Investigation of Transcription Start Site-Associated Regulatory Elements in Trypanosoma cruzi

Authors: ***Leticia, Pedro, David, John, Martin, Julia***

Trypanosoma cruzi, the causative agent of Chagas disease, exhibits a remarkable genomic organization in which protein-coding genes are arranged in polycistronic transcription units (PTUs), flanked by divergent and convergent strand switch regions (dSSRs and cSSRs). Its genome is organized into two main compartments: the core, enriched in housekeeping genes, and the disruptive, composed by virulence-associated genes. Transcription of protein-coding genes occurs polycistronically by RNA polymerase II, and gene regulation is considered to occur mainly at the post-transcriptional level. Although post-transcriptional regulation has long been considered the central mechanism, accumulating evidence indicates that additional regulatory layers, including chromatin organization and transcription initiation, may also play important roles. In this study, we aimed to identify transcription start sites (TSSs) and to investigate potential cis-regulatory elements that may contribute to gene regulation of transcription initiation in *T. cruzi.* To this end, we applied the triphosphorylated small RNA sequencing (3pRNA-seq) to map primary transcripts in the epimastigote form of the parasite. This strategy enabled the identification of nascent RNAs, particularly enriched within dSSRs, supporting the hypothesis that transcription in *T. cruzi* is preferentially initiated in regions of divergence. In addition, we investigate the epigenetic landscape and the presence of putative cis-regulatory elements across distinct classes of dSSRs. Notably, dSSRs located between disruptive PTUs differ significantly from those between core PTUs in terms of size, presence of 5-hydroxymethylcytosine, nucleosome density, and conserved motifs. Moreover, integration of the 3pRNA-seq data revealed that TSSs identified in different classes of dSSRs exhibit distinct patterns associated with transcriptional directionality. Collectively, all these findings suggest that distinct genomic compartments display differential nascent transcriptional activity, which may be associated with underlying structural and epigenetic features. To characterize potential cis-regulatory elements associated with TSSs, we integrated our 3pRNA-seq data and selected candidates sequences located in: (i) dSSRs from both core and disruptive compartments; (ii) non-dSSRs regions within PTUs of both compartments; and (iii) regions proximal to 3-dimensional genome features, such as chromatin boundaries. To functionally validate the transcriptional activity of the selected sequences, we have generated a luciferase-based reporter vector with a multiple cloning site upstream of a red-shifted *Photinus pyralis* luciferase, followed by an mNeonGreen cassette, establishing a dual-reporter system to evaluate the transcription initiation of the selected sequences in genetically modified *T. cruzi* strains. Our results provide new insights into the transcriptional landscape of *T. cruzi* and contribute to a deeper understanding of gene regulation mechanisms in this parasite.

Supported by FAPESP, CNPQ and CAPES.