



# Nuclear Compartmentalization Contributes to Stage-Specific Gene Expression Control in *Trypanosoma cruzi*

Lucía Pastro<sup>1,2</sup>, Pablo Smircich<sup>1,2</sup>, Andrés Di Paolo<sup>3</sup>, Lorena Becco<sup>1</sup>,  
María A. Duhagon<sup>1,2</sup>, José Sotelo-Silveira<sup>3</sup> and Beatriz Garat<sup>1\*</sup>

<sup>1</sup> Laboratorio de Interacciones Moleculares, Facultad de Ciencias, Universidad de la República, Montevideo, Uruguay, <sup>2</sup> Departamento de Genética, Facultad de Medicina, Universidad de la República, Montevideo, Uruguay, <sup>3</sup> Departamento de Genómica, Instituto de Investigaciones Biológicas Clemente Estable, Montevideo, Uruguay

## OPEN ACCESS

### Edited by:

Catalina Dirney Alba Soto,  
Instituto de Investigaciones en  
Microbiología y Parasitología Médica,  
UBA-CONICET, Argentina

### Reviewed by:

Santuza Maria Ribeiro Teixeira,  
Universidade Federal de Minas Gerais,

Brazil

Pedro José Alcolea,  
CSIC, Spain

### \*Correspondence:

Beatriz Garat  
bgarat@fcien.edu.uy

### Specialty section:

This article was submitted to  
Cellular Biochemistry,  
a section of the journal  
Frontiers in Cell and Developmental  
Biology

Received: 21 October 2016

Accepted: 25 January 2017

Published: 13 February 2017

### Citation:

Pastro L, Smircich P, Di Paolo A,  
Becco L, Duhagon MA,  
Sotelo-Silveira J and Garat B (2017)  
Nuclear Compartmentalization  
Contributes to Stage-Specific Gene  
Expression Control in *Trypanosoma  
cruzi*. *Front. Cell Dev. Biol.* 5:8.  
doi: 10.3389/fcell.2017.00008

In the protozoan parasite *Trypanosoma cruzi*, as in other trypanosomatids, transcription of protein coding genes occurs in a constitutive fashion, producing large polycistronic transcription units. These units are composed of non-functionally related genes which are pervasively processed to yield each mRNA. Therefore, post-transcriptional processes are crucial to regulate gene expression. Considering that nuclear compartmentalization could contribute to gene expression regulation, we comparatively studied the nuclear, cytoplasmic and whole cell transcriptomes of the non-infective epimastigote stage of *T. cruzi*, using RNA-Seq. We found that the cytoplasmic transcriptome tightly correlates with the whole cell transcriptome and both equally correlate with the proteome. Nonetheless, 1,200 transcripts showed differential abundance between the nuclear and cytoplasmic fractions. For the genes with transcript content augmented in the nucleus, significant structural and compositional differences were found. The analysis of the reported epimastigote transcriptome and proteome, revealed scarce ribosome footprints and encoded proteins for them. Ontology analyses unveiled that many of these genes are distinctive of other parasite life-cycle stages. Finally, the relocalization of transcript abundance in the metacyclic trypomastigote infective stage was confirmed for specific genes. While gene expression is strongly dependent on transcript steady-state level, we here highlight the importance of the distribution of transcripts abundance between compartments in *T. cruzi*. Particularly, we show that nuclear compartmentation is playing an active role in the developmental stage determination preventing off-stage expression.

**Keywords:** gene expression regulation, RNA-Seq, subcellular fractionation, trypanosomes, differentiation, life cycle

## INTRODUCTION

Chagas' disease is caused by the protozoan parasite *Trypanosoma cruzi* (Kinetoplastida, Trypanosomatidae), which is transmitted to humans by diverse species of the Reduviidae bug family (Chagas, 1909). The infected insect releases the parasite metacyclic trypomastigote forms with its feces while feeding host blood, entering the host through the wound area. Within the host, the metacyclic trypomastigotes can invade nearby cells and therein differentiate into the

intracellular amastigote forms. The latter multiply intracellularly and differentiate into bloodstream trypomastigotes, which are released into the circulation and infect cells from a variety of tissues, differentiating inside them into amastigotes. The bloodstream trypomastigotes can be ingested by bugs and transform in the vector's midgut into the replicative epimastigote and afterward, in the hindgut, into the infective metacyclic trypomastigotes.

Currently, there are approximately 6 to 8 millions of people suffering this disease and several millions are at risk of infection (WHO, 2015). Vector-borne transmission occurs in Central and South America but, since the parasite can also be transmitted by contaminated food, from mother to child and through contaminated blood or organ donations, Chagas' disease has spread to other continents (WHO, 2015).

Trypanosomatids show exceptional mechanisms for the expression of protein coding genes as the constitutive and polycistronic production of primary transcripts and mRNA maturation through the coordinate mechanisms of *trans*-splicing and polyadenylation. Since the mature messengers from the polycistronic transcriptional units (PTUs), are present in different levels, it is considered that post-transcriptional mechanisms should be actively controlling the gene expression in these organisms (Haile and Papadopoulou, 2007; Kramer and Carrington, 2011). Indeed, in *T. cruzi*, transcriptome studies have proved the existence of differences in mRNA abundance between genes which are consecutive in polycistronic transcriptional units (PTUs) (Minning et al., 2009; Smircich et al., 2015). These differences have been mainly attributed to mRNA stability and several *cis*-acting motifs and *trans*-acting factors have been identified (De Gaudenzi et al., 2005; Araujo and Teixeira, 2011; Bayer-Santos et al., 2012; Guerra-Slompo et al., 2012; Li et al., 2012). Besides, other post-transcriptional mechanisms occurring at the cytoplasm, such as those affecting mRNA localization (Cassola et al., 2007; Holetz et al., 2007; Cassola, 2011; Kramer, 2014) and translatability (da Silva Augusto et al., 2015; Smircich et al., 2015) provide additional steps for the control of mRNA availability for protein production.

The role of the nuclear compartment in the control of the cytoplasmic steady-state transcript levels has been recognized as crucial in eukaryotic gene expression regulation (Kohler and Hurt, 2007; Palazzo and Akef, 2012) and is also being studied in trypanosomatids (Nazer et al., 2011, 2012; Fadda et al., 2014; Antwi et al., 2016; De Gaudenzi et al., 2016; Kramer et al., 2016), for a very recent review see (Clayton, 2016). Although specific molecular actors involved in the nucleus-cytoplasmic transport in *T. cruzi* are being unraveled (Cassola and Frasch, 2009; Serpeloni et al., 2011a,b; Camara Mde et al., 2013; Inoue et al., 2014), no global analysis of the impact of nuclear compartmentation has been performed yet.

Transcriptome approaches in *T. cruzi*, have focused on the parasite life cycle, firstly studied by microarray analysis (Minning et al., 2009) and more recently, by RNA-Seq (Smircich et al., 2015; Houston-Ludlam et al., 2016; Li et al., 2016). In these cases, as for the majority of the transcriptome data from other organisms, the analyzed RNA is extracted from the whole cell disregarding the nuclear contribution to the total RNA population. This premise

is supported by the selection of polyA tailed RNA. However, it has been claimed that eukaryotic gene expression analyses using whole cell lysates, inadvertently measure a substantial number of mRNAs that are restrained into the nucleus (Trask et al., 2009). More recently, this assertion has been endorsed by the demonstration of nuclear retention of spliced polyadenylated mRNA (Bahar Halpern et al., 2015).

Since the characterization of nuclear and cytoplasmic RNA may contribute to further understand trypanosomatid gene expression regulation, a comparative *in masse* analysis of nuclear and cytoplasmic RNA in *T. cruzi* epimastigotes was performed and the impact of nuclear-cytoplasmic RNA partitioning in whole cell RNA was evaluated. We found that the use of the cytoplasmic transcriptome does not significantly improve the estimation of protein abundance obtained from whole cell parasite transcriptome. However, we detected that up to 20% of the genes have differential transcript levels between nucleus and cytoplasmic compartments. Transcripts with higher level in the nucleus are significantly longer and have a higher GC content than the ones with higher content in the cytoplasm. Using data from the reported transcriptome and proteome for *T. cruzi* epimastigotes (de Godoy et al., 2012; Smircich et al., 2015), we observed that these transcripts are also characterized by scarce depiction of ribosome footprints and encoded proteins. In addition, we found that the transcripts enriched in the cytoplasmic fraction correspond to genes expressed in the epimastigote stage, while the ones enriched in the nucleus are distinctive of other life cycle stages. For selected specific transcripts, fluorescent *in situ* hybridization (FISH) was used to study the RNA subcellular localization in non-infective epimastigotes and in infective metacyclic trypomastigotes. Altogether, these results support an active role of nuclear compartmentalization in stage-specific gene regulation in *T. cruzi*.

## MATERIALS AND METHODS

### Parasites Culture

The *T. cruzi* Dm28c clone (Contreras et al., 1988) was used. Epimastigotes were maintained at 28°C in liver infusion tryptose (LIT) medium supplemented with 10% heat inactivated fetal bovine serum (FBS). Metacyclic trypomastigotes were prepared as previously (Duhagon et al., 2009). Briefly, epimastigotes at stationary phase were incubated in TAU medium at 28°C for 2 h (Contreras et al., 1985). The parasites were then washed twice with PBS and immediately used.

### Parasite Fractionation

To obtain nucleus and cytoplasmic fractions we used a previous reported methodology (Gomez et al., 1989) with a few modifications. Briefly, approximately  $5 \times 10^9$  parasites were used for each isolation procedure; the parasites were harvested at 1,700 g for 5 min. Cell pellets were resuspended in phosphate-buffered saline, washed twice, and resuspended in 3 volumes of ice-cold hypotonic buffer (10 mM HEPES, pH 7.9; 1.5 mM MgCl<sub>2</sub>; 10 mM KCl; 0.5 mM DTT; 1 mg/mL pepstatin; 0.5 mg/mL leupeptin; 0.5 mM PMSF) and incubated on ice for

10 min. Nonidet P-40 was added to a final concentration of 0.2% and parasites were disrupted by 13 strokes in a 15-mL glass Potter-Elvehjem Dounce homogenizer (Wheaton). Then sucrose was added to a final concentration of 0.35 M. An aliquot was separated and named “Whole cell” (Wc). The lysate was centrifuged at 500 g for 15 min at 4°C. The supernatant was separated and named “Cytoplasm” (C). The pellet was resuspended in 5 volumes of sucrose buffer (0.35 M sucrose; 10 mM HEPES, pH 7.9; 3.3 mM MgCl<sub>2</sub>; 10 mM KCl; 0.5 mM DTT; 1 mg/mL pepstatin; 0.5 mg/mL leupeptin; 0.5 mM PMSF) and centrifuged in a swinging bucket rotor at 1,100 g for 15 min at 4°C. The pellet was resuspended in the same sucrose solution and named “Nucleus” (N). Four independent replicates of the fractioning procedure were done.

## Western Blot Assay

Each Wc, C, and N protein fractions (corresponding to  $5 \times 10^6$  parasites per lane) were separated by electrophoresis in 10 or 12% SDS-PAGE and electro-transferred onto Hybond C Extra membranes (GE Healthcare) following standard procedures. Membranes were blocked by incubation in 5% skim milk powder in PBS-0.1% Tween and were then incubated for 1 h at room temperature with polyclonal, anti-TcTXN1 (Pineyro et al., 2011), anti-TcH2A, FioCruz, Brazil), anti-Tc38 polyclonal antibody (Duhagon et al., 2009), and anti-TcRBP40 (Guerra-Slompo et al., 2012), were used as purification controls. Bound antibodies were detected using an IRDye 800CW and 680CW Goat anti-Rabbit or anti-Mouse IgG (H + L) (Li-Cor), diluted 1:2,500 and analyzed in a G-Box (Syngene).

## RNA Preparation and Sequencing

RNA was isolated from the Wc, C, and N fractions using MirVana kit (Ambion), according to Total RNA purification manufacturer's protocol. The obtained RNA was treated with DNase according to manufacturer's protocol (DNA-free, Ambion), quantified by Nanodrop (Thermo Scientific, USA) and Qubit 2.0 Fluorometer (Invitrogen) and its integrity was checked by Bioanalyzer (Agilent, USA). Genomic DNA control, in the cytoplasmic fraction, was done by qPCR against *gapdh* housekeeping gene. Then, RNA was treated with riboMinus (Invitrogen), the purified RNA was quantified by Nanodrop (Thermo Scientific, USA) and Qubit 2.0 Fluorometer (Invitrogen) and their integrity as well as rRNA depletion were analyzed by Bioanalyzer (Agilent, USA). Identical quantities of RNA derived from each of the four independent biological replicates were pooled to obtain the Whole cell, Nucleus and Cytoplasmic pooled samples. Pooled samples were sequenced by Illumina HiSeq2000 platform with a paired-end read running type and 100 bp cycle as a running condition at Macrogen (Korea).

## Data Analysis

The obtained data were analyzed by CLC Genomics Workbench. In first place a 3' quality trimming was done to the three sequence sets. To determine RNA transcript levels, the *T. cruzi* CLBrener Esmeraldo-like, genome release 4.2 from the TriTrypDB were used for RNA-Seq analysis included in the package. The reads

per kilobase of transcript per million mapped reads (RPKM) were obtained for each gene in each condition. Only genes with 10 reads or more per transcript were used for quantification purposes.

Functional annotation of the NET and CET genes was performed using the Blast2GO tool (Conesa and Gotz, 2008) using a gene cut off of 10. Overrepresentation analyses were calculated by DAVID (Database for Annotation, Visualization and Integrated Discovery) (Dennis et al., 2003).

The intergenic RNA contribution for each fraction was calculated through the differences of reads that map to genome minus those that map to transcriptome (intergenic reads) as percentages.

For the fine location of the *trans*-splicing sites, we used SLAP mapper (Fiebig et al., 2014) *Trypanosoma cruzi* CLBrener Esmeraldo-like V6.0 was set as genome reference. Reads containing at least 8 nucleotides of the splice leader sequence were retrieved and the highest length upstream to the CDS or the one with more read counts was selected for each gene. Meanwhile, for the fine location of the polyadenylation sites, reads containing a track of at least 8 A were listed and the highest length downstream to the CDS or the one with more read counts was selected for each gene. To avoid inconsistencies of UTR delimitation due to partial maturation, only the data derived from the cytoplasmic transcriptome were used for these purposes.

GC-content (GC%) and GC-content in the third codon position (GC3) were calculated using in-house python scripts.

The RNAfold Vienna RNA Package 2.0 algorithm was used to calculate the minimum free energy for the 5' and 3' UTR of transcripts. Default parameters were used and the temperature was set at 37°C (Lorenz et al., 2011).

## Quantitative RT-PCR

cDNA was synthesized from 1 µg of DNase treated RNA (DNA-free, Ambion) using Superscript III kit first strand synthesis (Invitrogen) and random hexamer primers. Quantification of specific products was done by qRT-PCR, using QuantiTect SYBR Green PCR Master Mix (Qiagen). Double stranded products were amplified using specific primers (Table S1) in a real time rotary analyzer RotorGene 6000 (Corbett). Relative amounts of target genes were calculated by normalization with the *gapdh* housekeeping gene that displayed similar values of absolute Ct in the quadruplicates: Wc ( $16.1 \pm 0.2$ ), N ( $16.1 \pm 0.3$ ) and C ( $16.5 \pm 0.1$ ) and of rpkm in the respective transcriptomes (216.5 for Wc, 263.6 for N and 266.5 for C). PCR reaction mixture containing 0.9 µM of each primer was carried out in a final volume of 10 µL. RNA levels were compared using the  $2^{-\Delta\Delta CT}$  method (Livak and Schmittgen, 2001).

## Fluorescence *In situ* Hybridization (FISH)

FISH analysis were done as previously described (Garcia-Silva et al., 2010) with some modifications. Briefly, epimastigotes were cultured, allowed to adhere to polylysine-coated microscope slides for 20 min at room temperature and, after washing twice in PBS, parasites were fixed with 4% paraformaldehyde in PBS for 10 min at room temperature, washed twice with PBS and further incubated in 25 mM NH<sub>4</sub>Cl for 10 min. Parasites were

permeabilized with 0.2% Triton X-100 in PBS for 5 min. Slides were then blocked and prehybridized for 2 h at room temperature in bovine serum albumin 2%,  $5 \times$  Denhardt,  $4 \times$  SSC and 35% deionized formamide (hybridization solution). Assays were performed under denaturing conditions by heating slides at  $75^\circ\text{C}$  for 3 min just prior to the hybridization step (see Table S1 for used probes). Hybridization was performed overnight at  $45^\circ\text{C}$  in a humid chamber in the presence of 1 ng/ml of the indicated oligonucleotide conjugated either to fluorescein amidite (6-FAM) or to the cyanine 3 dye (Cy3) (Table S1). After hybridization, the slides were washed once in  $2 \times$  SSC plus 50% deionized formamide, once in  $2 \times$  SSC, once in  $1 \times$  SSC, and twice in  $0.5 \times$  SSC. Slides were mounted with Prolong (Molecular Probes) containing DAPI. Confocal images were acquired at room temperature using an OLYMPUS FV 300 with lasers Ar 488 and He-Ne 633 (Melles Griot). Merged images were obtained by superimposing the indicated images files in ImageJ software (Schneider et al., 2012).

## Availability of Supporting Data

The data sets supporting the results of this article are available in the Sequence Read Archive repository, BioProject ID: PRJNA342509.

## RESULTS

### Transcriptomics of Nuclear and Cytoplasmic Fractions of *T. cruzi* Epimastigotes

In order to study the distinctive RNA contribution of the nuclear and cytoplasmic compartments to the total transcriptome of *T. cruzi* epimastigotes, a subcellular fractionation was performed. Proteins known to be specifically localized in: nucleus (histone TcH2A), cytoplasm (tryparedoxin TcTXN1), mitochondria (Tc38) and reservosomes (TcRBP40) were used in western blots to evaluate the purity of the fractions (Duhagon et al., 2009; Pineyro et al., 2011; Guerra-Slompo et al., 2012). The latter two markers were used to evaluate the frequent mitochondrial and reservosomal contamination of the nuclear fraction in this organism. As shown in **Figure 1**, the four independent replicates displayed very similar patterns. All the markers were present in the whole cell extract (Wc fraction), whereas the nuclear (N) fraction exhibited a conspicuous signal for the histone protein TcH2A and a faint to not detectable signals for the other three markers. On the contrary, the cytoplasmic (C) fraction showed strong bands for TcTXN1, Tc38, and TcRBP40, and a faint band for the nuclear marker.

Total RNA from each isolated fraction was extracted and qualitative and quantitatively analyzed (Figure S1). The four biological replicates yielded  $0.6 \pm 0.1$  pg (mean  $\pm$  S.E.) of total RNA per epimastigote cell. The nuclear RNA contribution represented  $10 \pm 1\%$  of the RNA in the cytoplasm fraction per epimastigote cell. The differential distribution of well-established nuclear confined RNAs, such as unprocessed rRNA and snoRNA, was analyzed by quantitative RT-PCR (qRT-PCR). A remarkable enrichment of snoRNA and intergenic rRNA in the nuclear fractions was revealed (**Figure 2**). The small standard error

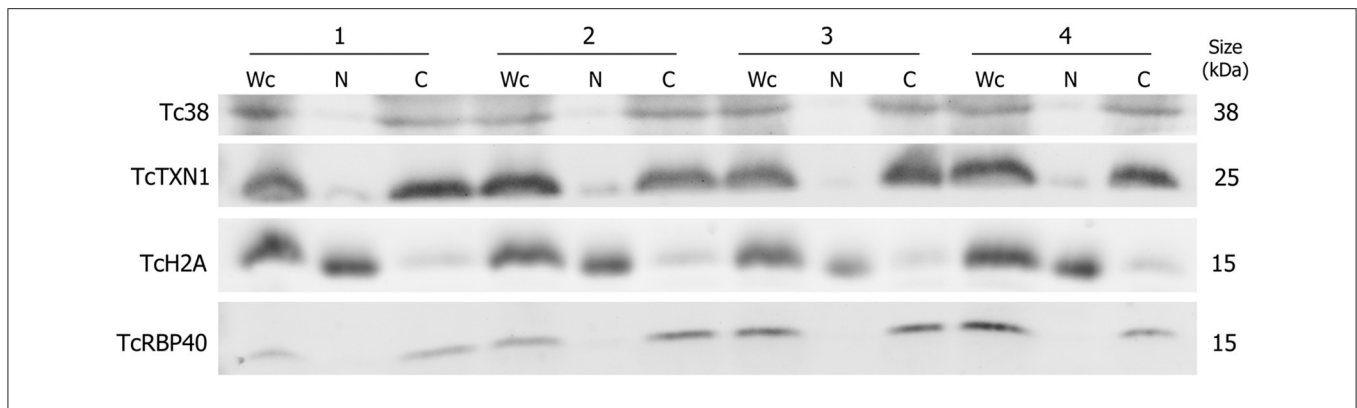
among the four independent experiments argues in favor of the similarity of the quadruplicates (also supported by additional analyses presented in the following section). Besides, no genomic DNA could be detected by qPCR in the non-retrotranscribed RNA from the cytoplasmic fraction. Overall, the protein and RNA markers profile supported the adequacy of the nucleus and cytoplasmic fractions from the four independent replicates for a further RNA-Seq analysis.

Given the reproducibility of the replicates, the RNA extractions of the four biological replicates of each fraction were then pooled and the electrophoretic profile was analyzed. As observed for each individual experiment, the whole cell and the cytoplasmic pooled fractions displayed similar profiles whereas the nuclear fraction was more dissimilar (Figure S2). The pooled fractions were then sequenced as indicated in Materials and Methods section and the raw statistic data are shown in Table S2. RNA-Seq analysis using the *T. cruzi* CLBrener transcriptome yielded roughly 17 million mapped reads for Wc, N, and C. When the reads were mapped to the genomic sequences, we observed a higher proportion of mapped reads in the nuclear than in cytoplasmic fraction. This result was expected because of the presence of intergenic regions in the polycistronic RNA which are lost during maturation in the nucleus previous to the export to cytoplasm (Clayton, 2016). When the reads were mapped to the transcriptome sequences (see Data analysis in Materials and Methods section), we found that most of the 10,600 annotated transcripts present at least one mapped read: 9,079 in the Wc, 9,225 in the N and 9,011 in the C RNA fractions.

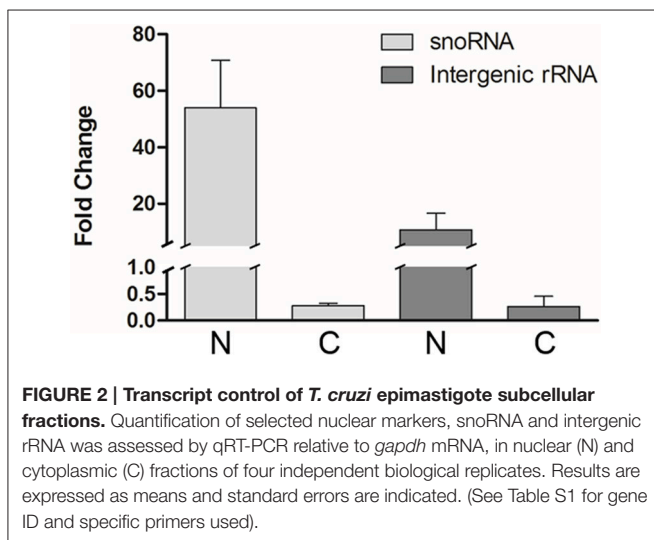
### Transcripts Showing Differential Distribution between the Nuclear and Cytoplasmic Fractions Are Found in *T. cruzi* Epimastigotes

A great proportion of the annotated transcripts were present in the cytoplasmic and nuclear transcriptomes, indicating that the majority of the transcripts are exported to the cytoplasm. This finding is in agreement with the recent hypothesis of pervasive RNA maturation in *T. cruzi* (Smircich et al., 2015). However, since there might be differences between transcript abundance in the cell compartments, we quantitatively compared the transcriptomes. We found a strong correlation between C and Wc transcriptomes (Pearson correlation coefficient  $r = 0.98$   $p < 0.0001$ , **Figure 3** upper panel). This indicates that the Wc transcriptome adequately represents the transcripts present in C, and therefore it is appropriate to describe the global mRNA steady-state level in this organism using whole cell transcriptome approaches. On the other hand, the nuclear transcriptome also correlated with the whole cell transcriptome, although with a smaller correlation coefficient (Pearson correlation coefficient  $r = 0.43$ , **Figure 3** middle panel). Similarly, when we analyzed the correlation between the N and C transcriptomes we determined a Pearson correlation coefficient  $r = 0.53$  (**Figure 3** lower panel). These findings pointed us to the existence of transcript abundance differences between compartments. Consistent with these results on transcriptomes' correlations, similar correlations for Wc and C transcriptome with the epimastigote proteome (de Godoy et al., 2012) were deduced. As expected, a minor





**FIGURE 1 | Protein control of *T. cruzi* epimastigote subcellular fractions.** Western blot analysis of whole cell (Wc), nucleus (N) and cytoplasm (C) fractions using specific antibodies for selected protein markers: Tc38 (mitochondria), TcTXN1 (cytoplasm), TcH2A (nucleus), TcRBP40 (reservosome) are shown for four independent biological replicates (1 to 4). The molecular weights (kDa) of the proteins are indicated on the right.



**FIGURE 2 | Transcript control of *T. cruzi* epimastigote subcellular fractions.** Quantification of selected nuclear markers, snoRNA and intergenic rRNA was assessed by qRT-PCR relative to *gapdh* mRNA, in nuclear (N) and cytoplasmic (C) fractions of four independent biological replicates. Results are expressed as means and standard errors are indicated. (See Table S1 for gene ID and specific primers used).

correlation coefficient with proteome was obtained for the nucleus transcriptome (Table S3).

In order to identify the sets of genes with differential transcript abundance between N and C, we compared their transcriptomes. A total of 1,182 genes (20% of the 6,039 genes with more than 10 reads in the Wc transcriptome) were found to have at least a 2-fold change between the two compartments ( $FC \geq 2$ , gray colored in **Figure 3** lower panel and Table S4 for the list of gene IDs). Among them, 444 were genes with cytoplasmic enriched transcript abundance (CET, dark gray in **Figure 3** lower panel) and 738 had nuclear enriched transcript abundance (NET, light gray in **Figure 3** lower panel), while 4,857 were not differentially distributed (NDT, black in **Figure 3** lower panel). These results indicate that even when the transcriptomes from Wc and C are highly correlated, a set of genes show significant differences of transcript abundance between the nuclear and cytoplasmic compartments in *T. cruzi* epimastigotes.

Robustness of the compartment distribution differences found by RNA-Seq was confirmed for eight genes by qRT-PCR in the

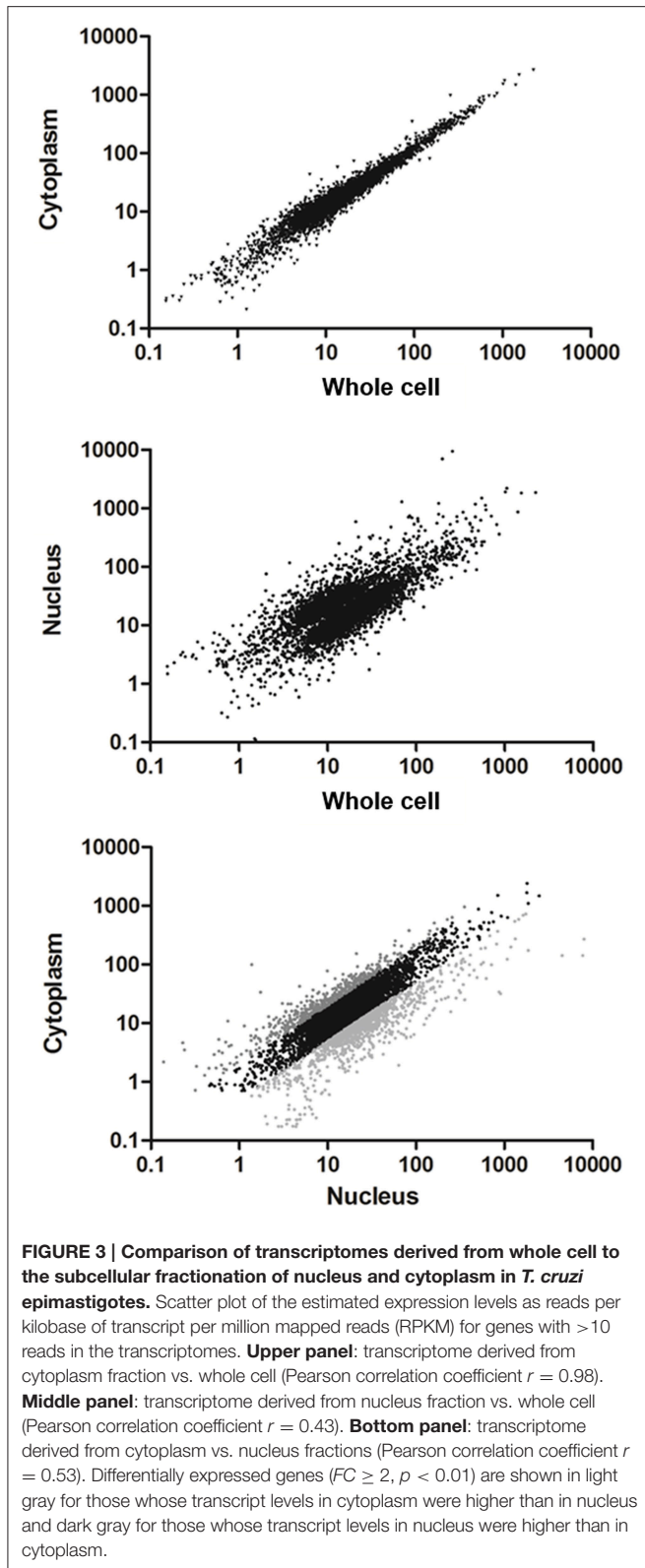
four independent biological replicates (**Figure 4**). Furthermore, the high correlation of transcript quantification by qRT-PCR and by RNA-Seq (Pearson  $r = 0.86$ ) strongly support the RNA-Seq results (**Figure S3**). These data together with the RNA markers (**Figure 2**) were used to compare the individual samples within each cellular fraction. The quadruplicates showed average correlations of  $r \geq 0.96$  with standard deviation  $\leq 0.02$  (**Figure S4A**). In addition, we found that the largest component of the variance of the samples, identified by PCA, comes from the identity of the cellular fraction (**Figure S4B**). These analyses validate the use of a pool of replicates of each subcellular fraction for the identification of differentially expressed genes.

### The Genes with Enriched Transcript Levels in the Nucleus of *T. cruzi* Epimastigotes Have Distinctive Compositional and Structural Characteristics

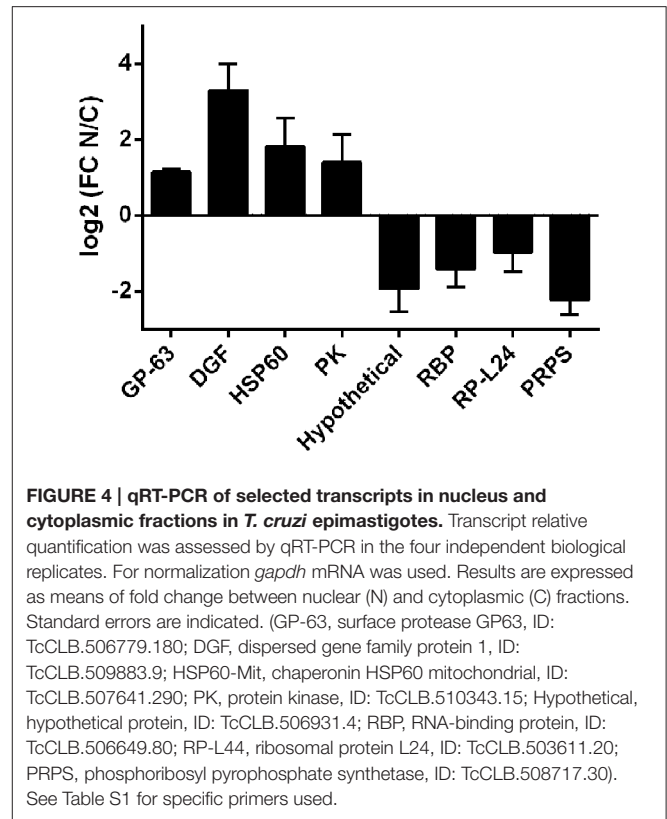
We then investigated the association of properties such as mRNA length and base composition with the nuclear-cytoplasmic differential transcript distribution. Firstly, we had to identify the 5' and 3' ends for the UTRs, since in *T. cruzi* only CDS are annotated. We sought out the UTRs boundaries using the strategy outlined in Materials and Methods section which is based on the detection of at least 8 nucleotides of the splice leader for the 5' UTRs and a track of at least 8 A for the 3' UTRs. For the 1,182 genes exhibiting differentially distributed transcripts, the 5' UTR and 3' UTR could be assigned to 505 and 162 genes respectively. Nonetheless, complete transcripts could only be determined for 81 genes.

The analysis of transcript length revealed that genes whose transcripts were enriched at the cytoplasm were significantly shorter than the non-differentially distributed transcripts (NDT) ( $p < 0.0001$ ) while those whose transcripts were enriched at the nucleus were slightly longer than the ones in the cytoplasm ( $p < 0.0001$ ) (**Figure 5A**). Although this difference mostly relied on CDS size a similar trend was observed for both UTRs.

Concerning the RNA composition, we observed that genes whose transcripts were enriched in the nucleus had a significantly



higher GC% content than NDT genes (Figure 5B). This difference was seen in the GC content at CDS and a similar trend was observed at the UTRs. In addition, the CET GC content at

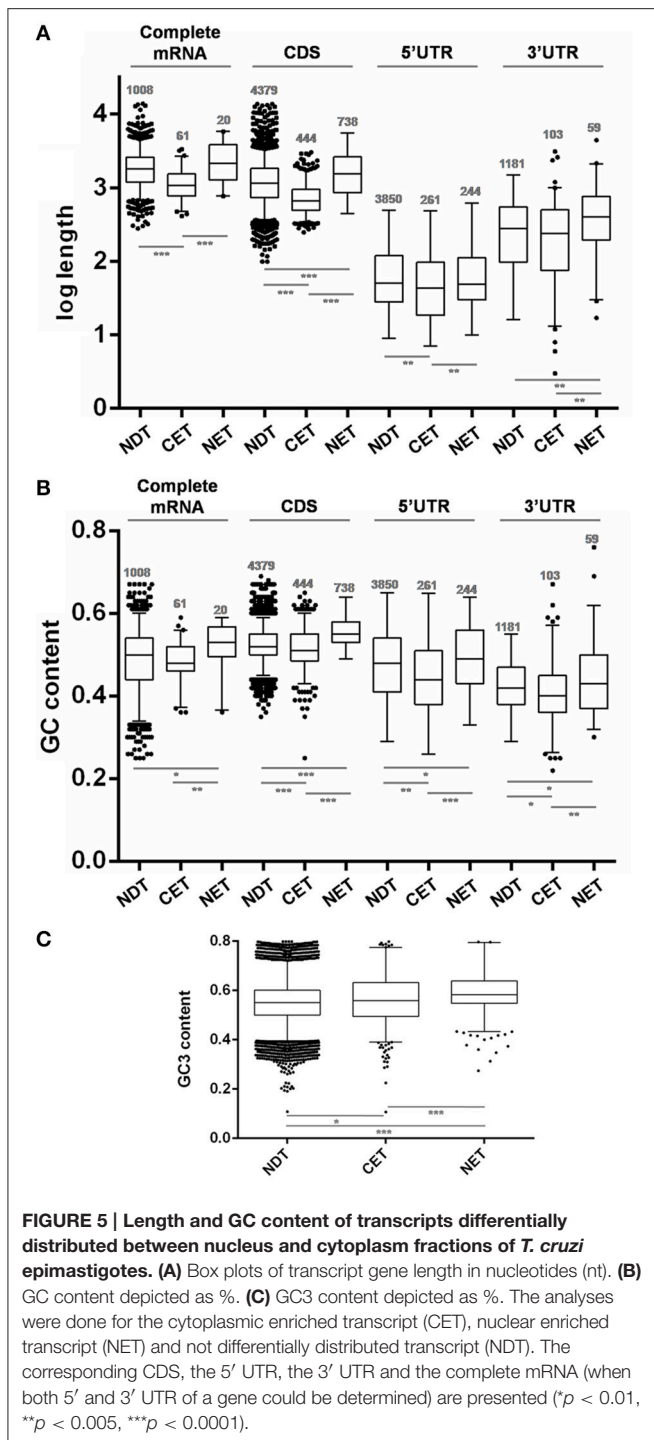


the CDS and also at the determined UTRs was significantly lower than the NET and NDT. For the complete mRNA sequences, a similar tendency, with restricted significances probably due to the low number of determined sequences, was observed. A good agreement between the CDS GC content and GC3 usage was also found (Figure 5C). In addition, since the GC content could affect the RNA conformation, we also investigated the existence of significant thermodynamic differences at the determined UTRs of these sets of transcripts. A significant difference in the free energy was only found for the 5' UTRs of the NET genes when compared to the NDT genes ( $p < 0.01$ ). The lower free energy of the 5' UTRs of the NET genes may be predicting the presence of more stable structures at these regions (Figure S4).

These findings support that gene compositional and structural characteristics may underlay the differential transcript partitioning between nucleus and cytoplasm.

### The Genes with Enriched Transcript Levels in the Nucleus of *T. cruzi* Epimastigotes Are Poorly Translated

Taking advantage of the availability of data derived from ribosome profiling of *T. cruzi* epimastigote (Smircich et al., 2015), we determined the ribosome occupancy of the nucleus-cytoplasm differentially distributed transcripts. A high number of ribosome footprints were observed for CET, while NET had significantly fewer ribosome footprints than either CET or NDT (Figure 6A). Using the quantitative proteomic data for *T. cruzi* epimastigotes (de Godoy et al., 2012) a similar result



was obtained (Figure 6B). Interestingly, virtually all the CET genes were present in the proteome (99%, 440 out of the 444 - dark gray in Figure 3 lower panel), on the contrary, only a few NET genes are detected (6%, 44 out of the 738 - light gray in Figure 3 lower panel). In addition, while CET levels, measured in Wc transcriptome, significantly correlate with their proteomic expression, no correlation is seen for NET genes (Table S5). Consistently, the presence of transcripts of pseudogenes in the

NET set was higher than expected by chance (122 out of 1,302 annotated pseudogenes, 36% increase, Fisher exact test  $p < 0.01$ ). Even more, pseudogenes were underrepresented in the CET genes (21 pseudogenes, 62% decrease, Fisher exact test  $p < 0.01$ ).

Taking together, these findings further support nuclear partitioning impact on protein content modulating transcript availability for translation.

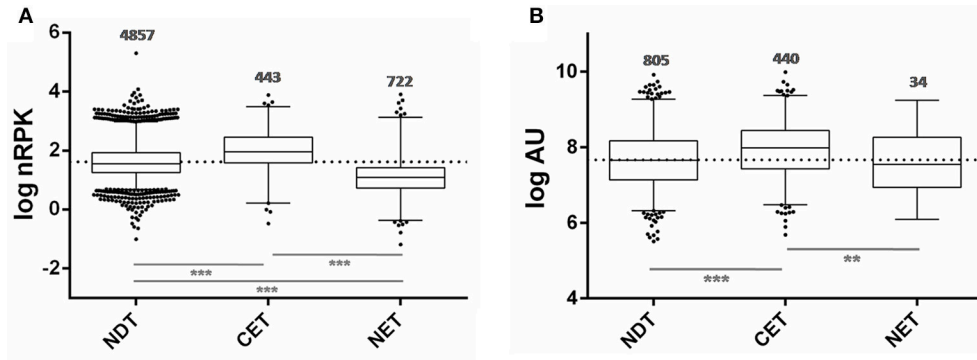
### Off-Stage Transcripts Are Enriched in the Nucleus Compartment of the *T. cruzi* Epimastigote Stage

In order to study the biological characteristics of the NET and CET genes, we performed an analysis of overrepresentation of gene ontology terms (Figure 7). Using the Blast2GO tool (Conesa and Gotz, 2008) we found that the CET genes are involved in biological processes of the replicative and non-infective epimastigote stage such as translation and metabolic pathways (Figure 7A). Conversely, genes encoding factors with active roles in life cycle stages different from the epimastigote stage, such as pathogenesis, have higher transcript abundance in the nucleus (Figure 7B). Concurrently, a detailed inspection of the NET genes exposed an overrepresentation of *trans*-sialidase, DGE, RHS, GP63 and MASP gene family's members ( $\chi^2$ ,  $p < 0.001$ ) and an underrepresentation of ribosomal proteins. The former families of genes have been extensively studied and their expression has been related to the infective metacyclic stage of the parasite (Cuevas et al., 2003; Bartholomeu et al., 2009; Kawashita et al., 2009; De Pablos and Osuna, 2012; Grynberg et al., 2012). Indeed, we found that the ribosome footprints of the epimastigote NET genes increase in the metacyclic trypomastigote stage while those of the epimastigote CET genes are diminished (Figure 8A). A similar pattern was obtained when comparing the metacyclic trypomastigote and epimastigote proteomic data from (de Godoy et al., 2012) (Figure 8B).

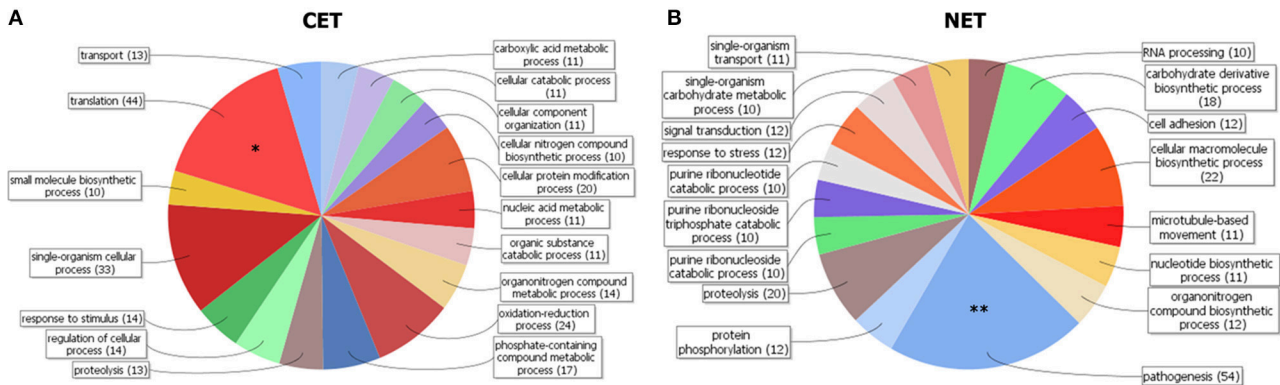
These findings support an active role of the nuclear compartment limiting the off-stage translation of genes from stages different from the actual that could occur due to the constitutive and massive characteristics of transcription in this parasite.

### Epimastigote Nuclear-Cytoplasmic Transcript Distribution Changes after the Transition to the Metacyclic Trypomastigote Stage

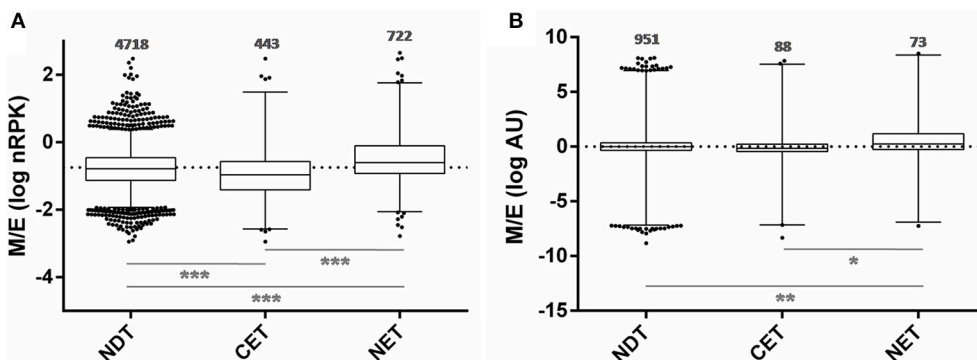
Taking into account all the above results, we hypothesized that transcript subcellular distribution is developmentally regulated. To test this hypothesis, we compared the localization of mRNAs belonging to CET (L44) and NET (GP63 and adenylate cyclase) in the non-infective epimastigote and the infective metacyclic trypomastigote, using fluorescent *in situ* hybridization (FISH). In contrast with the global mRNA localization, as seen by the detection of polyadenylated transcripts, these mRNAs showed developmental dependent localization (Figures 9, 10). The infection related gene, gp63, is faintly expressed in the epimastigote but increases its expression in the metacyclic trypomastigotes and amastigote stages (Cuevas et al., 2003). As



**FIGURE 6 | Ribosome occupancy and Protein abundance of transcripts differentially distributed between nucleus and cytoplasm of *T. cruzi* epimastigotes. (A)** Box plots of ribosomal footprints from Smirich et al. (2015) as normalized reads per kilobase (nRPK) for genes with cytoplasmic enriched transcript abundance (CET), with nuclear enriched transcript abundance (NET) and with not differentially distributed transcript abundance (NDT). **(B)** Box plots of protein abundance for the indicated set of genes obtained from de Godoy et al. (2012) expressed in arbitrary units (AU) (\*\* $p < 0.005$ , \*\*\* $p < 0.0001$ ).

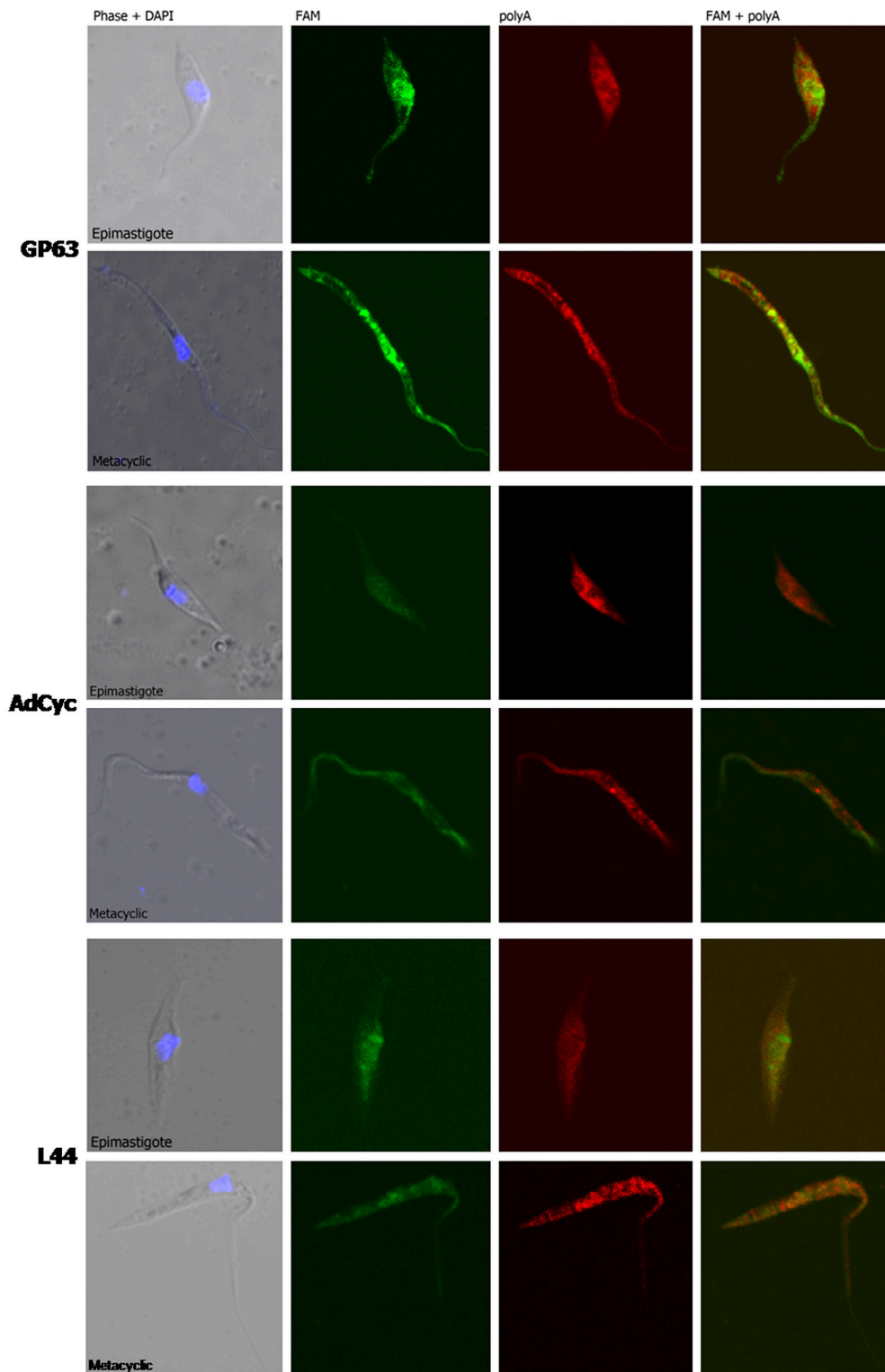


**FIGURE 7 | Gene clustering of nuclear-cytoplasmic differentially distributed transcripts in *T. cruzi* epimastigotes.** Grouping of genes according to Blast2GO biological processes for genes with cytoplasmic enriched transcript (CET, **A**) and nuclear enriched transcript (NET, **B**) is shown. Note the differential increase of genes associated with pathogenesis in the NET group (marked with two asterisks). DAVID overrepresentation is indicated, [ $p$  value:  $1e-10$  (\*) and  $1e-6$  (\*\*)].

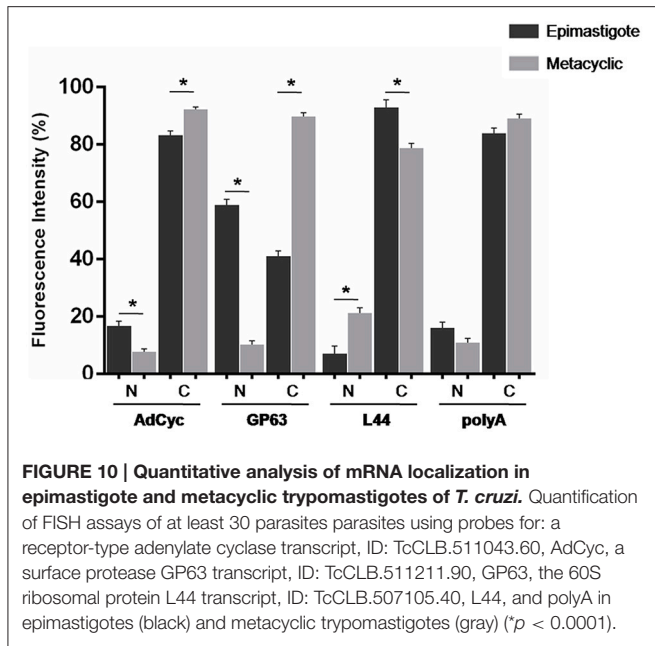


**FIGURE 8 | Translatability and protein abundance of *T. cruzi* epimastigote nuclear-cytoplasmic differentially distributed transcripts in the metacyclic trypomastigote stage. (A)** Box plots of the metacyclic trypomastigote to epimastigote ratio of the ribosomal footprints from Smirich et al. (2015) as normalized reads per kilobase (nRPK) for genes with cytoplasmic enriched transcript (CET), nuclear enriched transcript (NET) and not differentially distributed transcript (NDT) in *T. cruzi* epimastigotes. **(B)** Idem to (A) using the protein abundances expressed in arbitrary units (AU) from de Godoy et al. (2012) ( $p < 0.01$ , \*\* $p < 0.005$ , \*\*\* $p < 0.0001$ ).





**FIGURE 9 | mRNA localization in epimastigote and metacyclic trypomastigotes of *T. cruzi*.** Confocal images of FISH assays of representative parasites using probes for: a surface protease GP63 transcript, ID: TcCLB.511211.90, GP63, upper panel. A receptor-type adenylate cyclase transcript, ID: TcCLB.511043.60, AdCyc, middle panel; and the 60S ribosomal protein L44 transcript, ID: TcCLB.507105.40, L44 bottom. At least 30 parasites were analyzed for each probe in each condition. Staining of epimastigotes (upper panels) and metacyclic trypomastigotes (bottom panels) are presented. These probes (See Table S2) were 5' labeled with 6-FAM (green). For the detection of polyA tailed transcripts a 5' Cy5 labeled polyT probe was used (red). DNA was visualized by DAPI staining (blue). Phase contrast images were obtained. Magnification 120x.



shown in **Figure 9**, the gp63 transcripts were mainly localized in the nucleus in the epimastigote stage, while they acquired a cytoplasmic localization in the metacyclic trypomastigote stage. In addition, the superposition of polyA and the specific gp63 transcripts suggests an active translation in the metacyclic trypomastigote. A similar behavior was displayed by the adenylate cyclase transcript (also belonging to the NET gene set), being clear a depletion of the nuclear signal in the metacyclic trypomastigotes. For, the ribosomal protein L44 transcripts (belonging to the CET genes) a decrease of the colocalization with the polyA probe was observed in the transition from the epimastigote to the metacyclic trypomastigote stage. The quantitative analysis of the images for at least 30 parasites for each probe at the epimastigote or metacyclic trypomastigote stages, confirmed the displayed pattern changes (**Figure 10**).

These results support the existence of changes in the subcellular localization of transcript abundances accompanying the metacyclogenesis developmental process in *T. cruzi*.

## DISCUSSION

The steady-state level of proteins is the result of successive, alternative and combinatorial steps starting at transcription initiation. In *T. cruzi*, as well as in other trypanosomatids, the control of this step seems to be mostly absent. Therefore, trypanosomatids have been considered unique models for the study of post-transcriptional processes and their regulation. Indeed, processes like editing and *trans*-splicing were firstly described in trypanosomatids and later revealed to also occur, in other eukaryotes. However, while in eukaryotes the co-transcriptional RNA processing, surveillance and transport from nucleus to cytoplasm have been widely recognized as regulatory steps in gene expression (Kohler and Hurt, 2007), their survey in trypanosomes is very recent (Nazer et al., 2011, 2012; Fadda

et al., 2014; Antwi et al., 2016; De Gaudenzi et al., 2016; Kramer et al., 2016). Likewise, genome-wide approaches have been used to identify the RNA nature and abundance distribution in the nucleus and the cytoplasm in many eukaryotes (Liao et al., 2010; Solnestam et al., 2012; Tilgner et al., 2012; Zaghlool et al., 2013; Bai et al., 2014), but such analyses have not yet been reported in trypanosomatids.

Seeking to study the role of the nuclear compartment in gene expression control in trypanosomes, we undertook cellular fractionation to obtain RNA from purified nucleus and cytoplasm fractions from *T. cruzi* epimastigotes to perform RNA-Seq and find whether transcripts were differentially distributed between these two compartments.

The fractionation reproducibility and the quality of the fractions were confirmed by protein and RNA compartment specific markers. An expected profile of protein markers was obtained for each of the independent experiment. The whole cell and cytoplasm extracts exhibited similar content of the different markers except for the nuclear marker TcH2A, whose almost disappearance in the later fractions supports the appropriateness of the selected fractionation method. Analogously, the mitochondrion, the reservosome as well as the cytoplasm markers exhibited weak to null signals in the nuclear fractions. The observed pattern may be pointing out to a negligible cross contamination of the fractions, however, the faint signal of TcH2A in the cytoplasm may also be attributed to the protein synthesis occurring at this compartment.

The content of total RNA per cell obtained from *T. cruzi* epimastigotes (0.6 ± 0.1 pg) is in good accordance with reported measurements in *T. brucei* bloodstream (0.5 pg, Haanstra et al., 2008, and procyclic forms 0.8 pg, Pays et al., 1993, and 1.1 pg, Haanstra et al., 2008). Variations of total RNA per cell content in the same range (0.5 to 1.6 pg) have also been reported for *Saccharomyces cerevisiae*, depending on the growth conditions (von der Haar, 2008). While the amount of total RNA in the nuclear fraction may vary with the physiological state of the cell, a rough estimation of 5 to 10% is commonly accepted based on the contribution of rRNA, mRNA and tRNA plus pre-mRNAs of 85, 2, and 13% respectively (Finka et al., 2015). Thus, we considered that the distribution of total RNA between nucleus and cytoplasm that we obtained (10 and 90% respectively) was consistent with the literature.

Thereafter, the RNA derived from the fractions of four independent experiments were pooled, sequenced and analyzed. Despite the fact that the genome of CLBrenner Esmeraldo-like *T. cruzi* strain was used to map the RNA transcripts, the use of a paired end strategy strengthens the reliability of the assignments, and more than 9,000 genes (out of 10,600) were detected in each of the transcriptomes analyzed. Thus, in *T. cruzi* epimastigotes most of the genes are detected in the RNA extracted from whole cells or the nuclear or cytoplasmic fractions. While the good correlation between cytoplasmic and whole cell transcriptome (*r* = 0.98 *p* < 0.0001) was expected due to the reduced total RNA content of the nuclear fraction, the wide representation of transcribed genome in the cytoplasmic fraction (85%) further supports pervasive mechanisms for transcription and maturation as previously proposed (Smirich et al., 2015). However, though

almost all the genes are shared by nuclear and cytoplasmic transcriptomes (84%), the poor correlation obtained between nucleus and cytoplasm transcripts levels ( $r = 0.53$ ,  $p < 0.001$ ) points out to the existence of an active compartment role. Indeed, 20% of the transcripts detected (NET+CET) showed at least a 2-fold change in expression levels between the two compartments analyzed. The reliability of these changes was confirmed for eight genes by qRT-PCR in each of the four independent biological replicates. A high correlation between qRT-PCR and RNA-Seq data was found. In addition, the low dispersion of the obtained qRT-PCR results provides evidence for the similarity of the used replicates. Altogether these findings support the existence of differentially distributed transcripts between nucleus and cytoplasm.

In agreement with previous reports for three human cell lines (Solnestam et al., 2012), where longer transcripts in the nucleus were inferred, we found that NET genes are significantly longer either than the NDT or CET genes. While this profile is mostly due to the size of the CDSs, the size of the UTRs also accompanies this trend. Concordantly, human genome-wide studies have revealed that highly expressed genes are significantly smaller and produce shorter mRNAs with shorter 3' UTRs (Chiaromonte et al., 2003; Urrutia and Hurst, 2003). Similar results have been reported for *Arabidopsis thaliana* (Caldwell et al., 2010) and yeasts (Lackner et al., 2007; Lu et al., 2007). Recently, the length of mRNA has been related to developmental expression in *T. brucei* (Antwi et al., 2016). This finding is consistent with the role of nuclear compartmentalization in gene expression regulation during differentiation that we have here unraveled. Nonetheless, other features of gene sequences, such as G+C contents, are also associated with gene expression levels (Konu and Li, 2002). We found that NET genes have a significantly higher GC% content than NDT genes either at CDSs or UTRs. Since this trend is also accompanied by GC3 content at the CDS, codon usage may be contributing to further regulate gene expression of these transcripts. Codon usage has been identified as a major determinant of mRNA stability (Presnyak et al., 2015). In addition, high GC contents have been related to thermostability in different studies in bacteria (Nishio et al., 2003). Indeed, we found a significantly higher predicted thermodynamic stability for the 5' UTR of NET genes which may be further influencing their low translation rate. Therefore, although the subcellular fractionation to obtain a cytoplasmic transcriptome may not introduce a significant advantage comparing to the whole cell transcriptome, we here provide evidence for a putative bias estimation due to length and composition of the transcripts in *T. cruzi* epimastigotes.

Considering that the transcriptome derived from a purified cytoplasmic fraction, following separation of the nuclear contribution, could better reflect the translationally active transcripts than the whole cell transcriptome, we analyzed the correlations between the transcriptome and the proteome. Again, we found that cytoplasmic transcriptome does not significantly improve the estimation of protein abundance obtained from whole cell parasite transcriptome. However, for the set of genes which have differential transcript abundance between the nucleus and cytoplasmic compartments, a functional bias was observed. The transcripts enriched in the cytoplasm correspond to genes

which are: expressed in the epimastigote following ontology analyses; translationally active according to the ribosome footprint analysis (Smircich et al., 2015), and are widely represented in the epimastigote proteome (de Godoy et al., 2012). Meanwhile, an important number of the transcripts enriched in the nucleus correspond to pseudogenes. This is not surprising since several studies have demonstrated the functional roles of the pseudogene transcripts in post-transcription regulation mainly in the nucleus, either as competing endogenous RNAs or *trans*-acting RNA (Johnsson et al., 2014; Sen et al., 2014). A deeper analysis of the specific pseudogene transcript nature, sequence and structure could shed light to better understand their mechanism of action in this organism. In addition to pseudogenes, several *bona fide* gene transcripts showed a nuclear enrichment. From reported ribosome footprint studies, we found that these genes are less translated than the average in epimastigote. Likewise, very few of them are present in the epimastigote proteome. Ontology analyses revealed that the genes with nucleus enriched transcripts in the epimastigote are distinctive of other life cycle stages. Indeed, several genes belong to the extended multigenic families that codify for surface antigens such as the trans-sialidase (Schenkman et al., 1994; Nardy et al., 2016), the mucin associated surface proteins (Bartholomeu et al., 2009; dos Santos et al., 2012), the dispersed gene family (Kawashita et al., 2009; De Pablos and Osuna, 2012) and the glycoprotein family gp63 (Cuevas et al., 2003; Yao, 2010). Their function has been frequently associated with parasite protection and evasion of the host immune system (dos Santos et al., 2012; De Pablos et al., 2016; Houston-Ludlam et al., 2016). Since these gene families are absent in the phylogenetically closest free-living trypanosomatid *Bodo saltans*, it has been recently suggested that their acquisition could be related to the parasite adaptation to the hostile environments of the hosts (Jackson et al., 2016). Besides, the family of genes with hot spot for the insertion of retroelements, which also has nucleus enriched transcripts, is known to increase its expression upon exposure to  $\gamma$  radiation (Grynberg et al., 2012). Thus, the nuclear compartmentation may contribute to maintain the availability of transcripts necessary to readily respond to environmental challenges.

Using the available data for ribosome footprinting and proteome of *T. cruzi* metacyclic trypomastigote, we analyzed the fate of the NET and CET epimastigote genes in the infective non-replicative parasite stage. We could demonstrate a significant change of the ribosome occupancy of transcripts and encoded proteins from these two transcript sets during metacyclogenesis, that is in good agreement with a role of the nuclear compartment in determining which mRNAs will be retained or released to the translationally active cytoplasmic compartment.

We further studied whether this developmental relocalization could be also proved through direct visualization of selected transcripts. In epimastigotes, the ribosomal protein L44 mRNA is found in the cytoplasm, mainly at the nucleus boundaries, and colocalizing with the polyadenylated mRNAs. Meanwhile, in addition to its reported decreased expression in the metacyclic trypomastigote, a wider spread pattern along the parasite, with no nuclear delimitation and a lesser colocalization with polyadenylated mRNAs was observed. On the other hand, the adenylate cyclase, known to be expressed in the metacyclic

trypomastigote stage showed a diminished nuclear transcript content in this stage in comparison to the epimastigote stage. The change of location was even more evident for a gp63 gene involved in the infectivity process, whose transcript was markedly restricted to the nuclear translational inactive compartment in the non-infective epimastigote stage, moving to the translationally active cytoplasm, with a good colocalization with polyadenylated mRNAs, in the infective metacyclic trypomastigote stage.

These findings support an active role of the nuclear compartment confining the transcripts that are not immediately needed to produce proteins but may be rapidly required upon differentiation. The analysis of the actual contribution of each transcript species, i.e., unprocessed, partial or complete processed transcript to the nuclear content would shed light about the implied regulatory mechanisms. In addition, whether this differential RNA distribution is achieved by nuclear retention or differential nuclear-cytoplasm RNA decay should be investigated. Altogether these results provide strong evidence indicating that the nucleus-cytoplasm partitioning constitutes a control step that contributes to the differential regulation of a life cycle specific set of genes in *T. cruzi*.

## AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: MD, JS, and BG; Performed the experiments: LP; Analyzed the data: LP, PS, AD,

LB, MD, JS, and BG; Contributed reagents/materials/analysis tools: PS, AD, LB, JS, and BG; Wrote the paper: LP, PS, MD, JS, and BG.

## FUNDING

This work was supported by CSIC and CAP UdelaR, ANII and PEDECIBA, Uruguay.

## ACKNOWLEDGMENTS

We wish to thank Gisele Picchi and Eloise Guerra-Slomp (Instituto Carlos Chagas, FioCruz Paraná, Brasil) and Dolores Piñeyro (Institut Pasteur Montevideo, Uruguay) for kindly providing TcH2A, TcRBP40, and TcTXN1 specific antibodies, respectively. We acknowledge all members of the Laboratorio de Interacciones Moleculares at UDELAR and the Department of Genomics at IIBCE for constant discussion and technical support. We also thank the colleagues that have provided critical insight into this study.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fcell.2017.00008/full#supplementary-material>

## REFERENCES

- Antwi, E. B., Haanstra, J. R., Ramasamy, G., Jensen, B., Droll, D., Rojas, F., et al. (2016). Integrative analysis of the *Trypanosoma brucei* gene expression cascade predicts differential regulation of mRNA processing and unusual control of ribosomal protein expression. *BMC Genomics* 17:306. doi: 10.1186/s12864-016-2624-3
- Araujo, P. R., and Teixeira, S. M. (2011). Regulatory elements involved in the post-transcriptional control of stage-specific gene expression in *Trypanosoma cruzi*: a review. *Mem. Inst. Oswaldo Cruz* 106, 257–266. doi: 10.1590/S0074-02762011000300002
- Bahar Halpern, K., Caspi, I., Lemze, D., Levy, M., Landen, S., Elinav, E., et al. (2015). Nuclear retention of mRNA in mammalian tissues. *Cell Rep.* 13, 2653–2662. doi: 10.1016/j.celrep.2015.11.036
- Bai, B., Yegnasubramanian, S., Wheelan, S. J., and Laiho, M. (2014). RNA-Seq of the nucleolus reveals abundant SNORD44-derived small RNAs. *PLoS ONE* 9: e107519. doi: 10.1371/journal.pone.0107519
- Bartholomeu, D. C., Cerqueira, G. C., Leao, A. C., daRocha, W. D., Pais, F. S., Macedo, C., et al. (2009). Genomic organization and expression profile of the mucin-associated surface protein (masp) family of the human pathogen *Trypanosoma cruzi*. *Nucleic Acids Res.* 37, 3407–3417. doi: 10.1093/nar/gkp172
- Bayer-Santos, E., Gentil, L. G., Cordero, E. M., Correa, P. R., and da Silveira, J. F. (2012). Regulatory elements in the 3' untranslated region of the GP82 glycoprotein are responsible for its stage-specific expression in *Trypanosoma cruzi* metacyclic trypomastigotes. *Acta Trop.* 123, 230–233. doi: 10.1016/j.actatropica.2012.03.014
- Caldwell, R., Lin, Y. X., and Zhang, R. (2010). Assessment of length distributions between non-coding and coding sequences amongst two model organisms. *Int. J. Data Min. Bioinform.* 4, 535–552. doi: 10.1504/IJDMB.2010.035899
- Camara Mde, L., Bouvier, L. A., Canepa, G. E., Miranda, M. R., and Pereira, C. A. (2013). Molecular and functional characterization of a *Trypanosoma cruzi* nuclear adenylate kinase isoform. *PLoS Negl. Trop. Dis.* 7:e2044. doi: 10.1371/journal.pntd.0002044
- Cassola, A. (2011). RNA Granules Living a Post-transcriptional Life: the Trypanosomes' Case. *Curr. Chem. Biol.* 5, 108–117. doi: 10.2174/2212796811105020108
- Cassola, A., De Gaudenzi, J. G., and Frasch, A. C. (2007). Recruitment of mRNAs to cytoplasmic ribonucleoprotein granules in trypanosomes. *Mol. Microbiol.* 65, 655–670. doi: 10.1111/j.1365-2958.2007.05833.x
- Cassola, A., and Frasch, A. C. (2009). An RNA recognition motif mediates the nucleocytoplasmic transport of a trypanosome RNA-binding protein. *J. Biol. Chem.* 284, 35015–35028. doi: 10.1074/jbc.M109.031633
- Chagas, C. (1909). Nova tripanozomíaze humana. Estudos sobre a morfologia e o ciclo evolutivo do *Schizotrypanum cruzi*, agente etiologico da nova entidade mórbida do homen. *Mem. Inst. Oswaldo Cruz* 1, 159–219. doi: 10.1590/S0074-02761909000200008
- Chiaromonte, F., Miller, W., and Bouhassira, E. E. (2003). Gene length and proximity to neighbors affect genome-wide expression levels. *Genome Res.* 13, 2602–2608. doi: 10.1101/gr.1169203
- Clayton, C. E. (2016). Gene expression in Kinetoplastids. *Curr. Opin. Microbiol.* 32, 46–51. doi: 10.1016/j.mib.2016.04.018
- Conesa, A., and Gotz, S. (2008). Blast2GO: a comprehensive suite for functional analysis in plant genomics. *Int. J. Plant Genomics* 2008:619832. doi: 10.1155/2008/619832
- Contreras, V. T., Araujo-Jorge, T. C., Bonaldo, M. C., Thomaz, N., Barbosa, H. S., Meirelles Mde, N., et al. (1988). Biological aspects of the Dm 28c clone of *Trypanosoma cruzi* after metacyclogenesis in chemically defined media. *Mem. Inst. Oswaldo Cruz* 83, 123–133. doi: 10.1590/S0074-02761988000100016
- Contreras, V. T., Morel, C. M., and 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



- Cuevas, I. C., Cazzulo, J. J., and Sanchez, D. O. (2003). gp63 homologues in *Trypanosoma cruzi*: surface antigens with metalloprotease activity and a possible role in host cell infection. *Infect. Immun.* 71, 5739–5749. doi: 10.1128/IAI.71.10.5739-5749.2003
- da Silva Augusto, L., Moretti, N. S., Ramos, T. C., de Jesus, T. C., Zhang, M., Castilho, B. A., et al. (2015). A membrane-bound eIF2 alpha kinase located in endosomes is regulated by heme and controls differentiation and ROS levels in *Trypanosoma cruzi*. *PLoS Pathog.* 11:e1004618. doi: 10.1371/journal.ppat.1004618
- De Gaudenzi, J., Frasch, A. C., and Clayton, C. (2005). RNA-binding domain proteins in Kinetoplastids: a comparative analysis. *Eukaryot. Cell* 4, 2106–2114. doi: 10.1128/EC.4.12.2106-2114.2005
- De Gaudenzi, J. G., Jager, A. V., Izcovich, R., and Campo, V. A. (2016). Insights into the regulation of mRNA processing of polycistronic transcripts mediated by DRBD4/PTB2, a trypanosome homolog of the polypyrimidine tract-binding protein. *J. Eukaryot. Microbiol.* 63, 440–452. doi: 10.1111/jeu.12288
- de Godoy, L. M., Marchini, F. K., Pavoni, D. P., Rampazzo Rde, C., Probst, C. M., Goldenberg, S., et al. (2012). Quantitative proteomics of *Trypanosoma cruzi* during metacyclogenesis. *Proteomics* 12, 2694–2703. doi: 10.1002/pmic.201200078
- Dennis, G. Jr., Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C., et al. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol.* 4:P3. doi: 10.1186/gb-2003-4-5-p3
- De Pablos, L. M., Diaz Lozano, I. M., Jercic, M. I., Quinzada, M., Gimenez, M. J., Calabuig, E., et al. (2016). The C-terminal region of *Trypanosoma cruzi* MASP5 is antigenic and secreted via exovesicles. *Sci. Rep.* 6:27293. doi: 10.1038/srep27293
- De Pablos, L. M., and Osuna, A. (2012). Multigene families in *Trypanosoma cruzi* and their role in infectivity. *Infect. Immun.* 80, 2258–2264. doi: 10.1128/iai.06225-11
- dos Santos, S. L., Freitas, L. M., Lobo, F. P., Rodrigues-Luiz, G. F., Mendes, T. A., Oliveira, A. C., et al. (2012). The MASP family of *Trypanosoma cruzi*: changes in gene expression and antigenic profile during the acute phase of experimental infection. *PLoS Negl. Trop. Dis.* 6:e1779. doi: 10.1371/journal.pntd.0001779
- Duhagon, M. A., Pastro, L., Sotelo-Silveira, J. R., Perez-Diaz, L., Maugeri, D., Nardelli, S. C., et al. (2009). The *Trypanosoma cruzi* nucleic acid binding protein Tc38 presents changes in the intramitochondrial distribution during the cell cycle. *BMC Microbiol.* 9:34. doi: 10.1186/1471-2180-9-34
- Fadda, A., Ryten, M., Droll, D., Rojas, F., Farber, V., Haanstra, J. R., et al. (2014). Transcriptome-wide analysis of trypanosome mRNA decay reveals complex degradation kinetics and suggests a role for co-transcriptional degradation in determining mRNA levels. *Mol. Microbiol.* 94, 307–326. doi: 10.1111/mmi.12764
- Fiebig, M., Gluenz, E., Carrington, M., and Kelly, S. (2014). SLAP mapper: a webserver for identifying and quantifying spliced-leader addition and polyadenylation site usage in kinetoplastid genomes. *Mol. Biochem. Parasitol.* 196, 71–74. doi: 10.1016/j.molbiopara.2014.07.012
- Finka, A., Sood, V., Quadroni, M., Rios Pde, L., and Goloubinoff, P. (2015). Quantitative proteomics of heat-treated human cells show an across-the-board mild depletion of housekeeping proteins to massively accumulate few HSPs. *Cell Stress Chaperones* 20, 605–620. doi: 10.1007/s12192-015-0583-2
- Garcia-Silva, M. R., Frugier, M., Tosar, J. P., Correa-Dominguez, A., Ronaltes-Alves, L., Parodi-Talice, A., et al. (2010). A population of tRNA-derived small RNAs is actively produced in *Trypanosoma cruzi* and recruited to specific cytoplasmic granules. *Mol. Biochem. Parasitol.* 171, 64–73. doi: 10.1016/j.molbiopara.2010.02.003
- Gomez, M. L., Erijman, L., Arauzo, S., Torres, H. N., and Tellez-Inon, M. T. (1989). Protein kinase C in *Trypanosoma cruzi* epimastigote forms: partial purification and characterization. *Mol. Biochem. Parasitol.* 36, 101–108. doi: 10.1016/0166-6851(89)90182-5
- Grynberg, P., Passos-Silva, D. G., Mourao Mde, M., Hirata, R. Jr., Macedo, A. M., Machado, C. R., et al. (2012). *Trypanosoma cruzi* gene expression in response to gamma radiation. *PLoS ONE* 7:e29596. doi: 10.1371/journal.pone.0029596
- Guerra-Slomp, E. P., Probst, C. M., Pavoni, D. P., Goldenberg, S., Krieger, M. A., and Dallagiovanna, B. (2012). Molecular characterization of the *Trypanosoma cruzi* specific RNA binding protein TcRBP40 and its associated mRNAs. *Biochem. Biophys. Res. Commun.* 420, 302–307. doi: 10.1016/j.bbrc.2012.02.154
- Haanstra, J. R., Stewart, M., Luu, V. D., van Tuijl, A., Westerhoff, H. V., Clayton, C., et al. (2008). Control and regulation of gene expression: quantitative analysis of the expression of phosphoglycerate kinase in bloodstream form *Trypanosoma brucei*. *J. Biol. Chem.* 283, 2495–2507. doi: 10.1074/jbc.M705782200
- Haile, S., and Papadopolou, B. (2007). Developmental regulation of gene expression in trypanosomatid parasitic protozoa. *Curr. Opin. Microbiol.* 10, 569–577. doi: 10.1016/j.mib.2007.10.001
- Holetz, F. B., Correa, A., Avila, A. R., Nakamura, C. V., Krieger, M. A., and Goldenberg, S. (2007). Evidence of P-body-like structures in *Trypanosoma cruzi*. *Biochem. Biophys. Res. Commun.* 356, 1062–1067. doi: 10.1016/j.bbrc.2007.03.104
- Houston-Ludlam, G. A., Belew, A. T., and El-Sayed, N. M. (2016). Comparative Transcriptome Profiling of Human Foreskin Fibroblasts Infected with the Sylvio and Y Strains of *Trypanosoma cruzi*. *PLoS ONE* 11:e0159197. doi: 10.1371/journal.pone.0159197
- Inoue, A. H., Serpeloni, M., Hiraiwa, P. M., Yamada-Ogatta, S. F., Muniz, J. R., Motta, M. C., et al. (2014). Identification of a novel nucleocytoplasmic shuttling RNA helicase of trypanosomes. *PLoS ONE* 9:e109521. doi: 10.1371/journal.pone.0109521
- Jackson, A. P., Otto, T. D., Aslett, M., Armstrong, S. D., Bringaud, F., Schlacht, A., et al. (2016). Kinetoplastid phylogenomics reveals the evolutionary innovations associated with the origins of parasitism. *Curr. Biol.* 26, 161–172. doi: 10.1016/j.cub.2015.11.055
- Johnsson, P., Morris, K. V., and Grander, D. (2014). Pseudogenes: a novel source of trans-acting antisense RNAs. *Methods Mol. Biol.* 1167, 213–226. doi: 10.1007/978-1-4939-0835-6\_14
- Kawashita, S. Y., da Silva, C. V., Mortara, R. A., Burleigh, B. A., and Briones, M. R. (2009). Homology, paralogy and function of DGF-1, a highly dispersed *Trypanosoma cruzi* specific gene family and its implications for information entropy of its encoded proteins. *Mol. Biochem. Parasitol.* 165, 19–31. doi: 10.1016/j.molbiopara.2008.12.010
- Kohler, A., and Hurt, E. (2007). Exporting RNA from the nucleus to the cytoplasm. *Nat. Rev. Mol. Cell Biol.* 8, 761–773. doi: 10.1038/nrm2255
- Konu, O., and Li, M. D. (2002). Correlations between mRNA expression levels and GC contents of coding and untranslated regions of genes in rodents. *J. Mol. Evol.* 54, 35–41. doi: 10.1007/s00239-001-0015-z
- Kramer, S. (2014). RNA in development: how ribonucleoprotein granules regulate the life cycles of pathogenic protozoa. *Wiley Interdiscip. Rev. RNA* 5, 263–284. doi: 10.1002/wrna.1207
- Kramer, S., and Carrington, M. (2011). Trans-acting proteins regulating mRNA maturation, stability and translation in trypanosomatids. *Trends Parasitol.* 27, 23–30. doi: 10.1016/j.pt.2010.06.011
- Kramer, S., Piper, S., Estevez, A., and Carrington, M. (2016). Polycistronic trypanosome mRNAs are a target for the exosome. *Mol. Biochem. Parasitol.* 205, 1–5. doi: 10.1016/j.molbiopara.2016.02.009
- Lackner, D. H., Beilharz, T. H., Marguerat, S., Mata, J., Watt, S., Schubert, F., et al. (2007). A network of multiple regulatory layers shapes gene expression in fission yeast. *Mol. Cell* 26, 145–155. doi: 10.1016/j.molcel.2007.03.002
- Li, Y., Shah-Simpson, S., Okrah, K., Belew, A. T., Choi, J., Caradonna, K. L., et al. (2016). Transcriptome Remodeling in *Trypanosoma cruzi* and Human Cells during Intracellular Infection. *PLoS Pathog.* 12:e1005511. doi: 10.1371/journal.ppat.1005511
- Li, Z. H., De Gaudenzi, J. G., Alvarez, V. E., Mendiondo, N., Wang, H., Kissinger, J. C., et al. (2012). A 43-nucleotide U-rich element in 3'-untranslated region of large number of *Trypanosoma cruzi* transcripts is important for mRNA abundance in intracellular amastigotes. *J. Biol. Chem.* 287, 19058–19069. doi: 10.1074/jbc.M111.338699
- Liao, J. Y., Ma, L. M., Guo, Y. H., Zhang, Y. C., Zhou, H., Shao, P., et al. (2010). Deep sequencing of human nuclear and cytoplasmic small RNAs reveals an unexpectedly complex subcellular distribution of miRNAs and tRNA 3' trailers. *PLoS ONE* 5:e10563. doi: 10.1371/journal.pone.0010563
- Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta C_T}$  Method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262
- Lorenz, R., Bernhart, S. H., Höner Zu Siederdisen, C., Tafer, H., Flamm, C., Stadler, P. F., et al. (2011). ViennaRNA Package 2.0. *Algorithms Mol. Biol.* 6:26. doi: 10.1186/1748-7188-6-26

- Lu, P., Vogel, C., Wang, R., Yao, X., and Marcotte, E. M. (2007). Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nat. Biotechnol.* 25, 117–124. doi: 10.1038/nbt1270
- Minning, T. A., Weatherly, D. B., Atwood, J. III, Orlando, R., and Tarleton, R. L. (2009). The steady-state transcriptome of the four major life-cycle stages of *Trypanosoma cruzi*. *BMC Genomics* 10:370. doi: 10.1186/1471-2164-10-370
- Nardy, A. F., Freire-de-Lima, C. G., Perez, A. R., and Morrot, A. (2016). Role of *Trypanosoma cruzi* Trans-sialidase on the escape from host immune surveillance. *Front. Microbiol.* 7:348. doi: 10.3389/fmicb.2016.00348
- Nazer, E., Verdun, R. E., and Sanchez, D. O. (2011). Nucleolar localization of RNA binding proteins induced by actinomycin D and heat shock in *Trypanosoma cruzi*. *PLoS ONE* 6:e19920. doi: 10.1371/journal.pone.0019920
- Nazer, E., Verdun, R. E., and Sanchez, D. O. (2012). Severe heat shock induces nucleolar accumulation of mRNAs in *Trypanosoma cruzi*. *PLoS ONE* 7:e43715. doi: 10.1371/journal.pone.0043715
- Nishio, Y., Nakamura, Y., Kawarabayasi, Y., Usuda, Y., Kimura, E., Sugimoto, S., et al. (2003). Comparative complete genome sequence analysis of the amino acid replacements responsible for the thermostability of *Corynebacterium efficiens*. *Genome Res.* 13, 1572–1579. doi: 10.1101/gr.1285603
- Palazzo, A. F., and Akef, A. (2012). Nuclear export as a key arbiter of “mRNA identity” in eukaryotes. *Biochim. Biophys. Acta* 1819, 566–577. doi: 10.1016/j.bbagr.2011.12.012
- Pays, E., Hanocq-Quertier, J., Hanocq, F., Van Assel, S., Nolan, D., and Rolin, S. (1993). Abrupt RNA changes precede the first cell division during the differentiation of *Trypanosoma brucei* bloodstream forms into procyclic forms *in vitro*. *Mol. Biochem. Parasitol.* 61, 107–114. doi: 10.1016/0166-6851(93)90163-R
- Pineyro, M. D., Parodi-Talice, A., Portela, M., Arias, D. G., Guerrero, S. A., and Robello, C. (2011). Molecular characterization and interactome analysis of *Trypanosoma cruzi* tryparedoxin 1. *J. Proteomics* 74, 1683–1692. doi: 10.1016/j.jprot.2011.04.006
- Presnyak, V., Alhusaini, N., Chen, Y. H., Martin, S., Morris, N., Kline, N., et al. (2015). Codon optimality is a major determinant of mRNA stability. *Cell* 160, 1111–1124. doi: 10.1016/j.cell.2015.02.029
- Schenkman, S., Eichinger, D., Pereira, M. E., and Nussenzweig, V. (1994). Structural and functional properties of *Trypanosoma* trans-sialidase. *Annu. Rev. Microbiol.* 48, 499–523. doi: 10.1146/annurev.mi.48.100194.002435
- Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675. doi: 10.1038/nmeth.2089
- Sen, R., Ghosal, S., Das, S., Balti, S., and Chakrabarti, J. (2014). Competing endogenous RNA: the key to posttranscriptional regulation. *Sci. World J.* 2014:896206. doi: 10.1155/2014/896206
- Serpeloni, M., Moraes, C. B., Muniz, J. R., Motta, M. C., Ramos, A. S., Kessler, R. L., et al. (2011a). An essential nuclear protein in trypanosomes is a component of mRNA transcription/export pathway. *PLoS ONE* 6:e20730. doi: 10.1371/journal.pone.0020730
- Serpeloni, M., Vidal, N. M., Goldenberg, S., Avila, A. R., and Hoffmann, F. G. (2011b). Comparative genomics of proteins involved in RNA nucleocytoplasmic export. *BMC Evol. Biol.* 11:7. doi: 10.1186/1471-2148-11-7
- Smircich, P., Eastman, G., Bispo, S., Duhagon, M. A., Guerra-Slompo, E. P., Garat, B., et al. (2015). Ribosome profiling reveals translation control as a key mechanism generating differential gene expression in *Trypanosoma cruzi*. *BMC Genomics* 16:443. doi: 10.1186/s12864-015-1563-8
- Solnestam, B. W., Stranneheim, H., Hallman, J., Kaller, M., Lundberg, E., Lundeberg, J., et al. (2012). Comparison of total and cytoplasmic mRNA reveals global regulation by nuclear retention and miRNAs. *BMC Genomics* 13:574. doi: 10.1186/1471-2164-13-574
- Tilgner, H., Knowles, D. G., Johnson, R., Davis, C. A., Chakraborty, S., Djebali, S., et al. (2012). Deep sequencing of subcellular RNA fractions shows splicing to be predominantly co-transcriptional in the human genome but inefficient for lncRNAs. *Genome Res.* 22, 1616–1625. doi: 10.1101/gr.134445.111
- Trask, H. W., Cowper-Sal-lari, R., Sartor, M. A., Gui, J., Heath, C. V., Renuka, J., et al. (2009). Microarray analysis of cytoplasmic versus whole cell RNA reveals a considerable number of missed and false positive mRNAs. *RNA* 15, 1917–1928. doi: 10.1261/rna.1677409
- Urrutia, A. O., and Hurst, L. D. (2003). The signature of selection mediated by expression on human genes. *Genome Res.* 13, 2260–2264. doi: 10.1101/gr.641103
- von der Haar, T. (2008). A quantitative estimation of the global translational activity in logarithmically growing yeast cells. *BMC Syst. Biol.* 2:87. doi: 10.1186/1752-0509-2-87
- World Health Organization (WHO) (2015). *Chagas Disease (American trypanosomiasis)*. Fact sheet N°340. Geneva: WHO. Available online at: <http://www.who.int/mediacentre/factsheets/fs340/en/>
- Yao, C. (2010). Major surface protease of trypanosomatids: one size fits all? *Infect. Immun.* 78, 22–31. doi: 10.1128/IAI.00776-09
- Zaghlool, A., Ameer, A., Nyberg, L., Halvardson, J., Grabherr, M., Cavelier, L., et al. (2013). Efficient cellular fractionation improves RNA sequencing analysis of mature and nascent transcripts from human tissues. *BMC Biotechnol.* 13:99. doi: 10.1186/1472-6750-13-99

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Pastro, Smircich, Di Paolo, Becco, Duhagon, Sotelo-Silveira and Garat. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.