Published in final edited form as:

Mol Biochem Parasitol. 2013 October ; 191(2): . doi:10.1016/j.molbiopara.2013.09.005.

Kinetoplastid-specific histone variant functions are conserved in *Leishmania major*

Britta A. Anderson^a, Iris L. K. Wong^a, Loren Baugh^b, Gowthaman Ramasamy^b, Peter J. Myler^{b,c,d}, and Stephen M. Beverley^{a,*}

^aDepartment of Molecular Microbiology, Box 8230, Washington University School of Medicine, 600 S. Euclid Ave., St. Louis, MO 63110, USA

^bSeattle Biomedical Research Institute, 307 Westlake Avenue, Seattle, WA 98109-5219, USA

^cDepartment of Biomedical Informatics and Medical Education, University of Washington, Seattle WA 98195, USA

^dDepartment of Global Health, University of Washington, Seattle WA 98195, USA

Abstract

Regions of transcription initiation and termination in kinetoplastid protists lack known eukaryotic promoter and terminator elements, although epigenetic marks such as histone variants and the modified DNA base J have been localized to these regions in *Trypanosoma brucei*, *Trypanosoma cruzi*, and/or *Leishmania major*. Phenotypes of base J mutants vary significantly across trypanosomatids, implying divergence in the epigenetic networks governing transcription during evolution. Here, we demonstrate that the histone variants H2A.Z and H2B.V are essential in *L. major* using a powerful quantitative plasmid segregation-based test. In contrast, H3.V is not essential for viability or normal growth in *Leishmania*. Steady-state transcript levels and the efficiency of transcription termination at convergent strand switch regions (SSRs) in *H3V*-null parasites were comparable to WT parasites. Our genetic tests show a conservation of histone variant phenotypes between *L. major* and *T. brucei*, unlike the diversity of phenotypes associated with genetic manipulation of the DNA base J modification.

Keywords

epigenetics; histone variants; transcriptional read-through; chromatin; trypanosomatid protozoa; genetics

The generation of mature messenger RNAs (mRNAs) in *Leishmania* and other kinetoplastid parasites involves a bipartite mechanism of transcription by RNA polymerase II (RNAP II), unlike the majority of eukaryotes studied to date. All protein-coding genes are transcribed as pre-mRNAs arising from long head-to-tail arrays called polycistronic gene clusters (PGCs), while the RNAs encoding the capped 5' ends of mature transcripts are transcribed separately from the spliced leader (SL) RNA array (reviewed in [1]). Polycistronic pre-mRNAs are then processed by 5' trans-splicing of the SL RNA to generate the capped 5' end of the mRNA

© 2013 Elsevier B.V. All rights reserved.

*Corresponding author: beverley@borcim.wustl.edu, Phone: (314)-747-2630.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

and cleavage and polyadenylation to generate the 3' end. In this system, individual transcription units are mostly defined by the boundaries of PGCs: transcription primarily initiates within divergent strand-switch regions (dSSRs), where two PGCs are oriented head-to-head, and terminates in convergent strand-switch regions (cSSRs), where two PGCs meet tail-to-tail. These regions lack known eukaryotic promoter and terminator elements [2,3,4], and trypanosomatid genomes reveal the presence of general but not sequence-specific RNAP II transcription factors [5]. However, a number of epigenetic marks have been identified in trypanosomatids, including histone variants, histone modifications, and the trypanosomatid-specific DNA modification β -d-glucopyranosyloxymethyluracil (base J), many of which have been mapped to dSSRs or cSSRs in one or more trypanosomatid species [2,6,7,8,9,10]. However, transcription termination and re-initiation may also potentially occur within a PGC, since these chromatin signatures have been found within PGCs [2,6,11].

Despite the apparent conservation of epigenetic marks and their genomic localization amongst trypanosomatid species, recent data suggest that their functions may differ greatly. This is most clearly seen in studies of DNA base J, perturbations of which show widely varying consequences in the three lineages examined thus far. In *T. brucei*, *T. cruzi*, *L. tarentolae*, and *L. major*, base J has been localized to convergent and divergent SSRs as well as telomeres, including the inactive subtelomeric variant surface glycoprotein genes in *T. brucei* [7,8,9,12]. In *T. brucei*, deletion of the genes encoding the thymidine hydroxylases JBP1 and JBP2, which catalyze the first step in base J biosynthesis, generates viable parasites lacking J with no other observable phenotypes or changes in gene expression [13]. In *T. cruzi*, the *JBP1*-/*JBP2*- double null mutant was not viable, while individual *JBP1*- or *JBP2* mutants showed altered transcriptional rates and polymerase occupancy near dSSRs, but normal transcription termination at the cSSRs examined [9]. In *Leishmania tarentolae*, *JBP1* is essential [14], and *JBP2*-null mutants showed massive transcriptional read-through at cSSRs in addition to increased antisense transcription and use of alternative transcription start sites in dSSRs [7].

The evolutionary diversity evident from base J perturbations prompted us to ask whether other epigenetic marks might show functional divergence as well. In *T. brucei*, chromatin immunoprecipitation studies localized H2A.Z and H2B.V to dSSRs and H3.V and H4.V to cSSRs [2], implicating these proteins in regulation of transcription initiation and termination, respectively. Both *H2AZ* and *H2BV* are essential in *T. brucei*, while *H3V* and *H4V* are not [2]. The histone variants H2A.Z, H2B.V, and H3.V are readily identified in *L. major*, although a clear ortholog of H4.V has not been identified yet [5]. Here we focus on the *Leishmania* histone variants H2A.Z, H2B.V, and H3.V, and explore the functional consequences of their genetic inactivation on viability and transcription.

In anticipation that *H2AZ* and *H2BV* would be essential in *Leishmania*, we employed a definitive test for gene essentiality which relies on segregational loss of an episomal complementation vector [15]. In addition to separating the test of gene function from the process of transfection and allelic replacement, this technique allows for screening high number of events rapidly, improving the chances of isolating null mutants whose fitness may be compromised and mutants from loci where homologous recombination is less efficient. Furthermore, when null mutants are not obtained, one has a higher confidence in conclusions concerning the essentiality of the gene of interest [15,16]. We generated heterozygous (*H2AZ*/*HYG* [*pXNG-H2AZ*] and *H2BV*/*SAT* [*pXNG-H2BV*]) and chromosomal-null lines $\Delta h2az$ [*pXNG-H2AZ*] and $\Delta h2bv$ [*pXNG-H2BV*] bearing an episomal complementation vector expressing either *H2AZ* or *H2BV* along with GFP (see Supplemental Data 1 for methods and primer sequences). These lines were grown without selection to permit loss of the episomal complementation vector. We observed that only

0.2% of the $\Delta h2az$ [*pXNG-H2AZ*] were 'GFP-dim' compared to the 33.7% of the *H2AZ/HYG* [*pXNG-H2AZ*] heterozygote line, potentially heralding that this gene is essential (Fig. 1A). Similar results were obtained for $\Delta h2bv$ [*pXNG-H2BV*] (0.1% 'GFP-dim') and *H2BV/PAC* [*pXNG-H2BV*] (10.4% 'GFP-dim') (Supplemental Fig. S1A).

Single cells from the 'GFP-dim' and 'GFP-bright' populations of the chromosomal-null and heterozygote lines were sorted into 96-well microtiter plates and incubated with media until robust growth was seen in control wells. Sorting of the heterozygous lines bearing the episome (*H2AZ/HYG* [*pXNG-H2AZ*] and *H2BV/SAT* [*pXNG-H2BV*]) yielded growth in 70–80% of wells for both the 'GFP-dim' and 'GFP-bright' populations (Fig. 1B and Supplemental Fig. S1B, respectively); this provides a basal measure of cell survival and growth following sorting. As expected, all of the 'GFP-bright' clones retained episomes containing the streptothricin acetyltransferase (*SAT*) or hygromycin B phosphotransferase (*HYG*) markers, while most (80–100%) cells arising from the 'GFP-dim' populations completely lost the episome and became sensitive to the selective antibiotics (Fig. 1B, Supplemental Fig. S1B). Sorting of the 'GFP-bright' cells from the $\Delta h2az$ [*pXNG-H2AZ*] and $\Delta h2bv$ [*pXNG-H2BV*] populations yielded 70–85% growth, comparable to that of the heterozygous control populations. In contrast, only a small fraction (0.6–0.8%) of the 'GFP-dim' cells showed growth following sorting; none of these cells had lost the episome bearing *H2AZ* or *H2BV* as judged by retention of the selectable marker from the episome (Fig. 1B, Supplemental Fig. S1B). From the plating efficiency and numbers of wells tested, we estimated that approximately 610 events were scored in this assay for *H2AZ* and 740 events for *H2BV*, many more than typically screened by traditional non-segregational methods. Thus, we conclude from these experiments that both *H2AZ* and *H2BV* are essential in *L. major*.

In contrast to *H2A.Z* and *H2B.V*, we were able to delete both *H3V* alleles by the standard method of two rounds of allelic replacement, yielding homozygous null mutants ($\Delta h3v$). Colonies were readily obtained from the second round of allelic replacement, and out of six colonies screened five had lost the *H3V* gene. This was shown by the presence of the planned replacements as revealed by PCR using primers flanking and internal to the targeting fragment (data not shown), and the absence of the *H3V* ORF by PCR using primers within the *H3V* ORF (Fig. 2A). The absence of H3.V protein was shown by Western blotting with H3.V-specific antisera (Fig. 2B; see Supplemental Fig. S2 for characterization of antisera). Complemented lines were generated by transfection of an *H3V*-containing episome and showed restoration of H3.V protein levels to levels comparable to WT (Fig. 2B). Since typically episomes are present in multiple copies leading to overexpression of encoded genes, these data suggest the possibility that H3.V levels are regulated at the protein level. Analysis of several clonal $\Delta h3v$ lines showed that they were phenotypically normal, showing WT growth *in vitro* (Fig. S2C). Although we observed a significant increase in the fraction of metacyclic parasites in both $\Delta h3v$ clones, a similar increase was observed in the complementing lines (Fig. S2D). While it is possible that the episomally-encoded *H3V* is regulated differently than chromosomal copies of the gene, the WT H3.V levels in the complemented lines lead us to believe that this phenotype is unrelated to loss of H3.V expression. To identify defects in parasite virulence we inoculated BALB/c mice in the footpad with 10^7 stationary phase parasites from six independent clones. All lines generated lesions within one month, similar to WT *L. major* (data not shown). Together these data demonstrate that deletion of *H3V* does not alter viability of *L. major* in the promastigote stage or significantly impair the infectivity of amastigote stages in murine infections.

To elucidate potential roles for H3.V in transcription in *Leishmania*, we analyzed mRNA levels in WT and $\Delta h3v$ parasites by high-throughput sequencing of spliced leader (SL)-

primed cDNA libraries. This method quantifies steady-state RNA levels in a population of cells by specifically amplifying only transcripts with an SL sequence at their 5' end [7,11,17,18]. Importantly, studies in *L. tarentolae* demonstrate that this approach is also a sensitive method for detecting read-through transcription arising from defects in transcription termination, as these abnormal RNAs can give rise to stable RNAs after processing using cryptic splice acceptor sites (see Fig. 2 in reference [7] for an example arising from base J deficiency in *L. tarentolae*). The sensitivity of detection of both normal 'sense' and cryptic 'antisense' splice acceptors is very high, with ranges in the hundreds of reads per million reads mapped for both 'normal' and 'cryptic' splice acceptors [7].

We focus first on transcriptional read-through, a hallmark of defects in transcriptional termination. As in previous studies in *L. tarentolae* and *T. brucei* [7,11,18], in WT *L. major* the vast majority of SL-containing reads map to the coding strand, with very few mapping to antisense regions beyond cSSRs (see Fig. 2C–D, Supplemental Figs. S3, S4). Remarkably, this pattern was unchanged in $\Delta h3v$ parasites across the parasite genome (Supplemental Fig. S4). This included 'simple' cSSRs (Fig. 2C C, D show two representative examples), cSSRs containing RNA polymerase III-transcribed genes (which are known to suppress transcriptional read-through in the absence of base J in *L. tarentolae* [7]; Supplemental Fig. S3A,B), or the single cSSR known to lack base J in *L. major* (located on chromosome 28; Supplemental Fig. S3C). Quantitative measurement of transcriptional read-through (the antisense-to-sense ratio of reads mapping within 10 kb of a cSSR) shows a very similar distribution in the WT, heterozygous, and $\Delta h3v$ lines (Fig. 2E). These findings are in stark contrast to the results seen in *L. tarentolae* by SL-primed RNA sequencing, where perturbations of base J synthesis in *JBP2dKO* parasites led to high levels of transcriptional read-through [7].

Lastly, we compared mRNA levels by plotting the normalized number of reads mapping to the sense strand of individual *L. major* genes for WT against $\Delta h3v$ parasites. Again, normalized sense transcript levels were remarkably similar between WT and $\Delta h3v$ parasites, with R^2 values >0.96 for two independent $\Delta h3v$ clonal transfectants (Fig. 2F). Examination of all genes containing at least 50 mapped reads in the WT and/or $\Delta h3v$ datasets showed that only two genes showed greater than two-fold differences, occurring in both independent $\Delta h3v$ clonal lines. These genes are located in the middle of PGCs and would appear unlikely candidates to be unaffected by any potential alterations in regulation of transcription initiation or termination. The P27 protein (*Lmj28.0980*), a component of the cytochrome c oxidase complex [19,20], was up-regulated 2.3-fold in both $\Delta h3v$ lines tested relative to WT. In addition, a protein tyrosine phosphatase 1-like protein (*LmjF36.2180*) was up-regulated 2.2–2.3-fold in these lines. This protein has not been characterized to date in *Leishmania* but is an important regulator of cell differentiation in *T. brucei* [21]. Given the absence of detectable phenotypes in $\Delta h3v$ mutants, the significance of these small changes or whether they even result in changes in protein levels is uncertain. Together, these data suggest that H3.V is not required for defining transcriptional stops in *Leishmania* and likely does not play a critical role in controlling mRNA abundance.

While it is often standard practice to translate the functional aspects of epigenetic networks from one system to another based on localization patterns, studies of the hypermodified DNA base J in kinetoplastids demonstrates that assumptions of conserved function based on conserved localization patterns may be incorrect [7,8,9,22]. In our survey of histone variants in *Leishmania*, we found that *H2AZ* and *H2BV* were essential (much like in *T. brucei*), suggesting that their functions are likely conserved. Given their genomic distribution in *T. brucei* and the high degree of *H2A.Z* conservation among all eukaryotes, we suspect that these proteins are integral components of the epigenetic networks controlling transcription initiation. However, the dSSR-associated epigenetic network could differ in *T. cruzi*, as base

J mutants show a transcription initiation-related phenotype [9,22] that is not replicated in in base J mutants in *T. brucei* [8] or *L. tarentolae* [7]. In such a case, H2A.Z and/or H2B.V may play functionally different roles which could differ significantly from eukaryotes studied to date. Elucidation of the effects of H2A.Z/H2B.V incorporation on chromatin compaction and characterization of histone variant incorporation during parasite differentiation will allow us to more specifically define the roles of these proteins in kinetoplastids.

In contrast to H2A.Z and H2B.V, we found that *H3V*-null *L. major* were viable, morphologically normal, and infectious; moreover, they behaved as WT parasites with respect to transcriptional regulation and most interestingly, transcription termination. These data, when interpreted in the light of recent work demonstrating the deleterious effects of perturbation of transcription termination-associated epigenetic networks in *Leishmania* mediated by base J [7], suggests that H3.V is not an essential component of this epigenetic network. Transcription termination-associated phenotypes were not examined in *H3V*-null *T. brucei* [23] and no data exist regarding the essentiality of this protein in *T. cruzi*, so it remains unclear whether this protein is functioning redundantly with other components in the epigenetic network of these parasite species. Although H3.V may not be a critical component of the transcription termination-associated epigenetic networks, chromatin-based studies of H3.V mutants may further define the roles of this conserved, kinetoplastid-specific histone variant.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the Alvin J. Siteman Cancer Center at Washington University School of Medicine and Barnes-Jewish Hospital in St. Louis, Mo., for the use of the Siteman Flow Cytometry Core, which provided the single-cell sorting service. The Siteman Cancer Center is supported in part by NCI Cancer Center Support Grant #P30 CA91842. This work was funded by the National Institutes of Health (R0129646 to SMB and ILKW, T32 GM007067 to BAA) and the Washington University Berg-Morse and Schlesinger Graduate Fellowships to BAA. We thank Andrew Haydock and members of our laboratories for discussions.

References

1. Gunzl, A.; Vanhamme, L.; Myler, P. Transcription in trypanosomes: a different means to the end. In: Barry, J.; Mottram, J.; McCulloch, R.; Acosta-Serrano, A., editors. *Trypanosomes - After the Genome*. Horizon Bioscience. Wymondham; Norfolk, UK: 2007. p. 177-208.
2. Siegel TN, Hekstra DR, Kemp LE, Figueiredo LM, Lowell JE, Fenyo D, et al. Four histone variants mark the boundaries of polycistronic transcription units in *Trypanosoma brucei*. *Genes Dev.* 2009; 23:1063–76. [PubMed: 19369410]
3. Martínez-Calvillo S, Nguyen D, Stuart K, Myler PJ. Transcription Initiation and Termination on *Leishmania major* Chromosome 3. *Eukaryot Cell.* 2004; 3:506–17. [PubMed: 15075279]
4. Martínez-Calvillo S, Yan S, Nguyen D, Fox M, Stuart K, Myler PJ. Transcription of *Leishmania major* Friedlin chromosome 1 initiates in both directions within a single region. *Mol Cell.* 2003; 11:1291–9. [PubMed: 12769852]
5. Ivens AC, Peacock CS, Worthey EA, Murphy L, Aggarwal G, Berriman M, et al. The genome of the kinetoplastid parasite, *Leishmania major*. *Science.* 2005; 309:436–42. [PubMed: 16020728]
6. Thomas S, Green A, Sturm NR, Campbell DA, Myler PJ. Histone acetylations mark origins of polycistronic transcription in *Leishmania major*. *BMC Genomics.* 2009; 10:152. [PubMed: 19356248]

7. Van Luenen HGAM, Farris C, Jan S, Genest P-A, Tripathi P, Velds A, et al. Glucosylated hydroxymethyluracil, DNA base J, prevents transcriptional readthrough in *Leishmania*. *Cell*. 2012; 150:909–21. [PubMed: 22939620]
8. Cliffe LJ, Siegel TN, Marshall M, Cross GAM, 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*. 2010; 38:3923–35. [PubMed: 20215442]
9. Ekanayake DK, Sabatini R. Epigenetic regulation of polymerase II transcription initiation in *Trypanosoma cruzi*: modulation of nucleosome abundance, histone modification, and polymerase occupancy by O-linked thymine DNA glucosylation. *Eukaryot Cell*. 2011; 10:1465–72. [PubMed: 21926332]
10. Wright JR, Siegel TN, Cross GAM. Histone H3 trimethylated at lysine 4 is enriched at probable transcription start sites in *Trypanosoma brucei*. *Mol. Biochem. Parasitol*. 2010; 172:141–4.
11. Kolev NG, Franklin JB, Carmi S, Shi H, Michaeli S, Tschudi C. The transcriptome of the human pathogen *Trypanosoma brucei* at single-nucleotide resolution. *PLoS Pathog*. 2010; 6:e1001090. [PubMed: 20838601]
12. Van Leeuwen F, Wijsman ER, Kieft R, Van Der Marel GA, Van Boom JH, Borst P. Localization of the modified base J in telomeric VSG gene expression sites of *Trypanosoma brucei*. *Genes Dev*. 1997; 11:3232–41. [PubMed: 9389654]
13. Cliffe LJ, Kieft R, Southern T, Birkeland SR, Marshall M, Sweeney K, et al. JBP1 and JBP2 are two distinct thymidine hydroxylases involved in J biosynthesis in genomic DNA of African trypanosomes. *Nucleic Acids Res*. 2009; 37:1452–62. [PubMed: 19136460]
14. Genest P-A, ter Riet B, Dumas C, Papadopoulou B, van Luenen HGAM, Borst P. Formation of linear inverted repeat amplicons following targeting of an essential gene in *Leishmania*. *Nucleic Acids Research*. 2005; 33:1699–709. [PubMed: 15781496]
15. Murta SMF, Vickers TJ, Scott DA, Beverley SM. Methylene tetrahydrofolate dehydrogenase/cyclohydrolase and the synthesis of 10-CHO-THF are essential in *Leishmania major*. *Molecular Microbiology*. 2009; 71:1386–401. [PubMed: 19183277]
16. Feng X, Rodriguez-Contreras D, Polley T, Lye L-F, Scott D, Burchmore RJS, et al. ‘Transient’ genetic suppression facilitates generation of hexose transporter null mutants in *Leishmania mexicana*. *Molecular microbiology*. 2013; 87:412–29. [PubMed: 23170981]
17. Mittra B, Cortez M, Haydock A, Ramasamy G, PJM, Andrews N. Iron uptake controls the generation of *Leishmania* infective forms through regulation of ROS levels. *J Exp. Med*. 2013; 210:401–16.
18. Nilsson D, Gunasekara K, Mani J, Osteras M, Farinelli L, Baerlocher L, et al. Spliced leader trapping reveals widespread alternative splicing patterns in the highly dynamic transcriptome of *Trypanosoma brucei*. *PLoS Pathog*. 2010; 6:e1001037. [PubMed: 20700444]
19. Dey R, Dagur PK, Selvapandiyan A, McCoy JP, Salotra P, Duncan R, et al. Live attenuated *Leishmania donovani* p27 gene knockout parasites are nonpathogenic and elicit long-term protective immunity in BALB/c mice. *J. Immunol*. 2013; 190:2138–49.
20. Dey R, Meneses C, Salotra P, Kamhawi S, Nakhasi HL, Duncan R. Characterization of a *Leishmania* stage-specific mitochondrial membrane protein that enhances the activity of cytochrome c oxidase and its role in virulence. *Mol. Microbiol*. 2010; 77:399–414.
21. Szö r B, Wilson J, McElhinney H, Taberner L, Matthews KR. Protein tyrosine phosphatase TbPTP1: a molecular switch controlling life cycle differentiation in trypanosomes. *J Cell Biol*. 2006; 175:293–303. [PubMed: 17043136]
22. Ekanayake DK, Minning T, Weatherly B, Gunasekera K, Nilsson D, Tarleton R, et al. Epigenetic regulation of transcription and virulence in *Trypanosoma cruzi* by O-linked thymine glucosylation of DNA. *Mol. Cell. Biol*. 2011; 31:1690–1700.
23. Lowell JE, Cross GAM. A variant histone H3 is enriched at telomeres in *Trypanosoma brucei*. *J. Cell Sci*. 2004; 117:5937–47.
24. Thorvaldsdottir H, Robinson J, Mesirov J. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform*. 2013; 14:178–92. [PubMed: 22517427]

25. Späth GF, Beverley SM. A lipophosphoglycan-independent method for isolation of infective *Leishmania* metacyclic promastigotes by density gradient centrifugation. *Experimental parasitology*. 2001; 99:97–103. [PubMed: 11748963]

- Histone variants H2A.Z and H2B.V are essential in *Leishmania major*
- Histone variant H3.V is not essential in *L. major*
- H3.V loss does not affect read-through transcription, gene expression, viability nor virulence in a mouse model
- Genetic tests suggest that the roles of epigenetic networks can be conserved (histone variants) or can diverge (DNA base J) during evolution

Figure 1. *H2AZ* is essential in *L. major*

(A) Quantitation of pXNG-H2AZ levels by GFP flow cytometry following removal of nourseothricin selection. The dark gray shaded regions represent GFP fluorescence of the experimental lines, and the dotted line shows GFP fluorescence values from WT *L. major* FV1. Light gray shaded regions represent FACS gates used for recovery of ‘GFP-dim’ (left shaded region) and ‘GFP-bright’ (right shaded region) cells; parasites with a GFP fluorescence signal of 1 or less were not included in the ‘GFP-dim’ gate. The lines *H2AZ/HYG*[pXNG-H2AZ] (left panel) and *Δh2az*[pXNG-H2AZ] (right panel) were grown for 48 hours (~12 cell doublings) in the absence of nourseothricin to allow loss of the episome before GFP fluorescence was analyzed. Boxes show percent of parasites classified as ‘GFP-bright’ or ‘GFP-dim’. (B) Single cells from ‘GFP-dim’ and ‘GFP-bright’ *H2AZ/HYG*[pXNG-H2AZ] and *Δh2az*[pXNG-H2AZ] were sorted into 96-well plates containing supplemented Schneiders’ medium (see Materials and Methods). Boxes show the percentage of wells scored for robust growth after two weeks of incubation at 26°C; numbers in

parentheses represent the total number of cells sorted (total from two independent experiments). For these, retention of pXNG-H2AZ was tested by growth in the presence of nourseothricin, conferred by the plasmid *SAT* marker. Boxes show the percentage of cells demonstrating nourseothricin resistance; numbers in parentheses represent the total number of wells subjected to nourseothricin resistance testing.

Figure 2. Deletion of H3.V does not increase read-through transcription as observed by SL-RNA-seq

(A) Deletion of *H3V* was shown by PCR analysis using *H3V* ORF primers located as depicted in the upper figure in this panel. (B) Loss of H3.V expression shown by western blotting using anti-H3.V antisera. The migration position of H3.V is shown, as is a nonspecific band evident in all samples. The nonspecific band does not arise from cross-reactivity with H3 (Fig. S1). (C–D) Integrative Genomics Viewer [IGV; 24] screenshots demonstrating SL-RNA-seq coverage across ‘simple’ cSSRs. The Y-axes represent normalized read counts (per million reads mapped) and the X-axis represents physical location on each chromosome; a 20 kb window showing 10 kb flanking the transcription termination site (TTS) is shown (C: Chromosome 4, 118,903–138,903 bp; D: Chromosome 7, 49,636–69,636 bp). Unlike random RNA-seq reads, SL-RNA-seq results in clustering of

reads on a limited number of splice acceptor sites (regardless of whether they are 'normal' or 'cryptic' [7]). (E) Quantitative analysis of transcription termination assessed by SL-RNA-seq. Following previous studies [7], TTS within cSSRs were defined using the midpoint of base J 'peaks' associated with TTS; reads mapping to the 'sense' and 'antisense' strand within 10 kb of the TTS were quantitated and the ratio of antisense to sense reads is shown by a box plot. The middle line represents the median, while the box represents the 25th through 75th percentiles. Whiskers represent the 10th through 90th percentiles, and dots represent individual cSSRs which lie below the 10th or above the 90th percentile. (F) Total mRNA levels quantitated by SL-RNA-seq are unchanged in *Δh3v* parasites. Read counts were normalized to the median number of reads mapped to each gene (see [17] for methods used). The X- and Y-axes shows sense strand read counts for genes from WT and *Δh3v* clone 3, respectively. The solid line shows the slope (1) expected for no changes in transcript levels; dotted lines represent 2-fold higher and lower boundaries. The correlation coefficient (R^2) comparing WT and *Δh3v* clone 3 was 0.9646. Comparable results were obtained in comparisons of WT with *H3V/HYG* ($R^2 = 0.9905$) or *Δh3v* clone 4 ($R^2 = 0.9918$).