T) to hmU. TET enzymes are important factors for reprogramming during development and differentiation<sup>53,54</sup>. TET1/2 depleted trophoblast cells showed increased polyploidy due to endoreduplication, defective G2-M transition and centriole separation<sup>55</sup>. Regions adjacent to DNA damage accumulate 5hmC modifications in a TET2 dependent manner in human cells. Mouse ESC lacking TET1-3 enzymes exhibited abnormal chromosome segregation that was further aggravated by aphidicolin induced replication stress<sup>56</sup>. Here, data suggest that base J (JBP1/2) may bridge between chromosome abnormalities and cell-cycle control, which was unnoticeable under normal conditions. Although not identical, the experiments described here demonstrate shared roles between epigenetic modifiers from evolutionarily distant organisms.

H3v, but not H4v, bind telomere repeats<sup>5</sup>. Consistent with this, deletion of H3v, but not H4v, disrupts VSG silencing. If canonical H3-H4 nucleosomes occupy telomeres in H3v $\Delta$  mutant, the same VSG silencing phenotype should occur in the H3v $\Delta$  and H4v $\Delta$  H3v $\Delta$  mutants. However, the synergistic VSG silencing defect observed in Tet-treated DKO indicates that H3v $\Delta$  and Tet-treated DKO cells may have different telomeric chromatin structure. One possibility is that H4v nucleosomes occupy telomeres in the H3v $\Delta$  mutant. In Tet-treated DKO cells, telomeres may be nucleosome free or occupied by H3-H4 nucleosomes that cannot maintain telomeric heterochromatin. Telomere unprotection and

heterochromatin structure disturbance could lead to the loss of telomere silencing and activation of DDR (Fig. 8B). JBP2 expression in J $\Delta$  cells partially restored base J levels at telomere repeats<sup>7,10</sup>, and I found that JBP2 can partially rescue cell-cycle defects in the Tet-treated TKO cells. JBP2 or telomeric J may play critical roles in replication stress and DNA damage induced cell-cycle checkpoint control in *T. brucei*.

Upregulation of antisense transcription also occurs in H4v $\Delta$  and Tet-treated TKO mutants, but the increase in polyadenylated antisense RNA only occurred in Tet-treated TKO cells. Polyadenylation machinery may interact with TTS chromatin to distinguish coding region from non-coding within precursor RNAs, preventing polyadenylation of antisense transcripts. TFIS2-2 (transcription elongation factor) is enriched at TTSs and interacts with PAF1 complex in *T. brucei*<sup>51</sup>. PAF1 contributes to chromatin remodeling, transcription elongation and polyadenylation in yeast and mouse cells<sup>57,58</sup>. Interestingly, PAF1 complex interacts with JBP3 in *Leishmania tarentolae*<sup>59</sup>. JBP3 is a J-binding protein 3 recently identified in *T. brucei* and *L. tarentolae*. JBP3 depletion also led to transcription readthrough at TTSs in *T. brucei*<sup>59,60</sup>. In plant, polyadenylation of antisense transcripts affects sense transcript level by interacting with RNA-binding protein, RNA 3' processing factor, histone H3 methyltransferase and demethylase<sup>61</sup>. Future studies will elucidate whether polyadenylated antisense transcripts in the triple KO modulate cell viability and how TTS chromatin proteins contribute to polyadenylation regulation.

## Materials and Methods

### *Trypanosoma brucei* strains and plasmids

Bloodstream form *Trypanosoma brucei* (the Lister 427 antigenic type MITat1.2 clone 221a) was cultured in HMI-9 at 37 °C<sup>62</sup>. All cell lines were constructed in a 'Single Marker (SM)' background, which has a tetracycline-inducible system constitutively expressing T7 RNA polymerase and Tet repressor<sup>63</sup>. CRE gene under Tet Operator control is stably integrated at an rDNA spacer locus and its expression can be induced by tetracycline addition.

Genes encoding JBP1, JBP2, H3v, and H4v were deleted using knockout (KO) vectors. These vectors have gene-specific homology sequence arms that are homologous to upstream or downstream sequences of a target gene. Between upstream and downstream



homology arms, a selection marker flanked by loxP sites is located. A positive selection marker (hygromycin-resistance gene, HYG, or puromycin-resistance gene, PUR) conjugated with *Herpes Simplex Virus* Thymidine Kinase gene (HSVTK or TK) was used. After deleting both alleles of a gene, tetracycline (Sigma-Aldrich) can be added to express Cre to remove the floxed HYG-TK and PUR-TK markers, so these selection markers can be reused to delete the next gene. To generate a conditional null mutant, a 'conditional KO (cKO or floxed)' allele was introduced using a cKO vector, which is the same as a KO vector except that H3v-HA, H3v-Ty1, or Flag-H4v gene is inserted in between of the 5' loxP and a selection marker gene. Cre can remove a region between two loxP sites (including the H3v-HA or Flag-H4v allele), thereby generating cells that no longer express H3v or H4v gene upon tetracycline addition. WT (HSTB-904), JΔ (HSTB-778), H3vΔ (HSTB-881), and H3vΔ JΔ (HSTB-868) were generated previously<sup>19</sup>. In this study, following strains were generated.

H4vΔ strain: One H4v allele in the wild type strain (HSTB-904) was deleted with pKP26 vector containing a H4v-KO-HYG-TK cassette ('H4v-KO' indicates that the cassette contains homology arms for H4v deletion), creating a H4vΔ/+ heterozygote (HSTB-1039). The second H4v allele in HSTB-1039 was deleted with pKP27 containing a H4v-KO-PUR-TK, creating a H4vΔ/Δ homozygote (HSTB-1051). Selection markers in HSTB-1051 were removed by Cre expression (HSTB-1067).

H4vΔ JΔ strain: One H4v allele in the JΔ strain (HSTB-778) was deleted with pKP26 (H4v-KO-HYG-TK), creating a H4vΔ/+ JΔ (HSTB-1041). The second H4v allele in HSTB-1041 was deleted with pKP27 (H4v-KO-PUR-TK), creating a H4vΔ/Δ JΔ (HSTB-1055). Selection markers in HSTB-1055 were removed by Cre expression (HSTB-1069).

H4vΔ H3vΔ strain: One H4v allele in the H3vΔ strain (HSTB-881) was deleted with pKP26 (H4v-KO-HYG-TK), creating a H4vΔ/+ H3vΔ (HSTB-1043). The second H4v allele in HSTB-1043 was deleted with pKP27 (H4v-KO-PUR-TK), creating H4vΔ/Δ H3vΔ strains (HSTB-1260, 1261, 1262, 1263, 1264, and 1265).

An attempt to generate H4vΔ H3vΔ JΔ strain: One H4v allele in the H3vΔ JΔ strain (HSTB-868) was deleted with pKP26 (H4v-KO-HYG-TK), creating H4vΔ/+ H3vΔ JΔ (HSTB-1045). pKP27 (H4v-KO-PUR-TK) was transfected in HSTB-1045. Clones were obtained initially but did not expand further.

Floxed H3v-Ty1 in H4vΔ strain (DKO with floxed H3v-Ty1): One allele of H3vΔ was deleted in the H4vΔ strain (HSTB-1067) with pDS88 (H3v-KO-HYG-TK), creating a H3vΔ/+ H4vΔ strain (HSTB-1074). Cre was induced to remove a HYG-TK selection marker (HSTB-

1078). The remaining H3v allele in the H3v $\Delta$ /+ H4v $\Delta$  strain (HSTB-1078) was replaced with the floxed H3v-Ty1-PUR (pKP33 vector), creating a floxed H3v-Ty1/ $\Delta$  H4v $\Delta$  strain (HSTB-1082, DKO). H3v-Ty1 allele can be removed by tetracycline addition, generating H3v $\Delta$  H4v $\Delta$  double KO cells (Tet-treated DKO).

Floxed H3v-Ty1 in H4v $\Delta$  J $\Delta$  strain (TKO with floxed H3v-Ty1): One allele of H3v $\Delta$  was deleted in the H4v $\Delta$  J $\Delta$  strain (HSTB-1069) with pDS88 (H3v-KO-HYG-TK), creating a H3v $\Delta$ /+ H4v $\Delta$  J $\Delta$  strain (HSTB-1076). Cre was induced to remove a HYG-TK selection marker (HSTB-1080). The remaining H3v allele in the H3v $\Delta$ /+ H4v $\Delta$  J $\Delta$  strain (HSTB-1080) was replaced with the floxed H3v-Ty1-PUR (pKP33 vector), creating a floxed H3v-Ty1/ $\Delta$  H4v $\Delta$  J $\Delta$  strain (HSTB-1088, TKO). H3v-Ty1 allele can be removed by tetracycline addition, generating H3v $\Delta$  H4v $\Delta$  J $\Delta$  triple KO cells (Tet-treated TKO).

Floxed H3v-HA in H4v $\Delta$  J $\Delta$  strain (TKO with floxed H3v-HA): The remaining H3v allele in the H3v $\Delta$ /+ H4v $\Delta$  J $\Delta$  strain (HSTB-1080) was replaced with floxed H3v-HA-PUR-TK (pDS84 vector), creating a floxed H3v-HA/ $\Delta$  H4v $\Delta$  J $\Delta$  strain (HSTB-1142). H3v-HA allele can be removed by tetracycline addition, generating H3v $\Delta$  H4v $\Delta$  J $\Delta$  triple KO cells.

Floxed Flag-H4v in H3v $\Delta$  J $\Delta$  strain (TKO with floxed Flag-H4v): The remaining wild type H4v allele in the H4v $\Delta$ /+ H3v $\Delta$  J $\Delta$  strain (HSTB-1045) was replaced with a floxed Flag-H4v-PUR-TK (pKP49), creating a floxed Flag-H4v/ $\Delta$  H3v $\Delta$  J $\Delta$  strain (HSTB-1098). Flag-H4v allele can be removed by tetracycline addition, generating H4v $\Delta$  H3v $\Delta$  J $\Delta$  triple KO cells.

Floxed H3v-HA and floxed Flag-H4v in J $\Delta$  strain (TKO with floxed H3v-HA and floxed Flag-H4v): Floxed Flag-H4v construct was transfected in floxed H3v-HA/ $\Delta$  H4v $\Delta$ / $\Delta$  and J $\Delta$  strain (HSTB-1142), creating a strain with floxed H3v-HA/ $\Delta$  floxed Flag-H4v/ $\Delta$  J $\Delta$  (HSTB-1246).

TKO strain expressing HA-tagged TbRPA1: One allele of RPA1 was epitope tagged with 3xHA by one-step PCR integration using a pMOTag-3H vector<sup>64</sup>, in the floxed H3v-Ty1 H4v $\Delta$  J $\Delta$  (HSTB-1088), generating HSTB-1093.

Bloodstream form *T. brucei* strains were maintained in HMI-9 media containing necessary antibiotics at the following concentrations: 2.5  $\mu$ g/ml of G418 (Sigma-Aldrich); 5  $\mu$ g/ml blasticidin (InvivoGen); 1  $\mu$ g/ml phleomycin (InvivoGen); 5  $\mu$ g/ml hygromycin (InvivoGen); 0.1  $\mu$ g/ml puromycin (InvivoGen); 1  $\mu$ g/ml tetracycline (Sigma-Aldrich); 35 $\mu$ g/ml ganciclovir (Sigma-Aldrich). *T. brucei* strains, plasmids, and oligonucleotides used in this study are listed in Supplementary Table 1.

## Western blot

5 to 10 million cells were collected and suspended in the Laemmli buffer and separated on an SDS-PAGE gel (Bio-RAD). Following transfer to nitrocellulose membrane (GE Healthcare), the proteins were analyzed using mouse anti-Ty1, mouse anti-HA (Sigma-Aldrich), mouse-anti-Flag (Sigma-Aldrich), mouse anti-VSG3 (Antibody & Bioresource Core Facility, MSKCC), rabbit anti-VSG3, rabbit H3 (Abcam), or mouse anti-tubulin<sup>65</sup> antibodies.

## Flow cytometry analysis for cell-cycle, BrdU pulse, and VSG switching experiments

About 10 million cells of WT, DKO and TKO strain treated or untreated with tetracycline were collected and fixed with ice-cold 70% ethanol. To stain DNA, fixed cells were incubated with 50 µg/ml Propidium Iodide (PI, Sigma-Aldrich) and 200 µg/ml RNase A (Sigma-Aldrich) in PBS (Corning) at 37 °C for 30 min. Cell-cycle progression was analyzed using the LSR II or Via (BD Biosciences), FACSdiva software (BD Biosciences), and FlowJo software package (FlowJo).

BrdU pulse experiments were performed as described previously<sup>66</sup>, with some modifications<sup>32</sup>. BrdU (Sigma-Aldrich) was added to a final concentration of 500 µM to cell cultures and incubated for 40 min. Cells were then fixed with ice-cold 70% ethanol, incubated in 0.1N HCl (Fisher Scientific) with 0.5% Triton X-100 (Fisher Scientific) and washed with PBS containing 0.5% Tween 20 and 1% BSA (Jackson ImmunoResearch). To detect BrdU incorporation, fixed cells were incubated with 1 µg/ml mouse anti-BrdU antibody (Fisher Scientific; BD Pharmingen) at room temperature for 2 hours and incubated with donkey anti-mouse-Alexa 488 (Invitrogen Molecular Probes) at 4 µg/ml for 45 min at room temperature in the dark. Cells were incubated in PBS containing 5 µg/ml PI and 250 µg/ml RNase A for 30 min at room temperature in the dark, and then analyzed by flow cytometry using LSR II or Via (BD Biosciences), FACSdiva software (BD Biosciences), and FlowJo software package (FlowJo).

VSG switching from VSG2 to VSG3 was examined by staining cells with antibodies conjugated with different fluorophores. VSG2 expressing cells containing 0%, 0.01%, 0.1%, 1%, or 10% of VSG3 expressing cells, and Tet-untreated or treated WT, DKO and TKO cells were prepared. 6 million live cells were stained with mouse anti-VSG2 antibody (Antibody & Bioresource Core Facility, MSKCC) conjugated with Dylight 488 (Abcam) and mouse anti-VSG3 antibody (Antibody & Bioresource Core Facility, MSKCC) conjugated with Dylight 650 (Abcam). One million cells were analyzed with Via (BD Biosciences).

## Stranded RNA-seq with rRNA depletion

WT, KO mutants, Tet-untreated DKO and TKO, and Tet-treated (for 2 day) DKO and TKO cells in triplicate, three DS clones and three TS clones were prepared. About 50 million cells were collected and total RNA was extracted using the RNA Stat-60 (Tel-Test) according to the manufacturer's protocol and quantified on a NanoDrop2000c and further cleaned using RNeasy kit (Qiagen). rRNA was removed using the Ribo-Zero kit (Illumina) and stranded RNA-seq libraries were prepared using random hexamer and NEB directional RNA library prep kit (NEB) and sequenced on the NovaSeq 6000 PE150.

Read quality was analyzed using the FastQC program and reads were trimmed using the Trim Galore program from Galaxy (<https://usegalaxy.org>) (parameters: phred 20, stringency 1bp, error 0.1, length 20) and aligned with Bowtie 2<sup>33</sup> to the Lister 427<sup>41</sup>, *Tb927v5* ([www.tritrypdb.org](http://www.tritrypdb.org)), or VSGnome<sup>26</sup>, and analyzed using SeqMonk algorithm (<http://www.bioinformatics.babraham.ac.uk/projects/seqmonk/>) from Babraham Bioinformatics.

To examine transcription profiles at chromosome level, each chromosome was binned at 5 kb resolution with 1 kb step (or 10kb bin, 2.5kb step for DS and TS clone analyses) as described previously<sup>32</sup>. Read Per Million mapped reads (RPM) values from forward or reverse reads only were generated (Supplementary Table 5, Table 11) and fold changes between WT and mutant were analyzed. To compare gene expression between WT and mutants, reads mapping to about 8,428 CDSs were analyzed. Read Per Kilobase per Million mapped reads (RPKM) values were generated from all reads, reads mapping to opposite direction of genes (sense transcription) or reads mapping the same direction as genes (antisense transcription) (Supplementary Table 6). Box plots were generated. To examine VSG expression, trimmed reads were aligned to the VSGnome (retrieved from <http://tryps.rockefeller.edu>)<sup>26</sup>. RPKM values (Supplementary Table 9) were displayed in box plots. Alignment report, correlation between replicates, and PCA analysis are summarized in Supplementary Table 2-4. Statistical analyses for box plots are summarized in Supplementary Table 7 and 9.

## Stranded RNA-seq with poly A selection

WT, JΔ, H4vΔ, H4vΔ JΔ, and Tet-treated or untreated TKO strains were grown in triplicate. About 50 million cells were collected and total RNA was extracted using the RNA

Stat-60 (Tel-Test) according to the manufacturer's protocol and quantified on a NanoDrop2000c and further cleaned using RNeasy kit (Qiagen). RNA-seq libraries were prepared from 500ng of RNA samples using an oligo dT based method with 15 cycles of PCR amplification with the Illumina TruSeq mRNA stranded kit (Illumina) and sequenced on the Illumina HiSeq 2000 v4 (50-bp single end read). Reads were analyzed for transcription profile (Supplementary Table 8) and gene expression as above (Supplementary Table 7). Alignment report, correlation between replicates, and PCA analysis are summarized in Supplementary Table 2-4. Statistical analyses for box plots are summarized in Supplementary Table 7.

## **MFA-seq**

MFA-seq was performed as reported previously<sup>45</sup>, with some modifications<sup>32</sup>. Floxed H3v-HA TKO cells (HSTB-1142) were treated with tetracycline for 0, 1, and 2 days. About 100 million cells at  $\sim 1 \times 10^6$ /ml density were fixed with ice-cold 70% ethanol and DNA stained with 50  $\mu$ g/ml PI in PBS. Cells in G1, early S, late S and G2 phases were FACS-sorted (BD Aria3: BD Biosciences) and genomic DNA was isolated using the QIAamp DNA Blood mini kit (Qiagen). About 10 ng of genomic DNA was fragmented using the NEBNext dsDNA fragmentase (NEB) and sequencing libraries were generated using the NEBNext Ultra DNA Library prep kit (NEB) and NEB multiplex oligos (NEB) for Illumina according to the manufacturer's protocol. Sequencing was performed on an Illumina HiSeq 2000 sequencer (50-bp single end read). Read quality was analyzed as above; quality checked using FastQC, trimmed using Trim galore, aligned with Bowtie 2 to the Lister 427 genome. About 3 ~ 15 million reads were analyzed using the SeqMonk algorithm. To examine replication initiation profiles, chromosomes were binned at 10 kb resolution with 2.5 kb step and RPM values were generated with SeqMonk (Supplementary Table 10). RPM values from the S phase samples were normalized to values from G1 phase and plotted.

## **RPA1 immunofluorescence**

Floxed H3v-Ty1 TKO strain expressing *TbRPA1* tagged with HA (HSTB-1093) was treated with tetracycline for 0, 1, and 2 days. Cells were fixed with 0.5% paraformaldehyde (Thermo Fisher Scientific, Pierce) for 10 minutes, permeabilized with 0.2% NP-40 (Sigma-Aldrich) in PBS, incubated with mouse anti-HA antibody and then with the secondary antibody conjugated with Donkey-anti-mouse Alexa 567 (Invitrogen Molecular Probes). DNA

was stained with 0.5 mg/ml DAPI (Sigma-Aldrich). Images were captured using a Zeiss Axioplan 2 fluorescence microscope and edited with Adobe Photoshop.

### **Co-immunoprecipitation and western blot**

Nuclear extract was prepared from about  $10^8$  cells using Chromatin Extraction kit (Abcam). Extracts were immunoprecipitated with rabbit anti-HA (Sigma) or rabbit anti-Flag (Sigma) antibodies and precipitated proteins were analyzed by western blot using mouse anti-HA (Sigma) or mouse anti-Flag (Sigma), rabbit anti-H3 antibodies (Abcam).

### **Acknowledgements**

This work was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health (grant number R01AI127652 to H.K.). I would like to thank Svetlana Mazel and Songyan Han (Rockefeller University, Flow Cytometry Core) for help with cell sorting; Chingwen Yang (Rockefeller University, Genomics Core) for help with MFA-seq; Dewi Harjanto (Rockefeller University and Neon Labs) for discussions with regard to bioinformatics analysis tools; Nina Papavasiliou & Esteban Erben (DKFZ, Germany), and Genomics & Proteomics Core Facility at DKFZ for help with stranded RNA-seq (poly A); Novogene for NGS library prep and rRNA depleted stranded RNA-seq (Beijing, China); and Brandi Mattson at Life Science Editors for help with manuscript editing.

**Supplementary Information** is available online.

### **Data availability**

MFA-seq and stranded RNA-seq raw sequence files have been deposited to the Sequence Read Archive (accession number PRJNA727846).



## Figure legends

**Figure 1. Genetic interaction between H3v, H4v and base J in proliferation, cell-cycle progression and DNA replication. (A)** Strategy to generate a conditional double KO (DKO, floxed H3v-Ty1 in H4vΔ strain) and a conditional triple KO (TKO, floxed H3v-Ty1 in H4vΔ JΔ strain). The strain has one H3v allele deleted and the other H3v allele replaced with a floxed H3v-Ty1 and PUR marker. Cre protein expressed by tetracycline removes the floxed H3v-Ty1-PUR in H4vΔ or H4vΔ JΔ strain, generating H3vΔ H4vΔ (DKO + Tet) or H3vΔ H4vΔ JΔ (TKO + Tet) cells. **(B)** Cell growth in DKO and TKO mutants upon Tet induction. WT, DKO and TKO strains were treated with tetracycline and cell count was monitored daily in triplicate. **(C)** Cell-cycle progression in DKO and TKO mutants upon Tet induction. Cells from (B) were fixed, stained with Propidium Iodide (PI) and analyzed by flow cytometry. **(D)** Depletion of H3v-Ty1 protein and derepression of silent VSG3 protein in DKO and TKO mutants upon Tet induction. Denatured whole cells from (B) were analyzed by western blot. Tubulin served as a loading control. **(E)** BrdU incorporation in DKO and TKO mutants upon Tet induction. WT, DKO and TKO strains untreated or treated with tetracycline for 2 days were pulse-labeled with 500 μM BrdU for 40 min and fixed. Cells were stained with PI (bulk DNA) and anti-BrdU-Alexa 488, and then analyzed by flow cytometry. **(F)** Quantification of BrdU-positive or negative S phase population and, G1, G2, and sub-G1 cell population. **(G)** Western blot control for experiment E.

**Figure 2. H4v is the major transcription termination signal with a secondary role of H3v-J.** Global transcription was examined by rRNA-depleted stranded RNA-seq and mapping the sequence reads to the *T. brucei* Lister 427 genome. Forward and reverse reads were analyzed separately with sliding windows (5kb bin, 1kb step). Fold changes of RPM values between WT and each of KO mutant are plotted for all chromosomes. Chromosome 10 is shown as an example. Diagram of chromosome 10 is shown on top: PTUs are shown in red and blue bars with arrow heads indicating transcription direction. Dotted lines are regions containing multiple HT PTUs in the same direction. Primary y-axis is a fold change of RPM values of each mutant compared to WT. Secondary y-axis is RPM values obtained from H4K10ac ChIP seq. H4K10ac ChIP raw reads obtained from the Cross lab<sup>5</sup> were mapped to the Lister 427 genome with the Bowtie 2. Following KO mutants are compared in each plot



**(A)** J $\Delta$ , H4v $\Delta$ , and H3v $\Delta$ . **(B)** J $\Delta$ , H4v $\Delta$ , and H4v $\Delta$  J $\Delta$ . **(C)** J $\Delta$ , H3v $\Delta$ , and H3v $\Delta$  J $\Delta$ . **(D)** H3v $\Delta$ , H4v $\Delta$ , and Tet-treated DKO (H3v $\Delta$  H4v $\Delta$ ). **(E)** J $\Delta$ , H4v $\Delta$  J $\Delta$ , Tet-treated TKO (H3v $\Delta$  H4v $\Delta$  J $\Delta$ ).

**Figure 3. TTS-chromatin marks affect polyadenylation of antisense transcripts.** Global transcription was examined by polyA-selected stranded RNA-seq and mapping to the Lister 427 genome. **(A)** Forward and reverse reads were analyzed separately with sliding windows (5kb bin, 1kb step). RPM values from WT RNA-seq prepared by polyA selection (dotted black line) were compared with those from WT rRNA-depleted RNA-seq (solid black line). RPM values were plotted for all chromosomes. Shown chromosome 10 as an example. Diagram of chromosome 10 is depicted in the middle of forward and reverse plots. **(B-F)** Fold changes between wild type and each KO mutant are plotted. Shown chromosome 10: J $\Delta$  (B), H4v $\Delta$  (C), H4v $\Delta$  J $\Delta$  (D), TKO without Tet (E), TKO with Tet (F). Green dotted lines indicate locations of TTSs.

**Figure 4. TTS-chromatin marks are requisite for VSG transcription in a synergistic manner.** VSG RNA levels in WT and KO mutants were analyzed by rRNA-depleted stranded RNA-seq. RNA-seq reads were aligned to the VSGnome database<sup>26</sup> and LOG<sub>2</sub>(RPKM) values were obtained for each VSG in wild type and KO mutants. **(A)** Four types of VSG loci in the genome. BES VSG (telomeric, polycistronically transcribed by RNA pol I), MES VSG (telomeric, monocistronically transcribed by RNA pol I), and MC VSG (telomeric, no promoter). Chromosome-internal VSG in arrays. **(B)** Heat map comparison between WT and KO mutants for BES and MES VSG expression levels (darker color = higher expression levels). Shown: silent VSGs (red) and active VSG2 (green). **(C)** Bar graphs comparing WT and KO mutants for BES and MES VSG expression levels. Mutants are grouped by the loss of VSG silencing phenotype: mutants that showed no change (top) and mutants that increased the expression of silent VSGs (bottom). Error bars indicate standard deviation between three replicates. **(D)** Box plot comparing WT and KO mutants for MC VSG expression levels. **(E)** Box plot comparing WT and KO mutants for chromosome internal VSG expression levels. Statistically analyses are summarized in Supplementary Table 9.

**Figure 5. Global DNA replication impairment in TKO cells.** Replication was examined by MFA-seq. TKO strain with floxed H3v-HA allele was treated with tetracycline for 0, 1 or 2 days and stained with PI. Cells in G1, early S, late S and G2 phase were FACS-sorted

(Supplementary Figure 9). Genomic DNA was prepared and sequenced in the Illumina platform. Sequence reads were aligned to the Lister 427 genome and analyzed with sliding window (10kb bin, 2.5kb step). Read count ratio between early S to G1 was plotted for the whole genome, including subtelomeric regions. Shown: chromosome cores (pink) and subtelomeres (grey) in the chromosome diagram. Several of early replicating origins occur in subtelomeres of chromosome 10, 11, and 9 (red arrows). MFA-seq reads obtained from WT<sup>32</sup> was re-analyzed and mapped to the Lister 427 genome for comparison. Subtelomere 6A of chromosome 6 is lost in the TKO strain.

**Figure 6. Emergence of a surviving cell population lacking H3v and H4v. (A)** Plating efficiency after H3v-Ty1 removal in DKO and TKO strains. WT, DKO and TKO strains were treated with tetracycline for two days to remove H3v-Ty1 allele. 100 or 1000 cells were distributed in 96 well plates and the number of wells containing live cells was counted to determine the viability. Clones that are resistant to puromycin were excluded as they may grow because they still retain the H3v-Ty1-PUR allele. Error bars indicate standard deviation between three replicates. **(B)** Western blot confirming the absence of H3v-Ty1 protein in Tet-treated cells in (A). **(C)** Silent VSG3 protein expression in survivors emerged after removing H3v-Ty1. Seven of survivor clones from DKO (DS 1~7) or TKO (TS 1~7) were examined for silent VSG3 expression by western blot. **(D)** VSG silencing in DS and TS clones was examined by stranded RNA-seq (rRNA depletion). Heat maps compare DKO replicates (-/+ Tet) and DS clones, and TKO replicates (-/+ Tet) and TS clones for BES and MES VSG transcript levels (darker color = higher expression level). **(E)** Global transcription and antisense transcription in DS and TS clones by rRNA-depleted stranded RNA-seq (10kb bin, 2.5kb step). Fold changes compared to WT were plotted and compared with Tet-treated DKO or TKO mutant.

**Figure 7. Full and partial complementation of TKO phenotypes by H3v and JBP1/2. (A)** Strategy to generate complemented cell lines. WT-H3v, JBP1 or JBP2 gene was reintroduced to the H3v-HA TKO strain (floxed H3v-HA in H4vΔ JΔ strain) at the endogenous H3v locus where the H3v was deleted and the selection marker removed. Because the targeting vectors do not have loxP sites, Cre cannot remove the complementing gene, while it removes floxed H3v-HA allele. Resulting cells will express WT-H3v, JBP1 or JBP2 in H3vΔ H4vΔ JΔ background. **(B)** PCR genotyping confirming the removal of the floxed H3v-HA allele

and presence of complemented H3v, JBP1 or JBP2 gene in Tet-untreated or Tet-treated TKO strains. **(C)** Complementation of growth defect. TKO or TKO strain integrated with WT-H3v, JBP1, or JBP2 were treated with tetracycline for 0, 0.5, 1, and 2 days as the cell count was monitored. **(D)** Complementation of cell-cycle defect. Cells treated as in (C) were fixed and stained with PI and analyzed by flow cytometry. **(E)** Graphs showing percentage of G1, G2, S, and cells with abnormal DNA content (e.g., <2C or >4C). **(F)** Complementation of silent VSG3 derepression. H3v-HA and VSG3 protein expression by western blot.

**Figure 8. Roles of H3v, H4v and base J in transcription and genome maintenance. (A)** At TTS, J-modified DNA wraps around nucleosomes containing H3v or H4v. H4v nucleosomes signal the RNA pol II machinery to stop at TTSs. Pol II that missed the stop sign continues to progress until it reaches the next divergent TSS. Simultaneous absence of H3v and J affects H4v nucleosome function by changing the TTS chromatin structure. Upon removal of H3v in DKO or TKO cells, nucleosome-free TTS DNA becomes fragile, accumulates replication stress and DNA lesions, and activates DDR. When J is present, nucleosome-free TTSs may signal cells to slow down replication and cell-cycle progression to repair the damage. However, TKO cells without J go through mitosis and cytokinesis with incomplete replication and unrepaired DNA, producing daughter cells lacking DNA. A cell may deposit a functional pair of H3-H4 PTMs. If a survivor clone with new epigenetic marks can emerge, it could form stable nucleosomes and TTS chromatin structure and inherit. **(B)** The synergistic effect caused by H3v and H4v double KO in VSG silencing suggests that H4v nucleosome may bind telomere repeats in the absence of H3v. The loss of both H3v and H4v nucleosomes at telomeric regions could leave telomeres unprotected, leading to chromosome abnormalities, disruption of heterochromatin structure and loss of VSG silencing. Differing degrees of VSG derepression level between BES VSGs in a DS clone and also between DS clones suggests that individual telomere in survivors may not have the same epigenetic code. Partial complementation of cell-cycle defect by JBP1 or 2 suggests that JBP1/2 or base J at TTS and/or telomere may coordinate the completion of replication with cell-cycle progression.

# References

1. Matthews, K. R. The developmental cell biology of *Trypanosoma brucei*. *J. Cell. Sci.* **118**, 283-290 (2005).
2. Martinez-Calvillo, S., Vizuet-de-Rueda, J. C., Florencio-Martinez, L. E., Manning-Cela, R. G. & Figueroa-Angulo, E. E. Gene expression in trypanosomatid parasites. *J Biomed Biotechnol* **2010**, 525241 (2010).
3. Clayton, C. E. Gene expression in Kinetoplastids. *Curr Opin Microbiol* **32**, 46-51 (2016).
4. Clayton, C. Regulation of gene expression in trypanosomatids: living with polycistronic transcription. *Open Biol* **9**, 190072 (2019).
5. Siegel, T. N. et al. Four histone variants mark the boundaries of polycistronic transcription units in *Trypanosoma brucei*. *Genes Dev* **23**, 1063-1076 (2009).
6. Wright, J. R., Siegel, T. N. & Cross, G. A. Histone H3 trimethylated at lysine 4 is enriched at probable transcription start sites in *Trypanosoma brucei*. *Mol Biochem Parasitol* **172**, 141-144 (2010).
7. Cliffe, L. J., Siegel, T. N., Marshall, M., Cross, G. A. & Sabatini, R. Two thymidine hydroxylases differentially regulate the formation of glucosylated DNA at regions flanking polymerase II polycistronic transcription units throughout the genome of *Trypanosoma brucei*. *Nucleic Acids Res* **38**, 3923-3935 (2010).
8. Kraus, A. J. et al. Distinct roles for H4 and H2A.Z acetylation in RNA transcription in African trypanosomes. *Nat Commun* **11**, 1498 (2020).
9. Wedel, C., Förstner, K. U., Derr, R. & Siegel, T. N. GT-rich promoters can drive RNA pol II transcription and deposition of H2A.Z in African trypanosomes. *EMBO J* **36**, 2581-2594 (2017).
10. Cliffe, L. J. et al. JBP1 and JBP2 are two distinct thymidine hydroxylases involved in J biosynthesis in genomic DNA of African trypanosomes. *Nucleic Acids Res* **37**, 1452-1462 (2009).
11. Bullard, W., Lopes da Rosa-Spiegler, J., Liu, S., Wang, Y. & Sabatini, R. Identification of the glucosyltransferase that converts hydroxymethyluracil to base J in the trypanosomatid genome. *J Biol Chem* **289**, 20273-20282 (2014).
12. Genest, P. A. et al. Formation of linear inverted repeat amplicons following targeting of an essential gene in *Leishmania*. *Nucleic Acids Res* **33**, 1699-1709 (2005).
13. van Leeuwen, F. et al.  $\beta$ -d-glucosyl-hydroxymethyluracil is a conserved DNA modification in kinetoplastid protozoans and is abundant in their telomeres. *Proc Natl Acad Sci USA* **95**, 2366-2371 (1998).
14. van Leeuwen, F. et al. Localization of the modified base J in telomeric VSG gene expression sites of *Trypanosoma brucei*. *Genes & Dev* **11**, 3232-3241 (1997).
15. Genest, P. A., Ter Riet, B., Cijssouw, T., van Luenen, H. G. & Borst, P. Telomeric localization of the modified DNA base J in the genome of the protozoan parasite *Leishmania*. *Nucleic Acids Res* **35**, 2116-2124 (2007).
16. Borst, P. & Sabatini, R. Base J: discovery, biosynthesis, and possible functions. *Annu Rev Microbiol* **62**, 235-251 (2008).
17. van Luenen, H. G. et al. Glucosylated hydroxymethyluracil, DNA base J, prevents transcriptional readthrough in *Leishmania*. *Cell* **150**, 909-921 (2012).
18. Reynolds, D. L. et al. Base J represses genes at the end of polycistronic gene clusters in *Leishmania major* by promoting RNAP II termination. *Mol Microbiol* **101**, 559-574 (2016).
19. Schulz, D., Zaringhalam, M., Papavasiliou, F. N. & Kim, H.-S. Base J and H3.V Regulate Transcriptional Termination in *Trypanosoma brucei*. *PLoS Genet* **12**, e1005762 (2016).

20. Reynolds, D. et al. Histone H3 Variant Regulates RNA Polymerase II Transcription Termination and Dual Strand Transcription of siRNA Loci in *Trypanosoma brucei*. *PLoS Genet* **12**, e1005758 (2016).
21. Anderson, B. A. et al. Kinetoplastid-specific histone variant functions are conserved in *Leishmania major*. *Mol Biochem Parasitol* **191**, 53-57 (2013).
22. Borst, P. et al. The control of variant surface antigen synthesis in trypanosomes. *Eur. J. Biochem* **137**, 383-389 (1983).
23. van der Ploeg, L. H. T., Schwartz, D. C., Cantor, C. R. & Borst, P. Antigenic variation in *Trypanosoma brucei* analysed by electrophoretic separation of chromosome-sized DNA molecules. *Cell* **37**, 77-84 (1984).
24. Wickstead, B., Ersfeld, K. & Gull, K. The small chromosomes of *Trypanosoma brucei* involved in antigenic variation are constructed around repetitive palindromes. *Genome Res* **14**, 1014-1024 (2004).
25. Hertz-Fowler, C. et al. Telomeric expression sites are highly conserved in *Trypanosoma brucei*. *PLoS ONE* **3**, e3527 (2008).
26. Cross, G. A., Kim, H. S. & Wickstead, B. Capturing the variant surface glycoprotein repertoire (the VSGnome) of *Trypanosoma brucei* Lister 427. *Mol Biochem Parasitol* **195**, 59-73 (2014).
27. Kolev, N. G., Ramey-Butler, K., Cross, G. A., Ullu, E. & Tschudi, C. Developmental progression to infectivity in *Trypanosoma brucei* triggered by an RNA-binding protein. *Science* **338**, 1352-1353 (2012).
28. Machado, C. R., Augusto-Pinto, L., McCulloch, R. & Teixeira, S. M. R. DNA metabolism and genetic diversity in Trypanosomes. *Mut Res-Reviews in Mut Res.* **612**, 40-57 (2006).
29. Lowell, J. E. & Cross, G. A. M. A variant histone H3 is enriched at telomeres in *Trypanosoma brucei*. *J. Cell Sci.* **117**, 5937-5947 (2004).
30. Kim, H. S., Li, Z., Boothroyd, C. & Cross, G. A. Strategies to construct null and conditional null *Trypanosoma brucei* mutants using Cre-recombinase and loxP. *Mol Biochem Parasitol* **191**, 16-19 (2013).
31. Kim, H. S., Park, S. H., Gunzl, A. & Cross, G. A. MCM-BP is required for repression of life-cycle specific genes transcribed by RNA polymerase I in the mammalian infectious form of *Trypanosoma brucei*. *PLoS ONE* **8**, e57001 (2013).
32. Kim, H. S. Genome-wide function of MCM-BP in *Trypanosoma brucei* DNA replication and transcription. *Nucleic Acids Res* **47**, 634-647 (2019).
33. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-359 (2012).
34. Marquardt, S. et al. A chromatin-based mechanism for limiting divergent noncoding transcription. *Cell* **157**, 1712-1723 (2014).
35. Bagchi, D. N. & Iyer, V. R. The Determinants of Directionality in Transcriptional Initiation. *Trends Genet* **32**, 322-333 (2016).
36. Ottaviani, A., Gilson, E. & Magdinier, F. Telomeric position effect: from the yeast paradigm to human pathologies? *Biochimie* **90**, 93-107 (2008).
37. Janssen, A., Colmenares, S. U. & Karpen, G. H. Heterochromatin: Guardian of the Genome. *Annu Rev Cell Dev Biol* **34**, 265-288 (2018).
38. Cestari, I. & Stuart, K. Transcriptional Regulation of Telomeric Expression Sites and Antigenic Variation in Trypanosomes. *Curr Genomics* **19**, 119-132 (2018).
39. Gunzl, A. et al. RNA polymerase I transcribes procyclin genes and variant surface glycoprotein gene expression sites in *Trypanosoma brucei*. *Eukaryot Cell* **2**, 542-551 (2003).

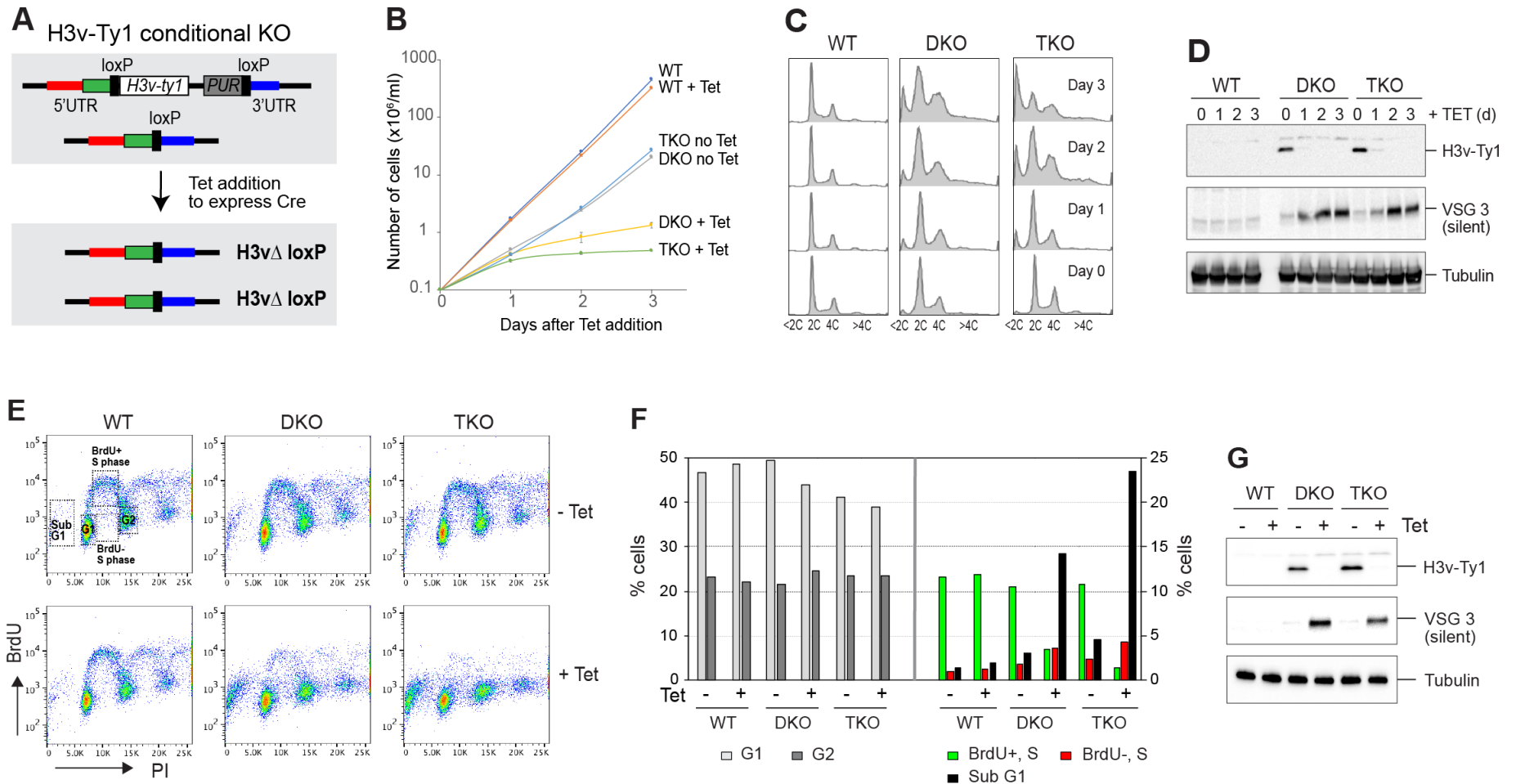


- 877 40. Horn, D. & McCulloch, R. Molecular mechanisms underlying the control of antigenic  
878 variation in African trypanosomes. *Curr Opin Microbiol* **13**, 700-705 (2010).
- 879 41. Müller, L. S. M. et al. Genome organization and DNA accessibility control antigenic  
880 variation in trypanosomes. *Nature* **563**, 121-125 (2018).
- 881 42. Kim, H. S. & Cross, G. A. TOPO3alpha influences antigenic variation by monitoring  
882 expression-site-associated VSG switching in *Trypanosoma brucei*. *PLoS Pathog* **6**,  
883 e1000992 (2010).
- 884 43. Hovel-Miner, G. A. et al. Telomere length affects the frequency and mechanism of  
885 antigenic variation in *Trypanosoma brucei*. *PLoS Pathog* **8**, e1002900 (2012).
- 886 44. Hovel-Miner, G., Mugnier, M. R., Goldwater, B., Cross, G. A. & Papavasiliou, F. N. A  
887 Conserved DNA Repeat Promotes Selection of a Diverse Repertoire of *Trypanosoma*  
888 *brucei* Surface Antigens from the Genomic Archive. *PLoS Genet* **12**, e1005994 (2016).
- 889 45. Tiengwe, C. et al. Genome-wide analysis reveals extensive functional interaction  
890 between DNA replication initiation and transcription in the genome of *Trypanosoma*  
891 *brucei*. *Cell Rep* **2**, 185-197 (2012).
- 892 46. Devlin, R. et al. Mapping replication dynamics in *Trypanosoma brucei* reveals a link with  
893 telomere transcription and antigenic variation. *Elife* **5**, (2016).
- 894 47. Marin, P. A., da Silva, M. S., Pavani, R. S., Machado, C. R. & Elias, M. C. Recruitment  
895 kinetics of the homologous recombination pathway in procyclic forms of *Trypanosoma*  
896 *brucei* after ionizing radiation treatment. *Sci Rep* **8**, 5405 (2018).
- 897 48. Allen, C., Ashley, A. K., Hromas, R. & Nickoloff, J. A. More forks on the road to  
898 replication stress recovery. *J Mol Cell Biol* **3**, 4-12 (2011).
- 899 49. Broderick, S., Rehmet, K., Concannon, C. & Nasheuer, H. P. Eukaryotic single-stranded  
900 DNA binding proteins: central factors in genome stability. *Subcell Biochem* **50**, 143-163  
901 (2010).
- 902 50. Brambilla, F. et al. Nucleosomes effectively shield DNA from radiation damage in living  
903 cells. *Nucleic Acids Res* **48**, 8993-9006 (2020).
- 904 51. Staneva, D. P. et al. A systematic analysis of *Trypanosoma brucei* chromatin factors  
905 identifies novel protein interaction networks associated with sites of transcription  
906 initiation and termination. Preprint at <https://doi.org/10.1101/2021.02.09.430399>(2021).
- 907 52. Waterman, D. P., Haber, J. E. & Smolka, M. B. Checkpoint Responses to DNA Double-  
908 Strand Breaks. *Annu Rev Biochem* **89**, 103-133 (2020).
- 909 53. Pastor, W. A., Aravind, L. & Rao, A. TETonic shift: biological roles of TET proteins in  
910 DNA demethylation and transcription. *Nat Rev Mol Cell Biol* **14**, 341-356 (2013).
- 911 54. Delatte, B., Deplus, R. & Fuks, F. Playing TETris with DNA modifications. *EMBO J* **33**,  
912 1198-1211 (2014).
- 913 55. Chrysanthou, S. et al. A Critical Role of TET1/2 Proteins in Cell-Cycle Progression of  
914 Trophoblast Stem Cells. *Stem Cell Reports* **10**, 1355-1368 (2018).
- 915 56. Kafer, G. R. et al. 5-Hydroxymethylcytosine Marks Sites of DNA Damage and Promotes  
916 Genome Stability. *Cell Rep* **14**, 1283-1292 (2016).
- 917 57. Yang, Y. et al. PAF Complex Plays Novel Subunit-Specific Roles in Alternative  
918 Cleavage and Polyadenylation. *PLoS Genet* **12**, e1005794 (2016).
- 919 58. Penheiter, K. L., Washburn, T. M., Porter, S. E., Hoffman, M. G. & Jaehning, J. A. A  
920 posttranscriptional role for the yeast Paf1-RNA polymerase II complex is revealed by  
921 identification of primary targets. *Mol Cell* **20**, 213-223 (2005).
- 922 59. Jensen, B. C. et al. Chromatin-Associated Protein Complexes Link DNA Base J and  
923 Transcription Termination in *Leishmania*. *mSphere* **6**, (2021).
- 924 60. Kieft, R. et al. Identification of a novel base J binding protein complex involved in RNA  
925 polymerase II transcription termination in trypanosomes. *PLoS Genet* **16**, e1008390  
926 (2020).

61. Fang, X. et al. The 3' processing of antisense RNAs physically links to chromatin-based transcriptional control. *Proc Natl Acad Sci U S A* **117**, 15316-15321 (2020).
62. Hirumi, H. & Hirumi, K. Continuous cultivation of *Trypanosoma brucei* bloodstream forms in a medium containing a low concentration of serum protein without feeder cell layers. *J. Parasitol.* **75**, 985-989 (1989).
63. Wirtz, E., Leal, S., Ochatt, C. & Cross, G. A. M. A tightly regulated inducible expression system for dominant negative approaches in *Trypanosoma brucei*. *Mol Biochem Parasitol.* **99**, 89-101 (1999).
64. Oberholzer, M., Morand, S., Kunz, S. & Seebeck, T. A vector series for rapid PCR-mediated C-terminal in situ tagging of *Trypanosoma brucei* genes. *Mol Biochem Parasitol* **145**, 117-120 (2006).
65. Sherwin, T. & Gull, K. Visualization of dephosphorylation along single microtubules reveals novel mechanisms of assembly during cytoskeletal duplication in trypanosomes. *Cell* **57**, 211-221 (1989).
66. Ryba, T., Battaglia, D., Pope, B. D., Hiratani, I. & Gilbert, D. M. Genome-scale analysis of replication timing: from bench to bioinformatics. *Nat Protoc* **6**, 870-895 (2011).



**Figure 1**



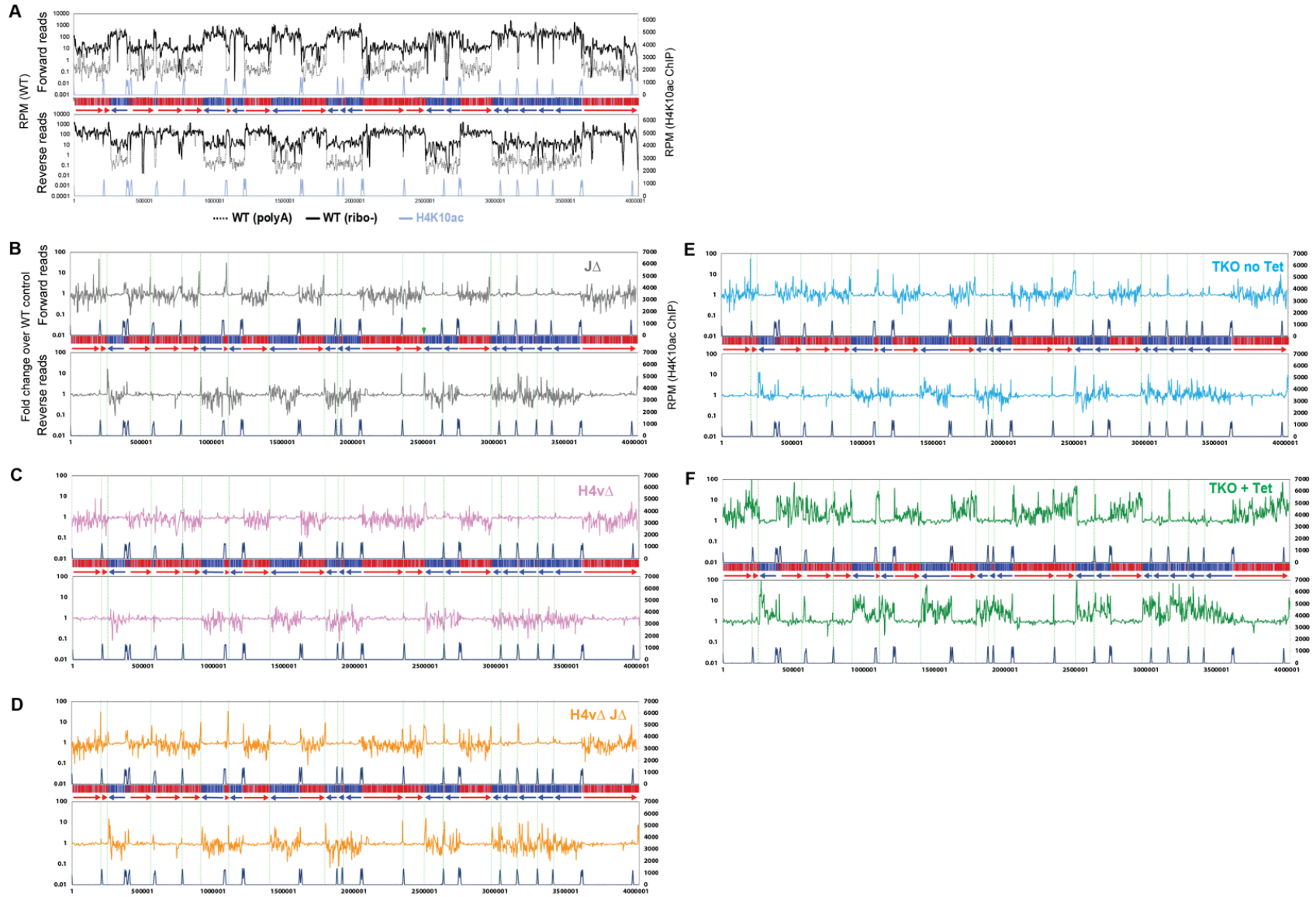
**Figure 1. Genetic interaction between H3v, H4v and base J in proliferation, cell-cycle progression and DNA replication. (A)** Strategy to generate a conditional double KO (DKO, floxed H3v-Ty1 in H4vΔ strain) and a conditional triple KO (TKO, floxed H3v-Ty1 in H4vΔ JΔ strain). The strain has one H3v allele deleted and the other H3v allele replaced with a floxed H3v-Ty1 and PUR marker. Cre protein expressed by tetracycline removes the floxed H3v-Ty1-PUR in H4vΔ or H4vΔ JΔ strain, generating H3vΔ H4vΔ (DKO + Tet) or H3vΔ H4vΔ JΔ (TKO + Tet) cells. **(B)** Cell growth in DKO and TKO mutants upon Tet induction. WT, DKO and TKO strains were treated with tetracycline and cell count was monitored daily in triplicate. **(C)** Cell-cycle progression in DKO and TKO mutants upon Tet induction. Cells from (B) were fixed, stained with Propidium Iodide (PI) and analyzed by flow cytometry. **(D)** Depletion of H3v-Ty1 protein and derepression of silent VSG3 protein in DKO and TKO mutants upon Tet induction. Denatured whole cells from (B) were analyzed by western blot. Tubulin served as a loading control. **(E)** BrdU incorporation in DKO and TKO mutants upon Tet induction. WT, DKO and TKO strains untreated or treated with tetracycline for 2 days were pulse-labeled with 500 μM BrdU for 40 min and fixed. Cells were stained with PI (bulk DNA) and anti-BrdU-Alexa 488, and then analyzed by flow cytometry. **(F)** Quantification of BrdU-positive or negative S phase population and, G1, G2, and sub-G1 cell population. **(G)** Western blot control for experiment E.

**Figure 2**



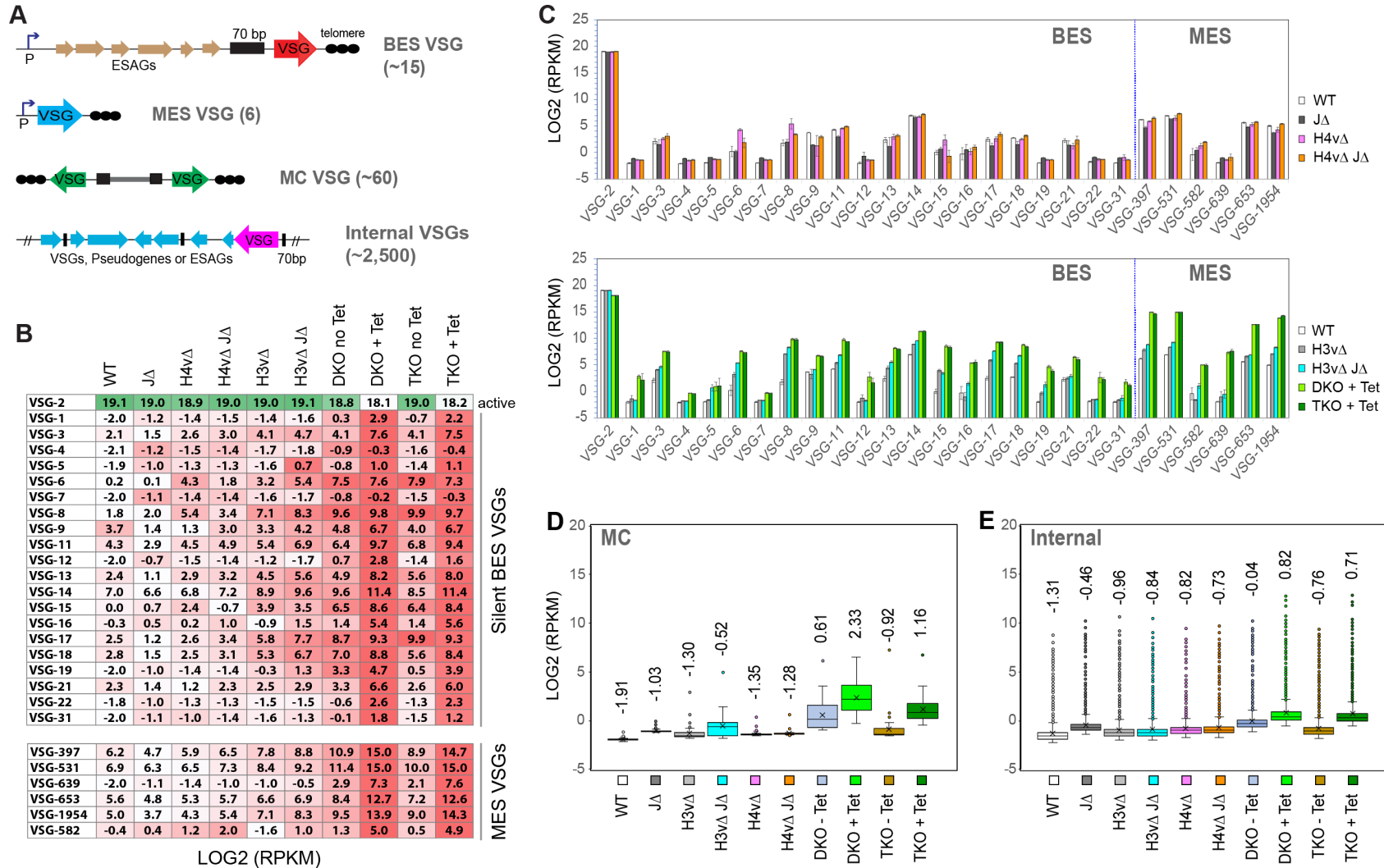
**Figure 2. H4v is the major transcription termination signal with a secondary role of H3v-J.** Global transcription was examined by rRNA-depleted stranded RNA-seq and mapping the sequence reads to the *T. brucei* Lister 427 genome. Forward and reverse reads were analyzed separately with sliding windows (5kb bin, 1kb step). Fold changes of RPM values between WT and each of KO mutant are plotted for all chromosomes. Chromosome 10 is shown as an example. Diagram of chromosome 10 is shown on top: PTUs are shown in red and blue bars with arrow heads indicating transcription direction. Dotted lines are regions containing multiple HT PTUs in the same direction. Primary y-axis is a fold change of RPM values of each mutant compared to WT. Secondary y-axis is RPM values obtained from H4K10ac ChIP seq. H4K10ac ChIP raw reads obtained from the Cross lab<sup>5</sup> were mapped to the Lister 427 genome with the Bowtie 2. Following KO mutants are compared in each plot **(A)** JΔ, H4vΔ, and H3vΔ. **(B)** JΔ, H4vΔ, and H4vΔ JΔ. **(C)** JΔ, H3vΔ, and H3vΔ JΔ. **(D)** H3vΔ, H4vΔ, and Tet-treated DKO (H3vΔ H4vΔ). **(E)** JΔ, H4vΔ JΔ, Tet-treated TKO (H3vΔ H4vΔ JΔ).

**Figure 3**



**Figure 3. TTS-chromatin marks affect polyadenylation of antisense transcripts.** Global transcription was examined by polyA-selected stranded RNA-seq and mapping to the Lister 427 genome. **(A)** Forward and reverse reads were analyzed separately with sliding windows (5kb bin, 1kb step). RPM values from WT RNA-seq prepared by polyA selection (dotted black line) were compared with those from WT rRNA-depleted RNA-seq (solid black line). RPM values were plotted for all chromosomes. Shown chromosome 10 as an example. Diagram of chromosome 10 is depicted in the middle of forward and reverse plots. **(B-F)** Fold changes between wild type and each KO mutant are plotted. Shown chromosome 10: JΔ (B), H4vΔ (C), H4vΔ JΔ (D), TKO without Tet (E), TKO with Tet (F). Green dotted lines indicate locations of TTSs.

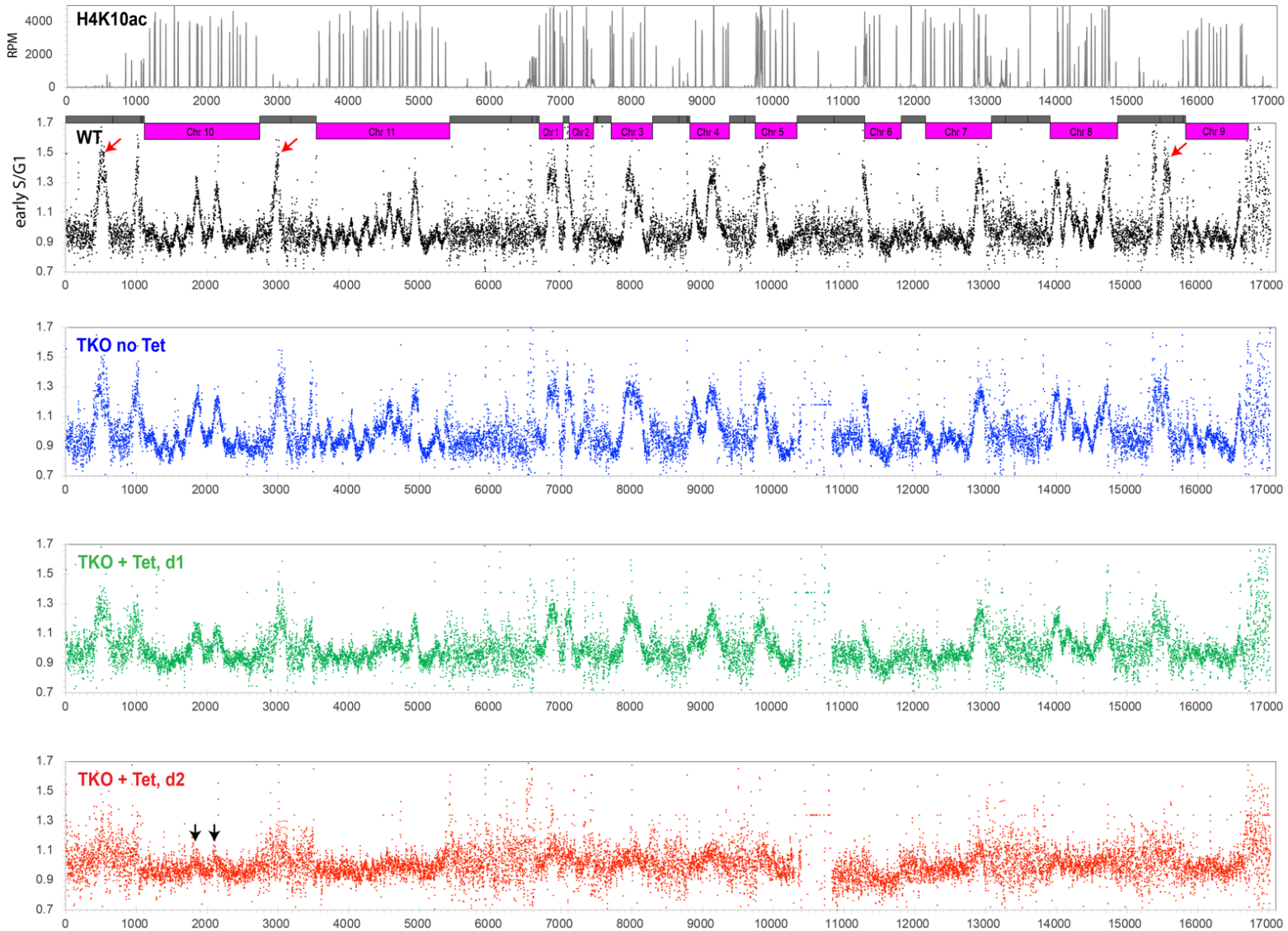
**Figure 4**



**Figure 4. TTS-chromatin marks are requisite for VSG transcription in a synergistic manner.** VSG RNA levels in WT and KO mutants were analyzed by rRNA-depleted stranded RNA-seq. RNA-seq reads were aligned to the VSGnome database<sup>26</sup> and LOG<sub>2</sub>(RPKM) values were obtained for each VSG in wild type and KO mutants. **(A)** Four types of VSG loci in the genome. BES VSG (telomeric, polycistronically transcribed by RNA pol I), MES VSG (telomeric, monocistronically transcribed by RNA pol I), and MC VSG (telomeric, no promoter). Chromosome-internal VSG in arrays. **(B)** Heat map comparison between WT and KO mutants for BES and MES VSG expression levels (darker color = higher expression levels). Shown: silent VSGs (red) and active VSG2 (green). **(C)** Bar graphs comparing WT and KO mutants for BES and MES VSG expression levels. Mutants are grouped by the loss of VSG silencing phenotype: mutants that showed no change (top) and mutants that increased the expression of silent VSGs (bottom). Error bars indicate standard deviation between three replicates. **(D)** Box plot comparing WT and KO mutants for MC VSG expression levels. **(E)** Box plot comparing WT and KO mutants for chromosome internal VSG expression levels. Statistically analyses are summarized in Supplementary Table 9.

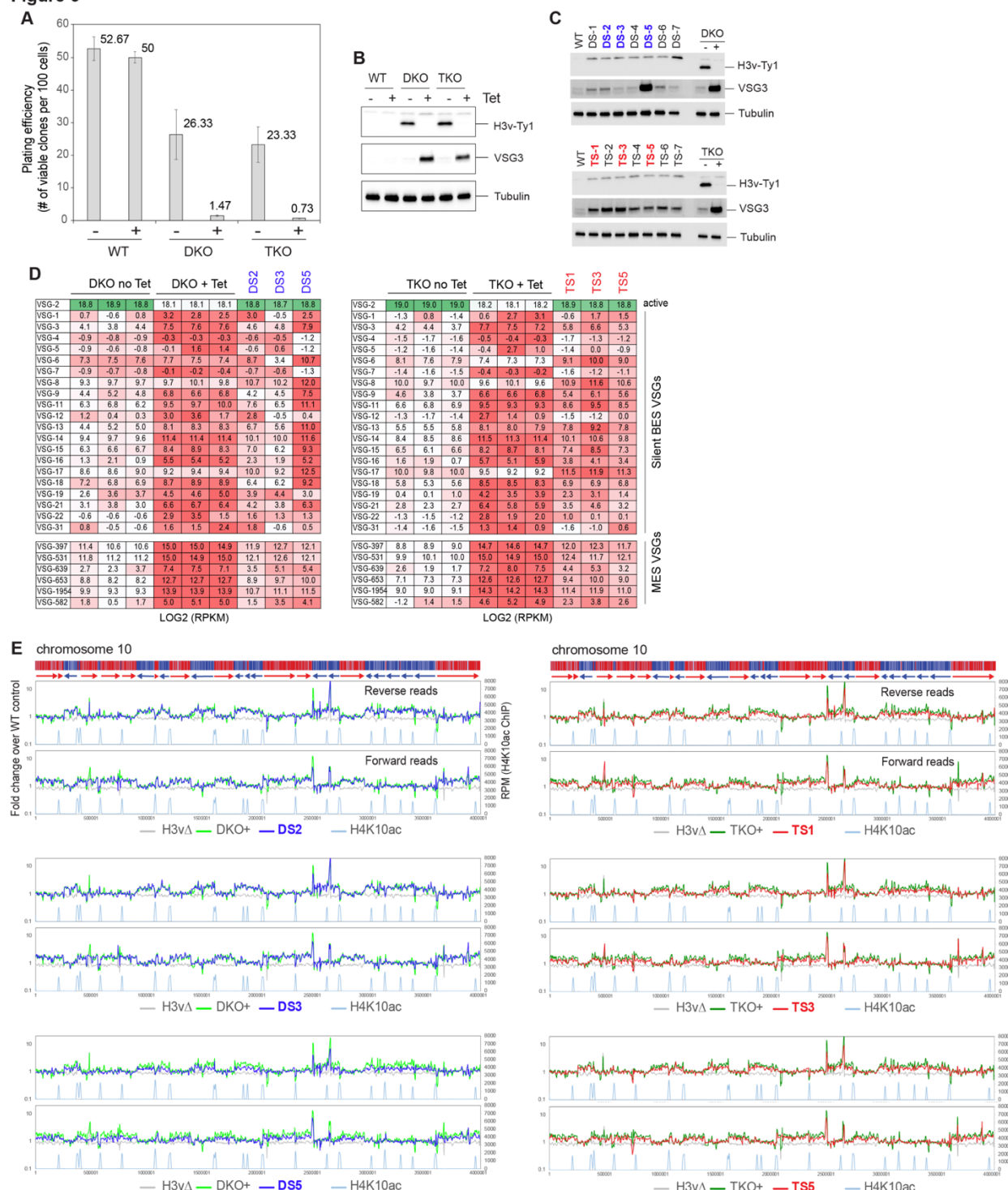


**Figure 5**



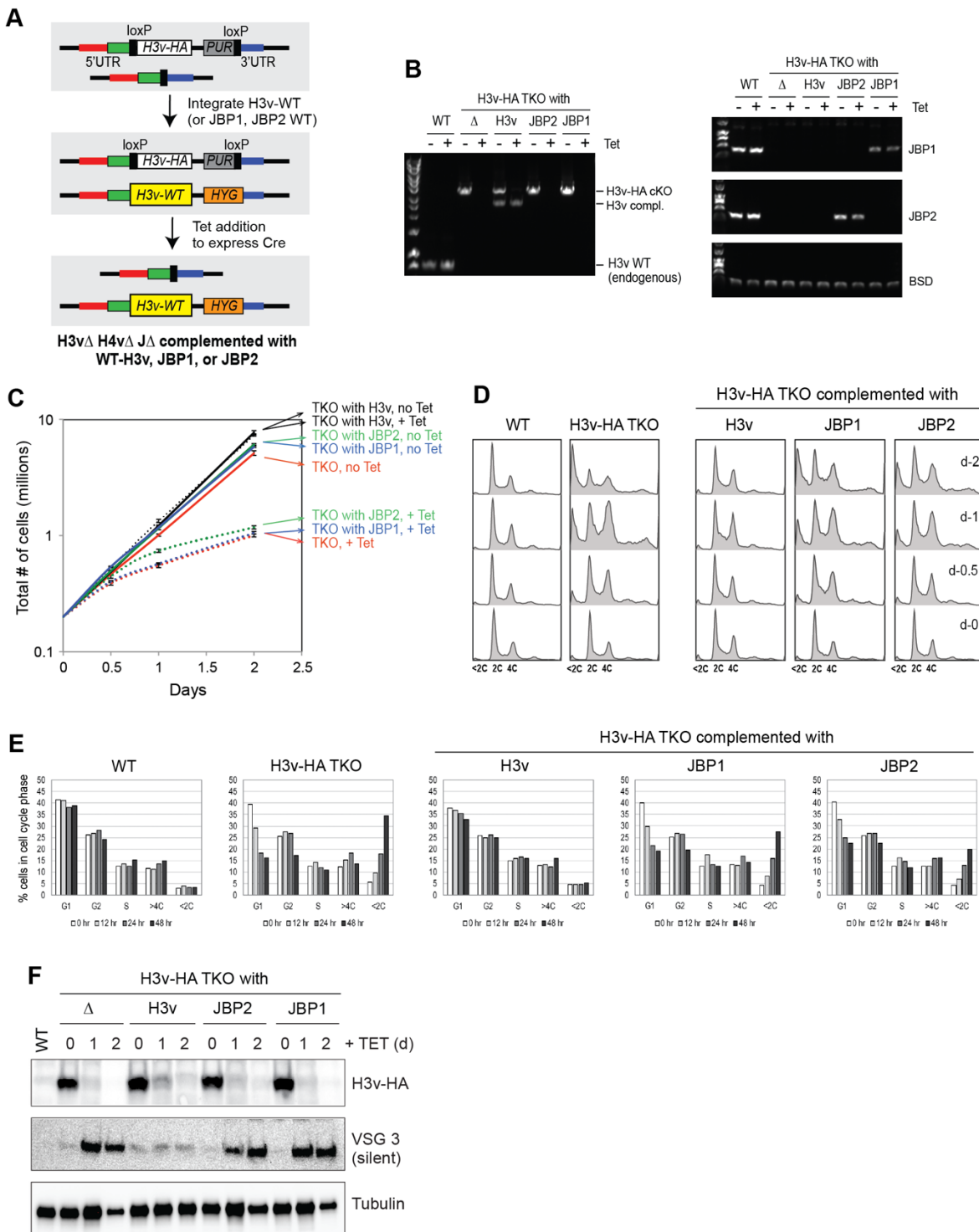
**Figure 5. Global DNA replication impairment in TKO cells.** Replication was examined by MFA-seq. TKO strain with floxed H3v-HA allele was treated with tetracycline for 0, 1 or 2 days and stained with PI. Cells in G1, early S, late S and G2 phase were FACS-sorted (Supplementary Figure 9). Genomic DNA was prepared and sequenced in the Illumina platform. Sequence reads were aligned to the Lister 427 genome and analyzed with sliding window (10kb bin, 2.5kb step). Read count ratio between early S to G1 was plotted for the whole genome, including subtelomeric regions. Shown: chromosome cores (pink) and subtelomeres (grey) in the chromosome diagram. Several of early replicating origins occur in subtelomeres of chromosome 10, 11, and 9 (red arrows). MFA-seq reads obtained from WT<sup>32</sup> was re-analyzed and mapped to the Lister 427 genome for comparison. Subtelomere 6A of chromosome 6 is lost in the TKO strain.

**Figure 6**



**Figure 6. Emergence of a surviving cell population lacking H3v and H4v. (A)** Plating efficiency after H3v-Ty1 removal in DKO and TKO strains. WT, DKO and TKO strains were treated with tetracycline for two days to remove H3v-Ty1 allele. 100 or 1000 cells were distributed in 96 well plates and the number of wells containing live cells was counted to determine the viability. Clones that are resistant to puromycin were excluded as they may grow because they still retain the H3v-Ty1-PUR allele. Error bars indicate standard deviation between three replicates. **(B)** Western blot confirming the absence of H3v-Ty1 protein in Tet-treated cells in (A). **(C)** Silent VSG3 protein expression in survivors emerged after removing H3v-Ty1. Seven of survivor clones from DKO (DS 1~7) or TKO (TS 1~7) were examined for silent VSG3 expression by western blot. **(D)** VSG silencing in DS and TS clones was examined by stranded RNA-seq (rRNA depletion). Heat maps compare DKO replicates (-/+ Tet) and DS clones, and TKO replicates (-/+ Tet) and TS clones for BES and MES VSG transcript levels (darker color = higher expression level). **(E)** Global transcription and antisense transcription in DS and TS clones by rRNA-depleted stranded RNA-seq (10kb bin, 2.5kb step). Fold changes compared to WT were plotted and compared with Tet-treated DKO or TKO mutant.

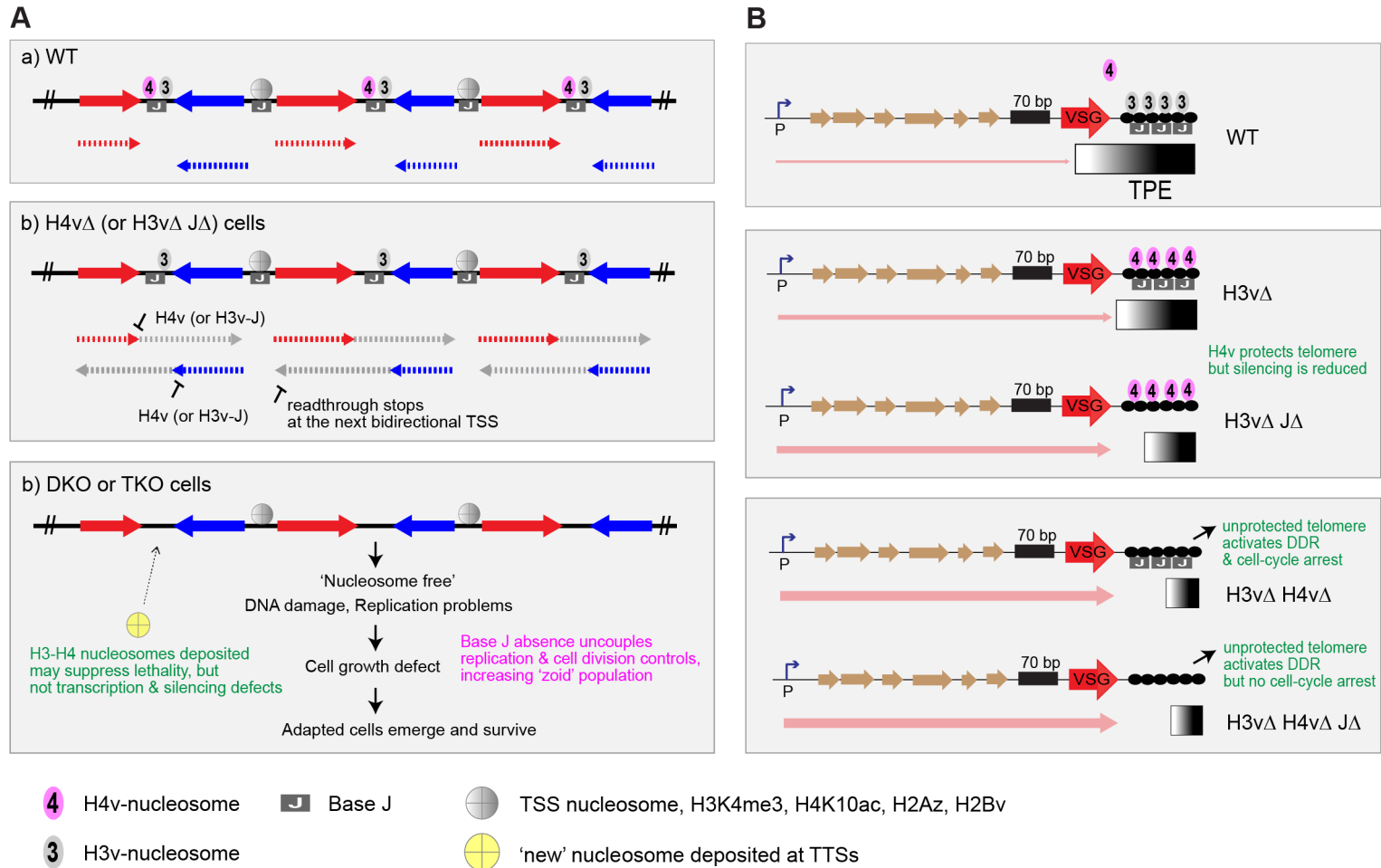
**Figure 7**



**Figure 7. Full and partial complementation of TKO phenotypes by H3v and JBP1/2. (A)** Strategy to generate complemented cell lines. WT-H3v, JBP1 or JBP2 gene was reintroduced to the H3v-HA TKO strain (floxed H3v-HA in H4vΔ JΔ strain) at the endogenous H3v locus where the H3v was deleted and the selection marker removed. Because the targeting vectors do not have loxP sites, Cre cannot remove the complementing gene, while it removes floxed H3v-HA allele. Resulting cells will express WT-H3v, JBP1 or JBP2 in H3vΔ H4vΔ JΔ background. **(B)** PCR genotyping confirming the removal of the floxed H3v-HA allele and presence of complemented H3v, JBP1 or JBP2 gene in Tet-untreated or Tet-treated TKO strains. **(C)** Complementation of growth defect. TKO or TKO strain integrated with WT-H3v, JBP1, or JBP2 were treated with tetracycline for 0, 0.5, 1, and 2 days as the cell count was monitored. **(D)** Complementation of cell-cycle defect. Cells treated as in (C) were fixed and stained with PI and analyzed by flow cytometry. **(E)** Graphs showing percentage of G1, G2, S, and cells with abnormal DNA content (e.g., <2C or >4C). **(F)** Complementation of silent VSG3 derepression. H3v-HA and VSG3 protein expression by western blot.



**Figure 8**



**Figure 8. Roles of H3v, H4v and base J in transcription and genome maintenance. (A)** At TTS, J-modified DNA wraps around nucleosomes containing H3v or H4v. H4v nucleosomes signal the RNA pol II machinery to stop at TTSs. Pol II that missed the stop sign continues to progress until it reaches the next divergent TSS. Simultaneous absence of H3v and J affects H4v nucleosome function by changing the TTS chromatin structure. Upon removal of H3v in DKO or TKO cells, nucleosome-free TTS DNA becomes fragile, accumulates replication stress and DNA lesions, and activates DDR. When J is present, nucleosome-free TTSs may signal cells to slow down replication and cell-cycle progression to repair the damage. However, TKO cells without J go through mitosis and cytokinesis with incomplete replication and unrepaired DNA, producing daughter cells lacking DNA. A cell may deposit a functional pair of H3-H4 PTMs. If a survivor clone with new epigenetic marks can emerge, it could form stable nucleosomes and TTS chromatin structure and inherit. **(B)** The synergistic effect caused by H3v and H4v double KO in VSG silencing suggests that H4v nucleosome may bind telomere repeats in the absence of H3v. The loss of both H3v and H4v nucleosomes at telomeric regions could leave telomeres unprotected, leading to chromosome abnormalities, disruption of heterochromatin structure and loss of VSG silencing. Differing degrees of VSG derepression level between BES VSGs in a DS clone and also between DS clones suggests that individual telomere in survivors may not have the same epigenetic code. Partial complementation of cell-cycle defect by JBP1 or 2 suggests that JBP1/2 or base J at TTS and/or telomere may coordinate the completion of replication with cell-cycle progression.