



# Guide RNA Repertoires in the Main Lineages of *Trypanosoma cruzi*: High Diversity and Variable Redundancy Among Strains

## OPEN ACCESS

Fanny Rusman, Noelia Florida-Yapur, Nicolás Tomasini\*<sup>†</sup> and Patricio Diosque\*<sup>†</sup>

### Edited by:

Juan David Ramírez,  
Rosario University, Colombia

### Reviewed by:

Philipp Schwabl,  
Harvard University, United States  
Carlos Robello,  
Universidad de la República, Uruguay  
Rodrigo Baptista,  
University of Georgia, United States

### \*Correspondence:

Nicolás Tomasini  
nicolas.tomasini@conicet.gov.ar  
Patricio Diosque  
patricio.diosque@unsa.edu.ar

<sup>†</sup>These authors have contributed  
equally to this work and  
share senior authorship

### Specialty section:

This article was submitted to  
Parasite and Host,  
a section of the journal  
Frontiers in Cellular and  
Infection Microbiology

Received: 02 February 2021

Accepted: 12 May 2021

Published: 31 May 2021

### Citation:

Rusman F, Florida-Yapur N,  
Tomasini N and 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

Unidad de Epidemiología Molecular (UEM), Instituto de Patología Experimental, Universidad Nacional de Salta-CONICET, Salta, Argentina

*Trypanosoma cruzi*, as other kinetoplastids, has a complex mechanism of editing of mitochondrial mRNAs that requires guide RNAs (gRNAs) coded in DNA minicircles in the kinetoplast. There are many variations on this mechanism among species. mRNA editing and gRNA repertoires are almost unknown in *T. cruzi*. Here, gRNAs were inferred based on deep-sequenced minicircle hypervariable regions (mHVRs) and editing cascades were rebuilt in strains belonging to the six main *T. cruzi* lineages. Inferred gRNAs were clustered according to their sequence similarity to constitute gRNA classes. Extreme diversity of gRNA classes was observed, which implied highly divergent gRNA repertoires among different lineages, even within some lineages. In addition, a variable gRNA class redundancy (i.e., different gRNA classes editing the same mRNA region) was detected among strains. Some strains had upon four times more gRNA classes than others. Such variations in redundancy affected gRNA classes of all mRNAs in a concerted way, i.e., there are correlated variations in the number of gRNAs classes editing each mRNA. Interestingly, cascades were incomplete for components of the respiratory complex I in several strains. Finally, gRNA classes of different strains may potentially edit mitochondrial mRNAs from other lineages in the same way as they edit their own mitochondrial mRNAs, which is a prerequisite for biparental inheritance of minicircle in hybrids. We propose that genetic exchange and biparental inheritance of minicircles combined with minicircle drift due to (partial) random segregation of minicircles during kDNA replication is a suitable hypothesis to explain the divergences among strains and the high levels of gRNA redundancy in some strains. In addition, our results support that the complex I may not be required in some stages in the life cycle as previously shown and that linkage (in the same minicircle) of gRNAs that edit different mRNAs may prevent gRNA class lost in such stage.

**Keywords:** Chagas disease, kinetoplastids, minicircles, RNA editing, DTU

## INTRODUCTION

Kinetoplastids are a phylogenetic group of flagellate protozoa that include the genus *Leishmania* and *Trypanosoma*, with three species causing neglected diseases: Leishmaniasis, Chagas disease and sleeping sickness. Interestingly, kinetoplastids have a single large mitochondrion with a disk shape structure called kinetoplast. This structure has a complex network of concatenated DNA rings (Cavalcanti and de Souza, 2018). Two different kinds of DNA rings are contained in such network: maxicircles and minicircles. Maxicircles are large molecules (20–30 kb) that appear in a few dozens of almost identical copies. They code for two mitochondrial ribosomal subunits and eighteen mitochondrial proteins (most of them participating in the respiratory chain) (Simpson et al., 1987; Ruvalcaba-Trejo and Sturm, 2011; Lin et al., 2015). Several of such genes require post-transcriptional modifications in their pre-mRNA to generate functional open reading frames. The process is known as mRNA editing and it is made by inserting or deleting uridines (U) (Benne et al., 1986; Koslowsky et al., 1991; Stuart, 1993; Kim et al., 1994; Simpson et al., 1996; Hajduk et al., 1997; Hajduk and Ochsenreiter, 2010; Ammerman et al., 2012; Aphasizheva et al., 2016). Some mRNAs require extensive editing, a phenomenon known as pan-editing [e.g. more than six hundred edited bases to generate the cytochrome c oxidase subunit III (COIII) open reading frame in *T. brucei* (Feagin et al., 1988)]. Instead, other mRNAs are just partially or non-edited.

Uridine insertion/deletion is directed by short RNAs called guide RNAs (gRNAs) and it works as a cascade of steps. Basically, the first gRNA binds the pre-mRNA at a specific and complementary site (non-canonic G::U base pairing is allowed) in the 3' extreme of the mRNA. Then, it determines the editing of the adjacent sites at the 5' based on mismatches (Estevez and Simpson, 1999). This modified region is recognized by another gRNA which drives the editing of the next adjacent region, and successively up to generate the initial codon in the 5' end. If gRNAs that edit a region are lost, the editing is stopped, and no translation is made (Estevez and Simpson, 1999; Lukes et al., 2005). Most editing cascades have been identified for *T. brucei* (Ochsenreiter et al., 2007; Koslowsky et al., 2014; Kirby et al., 2016), *T. vivax* (Greif et al., 2015) and *L. tarentolae* (Simpson et al., 2015). Different evolutionary hypotheses have been proposed to explain the origin and persistence of this expensive editing system, although no consensus has been achieved yet (Gray et al., 2010; Speijer, 2010; Flegontov et al., 2011).

Guide RNAs are coded in the minicircles. In *T. cruzi* there are about 20–30 thousand minicircles per parasite, 1.4 kb each one. Each minicircle contains four hypervariable regions (mHVRs) interspersed by four conserved regions located 90° apart each other. Every mHVR potentially encodes one gRNA (Degraeve et al., 1988; Velazquez et al., 2008).

Most of our knowledge on mitochondrial mRNA editing in trypanosomatids comes from studies in *T. brucei* and *L. tarentolae*, while little is known about editing in *T. cruzi* (Thomas et al., 2007). This parasite is a very diverse species and seven different lineages (also called Discrete Typing Units or

simply DTUs) have been described: the six main lineages TcI–TcVI (Zingales et al., 2009) plus TcBat (Marcili et al., 2009; Pinto et al., 2012; Lima et al., 2015). Strains belonging to the six main lineages of *T. cruzi* can be classified in three different mitochondrial clades according to their maxicircle sequences (Machado and Ayala, 2001; de Freitas et al., 2006; Westenberger et al., 2006; Ruvalcaba-Trejo and Sturm, 2011).

Recently, we deep-sequenced the mHVRs from 9 strains of this parasite at millions of paired-end reads (Rusman et al., 2019). The number of mHVR sequence clusters (according to pairwise identity) was quite different among most strains and the different DTUs shared few mHVR clusters. These results suggest that different DTUs, and even strains, may have different gRNA repertoires. In addition, such differences in the number of mHVR clusters may imply redundant gRNAs (gRNAs with relatively different sequences that edit the same region) or mHVRs not encoding gRNAs. We proposed that minicircle inheritance is biparental in hybrid DTUs (Rusman et al., 2019) and hybrid strains (Rusman et al., 2020) despite the uniparental inheritance of maxicircles (Tomasini, 2018). This scenario opens the question on how mixing minicircles from different parents could still lead to correct editing of mitochondrial mRNAs, since maxicircles are inherited from one parental. The mHVR sequences have previously been used to identify gRNAs and to build editing cascades by comparing them with edited mRNAs (Thomas et al., 2007; Simpson et al., 2015). Here, based on mHVR reads previously obtained, an algorithm was developed to identify and cluster gRNA sequences from big datasets. The aims of this study were to describe and compare editing cascades for representative strains of the *T. cruzi* diversity, to address differences in mRNA editing and to infer how gRNA repertoires evolve.

## MATERIAL AND METHODS

### mHVR Sequences

The strains analyzed in this study are listed in **Table 1**. The paired-end mHVR amplicon sequences were obtained by Rusman et al. (2019). Reads are available at the Sequence Read Archive (SRA) under the BioProject ID: PRJNA514922. Raw sequence reads for all samples were pre-processed. First, reads were quality filtered using the pair-end mode of Trimmomatic v0.36 (Bolger et al., 2014). Posteriorly, the preserved paired reads were merged into consensus sequences with their associated quality score and chimeras were removed using LeeHom software with default parameters (Renaud et al., 2014).

### Mitochondrial mRNA Estimation

Predicted edited mRNA sequences for strains Sylvio (mitochondrial clade A, TcI), CL Brener (mitochondrial clade B, TcVI) and Esmeraldo (mitochondrial clade C, TcII) (Ruvalcaba-Trejo and Sturm, 2011) were generated by manually inserting or deleting uridines (U's) into the unedited sequences of their genes as proposed by (Greif et al., 2015). Briefly, unedited sequences corresponding to Sylvio (FJ203996.1), CL Brener (DQ343645.1) and Esmeraldo (DQ343646.1) obtained from the

**TABLE 1** | Datasets used in this study.

Accession	Strain	DTU	Mitochondrial clade	Origin	Host
SRX5245771	PalDa20cl3	Tcl	A	El Palmar, Argentina	<i>Didelphis albiventris</i>
SRX5245770	TEV55cl1	Tcl	A	Tres Estacas, Argentina	<i>Triatoma infestans</i>
SRX5245773	Esmeraldo	Tcll	C	Sao Felipe, Brazil	<i>Homo sapiens</i>
SRX5245772	TU18cl93	Tcll	C	Potosí, Bolivia	<i>Triatoma infestans</i>
SRX5245767	X109/2	Tclll	B	Makthlawaiya, Paraguay	<i>Canis familiaris</i>
SRX5245766	CANllcl1	TclV	B	Belém, Brazil	<i>Homo sapiens</i>
SRX5245769	MNcl2	TcV	B	Región IV, Chile	<i>Homo Sapiens</i>
SRX5245768	LL014R1	TcV	B	Las Leonas, Argentina	<i>Triatoma infestans</i>
SRX5245774	L015P68R0cl4	TcVI	B	Las Leonas, Argentina	<i>Canis familiaris</i>

NCBI database were aligned using MEGA v7 software (Kumar et al., 2016). Then, the alignment was manually manipulated by inserting or deleting uridines (U's) at specific sites into the unedited sequences following the editing patterns of the corresponding *T. brucei* and *T. vivax* edited mRNAs, meanwhile the resulting amino acid sequence was preserved. Predicted sequences were generated for *mitochondrial unidentified reading frame 2* (MURF2), *NADH dehydrogenase subunit 3* (ND3) and *C-rich region 4* (CR4) following the known edited mRNAs for *T. vivax* and *T. brucei* (**Supplementary File 1**). On the other hand, sequence data of predicted edited mRNAs for the following genes were obtained from Ruvalcaba-Trejo et al. (Ruvalcaba-Trejo and Sturm, 2011) for the same three *T. cruzi* strains: *ATPase6*, *C-rich region 3* (CR3), *cytochrome b* (CyB), *ribosomal protein S12* (RPS12), *NADH dehydrogenase subunit 7* (ND7), *NADH dehydrogenase subunit 8* (ND8) and *NADH dehydrogenase subunit 9* (ND9). CL Brener *cytochrome oxidase subunit 3* (COIII) mRNA was downloaded from GenBank (accession number: EF058194.1). In addition, Sylvio and Esmeraldo COIII edited mRNA sequences were manually generated (**Supplementary File 1**). Additional maxicircle sequences described by Reis-Cunha et al. 2018) were downloaded from GenBank, manually edited, and also used for gRNA search. Predicting gRNAs based on such sequences (shown in **Table S1**) had similar results to the obtained by using Sylvio, Esmeraldo or CL-Brener edited mRNAs.

## gRNA Detection

An algorithm using c++ and the SeqAn library (Reinert et al., 2017) was built to detect potential gRNA in the mHVR reads. First, a hash-table was built from the reads by storing position of all the possible k-mers (seeds) of 15 bases. However, due to non-canonical base pair a hash-table was used considering just purines (R) and pyrimidines (Y) instead the four bases. Although C::A pairing is allowed with this approach, it is considered a mismatch in the following steps. An algorithm scheme is provided in **Figure S1**. The hash table only allows to reduce the search time in the next steps. The mRNA sequence for each edited gene was used to identify gRNAs. The estimated edited mRNA of the corresponding mitochondrial clade was used to infer gRNA for each strain. First, a window of 15 bases was selected at the 5' end in the mRNA and the reverse complement of the purine/pyrimidine pattern was used to determine which mHVRs contains a potential gRNA by using the hash-table. Then, the potential site was evaluated considering

canonical and non-canonical base pairing (G::U) with the mRNA and allowing only one mismatch. Finally, the pairing is extended to both sides using an x-drop extension allowing one mismatch. Mismatches were allowed considering some flexibility in editing and considering that mRNAs are estimated and minimal errors in estimations are possible (Thomas et al., 2007). A filter was then applied to discard the predicted gRNA if it has fewer bases than the parameter *f*. This parameter is the minimum number of bases that a gRNA candidate must have to be accepted. This parameter is required because short sequences with complementarity to mRNAs are expected by simple random. Then, the window is moved over the mRNA one base at a time and the process is repeated until the entire mRNA is scanned. Different values for *f* were evaluated in order to discard sequences that are not gRNAs. In a first approach, gRNAs with <40 bases were discarded (*f* = 40). This filter was too stringent but allowed to identify the most probable regions of the mHVR that contain the gRNA, that is the gRNA cassette. Then, the filter was lowered looking for hits outside the cassette and the minimum filter that had as maximum as 5% of the gRNA hits out of the cassette was selected (*f* = 30) (**Figure S3**). Alternatively, other approaches for gRNA detection were evaluated. The first one was based on (Koslowsky et al., 2014). Briefly, after looking for 15 bp seeds, the algorithm scored canonical (score = 2) and non-canonical (score = 1) base pairings and gRNAs with scores <45 were discarded. The second approach was based on (Cooper et al., 2019) by looking for gRNAs with 6 bp-anchors (canonical base pairing) in the 5' end and then discarding gRNAs with less than 25 bp (canonical plus non-canonical base pairing). Both approaches were too stringent, they recovered reduced gRNA coverage in ATPase 6 mRNA and a smaller number of gRNAs than using the above approach *f* = 40 (**Figure S1**). Consequently, *f* = 30 was established given its higher sensitivity than *f* = 40, and a reasonable rate of gRNA hits out of the cassette. The c++ source code for gRNA detection in mHVRs and example datasets are stored in <https://osf.io/kn34z/>.

## gRNA Clustering Analyses

Identified gRNAs were clustered according to sequence similarity and overlapping in the edited-mRNA target region. A SWARM-like algorithm (Mahe et al., 2014) was implemented to cluster gRNAs. SWARM is a fast unsupervised (*de novo*) single-linkage-clustering algorithm. An algorithm scheme is provided in **Figure S2**. The modified algorithm also included an overlapping threshold (*o*) in addition to a threshold of the

maximum number of base mismatches between sequences ( $d$ ). Basically, the algorithm selects one gRNA as a cluster seed and then it iteratively looks for gRNAs with a number of mismatches lower than  $d$  and an overlapping  $> o$  and adds them to the cluster. Then, it iteratively looks for unclustered gRNAs with less than  $d$  mismatches and overlapping  $> o$  to every gRNA previously included in the cluster. When no new gRNAs can be added to the cluster a new cluster is seeded using a gRNA that was not previously clustered. The process is iterated until the all gRNAs were joined to a cluster.  $o = 0.7$  and  $d = 4$  were implemented. This is a relaxed threshold to cluster moderately different gRNAs in the same cluster considering divergency in such kind of sequences. Considering the overlapping region of two gRNAs of ~30-50bp,  $d = 4$  implies 85-92% pairwise similarity. Similar percentages were used to cluster mHVRs in (Rusman et al., 2019). The second phase of breakage in the original SWARM algorithm was not required (Mahe et al., 2014) because of the high distances and reduced overlapping among clusters. Identified gRNA clusters based on SWARM were termed as gRNA classes here. Finally, a consensus gRNA was built for every gRNA class to represent editing cascades. In order to fast calculation, the consensus gRNA was built using a base-by-base majority-rule criterium (i.e. the most abundant base in each position of gRNA alignment was used to build the consensus gRNA). Because artifactual gRNA may be generated by this method the algorithm also calculated the number of gRNAs sharing such sequence and informed it in the output file to manual inspection. The consensus gRNAs were graphed aligned to the edited mRNA and the cluster abundance was represented with a color scale.

In order to assess linkage disequilibrium as a measure of genetic structuring among strains, the index of association ( $I_a$ ) and the proportion of compatible pairs of loci with strict clonality (PrCompat) were calculated. Briefly, gRNA classes were coded as present ( $>20$  reads) or absent ( $\geq 20$  reads) emulating RFLP pattern data for each strain and the coded matrix was analyzed as haploid data in Mulilocus v1.3 (Agapow and Burt, 2001) with 10,000 iterations to calculate statistical significance.

A subset of gRNA classes was selected to predict secondary structure. Briefly, gRNA classes with more than 200 reads (high abundance), with more than 90% of the reads being identical to the consensus sequence of the class and with at least 40 bases were selected. The selected class representative gRNAs were analyzed using RNAstructure software (Reuter and Mathews, 2010). The maximum accuracy structures were calculated using default parameters but considering a temperature of 27°C (optimal to temperature for epimastigotes and previously used in *T. brucei* analysis of gRNA secondary structure) (Schmid et al., 1995).

## mHVR Clustering and Comparison With gRNA Classes

mHVR reads were processed and clustered according to (Rusman et al., 2019) using a UCLUST algorithm with a threshold of 85% of identity for cluster definition. Clustering at

85% was previously used to define mHVR clusters (Rusman et al., 2019) and such clusters were used to determine relationships between mHVR clusters and gRNA classes. Briefly, a c++ algorithm was designed to add a string with the mHVR cluster ID to each read name in the raw fastq file. Then, this fastq file was used to determine gRNAs using the above-described algorithm. This way, each identified gRNA could be assigned to an mHVR cluster. After clustering gRNAs in classes using the SWARM algorithm, it was evaluated whether each mHVR cluster coded none, one or more than one gRNA. Around 9.5% of the clusters corresponded to two gRNA classes in the same mHVR cluster. Although it is expected none or just one gRNA in an mHVR, such possible method inaccuracy was allowed in order to gain sensitivity to obtain the maximum coverage on mRNA editing cascades.

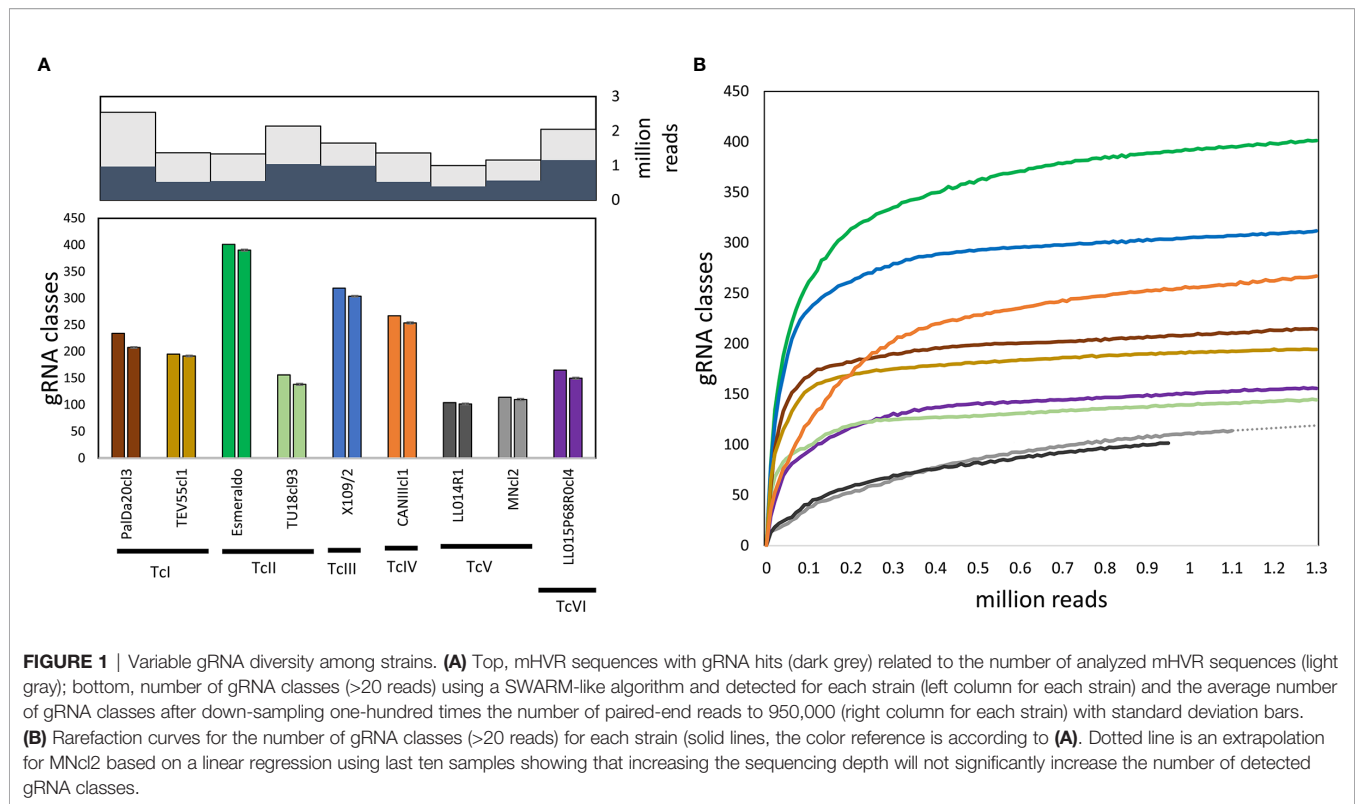
## RESULTS

### Fast Evolutionary Changes in gRNA Repertoires

Guide RNAs coded in mHVRs were inferred based on the available sequences or predicted edited mRNAs of *T. cruzi*. A total of 7,476,003 gRNAs were detected for nine *T. cruzi* strains belonging to the six DTUs. The percentage of detected gRNAs from the total number of mHVR reads for each studied strain varied from 38% to 60% (Figure 1A), suggesting that some mHVRs would not encode gRNAs. Posteriorly, the gRNAs were clustered in gRNA classes according to the editing region and their sequence similarity (Supplementary File 2). A total of 1334 gRNA classes were identified (every class with  $>20$  reads). The number of gRNA classes *per* strain was variable ranging from 104 (LL014R1-TcV) to 401 (Esmeraldo-TcII). Particularly, TcV strains had up to three or four times fewer gRNA classes than other DTUs (Figure 1). The number of gRNA classes was also variable within DTUs (Figure 1 and Figure 2) suggesting fast changes in gRNA repertoires and probably gRNA redundancy (i.e., different gRNA classes editing the same mRNA region). Rarefactions of each dataset showed that these differences between strains are not the effect of different sequencing depths (Figure 1B).

Just 22.8% of the whole identified gRNA classes in all strains were shared among two or more of them and only four gRNA classes were shared among all strains (sequences provided in Supplementary File 2) revealing high divergence. However, the percentage of shared gRNA classes between strains of the same DTU was variable (Figure 2). TcI strains shared a relatively low number of gRNA classes (19.2%) despite that the studied strains are very close phylogenetically (according to nuclear genes) (Schmid et al., 1995) and they were isolated in the same geographical area and time (Diosque et al., 2003). In addition, TcII strains shared even few gRNA classes (9.1%) (Figure 2). Instead, TcV strains (LL014R1 and MNcl2) shared a high proportion of gRNA classes (73.8%) despite they were isolated in geographically distant places and with a difference of around 30 years between the isolation dates. All these results suggest





significant changes in gRNA repertoires in relatively short evolutionary times. In addition, linkage disequilibrium was analyzed to address the structuring of such diversity. The index of association ( $I_a$ ) was 49.3 and showed high significance ( $p < 0.001$ ). Furthermore, the ratio of pairs of gRNA classes compatible (PrCompat) with unmixed gRNA classes was 0.98 suggesting structuring of diversity among strains. Finally, it was addressed whether probable predicted secondary structures are common in such gRNA sequence diversity. A subset of 89 gRNA classes defined by a stringent criterion (see Material and methods) was studied. Although variable structures were observed (**Supplementary File 4**), 70.7% of the gRNAs have a stem region intercalated by one or two internal loops (or less frequently bulges) and ending in a hairpin loop (**Figure S4**). Only 26% of the structures have a double hairpin (**Figure S4**).

## Silent mHVRs

In a previous work, entire mHVR sequences were clustered according to sequence identity (Rusman et al., 2019). Here, the gRNA classes were assigned to mHVR clusters. First, most (97.1%) mHVR clusters of all strains had one of two states: (i) high proportion (> 75%) of the sequences in the mHVR cluster coding one gRNA class; or (ii) low proportion (< 10%) of sequences in the mHVR cluster coding a gRNA class (**Table S2**). Most exceptions to the rule were observed in TcV strains (**Table S1**). The term “silent” mHVR clusters was used for those in which less than 20 gRNAs were identified. Silent mHVR clusters were from 25% to 54% of the total mHVR clusters in the studied strains (**Table 2**). Curiously, strains with the lowest

number of gRNA classes (TcV and TcVI strains) also had the lowest percentage of silent mHVR clusters (**Table 2**).

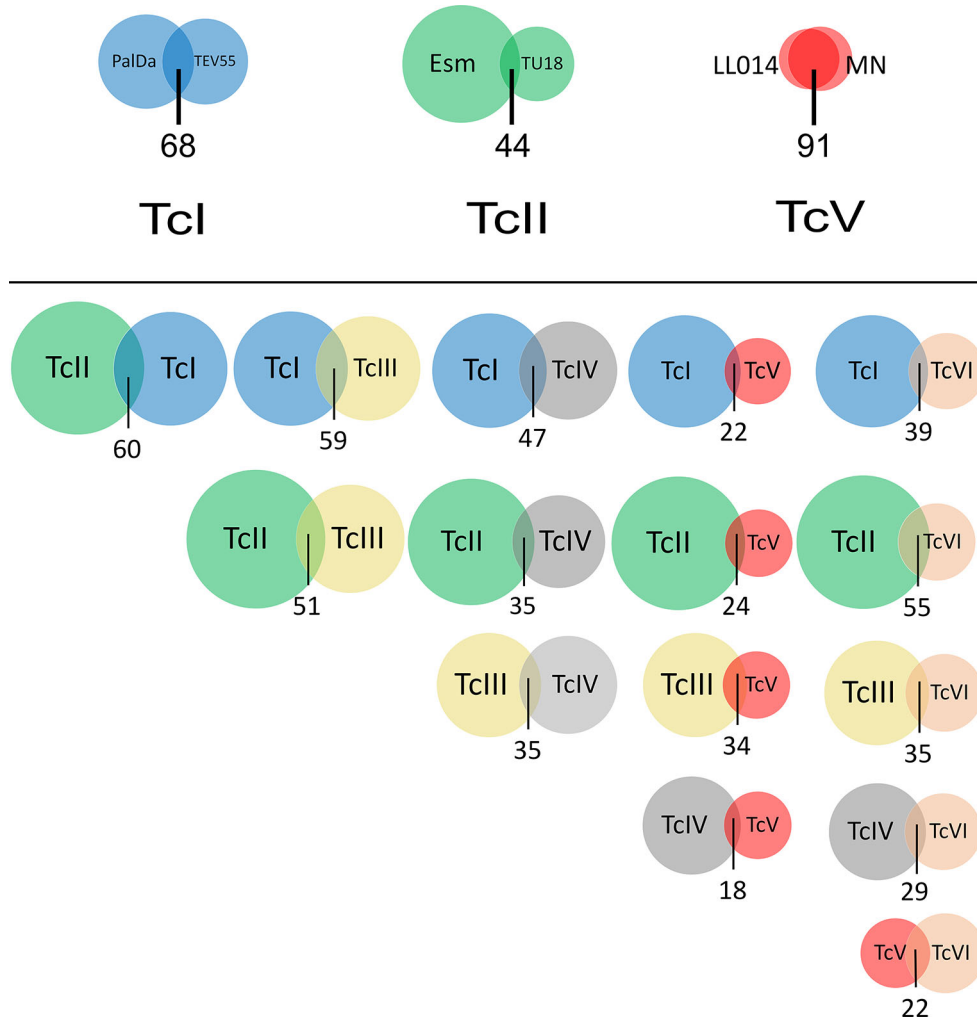
Furthermore, the analyses showed there is no clear relationship between mHVR cluster size and whether it encodes a gRNA or not since silent mHVR clusters were of quite different abundances (**Figure 3**).

## Concerted Changes in the Number of gRNA Classes

The gRNA classes identified for each strain were aligned with the edited mRNAs of the corresponding mitochondrial clade. The number of gRNA classes editing each gene for all strains is represented in **Figure 4** and **Table S3**. Most gRNA classes (more than 50%) edit ATPase 6 and COIII mRNAs in all strains (**Figure 4A**). Interestingly, the percentage of gRNA classes editing each mRNA was nearly conserved among strains (**Figure 4B**). This implies correlated variations in the number of gRNAs classes editing each mRNA (see correlation coefficients in **Table S4**). These results show that observed variations in the number of gRNA classes are balanced, i.e., if one strain decreased gRNA classes for one mRNA also decreased the number of gRNAs classes for editing other mRNAs. This would be an expected result whether each minicircle codes for different gRNAs, each one editing different mRNAs.

## Editing mRNA Cascades for ATPase6 and COIII

The algorithms identified a complete or almost complete set of gRNA classes necessary to edit the mRNAs for the ATPase6



**FIGURE 2** | Venn diagrams showing overlapping gRNA repertoires between strains of the same DTU and between DTUs. Circle size represents the number of gRNA classes identified in a strain (above) or in a DTU (below). The overlapping area among circles indicates gRNA classes with more than 20 reads that are shared between strains or DTUs. The number below circles represents the percentage of shared gRNA classes. PalDa=PalDa20c3, TEV=TEV55c1, Esm=Esmeraldo, X109=X109/2, Tu18=Tu18c93, MN=MNc12 and LL014=LL014R1.

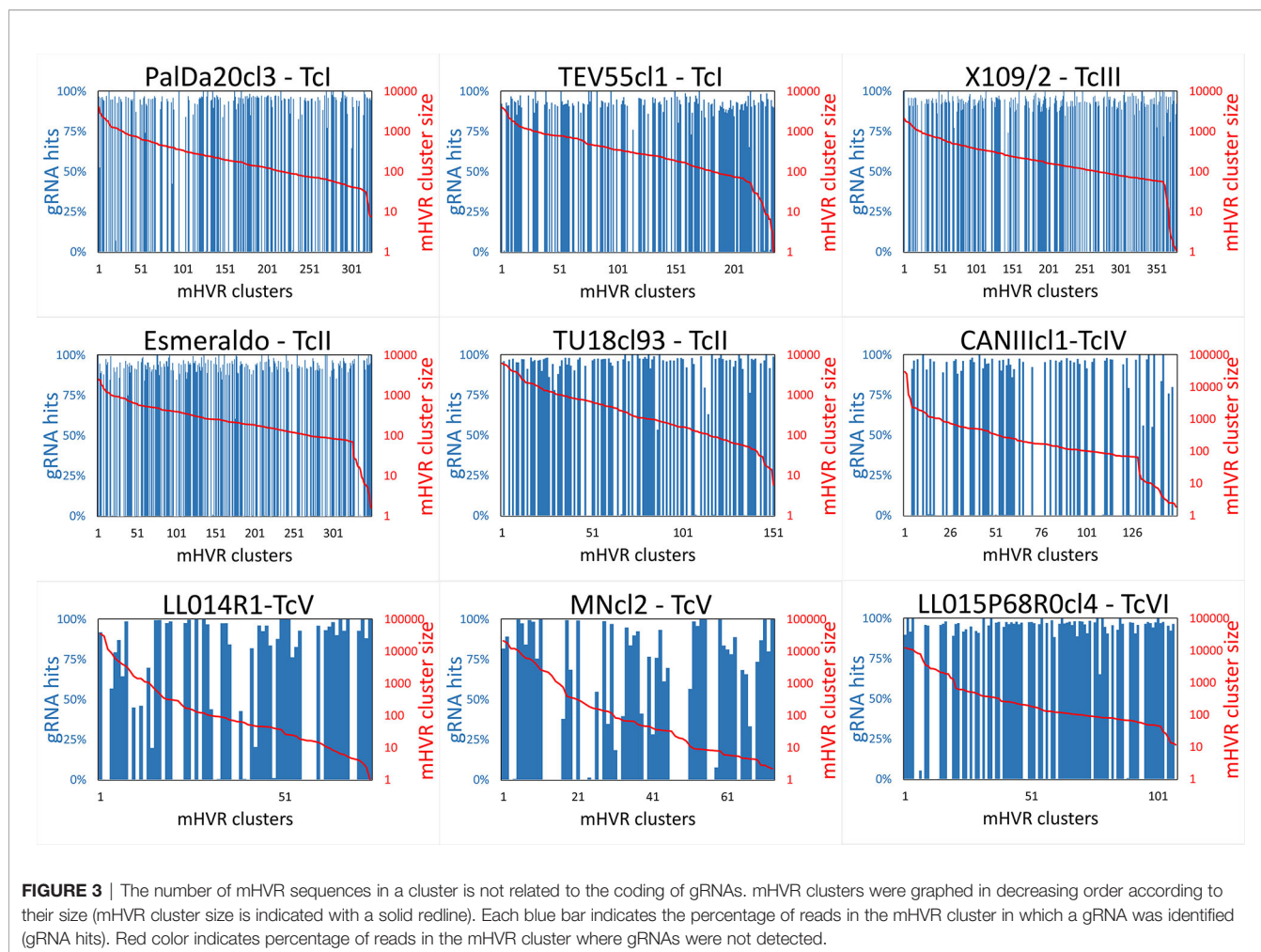
(complex V) and COIII (complex IV) in every strain despite the highly variable number of gRNA classes editing such mRNAs. This clearly shows redundant gRNA classes for some strains. The editing mRNA cascades for ATPase6 and COIII pan-edited genes are shown in **Figure 5** and **Figure S5**, respectively.

Aligned gRNAs-mRNAs are provided in the **Supplementary File 2**. Strains PalDa20c3, Esmeraldo, X109/2 and CANIIIc1 showed a greater redundancy of gRNA classes than other strains which implied strong variations among and within some DTUs, mainly TcII strains (**Figure 5** and **Figure S5**). On the other side,

**TABLE 2** | Relationship between mHVR clusters and gRNA classes.

	TcI		TcII		TcIII	TcIV	TcV		TcVI
	PalDa20c3	TEV55c1	Esmeraldo	TU18c93	X109/2	CANIIIc1	LL014R1	MNc12	LL015P68R0c4
mHVR clusters <sup>1</sup>	324	234	347	151	373	149	72	71	108
Silent mHVR clusters <sup>2</sup>	47%	39%	42%	37%	43%	54%	32%	34%	25%

<sup>1</sup>mHVR clusters were defined in Rusman et al. (2019).  
<sup>2</sup>mHVR clusters for which less than 20 gRNAs were detected.



TcV strains - MNc12 and LL014R1 - had the lowest redundancy of gRNA classes (Figure 5 and Figure S5).

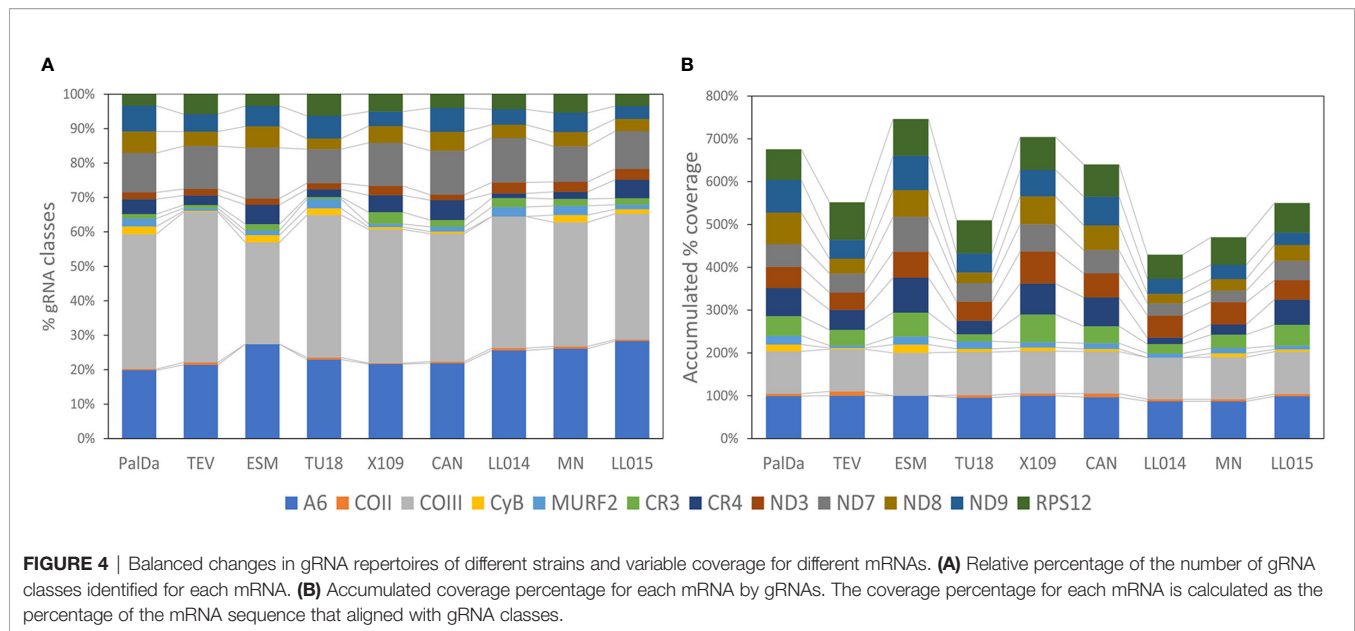
## Respiratory Complex I: Incomplete Editing Cascades

The percentage of coverage by gRNAs for each mRNA is shown in Figure 4B and Table S5. Interestingly, the coverage was almost complete for ATPase6 and COIII mRNAs but highly variable for the remaining mRNAs, especially in the complex I members (Figure 4B). The editing cascades of the mRNAs coding complex I members —ND9, CR3, CR4, ND3, ND7, ND8— of the respiratory chain are shown in Figure 6 and Figures S6–S10. Even strains showing high mRNA class redundancy in COIII and ATPase6, such as Esmeraldo and X109/1, have incomplete cascades for complex I members. The exception was ND9 mRNA, in which both strains (X109/2 and Esmeraldo) presented almost a complete coverage (Table S5), although most of the gRNA classes (82.1% and 74.8%, respectively) had less than 20 reads (see also Table S6). Breaking points in ND3, ND7 and ND8 editing cascades in these two strains were also observed. Interestingly, TcV strains —LL014R1 and MNc12— presented

relatively few gRNA classes that edit the complex I components messengers (Table S3, Figure 6 and Figures S6–S10) and relatively low coverage of the mRNAs (Table S5 and Table S6).

## Potential Cross-Editing of Mitochondrial mRNAs Among Different DTUs

Biparental inheritance of minicircles in hybrid DTUs was previously proposed (Rusman et al., 2019). Here, it was analyzed whether gRNAs from one strain may potentially edit mRNAs from a strain from another DTU, particularly another mitochondrial clade. This ability for “cross-editing” is a requirement in hybrid DTUs, where minicircles come from both parentals but maxicircles come from just one parent. Editing cascades were again generated for Esmeraldo gRNAs and X109/2 gRNAs but based on mRNAs of different mitochondrial clades. mRNA coverage by gRNAs was evaluated and both gRNA repertoires have similar coverages on mRNAs from different mitochondrial clades (Figure 7). Cross-editing was also evaluated in more stringent conditions by lowering the allowed number of mismatches in cluster definition ( $d = 3$ ,  $d = 2$  and  $d = 1$ ) and it was not detected a significant difference in mRNA coverage percentage. Minor differences were observed in the mRNA



coverage percentages for genes *CyB* and *ND9* in  $d = 1$  settings (**Figure S11**). These results suggest that biparental inheritance of the kDNA would not generate mRNA editing problems even among phylogenetically distant DTUs.

## DISCUSSION

mRNA editing in kinetoplasts is an intriguing evolutionary mechanism. Questions about how and why such complex and energy-expensive mechanism has evolved are still controversial (Covello and Gray, 1993; Gray et al., 2010; Flegontov et al., 2011). The elucidation of editing cascades in mRNA is a prerequisite to start filling the gaps around such questions. We have previously reported the DNA sequence diversity of the hypervariable region of kinetoplastid minicircles (mHVR) in the main lineages of *T. cruzi*, which were virtually unknown until then (Rusman et al., 2019). Here, we deeply addressed gRNA repertoires (coded in the mHVRs) of nine *T. cruzi* strains belonging to the main lineages of this parasite.

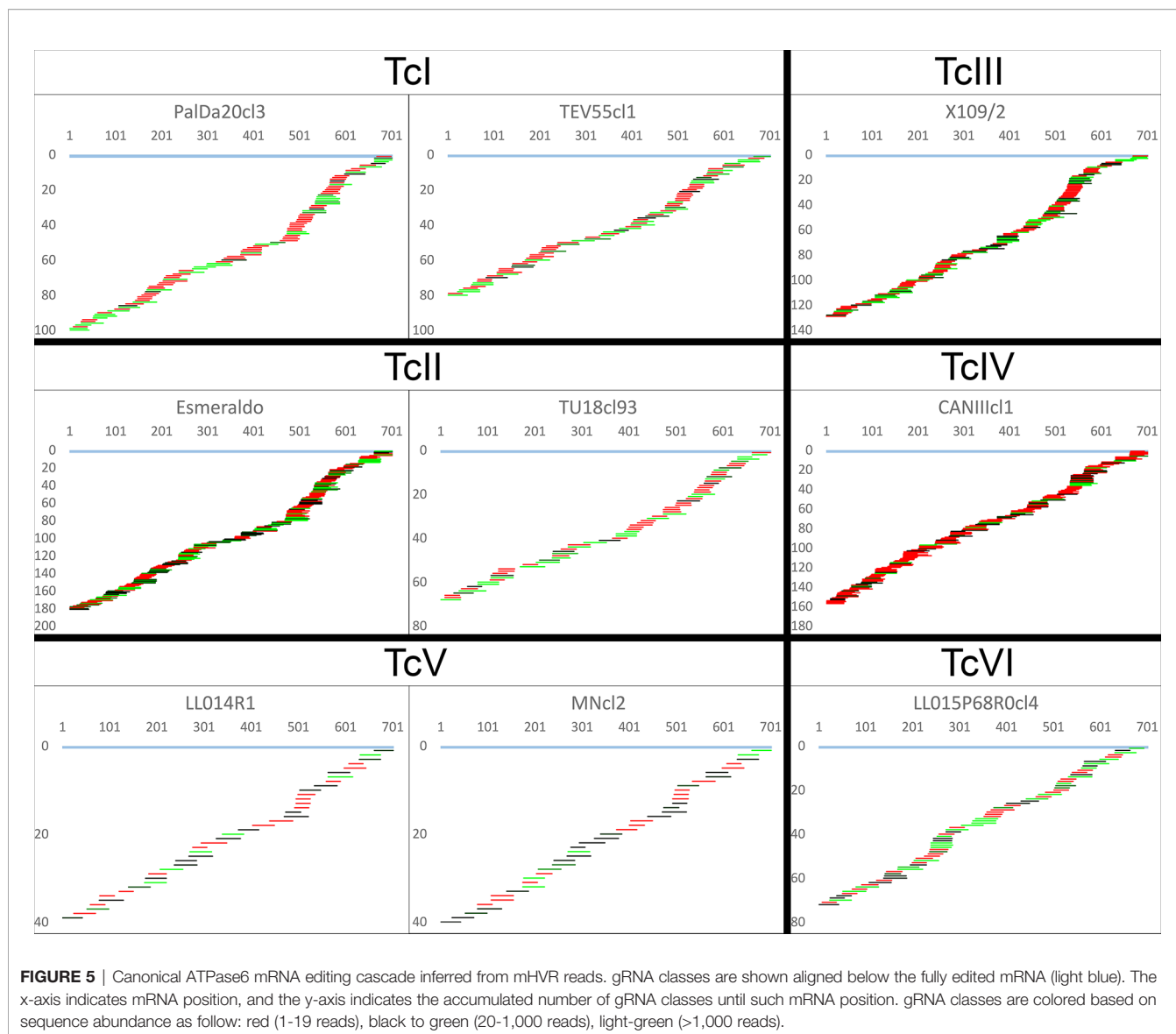
Our results showed that gRNA diversity was enormous, with more than 1,300 different gRNA classes, each one detected in >20 reads. Such diversity was structured among the strains according to linkage disequilibrium indexes. However, further analyses are required to address how good the linkage disequilibrium measures perform on this kind of data. In addition, the number of gRNA classes was highly variable among strains. In addition, it was addressed whether shared common secondary structures exist despite the high sequence diversity for gRNAs, as it has been described for *T. brucei* (Schmid et al., 1995). Double-hairpin structures observed in *T. brucei* were not the most frequent structures in *T. cruzi* predictions. However, the predictions should be considered cautiously because they do not include the poly-U commonly added to 3' of gRNAs (Schmid et al., 1995) and

predictions could change if constraints based on experimental data are added to the analysis (Reuter and Mathews, 2010). Finally, despite this study analyzed the mitochondrial editomes of different strains in a comparative and large scale, experimental studies are needed to determine gRNA complete sequences and the abundances of each one. Nevertheless, studies in *T. vivax* and *T. brucei* have revealed that gRNA abundance is related to the coding minicircle abundance (Koslowsky et al., 2014; Greif et al., 2015) although such correlation is not so clear in *Leishmania tarentolae* (Simpson et al., 2015).

As expected from mHVR sequence analyses (Telleria et al., 2006; Rusman et al., 2019), different DTUs shared a relatively low number of gRNA classes, but it was surprising that even strains of the same DTU (here TcI and TcII) shared few gRNA classes. Such strains showed not only different gRNA classes but also very divergent abundances in shared gRNA classes.

The TcI strains analyzed here are phylogenetically close according to nuclear and mitochondrial sequences (Tomasini et al., 2014; Rusman et al., 2020). However, they shared just 19.2% of their gRNA classes. Such results suggest that gRNA repertoires may vary extremely fast. Even TcV strains which are identically at most nuclear analyzed DNA sequences have a poor correlation in the abundances of shared gRNAs, even though TcV strains share most of their mHVR clusters at 95% identity threshold (Rusman et al., 2019) and most of their gRNA classes. Such results suggest that fast changes in minicircle repertoires are not mainly caused by mutation. It is important to mention that the variations in gRNA classes between strains described above would not be a result of minicircle loss after long periods in culture. The strain with more gRNA diversity (Esmeraldo, TcII) was isolated in the eighties and maintained in culture for several years; whereas, for example, LL015P68R0c14 (TcVI) was isolated ten years ago, cultured for less than a year and then frozen until this study. In addition, LL014R1 (TcV) which showed the lowest



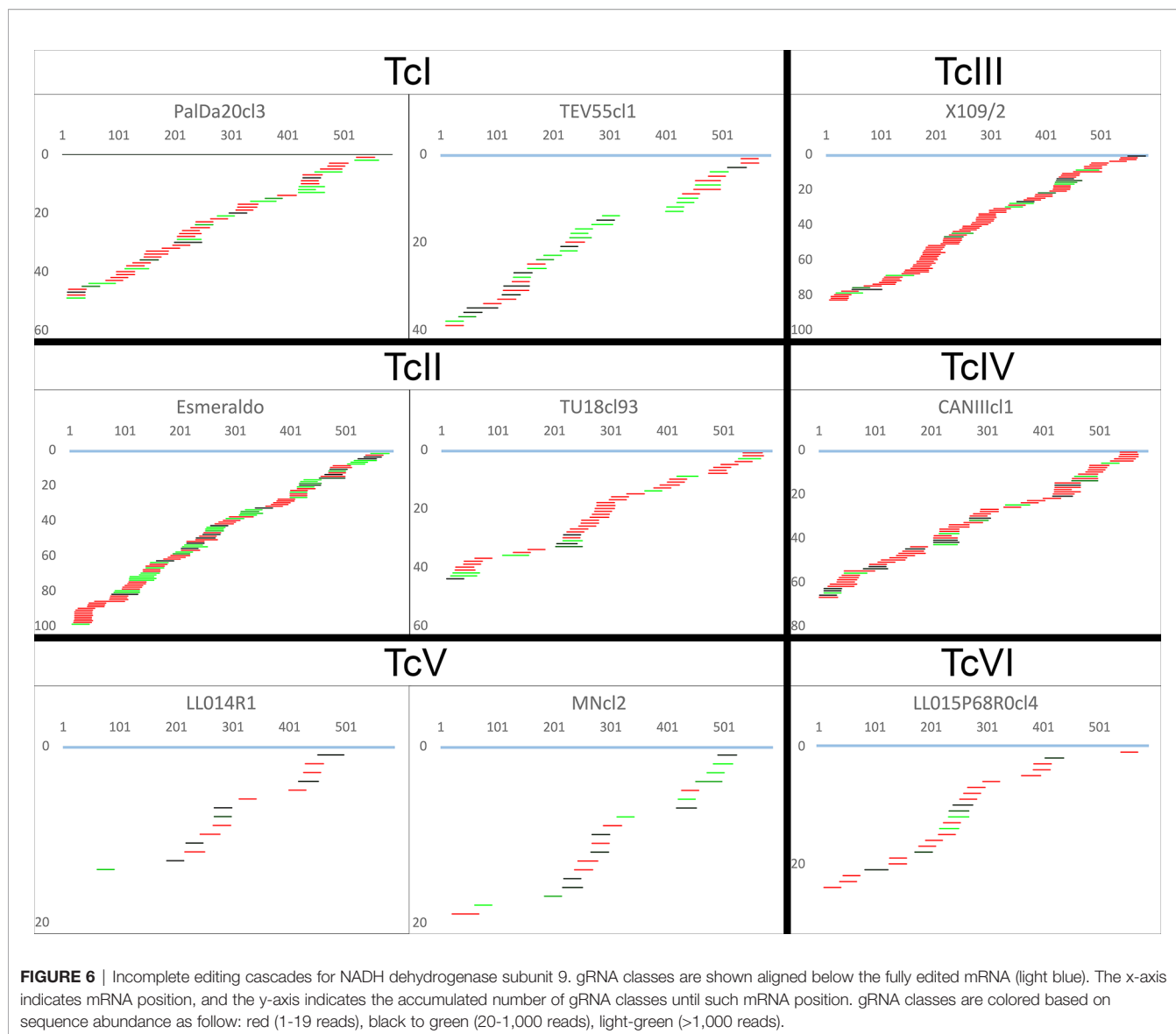


number of gRNA classes, was maintained mostly through mouse-triatomine passages since isolation in 2009. In addition, such variations in gRNA class diversity between strains correlate well with mHVR cluster diversity. This result supports that changes in gRNA diversity among strains is not caused by differences in gRNA estimation for different DTUs (caused by edited mRNA prediction).

The models of random or partially random segregation of minicircles (minicircle drift) during kDNA division predicts changes in gRNA composition by variations in gRNA class abundance (Savill and Higgs, 1999). “Minicircle drift” is caused by an error during kDNA network division resulting in the two copies of a replicated minicircle being inherited to the same descendant kinetoplast. Such random changes are analogous to genetic drift in populations (e.g., allele frequency varies randomly until fixation or lost, reducing allele diversity in absence of any other

evolutionary force). In the same way, partially random segregation of minicircles predicts loss of gRNA redundancy (Savill and Higgs, 1999). Linkage of gRNA classes (that is, their presence in the same minicircle) may conserve some redundancy. This would occur in the case that redundant gRNA classes were linked with different essential gRNAs (i.e., gRNA classes without redundancy). However, gRNA linkage is still not compatible with spread redundancy observed in some strains because it is expected that the greater the number of redundant classes the greater the number of essential gRNA classes necessary to support them, which strongly limits redundancy.

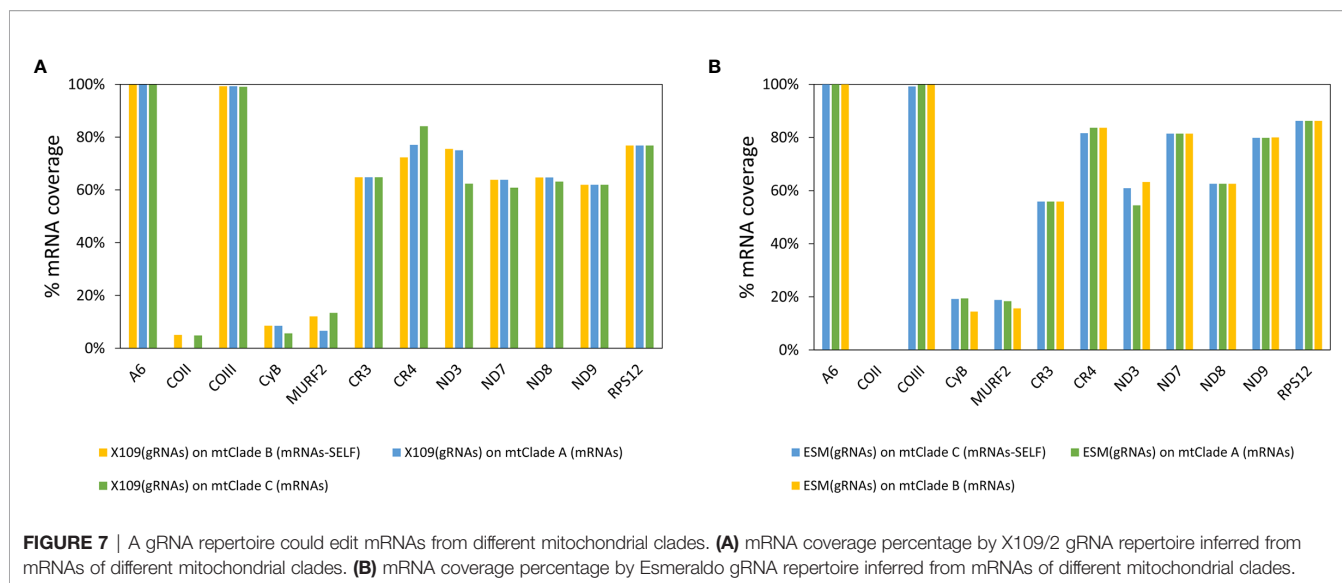
Since the model of (partial) random segregation as the unique evolutionary force does not fit the high levels of gRNA class redundancy observed in some strains, the genetic exchange would be the most probable force generating gRNA class redundancy. In a previous paper, we proposed biparental



inheritance of minicircles in hybrid DTUs of *T. cruzi* (Rusman et al., 2019) and in an intra TcI hybrid (Rusman et al., 2020). Biparental inheritance has also been proposed for *T. brucei* (Turner et al., 1995; Gibson et al., 2008). Maxicircles and minicircles may be biparentally inherited but maxicircles of one parent are lost by drift in few generations (Turner et al., 1995). Despite, minicircles of both parents are maintained by many more generations. In this sense, genetic exchange may generate gRNA redundancy (Savill and Higgs, 1999). In addition, the observed correlated variations in the number of gRNAs classes editing each mRNA (see **Tables 2** and **S4**) may also be explained by a biparental inheritance of minicircles which is a bulky change in gRNA classes. Conversely, clonal propagation would generate a loss of gRNA classes by minicircle drift. Under this hypothetical dynamic, DTUs with less gRNA redundancy (i.e. TcV and TcVI) should have a typical clonal propagation (Lewis et al., 2011;

Diosque et al., 2014; Tomasini and Diosque, 2015); while, most gRNA class redundancy has been observed in lineages in which genetic exchange has been reported (TcI, TcII and TcIII) (Llewellyn et al., 2009; Ocana-Mayorga et al., 2010; Baptista Rde et al., 2014; Tomasini and Diosque, 2015; Berry et al., 2019).

Biparental inheritance of minicircles has another challenge, i.e., to make the right editing of mRNA when maxicircles come from just one parent. However, *in silico* analyses showed that gRNA classes from one strain may edit mRNAs from the other mitochondrial clades in a similar way than in the self-clade. This result suggests that genetic exchange and biparental inheritance of minicircles should not cause editing conflicts. In this sense, genetic exchange may increase gRNA class redundancy, avoiding that an editing site being near to lost gRNAs, even restoring broken editing cascades. Consequently, strict clonal lineages would not persist in long evolutionary times (a Muller ratchet).



Interestingly, gRNA class redundancy was mainly detected in ATPase subunit 6 and Cytochrome c oxidase subunit III. Editing cascades of both mRNAs were complete or almost complete in all strains suggesting that such genes are essential for the parasite. However, editing cascades of the complex I subunits have less gRNA class redundancy and even interrupted cascades at several positions in some strains. Functions of the complex I are debatable in trypanosomatids (Opperdoes and Michels, 2008) and it has been proposed that electrons bypass this complex in the respiratory chain of *T. cruzi* epimastigotes (Denicola-Seoane et al., 1992; Carranza et al., 2009). In addition, the absence of two key subunits suggests that if functional for electron transport, the complex I cannot pump protons to intermembranous space and may function to renewing NAD<sup>+</sup> (Opperdoes and Michels, 2008) although other studies showed that this activity has no differences between ND mutants and wild type strains in *T. cruzi* epimastigotes (Carranza et al., 2009) although only log-phase epimastigotes were evaluated. Considering the fast evolutionary rates shown for gRNA class frequencies, a completely non-functional complex I should imply highly impaired cascades of gRNA elements editing such sequences. Instead, almost complete editing cascades recovered for some complex I elements do not fit with the hypothesis of non-functional complex I at least in some strains like Esmeraldo or X109 (Figure 6). Despite, it is a more suitable hypothesis that time in culture without infecting a mammal allowed that some gRNA classes were lost, or they reduced their abundance as observed. However, it cannot be discarded that some DTUs like TcV may have lost complex I components since the cascades are highly incomplete in both strains in a similar manner. Further studies on this topic are required, especially because deletions in genes coding the complex I subunits were associated with the indeterminate form of the Chagas disease in TcII (Baptista et al., 2006). Nevertheless, our results support that complex I is not required in all stages of the parasite. In this sense, the linkage of different gRNA classes editing different mRNAs may help to protect

against gRNA class loss in stages where the complex I is not required.

Another interesting observation was the high frequency of mHVR clusters without gRNA hits. Such “silent” mHVR has been observed in *L. tarentolae* (Simpson et al., 2015) and *T. brucei* (Hajduk et al., 1997) although in minor proportions. There are several potential non-exclusive explanations. Some mHVR clusters may code for gRNAs involved in alternative editing of the mRNAs (which cannot be detected by our current algorithm), or code gRNAs that accumulated mutations or even code for shorter gRNAs (less than 30 bp). Alternatively, many mHVRs may not code for gRNAs and such mHVRs persist by linkage to others that code for essential gRNAs (Savill and Higgs, 2000). However, it is not clear if such situations are enough to explain that around a half of the mHVR clusters did not code for gRNAs in some strains. Interestingly, TcV and TcVI strains, which have relatively less redundancy in gRNA classes, also have relatively less silent mHVR clusters, although the percentage is still not negligible. Consequently, unknown functions of these silent mHVR sequences it is a possibility that cannot be ruled out.

Finally, a model is proposed based on data presented here, from which hypotheses can be derived for further testing. Each minicircle in *T. cruzi* has mHVRs coding for different gRNAs that edit different regions and probably (by random) from different mRNAs (which is also supported by the correlated variations in the number of gRNAs classes editing each mRNA). Such linkage may reduce the chances of losing a gRNA class which is only required in some stages of the lifecycle. Such linkage may also explain the existence of silent mHVRs (although other functions cannot be discarded). In addition, minicircle drift cause loss of gRNA classes and low levels of gRNA class redundancy. Consequently, genetic exchange and biparental inheritance of minicircles may restore gRNA abundances for editing each mRNA site and redundancy which may reduce the chances of lethal loss of essential gRNA

classes. The conjunction of minicircle drift and occasional biparental inheritance of kDNA may explain the divergence of gRNA repertoires among strains even within the same DTU.

## DATA AVAILABILITY STATEMENT

The datasets analyzed for this study can be found in the Sequence Read Archive at NCBI with the following accession number PRJNA514922. The source code for gRNA inference is stored at the Open Science Framework <https://osf.io/kn34z/>.

## AUTHOR CONTRIBUTIONS

FR: Bioinformatic Analysis, Writing – original draft. NF-Y: Bioinformatic Analysis, Writing – Review and Editing. NT: Conceptualization, Programming, Bioinformatic analysis, Supervision, Writing – Review and Editing. PD: Conceptualization, Funding acquisition, Writing – Review and Editing. All authors contributed to the article and approved the submitted version.

## REFERENCES

- Agapow, P. M., and Burt, A. (2001). Indices of Multilocus Linkage Disequilibrium. *Mol. Ecol. Notes* 1, 101–102. doi: 10.1046/j.1471-8278.2000.00014.x
- Ammerman, M. L., Downey, K. M., Hashimi, H., Fisk, J. C., Tomasello, D. L., Faktorova, D., et al. (2012). Architecture of the Trypanosome RNA Editing Accessory Complex, MRB1. *Nucleic Acids Res.* 40 (12), 5637–5650. doi: 10.1093/nar/gks211
- Aphasizheva, I., Zhang, L., and Aphasizhev, R. (2016). Investigating RNA Editing Factors From Trypanosome Mitochondria. *Methods* 107, 23–33. doi: 10.1016/j.jmeth.2016.03.020
- Baptista Rde, P., D'Avila, D. A., Segatto, M., do Valle, I. F., Franco, G. R., Valadares, H. M., et al. (2014). Evidence of Substantial Recombination Among *Trypanosoma Cruzi* II Strains From Minas Gerais. *Infect. Genet. Evol.* 22, 183–191. doi: 10.1016/j.meegid.2013.11.021
- Baptista, C. S., Vencio, R. Z., Abdala, S., Carranza, J. C., Westenberger, S. J., Silva, M. N., et al. (2006). Differential Transcription Profiles in *Trypanosoma Cruzi* Associated With Clinical Forms of Chagas Disease: Maxicircle NADH Dehydrogenase Subunit 7 Gene Truncation in Asymptomatic Patient Isolates. *Mol. Biochem. Parasitol* 150 (2), 236–248. doi: 10.1016/j.molbiopara.2006.08.008
- Benne, R., Van den Burg, J., Brakenhoff, J. P., Sloof, P., Van Boom, J. H., and Tromp, M. C. (1986). Major Transcript of the Frameshifted coxII Gene From Trypanosome Mitochondria Contains Four Nucleotides That are Not Encoded in the DNA. *Cell* 46 (6), 819–826. doi: 10.1016/0092-8674(86)90063-2
- Berry, A. S. F., Salazar-Sanchez, R., Castillo-Neyra, R., Borrini-Mayori, K., Chipana-Ramos, C., Vargas-Maquera, M., et al. (2019). Sexual Reproduction in a Natural *Trypanosoma Cruzi* Population. *PLoS Negl. Trop. Dis.* 13 (5), e0007392. doi: 10.1371/journal.pntd.0007392
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* 30 (15), 2114–2120. doi: 10.1093/bioinformatics/btu170
- Carranza, J. C., Kowaltowski, A. J., Mendonca, M. A., de Oliveira, T. C., Gadelha, F. R., and Zingales, B. (2009). Mitochondrial Bioenergetics and Redox State are Unaltered in *Trypanosoma Cruzi* Isolates With Compromised Mitochondrial Complex I Subunit Genes. *J. Bioenerget Biomembranes* 41 (3), 299–308. doi: 10.1007/s10863-009-9228-4
- Cavalcanti, D. P., and de Souza, W. (2018). The Kinetoplast of Trypanosomatids: From Early Studies of Electron Microscopy to Recent Advances in Atomic Force Microscopy. *Scanning* 2018:9603051. doi: 10.1155/2018/9603051

## FUNDING

The current study is funded by Bunge and Born foundation and the National Scientific and Technical Research Council (D.2555/16.22920160100063CO) (CONICET, Argentina) to Patricio Diosque.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcimb.2021.663416/full#supplementary-material>

**Supplementary File 1** | Zip file containing fasta formatted files with mRNA sequences used for gRNA inference.

**Supplementary File 2** | Zip file containing gRNA classes aligned to different mRNAs for each strain in this study.

**Supplementary File 3** | **Tables S1–S6** and **Figures S1–S11**.

**Supplementary File 4** | Zip file containing predicted gRNA structures in postscript format and the gRNA sequences used for the analysis.

- Cooper, S., Wadsworth, E. S., Ochsenreiter, T., Ivens, A., Savill, N. J., and Schnauffer, A. (2019). Assembly and Annotation of the Mitochondrial Minicircle Genome of a Differentiation-Competent Strain of *Trypanosoma Brucei*. *Nucleic Acids Res.* 47 (21), 11304–11325. doi: 10.1093/nar/gkz928
- Covello, P. S., and Gray, M. W. (1993). On the Evolution of RNA Editing. *Trends Genet.* 9 (8), 265–268. doi: 10.1016/0168-9525(93)90011-6
- de Freitas, J. M., Augusto-Pinto, L., Pimenta, J. R., Bastos-Rodrigues, L., Goncalves, V. F., Teixeira, S. M., et al. (2006). Ancestral Genomes, Sex, and the Population Structure of *Trypanosoma Cruzi*. *PLoS Pathog.* 2 (3), e24. doi: 10.1371/journal.ppat.0020024
- Degrave, W., Fragoso, S. P., Britto, C., van Heuverswyn, H., Kidane, G. Z., Cardoso, M. A., et al. (1988). Peculiar Sequence Organization of Kinetoplast DNA Minicircles From *Trypanosoma Cruzi*. *Mol. Biochem. Parasitol* 27 (1), 63–70. doi: 10.1016/0166-6851(88)90025-4
- Denicola-Seoane, A., Rubbo, H., Prodanov, E., and Turrens, J. F. (1992). Succinate-Dependent Metabolism in *Trypanosoma Cruzi* Epimastigotes. *Mol. Biochem. Parasitol* 54 (1), 43–50. doi: 10.1016/0166-6851(92)90093-Y
- Diosque, P., Barnabe, C., Padilla, A. M., Marco, J. D., Cardozo, R. M., Cimino, R. O., et al. (2003). Multilocus Enzyme Electrophoresis Analysis of *Trypanosoma Cruzi* Isolates From a Geographically Restricted Endemic Area for Chagas' Disease in Argentina. *Int. J. Parasitol* 33 (10), 997–1003. doi: 10.1016/s0020-7519(03)00139-5
- Diosque, P., Tomasini, N., Lauthier, J. J., Messenger, L. A., Monje Rumi, M. M., Ragone, P. G., et al. (2014). Optimized Multilocus Sequence Typing (MlSt) Scheme for *Trypanosoma Cruzi*. *PLoS Negl. Trop. Dis.* 8 (8), e3117. doi: 10.1371/journal.pntd.0003117
- Estevez, A. M., and Simpson, L. (1999). Uridine Insertion/Deletion RNA Editing in Trypanosome Mitochondria—a Review. *Gene* 240 (2), 247–260. doi: 10.1016/s0378-1119(99)00437-0
- Feagin, J. E., Abraham, J. M., and Stuart, K. (1988). Extensive Editing of the Cytochrome C Oxidase III Transcript in *Trypanosoma Brucei*. *Cell* 53 (3), 413–422. doi: 10.1016/0092-8674(88)90161-4
- Flegontov, P., Gray, M. W., Burger, G., and Lukes, J. (2011). Gene Fragmentation: A Key to Mitochondrial Genome Evolution in Euglenozoa? *Curr. Genet.* 57 (4), 225–232. doi: 10.1007/s00294-011-0340-8
- Gibson, W., Peacock, L., Ferris, V., Williams, K., and Bailey, M. (2008). The Use of Yellow Fluorescent Hybrids to Indicate Mating in *Trypanosoma Brucei*. *Parasit Vectors* 1 (1):4. doi: 10.1186/1756-3305-1-4
- Gray, M. W., Lukes, J., Archibald, J. M., Keeling, P. J., and Doolittle, W. F. (2010). Cell Biology. Irremediable Complexity? *Science* 330 (6006), 920–921. doi: 10.1126/science.1198594



- Greif, G., Rodriguez, M., Reyna-Bello, A., Robello, C., and 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
- Hajduk, S. L., Adler, B., Madison-Antenucci, S., McManus, M., and Sabatini, R. (1997). Insertional and Deletional RNA Editing in Trypanosome Mitochondria. *Nucleic Acids Symposium Ser.* 36, 15–18.
- Hajduk, S., and Ochsenteiter, T. (2010). RNA Editing in Kinetoplastids. *RNA Biol.* 7 (2), 229–236. doi: 10.4161/rna.7.2.11393
- Kim, K. S., Teixeira, S. M., Kirchhoff, L. V., and Donelson, J. E. (1994). Transcription and Editing of Cytochrome Oxidase II RNAs in *Trypanosoma Cruzi*. *J. Biol. Chem.* 269 (2), 1206–1211. doi: 10.1016/S0021-9258(17)42243-5
- Kirby, L. E., Sun, Y., Judah, D., Nowak, S., and Koslowsky, D. (2016). Analysis of the *Trypanosoma Brucei* EATRO 164 Bloodstream Guide RNA Transcriptome. *PLoS Negl. Trop. Dis.* 10 (7), e0004793. doi: 10.1371/journal.pntd.0004793
- Koslowsky, D. J., Bhat, G. J., Read, L. K., and Stuart, K. (1991). Cycles of Progressive Realignment of gRNA With mRNA in RNA Editing. *Cell* 67 (3), 537–546. doi: 10.1016/0092-8674(91)90528-7
- Koslowsky, D., Sun, Y., Hindenach, J., Theisen, T., and Lucas, J. (2014). The Insect-Phase gRNA Transcriptome in *Trypanosoma Brucei*. *Nucleic Acids Res.* 42 (3), 1873–1886. doi: 10.1093/nar/gkt973
- Kumar, S., Stecher, G., and Tamura, K. (2016). Mega7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* 33 (7), 1870–1874. doi: 10.1093/molbev/msw054
- Lewis, M. D., Llewellyn, M. S., Yeo, M., Acosta, N., Gaunt, M. W., and Miles, M. A. (2011). Recent, Independent and Anthropogenic Origins of *Trypanosoma Cruzi* Hybrids. *PLoS Negl. Trop. Dis.* 5 (10), e1363. doi: 10.1371/journal.pntd.0001363
- Lima, L., Espinosa-Alvarez, O., Ortiz, P. A., Trejo-Varon, J. A., Carranza, J. C., Pinto, C. M., et al. (2015). Genetic Diversity of *Trypanosoma Cruzi* in Bats, and Multilocus Phylogenetic and Phylogeographical Analyses Supporting TcBat as an Independent DTU (Discrete Typing Unit). *Acta Trop.* 151, 166–177. doi: 10.1016/j.actatropica.2015.07.015
- Lin, R. H., Lai, D. H., Zheng, L. L., Wu, J., Lukes, J., Hide, G., et al. (2015). Analysis of the Mitochondrial Maxicircle of *Trypanosoma Lewisi*, a Neglected Human Pathogen. *Parasit Vectors* 8, 665. doi: 10.1186/s13071-015-1281-8
- Llewellyn, M. S., Lewis, M. D., Acosta, N., Yeo, M., Carrasco, H. J., Segovia, M., et al. (2009). *Trypanosoma Cruzi* IIC: Phylogenetic and Phylogeographic Insights From Sequence and Microsatellite Analysis and Potential Impact on Emergent Chagas Disease. *PLoS Negl. Trop. Dis.* 3 (9), e510. doi: 10.1371/journal.pntd.0000510
- Lukes, J., Hashimi, H., and Zikova, A. (2005). Unexplained Complexity of the Mitochondrial Genome and Transcriptome in Kinetoplastid Flagellates. *Curr. Genet.* 48 (5), 277–299. doi: 10.1007/s00294-005-0027-0
- Machado, C. A., and Ayala, F. J. (2001). Nucleotide Sequences Provide Evidence of Genetic Exchange Among Distantly Related Lineages of *Trypanosoma Cruzi*. *Proc. Natl. Acad. Sci. U.S.A.* 98 (13), 7396–7401. doi: 10.1073/pnas.121187198
- Mahe, F., Rognes, T., Quince, C., and Vargas, C., and Dunthorn, M. (2014). Swarm: Robust and Fast Clustering Method for Amplicon-Based Studies. *PeerJ* 2, e593. doi: 10.7717/peerj.593
- Marcili, A., Lima, L., Cavazzana, M., Junqueira, A. C., Veludo, H. H., Maia Da Silva, F., et al. (2009). A New Genotype of *Trypanosoma Cruzi* Associated With Bats Evidenced by Phylogenetic Analyses Using SSU rDNA, Cytochrome B and Histone H2B Genes and Genotyping Based on ITS1 rDNA. *Parasitology* 136 (6), 641–655. doi: 10.1017/S0031182009005861
- Ocana-Mayorga, S., Llewellyn, M. S., Costales, J. A., Miles, M. A., and Grijalva, M. J. (2010). Sex, Subdivision, and Domestic Dispersal of *Trypanosoma Cruzi* Lineage I in Southern Ecuador. *PLoS Negl. Trop. Dis.* 4 (12), e915. doi: 10.1371/journal.pntd.0000915
- Ochsenteiter, T., Cipriano, M., and Hajduk, S. L. (2007). KISS: The Kinetoplastid RNA Editing Sequence Search Tool. *RNA* 13 (1), 1–4. doi: 10.1261/rna.232907
- Opperdoes, F. R., and Michels, P. A. (2008). Complex I of Trypanosomatidae: Does it Exist? *Trends Parasitol* 24 (7), 310–317. doi: 10.1016/j.pt.2008.03.013
- Pinto, C. M., Kalko, E. K., Cottontail, L., Wellinghausen, N., and Cottontail, V. M. (2012). TcBat a Bat-Exclusive Lineage of *Trypanosoma Cruzi* in the Panama Canal Zone, With Comments on its Classification and the Use of the 18S rRNA Gene for Lineage Identification. *Infect. Genet. Evol.* 12 (6), 1328–1332. doi: 10.1016/j.meegid.2012.04.013
- Reinert, K., Dadi, T. H., Ehrhardt, M., Hauswedell, H., Mehringer, S., Rahn, R., et al. (2017). The SeqAn C++ Template Library for Efficient Sequence Analysis: A Resource for Programmers. *J. Biotechnol.* 261, 157–168. doi: 10.1016/j.jbiotec.2017.07.017
- Reis-Cunha, J. L., Baptista, R. P., Rodrigues-Luiz, G. F., Coqueiro-Dos-Santos, A., Valdivia, H. O., de Almeida, L. V., et al. (2018). Whole Genome Sequencing of *Trypanosoma Cruzi* Field Isolates Reveals Extensive Genomic Variability and Complex Aneuploidy Patterns Within TcII Dtu. *BMC Genomics* 19 (1), 816. doi: 10.1186/s12864-018-5198-4
- Renaud, G., Stenzel, U., and Kelso, J. (2014). leeHom: Adaptor Trimming and Merging for Illumina Sequencing Reads. *Nucleic Acids Res.* 42 (18), e141. doi: 10.1093/nar/gku699
- Reuter, J. S., and Mathews, D. H. (2010). Rnastructure: Software for RNA Secondary Structure Prediction and Analysis. *BMC Bioinf.* 11:129. doi: 10.1186/1471-2105-11-129
- Rusman, F., Florida-Yapur, N., Ragone, P. G., Diosque, P., and Tomasini, N. (2020). Evidence of Hybridization, Mitochondrial Introgression and Biparental Inheritance of the kDNA Minicircles in *Trypanosoma Cruzi* I. *PLoS Negl. Trop. Dis.* 14 (1), e0007770. doi: 10.1371/journal.pntd.0007770
- Rusman, F., Tomasini, N., Yapur, N. F., Puebla, A. F., Ragone, P. G., and Diosque, P. (2019). Elucidating Diversity in the Class Composition of the Minicircle Hypervariable Region of *Trypanosoma Cruzi*: New Perspectives on Typing and kDNA Inheritance. *PLoS Negl. Trop. Dis.* 13 (6), e0007536. doi: 10.1371/journal.pntd.0007536
- Ruvalcaba-Trejo, L. I., and Sturm, N. R. (2011). The *Trypanosoma Cruzi* Sylvio X10 Strain Maxicircle Sequence: The Third Musketeer. *BMC Genomics* 12:58. doi: 10.1186/1471-2164-12-58
- Savill, N. J., and Higgs, P. G. (1999). A Theoretical Study of Random Segregation of Minicircles in Trypanosomatids. *Proc. Biol. Sci.* 266 (1419), 611–620. doi: 10.1098/rspb.1999.0680
- Savill, N. J., and Higgs, P. G. (2000). Redundant and non-Functional Guide RNA Genes in *Trypanosoma Brucei* are a Consequence of Multiple Genes Per Minicircle. *Gene* 256 (1-2), 245–252. doi: 10.1016/S0378-1119(00)00345-0
- Schmid, B., Riley, G. R., Stuart, K., and Goring, H. U. (1995). The Secondary Structure of Guide RNA Molecules From *Trypanosoma Brucei*. *Nucleic Acids Res.* 23 (16), 3093–3102. doi: 10.1093/nar/23
- Simpson, L., Douglass, S. M., Lake, J. A., Pellegrini, M., and Li, F. (2015). Comparison of the Mitochondrial Genomes and Steady State Transcriptomes of Two Strains of the Trypanosomatid Parasite, *Leishmania Tarentolae*. *PLoS Negl. Trop. Dis.* 9 (7), e0003841. doi: 10.1371/journal.pntd.0003841
- Simpson, L., Frech, G. C., and Maslov, D. A. (1996). RNA Editing in Trypanosomatid Mitochondria. *Methods Enzymol.* 264, 99–121. doi: 10.1016/S0076-6879(96)64012-9
- Simpson, L., Neckelmann, N., de la Cruz, V. F., Simpson, A. M., Feagin, J. E., Jasmer, D. P., et al. (1987). Comparison of the Maxicircle (Mitochondrial) Genomes of *Leishmania Tarentolae* and *Trypanosoma Brucei* at the Level of Nucleotide Sequence. *J. Biol. Chem.* 262 (13), 6182–6196. doi: 10.1016/S0021-9258(18)45555-X
- Speijer, D. (2010). Constructive Neutral Evolution Cannot Explain Current Kinetoplastid Panediting Patterns. *Proc. Natl. Acad. Sci. U.S.A.* 107 (7), E25. doi: 10.1073/pnas.0909867107
- Stuart, K. (1993). The RNA Editing Process in *Trypanosoma Brucei*. *Semin. Cell Biol.* 4 (4), 251–260. doi: 10.1006/scel.1993.1030
- Telleria, J., Lafay, B., Virreira, M., Barnabe, C., Tibayrenc, M., and Svoboda, M. (2006). *Trypanosoma Cruzi*: Sequence Analysis of the Variable Region of Kinetoplast Minicircles. *Exp. Parasitol* 114 (4), 279–288. doi: 10.1016/j.exppara.2006.04.005
- Thomas, S., Martinez, L. L., Westenberger, S. J., and Sturm, N. R. (2007). A Population Study of the Minicircles in *Trypanosoma Cruzi*: Predicting Guide RNAs in the Absence of Empirical RNA Editing. *BMC Genomics* 8:133. doi: 10.1186/1471-2164-8-133
- Tomasini, N. (2018). Introgression of the Kinetoplast Dna: An Unusual Evolutionary Journey in *Trypanosoma Cruzi*. *Curr. Genomics* 19 (2), 133–139. doi: 10.2174/1389202918666170815124832
- Tomasini, N., and Diosque, P. (2015). Evolution of *Trypanosoma Cruzi*: Clarifying Hybridisations, Mitochondrial Introgressions and Phylogenetic Relationships Between Major Lineages. *Mem Inst Oswaldo Cruz* 110 (3), 403–413. doi: 10.1590/0074-02760140401

- Tomasini, N., Lauthier, J. J., Monje Rumi, M. M., Ragone, P. G., Alberti D'Amato, A. M., Brandan, C. P., et al. (2014). Preponderant Clonal Evolution of *Trypanosoma Cruzi* I From Argentinean Chaco Revealed by Multilocus Sequence Typing (Mlst). *Infect. Genet. Evol.* 27C, 348–354. doi: 10.1016/j.meegid.2014.08.003
- Turner, C. M., Hide, G., Buchanan, N., and Tait, A. (1995). *Trypanosoma Brucei*: Inheritance of Kinetoplast DNA Maxicircles in a Genetic Cross and Their Segregation During Vegetative Growth. *Exp. Parasitol* 80 (2), 234–241. doi: 10.1006/expr.1995.1029
- Velazquez, M., Diez, C. N., Mora, C., Diosque, P., and Marcipar, I. S. (2008). *Trypanosoma Cruzi*: An Analysis of the Minicircle Hypervariable Regions Diversity and its Influence on Strain Typing. *Exp. Parasitol* 120 (3), 235–241. doi: 10.1016/j.exppara.2008.07.016
- Westenberger, S. J., Cerqueira, G. C., El-Sayed, N. M., Zingales, B., Campbell, D. A., and Sturm, N. R. (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
- Zingales, B., Andrade, S. G., Briones, M. R., Campbell, D. A., Chiari, E., Fernandes, O., et al. (2009). A New Consensus for *Trypanosoma Cruzi* Intraspecific Nomenclature: Second Revision Meeting Recommends TcI to TcVI. *Mem Inst Oswaldo Cruz* 104 (7), 1051–1054. doi: 10.1590/s0074-02762009000700021

**Conflict of Interest:** 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 © 2021 Rusman, Florida-Yapur, Tomasini and Diosque. 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) and the copyright owner(s) 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.