

Trypanosoma cruzi strain and starvation-driven mitochondrial RNA editing and transcriptome variability

EVGENY S. GERASIMOV,^{1,2} ROGER RAMIREZ-BARRIOS,^{3,6} VYACHESLAV YURCHENKO,^{4,5}
and SARA L. ZIMMER³

¹Faculty of Biology, M.V. Lomonosov Moscow State University, Moscow 119991, Russia

²Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow 127051, Russia

³Department of Biomedical Sciences, University of Minnesota Medical School, Duluth Campus, Duluth, Minnesota 55812, USA

⁴Life Science Research Centre, Faculty of Science, University of Ostrava, 710 00 Ostrava, Czech Republic

⁵Martsinovskiy Institute of Medical Parasitology, Tropical and Vector Borne Diseases, Sechenov University, Moscow 119435, Russia

ABSTRACT

Trypanosoma cruzi is a unicellular protistan parasitic species that is comprised of strains and isolates exhibiting high levels of genetic and metabolic variability. In the insect vector, it is known to be highly responsive to starvation, a signal for progression to a life stage in which it can infect mammalian cells. Most mRNAs encoded in its mitochondrion require the targeted insertion and deletion of uridines to become translatable transcripts. This study defined differences in uridine-insertion/deletion RNA editing among three strains and established the mechanism whereby abundances of edited (and, thus, translatable) mitochondrial gene products increase during starvation. Our approach utilized our custom T-Aligner toolkit to describe transcriptome-wide editing events and reconstruct editing products from high-throughput sequencing data. We found that the relative abundance of mitochondrial transcripts and the proportion of mRNAs that are edited varies greatly between analyzed strains, a characteristic that could potentially impact metabolic capacity. Starvation typically led to an increase in overall editing activity rather than affecting a specific step in the process. We also determined that transcripts *CR3*, *CR4*, and *ND3* produce multiple open reading frames that, if translated, would generate different proteins. Finally, we quantitated the inherent flexibility of editing in *T. cruzi* and found it to be higher relative to that in a related trypanosomatid lineage. Over time, new editing domains or patterns could prove advantageous to the organism and become more widespread within individual transcriptomes or among strains.

Keywords: Chagas disease; RNA editing; metabolism; epimastigote; electron transport chain

INTRODUCTION

Trypanosoma cruzi, the causative agent of Chagas disease in humans and other mammals, is an insect-transmitted protistan parasite of the family Trypanosomatidae (Kostygov et al. 2021). These parasites have evolved numerous novel features, some of which relate to the uniquely structured genome of the single mitochondrion residing in each cell (Maslov et al. 2019). A sufficient understanding of the molecular biology and biochemistry of *T. cruzi* is complicated by its genetic diversity, encompassing both its nuclear and mitochondrial genome (Ruvalcaba-Trejo and Sturm 2011; Zingales 2018). It requires progression

through various life stages to cause human disease as part of its normal life cycle (Tyler and Engman 2001).

Several key cellular processes occur in the mitochondrion. One of the main processes is oxidative phosphorylation coupled with ATP generation, which is executed by a complex of multiple proteins of the electron transport chain (ETC) termed complexes I–IV (cI–cIV), and ATP synthase (Vercellino and Sazanov 2021). While cII in trypanosomatids is comprised of subunits that are entirely nuclear-encoded, those of cI, cIII, cIV, and ATP synthase all have at least one subunit that is encoded in the mitochondrial genome (Škodová-Sveráková et al. 2015). Often the subunits encoded in the mitochondrion (numbering up to 16 in *T. cruzi*) are key catalytic components. They are highly

⁶**Present address:** Department of Biomedical Sciences and Pathobiology, Virginia-Maryland College of Veterinary Medicine, Virginia Tech, Blacksburg, VA, USA

Corresponding author: szimmer3@d.umn.edu

Article is online at <http://www.majournal.org/cgi/doi/10.1261/ma.079088.121>.

© 2022 Gerasimov et al. This article is distributed exclusively by the RNA Society for the first 12 months after the full-issue publication date (see <http://majournal.cshlp.org/site/misc/terms.xhtml>). After 12 months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

hydrophobic and typically embedded in the mitochondrial membrane. It is these physical properties that are thought to be responsible for the difficulty in detecting them through means of mass spectrometry or specific antibodies, making their study very challenging.

In *T. cruzi*, as in other trypanosomatids, the protein-coding loci are located on dozens of near-identical maxicircle molecules of the mitochondrial genome (Callejas-Hernández et al. 2021). The study of their products is also challenging at the RNA level. This is because 12 of the 18 protein-encoding loci of its mitochondrial genome do not possess the correct coding sequence (these loci are called cryptogenes) and must be edited post transcriptionally with the targeted insertion and deletion of one or more uridines (U's) (Sturm and Simpson 1990). This U-insertion/deletion (indel) editing is unique to the mitochondrial genomes of kinetoplastids, to which trypanosomes belong (Lukeš et al. 2018, 2021). A few transcripts are edited at only a small domain, but most are edited throughout their length (pan-edited) except for their very ends (which are usually their very short UTRs). The enzymatic steps and processivity of U-indel editing rely on a large group of proteins encompassing both enzymatic and regulatory components (Cruz-Reyes et al. 2018; Aphasizheva et al. 2020).

The targeting and specific nature of each editing event (specific site and number of U's inserted or deleted) relies on information inherent to the sequences of guide (g) RNAs, small noncoding RNAs encoded on minicircles (Camacho et al. 2019; Rusman et al. 2021). These mitochondrial DNA molecules number in the thousands and are categorized into distinct classes of molecules of nearly identical sequence (Cooper et al. 2019; Callejas-Hernández et al. 2021; Gerasimov et al. 2021). RNA editing begins at the 3' end of a transcript or edited domain and proceeds towards its 5' end, typically utilizing multiple successive gRNAs to achieve a completely edited product. Therefore, a trypanosomatid mitochondrial transcriptome consists of mRNAs correctly encoded as transcribed (termed never-edited), transcripts prior to entering the editing process (preedited), those in the process of being edited (partially edited), and completely edited. Additionally, U's may be inserted or deleted in different numbers or positions than what is required to generate the expected final edited product. These might be dead-end products, normal intermediates to a correctly edited product, or, if translatable, may represent an alternative mature product (Zimmer et al. 2018). The overall or transcript-specific regulation of the initiation, speed, efficiency, or nature of RNA editing could certainly impact abundances of the mitochondrial mRNA products (Zimmer et al. 2018). Indeed, we previously determined that fully edited or near fully edited mitochondrial mRNAs increase in abundance during starvation in *T. cruzi* insect stage (epimastigote) cells. Such changes were far smaller when the primers for the reverse transcription quantitative polymerase chain reaction

(RT-qPCR) assay were targeted to the preedited transcript rather than its edited form, suggesting that editing was stimulated upon starvation (Shaw et al. 2016; Ramirez-Barrios et al. 2020). As nutrient depletion or absence is an essential step for transition from the replicative epimastigote to the infective trypomastigote *T. cruzi* life stage (Goldenberg and Avila 2011), understanding both the mechanism behind and reason for this starvation-induced change in gene expression is critical.

When investigating molecular and metabolic pathways in *T. cruzi*, its high degree of intra species genetic heterogeneity of both its nuclear and mitochondrial genomes must be accounted for. The species is comprised of clusters of genetically similar groups termed discrete typing units (DTUs). DTUs I-VI all harbor strains that have been obtained from human infections to differing extents (Brenière et al. 2016). Adding further complexity, the diploid nuclear genomes of DTUs V and VI are hybrid (Majeau et al. 2021). A species' particular genetic identity is important because it almost certainly impacts its fitness in both its insect and human hosts. The relationships between the genetic identity of a species and its ability to cause disease or progress through a sylvatic or domestic transmission cycle has barely begun to be adequately addressed (Zingales 2018). Specific to our study, it will be important to determine which impacts of starvation on mitochondrial mRNA are specific to a strain or strains, and which are conserved across the species. Differences in maxicircle sequence and abundance of nuclear transcripts encoding mitochondrial ETC subunits between strains have been previously identified (Ruvalcaba-Trejo and Sturm 2011; Kalem et al. 2018). Notably, cell line differences in editing were noted to be a concern for the model trypanosome *Trypanosoma brucei* (Kirby and Koslowsky 2020) even though the species possesses a much lesser degree of genetic heterogeneity.

This current study ultimately focuses on the mechanisms by which abundances of the edited mitochondrial gene products are increased during starvation. This first required an understanding of any differences in editing within the *T. cruzi* species, and between *T. cruzi* and other analyzed trypanosomatids. As there was evidence of differences in expression of nuclear-encoded components of the ETC among strains (Kalem et al. 2018), we compared editing and mRNA abundance among them. We found that the relative abundance of the various mitochondrial transcripts and the proportion of mRNAs that are edited varies greatly within the species. Specifically, editing in cultured epimastigotes of the reference strain CL Brener is radically limited, while Sylvio X/10 possesses the most heterogeneous mRNA population. We also demonstrated that editing of the mitochondrial transcripts *CR3*, *CR4*, and *ND3* generates multiple molecules with edited translatable open reading frames (ORFs), and uncovered instances of unusual or strain-specific editing patterns in the mitochondrial

mRNAs CO2 and A6. To address how *T. cruzi* editing may be similar or different to other species, we presented a method to quantitate the degree to which editing events in a kinetoplastid mitochondrial genome represent functional editing over transient, nonproductive, or alternative editing events.

Prior to testing mechanisms by which abundances of the edited mitochondrial gene products are increased during starvation, we recognized theoretical possibilities. One possibility is that the edited mRNA species are selectively stabilized during starvation. Another is that overall rate of the editing process is upregulated during starvation, shifting the balance of transcripts towards fully edited mRNAs. A third possibility is that the rate of editing remains the same, but the accuracy of the process is increased, with fewer nonprocessive editing events occurring that are inconsistent with a single canonical and translatable fully edited sequence. Using our high-throughput sequencing approach, we determined that at least for *T. cruzi* strain Sylvio X/10, the sheer number of editing events per transcriptome increases during starvation.

RESULTS

Putative products of different strains of *T. cruzi* can be assembled, but the primary products of several loci are ambiguous

We analyzed *T. cruzi* editing in three strains with known differences in maxicircle sequence and differences in abundances of relevant nuclear transcripts (Ruvalcaba-Trejo and Sturm 2011; Kalem et al. 2018). These are: Sylvio X/10 of DTU I, Esmeraldo of DTU II, and the reference sequenced strain CL Brener of DTU VI. We collected cells at medium-high logarithmic growth density. Initially, we obtained reads for two biological replicate samples of each strain; however, their subsequent analysis strongly suggested that they can be merged (Pearson correlation coefficient r ranged between 0.95 and 1 for duplicates, Supplemental Fig. S1). In addition, we analyzed maxicircle-derived reads from sequencing of total and mitochondrial RNA samples. As previously reported for *L. pyrrhocoris* (Gerasimov et al. 2018), the samples enriched for mitochondrial RNA contained considerably more reads in total that were mitochondrial in origin. Thus, they are preferred for collection despite the extra steps and potential variability involved. The correlation plot indicates segregation of the strains yet tight clustering of replicates derived from total and mitochondrial samples (Supplemental Fig. S1). Thus, biological replicates and the total/mitochondrial-enriched samples were merged for each strain, unless otherwise noted.

Until now, the sequences of many fully edited mRNAs have not been determined for the three strains analyzed. In fact, a full complement of experimentally derived edited maxicircle translatable sequences is unavailable for any

T. cruzi strain. Therefore, our first task was to reconstruct as many *T. cruzi* ORFs as possible for the analyzed strains. To do that we employed T-Aligner (Gerasimov et al. 2018), a program previously used for successful reconstruction of edited mRNAs of *L. pyrrhocoris* (Gerasimov et al. 2021). Results of this reconstruction are summarized in Supplemental Table S1 and discussed below.

A first finding is that we were unable to assemble edited MURF2 (maxicircle unidentified reading frame 2) for the strains CL Brener and Esmeraldo, despite it being edited at only a small domain. This was due to insufficient coverage across the transcript, although reads spanning its small, edited domain were available. In general, we were less successful in assembling edited mRNAs of these two strains than those of the strain Sylvio X/10. This is because (1) the CL Brener mitochondrial transcriptome has by far the lowest percentage of edited reads, and (2) we had fewer libraries of the strain Esmeraldo to work with, as the starvation experiments were performed in the other two strains only. T-Aligner reconstructions utilized all available read populations collected for this study, including reads from starved populations. Coverage of each reference loci for the Esmeraldo libraries was typically more uneven, making it difficult to reconstruct full ORFs, even when the percentage of edited reads was high.

For CL Brener, we assembled minimally edited CO2 (cIV subunit) and CYb (cIII subunit), as well as pan-edited CO3 (cIV subunit) and RPS12 (encoding a mitoribosome subunit) that are among the smaller transcripts. For Esmeraldo, we were able to assemble CO3 and RPS12. For Sylvio X/10, ORF assembly of a single major edited product was possible for all cryptogenes except for CR3 (C-rich locus 3), CR4 (C-rich locus 4), and ND3 (encoding a cI subunit). If there were more than one full-length ORF identified for each of these pan-edited cryptogenes listed above, the ORFs were conserved across their length, differing in a narrow locus only. Nevertheless, what we are hereby defining as the “canonical” editing product for each of these pan-edited cryptogenes must be acknowledged to merely be the most likely primary edited isoform.

Due to editing 3' to 5' directionally, the proportion of reads that are edited in the downstream part of a transcript is much higher than that of its upstream portion. Edited reads covering the extreme 5'-edited transcripts have very low representation in the total read pool of each transcript. This challenge to ORF reconstruction is illustrated in coverage plots of cryptogenes throughout this study. Enhancing the total number of edited reads for extreme 5' ends of transcripts is the main reason we merged all Illumina reads for each transcript for ORF reconstruction.

T-Aligner is designed to return the longest possible ORFs that can be assembled from reads derived from each maxicircle loci. However, if there are insufficient edited reads within the entire maxicircle mRNA population to cover upstream parts of edited transcripts, T-Aligner will

often fail to return even a truncated transcript. This is because ORFs are not generated until a start codon, either canonical or noncanonical, has been identified in the course of assembly. This phenomenon of fewer edited reads in upstream regions of pan-edited transcripts is common to all species studied thus far (Zimmer et al. 2018). The assembly challenges for *CR3*, *CR4* and *ND3* were different. In these cases, T-Aligner assembled many full-length ORFs. These ORF collections often had high sequence variability at their 5' and 3' ends and one of several consensus regions in the middle of the mRNA. Therefore, we could not report a "single" canonical product for these cryptogenes.

In the case of *CR3*, *CR4*, and *ND3*, a relevant question is the degree to which the various putative edited ORFs are read supported for a specific cryptogene. This is challenging to determine because the read support for the pattern of editing varies across an ORFs length. Therefore, we rely on a visualization that we term a "coding capacity array" to convey an estimation of relative support among dozens or more T-Aligner-generated ORFs. For each transcript, an array was generated by first translating all ORFs and breaking the translations into a series of peptides (successive 6-mers in Fig. 1) that become the "nodes" in the coding capacity array. Arrows connect all 6-mers with the previous and successive amino acid sequence for all translated

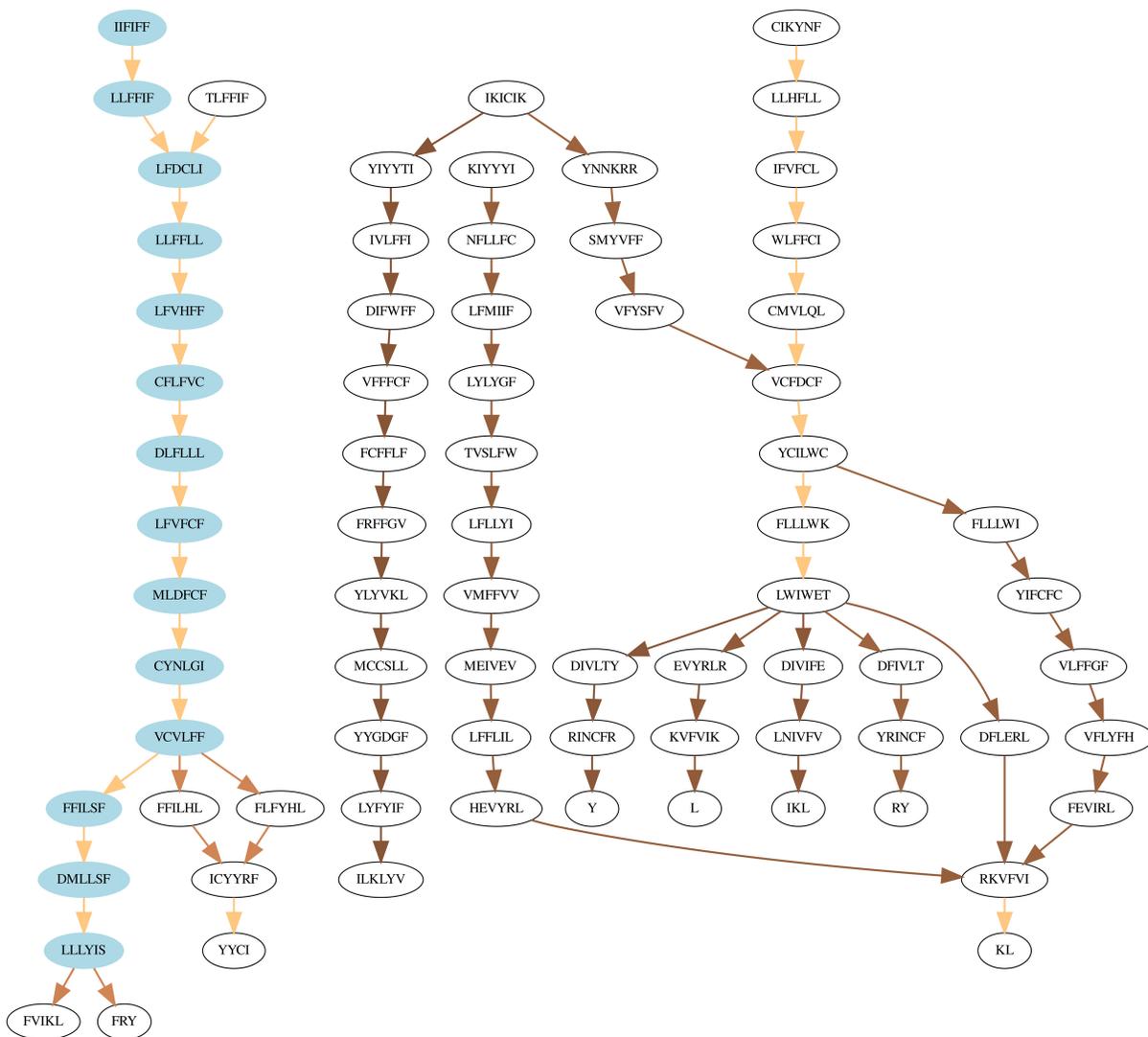


FIGURE 1. T-Aligner reconstructions of *Trypanosoma cruzi* edited *CR3* reads. Coding capacity graph for strain Sylvio X/10 *CR3* based on T-Aligner reconstructions with Illumina sequencing reads. T-Aligner translated ORFs over 78 amino acids long were separated into consecutive series of amino acid 6-mers. The coding capacity is presented as arrows connecting the successive amino acid 6-mer nodes that, when traced, represent the product of an ORF reconstructed by the program. The amino acid pattern colored in blue depicts the longest ORF generated by the T-Aligner (thus, chosen as the "canonical" peptide). Arrow intensities reflect cumulative read support based on the median read support for the entire translated ORFs. Light color: highest read support. Dark color: lowest read support.

ORFs. In this scheme, the protein translation of any reconstructed ORF is the successive addition of all amino acid 6-mers in its path through the coding capacity array. To estimate the relative support of each of these paths, we first calculated a median read support value for each entire ORF reconstructed with T-Aligner. These values can then be cumulatively applied to each segment of the constructed coding capacity graph and visualized by the color intensity of the arrows surrounding each segment (node).

An example is provided in the coding capacity graph in Figure 1 for the small ORF CR3. In no species has a translated CR3 product been assigned a function. This mRNA is known to be the subject of alternative editing in *T. brucei* (Kirby and Koslowsky 2020). T-Aligner preferentially weights longer products in the reconstruction process, so the preferred ORF displayed in Figure 1 (in blue with an isoleucine start codon) is simply the longest one reconstructed. It has both multiple amino- and carboxyl termini. T-Aligner also reconstructed additional ORFs coding a separate array of proteins (right side of Fig. 1). Nearly two dozen paths can be taken through the entire coding capacity graph from the various amino- to the various carboxyl termini, each representing a reconstructed ORF. However, the situation is somewhat simpler because all these paths utilize one of four main routes, with each main route yielding a completely different sequence. Differing termini on each of these four predicted protein groups is where a high degree of amino acid diversity is localized.

Notably, the central domain of the longest translated ORF (in blue) has approximately the same level of support as the amino terminus and central region of the translated ORFs initiating with the 6-mer "CIKYNF" (Fig. 1). If all CR3 ORFs are translated, we would expect both protein products to be well represented. However, each would be present as multiple isoforms varying in the terminal ~1–20 amino acids. No isoform of either ORF translation appears to be predominant. As the coding region of edited CR3 is scarcely larger than the Illumina read length, it is unlikely that T-Aligner is making artifactual reconstructions from partially edited mRNAs. The assembly of multiple products likely reflects the biological state of the transcriptome. Our data do not support the premise that a single edited canonical product exists for CR3, CR4, and ND3.

Findings of unexpected editing

CO2

The CO2 mRNA presented some interesting features related to its editing. In *T. cruzi* (Kim et al. 1994), as is also the case for *T. brucei* (Golden and Hajduk 2005, 2006), CO2 is edited in a very narrow domain near the 3' terminus, with guiding thought to be directed by a *cis*-acting gRNA in its 3' UTR. Only four U's are added in three successive editing sites, with a +2, +1, +1 addition pattern. The editing

domain is followed by several in-frame stop codons in that species, so if no editing occurs, translation of the mRNA would lead to premature termination (Benne et al. 1986). The situation in the analyzed *T. cruzi* strains is different. If unedited CO2 were to be translated, a longer amino acid sequence than that of the canonical edited sequence would be generated, as no stop codon was apparent within CO2 and its adjacent flanking loci. Any unedited translated CO2 transcripts would produce a minor protein product, as among reads covering this editing domain, ~95% appear to be fully edited.

The most read-supported ORFs for the CL Brener and Sylvio X/10 strains are the sequences that are edited at the three consecutive expected positions with no additional editing at any other domain. Their protein products differ by three conservative amino acid substitutions because of maxicircle sequence variation. T-Aligner could not assemble the expected ORF for Esmeraldo, but that is likely because of unusually poor read coverage in this strain at the 5' end of the mRNA. A feature of editing detected in all three strains was a low level of specific editing activity at a region starting ~100 nt from the CO2 start codon. Some level of editing in normally unedited domains was previously mentioned as likely nonfunctional (Gerasimov et al. 2018). However, for *T. cruzi* CO2, the editing converged on a single pattern that affected a region of less than ten amino acids, ultimately resulting in conservation of the reading frame upstream and downstream of the edited region (Supplemental Fig. S2). This was observed in both CL Brener and Sylvio X/10 and detected in the few reads covering this region in Esmeraldo. Thus, some protein products may have an alternative amino terminus.

A6

The strain Sylvio X/10 was the only one for which we were able to reconstruct the translatable mRNA for pan-edited A6 (ATP synthase subunit). In examining the read and editing coverage across this transcript, the reasons for failure to assemble a single product may be strain-specific. In the case of CL Brener editing, when compared to Sylvio X/10 editing, a more severe 3'–5' drop-off in level of editing is apparent (Fig. 2). A loss of sufficient processivity of editing needed to generate the minimal number of upstream editing reads may be responsible for this. In contrast, in the strain Esmeraldo, the incidence of edited reads extends further into the 5' region than it does in CL Brener. However, there may be another impediment to assembly of a single product in this strain—a heavily edited domain in the central region of the transcript. While editing in this region is also apparent for the other two strains, it is far less abundant. We surmise that the editing occurring in Esmeraldo in this domain may be unique to the strain. It is possible that this strain possesses several unique gRNAs that can divert A6 transcripts in the process of being canonically edited to a

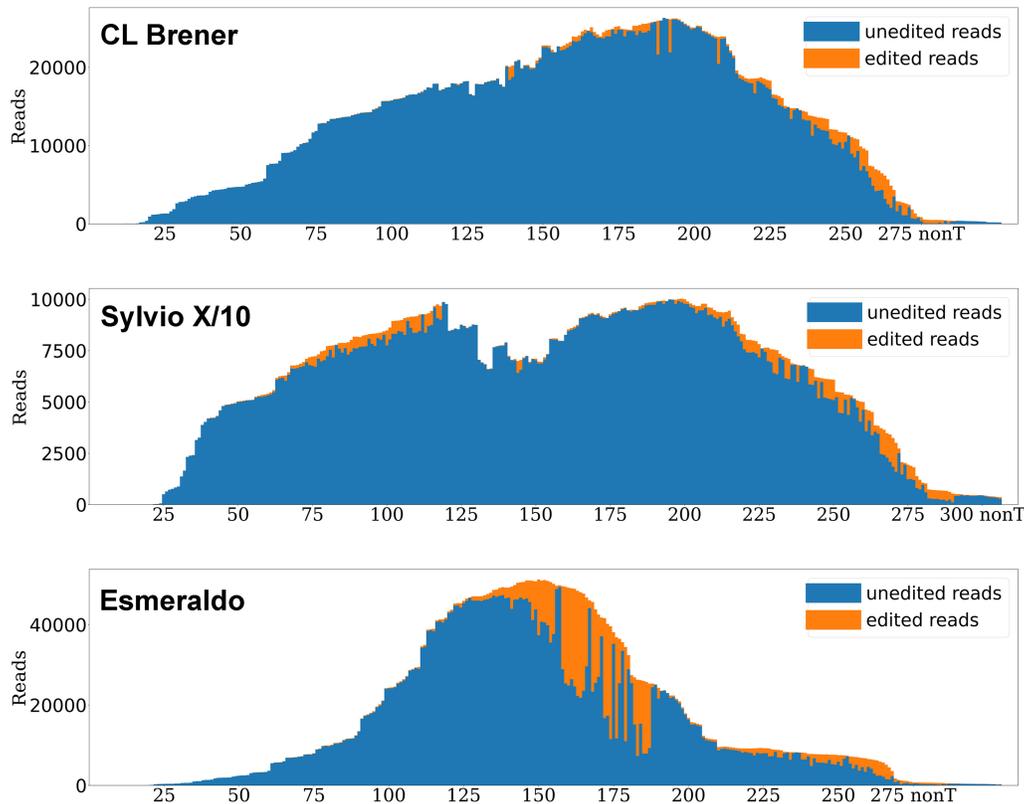


FIGURE 2. A6 transcript read coverage and editing for three *Trypanosoma cruzi* strains. Read coverage of the A6 maxicircle locus is presented with coverage level on the Y-axis. For every non “T” position on the locus represented 5’ to 3’ on the X-axis, the total number of reads containing that nucleotide position are shown. The fraction of them that are edited at that position are displayed in orange (top), while the fraction of unedited reads is depicted in blue (bottom).

noncanonical pattern, thus greatly reducing the portion of A6 that will be translatable. Such gRNAs may be missing in CL Brener and Sylvio X/10. In summary, examining patterns of editing across *T. cruzi* strains and mitochondrial transcripts allowed us to discover editing anomalies for CO2 and A6 that reveal an underlying flexibility in this species.

***Trypanosoma cruzi* strain identity impacts abundance of mitochondrial mRNAs**

We next examined the degree to which mitochondrial mRNAs are differentially expressed in *T. cruzi*, that is, the degree to which abundances vary between transcripts. We analyzed the abundance of total read coverage per transcript and that for the transcripts possessing at least three editing events (U insertion or deletion relative to its maxicircle cryptogene sequence). These “edited reads” constituted 24%–94% of total reads depending on the strain analyzed. The Sylvio X/10 reads had been previously used to demonstrate the features of the T-Aligner program that we were using for assignment of edited reads to maxicircle loci (Gerasimov et al. 2018). The Esmeraldo and CL Brener reads were obtained at the same time as the reads

for the Sylvio X/10 strain. The major expressed mRNA products of the kinetoplast genome for all analyzed strains are the transcripts *ND7* and *ND8* (cl subunits), *RPS12*, and *CO3*. The mRNA *CO2* is fairly abundant in the Sylvio X/10 and Esmeraldo, but not in the CL Brener strain (Fig. 3). This abundance pattern may be partially explained by a transcriptional positional effect: The mRNA positioned immediately after the two highly transcribed rRNAs and encoded in the same orientation is *ND8*. Further, the next transcripts in that same orientation are highly expressed *ND7* and *CO3* (cl-encoding *ND9* and mitoribosome subunit *RPS3* are encoded in the opposite orientation). In a single polycistronic transcriptional unit (rRNAs, *ND8*, and, perhaps, also *ND7*, *CO3*, *CYb* and *A6*), the tail-off abundance of the later mRNAs would be the expected expression pattern. The *ND8* transcript is also highly abundant in *Leptomonas pyrrocoris* (Gerasimov et al. 2021), but that is not the case for the more closely related *Trypanosoma vivax* Y486 (Greif et al. 2015). Differences in starting material, library preparation, mapping protocols, or relative maxicircle transcript abundances may all underlie this discrepancy. Another evident pattern was the lower abundance of all the transcripts not requiring editing relative to those that do (the single exception being Sylvio X/10 *CO1* (cIV

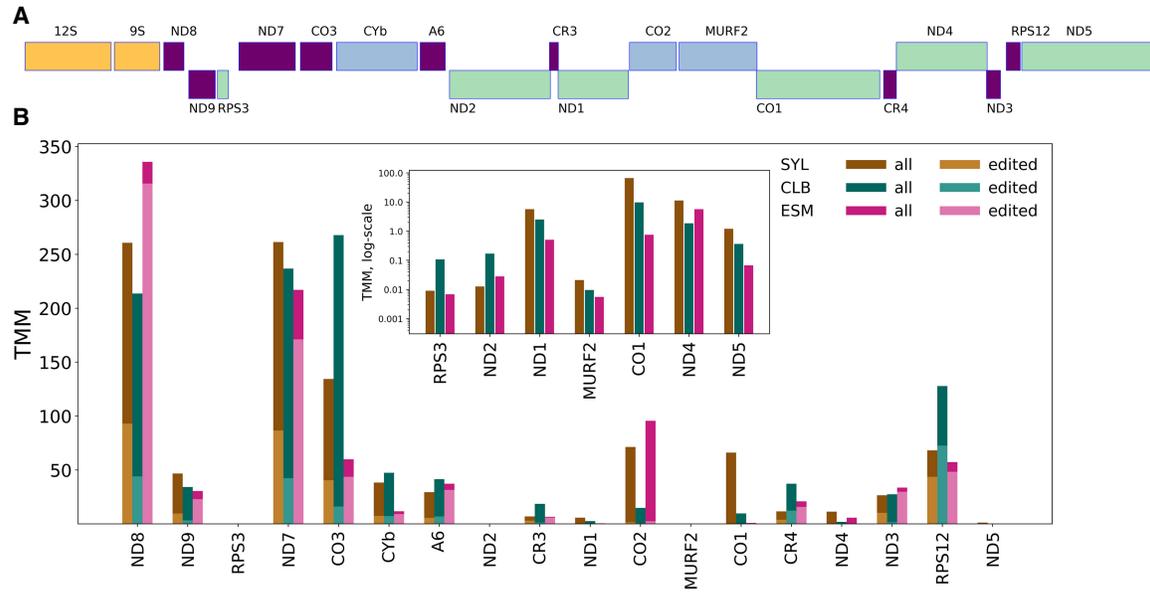


FIGURE 3. Overall patterns of mitochondrial mRNA abundances show similarity across analyzed strains of *Trypanosoma cruzi*. (A) Visual representation of the coding region of the *T. cruzi* maxicircle. The orientation of transcripts is indicated by their position on the top (forward) or bottom (reverse) of the maxicircle coding region depiction. rRNA loci are presented in yellow. Loci that correctly encoded their products are indicated in green, loci producing transcripts requiring only a small domain of RNA editing are shown in blue, and those that produce transcripts requiring editing throughout most of its length are shown in violet. (B) Mitochondrial mRNA abundances derived from total reads or reads containing at least three editing events. The transcripts that are edited are ND8, ND9, ND7, CO3, CYb, A6, MURF2, CR3, CO2, CR4, ND3, and RPS12. Mitochondrial transcripts are ordered from left to right by the order at which they appear on the maxicircle, regardless of orientation. The inset shows all never-edited transcripts and low-abundant MURF2 on a log-scale Y-axis. The low proportions of reads containing editing events aligning to transcripts in the inset (Table 1) are hidden. SYL, strain Sylvio X/10. CLB, strain CL Brener. ESM, strain Esmeraldo.

subunit) (Fig. 3 and inset). In fact, despite being correctly encoded on the maxicircle, the never-edited ND2 (cl subunit) and RPS3 reads are not present in high enough abundance to cover the entire putative ORF. The transcripts per million (TMM) values used in Figure 3 are reported in Supplemental Table S2.

Although ratios of the main expressed transcripts are largely conserved among the strains, strain-specific differences are apparent. Of the three investigated strains, Sylvio X/10 has the most diverse transcriptome, with more reads originating from less abundant transcripts (e.g., ND1, ND4, and ND5, encoded cl subunits, as well as CO1). Reads for the CO1 transcript are largely absent in the strain Esmeraldo, while CO2 reads are relatively rare in the strain CL Brener. Also, the levels of CO3 and RPS12 in CL Brener are unusually high relative to the other two analyzed strains. Taking all our findings together, as with other deep sequencing mitochondrial transcriptomes of culture-grown trypanosomatids (Cooper et al. 2019; Callejas-Hernández et al. 2021; Gerasimov et al. 2021), individual transcript abundances vary greatly and differ to a lesser extent by strain.

One thing that is consistent among all the strains is that transcripts that are relatively abundant are also abundant when only their edited populations were considered (Fig. 3), reducing the probability of widespread strain-specific

differential editing of entire transcripts. This is consistent with our previous results by RT-qPCR analysis (Shaw et al. 2016; Ramirez-Barrios et al. 2020). However, we also wanted to understand whether the proportion of edited reads is different among strains. We therefore determined the proportion of edited reads for each mRNA for *T. cruzi* Sylvio X/10, CL Brener, and Esmeraldo strains (Table 1) from raw TMMs (Supplemental Table S2).

Finally, there are considerable overall differences in the degree to which editing of maxicircle-derived mRNAs occurs between *T. cruzi* strains. Most maxicircle-derived reads show little editing in the strain CL Brener, with the exception of RPS12 (56% of edited reads) and CR4 (34%). In total, only 24% of the CL Brener reads are edited. In contrast, 42% and 94% of the maxicircle-derived reads in Sylvio X/10 and Esmeraldo strains, respectively, were edited.

***Trypanosoma cruzi* editing is particularly stochastic, a feature conserved among strains, but variable among individual transcripts**

Our inability to determine a single canonical edited product for several Sylvio X/10 transcripts highlights unresolved questions related to editing. First, what is the relative representation of productive compared to nonproductive editing events in *T. cruzi*? Second, is this ratio conserved

TABLE 1. Percentage of reads of each maxicircle locus that contain three or more U insertion or deletions relative to the corresponding maxicircle DNA sequence (i.e., edited)

Transcript	CL Brener		Esmeraldo		Sylvio X/10		CL Brener			Sylvio X/10		
	Mito	Total	Mito	Total	Mito	Total	Fed	Abrupt	Slow	Fed	Abrupt	Slow
ND8	22	14	93	93	35	33	18	21	19	52	71	77
ND9	9	10	75	72	20	17	9	4	6	17	18	29
<i>RPS3*</i>	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
ND7	18	14	78	76	32	30	24	13	22	35	46	63
CO3	6	6	73	85	29	26	13	5	6	45	47	73
<u>CYb</u>	14	16	78	75	19	17	7	5	6	10	8	7
A6	16	16	82	82	19	20	23	22	21	30	52	71
ND2	0	3	4	0	2	0	1	0	0	0	5	9
CR3	7	6	87	85	52	74	3	2	2	81	88	98
ND1	1	1	23	20	0	0	3	1	1	0	1	1
<u>CO2</u>	3	3	2	2	2	1	6	5	4	4	4	4
MURF2*	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
CO1	0	0	17	27	0	0	0	0	0	0	1	1
CR4	32	34	75	72	33	32	45	36	36	41	48	72
ND4	0	0	1	0	0	0	0	0	0	0	1	1
ND3	5	4	88	87	38	38	5	3	4	47	63	69
RPS12	56	56	83	84	65	65	63	51	58	61	84	88
ND5	10	7	22	10	0	0	6	6	5	1	1	1
Total	24	21	94	93	42	35	24	19	22	37	35	51

Left-most grouping of three strains: *Trypanosoma cruzi* mitochondrial transcriptome libraries used in strain comparisons. Right-most grouping of two strains: for *T. cruzi* strains CL Brener or Sylvio X/10, percentage of reads that are edited for *T. cruzi* in Fed, Abrupt starvation, and Slow starvation cultivation conditions. Mitochondrial transcripts are ordered from top to bottom by the order in which they appear on the maxicircle, regardless of orientation. Bolded loci (left-most column) are those transcripts requiring editing to generate ORFs; those that are additionally underlined only require editing at a small domain; and loci that are neither bolded nor underlined are not typically edited. Their products are correctly encoded on the maxicircle. Asterisked loci lack sufficient read coverage for assessment of percentage of edited reads (not available, N/A).

among other trypanosomatids? We attempted to answer these questions and compared our findings to the previously analyzed *L. pyrrocoris* (Gerasimov et al. 2021) as a model phylogenetically distant trypanosomatid species.

Simple numerical answers to these questions rely on our ability to definitively identify the mitochondrial edited mRNAs that are translated, which we currently lack. However, we can begin answering them by focusing on the mRNAs for which we were able to reconstruct a predominant ORF in one or more species: *A6*, *CO2*, *CO3*, *CYb*, *ND7*, *ND8*, *ND9*, *MURF2*, and *RPS12*. For every site that was edited in these major reconstructed ORFs, we asked what fraction of the reads that exhibited editing at that site was edited with the correct number of U's inserted or deleted to match that in what we previously identified as the canonical ORF for that cryptogene. Noncanonical editing at the same site was defined as a lack of editing or else U insertions or deletions to numbers that are inconsistent with the canonical ORF. Therefore, noncanonical editing events at any site could potentially include both editing events that would result in alternative isoforms and editing events that would be nonproductive for generating a translatable

mRNA. However, this analysis does not provide a way to distinguish between these options.

Examples of this analysis are presented in Figure 4 for *CO3* and Figure 5 for *RPS12*, with the rest available in Supplemental Figure S3. For many sites, the vast majority of the edited output was canonical. However, there were also numerous sites for which a quarter or more of the mapped reads possessed noncanonical edited states. For a few sites, nearly all editing was noncanonical (asterisked in Fig. 4). For *CO3* specifically, we observed that the total level of editing differed among strains. These differences were concentrated at the 5' termini. While the strain Esmeraldo displays a consistently high percentage of canonical editing across the transcript, the strain Sylvio has a dip in the canonical editing at its 3' end, and the percentage of canonical editing for CL Brener was lower in the upstream regions of *CO3*. Due to uneven coverage of the longer transcripts in these libraries, we were only able to compare editing across all three strains for a handful transcripts and could not determine whether canonical to non-canonical editing ratios across transcripts are universally strain-specific.

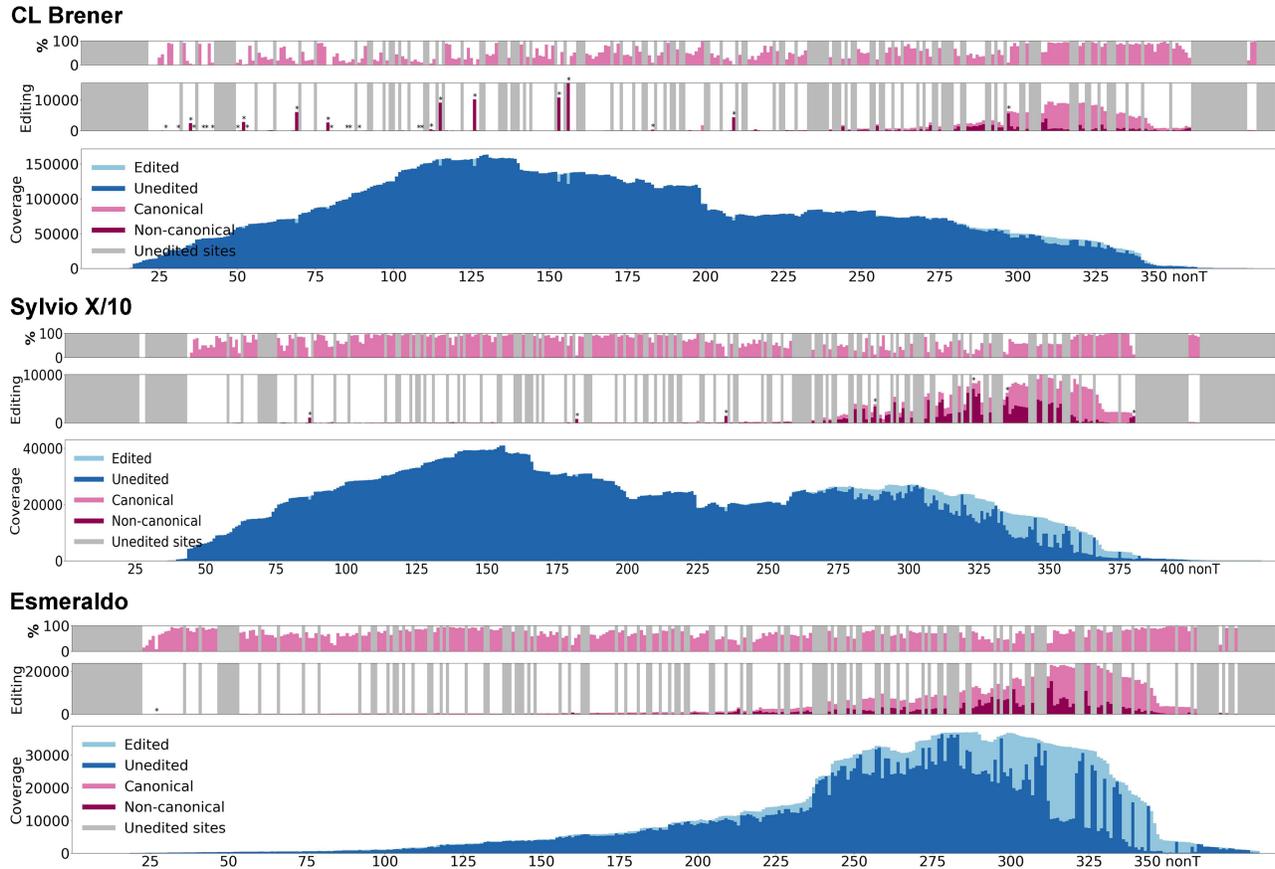


FIGURE 4. Relative levels of canonical and noncanonical editing of *CO3* in different strains of *Trypanosoma cruzi*, CL Brener, Sylvio X/10, and Esmeraldo. For each strain, the *bottom* plot depicts edited and nonedited read coverage across the *CO3* transcript, with the X-axis depicting every non “T” position of the *CO3* loci from 5’ to 3’. The portion of reads edited at each location along the transcript is shown in a lighter blue. The *middle* plot is a bar plot with the X-axis as above. At each editing site the bar distinguishes the total number of reads, in which an editing event at that site is observed and breaks down the number by canonical and noncanonical editing events. The A, C, and G positions along the transcript that are not locations of editing for canonical edited sequence are blocked out in gray and not analyzed. The *top* plot shows a similar bar graph, but the Y-axis represents percentage of editing that is canonical rather than absolute numbers of reads possessing editing at each site. Asterisks indicate positions where noncanonical editing exceeds 90% of the total.

Across all transcripts in all libraries, there was no obvious pattern for position in the transcript where the majority-noncanonical sites appear (5’ or 3’ region of the editing domain). Nor was there any correlation between how highly edited a site was and the degree to which it was canonically edited. One clear pattern was that a very high percentage of editing occurred at sites within minimal edited domains (those of *CO2* and *CYb*). This phenomenon of high percentage of editing of short domains has been described in *T. brucei*, as transcripts with shorter edited domains have a simpler path to editing completion (Tylec et al. 2019).

In our previous analysis of noncanonical RNA editing in *L. pyrrocoris*, we examined portions of edited reads aligning to each specific gRNA editing region for the entire transcriptome. We determined that ~85% of the sequences in a region aligning to a particular gRNA possess canonical sequence, while ~15% possess an alternative editing

pattern (Gerasimov et al. 2021). As apparent for the *L. pyrrocoris* transcript *RPS12* (Fig. 5, top), most editing (>80%) is indeed consistent with the single canonical edited *RPS12* product, in strong agreement with the previous overall calculations. However, for the *T. cruzi* transcripts for which a canonical ORF is apparent, most have at least a few editing sites at which the proportions of canonical to noncanonical editing is reversed (over 80% of edited reads possess a U indel of a length inconsistent with that necessary for the canonical sequence; Fig. 5, middle, bottom; and Supplemental Fig. S3). Overall, under normal culture growth conditions, the degree to which *T. cruzi* Sylvio X/10 edited sites in transcripts display canonical versus noncanonical numbers of U’s inserted or deleted is transcript dependent. For pan-edited transcripts where a single canonical editing product was identified, the percentage of editing that is canonical (averaged across the editing sites of the transcript) varies from 29% (*ND9*)

to 58% (ND8) (Table 2; average of Tot, Mt, and Fed library numbers). In summary, the *T. cruzi* edited mitochondrial transcriptome appears to possess a malleable editing process that is extraordinarily stochastic, even at “steady state” culture conditions.

Trypanosoma cruzi starvation impacts abundance of mitochondrial mRNAs

Finally, we analyzed whether the state of nutrient availability and stress can impact mRNA abundance and editing as might be expected based on the previous results (Shaw et al. 2016; Ramirez-Barrios et al. 2020). The mRNA reads were obtained in biological replicates from the mitochondrion-enriched cell fractions of epimastigote cultures at a density where glucose and other nutrients were plentiful (Kalem et al. 2018) (hereafter marked “Fed”); after 8 d of undisturbed culture growth, when the cells were in stationary phase (hereafter marked “Slow”); and 2 d after transfer of fed epimastigotes to a medium entirely lacking nutrient sources (triatomine artificial urine; TAU) (hereafter marked “Abrupt”). Starvation influences the abundance of total reads in the Sylvio X/10 strain. Total maxicircle reads for *CO1*, *CO2*, *CYb*, *ND1*, *ND8*, and *ND9* are increased, while those for *A6*, *CO3*, *ND4*, *ND9*, and *ND7* are decreased in at least one starvation condition, albeit sometimes modestly (Fig. 6A; TMM values used are reported in Supplemental Table S2). The abundance of the *CR3*, *CR4*, *MURF2*, *RPS3*, and *RPS12* transcripts remained virtually unaffected by starvation. As three already abundant mRNAs (*CO1*, *CO2*, and *ND8*) robustly increased in abundance upon starvation, it results in a reduced complexity of the mitochondrial transcriptome, if the editing state of reads is not considered. When editing is examined, there is a consistent increase of percentage of reads that are edited upon starvation in the Sylvio X/10 strain (from 37% edited reads in replete medium to ~47% in starved cells). This trend holds true for reads from loci individually in Sylvio X/10 (Table 1).

In contrast to the changes observed in Sylvio X/10 cells, mitochondrial gene expression remodeling for the strain

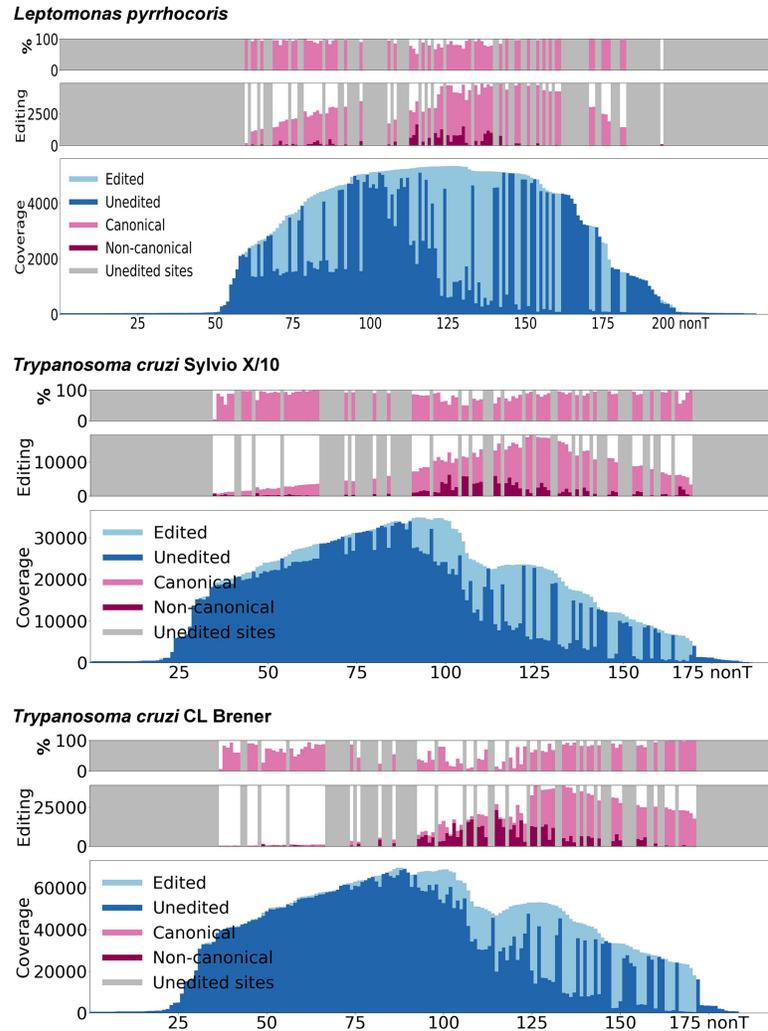


FIGURE 5. A lower percentage of *RPS12* sites in *Trypanosoma cruzi* are edited, and a higher percentage of noncanonical editing states is present when compared with *Leptomonas pyrrocoris*. Presented is editing of the *RPS12* transcript for *L. pyrrocoris*, *T. cruzi* Sylvio X/10, and *T. cruzi* CL Brener at every non “T” position from 5’ to 3’ along the X-axis. For each species/strain, the bottom plot depicts read coverage across the *RPS12* transcript, with the portion of reads edited at each location along the transcript shown in a lighter blue tone. The middle plot is a bar plot with an X-axis consistent with the bottom coverage plot. At each editing site, the bar distinguishes the total number of reads in which an editing event at that site is observed and breaks down the number by canonical and noncanonical editing events. A, C, and G positions along the transcript that are not locations of editing for canonical edited sequence are blocked out in gray and not analyzed. The top plot shows a similar bar graph, but the Y-axis represents percentage of editing that is canonical rather than absolute numbers of reads possessing editing at each site.

CL Brener was incremental upon starvation (Table 1; Fig. 6B; Supplemental Fig. S1). This is not entirely surprising, given the lesser extent that editing in general occurs in this strain in fed conditions. The strain CL Brener parasites’ potential to respond to the environmental cues (e.g., starvation) could be tempered by long-term axenic cultivation of the epimastigote stage in media rich in nutrients. In summary, whether relative abundance or editing of the CL Brener maxicircle transcript reads is measured, high-

TABLE 2. The percentage of edited reads that are edited to a canonical pattern at each edited site for a transcript, averaged

Transcript	Tot	Mt	Fed	Slow	Abrupt
A6	32	32	34	33	35
CO2	99	99	99	99	99
CO3	36	36	42	40	49
CYb	98	98	98	99	99
MURF2	100	89	88	92	92
ND7	56	55	58	58	65
ND8	58	58	58	63	69
ND9	29	29	28	28	30
RPS12	54	55	58	72	65

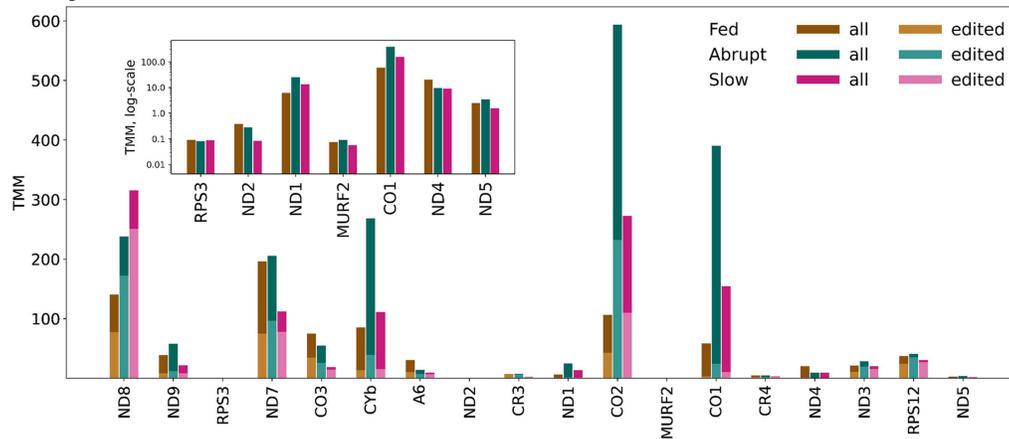
Analysis was performed separately for replicate merged Sylvio X/10 libraries. (Tot) Libraries prepared from total RNA. (Mt) Libraries prepared from RNA extracted from subcellular fractions enriched in mitochondrial vesicles. Other libraries are derived from total reads obtained in three feeding conditions: Fed, Slow starvation, and Abrupt starvation.

throughput sequencing reads largely exhibit insensitivity to starvation. In contrast, the same approach shows Sylvio X/10 transcripts to be quite responsive to starvation.

Starvation increases canonical and noncanonical editing events at discrete transcript regions

With ORFs reconstructed for Sylvio X/10 edited mRNAs, we were able to characterize the starvation-induced increase in editing in this strain observed in Table 1 and in (Shaw et al. 2016; Ramirez-Barrios et al. 2020). T-Aligner creates matrices that describe both the position and the number of U's inserted or deleted within the entire population of reads for the length of a given transcript. Both canonical and noncanonical editing (utilizing the same definition as for Figs. 4, 5) at each site are represented by dots positioned above and below the horizontal neutral position, respectively, on the T-Aligner visual output. In

A Sylvio X/10



B CL Brener

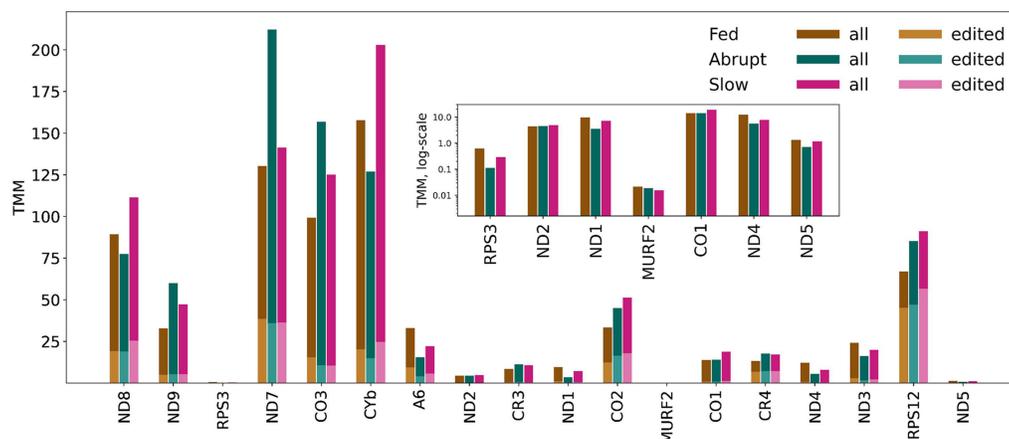


FIGURE 6. *Trypanosoma cruzi* strain Sylvio X/10 remodels its mitochondrial gene expression upon starvation. Relative mitochondrial mRNA abundances derived from total reads obtained in three feeding conditions: Fed, Abrupt starvation, and Slow starvation. Mitochondrial transcripts are ordered based on their order on the maxicircle, regardless of orientation. The inset shows the transcripts of lowest abundance, never-edited transcripts and MURF2, on a log-scale Y-axis. Abundances are shown for strains Sylvio X/10 (A) and CL Brener (B).

these matrices, each position is termed an editing event. The matrix reports the extent of representation within the read population of specific numbers of U inserted or deleted at every editing position (Gerasimov et al. 2018) by the size or color intensity of a dot at that matrix position (Figs. 7A, 8A). Any single editing site is typically covered by mapped reads of the following categories: reads containing the number of U's inserted or deleted consistent with the best supported assembled ORF, reads that contain no insertions/deletions (not edited), and reads that contain a noncanonical number of U's inserted or deleted. T-Aligner can represent any individual ORF reconstruction on a matrix with a line connecting the dots corresponding to editing states necessary at each site to result in that ORF (e.g., Figs. 7B,C, 8B,C).

Our approach to examining the effect of starvation on editing was to determine fold-change of fraction of reads within the read population that possessed each possible editing event upon starvation, for each of the cryptogenes. If the overall rate of editing was increased during starvation, we expected positive fold-change exhibited at editing events throughout the transcript. We anticipated

increased initiation of editing to be reflected by positive fold-change starting immediately at sites close to the position of the first editing event on the transcript. More efficient progression of editing across the transcript would appear as positive fold-change increasingly accumulating from 3' to 5' and possibly at key positions following common sites where editing stalls. Strong positive fold-change focused at positions closer to the 5' terminus of the edited domain for each transcript would be expected if only the fully edited form was stabilized. Finally, an increased ratio of canonical editing events relative to noncanonical events, distinguishable on the editing matrices, could also result in more product. Fold-change for each editing event was plotted on a differential editing comparison matrix (Figs. 7, 8; Supplemental Fig. S4).

Several trends were evident in our differential editing analysis, which is represented in Figures 7 and 8 for the Sylvio X/10 CO3 and A6 transcripts, respectively. Positive fold-changes following either abrupt or slow starvation are indicated with red dots, with more intense coloring indicating greater fold-change (Fig. 7B,C; Supplemental Fig. S4). The increases in percentage of reads from each locus

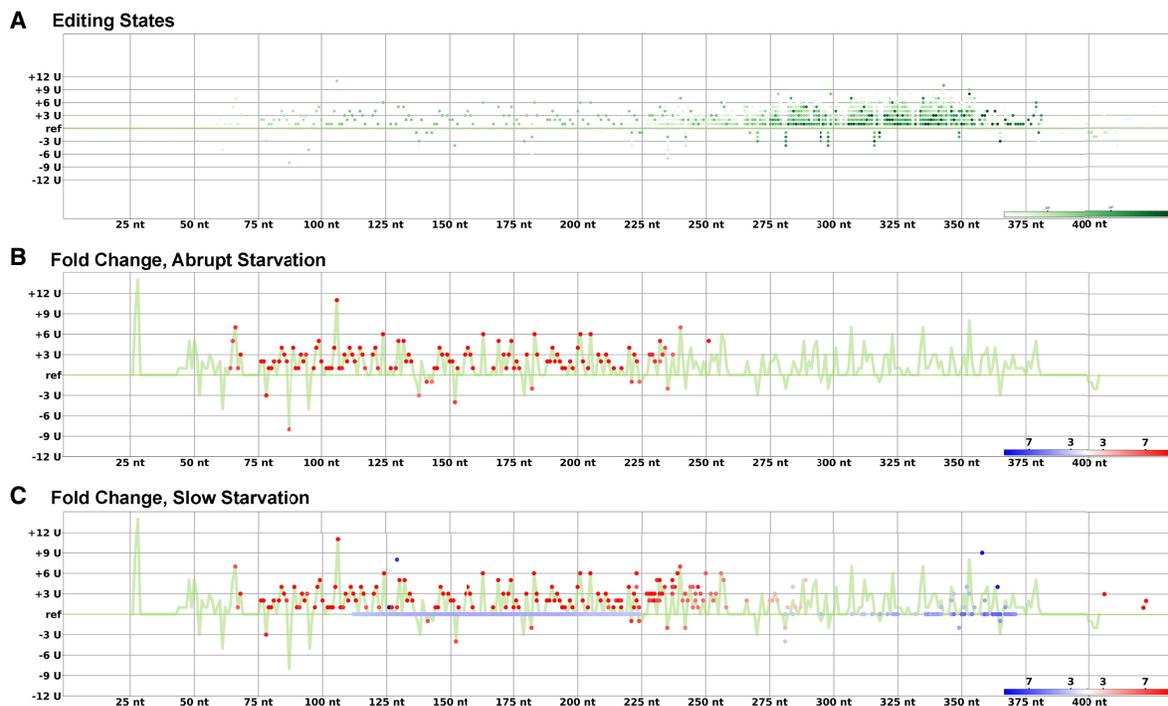


FIGURE 7. Differences in CO3 editing upon starvation of cultured *Trypanosoma cruzi* strain Sylvio X/10 epimastigotes. Each scheme represents the numbers of inserted/deleted U's as dots, located above (insertion) or below (deletion) the reference (ref) line, with the distance from neutral distinguishing the number of insertions or deletions. Each dot defines an editing state. The X-axis (in green) represents an A, G, or C nucleotide position in the maxicircle sequence for the CO3 cryptogene from 5' to 3' where editing can potentially occur. (A) Visualization of CO3 editing states found in at least two reads represented by dots on the matrix. The relative abundance of reads supporting an editing state is calculated for each site separately and is presented as the intensity of the green dot at that position. A more intense green color reflects a higher fraction of reads displaying that editing state. (B,C) Fold increase (red) or decrease (blue) of each specific editing event represented by positions on the CO3 transcript matrix for parasites abruptly starved of nutrients (B) or parasites allowed to gradually deplete medium nutrients over 8 d (C). The jagged green line shows the "path" of the canonical edited mRNA for the cryptogene, the number of U insertions or deletions at each editing site that will result in the canonical ORF.

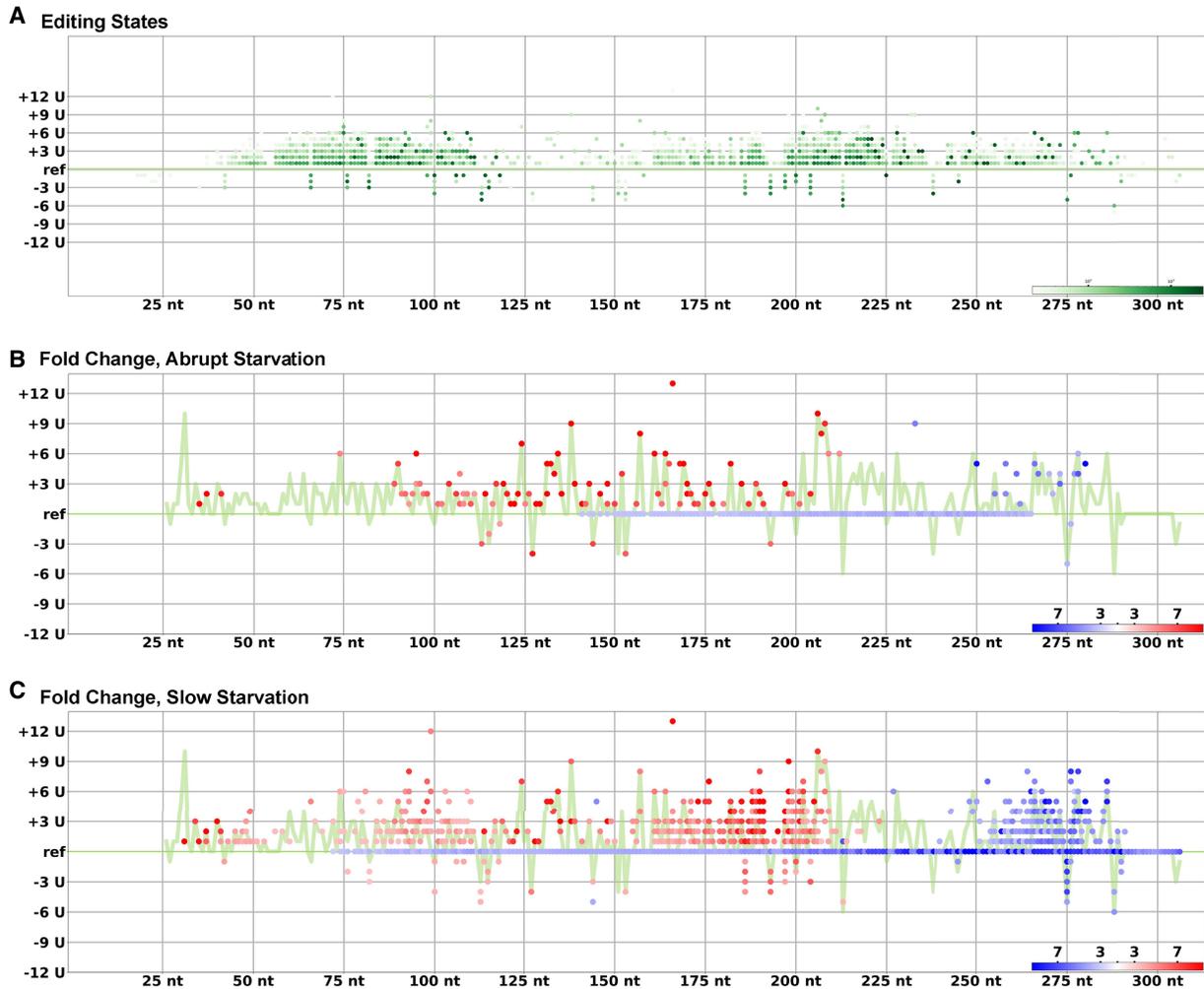


FIGURE 8. Differences in *A6* editing upon starvation of cultured *Trypanosoma cruzi* strain Sylvio X/10 epimastigotes. Each scheme represents the numbers of inserted/deleted U's as dots, located above (insertion) or below (deletion) the reference (ref) line, with the distance from neutral distinguishing the number of insertions or deletions. Each dot defines an editing state. The X-axis (in green) represents an A, G, or C nucleotide position in the maxicircle sequence for the *A6* cryptogene from 5' to 3' where editing can potentially occur. (A) Visualization of *A6* editing states found in at least two reads represented by dots on the matrix. The relative abundance of reads supporting an editing state is calculated for each site separately and is presented as the intensity of the green dot at that position. A more intense green color reflects a higher fraction of reads displaying that editing state. (B,C) Fold increase (red) or decrease (blue) of each specific editing event represented by positions on the *A6* transcript matrix for parasites abruptly starved of nutrients (B), or parasites allowed to gradually deplete medium nutrients over 8 d (C). The jagged green line shows the "path" of the canonical edited mRNA for the cryptogene, the number of U insertions or deletions at each editing site that will result in the canonical ORF.

that are edited (Table 1) manifest as almost entirely positive fold-change in editing events. The only exception is *ND9* under abrupt starvation, in which global fold-change in editing was not apparent. Starvation rarely results in a loss of editing, canonical or noncanonical, insertion or deletion, at any position, with two exceptions (see below). Interestingly, for *CO3*, increased fold-change in editing during starvation was observed at a position in the transcript where the processivity of editing has tailed off (as detected by the lower intensity of upstream dots in Fig. 7A).

Figures 7, 8 and Supplemental Figure S4 show that upon starvation there were lesser fold-changes in editing events

in the extreme 5' terminus (the last portion of transcript to be edited) and 3' terminus (the first portion to be edited) of editing domains. Typically, starvation-induced editing event increases were concentrated in the middle of an editing domain. Often, there were one or more narrower central regions that could be considered "hotspots" of starvation-induced editing. Where negative fold-change appeared at positions within the matrix, it was often along the axis that represents an absence of U insertion or deletion. This was expected to correlate with an increase in editing events at those sites. The other situation where negative fold changes in editing appeared was at the site

of editing initiation in some transcripts—particularly A6, where the positive fold-change in editing positions occurs further into the edited domain (Fig. 8). This pattern is consistent with starvation-induced change manifested as an increase in efficiency of editing progression, rather than increases in initiation of editing on nascent transcripts. Further, a lack of robust positive fold-change in the extreme 5' of the edited domain of all transcripts suggests that fully edited transcripts are not preferentially stabilized to bring about their increased abundance upon starvation.

Critically, the type of starvation experienced by the parasite appears to make a difference in the resulting changes in editing. We find that increases in editing events upon “Abrupt” starvation are largely canonical (visualized as the red dots through which pass the line representing the editing events of the canonical ORF in Fig. 7A,B). This is consistent across all analyzed pan-edited loci except for *RPS12* and *ND9*. The same positive fold-changes in canonical editing events were also observed under the conditions of “Slow” starvation. Additionally, however, under “Slow” starvation we observed positive fold-changes for many noncanonical events as well (Figs. 7, 8; Supplemental Fig. S4; this phenotype is particularly strong for A6). This is especially true for hotspots.

Increased editing for minimally edited *CO2* and *CYb* sites was not detectable in these matrices. This is because in fed parasites, most of the reads are already edited at those sites across all strains, as previously shown (Supplemental Fig. S3). Thus, fold-increase in editing of the scale that would be visible on the matrix is not possible. For example, of *CO2* reads spanning the edited domain, over 99% of them exhibited editing in the CL Brener and Sylvio X/10 “Fed” libraries. We also note that the portions of *CO2* and *CYb* which are not normally edited exhibit islands of noncanonical editing in parasites from fed conditions. There is a reduction of editing at these positions upon starvation, even in CL Brener, which is less sensitive to starvation (Supplemental Fig. S4). Taken together, T-Aligner is best at revealing starvation-induced changes in editing for pan-edited mRNAs in the strain Sylvio X/10 out of the two strains that were analyzed. Increases in editing upon starvation are seen largely throughout the transcript, but canonical editing increases are favored when the starvation is abrupt in nature.

DISCUSSION

Strain-specific and starvation-specific maxicircle expression and editing

In this study, we identified differences in relative abundances of mitochondrial mRNAs and RNA editing among *T. cruzi* strains. The long-term goal of this research direction is to understand how such mRNA editing differences and perturbations affect the functional output of the mitochondrial genome. Most maxicircle genes and crypto-

genes encode ETC components. In particular, we want to know whether ETC function can be impacted by changes in mitochondrial gene expression or RNA editing. The strains utilized in this work were selected because their maxicircle coding regions had already been sequenced and compared (Ruvalcaba-Trejo and Sturm 2011). Further, we had previously seen that multiple aspects of metabolic function varied widely between these strains, as did the abundances of mRNAs encoding nuclear-encoded components of the ETC components that are the partners of the mitochondrial maxicircle protein products (Kalem et al. 2018). Our current inability to ascertain which mitochondrial gene products are translated is a barrier to our efforts to tie these differences to a metabolic or other mitochondrial phenotype. Here we have taken the first step by showing that mitochondrial gene expression and RNA editing differences clearly exist between strains (Fig. 3). It is possible that a larger study including many more strains would reveal correlations between mitochondrial gene expression/RNA editing and mitochondrial metabolism despite the dearth of data at the protein level.

Starvation-induced differences in *T. cruzi* mitochondrial gene expression and RNA editing are potentially linked to the parasite’s life stage transitions and, thus, may be critical to its survival (Ramirez-Barrios et al. 2020). In this work we illuminated mechanisms of starvation-induced changes to *T. cruzi* epimastigotes (Figs. 7, 8) and on the level of the entire transcriptome, rather than for a few select mRNAs (Table 1; Fig. 6). By and large, what we observed agreed with what we previously determined by RT-qPCR in previous studies (Shaw et al. 2016; Ramirez-Barrios et al. 2020). For instance, overall levels of transcripts were largely unresponsive to starvation with the exception of a few transcripts, such as *CO1* and *CO2*. The mitochondrial transcriptome of Sylvio X/10 was more responsive to starvation in terms of the magnitude of the changes it elicited. The pattern of starvation-induced increases in edited products accompanied by lesser or no increases in pre-edited versions of those same transcripts was also mostly consistent.

Finally, we were able to determine that for the Sylvio X/10 strain, the increases in fully and near fully edited products detected by RT-qPCR were caused by an overall increase in the amount of RNA editing occurring in the mitochondrion, rather than an increase in the number of transcripts on which editing was initiated, or by a preferential stabilization of transcripts once they were fully edited (Figs. 7, 8). Of course, what looks to be an increase in the overall enzymatic activity of editing may instead be some form of RNA stabilization that indirectly impacts specifically RNAs in the editing pathway. For instance, overall stabilization of gRNAs would increase their abundance and thus their availability to the editing machinery. This could result in more efficient editing in general. An increase in minor, alternate gRNAs due to their selective

stabilization could result in the observed effects specific to slowly starved parasites: namely, an increase in noncanonical editing events. It is also possible that under starvation, and only in this condition, any editing activity on an mRNA triggers increased stability leading to the observed result. However, this explanation is more convoluted and therefore we prefer models that invoke increased enzymatic activity or gRNA availability during starvation.

One point of incomplete agreement between these and our previous results are some CL Brener editing patterns. Notably, our current results indicate that editing appears to be restricted in this strain. It is possible that most mRNAs of the mitochondrial genome, except for *CYb*, *CO2*, *CO3*, and *RPS12*, are no longer edited completely enough to generate functional ORFs. We could not assemble them, owing to both editing progression drop-off and low abundance of edited reads in the entire mitochondrial transcriptome. This failure includes an inability to generate edited *A6*. As the product of edited *A6* is essential to ATP generation via oxidative phosphorylation, its absence would most likely yield a nonrespiring strain. However, we have firm evidence that these CL Brener parasites respire, even when continuously grown in culture (Kalem et al. 2018). Therefore, it is most likely that levels of translatable mitochondrial mRNAs, expressed below the limits of detection in our read populations, are in fact sufficient to generate protein products. The lack of proportional increase of edited over unedited reads in the starved CL Brener parasites is surprising because increases in the abundances of edited mRNAs upon starvation were first detected in this strain (Shaw et al. 2016) (although they were later found to be lesser in amplitude than those of Sylvio X/10, Ramirez-Barrios et al. 2020). The CL Brener differential editing matrices in Supplemental Figure S4 may provide the clue. Unlike in the Sylvio X/10 strain, where canonical patterns increase upon starvation, in CL Brener, noncanonical editing appears to decrease. It is possible that with less editing diverting from the canonical track during starvation, more fully edited canonical product is produced. Thus, this unusual strain with very low levels of editing may have evolved an alternative mechanism to be able to respond to starvation.

Continued development of our concepts of trypanosomatid mRNA editing

The quality of evidence that underlies U-indel edited mRNA sequences in public repositories can be cryptic (Zimmer et al. 2018). Some publicly available sequences are in silico reconstructions based on a kinetoplastid maxicircle sequence and its alignment with an edited mRNA sequence that has been determined experimentally, usually that of *T. brucei* or *L. tarentolae*. Those determined experimentally are nearly all based on Sanger sequencing of cloned PCR amplified cDNAs, some with only a single read

to cover the most extreme 5' end of the molecule. Considering the tools available for the discovery and initial analysis of U-indel edited mRNAs, it is not surprising that confidence in the existence of multiple ORFs generated from transcripts of a single cryptogene has emerged only with the advent of deep sequencing technologies. However, our continued reliance on inter-species comparisons to identify single predominant edited products in novel kinetoplastids, utilized to a degree even in this study, does present a potential experimental limitation.

Trypanosoma cruzi *CYb* has a short 5' terminal edited domain. It is a revealing example of how editing studies may be impacted by assumptions of a single edited translatable product, the identity of which may be based on limited cloning-and-sequencing-derived early data. An in silico prediction of the *T. cruzi* ORF (Ruvalcaba-Trejo and Sturm 2011) was first established based on the experimentally determined *T. brucei* *CYb* ORF. Subsequently, we cloned and Sanger-sequenced PCR products to discover that the editing pattern for *T. cruzi* was slightly different than that predicted (Shaw et al. 2016). We used our updated sequence to establish PCR primers for edited *CYb*. However, our deeper analysis here reveals that the most abundant ORF-generating editing pattern differs from either of these previously determined sequences. The small sequencing discrepancy may have led to RT-qPCR artifacts that overestimated the degree to which the abundance of fully edited *CYb* increased during starvation.

Our recent identification of *L. pyrrhocoris* mitochondrial ORFs yielded a single canonical product for all edited loci (Gerasimov et al. 2018). This differs from our *T. cruzi* T-Aligner output that suggests multiple products for *CR3*, *CR4*, and *ND3*. Previously, the Koslowsky laboratory presented evidence of sequence permissiveness for dual coding reading frames for *T. brucei* edited mRNAs accessible with alternative 5' editing (Kirby and Koslowsky 2017). Subsequently, they reported extensive alternative sequences for the transcript *CR3* (Kirby and Koslowsky 2020), for which they also utilized a linked node array to describe the ultimate translatable editing output. Their study utilized PCR to obtain deep sequencing libraries for *CR3* and other transcripts. We performed our *T. cruzi* analysis on Illumina reads mapped to T-less maxicircles in a total transcriptome approach. Thus, here we have validated the prior evidence of dual-coding transcripts with both another *Trypanosoma* spp. and an alternative, PCR-less method.

Despite limitations in accurately describing the population of translated mRNAs, similarities and differences of U-indel edited transcriptomes may be coming into focus. Our current findings appear to build on gRNA-related differences in editing output between kinetoplastid subfamilies Leishmaniinae (in which reside the *Leishmania* and *Leptomonas* spp.) and Trypanosomatinae (in which reside *T. cruzi* and *T. brucei*) (Kostygov et al. 2021). We recently characterized the minicircles of *L. pyrrhocoris* and defined

the relationship of their gRNA products to the maxicircle transcriptome (Gerasimov et al. 2021). We discovered that *L. pyrrhocoris* possesses only 67 minicircle classes. The number of minicircle classes is a rough measure of the complexity capacity of the organism's gRNA repertoire. The related *Leishmania tarentolae* has more minicircle classes (114, Camacho et al. 2019) but even this is dwarfed by the number of minicircle classes for *T. cruzi* and *T. brucei* in the range of 300–400 (Cooper et al. 2019; Callejas-Hernández et al. 2021). As the Trypanosomatinae species also encode more gRNAs on each minicircle, the gRNA population complexity of this subfamily appears appreciably higher than that of the Leishmaniinae.

Compared to other studied species, the anchor regions of *L. pyrrhocoris* gRNAs have fairly relaxed base-pairing “rules” to bind to maxicircle transcripts and serve as editing templates. It appears that this binding flexibility of their gRNAs is the primary source of the observed noncanonical editing events in this species, as this flexibility also permits binding to noncanonical locations within other transcripts (Gerasimov et al. 2021). In contrast, *T. brucei* gRNAs appear to require higher levels of anchor-region sequence similarity with their maxicircle targets. This would restrict their ability to bind to more than one maxicircle-derived transcript. Therefore, we may postulate that the more complex gRNA populations of *Trypanosoma* spp. are responsible for the expansion of editing patterns in certain cryptogene loci that result in multiple ORFs, as well as an increase in percentage of noncanonical editing of cryptogene products that do appear to have single canonical products (Figs. 1, 2, 4, 5; Supplemental Fig. S2). Ultimately, gRNA complexity may reflect an inherent ability of kinetoplastids to modify editing for their evolutionary advantage. Fewer gRNAs means more “looseness” in the gRNA:mRNA binding during editing, but ultimately this results in fewer sites edited to noncanonical patterns.

Future directions

We have developed a considerable suite of tools to characterize U-indel transcriptomes from Illumina sequencing reads. However, it is clear that tool expansion will continue, as ever-more-detailed analyses of editing are desired. For instance, a caveat of our Figure 1 coding capacity analysis is that median read support values can mask large variation in read support values for individual editing events across the entirety of a reconstructed ORF. More rigorous support analytics can and will eventually be developed. Another subject to tackle is the current assumption of the identity of alternative start codons in kinetoplastid mitochondrial transcriptomes, originating primarily from our inability to identify AUG start codons in early sequencing efforts. Inclusion or rejection of specific non-AUG start codons greatly influences T-Aligner's output of full-length ORFs. A third unanswered question to pursue quantitatively is

whether editing events that do not appear to contribute to any possible ORF are actual editing errors. Do they instead represent editing intermediates on transcripts caught in the process of editing at the noncanonically edited sites? They might also represent “dead end” products generated by the alignment of a noncognate gRNA binding imperfectly to a location and incorrectly directing editing (Gerasimov et al. 2021).

In summary, we present here the first differential editing analysis performed on next-generation sequencing data with PCR-independent methodologies. This allows for comparisons of editing levels with a high level of statistical significance ensured by usage of trusted differential expression tools. Also, a transcriptome-wide approach in differential editing analysis does not force the experimenter to focus on a subset of U-indel edited transcripts. This is particularly important when differential editing is confined to a subset of transcripts that cannot be predicted in advance. We anticipate an acceleration of the discovery of editing differences in specific transcripts across multiple transcriptomes, in response to altered growth conditions, or between parasite life stages. Further, exploration of the impact of genetic mutation, silencing or knockout of known or suspected components of mitochondrial editing or mRNA processing and stability pathways can now be analyzed at the level of the entire mitochondrial transcriptome. Interesting changes that are identified can subsequently be followed up with closer examination by more targeted methods, such as the Trypanosome RNA Editing Alignment Tool (TREAT) (Simpson et al. 2017).

MATERIALS AND METHODS

Strains and parasite growth

Trypanosoma cruzi strains CL Brener (original nonclonal strain utilized in *T. cruzi* genome project, acquired from the laboratory of Roberto Docampo, University of Georgia), Sylvio X10 (ATCC 50823) and Esmeraldo (ATCC 50794) were cultured as epimastigotes in Liver infusion tryptose medium (Castellani et al. 1967) supplemented with 10% fetal bovine serum and 20 µg/mL hemin (Sigma-Aldrich) at 27°C. The continually propagated source cells for all experiments were diluted 1:10 every 2–3 d to maintain the growth in exponential phase (Shaw et al. 2016). A starting number of 5×10^8 epimastigotes were collected and processed for each sample. For “Fed” or normal growth samples, CL Brener or Sylvio X/10 parasites were collected 2 d after they were diluted to a concentration of 5×10^6 cells/mL (typically 1×10^7 cells/mL at time of collection). Slow starvation studies were initiated with 2×10^7 cells/mL of exponentially growing cells and allowed to grow undisturbed for 8 d. Abrupt starvation conditions were achieved with a starting concentration of 1.5×10^7 cells/mL and incubation for 2 d in triatomine artificial urine (TAU) medium (Contreras et al. 1985). Strains were validated by comparing obtained maxicircle sequences from the growing parasites with previously published sequences for these strains (Westenberger et al.

2006; Ruvalcaba-Trejo and Sturm 2011). Maxicircle sequence GenBank accession numbers are FJ203996 (Sylvio X/10), DQ343645 (CL Brener) and DQ343646 (Esmeraldo). No variations were detected for CL Brener or Esmeraldo coding regions, and a total of 41 single nucleotide variants were detected and corrected in the Sylvio X/10 maxicircle. The corrected sequence of the coding portion of the Sylvio X/10 maxicircle is incorporated in the mRNA GenBank submissions (Supplemental Table S1).

DNA and RNA purification, library generation, and sequencing

DNA was purified using the Qiagen DNeasy Blood and Tissue kit (Qiagen). Crude mitochondrial preparations, RNA purification, library preparation (including mRNA purification on an oligo(dT) column) and sequencing at the University of Minnesota Genomics Center were performed as described in Gerasimov et al. (2018) on biological replicate duplicate samples. Raw sequencing output is available in the GenBank Sequence Read Archive as BioProjects PRJNA792819: *Trypanosoma cruzi* starvation-driven mitochondrial RNA editing and transcriptome variability, and PRJNA395140: Study of RNA editing process in *Trypanosoma cruzi* strains.

Read alignment and relative mitochondrial mRNA abundance analysis

For each sample, T-less reads were mapped to the T-less maxicircle reference loci using the alignlib program in the T-Aligner toolkit with default settings (Gerasimov et al. 2021). The vast majority of edited reads mapped to transcripts that are known to be edited in *Trypanosoma* spp. Our method also resulted in non-zero edited read counts for many never-edited transcripts (Table 1; Supplemental Table S1). Most of them are mapped to flanking regions of the never-edited genes that must be included in maxicircle reference loci sequences but are actually the termini of neighboring edited genes. Some of the remaining examples may be mapping artifacts and sequencing errors, but there may also be a very low level of noncanonical editing of never-edited mRNAs. The rate that such reads appear in the *T. cruzi* mitochondrial transcriptome is comparable with that of *L. pyrrhocoris* that we have previously analyzed (Gerasimov et al. 2018).

To determine the appropriateness of merging reads for all experiments, a correlation plot was generated from all samples once they were normalized with EdgeR (Supplemental Fig. S1; Robinson et al. 2010). For transcript abundance analysis among strains, reads from biological replicate libraries from RNA from both total and mitochondria-enriched lysates (four in total) per strain were merged. For transcript abundance analysis among starvation and fed conditions, the biological replicates were merged (two in total per condition/strain). For both abundance analyses, TMM values were used.

Initial assembly of edited mRNAs

For each strain all Illumina reads (from starved cells as well as those grown in normal conditions) were merged and used to reconstruct ORFs with the "findorf" program from the T-Aligner

toolkit (Gerasimov et al. 2021). For Sylvio X/10, the corrected cryptogene reference sequences were used. For most pan-edited genes, the ORF finding mode was set to prefer edited reads or best ORF extension ("editing" or "extension" options). ORFs were only considered valid if they covered at least 80% of the length of the cryptogene. For minimally edited genes, the mode was "coverage_unedited" or "overlap." Of the ORFs generated, the canonical ORF was determined by manually inspecting translated ORF best blastp hits with respective GenBank mRNAs for *Trypanosoma vivax*, *T. brucei*, *T. lewisi*, or *L. pyrrhocoris*. The ORF with the highest amino acid sequence similarity was deemed "canonical." For cryptogenes for which one or fewer GenBank submissions for comparison were available, the "canonical" ORF was defined as the isoform with the highest read support at each editing site. In most cases, the isoform with the best read support was also that with the highest blastp hit. For all ORF reconstructions, "orf_search_depth" was set to 2, and the default T-Aligner genetic code table was used in which initial attempts to trace ORFs allow initiation with AUG, GUG, AUA, UUG, or CUG, preferring AUG over others.

Analysis of differential editing during starvation

Biological replicate libraries from fed, abrupt starvation, and slow starvation conditions were used separately. Counts of editing events at each T-Aligner plot position (corresponding to the location, nature [insertion/deletion] and number of U's inserted or deleted) were inputs for EdgeR's differential expression algorithm. Fold-change incidence of each editing event upon abrupt or slow starvation was plotted using a custom script. The *P*-value cutoff was set to <0.001 FDR-adjusted *P*-value from EdgeR's output table.

Coverage plots

Read coverage across transcripts was analyzed for each replicate separately using raw reads mapped with T-Aligner omitting EdgeR normalization and displayed (as the coverage at each A/G/C nucleotide position of the pre-edited transcript that was contributed by reads with editing at that position, and the coverage by reads not edited at that position) using a custom script. Each edited position of the transcript was further presented in two additional plots. In one, the raw number of canonical and noncanonical events at each editing position were displayed. In the second, the percentage of editing events from the reads that possess the canonical editing event at each editing site was shown. For Figure 5, CO3 from replicate 1 from the strain comparison libraries derived from RNA from mitochondrial-enriched lysate is presented. For Figure 6, RPS12 analysis of replicate 1 of fed *T. cruzi* from the starvation experiments is presented. For *L. pyrrhocoris* edited RNAs, datasets SRR5796658 and SRR5796659 from our previous study (Gerasimov et al. 2018) were used to derive the RPS12 results for this species using the identical methodology.

Coding capacity graph

A custom python script was used to plot coding capacity graphs from T-Aligner output. For each assembled protein product, the

median edited read support level was determined by mapping raw mapped reads on the assembled mRNA sequence. Assembled protein sequences were split into *k*-mers of fixed size (6) and *k*-mers that could be extracted from canonical protein sequence were highlighted blue. The final direct graph (Fig. 1) was built from *k*-mers using the support value for edge (arrow) colors.

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

This work was supported by American Heart Association Scientist Development grant 16SDG26420019 to S.L.Z., the Grant Agency of the Czech Republic (22-01026S) and the European Regional Funds CZ.02.1.01/16_019/0000759 to V.Y., and a Russian Science Foundation grant (RSF 19-74-10008) to E.S.G.

Received December 28, 2021; accepted April 7, 2022.

REFERENCES

- Aphasizheva I, Alfonso J, Carnes J, Cestari I, Cruz-Reyes J, Goring HU, Hajduk S, Lukeš J, Madison-Antenucci S, Maslov DA, et al. 2020. Lexis and grammar of mitochondrial RNA processing in trypanosomes. *Trends Parasitol* **36**: 337–355. doi:10.1016/j.pt.2020.01.006
- Benne R, Van den Burg J, Brakenhoff JP, Sloof P, Van Boom JH, Tromp MC. 1986. Major transcript of the frameshifted *coxII* gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. *Cell* **46**: 819–826. doi:10.1016/0092-8674(86)90063-2
- Brenière SF, Waleckx E, Brenière C. 2016. Over six thousand *Trypanosoma cruzi* strains classified into discrete typing units (DTUs): attempt at an inventory. *PLoS Negl Trop Dis* **10**: e0004792. doi:10.1371/journal.pntd.0004792
- Callejas-Hernández F, Herreros-Cabello A, Del Moral-Salmoral J, Fresno M, Gironès N. 2021. The complete mitochondrial DNA of *Trypanosoma cruzi*: maxicircles and minicircles. *Front Cell Infect Microbiol* **11**: 672448. doi:10.3389/fcimb.2021.672448
- Camacho E, Rastrojo A, Sanchiz A, Gonzalez-de la Fuente S, Aguado B, Requena JM. 2019. *Leishmania* mitochondrial genomes: maxicircle structure and heterogeneity of minicircles. *Genes (Basel)* **10**: 758. doi:10.3390/genes10100758
- Castellani O, Ribeiro LV, Fernandes JF. 1967. Differentiation of *Trypanosoma cruzi* in culture. *J Protozool* **14**: 447–451. doi:10.1111/j.1550-7408.1967.tb02024.x
- Contreras VT, Morel CM, Goldenberg S. 1985. Stage specific gene expression precedes morphological changes during *Trypanosoma cruzi* metacyclogenesis. *Mol Biochem Parasitol* **14**: 83–96. doi:10.1016/0166-6851(85)90108-2
- Cooper S, Wadsworth ES, Ochsenreiter T, Ivans A, Savill NJ, Schnauffer A. 2019. Assembly and annotation of the mitochondrial minicircle genome of a differentiation-competent strain of *Trypanosoma brucei*. *Nucleic Acids Res* **47**: 11304–11325. doi:10.1093/nar/gkz928
- Cruz-Reyes J, Mooers BHM, Doharey PK, Meehan J, Gulati S. 2018. Dynamic RNA holo-editosomes with subcomplex variants: insights into the control of trypanosome editing. *Wiley Interdiscip Rev RNA* **9**: e1502. doi:10.1002/wrna.1502
- Gerasimov ES, Gasparyan AA, Afonin DA, Zimmer SL, Kraeva N, Lukeš J, Yurchenko V, Kolesnikov A. 2021. Complete minicircle genome of *Leptomonas pyrrocoris* reveals sources of its non-canonical mitochondrial RNA editing events. *Nucleic Acids Res* **49**: 3354–3370. doi:10.1093/nar/gkab114
- Gerasimov ES, Gasparyan AA, Kaurav I, Tichý B, Logacheva MD, Kolesnikov AA, Lukeš J, Yurchenko V, Zimmer SL, Flegontov P. 2018. Trypanosomatid mitochondrial RNA editing: dramatically complex transcript repertoires revealed with a dedicated mapping tool. *Nucleic Acids Res* **46**: 765–781. doi:10.1093/nar/gkx1202
- Golden DE, Hajduk SL. 2005. The 3'-untranslated region of cytochrome oxidase II mRNA functions in RNA editing of African trypanosomes exclusively as a *cis* guide RNA. *RNA* **11**: 29–37. doi:10.1261/ma.7170705
- Golden DE, Hajduk SL. 2006. The importance of RNA structure in RNA editing and a potential proofreading mechanism for correct guide RNA:pre-mRNA binary complex formation. *J Mol Biol* **359**: 585–596. doi:10.1016/j.jmb.2006.03.041
- Goldenberg S, Avila AR. 2011. Aspects of *Trypanosoma cruzi* stage differentiation. *Adv Parasitol* **75**: 285–305. doi:10.1016/B978-0-12-385863-4.00013-7
- Greif G, Rodriguez M, Reyna-Bello A, Robello C, Alvarez-Valin F. 2015. Kinetoplast adaptations in American strains from *Trypanosoma vivax*. *Mutat Res* **773**: 69–82. doi:10.1016/j.mrfmmm.2015.01.008
- Kalem MC, Gerasimov ES, Vu PK, Zimmer SL. 2018. Gene expression to mitochondrial metabolism: variability among cultured *Trypanosoma cruzi* strains. *PLoS ONE* **13**: e0197983. doi:10.1371/journal.pone.0197983
- Kim KS, Teixeira SM, Kirchhoff LV, Donelson JE. 1994. Transcription and editing of cytochrome oxidase II RNAs in *Trypanosoma cruzi*. *J Biol Chem* **269**: 1206–1211. doi:10.1016/S0021-9258(17)42243-5
- Kirby LE, Koslowsky D. 2017. Mitochondrial dual-coding genes in *Trypanosoma brucei*. *PLoS Negl Trop Dis* **11**: e0005989. doi:10.1371/journal.pntd.0005989
- Kirby LE, Koslowsky D. 2020. Cell-line specific RNA editing patterns in *Trypanosoma brucei* suggest a unique mechanism to generate protein variation in a system intolerant to genetic mutations. *Nucleic Acids Res* **48**: 1479–1493. doi:10.1093/nar/gkz1131
- Kostygov AY, Karnkowska A, Votýpka J, Tashyreva D, Maciszewski K, Yurchenko V, Lukeš J. 2021. Euglenozoa: taxonomy, diversity and ecology, symbioses and viruses. *Open Biol* **11**: 200407. doi:10.1098/rsob.200407
- Lukeš J, Butenko A, Hashimi H, Maslov DA, Votýpka J, Yurchenko V. 2018. Trypanosomatids are much more than just trypanosomes: clues from the expanded family tree. *Trends Parasitol* **34**: 466–480. doi:10.1016/j.pt.2018.03.002
- Lukeš J, Kaur B, Speijer D. 2021. RNA editing in mitochondria and plastids: weird and widespread. *Trends Genet* **37**: 99–102. doi:10.1016/j.tig.2020.10.004
- Majeau A, Murphy L, Herrera C, Dumonteil E. 2021. Assessing *Trypanosoma cruzi* parasite diversity through comparative genomics: implications for disease epidemiology and diagnostics. *Pathogens* **10**: 212. doi:10.3390/pathogens10020212
- Maslov DA, Opperdoes FR, Kostygov AY, Hashimi H, Lukeš J, Yurchenko V. 2019. Recent advances in trypanosomatid research: genome organization, expression, metabolism, taxonomy and evolution. *Parasitology* **146**: 1–27. doi:10.1017/S0031182018000951
- Ramirez-Barrios R, Susa EK, Smoniewski CM, Faacks SP, Liggett CK, Zimmer SL. 2020. A link between mitochondrial gene expression and life stage morphologies in *Trypanosoma cruzi*. *Mol Microbiol* **113**: 1003–1021. doi:10.1111/mmi.14466
- Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**: 139–140. doi:10.1093/bioinformatics/btp616

- Rusman F, Florida-Yapur N, Tomasini N, Diosque P. 2021. Guide RNA repertoires in the main lineages of *Trypanosoma cruzi*: high diversity and variable redundancy among strains. *Front Cell Infect Microbiol* **11**: 663416. doi:10.3389/fcimb.2021.663416
- Ruvalcaba-Trejo LI, Sturm NR. 2011. The *Trypanosoma cruzi* Sylvio X10 strain maxicircle sequence: the third musketeer. *BMC Genomics* **12**: 58. doi:10.1186/1471-2164-12-58
- Shaw AK, Kalem MC, Zimmer SL. 2016. Mitochondrial gene expression is responsive to starvation stress and developmental transition in *Trypanosoma cruzi*. *mSphere* **1**: e00051-00016. doi:10.1128/mSphere.00051-16
- Simpson RM, Bruno AE, Chen R, Lott K, Tylec BL, Bard JE, Sun Y, Buck MJ, Read LK. 2017. Trypanosome RNA editing mediator complex proteins have distinct functions in gRNA utilization. *Nucleic Acids Res* **45**: 7965–7983. doi:10.1093/nar/gkx458
- Škodová-Sveráková I, Verner Z, Skalický T, Votýpka J, Horváth A, Lukeš J. 2015. Lineage-specific activities of a multipotent mitochondrion of trypanosomatid flagellates. *Mol Microbiol* **96**: 55–67. doi:10.1111/mmi.12920
- Sturm NR, Simpson L. 1990. Partially edited mRNAs for cytochrome b and subunit III of cytochrome oxidase from *Leishmania tarentolae* mitochondria: RNA editing intermediates. *Cell* **61**: 871–878. doi:10.1016/0092-8674(90)90197-M
- Tylec BL, Simpson RM, Kirby LE, Chen R, Sun Y, Koslowsky DJ, Read LK. 2019. Intrinsic and regulated properties of minimally edited trypanosome mRNAs. *Nucleic Acids Res* **47**: 3640–3657. doi:10.1093/nar/gkz012
- Tyler KM, Engman DM. 2001. The life cycle of *Trypanosoma cruzi* revisited. *Int J Parasitol* **31**: 472–481. doi:10.1016/S0020-7519(01)00153-9
- Vercellino I, Sazanov LA. 2021. The assembly, regulation and function of the mitochondrial respiratory chain. *Nat Rev Mol Cell Biol* **23**: 141–161. doi:10.1038/s41580-021-00415-0
- Westenberger SJ, Cerqueira GC, El-Sayed NM, Zingales B, Campbell DA, Sturm NR. 2006. *Trypanosoma cruzi* mitochondrial maxicircles display species- and strain-specific variation and a conserved element in the non-coding region. *BMC Genomics* **7**: 60. doi:10.1186/1471-2164-7-60
- Zimmer SL, Simpson RM, Read LK. 2018. High throughput sequencing revolution reveals conserved fundamentals of U-indel editing. *Wiley Interdiscip Rev RNA* **9**: e1487. doi:10.1002/wrna.1487
- Zingales B. 2018. *Trypanosoma cruzi* genetic diversity: something new for something known about Chagas disease manifestations, serodiagnosis and drug sensitivity. *Acta Trop* **184**: 38–52. doi:10.1016/j.actatropica.2017.09.017

MEET THE FIRST AUTHOR



Evgeny S. Gerasimov

Meet the First Author(s) is a new editorial feature within *RNA*, in which the first author(s) of research-based papers in each issue have the opportunity to introduce themselves and their work to readers of *RNA* and the *RNA* research community. Evgeny Gerasimov is the first author of this paper, "*Trypanosoma cruzi* strain and starvation-driven mitochondrial RNA editing and transcriptome variability." Evgeny is a molecular biologist and bioinformatician working in the field of protist genomics at the Senior Researcher level, who is the author and developer of the RNA editing analysis software "T-Aligner."

What are the major results described in your paper and how do they impact this branch of the field?

One of the major points of our recent paper is that potentially translatable alternatively edited products in *Trypanosoma cruzi* are confirmed for at least three mitochondrial genetic loci. This

points to a possible role of RNA editing as a molecular mechanism of genetic variability. Another important finding is that RNA editing fine-tunes the expression of mitochondrial transcripts in a strain-specific manner and in response to starvation.

What led you to study RNA or this aspect of RNA science?

U-insertion/deletion RNA editing in the mitochondrion is one of the most fascinating features of trypanosomatids which, in spite of a long course of study, remains an open question in protist biology. While enzymatic activities involved in this process are well-studied, the role of this complex molecular system for the cell's evolution remains mostly unclear.

What are some of the landmark moments that provoked your interest in science or your development as a scientist?

The rapid development of NGS technologies gave me the opportunity to apply my skills in computer science to protist biology. Being passionate about programming, I found my place in biology as a software developer and data scientist. Very unusual trypanosomatid genomes and their expression demanded the development of new algorithms and software for data processing. I consider creating these to be the most exciting part of my job.

Are there specific individuals or groups who have influenced your philosophy or approach to science?

Trypanosomatid genomics is definitely a field with excellent investigators, but I would mention Dr. Larry Simpson's lab (UCLA, now retired), whose works I admired when I started to work in this field.

Continued

Gerasimov et al.

What are your subsequent near- or long-term career plans?

Our T-Aligner software makes it possible to perform a variety of different studies on U-insertion/deletion edited transcriptomes in nonmodel kinetoplast genomes. I can now obtain single-cell and

population data sets to analyze for many species of trypanosomatids, which can greatly impact our understanding of evolution of RNA editing.



RNA

A PUBLICATION OF THE RNA SOCIETY

***Trypanosoma cruzi* strain and starvation-driven mitochondrial RNA editing and transcriptome variability**

Evgeny S. Gerasimov, Roger Ramirez-Barrios, Vyacheslav Yurchenko, et al.

RNA 2022 28: 993-1012 originally published online April 25, 2022

Access the most recent version at doi:[10.1261/rna.079088.121](https://doi.org/10.1261/rna.079088.121)

Supplemental Material

<http://rnajournal.cshlp.org/content/suppl/2022/04/25/rna.079088.121.DC1>

References

This article cites 38 articles, 2 of which can be accessed free at:
<http://rnajournal.cshlp.org/content/28/7/993.full.html#ref-list-1>

Creative Commons License

This article is distributed exclusively by the RNA Society for the first 12 months after the full-issue publication date (see <http://rnajournal.cshlp.org/site/misc/terms.xhtml>). After 12 months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

A banner advertisement for Horizon. On the left, there is a colorful 3D molecular model. The text in the center reads: "Use CRISPRmod for targeted modulation of endogenous gene expression to validate siRNA data". On the right, there is the Horizon logo, which includes the word "horizon" and "a PerkinElmer company" below it.

Use CRISPRmod for targeted modulation of endogenous gene expression to validate siRNA data

To subscribe to *RNA* go to:

<http://rnajournal.cshlp.org/subscriptions>
