



Depletion of the RNA-Binding Protein RBP33 Results in Increased Expression of Silenced RNA Polymerase II Transcripts in *Trypanosoma brucei*

Sandra M. Fernández-Moya^{1‡}, Mark Carrington², Antonio M. Estévez^{1*}

1 Instituto de Parasitología y Biomedicina "López-Neyra", IPBLN-CSIC, Armilla, Granada, Spain, **2** Department of Biochemistry, University of Cambridge, Cambridge, United Kingdom

Abstract

We have characterized the RNA-binding protein RBP33 in *Trypanosoma brucei*, and found that it localizes to the nucleus and is essential for viability. The subset of RNAs bound to RBP33 was determined by immunoprecipitation of ribonucleoprotein complexes followed by deep sequencing. Most RBP33-bound transcripts are predicted to be non-coding. Among these, over one-third are located close to the end of transcriptional units (TUs) or have an antisense orientation within a TU. Depletion of RBP33 resulted in an increase in the level of RNAs derived from regions that are normally silenced, such as strand-switch regions, retroposon and repeat sequences. Our work provides the first example of an RNA-binding protein involved in the regulation of gene silencing in trypanosomes.

Citation: Fernández-Moya SM, Carrington M, Estévez AM (2014) Depletion of the RNA-Binding Protein RBP33 Results in Increased Expression of Silenced RNA Polymerase II Transcripts in *Trypanosoma brucei*. PLoS ONE 9(9): e107608. doi:10.1371/journal.pone.0107608

Editor: Akio Kanai, Keio University, Japan

Received: June 9, 2014; **Accepted:** August 18, 2014; **Published:** September 12, 2014

Copyright: © 2014 Fernández-Moya et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. The dataset was deposited in the European Nucleotide Archive (<http://www.ebi.ac.uk/ena/data/view/PRJEB6768>).

Funding: AME was funded by Spanish Plan Nacional, grant BFU 2009-07510. MC was funded by The Wellcome Trust, grant 085956/Z/08/Z. AME and MC were jointly funded by the Royal Society Joint Project grant, grant 2008/R2. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* Email: aestevez@ipb.csic.es

‡ Current address: Institute for Cell Biology, Ludwig-Maximilians-Universität München, München, Germany

Introduction

Trypanosomatids are single-celled eukaryotes that are responsible for major human and livestock diseases. Genome organization in these organisms is unusual for a eukaryote, as open reading frames are organized in long polycistronic transcription units (TUs) that are transcribed by RNA polymerase II from only a few initiation sites in the chromosome [1]. Individual mRNAs are co-transcriptionally processed from polycistronic precursors by coupled *trans*-splicing at the 5'-end and polyadenylation at the 3'-end [2]. Adjacent TUs are often convergent or divergent, and are separated by strand switch regions (SSRs). RNA polymerase II transcription usually begins at divergent SSRs and ends at convergent SSRs [1,3]. Histone variants present in convergent SSRs are distinct from other variants associated with divergent SSRs, and it is assumed that the limits of polycistronic TUs are defined by distinct chromatin conformations [4]. Therefore, RNA polymerase II-dependent transcription initiation and termination seems to be regulated simply by histone modifications and chromatin structure, but the mechanism is currently unknown. In addition, chromatin conformation could be important in insulating RNA polymerase II TUs from RNA polymerase I and III expression sites [4,5].

There are numerous repetitive sequences within the *Trypanosoma brucei* genome, including retroposons INGI/RIME and SLACs, and satellite-like repeats such as CIR147 [6]. Transcripts

derived from these elements are selectively degraded by the RNA interference (RNAi) pathway [6]. It is currently unknown whether chromatin structure also plays a role in the transcriptional silencing of retroposons and repeats. That could well be the case, since some histone variants that are known to be involved in heterochromatin-dependent silencing in other organisms, such as H2AZ, are associated with repetitive DNA in *T. brucei* [7].

In other eukaryotes, non-coding RNAs (ncRNAs) and RNA-binding proteins (RBPs) are emerging as important factors regulating gene silencing and chromatin structure through RNAi-dependent and independent mechanisms. Examples include RNAi-mediated formation of heterochromatin at non-coding and repeated DNA in *Saccharomyces pombe* [8], or transcriptional silencing promoted by the association of chromatin-remodeling complexes with ncRNA-binding proteins such as IDN2 in *Arabidopsis thaliana* [9] or Nrd1 in *S. pombe* and *S. cerevisiae* [10]. Whether a similar scenario also occurs in trypanosomes is not known at present.

In this work we describe the characterization of the RNA-binding protein RBP33, which was first identified in a search for proteins that were able to bind to a uridine-rich RNA element *in vitro* [11]. RBP33 depletion results in the accumulation of transcripts derived from silenced regions, such as SSRs and retroposon genes, and thus it could have a role in regulating gene silencing in trypanosomes.

Results

RBP33 is an essential nuclear protein in trypanosomes

The *T. brucei* RNA-binding protein RBP33 (Tb927.8.990) is 317 amino acids long, and contains an RNA-recognition motif, a putative nuclear localization signal and two intrinsically disordered regions (Fig. 1A). Its carboxyl half has no apparent similarity to any known protein outside the *Trypanosomatidae* family, and it is predicted to have intrinsically disordered regions also in other trypanosomatid species (Fig. S1). Western blot analysis of total cell extracts using a polyclonal antiserum raised against the carboxyl half of RBP33 [11] revealed a single band with an apparent electrophoretic mobility of 40 kDa, which was expressed at similar levels in the two readily cultured developmental forms, the mammalian bloodstream and the insect procyclic forms (Fig. 1B and Fig. S2A). RBP33 localized mainly in the nucleus, as judged by immunofluorescence assays (Fig. 1C). The abundance of RBP33 was estimated to be $\sim 50,000$ molecules per cell, as judged by Western blot analysis with procyclic cell extracts and recombinant RBP33 (Fig. S2), which is within the range described for other RNA-binding proteins in trypanosomes (reviewed in [12]). RBP33 expression was silenced *in vivo* using RNA interference (RNAi) in a tetracycline-inducible fashion. Although RNAi was not very efficient at the mRNA level, a marked reduction in protein abundance was observed after 48 h of tetracycline induction (Fig. 2A–B and Fig. S2B). The presence of two bands in the Northern blot assays most likely represents transcripts with different 3'UTRs, as suggested by global transcriptome surveys [13]. Depletion of RBP33 resulted in cell death, indicating that this protein is essential in both the developmental forms tested (Fig. 2C).

RBP33 binds to a specific cohort of RNAs

To gain insight into RBP33 function, we sought to identify the set of RNAs associated to RBP33 using immunoprecipitation of RBP33 ribonucleoprotein complexes followed by RNA-seq deep

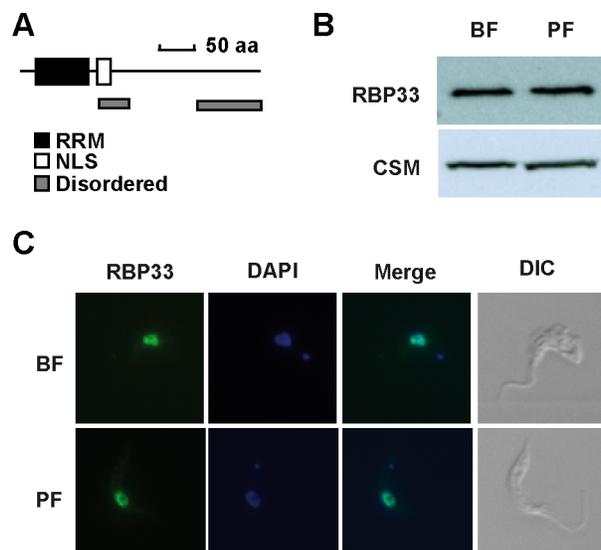


Figure 1. Expression and subcellular localization of RBP33 in bloodstream and procyclic trypanosomes. (A) Schematic diagram of *T. brucei* RBP33. An anti-RBP33 antiserum was used to detect the protein by (B) western blot analysis of total cell extracts or (C) immunofluorescence assays. CSM protein [39] was included as loading control, and DAPI was used to stain nuclei. doi:10.1371/journal.pone.0107608.g001

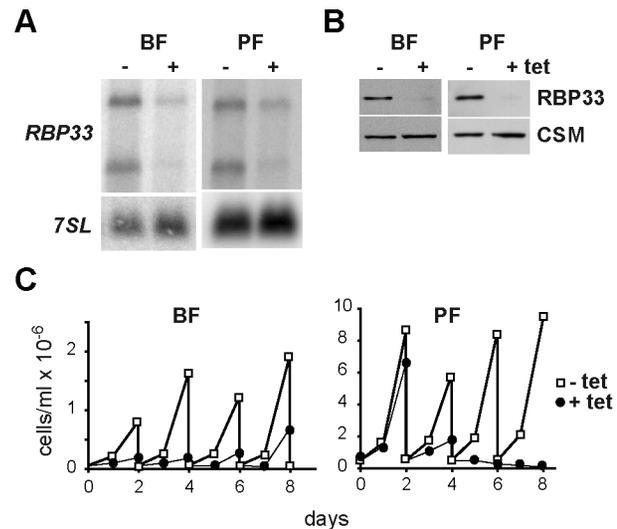


Figure 2. RBP33 is an essential protein in trypanosomes. Cell lines were generated that expressed *RBP33*-specific dsRNA in a tetracycline-inducible manner. The levels of mRNA (A) or protein (B) were analyzed in cells incubated with tetracycline for 48 h. Loading controls were the 7SL RNA [40] or the CSM protein [39]. (C) RNAi cell lines were grown in the absence (open squares) or presence (filled circles) of 1 μ g/ml tetracycline to induce RNAi. Cultures were followed up to 8 days and diluted to 0.5×10^5 cells/ml (bloodstream) or 0.4×10^6 cells/ml (procyclic) every 2 days as required. doi:10.1371/journal.pone.0107608.g002

sequencing. In the immunoprecipitated material, we could identify ~ 430 transcripts that were enriched 4-fold or more ($\log_2FC > 2$) with respect to input RNA, and ~ 170 transcripts enriched 8-fold or more ($\log_2FC > 3$) (Table S1). Examples of RNA-seq alignments visualized using the Artemis software can be seen in Fig. S3A. Only a minority of these enriched RNAs encoded *bona fide* proteins according to the annotated reference genome (<http://www.geneDB.org/Homepage/Tbruceibrucei927>) and available ribosome profiling data [14]. There was no obvious functional relationship between the proteins encoded by these mRNAs. Interestingly, RBP33 binds to its own mRNA (Table S1 and Fig. S3A). Most RBP33-associated RNAs are either non-coding or annotated as encoding ‘hypothetical proteins’ or ‘hypothetical proteins, unlikely’ (Fig. 3A), and are present at minimal levels within the cell, according to the low number of reads obtained in our input RNA sample (Table S1) and in other genome-wide analyses [13,15]. When we analyzed the chromosome location of these genes, the majority (60%) were found within polycistronic units in sense orientation, whereas 17% localized to the vicinity of strand switch regions and 13% showed an antisense orientation within a transcription unit (Fig. 3B). We could confirm binding of RBP33 to transcripts derived from both protein-coding genes and convergent SSR regions by quantitative RT-PCR analysis of the immunoprecipitated material (Fig. 4). Although RBP33 could interact directly with RNA, we cannot rule out the possibility that RBP33 is in a complex with other proteins that directly bind to target transcripts.

RBP33 silencing increases the levels of SSR-derived transcripts

We next analyzed the effect that RBP33 depletion had on the levels of two RBP33 targets encoding *bona fide* proteins, the pteridine transporter Tb927.10.9080 and the flagellum protein FTZC (Tb927.10.8590). Silencing of RBP33 did not have a major

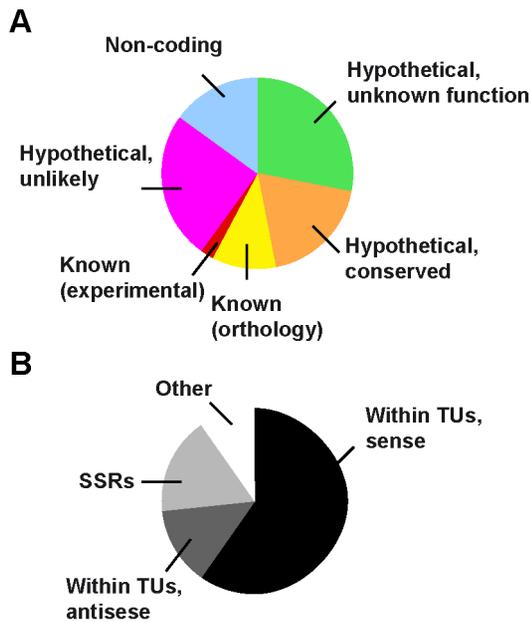


Figure 3. Distribution of RBP33-bound transcripts. (A) RNAs associated with RBP33 ($\log_2FC > 3$) were categorized according to whether they are predicted to encode functional proteins. Color code is based on that of GeneDB (<http://www.genedb.org/Homepage/Tbruceibrucei927>). (B) Localization of the genes encoding RBP33 targets ($\log_2FC > 3$) on the genome. 'SSR' refers to genes localized in the vicinity of strand-switch regions, whereas 'Other' indicates genes found close to the beginning or end of chromosomes, or close to snoRNA or pseudogene clusters.
doi:10.1371/journal.pone.0107608.g003

effect on the abundance of either transcript; FTZC protein levels were not altered either (Fig. S4). In contrast, the levels of several SSR-derived transcripts increased dramatically upon RBP33 silencing. As shown in Fig. 5A, a probe designed to detect transcripts derived from the convergent SSR between genes Tb927.7.1940 and Tb927.7.1950 (probe A), hybridized with two RNA species of 4.8 kb and 2.8 kb in both bloodstream and procyclic trypanosomes when RBP33 was ablated by RNAi. Accumulation of SSR-derived transcripts was not due to the slowing in growth and cell death associated with RBP33 RNAi as it was not observed upon depletion of DRBD3, an unrelated and essential RNA-binding protein [11]. In the case of the SSR between genes Tb927.10.5710 and Tb927.10.5720 (Fig. 5C), a probe hybridizing to the latter gene (probe B) detected three RNA species of 3.8 kb, 2.5 kb and 1.0 kb, whereas a probe hybridizing upstream (probe C) only detected the two larger RNAs. The three transcripts detected with probe B are likely to be polyadenylated, since they decreased in size upon incubation with RNaseH in the presence of oligo(dT) (Figure S5). We could map the splice-acceptor site for transcript Tb927.10.5720 to a canonical AG dinucleotide located 396 nt upstream the initiation codon (Figure 5B).

Retroposon- and repeat-derived transcripts also increase in abundance upon RBP33 depletion

Since convergent SSR regions are not transcribed in physiological conditions, we wondered whether RBP33 ablation could also promote the accumulation of other transcripts that are likewise silenced, such as retroposons INGI/RIME and SLACs, and CIR147 repeat transcripts [6]. Indeed, upon depletion of

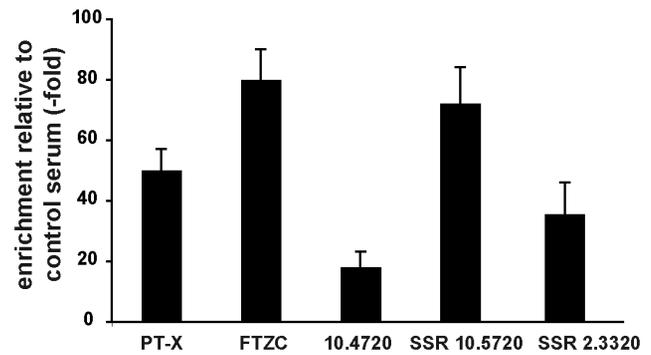


Figure 4. Validation of RBP33 targets. Quantitative RT-PCR assays were performed to confirm binding of three coding RNAs (pteridine transporter *PT-X*, Tb927.10.9080; flagellar transition zone component *FTZC*, Tb927.10.8590; and Tb927.10.4720) and transcripts derived from the SSR downstream gene Tb927.10.5720 or from the SSR downstream gene Tb927.2.3320.
doi:10.1371/journal.pone.0107608.g004

RBP33 there was an evident increase in the abundance of transcripts derived from both INGI and RIME retroposons, SLACs retroposons and CIR147 repeats (Fig. 6). A probe designed to detect only INGI-derived transcripts yielded similar results (data not shown). Interestingly, this phenomenon appears to be stage-specific, as the accumulation of transcripts was immediately apparent in bloodstream form cells but only marginal in procyclic form cells (Fig. 6).

Discussion

In this work, we have undertaken the characterization of the RNA-binding protein RBP33. We found that the protein is localized in the nucleus, and is essential in both bloodstream and procyclic trypanosomes. Using RNA immunoprecipitation and deep sequencing, we were able to identify a cohort of transcripts associated with RBP33 and thus obtain valuable information about the function of this protein. Although RBP33 binds to mRNAs encoding known or conserved proteins, most of its RNA targets are likely to be non-coding and/or present at minimal levels in the cell. Indeed, 90% of the RBP33-associated RNAs annotated as coding for 'hypothetical proteins' or 'hypothetical proteins, unlikely' were not detected by ribosome profiling [14], and 75% were not detected in a global transcriptome survey [13] (Table S1). Moreover, over one-third of the corresponding genes are located next to SSRs, close to the ends of chromosomes or have an antisense orientation within a transcription unit (Fig. 3 and Table S1). Interestingly, depletion of RBP33 results in an accumulation of SSR-derived transcripts. In the related parasite *Leishmania tarentolae*, it has been shown that base J (β -D-glucosyl-hydroxymethyluracil) prevents read-through at RNA polymerase II termination sites [16]. *L. tarentolae* cells seem to require base J for viability [17], but *T. brucei* can survive without J [18]. Two thymidine hydrolases are involved in the synthesis of J [18], but none is encoded by RBP33-bound mRNAs. To our knowledge, this is the first instance of a regulatory protein that has a role in silencing gene expression at RNA polymerase II termination regions. In addition, RBP33 silencing promotes the accumulation of retroposon and repeat-derived RNAs. This phenomenon has been also observed upon disruption of the RNAi machinery in *T. brucei* [6], raising the question of whether RBP33 could have a role in the regulation of the RNAi pathway in trypanosomes. In fact, convergent SSRs have been shown to produce small

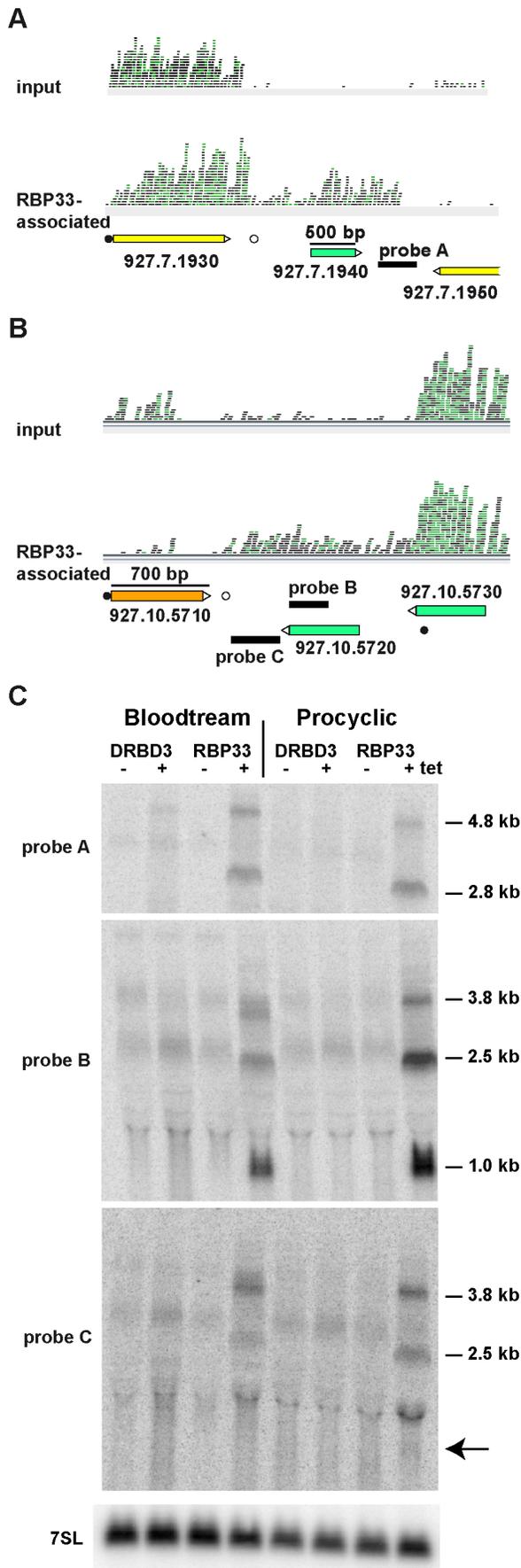


Figure 5. Effect of RBP33 depletion on the levels of SSR-derived transcripts. (A) Artemis visualization of RNA-seq data corresponding to the SSR downstream gene Tb927.7.1930 and to the SSR downstream gene Tb927.10.5710. Probes used for Northern blot analysis are indicated. 'Input' refers to the RNA-seq data obtained from total RNA before immunoprecipitation, whereas 'RBP33-bound' refers to the immunoprecipitated RNA using RBP33 specific antibodies. Annotated splice-acceptor and polyadenylation sites [13] are indicated as filled or empty circles, respectively. The splice-acceptor site for Tb927.10.5720 has been mapped in this study, and lies within the downstream open-reading frame. (B) Northern hybridizations using radiolabeled double-stranded DNA probes as indicated in panel A. An RNAi cell line expressing dsRNA against the essential RNA-binding protein DRBD3 [11] was used as a control. Total RNA was obtained from parasites grown in the absence (-) or the presence (+) of tetracycline for 48 h, transferred to nylon membranes and hybridized with specific probes. The 7SL RNA was used as a loading control. The arrow indicates the position of the 1.0 kb transcript detected with probe B. Representative blots are shown. doi:10.1371/journal.pone.0107608.g005

interfering RNAs (siRNAs) in *T. brucei* [19]. However, the molecular phenotype observed upon depletion of RBP33 presents some differences with that observed in trypanosomes lacking a functional RNAi pathway. Thus, deletion of the genes encoding the dicer-like protein TbDCL2 or the argonaute protein TbAGO1 results in the accumulation of full-length INGI transcripts which are ~6 kb long and CIR147-derived transcripts of ~10 kb [6], whereas the transcripts observed upon RBP33 silencing were smaller (Fig. 6). We did, however, detect full-length transcripts corresponding to SLACs retroposons. Second, the accumulation of retroposon and repeat transcripts is stage-dependent, being almost undetected in procyclic trypanosomes. And third, deletion of

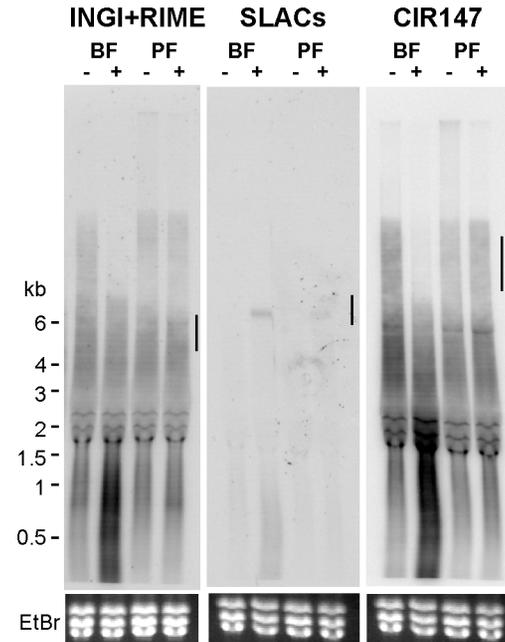


Figure 6. Effect of RBP33 depletion on the levels of transcripts derived from retroposons and CIR147 repeats. Probes designed to detect both INGI/RIME-derived transcripts, SLACs retroposons or CIR147-derived transcripts were used in Northern hybridizations as described in Fig. 5 legend. Ethidium bromide staining was used as a loading control. Vertical lines indicate the expected migration of the corresponding transcripts upon disruption of the RNAi machinery observed in [6]. doi:10.1371/journal.pone.0107608.g006

TbDCL2 did not result in the accumulation of convergent SSR-derived transcripts, whereas this sort of RNAs was readily visible upon RBP33 ablation. For instance, we could detect transcripts derived from the SSR located in between genes Tb927.10.2390 and Tb927.10.2400 (Figure S3B), whereas no accumulation of RNAs transcribed from this region was observed when TbDCL2 was deleted [19].

SSRs are enriched in specific histone variants in trypanosomes [4] but whether these variants promote heterochromatin assembly and thus transcription silencing is not known at present. Likewise, it is unknown whether the silent transcriptional status of CIR147 repeats and retroposon genes is due to heterochromatinization, although some histone variants that are known to be involved in heterochromatin-dependent silencing, such as H2AZ, have been found associated to repetitive DNA [7]. In other organisms, heterochromatin formation depends on both RNAi-dependent and RNAi-independent mechanisms. Among the latter, the yeast RNA-binding protein Nrd1 has been shown to bind to unstable non-coding RNAs (ncRNAs), and to promote heterochromatin assembly and transcriptional silencing, thus providing a failsafe termination mechanism in case of read through transcription [20,21]. RBP33 could perform a similar function in trypanosomes by binding to non-coding RNAs and promoting gene silencing *via* chromatin conformation. The source of ncRNAs could be SSRs, genes whose products are annotated as ‘hypothetical proteins (unlikely)’ or even transcripts coding for known proteins. In this regard, we have shown that although RBP33 binds to translatable transcripts, its depletion has no apparent effect on the levels of target mRNAs or in the protein product. However, RBP33 could be involved in the processing of cryptic ncRNAs located within the untranslated regions of a defined set of mRNAs. Alternatively, RBP33 could be part of a surveillance mechanism that selectively degrades RNAs transcribed from silenced regions in the genome. Another plausible interpretation is that RBP33 depletion results in an increased maturation or alternative processing of transcripts derived from SSR and retroposons, which is supported by the observation that slower-migrating INGI/RIME and CIR147 repeats transcripts decrease in abundance upon RBP33 silencing in bloodstream forms (Fig. 6). The actual mechanism remains to be determined. Our findings support the hypothesis that gene silencing is, at least in part, controlled at the RNA level in trypanosomes, and uncovers an unprecedented role for RNA-binding proteins in this process.

Materials and Methods

Trypanosome culture and RNA interference

Trypanosoma brucei 449 procyclic cells [22] were grown at 27°C in SDM-79 medium [23] containing 10% fetal bovine serum. ‘Single marker’ *T. brucei* Lister 427 bloodstream cell line S16 [24] were maintained in HMI-9 medium [25] containing 10% fetal bovine serum at 37°C with 5% CO₂. For RNAi studies, a stem-loop strategy was followed to generate dsRNA [11]. A fragment of the RBP33 ORF corresponding to the carboxyl half of the protein was PCR-amplified and cloned into pGR19 [26] to yield pGR70. Procyclic and bloodstream trypanosomes were transfected with pGR70 linearized with *Not* I and selected in the presence of 50 µg/ml (procyclic) or 5 µg/ml (bloodstream trypanosomes) hygromycin. For RNAi induction, tetracycline was added to the culture medium at a concentration of 1 µg/ml.

Protein analysis

Polyclonal antibodies against the last 160 amino acids of *T. brucei* RBP33 were raised in rabbits and affinity-purified using a

full-length version of the protein (described in [11]). Immunoblotting and immunofluorescence analysis were carried out as in [27,28]. Protein motifs were analyzed using PSORT [29] and SMART [30]. Intrinsically disordered regions were identified using GlobPlot [31], IUPred [32] and PrDOS [33] software.

RNA analysis

Northern hybridizations and quantitative RT-PCR analysis were done as described [11,34]. The oligodeoxynucleotides used in this work are listed in Table S2. To map the splice-acceptor site for Tb927.10.5720 transcript, cDNA was obtained from RBP33-depleted RNA samples using oligodeoxynucleotide AE340, and conventional RT-PCR reactions were carried out using AE340 and a sense oligodeoxynucleotide corresponding to the *T. brucei* spliced leader sequence. To remove poly(A) tails, 20 µg of total RNA obtained from RBP33-depleted samples were incubated with 300 pmol of oligo(dT)₂₀, heated at 65°C for 10 min, transferred to a water bath at 42°C, cooled to 30°C, treated with 5 U of RNaseH (Invitrogen) for 30 min at 37°C and subjected to Northern hybridization. Samples lacking oligo(dT)₂₀ were treated identically and analyzed in parallel.

To identify the cohort of transcripts bound to RBP33, 2 × 10⁹ procyclic trypanosomes were washed in ice-cold Voorheis’s-modified phosphate-buffered saline (vPBS; PBS supplemented with 10 mM glucose and 46 mM sucrose), resuspended in 1 ml of vPBS and UV-irradiated twice at 400 mJ/cm² in a Stratalinker apparatus (Stratagene). Cells were resuspended in 2 ml of 10 mM Tris-HCl, pH 7.4, 2 mM DTT, 1 mM EDTA, 0.1% Igepal, 50 U/ml RNaseOUT (Invitrogen) and protease inhibitors (complete mini EDTA-free cocktail, Roche) and lysed by passing the suspension through a 27-gauge syringe thrice on ice. The cell extract was centrifuged at 16,000 ×g for 10 min at 4°C, and NaCl was added to the supernatant at a final concentration of 150 mM. This final lysate was incubated with 2 µg of affinity-purified anti-RBP33 antibody for 3 h at 4°C. To capture RBP33 ribonucleoprotein particles, 150 µl of 50% protein A agarose slurry (Calbiochem) were added to the lysate. After an incubation of 1.5 h at 4°C, beads were transferred to a disposable column (Biorad), washed twice with 10 ml of washing buffer (10 mM Tris-HCl, pH 7.6, 150 mM NaCl, 1 mM EDTA) and once with 2 ml of washing buffer containing 1 M urea, resuspended in 450 µl of 10 mM Tris-HCl, pH 8, 100 mM NaCl, 0.5% SDS, 1 mM EDTA and incubated in the presence of 100 µg of proteinase K (Ambion) for 30 min at 50°C. RBP33-associated RNAs were obtained by phenol:chloroform extraction and ethanol precipitation, and were DNaseI-treated prior to sequencing.

cDNA libraries for high-throughput sequencing were obtained from two biological replicates using the Illumina mRNA-seq kit according to manufacturer’s instructions. Libraries were sequenced on an Illumina platform at EASIH (Eastern Sequencing and Informatics Hub, Cambridge, UK) and the reads of 36 nucleotides in length were aligned to the *T. brucei* TREU 927 genome v4 using Bowtie [15,35]. Manipulation and processing of the aligned tags were done using SAMtools [36]. Aligned sequence tags were visualized using Artemis [37]. The number of hits per ORF was determined using HT-Seq software (<http://www-huber.embl.de/users/anders/HTSeq/doc/index.html>). Replicates were compared with the input RNA sample using the edgeR package [38]. The dataset was deposited in the European Nucleotide Archive (<http://www.ebi.ac.uk/ena/data/view/PRJEB6768>).

Supporting Information

Figure S1 Alignment of RBP33 from *T. brucei*, *T. cruzi* and *Leishmania major*. (PDF)

Figure S2 Assessment of anti-RBP33 antiserum specificity and RBP33 abundance. (A) A single immunoreactive band is visible in total cell extracts of bloodstream and procyclic trypanosomes in western blot assays. (B) RBP33 specific staining is no longer detected when the protein is ablated by RNAi in immunofluorescence analysis. (C) Quantitation of RBP33 protein levels in procyclic trypanosomes. The abundance of RBP33 present in different amounts of trypanosome cells (left) was compared to known amounts of recombinant His₆-CBP-RBP33 [11] (right) in western blot assays. (PDF)

Figure S3 Examples of Artemis plots of RNA-seq data. (A) Regions corresponding to genes *FTZC* (Tb927.10.8590), *RBP33* (Tb927.8.990) and Tb927.7.6010 are shown. (B) Transcripts derived from the SSR located in between genes Tb927.10.2390 and Tb927.10.2400 increase in abundance upon RBP33 depletion. Probes used for Northern hybridizations are indicated. Probe A is identical to that used in [19]. The annotated splicing acceptor site for Tb927.10.2390 is indicated with a filled circle, and the most proximal and most distal annotated polyadenylation sites are represented with empty circles. (PDF)

Figure S4 RBP33 depletion has no effect on the levels of *PT-X* and *FTZC* mRNAs or *FTZC* protein. (A) Total RNA was obtained from trypanosomes grown in the presence or absence of tetracycline for 48 h, transferred to Nylon membranes

and hybridized with probes designed to detect the *PT-X* and *FTZC* mRNAs. *7SL* RNA was used as a loading control. (B) Total cell extracts were obtained from parasites grown in the presence or absence of tetracycline for 48 h and subjected to western blot analysis using an antiserum against *FTZC* [41]. CSM protein was used as a loading control.

(PDF)

Figure S5 RNaseH treatment. To assess whether the transcripts detected with probe B in Figure 5 are polyadenylated, RNA samples obtained from RBP33-depleted cells were incubated with RNase H in the absence (–) or presence (+) of oligo(dT).

(PDF)

Table S1 List of transcripts associated with RBP33. Transcripts with a log₂FC enrichment >3 (Sheet 1) or >2 (Sheet 2) with respect to the input sample are shown.

(XLSX)

Table S2 Oligonucleotides used in this work. List of the oligonucleotides designed for Northern and quantitative RT-PCR assays.

(XLSX)

Acknowledgments

We thank Frédéric Bringaud for anti-*FTZC* antiserum and advice in the design of retroposon probes.

Author Contributions

Conceived and designed the experiments: SMFM MC AME. Performed the experiments: SMFM MC AME. Analyzed the data: SMFM MC AME. Contributed to the writing of the manuscript: AME.

References

- Martínez-Calvillo S, Yan S, Nguyen D, Fox M, Stuart K, et al. (2003) Transcription of *Leishmania major* Friedlin chromosome 1 initiates in both directions within a single region. *Mol Cell* 11: 1291–1299.
- Liang XH, Haritan A, Uliel S, Michaeli S (2003) *trans* and *cis* splicing in trypanosomatids: mechanism, factors, and regulation. *Eukaryot Cell* 2: 830–840.
- Martínez-Calvillo S, Nguyen D, Stuart K, Myler PJ (2004) Transcription initiation and termination on *Leishmania major* chromosome 3. *Eukaryot Cell* 3: 506–517.
- Siegel TN, Hekstra DR, Kemp LE, Figueiredo LM, Lowell JE, et al. (2009) Four histone variants mark the boundaries of polycistronic transcription units in *Trypanosoma brucei*. *Genes Dev* 23: 1063–1076.
- Haenni S, Studer E, Burkard GS, Roditi I (2009) Bidirectional silencing of RNA polymerase I transcription by a strand switch region in *Trypanosoma brucei*. *Nucleic Acids Res* 37: 5007–5018.
- Patrick KL, Shi H, Kolev NG, Ersfeld K, Tschudi C, et al. (2009) Distinct and overlapping roles for two Dicer-like proteins in the RNA interference pathways of the ancient eukaryote *Trypanosoma brucei*. *Proc Natl Acad Sci U S A* 106: 17933–17938.
- Lowell JE, Kaiser F, Janzen CJ, Cross GA (2005) Histone H2AZ dimerizes with a novel variant H2B and is enriched at repetitive DNA in *Trypanosoma brucei*. *J Cell Sci* 118: 5721–5730.
- Creamer KM, Partridge JF (2011) RITS-connecting transcription, RNA interference, and heterochromatin assembly in fission yeast. *Wiley Interdiscip Rev RNA* 2: 632–646.
- Zhu Y, Rowley MJ, Bohmdorfer G, Wierzbicki AT (2013) A SWI/SNF chromatin-remodeling complex acts in noncoding RNA-mediated transcriptional silencing. *Mol Cell* 49: 298–309.
- Schulz D, Schwalb B, Kiesel A, Baejen C, Torkler P, et al. (2013) Transcriptome surveillance by selective termination of noncoding RNA synthesis. *Cell* 155: 1075–1087.
- Estevez AM (2008) The RNA-binding protein *TbDRBD3* regulates the stability of a specific subset of mRNAs in trypanosomes. *Nucleic Acids Res* 36: 4573–4586.
- Clayton CE (2014) Networks of gene expression regulation in *Trypanosoma brucei*. *Mol Biochem Parasitol*.
- Kolev NG, Franklin JB, Carmi S, Shi H, Michaeli S, et al. (2010) The transcriptome of the human pathogen *Trypanosoma brucei* at single-nucleotide resolution. *PLoS Pathog* 6: e1001090.
- Vasquez JJ, Hon CC, Vanselow JT, Schlosser A, Siegel TN (2014) Comparative ribosome profiling reveals extensive translational complexity in different *Trypanosoma brucei* life cycle stages. *Nucleic Acids Res* 42: 3623–3637.
- Siegel TN, Hekstra DR, Wang X, Dewell S, Cross GA (2010) Genome-wide analysis of mRNA abundance in two life-cycle stages of *Trypanosoma brucei* and identification of splicing and polyadenylation sites. *Nucleic Acids Res* 38: 4946–4957.
- van Luenen HG, Farris C, Jan S, Genest PA, Tripathi P, et al. (2012) Glucosylated hydroxymethyluracil, DNA base J, prevents transcriptional readthrough in *Leishmania*. *Cell* 150: 909–921.
- Genest PA, ter Riet B, Dumas C, Papadopoulou B, van Luenen HG, et al. (2005) Formation of linear inverted repeat amplicons following targeting of an essential gene in *Leishmania*. *Nucleic Acids Res* 33: 1699–1709.
- Cliffe IJ, Kieft R, Southern T, Birkeland SR, Marshall M, et al. (2009) JBP1 and JBP2 are two distinct thymidine hydroxylases involved in J biosynthesis in genomic DNA of African trypanosomes. *Nucleic Acids Res* 37: 1452–1462.
- Tschudi C, Shi H, Franklin JB, Ullu E (2012) Small interfering RNA-producing loci in the ancient parasitic eukaryote *Trypanosoma brucei*. *BMC Genomics* 13: 427.
- Webb S, Hector RD, Kudla G, Granneman S (2014) PAR-CLIP data indicate that *Nrd1-Nab3*-dependent transcription termination regulates expression of hundreds of protein coding genes in yeast. *Genome Biol* 15: R8.
- Marina DB, Shankar S, Natarajan P, Finn KJ, Madhani HD (2013) A conserved ncRNA-binding protein recruits silencing factors to heterochromatin through an RNAi-independent mechanism. *Genes Dev* 27: 1851–1856.
- Biebinger S, Wirtz LE, Lorenz P, C C (1997) Vectors for inducible expression of toxic gene products in bloodstream and procyclic *Trypanosoma brucei*. *Mol Biochem Parasitol* 85: 99–112.
- Brun R, Schönerberger M (1979) Cultivation and *in vitro* cloning of procyclic culture forms of *Trypanosoma brucei* in a semi-defined medium. 36: 289–292.
- Wirtz E, Leal S, Ochart C, Cross GAM (1999) A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in *Trypanosoma brucei*. *Mol Biochem Parasitol* 99: 89–102.
- Hirumi H, Hirumi K (1989) 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.
- Clayton C, Estevez AM, Hartmann C, Alibu VP, Field MC, et al. (2005) Down-regulating gene expression by RNA interference in *Trypanosoma brucei*. *Methods Mol Biol* 309: 39–60.

27. Estevez AM, Lehner B, Sanderson CM, Ruppert T, Clayton C (2003) The roles of intersubunit interactions in exosome stability. *J Biol Chem* 278: 34943–34951.
28. Haile S, Cristodero M, Clayton C, Estevez AM (2007) The subcellular localisation of trypanosome RRP6 and its association with the exosome. *Mol Biochem Parasitol* 151: 52–58.
29. Nakai K, Horton P (1999) PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends Biochem Sci* 24: 34–36.
30. Letunic I, Doerks T, Bork P (2012) SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res* 40: D302–305.
31. Linding R, Russell RB, Neduva V, Gibson TJ (2003) GlobPlot: Exploring protein sequences for globularity and disorder. *Nucleic Acids Res* 31: 3701–3708.
32. Dosztanyi Z, Csizmok V, Tompa P, Simon I (2005) IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* 21: 3433–3434.
33. Ishida T, Kinoshita K (2007) PrDOS: prediction of disordered protein regions from amino acid sequence. *Nucleic Acids Res* 35: W460–464.
34. Fernandez-Moya SM, Garcia-Perez A, Kramer S, Carrington M, Estevez AM (2012) Alterations in DRBD3 ribonucleoprotein complexes in response to stress in *Trypanosoma brucei*. *PLoS One* 7: e48870.
35. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10: R25.
36. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25: 2078–2079.
37. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, et al. (2000) Artemis: sequence visualization and annotation. *Bioinformatics* 16: 944–945.
38. Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26: 139–140.
39. Guerra-Giraldez C, Quijada L, Clayton CE (2002) Compartmentation of enzymes in a microbody, the glycosome, is essential in *Trypanosoma brucei*. *J Cell Sci* 115: 2651–2658.
40. Michaeli S, Podell D, Agabian N, Ullu E (1992) The 7SL RNA homologue of *Trypanosoma brucei* is closely related to mammalian 7SL RNA. *Mol Biochem Parasitol* 51: 55–64.
41. Bringaud F, Robinson DR, Barradeau S, Biteau N, Baltz D, et al. (2000) Characterization and disruption of a new *Trypanosoma brucei* repetitive flagellum protein, using double-stranded RNA inhibition. *Mol Biochem Parasitol* 111: 283–297.