



Data Article

RNA-seq data exploration after trypanosome RNA-binding protein UBP1 expression is altered by CRISPR-Cas9 gene editing and overexpression



Karina B. Sabalette^{a,b,1}, Vanina A. Campo^{a,b},
José R. Sotelo-Silveira^{c,d}, Pablo Smircich^{c,d}, Javier G. De
Gaudenzi^{a,b,*}

^a Instituto de Investigaciones Biotecnológicas, Universidad Nacional de San Martín - Consejo Nacional de Investigaciones Científicas y Técnicas, 1650 General San Martín, Prov. de Buenos Aires, Argentina

^b Escuela de Bio y Nanotecnologías (EBYN), Universidad Nacional de San Martín, 1650 General San Martín, Prov. de Buenos Aires, Argentina

^c Instituto de Investigaciones Biológicas Clemente Estable, Av. Italia 3318, Montevideo, CP 11600, Uruguay

^d Instituto de Biología, School of Sciences, Universidad de la República, Montevideo, Uruguay

ARTICLE INFO

Article history:

Received 11 January 2024

Revised 29 January 2024

Accepted 30 January 2024

Available online 6 February 2024

Dataset link: [RNA-seq of TcUBP1 knockdown in Trypanosoma cruzi CL-Brener epimastigotes \(Original data\)](#)
Dataset link: [RNA-seq of TcUBP1 transgenic Trypanosoma cruzi CL-Brener epimastigotes \(Original data\)](#)

ABSTRACT

Previous studies have shown that overexpression of the *Trypanosoma cruzi* U-rich RNA-binding protein 1 (TcUBP1) in insect-dwelling epimastigotes results in a gene expression pattern resembling that of the infective form of the pathogen. Here, we used CRISPR-Cas9-induced edition of TcUBP1 and full-length protein overexpression in epimastigote cells to monitor transcriptomic changes during the epimastigote-to-metacyclic trypomastigote stage transition of *T. cruzi*. This dataset includes the bioinformatics analysis of three different RNA-seq samples, each with three biological replicates, showing differential mRNA abundances. The current transcriptome report has the potential to shed light on the quantitative variances in the expression of significant

* Corresponding author at: Instituto de Investigaciones Biotecnológicas, Universidad Nacional de San Martín - Consejo Nacional de Investigaciones Científicas y Técnicas, 1650 General San Martín, Prov. de Buenos Aires, Argentina.

E-mail address: jdegauzenzi@iib.unsam.edu.ar (J.G. De Gaudenzi).

¹ Present address: European Molecular Biology Laboratory, Grenoble, 71 Avenue des Martyrs, CS 90181, 38042 Grenoble Cedex 9, France.

Keywords:
RNA-binding protein
Transcriptomics
Trypanosoma cruzi
Differential gene expression
Parasite differentiation

up- or down-regulated mRNAs as a consequence of the levels of the UBP1 protein. Raw data files were deposited at the NCBI Sequence Read Archive - SRA at <http://ncbi.nlm.nih.gov/Traces/sra/sra.cgi> with accession numbers PRJNA907231 and PRJNA949967.

© 2024 The Authors. Published by Elsevier Inc.
This is an open access article under the CC BY license
(<http://creativecommons.org/licenses/by/4.0/>)

Specifications Table

Subject	Transcriptomics
Specific subject area	Molecular Parasitology
Data format	Transcriptome dataset of TcUBP1-CRISPR/Cas9 gene editing and TcUBP1-GFP overexpressing epimastigote cells in <i>Trypanosoma cruzi</i>
Type of data	Raw, Analyzed
Data collection	Table, Figure
Data source location	High-throuput sequencing carried out by BGI using using 2 × 100 PE chemistry on DNBSeg platform
Data accessibility	Buenos Aires, Argentina RNA-seq data from this study has been submitted to the NCBI Sequence Archive - SRA at http://ncbi.nlm.nih.gov/Traces/sra/sra.cgi with accession numbers PRJNA907231 and PRJNA949967 (available at https://www.ncbi.nlm.nih.gov/bioproject/PRJNA907231 and https://www.ncbi.nlm.nih.gov/bioproject/PRJNA949967).

1. Value of the Data

- This is the first transcriptome study of *Trypanosoma cruzi* RNA-binding protein UBP1-overexpressing and knockdown parasites.
- The RNA-Seq dataset underlines the quantitative differences of significant up- or down-regulated mRNAs as a direct effect of UBP1 protein levels.
- Further gene expression analysis will provide a data source for understanding of how *T. cruzi* epimastigotes differentiate to infective metacyclic trypomastigotes.

2. Background

The present study focused on the U-rich RNA-binding protein 1 (UBP1, TcCLB.507093.220), one of the first trypanosome RNA-recognition motif-containing proteins to be described. Remarkably, *UBP1* gene is encoded in a single stable dicistronic unit and its protein product regulates the abundance of several genes that have U-rich elements in *Trypanosoma cruzi* [1,2]. We have already shown that when UBP1 is overexpressed in insect-dwelling epimastigotes, the microorganism shows an expression pattern that looks like the infective form [3]. To further investigate the regulatory role of this protein, we used the CRISPR-Cas9 technology to generate a population of parasites that did not express this protein. However, we obtained an N-terminal mutated protein that is significantly less expressed than the endogenous form found in normal cells [4].

3. Data Description

This data describes a transcriptomic experiment to investigate the differential gene expression profiles between the *T. cruzi* N-terminal mutated UB1P1 knockdown [4] and full-length protein overexpression in epimastigote cells. The experimental design consists of three high-throughput sequencing samples with three biological replicates each of wild-type RNA-binding protein TcUBP1 (WT), UB1P1-GFP tetracycline-induced epimastigotes for four days (UB1P1-OE), and CRISPR/Cas edited epimastigotes (UB1P1mut-KD). Differential gene analysis was conducted using DESeq2 [5], with the criteria of at least 2-fold change ($|\text{Log}_2 \text{ fold change}| > 1$, FDR-adjusted p value < 0.05). The expression values listed in Supplementary Table 1 show the average Log_2 fold change for all mapped genes in UB1P1-OE or UB1P1mut-KD versus WT. TcUBP1 overexpression in epimastigote cells led to a significant increase in the abundance of 801 genes and a downregulation of 362 genes. In the population of UB1P1mut-KD parasites, a significant decrease of 26 genes and an increase of 67 genes was observed (Fig. 1). Next, the transcriptome dataset was examined to check the amount of UB1P1 transcript levels within the two conditions, with UB1P1 being the experimental internal control of our RNA-Seq analysis. As expected, the Log_2 fold change obtained for UB1P1-OE/WT was 6.16 (FDR-adjusted p value = $2.1\text{E-}84$) and the Log_2 fold change of UB1P1mut-KD/WT was -0.90 (FDR-adjusted p value = $2.1\text{E-}02$). In addition, the oppositely affected genes after UB1P1 overexpression and knockdown were recently reported by Sabalette *et al.* (see Ref. [4]). Raw data is available through NCBI's Sequence Read Archive at <http://www.ncbi.nlm.nih.gov/sra.cgi> with accession numbers: PRJNA907231 and PRJNA949967.

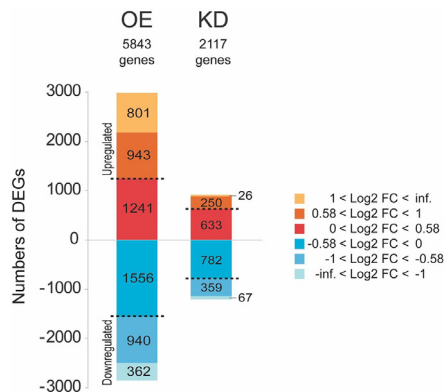


Fig. 1. A bar graph for the number of genes identified in the two conditions. Stacked bar charts show the distribution of the number of significant genes and DEGs identified in each pairwise comparison between OE (UB1P1-OE) and KD (UB1P1mut-KD) versus WT.

4. Experimental Design, Materials and Methods

Control wild-type UB1P1, four-day tetracycline-induced UB1P1 parasites [6], and UB1P1-CRISPR/Cas9 edited cells [4] were cultured in epimastigote form conditions at 28 °C. Total RNA from three biological replicates was prepared from approximately 10^7 epimastigote cells of each sample (WT, UB1P1-OE, and UB1P1mut-KD) using the TRIzol reagent following the manufacturer's instructions (Invitrogen). Isolation of poly(A)+ mRNA, library preparation, and deep-sequencing on an DNBSseq platform were performed at BGI. 100 nt reads were mapped to the *T. cruzi* reference genome CL Brener Esmeraldo-like strain (TriTrypDB-59_TcruziCLBrenerEsmeraldo-like_Genome.fasta), as described [3]. Differential gene analysis was conducted using DESeq2 [5]. All RNA-seq raw data files for WT, UB1P1-overexpression, and UB1P1-CRISPR/Cas9 edited samples, used in this study, are available as FASTQ files of 100 bp paired-end

reads in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) database with the following study numbers: PRJNA907231 and PRJNA949967.

Limitations

Not applicable.

Ethics statement

The authors have read and follow the ethical requirements for publication in Data in Brief and confirming that the current work does not involve human subjects, animal experiments, or any data collected from social media platforms.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT (OpenAI) in order to improve language and readability. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Data availability

[RNA-seq of TcUBP1 knockdown in Trypanosoma cruzi CL-Brener epimastigotes \(Original data\)](#) (SRA)

[RNA-seq of TcUBP1 transgenic Trypanosoma cruzi CL-Brener epimastigotes \(Original data\)](#) (SRA)

CRedit Author Statement

Karina B. Sabalette: Methodology; **Vanina A. Campo:** Methodology, Formal analysis; **José R. Sotelo-Silveira:** Data curation, Formal analysis; **Pablo Smircich:** Data curation, Formal analysis, Funding acquisition; **Javier G. De Gaudenzi:** Formal analysis, Visualization, Writing – original draft, Conceptualization, Funding acquisition, Supervision.

Acknowledgments

This work was supported by Agencia Nacional de Promoción Científica y Tecnológica Grant PICT [2019-00737](#) and CONICET Grant [PIP-GI 0302](#) (to J.G.D.G.), International Centre For Genetic Engineering and Biotechnology Research: Grants Programme 2022 Grant [CRP/URY22-03](#) (to P.S. and J.G.D.G.), and funds from the Research Career of CONICET (to V.A.C. and J.G.D.G.).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.dib.2024.110156](https://doi.org/10.1016/j.dib.2024.110156).

References

- [1] J.G. De Gaudenzi, A.V. Jäger, R. Izcovich, V.A. Campo, 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 (4) (2016) 440–452 JulEpub 2016 Jan 5, doi:[10.1111/jeu.12288](https://doi.org/10.1111/jeu.12288).
- [2] Z.H. Li, J.G. De Gaudenzi, V.E. Alvarez, N. Menciondo, H. Wang, J.C. Kissinger, A.C. Frasch, R. Docampo, 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 (23) (2012) 19058–19069 Jun 1Epub 2012 Apr 12, doi:[10.1074/jbc.M111.338699](https://doi.org/10.1074/jbc.M111.338699).
- [3] K.B. Sabalette, J.R. Sotelo-Silveira, P. Smircich, J.G. De Gaudenzi, RNA-Seq reveals that overexpression of TcUBP1 switches the gene expression pattern toward that of the infective form of *Trypanosoma cruzi*, *J. Biol. Chem.* 299 (5) (2023) 104623 MayEpub 2023 Mar 17, doi:[10.1016/j.jbc.2023.104623](https://doi.org/10.1016/j.jbc.2023.104623).
- [4] K.B. Sabalette, V.A. Campo, J.R. Sotelo-Silveira, P. Smircich, J.G. De Gaudenzi, Transcriptomic analysis of N-terminal mutated *Trypanosoma cruzi* UB1 knockdown underlines the importance of this RNA-binding protein in parasite development, *bioRxiv* (2023), doi:[10.1101/2023.12.07.570581](https://doi.org/10.1101/2023.12.07.570581).
- [5] M.I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, *Genome Biol.* 15 (12) (2014) 550, doi:[10.1186/s13059-014-0550-8](https://doi.org/10.1186/s13059-014-0550-8).
- [6] K.B. Sabalette, M.A. Romaniuk, G. Noé, A. Cassola, V.A. Campo, J.G. De Gaudenzi, The RNA-binding protein TcUBP1 up-regulates an RNA regulon for a cell surface-associated *Trypanosoma cruzi* glycoprotein and promotes parasite infectivity, *J. Biol. Chem.* 294 (26) (2019) 10349–10364 Jun 28Epub 2019 May 21, doi:[10.1074/jbc.RA118.007123](https://doi.org/10.1074/jbc.RA118.007123).