International Journal for Parasitology 48 (2018) 179–189

Contents lists available at ScienceDirect

International Journal for Parasitology

journal homepage: www.elsevier.com/locate/ijpara

Tail characteristics of *Trypanosoma brucei* mitochondrial transcripts are developmentally altered in a transcript-specific manner

Vahid H. Gazestani [a,1], Marshall Hampton [b], Aubie K. Shaw [c], Reza Salavati [a,*], Sara L. Zimmer [c,*]

[a] Institute of Parasitology, McGill University, Montreal, QC H9X 3V9, Canada
[b] Department of Mathematics and Statistics, University of Minnesota Duluth, Duluth, MN 55812, USA
[c] Department of Biomedical Sciences, University of Minnesota Medical School, Duluth, MN 55812, USA

ARTICLE INFO

Article history:
Received 16 June 2017
Received in revised form 22 August 2017
Accepted 26 August 2017
Available online 1 November 2017

Keywords:
RNA tail
Mitochondrion
Trypanosoma
Uridylation
Polyadenylation
Hidden Markov modelling

ABSTRACT

The intricate life cycle of *Trypanosoma brucei* requires extensive regulation of gene expression levels of the mtRNAs for adaptation. Post-transcriptional gene regulatory programs, including unencoded mtRNA 3′ tail additions, potentially play major roles in this adaptation process. Intriguingly, *T. brucei* mitochondrial transcripts possess two distinct unencoded 3′ tails, each with a differing functional role; i.e., while one type is implicated in RNA stability (in-tails), the other type appears associated with translation (ex-tails). We examined the degree to which tail characteristics differ among cytochrome c oxidase subunits I and III (CO1 and CO3), and NADH dehydrogenase subunit 1 (ND1) transcripts, and to what extent these characteristics differ developmentally. We found that CO1, CO3 and ND1 transcripts possess longer in-tails in the mammalian life stage. By mathematically modelling states of in-tail and ex-tail addition, we determined that the typical length at which an in-tail is extended to become an ex-tail differs by transcript and, in the case of ND1, by life stage. To the best of our knowledge, we provide the first evidence that developmental differences exist in tail length distributions of mtRNAs, underscoring the potential involvement of in-tail and ex-tail populations in mitochondrial post-transcriptional regulation mechanisms.

1. Introduction

Vector-borne *Trypanosoma brucei* parasites cause sleeping sickness in humans and major production losses (e.g., meat, milk and fertility) in livestock. These pathogens shuttle between mammalian hosts (bloodstage form, BSF) and the tsetse fly vector (procyclic form, PCF), encountering extremely different environmental conditions, particularly in available energy sources. To cope with such extreme changes, *T. brucei* possesses a complicated, multilayer mitochondrial gene regulatory pathway that is essential for its survival in both life stages (Aphasizhev and Aphasizheva, 2011). Such dependence on unique mitochondrial gene regulatory pathways presents an attractive target for drug development. Accordingly, several compounds with trypanosomacidal activity, including pentamidine and ethidium bromide, target the mitochondrial genome or its fitness (Roy Chowdhury et al., 2010; Fidalgo and Gille, 2011). However, many aspects of *T. brucei* mitochondrial gene regulation are still poorly understood.

Trypanosoma brucei mitochondrial gene expression is significantly divergent from that of yeast or mammals (Verner et al., 2015). Illustrated in Fig. 1A, gene expression starts with the transcription of two rRNAs and 18 protein-encoding transcripts (mtRNAs) from mitochondrial DNA "maxicircles" in a polycistronic manner. Post-transcriptional events are the main mechanism of mtRNA expression regulation, mediating cleavage/trimming of RNA precursors into monocistrons, and mtRNA decay and translation rates (Aphasizhev and Aphasizheva, 2011; Carnes et al., 2015; Suematsu et al., 2016; Zhang et al., 2017). Twelve of the 18 mtRNAs undergo RNA editing in which uridine(s) are inserted in and/or deleted from specific positions of mtRNAs to form correct translatable transcripts. This is a transcript-specific process that is directed by a set of template-containing guide RNAs (gRNAs) (Hashimi et al., 2013; Aphasizheva et al., 2014; Aphasizheva and Aphasizhev, 2016; Read et al., 2016). While the role of RNA editing in mtRNA post-transcriptional regulation has been studied extensively, the importance of mtRNA stability and translation regulation, and how those relate to RNA editing, has been analysed less thoroughly.

* Corresponding authors.
E-mail addresses: reza.salavati@mcgill.ca (R. Salavati), szimmer3@d.umn.edu (S.L. Zimmer).
[1] Present address: Department of Pediatrics, University of California, San Diego School of Medicine, La Jolla, CA 92093, USA.

https://doi.org/10.1016/j.ijpara.2017.08.012
0020-7519/

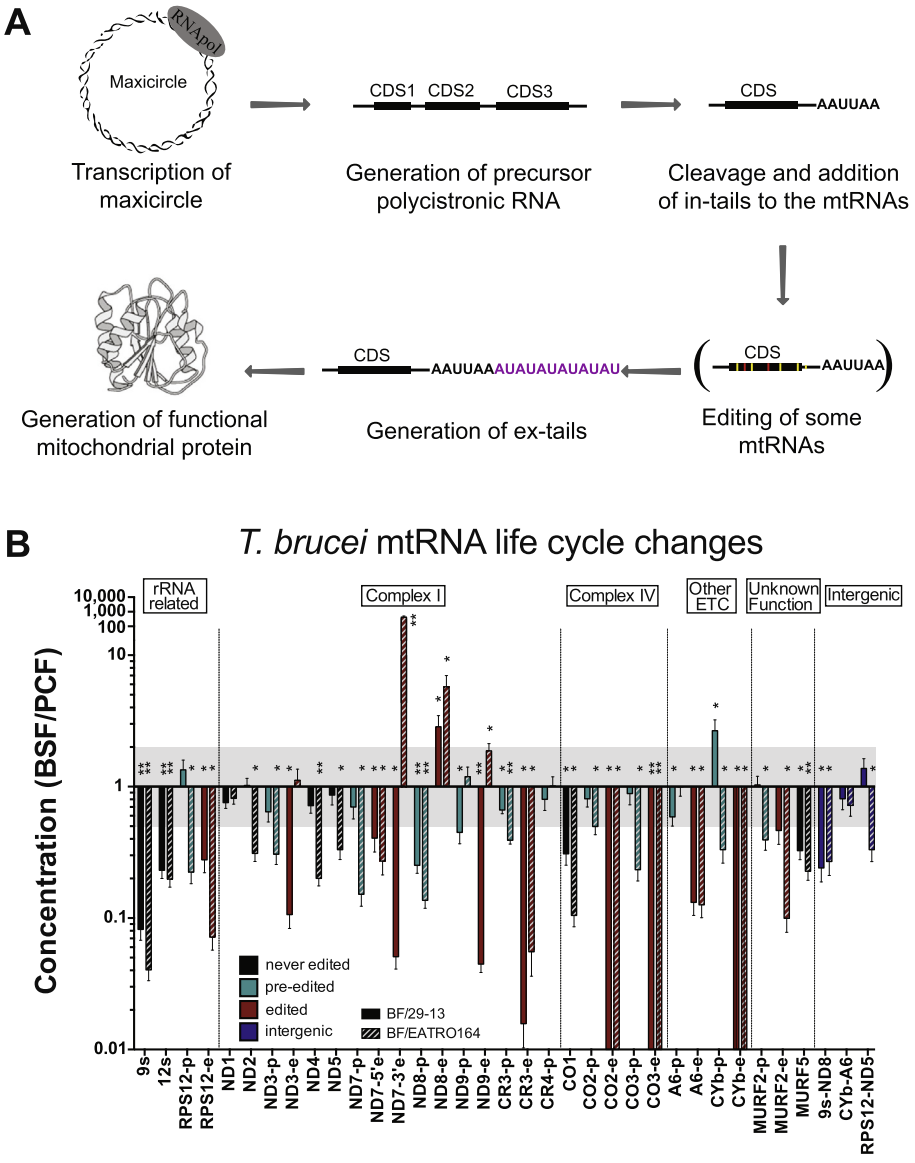

Fig. 1. *Trypanosoma brucei* maxicircle gene expression. (A) Pathway of mtRNA processing and expression. CDS, coding sequence. (B) Log-scale comparison of relative abundances of maxicircle-derived RNAs (arranged on the *X* axis) from cultured *T. brucei* in the bloodstage form compared with the procyclic form life stage. RNAs are arranged by gene product function. Never edited RNAs are shown in black, pre-edited (-p) RNAs in green, and edited (-e) RNAs in brown. ND7 contains a "constitutively edited" 5′ region and a 3′ region believed to be edited in the bloodstage form stage only; hence, each site was analysed separately. Three primer pairs spanning from one gene product to the next capture abundances of polycistronic RNAs containing those two sequences prior to cleavage into monocistrons (shown in blue). The relative expression of RNAs from the 427-derived BSF SM strain with both 427-derived PCF 29-13 strain (solid) and 927-derived PCF EATRO164 Istar1 strain (hatched) are shown. Error bars represent S.E.M. $^{**}P < 0.001$, $^{*}P < 0.05$. Where a ratio is displayed as reaching 0.01 (CO2-p, CO3-p and CYb-e), the transcript was undetectable in the bloodstage form cells.

In vitro and in vivo experiments have revealed distinguishable tailing processes that impact mtRNA stability and translation. mtRNA tails can be composed of adenine (A), uridine (U), or combinations thereof (Verner et al., 2015). mtRNA tails further fall into distinct categories of in-tails (initially added tails) and ex-tails (extended tails) with differences in their length and biological functions (Fig. 1A). In-tails are fairly ubiquitous oligomer tails added to mtRNAs. The As in in-tails are added by the poly(A) polymerase KPAP1 (Etheridge et al., 2008) and Us by the terminal uridyltransferase RET1 (Aphasizhev et al., 2002, 2003) in a manner influenced by the pentatricopeptide protein KPAF3 (Zhang et al., 2017). In-tails are recognised as stability elements that are ubiquitously added to mtRNAs upon maturation of their 3′ ends, but other potential roles for them have not been explored (Ryan et al., 2003; Kao and Read, 2005; Etheridge et al., 2008; Aphasizheva and Aphasizhev, 2010; Zhang et al., 2017). Transcript-specific variation in sequence composition and length of in-tails have been suggested by studies that either had low sequencing depth or were limited to one life stage of parasite (Decker and Sollner-Webb, 1990; Souza et al., 1992; Kao and Read, 2007; Zimmer et al., 2012; Gazestani et al., 2016, Zhang et al., 2017). However, systematic studies of adequate depth to explore the relationship between transcript-specific in-tail variation and mtRNA regulation are lacking.

Ex-tails are longer, with potential roles in translational regulation. Ex-tails are generated by extensions appended to a subset of in-tails of translatable mtRNAs. The nucleotide extensions appear to be fairly homogenous in A/U composition with KPAP1/RET1 addition of A and U (a 7:3 A/U ratio (Etheridge et al., 2008)) controlled by the pentatricopeptide protein KPAF1 and possibly KPAF2 (Aphasizheva et al., 2011). Extensions exhibit a fairly consistent frequency in switching of addition from A to U and back.

Potentially, ex-tails could be recognised by the translational machinery and thereby mark the transcript for the translation (Aphasizheva et al., 2011, 2016). Although the ability to distinguish ex-tails from in-tails is important for differentiating their roles in mitochondrial post-transcriptional events, a systematic way of identifying the type of a tail as an ex-tail or in-tail is still lacking, hindering quantitative comparisons of in-tail to ex-tail ratios among transcripts, or understanding the in-tail to ex-tail transition. Furthermore, very few ex-tail sequences are published, so whether differences exist in the extensions themselves between translatable transcript tail populations is not known.

Current knowledge of the role of tails in mitochondrial gene regulatory processes is also limited because most previous reports investigated the functional roles of mtRNA tails by focusing on either a single life stage or in vitro experiments. Interestingly, studies support mtRNA abundance differences between insect and mammalian life stages of *T. brucei* (Feagin and Stuart, 1985; Feagin et al., 1985, 1986, 1987; Jasmer et al., 1985; Michelotti and Hajduk, 1987; Bhat et al., 1992; Michelotti et al., 1992; Read et al., 1992, 1994; Souza et al., 1992, 1993; Corell et al., 1994), as its metabolism dramatically changes in these two environments (Tielens and van Hellemond, 2009). A fascinating and plausible possibility is that tails play a role in this developmental regulation. One way that they could exert this control is by possessing altered characteristics between the life stages that would impact their function.

We recently developed a deep sequencing-based approach, termed circTAIL-seq, to capture the 3′ tails of kinetoplastid mtRNAs (Gazestani et al., 2016). In reporting this new method, we proved its capacity to detect differences in tail populations by cursorily demonstrating limited tail population characteristics for the cytochrome oxidase subunits III and I (CO3 and CO1) transcripts obtained in the PCF life stage. The primary strength of circTAIL-seq, which utilises circularisation of RNA molecules followed by reverse transcription (RT)-PCR prior to deep sequencing, is that it returns tens of thousands to hundreds of thousands of raw reads. This depth provides a comprehensive picture of differences in tail characteristics with their potential impact on expression levels of mtRNAs.

CircTAIL-seq does, however, have limitations, the major one being length bias towards capture of shorter reads. As a result, ex-tails will be underrepresented in circTAIL-seq captured populations. Electrophoresis methods may often be superior for distinguishing absolute ratios of in-tails and ex-tails. In this study, we are particularly interested in exploring potential developmental and transcript-specific in-tail differences, the transition point from in-tail to ex-tail, and differences in 5′ and 3′ termini that are also quantifiable by circTAIL-seq. For these purposes, circTAIL-seq is well suited. The dataset presented here expands that which was originally utilised for method validation in Gazestani et al. (2016), to include data on the NADH dehydrogenase subunit 1 (ND1) transcript and the BSF life stage.

In this study, we first globally analysed relative abundances of mtRNAs and rRNAs in widely used culture-grown laboratory strains. Our results confirmed the existence of developmental gene regulation pathways affecting transcript stability in a target-specific manner, and showed differences in strain-specific mtRNA abundance. Second, we assessed whether and how tail populations of mtRNAs differ between BSF and PCF life stages and across three representative transcripts. Third, we integrated computational modelling into our descriptive toolbox that, among other things, allowed us to distinguish ex-tails from in-tails within tail populations. In summary, our results demonstrated that tail lengths of mtRNAs change during the life cycle of the parasite. We also quantified transcript and life stage differences in tail composition, and estimated relative abundances of ex-tails between tail populations. Finally, life stage-specific differences in 5′ and 3′ untranslated region (UTR) lengths of CO3 and ND1 were observed that suggested developmental differences in maxicircle polycistron processing.

2. Materials and methods

2.1. Cell culture

The *Trypanosoma brucei* 29-13 cell line was grown in 5% CO_2 at 27 °C in SDM-79 medium supplemented with G418 and hygromycin (EATRO164 Istar1 cells were grown the same way with the omission of the two drugs) and harvested at late log stage (1.5×10^7 cells/ml). *Trypanosoma brucei* Single Marker cells were grown in HMI-9 medium with addition of G418 in 5% CO_2 at 37 °C and harvested when they attained a density of $1–2 \times 10^6$ cells/ml.

2.2. Real-time RT-PCR

RNA was isolated from cell pellets with TRIzol reagent (Life Technologies, Grand Island, NY, USA), with an additional extraction in phenol:chloroform pH 5.2. RNA was treated with DNase using the DNAfree kit (Life Technologies), and the integrity was verified via electrophoresis. Total RNA (4 μg) was reverse transcribed using the TaqMan Reverse Transcription kit (Life Technologies). Relative RNA abundances were determined by real-time quantitative (q) RT-PCR using LightCycler 480 instrumentation and software version 1.5.1.62 (Roche, Indianapolis, IN, USA). "No RNA" and "no reverse transcriptase" cDNA control reactions, and water, were negative controls in the quantitative PCRs. Samples were quantitated using the second derivative maximum method that is fitted to a standard curve of five fourfold serial dilutions of cDNA generated concurrently with sample data. Samples were normalised to the mean of three nuclear-encoded genes (*TERT*, *ACTIN,* and *PFR2)* that were specifically tested for use in comparison studies across life stages (Brenndorfer and Boshart, 2010) and appeared reliable in our hands. We performed biological replicates ($n = 5$) in duplicate wells (two biological replicates for EATRO164 Istar1). Statistical differences were compared using unpaired *t*-tests assuming unequal variance in populations and multiple comparisons were completed using the Holm-Sidak method with a definition of alpha = 0.05 to indicate significance. Primer sets can be found in other studies (Carnes et al., 2005, 2008; Etheridge et al., 2008; Acestor et al., 2009; Aphasizheva et al., 2009; Aphasizheva and Aphasizhev, 2010), with the exception of those listed in Table 1.

2.3. Generation of circTAIL amplicon libraries, sequencing, and read processing

The generation and processing of the dataset used here are described in Gazestani et al. (2016), where primers for the CO1 and pre-edited version of the cryptogene for CO3 (CO3p) amplicons are provided. Primers for generation of the ND1 amplicons are provided in Table 1. Supplementary Table S1 (expanded from Table 3 in Gazestani et al. (2016)), provides information on total reads analysed and the complexity of each sample population. Distributions of length and composition in Fig. 2 were inferred using the R statistical package. The area under each curve is 1. Raw reads are deposited in Sequence Read Archive, accession number SRP064265.

2.4. Positional probabilities

Single position frequency plots shown in Fig. 3A were produced with WebLogo (Crooks et al., 2004).

Table 1
Oligomers used to amplify either *Trypanosoma brucei* maxicircle-derived quantitative reverse transcription-PCR (qRT-PCR) fragments or generate *T. brucei* NADH dehydrogenase subunit 1 (ND1) transcript amplicons for circTAIL-seq analysis. Nucleotides of Illumina PCR primers which are regions complementary to the transcript being amplified are in uppercase (non-bold). Nucleotides which are bar codes used for multiplexed samples are in uppercase bold. Bold "NNNN" sequences represent the number of nucleotides of mixed composition at that position in the primer, used to offset homogeneity of the first sequenced nucleotides.

Primer	Sequence
ND7 3′ edit fwd, qRT-PCR	TGCCGGGTATATCATTTGCT
ND7 3′ edit rev, qRT-PCR	ACCACGCAAACAAACATCCA
ND9 edited fwd, qRT-PCR	GTTGTTGGAACGCGAATGT
ND9 edited rev, qRT-PCR	CTCCACCAACACACAAAATAACA
ND1 cDNA primer	ACGTTCACATAAACTAACATACC
ND1 5′ end, rev Illumina	aatgatacggcgaccaccgagatctacactctttccctacacgacgctcttccgatct**NNNN**GTATGCATATATCTAAATG
ND1 3′ end, fwd A Illumina	caagcagaagacggcatacgagat**GCCAAT**gtgactggagttcagacgtgtgctcttccgatct**NNNNN**TCATTTATTGCTATTACAAA
ND1 3′ end, fwd B Illumina	caagcagaagacggcatacgagat**CTTGTA**gtgactggagttcagacgtgtgctcttccgatct**NNNNNN**TCATTTATTGCTATTACAAA

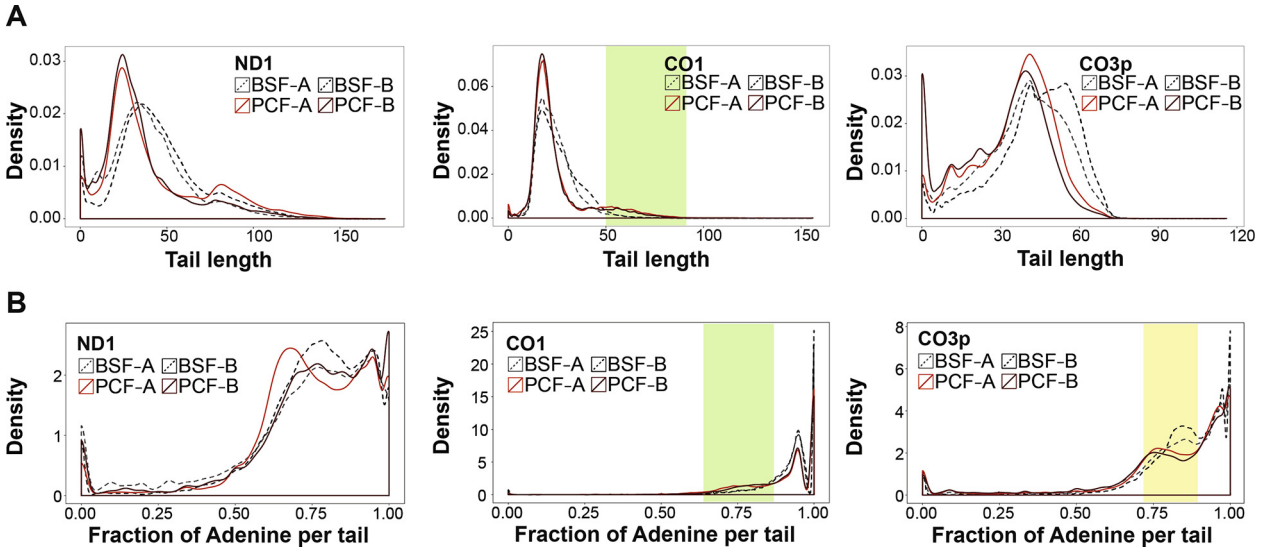

Fig. 2. Transcript tail length and compositional differences for *Trypanosoma brucei* transcripts coding for cytochrome c oxidase subunits I and III (CO1 and CO3) and NADH dehydrogenase subunit 1 (ND1). (A) Density curves of tail lengths of tails on indicated transcripts between populations from procyclic form and bloodstage form of *T. brucei*. Note scale differences of the X axes. CO1 tails of over 100 nucleotides (nt) are present in both procyclic form and bloodstage form populations but are not abundant enough to be observed on a chart of this scale. (B) Density curves of fraction nucleotides that are "A" in each tail (0 = tails that are oligo (U), 1 = tails that are oligo(A)) on indicated transcripts between tail populations from procyclic form and bloodstage form stage *T. brucei*. Yellow region indicates a population of tails of life stage shifted nucleotide composition in CO3p only. Green regions on CO1 plots indicate subpopulations of tails that were hypothesised to be the same tails in plots in (A) and (B). "A" and "B" within each plot are biological replicates of the same life stage. Represented results for CO3p and CO1 transcripts in the procyclic form life stage were obtained by re-analysis of previously published circTAIL-seq data (Gazestani et al., 2016). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.5. Modelling

Hidden Markov models (HMMs) were computed using the computational framework Sage (SageMath, the Sage Mathematics Software System (Version 6.1.1), The Sage Developers, 2015, http://www.sagemath.org) using the Baum-Welch algorithm for training to each set of sequences. Our HMMs allowed the possibility for each tail to possess states of addition that are consistent with existence as an in-tail only, or to initiate in in-tail state(s) and then transition to a state consistent with ex-tail addition. HMMs for each type of sample were trained using each of six tail populations separately, starting from the initial model shown in Supplementary Fig. S1B. If the most likely path through the model (using the Viterbi algorithm) contained the ex-tail state (grey circle), the tail was categorised as an ex-tail.

3. Results

3.1. Life stage differences in mtRNA abundances are varied and sometimes strain-specific

Our knowledge of stage-specific mtRNA abundances stems from multiple independent studies. Prior to investigating potential differences in tails, we first sought to gain a comprehensive picture of mtRNA abundance differences between the two proliferative life stages of cells grown in axenic culture. It was important to compare abundances between the stages of the strains used for inducible genetic studies, since the majority of molecular studies of trypanosome mitochondrial gene expression pathways that include life stage comparisons, including this one, use these strains (e.g. Ridlon et al., 2013). By virtue of primer selection, qRT-PCR is often more precise in the separate detection of edited and pre-edited forms of mitochondrial mtRNAs than RNA blots previously used in life stage comparisons. Furthermore, it can detect product over a wider range than the traditional RNA blots, allowing relative abundance calculations for even low abundance mtRNAs. The caveat is that for qRT-PCR, the abundance of only the ~100 nt amplified target is known for certain; the size information obtained in RNA blots is lost. To survey trypanosome mtRNA relative abundances under two different conditions, a qRT-PCR panel of mtRNAs is the preferred analytical method in even the most recent studies (Carnes et al., 2015; Suematsu et al., 2016; Zhang et al., 2017). Fig. 1B illustrates the qRT-PCR results for the common Lister 427-derived 29-13 (PCF) and Single Marker (BSF) cell lines.

Presumably mtRNAs are not transcriptionally regulated (Michelotti et al., 1992). Nevertheless, Fig. 1B reveals extensive

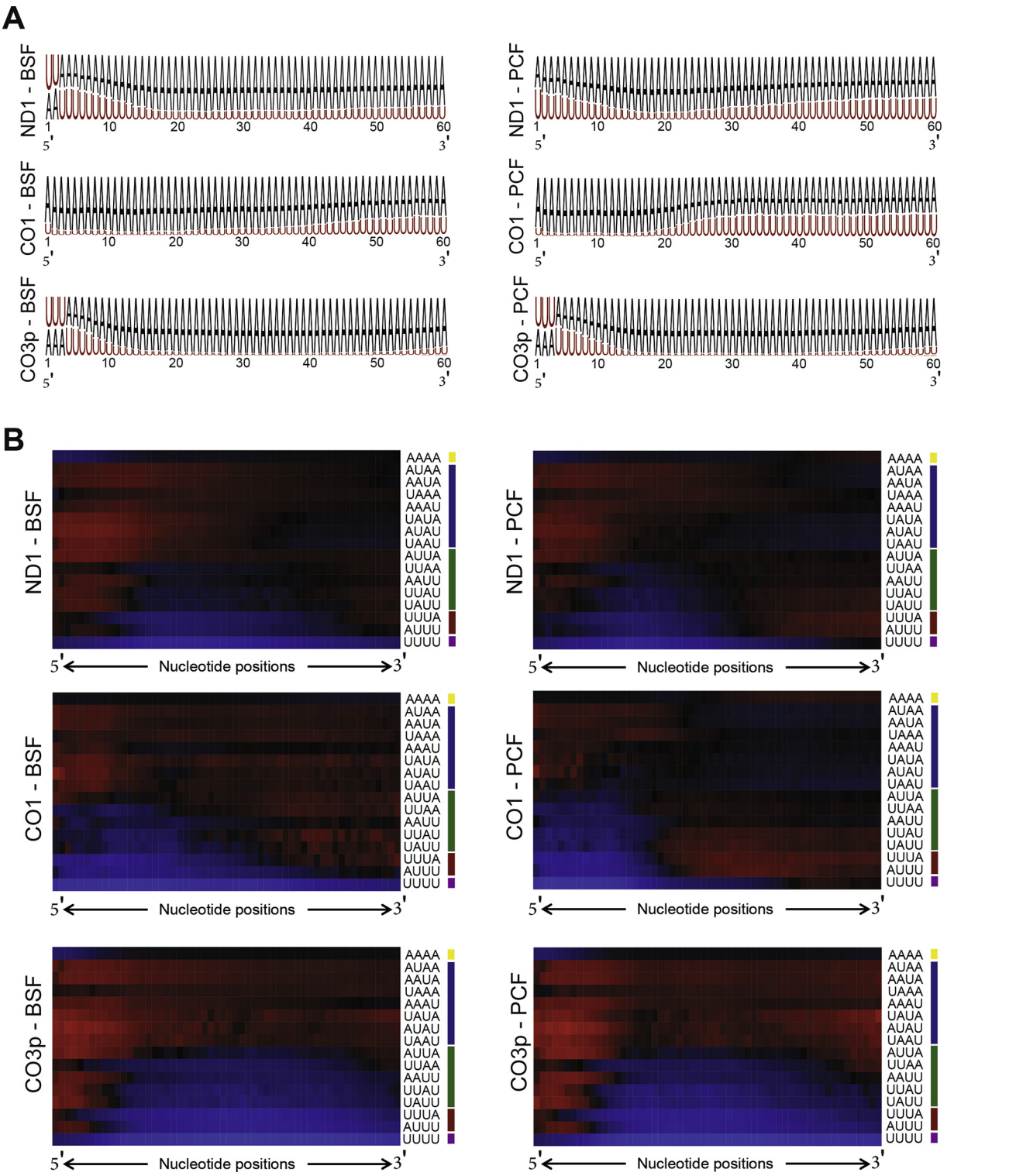

Fig. 3. Tail positional differences for *Trypanosoma brucei* transcripts coding for cytochrome c oxidase subunits I and III (CO1 and CO3) and NADH dehydrogenase subunit 1 (ND1). (A) Relative abundance of each nucleotide at each position from tail positions 1–60. All tails possessing a nucleotide at the analysed position in the total population were considered. (B) Heat map describing the occurrence of the indicated tetramer (nucleotide position 1 of the tetramer at the position indicated at the bottom), relative to the likelihood of that tetramer given the average nucleotide compositions at those positions under an independently distributed model. "Nucleotide positions" are tail positions 1–60 from 5′ to 3′. All plots are coloured using the same scale; the lowest (most intense red) value was 0.0117 (occurring 1/85 th as often as expected under an independently distributed model) and the highest (most intense blue) value was 4532 (occurring 4532 times as often as expected under an independently distributed model). At each position, the entire population of tails possessing a full tetramer starting at that position was analysed. The colour-coded bar to the right of the tetramers classifies those by how many consecutive Us they contain, an indicator of nucleotide switching frequency: yellow, 0 Us; blue, 1 U; green, 2 Us; red, 3 Us; and violet, 4 Us. Tails from biological replicates were combined for this analysis. The positional composition and tetramer probability plots for procyclic form CO3p and CO1 samples were drawn by the re-analysis of previously reported circTAIL-seq results (Gazestani et al., 2016).

Table 2
Trypanosoma brucei mtRNA relative life stage abundances determined by quantitative reverse transcription-PCR (qRT-PCR) that differ from relative abundances determined in previous studies by RNA blotting. Previous studies were performed prior to common use of a reference gene for normalisation.

mtRNA	Stage with highest abundance of mature form of mtRNA according to previous studies	References	Stage with highest abundance of mature form of mtRNA in 427-derived strain in this analysis (Fig. 1)
ND2	BSF[a]	Bhat et al. (1992)	BSF and PCF equivalent
ND3	BSF and PCF equivalent	Read et al. (1994)	PCF
ND5	BSF	Jasmer et al. (1985)	BSF and PCF equivalent
ND7 5′ region	BSF and PCF equivalent[a]	Koslowsky et al. (1990)	PCF
ND7 3′ region	BSF[a]	Koslowsky et al. (1990)	PCF[b]
ND9	BSF[a]	Souza et al. (1993)	PCF[b]
CO3	BSF and PCF equivalent[a]	Feagin et al. (1988)	PCF
A6	BSF and PCF equivalent[a]	Bhat et al. (1990)	PCF
RPS12	BSF	Jasmer et al. (1985) and Read et al. (1992)	PCF
CR4	BSF	Corell et al. (1994)	Unable to amplify published sequence of edited CR4 in either life stage

[a] RNA blots exhibited non-specific hybridisation of the maxicircle probes to cytosolic rRNAs for some measure of loading consistency (normalisation).
[b] Bloodstage form (BCF) when comparing with procyclic form (PCF) Lister 927-derived rather than Lister 427-derived line.

developmental gene expression regulation with a decreasing abundance trend for most tested RNA species in the BSF life stage. The fact that rRNAs and mtRNAs are differentially expressed – including those that do not undergo editing – confirms that RNA stability regulation can impact developmental mitochondrial gene expression. Interestingly, some amplicons primed with oligonucleotides of adjacent genes that capture abundance of preprocessed precursor mtRNA also exhibited abundance differences (shown in blue in Fig. 1B). This suggests that there may be multiple post-transcriptional regulatory layers contributing to *T. brucei* developmental mitochondrial gene expression.

Relative abundances of the edited ND7 3′ region (predicted to only exist in BSF) and edited ND9 contradicted those previously published (Table 2), so we confirmed amplification product identities by Sanger sequencing (not shown). To assess whether unexpected results were due to strain-specific differences, we compared abundances of mtRNA from PCF EATRO164 Istar1 strain that was utilised in most previous life stage abundance comparisons with those of BSF cells (Fig. 1B). EATRO 164 relative abundances trended similarly to those of 29-13, except for edited ND9, and the ND7 edited 3′ region that was undetectable in the PCF EATRO164 cells. These exceptions were in accordance with previous reports. Furthermore, two of our sequenced ND7 3′ region edited reads contained a different editing pattern than the one published. Therefore, stage-specific differences in editing may change over time in cultured cells, as possibly can occur with *Leishmania* spp. (Simpson et al., 2000), or may be the result of strain-specific differences in gRNA expression and editing (Koslowsky et al., 2014).

Additional differences between previous findings and ours are included in Table 2. These discrepancies may arise from the fact that earlier studies compared cultured PCF cells with BSF cells propagated in the rat or mouse bloodstream rather than growing both stages in culture as we have done. Indeed, mtRNA abundances vary in BSF cells during the course of infection (Feagin et al., 1986; Michelotti and Hajduk, 1987), and additionally, gene expression may alter when trypanosomes are on DEAE columns used for purification from animals (Mulindwa et al., 2014). Furthermore, in all cases where our results differed from those of prior studies, the prior studies lacked normalisation with an internal reference gene, and often they relied on imperfect probes (see studies in Table 2). Thus, it is possible that loading of RNA blots or the identity of material identified on the blots in at least some prior studies may have been skewed. In summary, we show that most mtRNAs decrease in abundance in the BSF compared with PCF stage with exceptions. We also observed that life stage-specific regulation varies between strains.

3.2. Tail populations exhibit inter-transcript and life stage-specific differences

Meta-analysis of multiple small-scale studies provides some evidence that in-tail inter-transcript differences exist (e.g. Kao and Read, 2007; Zimmer et al., 2012). However, they fail to examine whether characteristics of in-tails, regardless of the presence or absence of an ex-tailed population, are developmentally regulated. We hypothesised that in-tail differences could act as regulatory elements that govern developmental regulation of mtRNAs as observed in Fig. 1B. To test this hypothesis and more rigorously quantify tail differences between transcripts, we focused on the three transcripts of CO1, ND1, and CO3p. As detailed below, these transcripts differed in their developmental regulation based on our qRT-PCR results and prior knowledge.

CO1 is a never edited mtRNA with lower abundance in the BSF stage (Fig. 1B). As BSF cells lack cytochrome-containing electron transport chain complexes (Tielens and van Hellemond, 2009), the low numbers of CO1 mtRNAs present in BSF cells are possibly not translated, and hence, are likely not ex-tailed (Bhat et al., 1992). In contrast, the never-edited ND1 transcript does not change in abundance between the two life stages (Fig. 1B). We would expect some ND1 to be ex-tailed and translated in both life stages, since Complex I is present in both (Duarte and Tomas, 2014). Finally, CO3p also exists at consistent abundances in our developmental analysis. Unlike CO1 and ND1, CO3 lacks a translatable open reading frame prior to editing. Thus, CO3p should not be associated with the ribosome and is not likely ex-tailed in either proliferative life stage (Aphasizheva et al., 2011).

Tails of CO1, ND1 and CO3p transcripts were captured in the BSF life stage using circTAIL-seq, despite of the low abundance of the CO1 transcript in this stage. We also captured the tails of ND1 in the PCF life stage. To ensure the robustness of results, experiments were conducted using two independent biological replicates. These data were combined with our previous data on tails of CO1 and CO3p transcripts in the PCF life stage. The numbers of reads included in analyses ranged from 101,048 (one of the PCF ND1 replicates) to 1,003,267 (one of the BSF CO3p replicates). The

summary of sequencing depth and characteristics of each sample can be found in Supplementary Table S1.

Basic comparisons of population tail lengths were made using probability density curves (which show the relative frequency of tail lengths in a normalised manner) from the collected tails in each sample, with the area under each sample curve set to 1 (Fig. 2A). We observed transcript-specific differences in tail characteristics in BSF and PCF life stages, recapitulating and extending previous results on the PCF life stage from low depth experiments and our previous analysis (Kao and Read, 2007; Zimmer et al., 2012; Gazestani et al., 2016) and characterizing, to our knowledge for the first time, those in the BSF life stage. In both life stages, CO3p transcripts as a population have the longest tails followed by those of ND1, with CO1 having the shortest tails (*P* values <2.2E−16, Wilcoxon–Mann–Whitney rank sum test after pooling the replicates). More importantly, BSF tails were, overall, longer than that of the PCF stage for all three transcripts (*P* value <2.2E−16, Wilcoxon–Mann–Whitney rank sum test after pooling the replicates), demonstrating the existence of developmental differences in tail length distributions of mtRNAs.

Analysing probability density functions rather than median or mean tail length allows us to home in on additional but equally important length differences in the population. For example, as illustrated in Fig. 2A, CO3p and CO1 tail populations show a peak length density that is similar between life stages (17 nucleotides (nt) for CO1 and 42 nt for CO3p), yet, the tail population plots demonstrate that the BSF tails show a skewed distribution in the direction of increasing length. In contrast, the peak density length value for ND1 is shifted in BSF cells (25 nt in PCF cells and 33 nt in BSF cells). The tail population distributions also show the degree of tail length diversities can vary between transcripts, with the ND1 curve possessing a broadness of its peak that is intermediate to CO1 and CO3p. A wider distribution of tail lengths can imply a reduced stringency of length regulation for the tail addition machinery.

We also generated probability density functions to determine developmental and transcript-specific differences in tail composition of collected transcripts in each sample by reporting the percentage of A nucleotide, ranging from none (value of 0) to 100% (value of 1) per tail (Fig. 2B). U was the second most frequent nucleotide in tails and the percentages of C and G nucleotides were negligible. Nucleotide composition differences were dramatic between transcripts, with CO3p and ND1 incorporating far more Us than CO1. However, stage-specific differences between nucleotide compositions were subtle, with only a CO3p tail nucleotide composition peak at approximately 75% A in PCF cells shifted to 86% A in BSF cells (Fig. 2B, yellow region). This does not mean that there are no critical differences in the qualities of tail composition between transcripts, only that they are less likely to be found by analysing the total percentage of A per tail as the metric.

The combination of tail length and composition in tail populations can be an indicator of whether or not there are differences in ex-tail abundances between transcripts and life stages. For instance, CO1 tails in the PCF compared with the BSF life stage exhibit a significantly higher portion of tails with the expected compositional characteristics of ex-tails (Supplementary Fig. S2); i.e. longer tails ranging from 50 to 90 nt (green region, Fig. 2A) and possessing compositions in the neighbourhood of 70–75% A (green region, Fig. 2A). Therefore, CO1 tails with ex-tail characteristics are present in the PCF stage while nearly absent in the BSF stage. As shown in Supplementary Fig. S2, the ND1 density plots overlaying length and compositional characteristics showed a similar concentration of tails with ex-tail-like length and composition between life stages as expected. Supplementary Fig. S2 plots were less informative for CO3p, which possesses tails long enough to be ex-tails, despite being non-translatable.

3.3. Position and pattern of addition of As and Us in tails is population-specific

Tracking A and U composition as a function of nucleotide position for the tail population (Fig. 3A) is one way to determine whether inclusion of U is more likely to be occurring in in-tails or ex-tails within the population. Due to the observed high similarity of biological replicates in terms of tail characteristics, we combined both replicates into a single sample set for remaining analyses presented here. Positional composition differs in CO1 tails depending on the life stage analysed. In the PCF stage, Us are rare until approximately nt position 17, where those increase in frequency until they occur approximately 1/3 of the time at each position for the rest of the positional sites analysed. This observation is consistent with an in-tail to ex-tail transition at these sites around nt 17 within the tail population. The transition to higher U percentages later in the tails is much more gradual in BSF stage tails. The gradually increasing U composition at later positions in this instance could reflect the addition of a different type of tail other than ex-tails. ND1 also displays life stage-specific differences across nucleotide positions, with a frequency of U addition being more heavily weighted towards earlier positions in BSF cells than in PCF cells. In contrast, there are no life stage-specific CO3p positional differences in U composition, despite BSF tails being longer (Fig. 2A), and showing some overall compositional variation (Fig. 2B).

In our previous work (Gazestani et al., 2016), we also developed tetramer mapping as a way to trace the overall patterns of nucleotide switching along tail positions in each tail population. We now extend it to a different transcript and life stage to identify transcript and developmental differences that might exist between the tail populations. Fig. 3B shows deviation from the expected probability of each listed nucleotide tetramer at every tail position through nucleotide 60, tetramers being composed of strings of the same nucleotide and/or interspersed A and U combinations as shown. Blue represents tetramers occurring at a higher frequency in the population than would be predicted considering the total nucleotide compositions at those positions, with red representing lower frequency. Since the KPAF1-regulated A and U addition in ex-tails results in a high frequency of nucleotide switching, this metric is an additional indicator of whether an ex-tail subpopulation exists within a population, and it also could provide information about how nucleotides are added in in-tail regions. As an example, on CO3 tail populations from both BSF and PCF that are not expected to include ex-tails, the upper right region of the tetramer map is red, not blue, indicating a dearth of high frequency switching in the latter section of the long tails within these populations.

Interestingly, we observed over-representation of U polymer-containing tetramers (indicating less frequent switching of U to A) in CO1 and ND1 tails, occurring in certain regions of the tails. In CO1, U polymers (tetramers with two or more consecutive Us; bottom section of map) are over-represented in the initial part of the tail, while in ND1 the U polymer region of over-representation is shifted 3′ relative to that of CO1. These differences could reflect differences in access of A- versus U-adding enzymatic complexes to the nascent processed 3′ ends of these two mRNAs. In both cases, we also note that U polymers are more tightly confined to the beginning of tails in PCF cells compared with those in BSF cells. As U polymer tail initiation may be linked to processing of a monocistron newly excised from its polycistronic precursor, and to degradation pathways (Aphasizhev et al., 2016; Zhang et al., 2017), this may reflect a meaningful difference between the life stages. Overall, for populations that should not

include ex-tails (CO3p, and possibly CO1 from BSF cells), and within initial tail regions of tail populations expected to include ex-tails (ND1, and CO1 PCF), U homopolymers are over-represented. In contrast, tetramers containing single Us become over-represented only in the 3′ ends of longer CO1 (PCF) and ND1 tails where ex-tail additions with their frequent nucleotide switching are expected.

3.4. Nucleotides added early in tail formation do not influence ex-tail transition

Nucleotides added early in tail formation could influence downstream nucleotide addition including the transformation to an ex-tail. Therefore, we investigated short and long range associations of nucleotides within tails (Fig. 4, Supplementary Fig. S3). Shown are tail population heat maps of the deviation from the predicted frequency of A correlated with U at every nucleotide combination from position 1 to position 60 given the average nucleotide composition at the two sites being compared. The lack of bright blue or red on the heat maps indicates that no strong long range correlations exist for tails in either life stage. Although we did not detect strong long range associations, weak short range associations were observed that corroborate our other analyses. For instance, for some tail populations including PCF ND1, Fig. 3B suggests that it is rare to switch nucleotides throughout the initial 10 nt of the tail. In the ND1 association plot (Fig. 4), this is captured in negative A to U associations, as evidenced by red colour immediately adjacent to the diagonal in the first 10–20 nt of a tail. In summary, our association study suggests that initial sequences of tails are not determining factors for their final fates, at least for the three transcripts analysed.

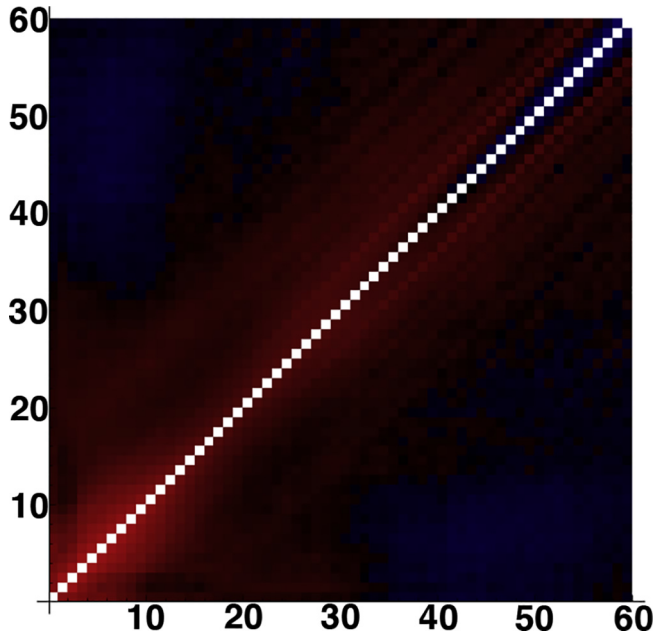

Fig. 4. *Trypanosoma brucei* heat map for procyclic form NADH dehydrogenase subunit 1 (ND1) transcript tail population showing the difference between the observed joint distribution of an A (*X* axis) and a U (*Y* axis), versus the product distribution given by the individual nucleotide frequencies at each pair of positions. Colour scale values ranged from −0.21 (most intense red; 21% less likely than expected under the product distribution) to 0.7 (most intense blue; 7% more likely than expected under the product distribution). Each pair takes into account all tails long enough to possess the positions being compared. BSF, bloodstage form.

3.5. HMM provides an unbiased measure of differences in U and A tail addition states

Presentations of tail characteristics in Figs. 2 and 3 thoroughly describe characteristics of each tail population. Nevertheless, succinct comparison of tail reads remained a challenge. It was also highly desirable to establish a procedure to determine the proportion of tails in a population that are ex-tails. We asked whether HMM could define states of A and U addition to address these challenges. The models would represent a largely unbiased view of differences in in-tails and tail extensions between samples. HMM for all six tail populations revealed homopolymer-adding states, and for all CO3p tail populations this is the only type of state appearing. However, some models possessed heteropolymer-adding states occurring after homopolymer-adding states that we define as an "ex-tail state". Models are depicted in Supplementary Fig. S1A, and data derived from the models are presented in Table 3.

As expected, states resembling ex-tails appear in ND1 tail models from both life stages, and in PCF-derived CO1 tail populations only. Overall, states that resembled ex-tails appeared in 22% of total tails of ND1 in BSF. A similar fraction was observed for ND1 and CO1 transcripts in PCF (Table 3). However, the median transition site of homopolymer to ex-tail addition state differed: 25 nt (BSF ND1) versus 21 nt (PCF ND1) versus 18 (PCF CO1). For both CO1 and ND1 PCF-derived tails, the modelled ex-tail transition position aligns remarkably well with the position in which U becomes increasingly abundant in Fig. 3A and at which likelihood of high-frequency nucleotide switching increases in Fig. 3B. For BSF tails of ND1, where these positions are ambiguous in Fig. 3A and B, modelling may the only way to identify ex-tail addition sites.

We also observed a tail quality that can only be determined with deep sequencing. We found that although the extensions that define ex-tails have A:U ratios approaching 7:3, they do differ between tail populations. For example, extensions from ND1 exhibited a ratio of 6.6:3.4 (A:U) from PCF cells and 7.5:2.5 from BCF cells. The concept of a single "style" of long tail addition mediated by KPAF1/2 importantly summarises an overall pathway. However, this idea may require expansion to include the possibility that different nucleotide ratios and patterns may strengthen or weaken the role of these tails in translation. Further, with HMM we can detect tails that do not conveniently fall into either the in-tail or ex-tail categories as we have defined them, such as the small fraction of BSF-derived CO1 tails that exhibit an addition state of 8.8:1.2 A:U ratio following homopolymer tail addition states. These tails with an average length of 31.5 nt are long relative to the typical population of CO1 tails that peaks at approximately 20 nt (Fig. 2A). Potentially, these may exist on CO1 mRNAs associated with a different type of enzymatic machinery such as a turnover complex. Finally, we also note that other differences such as the number of addition cycles that tails spend in one homopolymer state before transitioning to the other exist between these models (Supplementary Fig. S1A). Our long range association analysis (Fig. 4) suggested that these differences are likely unrelated to the addition of extensions. With sequencing of additional transcript tail populations, we expect that patterns among these state transitions will emerge.

3.6. Developmental differences exist in CO3p and ND1 5′ and 3′ lengths

circTAIL-seq reads can be exploited to identify 5′ and 3′ termini sites for a transcript population (Gazestani et al., 2016). In the case of 3′ termini, this directly influences the coded position at which the tail attaches to the transcript. This in turn could influence the structure and/or function of tails, particularly in light of recent links discovered between processing and tail addition (Zhang et al., 2017). mtRNA 5′ and 3′ termini sites were subtly different between life stages (Fig. 5). For the ND1 transcript, start site

Table 3
Summary of potential ex-tail states modelled in six tail populations from *Trypanosoma brucei* transcripts coding for cytochrome c oxidase subunits I and III (CO1 and CO3) and NADH dehydrogenase subunit 1 (ND1). Visual diagrams depicting the initial state model used in Hidden Markov models and final models are provided in Supplementary Fig. S1.

Transcript and life stage	Percentage of tails that modelled an additional state[a] (%)	Final A:U ratio of additional state[a]	Final status of additional state[a]	Median tail nucleotide position where tail extension begins
CO3p, BSF	0	9.6:0.4	No ex-tail state	–
CO3p, PCF	0	9.8:0.2	No ex-tail state	–
ND1, BSF	22	7.5:2.5	Ex-tail state	25
ND1, PCF	22	6.6:3.4	Ex-tail state	21
CO1, BSF	8	8.8:1.2	No ex-tail state	–
CO1, PCF	22	6.8:3.2	Ex-tail state	18

BSF, bloodstage form; PCF, procyclic form.
Shading indicates that A:U ratios and percentages of tails that modelled an additional state indicate that an ex-tailed state exists in the population.
[a]"Additional state" is a provided state option in the initial model that may or may not be heteropolymeric or even present in the modelled population.

heterogeneity was increased in the PCF compared with the BSF. A subtler CO3p UTR 3′ life stage difference, an increase in the subpopulation of the longest 3′UTR-containing transcripts, was also observed (seen primarily as a difference in termini within the sequence UUAUAUAAGGAAGA of CO3p 3′ UTR in Fig. 5). These data suggest that processing of maxicircle precursors differs between life stages. Also, they are consistent with reports where adjacent transcripts in a polycistron RNA compete for the UTR length; i.e., the longer 3′ region of the upstream can lead to a shorter 5′ UTR of the downstream gene or vice versa (Aphasizhev and Aphasizheva, 2011). However, UTR life stage differences are specific to the transcript analysed, as similar differences in the CO1 UTR populations between life stages were not observed. As previously reported (Gazestani et al., 2016), the three transcript populations analysed here in both stages demonstrate great transcript-specific variability in heterogeneity of termini.

4. Discussion

In this study, we undertook a global analysis of relative life stage mtRNA abundances and systematically analysed the characteristics of in-tails and ex-tails for the transcripts of CO1, ND1, and CO3 in two major life stages of *T. brucei*. Our data confirm that transcript-specific variations exist in both life stages and demonstrate how these variations differ developmentally. We also provide evidence of developmental changes in 5′- and 3′ UTRs of two of the mtRNAs. By using HMM as a tool, we distinguished differing states of non-templated tail populations, and revealed the average length of an in-tail at which it is extended into an ex-tail. Analysis of three mtRNAs with distinct developmental regulation of abundance or editing demonstrated that some aspects of developmental changes are transcript-specific including tail composition, presence of ex-tail, and UTR length, while another aspect was shared between the three: in-tail length. It should be noted that the observed developmental changes can be due to active regulatory mechanisms in the cell or passive factors that can affect enzymatic or RNA binding activity of proteins such as life-cycle mediated changes in the environmental temperature experienced by the parasite. However, our results suggest that there may be at least partial contribution of the former, as some aspects of developmental changes were transcript-specific.

When we examined life stage relative abundances of mtRNAs between the two strains that were also used for tail analysis, our results in large part modelled previous results from various studies. However, many transcripts that were thought more abundant in the BSF we found to be equally expressed in both stages; others thought to be equally expressed or expressed more abundantly in the BSF we found to be more abundant in the PCF (Table 2). While differences may be due to any of the technical reasons we listed in Section 3, we cannot rule out the possibility that our results which deviate from those expected, reflect the analysis of two laboratory strains grown independently for some time, rather than the true result of expression control from one stage to another. Mitochondrial mRNA abundance analysis by qRT-PCR of culture differentiation of a BSF polymorphic cell line to the PCF may be informative in this regard.

Oligo(U) regions were identified in tail populations of all three analysed transcripts from both life stages, but the number and position of oligo(U) regions were transcript-specific. Noted in previous studies (Aphasizheva et al., 2011; Zimmer et al., 2012) but quantified here is in-tail initiation with a short oligo(U) before poly(A) or A-rich addition. The oligo(U) tail initiation occurs rarely on CO1, more frequently on ND1 tails, but is the predominant tail type for CO3p tails. We also observed varying degrees of oligo(U) addition following oligo(A) addition occurring within in-tails. These results suggest that prior to ex-tailing, the order and strength of mtRNA associations with RET1 and KPAP1-containing complexes vary by transcript.

In general, it is possible that specific characteristics of in-tails or even ex-tails serve as regulatory elements. It is known that enzymes controlling tail addition, stability, editing and translation often reside in complexes that themselves are part of larger interconnected complex networks (Maslov and Agrawal, 2012; Aphasizheva et al., 2013, 2016; Hashimi et al., 2013; Ridlon et al., 2013; Aphasizhev et al., 2016; Zhang et al., 2017). For instance, certain pentatricopeptide (PPR) proteins complexed with KPAP1 and RET1 regulate their activities, and other PPR proteins appear to be involved in stability of specific mtRNAs (Mingler et al., 2006; Aphasizhev and Aphasizheva, 2013; Ridlon et al., 2013; Zhang et al., 2017) and/or play roles in translation (Pusnik et al., 2007; Ridlon et al., 2013; Aphasizheva et al., 2016). This regulatory network must be correctly targeted to specific mtRNAs in a developmentally appropriate manner. One possible mechanism for this intricate targeting is that tails of different compositions could have different affinities to the machinery or assume different secondary structures. In-tails that are primarily oligo(A) possess an inherently limited ability to form secondary structures. They will differ in accessibility from a population where most tails have regions of oligo(U) that will allow formation of hairpin structures. This in turn could influence the fate of the species of mtRNA carrying that tail, by influencing binding of regulatory proteins and complexes or even accessibility to other tail-adding enzymes.

Identification of life stage-specific differences in tail populations was a major aim of this work. The life stage-dependent difference consistent among the three analysed transcript populations was longer in-tails in BSF-derived populations. Lengths obtained from Fig. 2A paired with positional tail characteristics described in Fig. 3 and Supplementary Fig. S3 suggest that tails that are longer in BSF cells are in-tails, not ex-tails. Furthermore, HMM

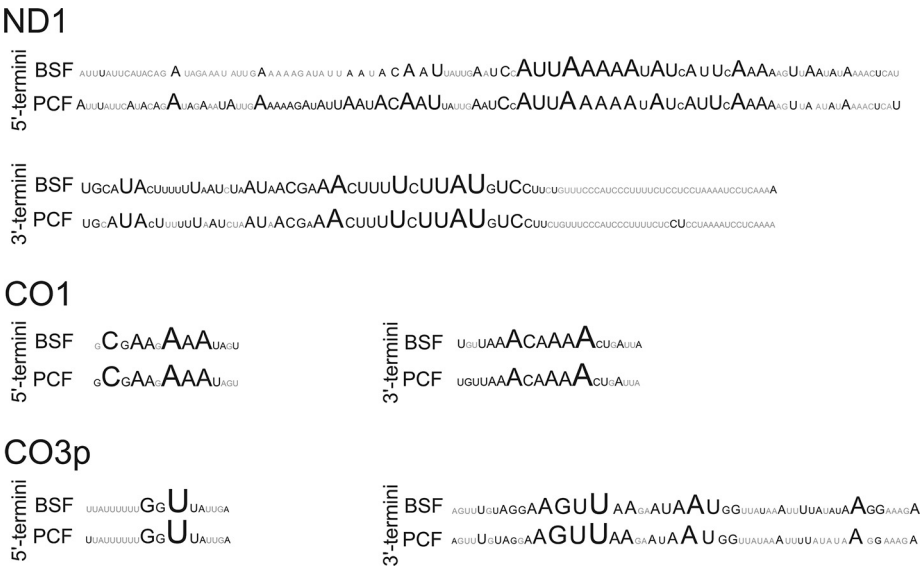

Fig. 5. Untranslated region (UTR) termini (3′ and 5′) derived from circTAIL-seq reads for *Trypanosoma brucei* transcripts coding for cytochrome c oxidase subunits I and III (CO1 and CO3) and NADH dehydrogenase subunit 1 (ND1). Replicate experiments were pooled by first normalizing for tail counts, so average density for each nucleotide is shown. Black represents the termini with occurrence probabilities higher than 0.01%, with the size of the nucleotide corresponding to its frequency as a terminus. The grey nucleotides represent positions with probabilities between 0.1% and zero that exist between nucleotide positions which are more frequently termini. Other nucleotides that are termini with a probability lower than 0.1% are not shown. Procyclic form CO3p and CO1 5′ and 3′ termini were extracted by re-analysis of previously reported circTAIL-seq data (Gazestani et al., 2016).

demonstrates that ND1 BSF-derived in-tails are longer before KPAF1/2-mediated extension, even when the percentage of tails that are ex-tails is equivalent in both life stages. This may indicate a delay or block in KPAF1/2 association with the tail-adding enzymes in the BSF in general. For CO1 but not ND1, this results in few or no ex-tailed molecules in BSF cells. Consequently, a smaller fraction of BSF CO1 is ribosome associated. CO1 ribosome association may be required for its stability and/or protection from degradation. Furthermore, we observed a developmental difference in 5′ UTR lengths within the ND1 transcript population. While this may have little to do with any developmental regulatory role for tails, it highlights how little we still understand about developmental regulation of mitochondrial gene expression at many levels.

Our study illustrates that we must continually evaluate and expand conclusions from earlier reports. Evaluation of KPAP1 and RET1 silencing studies must now include the consideration that different transcripts are targets of these enzymes to varying degrees, and that both in-tails and ex-tails are affected by silencing. Application of in vitro biochemical analyses to the situation within the mitochondrion must consider that in-tails from different transcripts will be more or less similar to the lengths and compositions of tails analysed in vitro. Silencing of a second putative mitochondrial poly(A)polymerase, kPAP2 (Kao and Read, 2007), has a very subtle phenotype. The role of KPAP2 in tail addition can now be better understood with experiments that were not technically feasible before. circTAIL-seq will prove very useful to enhance existing strategies of protein co-immunoprecipitations, over-expression and silencing studies, and in vitro assays to investigate how basic RNA editing, degradation and non-templated polymerisation activities are co-controlled by PPRs and other regulatory proteins.

Finally, using HMMs to pinpoint differences in populations of non-encoded tails will be useful in future studies that examine the nature and function of trypanosomatid mitochondrial tails. For example, we present here evidence of variability between tail populations (Figs. 2–6), and strain differences in developmental mitochondrial gene expression regulation (Fig. 1B). A natural extension of our work would be to determine whether the observed strain-specific differences are caused by differences in tail characteristics.

Acknowledgements

This study was supported through University of Minnesota, USA laboratory start-up funds to SLZ and Natural Sciences and Engineering Research Council of Canada (NSERC) grant #328186 to RS. VHG is supported by a Canadian Institutes of Health Research Systems Biology Fellowship. The Minnesota Supercomputing Institute (MSI) at the University of Minnesota provided resources that contributed to the research results reported within this paper. We wish to thank Rebecca A. Madden for her support in cloning qPCR amplicons for Sanger sequencing. The authors declare no competing financial interests.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.ijpara.2017.08.012.

References

Acestor, N., Panigrahi, A.K., Carnes, J., Zikova, A., Stuart, K.D., 2009. The MRB1 complex functions in kinetoplastid RNA processing. RNA 15, 277–286.
Aphasizhev, R., Aphasizheva, I., 2011. Mitochondrial RNA processing in trypanosomes. Res. Microbiol. 162, 655–663.
Aphasizhev, R., Aphasizheva, I., 2013. Emerging roles of PPR proteins in trypanosomes: Switches, blocks, and triggers. RNA Biol. 10, 1495–1500.
Aphasizhev, R., Aphasizheva, I., Simpson, L., 2003. A tale of two TUTases. Proc. Natl. Acad. Sci. U.S.A. 100, 10617–10622.
Aphasizhev, R., Sbicego, S., Peris, M., Jang, S.H., Aphasizheva, I., Simpson, A.M., Rivlin, A., Simpson, L., 2002. Trypanosome mitochondrial 3′ terminal uridylyl transferase (TUTase): the key enzyme in U-insertion/deletion RNA editing. Cell 108, 637–648.
Aphasizhev, R., Suematsu, T., Zhang, L., Aphasizheva, I., 2016. Constructive edge of uridylation-induced RNA degradation. RNA Biol. 13, 1078–1083.
Aphasizheva, I., Aphasizhev, R., 2010. RET1-catalyzed uridylylation shapes the mitochondrial transcriptome in *Trypanosoma brucei*. Mol. Cell. Biol. 30, 1555–1567.
Aphasizheva, I., Aphasizhev, R., 2016. U-insertion/deletion mRNA-editing holoenzyme: definition in sight. Trends Parasitol. 32, 144–156.

Aphasizheva, I., Maslov, D., Wang, X., Huang, L., Aphasizhev, R., 2011. Pentatricopeptide repeat proteins stimulate mRNA adenylation/uridylation to activate mitochondrial translation in trypanosomes. Mol. Cell 42, 106–117.
Aphasizheva, I., Maslov, D.A., Aphasizhev, R., 2013. Kinetoplast DNA-encoded ribosomal protein S12: a possible functional link between mitochondrial RNA editing and translation in *Trypanosoma brucei*. RNA Biol. 10, 1679–1688.
Aphasizheva, I., Maslov, D.A., Qian, Y., Huang, L., Wang, Q., Costello, C.E., Aphasizhev, R., 2016. Ribosome-associated pentatricopeptide repeat proteins function as translational activators in mitochondria of trypanosomes. Mol. Microbiol. 99, 1043–1058.
Aphasizheva, I., Ringpis, G.E., Weng, J., Gershon, P.D., Lathrop, R.H., Aphasizhev, R., 2009. Novel TUTase associates with an editosome-like complex in mitochondria of *Trypanosoma brucei*. RNA 15, 1322–1337.
Aphasizheva, I., Zhang, L., Wang, X., Kaake, R.M., Huang, L., Monti, S., Aphasizhev, R., 2014. RNA binding and core complexes constitute the U-insertion/deletion editosome. Mol. Cell. Biol. 34, 4329–4342.
Bhat, G.J., Koslowsky, D.J., Feagin, J.E., Smiley, B.L., Stuart, K., 1990. An extensively edited mitochondrial transcript in kinetoplastids encodes a protein homologous to ATPase subunit 6. Cell 61, 885–894.
Bhat, G.J., Souza, A.E., Feagin, J.E., Stuart, K., 1992. Transcript-specific developmental regulation of polyadenylation in *Trypanosoma brucei* mitochondria. Mol. Biochem. Parasitol. 52, 231–240.
Brenndorfer, M., Boshart, M., 2010. Selection of reference genes for mRNA quantification in *trypanosoma brucei*. Mol. Biochem. Parasitol. 172, 52–55.
Carnes, J., Trotter, J.R., Ernst, N.L., Steinberg, A., Stuart, K., 2005. An essential RNase III insertion editing endonuclease in *Trypanosoma brucei*. Proc. Natl. Acad. Sci. U. S.A. 102, 16614–16619.
Carnes, J., Trotter, J.R., Peltan, A., Fleck, M., Stuart, K., 2008. RNA editing in *Trypanosoma brucei* requires three different editosomes. Mol. Cell. Biol. 28, 122–130.
Carnes, J., Lerch, M., Kurtz, I., Stuart, K., 2015. Bloodstream form *Trypanosoma brucei* do not require RPN1 for gRNA processing. RNA 21, 28–35.
Corell, R.A., Myler, P., Stuart, K., 1994. *Trypanosoma brucei* mitochondrial CR4 gene encodes an extensively edited mRNA with completely edited sequence only in bloodstream forms. Mol. Biochem. Parasitol. 64, 65–74.
Crooks, G.E., Hon, G., Chandonia, J.M., Brenner, S.E., 2004. WebLogo: a sequence logo generator. Genome Res. 14, 1188–1190.
Decker, C.J., Sollner-Webb, B., 1990. RNA editing involves indiscriminate U changes throughout precisely defined editing domains. Cell 61, 1001–1011.
Duarte, M., Tomas, A.M., 2014. The mitochondrial complex I of trypanosomatids–an overview of current knowledge. J. Bioenerg. Biomembr. 46, 299–311.
Etheridge, R.D., Aphasizheva, I., Gershon, P.D., Aphasizhev, R., 2008. 3′ adenylation determines mRNA abundance and monitors completion of RNA editing in *T. brucei* mitochondria. EMBO J. 27, 1596–1608.
Feagin, J.E., Abraham, J.M., Stuart, K., 1988. Extensive editing of the cytochrome c oxidase III transcript in *Trypanosoma brucei*. Cell 53, 413–422.
Feagin, J.E., Jasmer, D.P., Stuart, K., 1985. Apocytochrome b and other mitochondrial DNA sequences are differentially expressed during the life cycle of *Trypanosoma brucei*. Nucleic Acids Res. 13, 4577–4596.
Feagin, J.E., Jasmer, D.P., Stuart, K., 1986. Differential mitochondrial gene expression between slender and stumpy bloodforms of *Trypanosoma brucei*. Mol. Biochem. Parasitol. 20, 207–214.
Feagin, J.E., Jasmer, D.P., Stuart, K., 1987. Developmentally regulated addition of nucleotides within apocytochrome b transcripts in *Trypanosoma brucei*. Cell 49, 337–345.
Feagin, J.E., Stuart, K., 1985. Differential expression of mitochondrial genes between life cycle stages of *Trypanosoma brucei*. Proc. Natl. Acad. Sci. U.S.A. 82, 3380–3384.
Fidalgo, L.M., Gille, L., 2011. Mitochondria and trypanosomatids: targets and drugs. Pharm. Res. 28, 2758–2770.
Gazestani, V.H., Hampton, M., Abrahante, J.E., Salavati, R., Zimmer, S.L., 2016. circTAIL-seq, a targeted method for deep analysis of RNA 3' tails, reveals transcript-specific differences by multiple metrics. RNA 22, 477–486.
Hashimi, H., Zimmer, S.L., Ammerman, M.L., Read, L.K., Lukes, J., 2013. Dual core processing: MRB1 is an emerging kinetoplast RNA editing complex. Trends Parasitol. 29, 91–99.
Jasmer, D.P., Feagin, J.E., Stuart, K., 1985. Diverse patterns of expression of the cytochrome c oxidase subunit I gene and unassigned reading frames 4 and 5 during the life cycle of *Trypanosoma brucei*. Mol. Cell. Biol. 5, 3041–3047.
Kao, C.Y., Read, L.K., 2005. Opposing effects of polyadenylation on the stability of edited and unedited mitochondrial RNAs in *Trypanosoma brucei*. Mol. Cell. Biol. 25, 1634–1644.
Kao, C.Y., Read, L.K., 2007. Targeted depletion of a mitochondrial nucleotidyltransferase suggests the presence of multiple enzymes that polymerize mRNA 3' tails in *Trypanosoma brucei* mitochondria. Mol. Biochem. Parasitol. 154, 158–169.
Koslowsky, D., Sun, Y., Hindenach, J., Theisen, T., Lucas, J., 2014. The insect-phase gRNA transcriptome in *trypanosoma brucei*. Nucleic Acids Res. 42, 1873–1886.
Koslowsky, D.J., Bhat, G.J., Perrollaz, A.L., Feagin, J.E., Stuart, K., 1990. The MURF3 gene of *T. brucei* contains multiple domains of extensive editing and is homologous to a subunit of NADH dehydrogenase. Cell 62, 901–911.
Maslov, D., Agrawal, R., 2012. Mitochondrial translation in trypanosomatids. In: Bindereif, A. (Ed.), RNA Metabolism in Trypanosomes. Springer-Verlag Berlin Heidelberg, Germany, pp. 215–236.
Michelotti, E.F., Hajduk, S.L., 1987. Developmental regulation of trypanosome mitochondrial gene expression. J. Biol. Chem. 262, 927–932.
Michelotti, E.F., Harris, M.E., Adler, B., Torri, A.F., Hajduk, S.L., 1992. *Trypanosoma brucei* mitochondrial ribosomal RNA synthesis, processing and developmentally regulated expression. Mol. Biochem. Parasitol. 54, 31–41.
Mingler, M.K., Hingst, A.M., Clement, S.L., Yu, L.E., Reifur, L., Koslowsky, D.J., 2006. Identification of pentatricopeptide repeat proteins in *Trypanosoma brucei*. Mol. Biochem. Parasitol. 150, 37–45.
Mulindwa, J., Fadda, A., Merce, C., Matovu, E., Enyaru, J., Clayton, C., 2014. Methods to determine the transcriptomes of trypanosomes in mixtures with mammalian cells: the effects of parasite purification and selective cDNA amplification. PLoS Negl. Trop. Dis. 8, e2806.
Pusnik, M., Small, I., Read, L.K., Fabbro, T., Schneider, A., 2007. Pentatricopeptide repeat proteins in *Trypanosoma brucei* function in mitochondrial ribosomes. Mol. Cell. Biol. 27, 6876–6888.
Read, L.K., Lukes, J., Hashimi, H., 2016. Trypanosome RNA editing: The complexity of getting U in and taking U out. Wiley Interdiscip. Rev. RNA 7, 33–51.
Read, L.K., Myler, P.J., Stuart, K., 1992. Extensive editing of both processed and preprocessed maxicircle CR6 transcripts in *Trypanosoma brucei*. J. Biol. Chem. 267, 1123–1128.
Read, L.K., Wilson, K.D., Myler, P.J., Stuart, K., 1994. Editing of *Trypanosoma brucei* maxicircle CR5 mRNA generates variable carboxy terminal predicted protein sequences. Nucleic Acids Res. 22, 1489–1495.
Ridlon, L., Skodova, I., Pan, S., Lukes, J., Maslov, D.A., 2013. The importance of the 45 S ribosomal small subunit-related complex for mitochondrial translation in *Trypanosoma brucei*. J. Biol. Chem. 288, 32963–32978.
Roy Chowdhury, A., Bakshi, R., Wang, J., Yildirir, G., Liu, B., Pappas-Brown, V., Tolun, G., Griffith, J.D., Shapiro, T.A., Jensen, R.E., Englund, P.T., 2010. The killing of African trypanosomes by ethidium bromide. PLoS Pathog. 6, e1001226.
Ryan, C.M., Militello, K.T., Read, L.K., 2003. Polyadenylation regulates the stability of *Trypanosoma brucei* mitochondrial RNAs. J. Biol. Chem. 278, 32753–32762.
Simpson, L., Thiemann, O.H., Savill, N.J., Alfonzo, J.D., Maslov, D.A., 2000. Evolution of RNA editing in trypanosome mitochondria. Proc. Natl. Acad. Sci. U.S.A. 97, 6986–6993.
Souza, A.E., Myler, P.J., Stuart, K., 1992. Maxicircle CR1 transcripts of *Trypanosoma brucei* are edited and developmentally regulated and encode a putative iron-sulfur protein homologous to an NADH dehydrogenase subunit. Mol. Cell. Biol. 12, 2100–2107.
Souza, A.E., Shu, H.H., Read, L.K., Myler, P.J., Stuart, K.D., 1993. Extensive editing of CR2 maxicircle transcripts of *Trypanosoma brucei* predicts a protein with homology to a subunit of NADH dehydrogenase. Mol. Cell. Biol. 13, 6832–6840.
Suematsu, T., Zhang, L., Aphasizheva, I., Monti, S., Huang, L., Wang, Q., Costello, C.E., Aphasizhev, R., 2016. Antisense transcripts delimit exonucleolytic activity of the mitochondrial 3′ processome to generate guide RNAs. Mol. Cell 61, 364–378.
Tielens, A.G., van Hellemond, J.J., 2009. Surprising variety in energy metabolism within trypanosomatidae. Trends Parasitol. 25, 482–490.
Verner, Z., Basu, S., Benz, C., Dixit, S., Dobakova, E., Faktorova, D., Hashimi, H., Horakova, E., Huang, Z., Paris, Z., Pena-Diaz, P., Ridlon, L., Tyc, J., Wildridge, D., Zikova, A., Lukes, J., 2015. Malleable mitochondrion of *Trypanosoma brucei*. Int. Rev. Cell Mol. Biol. 315, 73–151.
Zhang, L., Sement, F.M., Suematsu, T., Yu, T., Monti, S., Huang, L., Aphasizhev, R., Aphasizheva, I., 2017. PPR polyadenylation factor defines mitochondrial mRNA identity and stability in trypanosomes. EMBO J. 36, 2435–2454.
Zimmer, S.L., McEvoy, S.M., Menon, S., Read, L.K., 2012. Additive and transcript-specific effects of KPAP1 and TbRND activities on 3′ non-encoded tail characteristics and mRNA stability in *Trypanosoma brucei*. PLoS One 7, e37639.