



UNIVERSITY OF LEEDS

This is a repository copy of *Exploring TERRA during Leishmania major developmental cycle and continuous in vitro passages*.

White Rose Research Online URL for this paper:
<https://eprints.whiterose.ac.uk/172116/>

Version: Accepted Version

Article:

Morea, EGO, Vasconcelos, EJR orcid.org/0000-0001-5130-6622, Alves, CDS et al. (5 more authors) (2021) Exploring TERRA during Leishmania major developmental cycle and continuous in vitro passages. *International Journal of Biological Macromolecules*, 174. pp. 573-586. ISSN 0141-8130

<https://doi.org/10.1016/j.ijbiomac.2021.01.192>

© 2021, Elsevier. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

Highlights

- . TERRA expression was detected in some but not all chromosomes of *L. major* life stages
- . TERRA expression is regulated during *L. major* developmental cycle and continuous passages
- . In *L. major* TERRA transcripts are polyadenylated and processed by *trans*-splicing
- . Increased expression of TERRA and TERRA R-loops were mostly detected in the infective forms
- . TRF profiles varied during parasite development and may be directly involved with TERRA transcription regulation

Exploring TERRA during *Leishmania major* developmental cycle and continuous *in vitro* passages

Edna Gicela Ortiz Morea, Elton Jose Rosas Vasconcelos, Cristiane de Santis Alves, Selma Giorgio, Peter J. Myler, Helio Langoni, Claus Maria Azzalin, Maria Isabel Nogueira Cano

Abstract

Telomeres from different eukaryotes, including trypanosomatids, are transcribed into TERRA noncoding RNAs, crucial in regulating chromatin deposition and telomere length. TERRA is transcribed from the C-rich subtelomeric strand towards the 3'-ends of the telomeric array. Using bioinformatics, we confirmed the presence of subtelomeric splice acceptor sites at all *L. major* chromosome ends. Splice leader sequences positioned 5' upstream of *L. major* chromosomes subtelomeres were then mapped using SL-RNA-Seq libraries constructed from three independent parasite life stages and helped confirm TERRA expression from several chromosomes ends. Northern blots and RT-qPCR validated the results showing that *L. major* TERRA is processed by *trans*-splicing and polyadenylation coupled reactions. The number of transcripts varied with the parasite's life stage and continuous passages, being more abundant in the infective forms. However, no putative subtelomeric promoters involved in TERRA's transcriptional regulation were detected. In contrast, the observed changes in parasite's telomere length during development, suggest that differences in telomeric base J levels may control TERRA transcription in *L. major*. Also, TERRA-R loops' detection, mainly in the infective forms, was suggestive of TERRA's involvement in telomere protection. Therefore, *Leishmania* TERRA shares conserved features with other eukaryotes and advances new telomere specific functions in a Public Health-impacting parasite.

Keywords: *Leishmania* sp. life stages and *in vitro* passages; telomere transcription; TERRA R-loops

Exploring TERRA during *Leishmania major* developmental cycle and continuous *in vitro* passages

Edna Gicela Ortiz Morea¹, Elton Jose Rosas Vasconcelos², Cristiane de Santis Alves¹, Selma Giorgio³, Peter J. Myler⁴, Helio Langoni⁵, Claus Maria Azzalin⁶, Maria Isabel Nogueira Cano^{1*}

¹Department of Chemical and Biological Sciences, Biosciences Institute, São Paulo State University, UNESP, Botucatu, São Paulo, Brazil;

²Leeds Omics, University of Leeds, United Kingdom;

³Department of Animal Biology, Biology Institute, State University of Campinas, UNICAMP;

⁴Department of Global Health and Department of Biomedical Informatics & Medical Education, University of Washington, Seattle, Washington, United States of America;

⁵Department of Public Health, Veterinary Medical School, São Paulo State University, UNESP, Botucatu, São Paulo, Brazil;

⁶Institute of Molecular Medicine, Lisbon, Portugal.

*Corresponding author: maria.in.cano@unesp.br.

Declarations of interest: none'.

Abstract

1 Telomeres from different eukaryotes, including trypanosomatids, are transcribed into TERRA
2 noncoding RNAs, crucial in regulating chromatin deposition and telomere length. TERRA is transcribed from
3 the C-rich subtelomeric strand towards the 3'-ends of the telomeric array. Using bioinformatics, we confirmed
4 the presence of subtelomeric splice acceptor sites at all *L. major* chromosome ends. Splice leader sequences
5 positioned 5' upstream of *L. major* chromosomes subtelomeres were then mapped using SL-RNA-Seq
6 libraries constructed from three independent parasite life stages and helped confirm TERRA expression from
7 several chromosomes ends. Northern blots and RT-qPCR validated the results showing that *L. major* TERRA
8 is processed by *trans*-splicing and polyadenylation coupled reactions. The number of transcripts varied with
9 the parasite's life stage and continuous passages, being more abundant in the infective forms. However, no
10 putative subtelomeric promoters involved in TERRA's transcriptional regulation were detected. In contrast,
11 the observed changes in parasite's telomere length during development, suggest that differences in telomeric
12 base J levels may control TERRA transcription in *L. major*. Also, TERRA-R loops' detection, mainly in the
13 infective forms, was suggestive of TERRA's involvement in telomere protection. Therefore, *Leishmania*
14 TERRA shares conserved features with other eukaryotes and advances new telomere specific functions in a
15 Public Health-impacting parasite.
16
17
18
19

20 Keywords: *Leishmania* sp. life stages and *in vitro* passages; telomere transcription; TERRA R-loops
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1. Introduction

1 Protozoan parasites of the *Leishmania* genus (Trypanosomatidae family) belong to the Kinetoplastida
2 order, characterized by the kinetoplast's presence, a disk-shaped structure containing a network of circular
3 DNA (called kDNA), located in a unique and highly ramified mitochondrion. Some *Leishmania* species can
4 cause leishmaniasis, a neglected tropical disease that presents different clinical manifestations [1–3]. During
5 its life cycle, *Leishmania* spp. undergoes three main morphologically distinct developmental stages
6 (amastigote, promastigote, and metacyclic promastigote). The amastigote form resides within
7 phagolysosomes of the vertebrate host's (mainly mammals) mononuclear phagocytic system. The parasite
8 is transmitted to the mammalian host during the insect (invertebrate host) blood meal. Insects of the genus
9 Phlebotomine (Old world) or Lutzomyia (New world) ingest amastigote forms of the parasite, and after about
10 12 to 18 h of being in the insect's digestive system, the amastigotes transform into promastigotes. These
11 newly-transformed promastigotes are known as procyclic promastigotes. Promastigotes multiply quickly in
12 the insect's midgut, and after 30-60 hours, they transform into nectomonad promastigotes. After 7-10 days,
13 nectomonads migrate to the insect's pharyngeal valve and differentiate into the infective metacyclic form
14 through metacyclogenesis. Reports show that only metacyclic promastigotes get injected by the infected
15 insect, during a blood meal, into the vertebrate host skin [4–9]. The entire parasite developmental cycle can
16 be replicated in vitro by maintaining promastigotes in exponential growth in axenic cultures. Promastigote
17 cultures that reach the stationary growth phase contain metacyclics forms that can be selected and used to
18 infect macrophage cultures in vitro to transform into amastigotes [4,7,9,10].

19 Since there are no vaccines or efficient treatments available, several ongoing studies are focused on
20 finding new and better strategies to eradicate leishmaniasis [11]. Identifying new functional sequences in the
21 parasite's genome and transcriptome may provide sources for novel specific therapeutic targets. The
22 telomeric environment, which plays an important role in genome stabilization and cell proliferation, presents
23 many peculiar and species-specific features in *Leishmania* and is considered a potential target for exploitation
24 against the parasite. [12–15].

25 Telomeres are nucleoprotein structures formed by proteins and noncoding repetitive DNA sequences
26 located at chromosomes' extremities [12,13]. Telomeric DNA is composed of DNA in double- and single-
27 strand structures, with the G-rich strand forming a protrusion towards the end of the chromosome, the 3' G-
28 overhang [14–17]. Long noncoding RNAs (lncRNA) originated from both subtelomeric/telomeric strands have
29

already been described. One of these transcripts, dubbed TERRA, Telomeric Repeat-containing RNA, derives from the C-rich subtelomeric strand. TERRA was described in a range of eukaryotes, including mammals yeast and protozoan trypanosomatids (e.g., *Leishmania* sp.) [18–24]. TERRA molecules are mainly transcribed by RNA polymerase II (RNAPII) and, in mammals and yeast, carry a 7-methyl-guanosine 5' end cap. In budding yeast, all TERRA transcripts are polyadenylated, whereas, in humans and fission yeast, only a fraction of them undergo polyadenylation [18–21,25].

In protozoa, including *Leishmania* spp., it was previously reported that different types of RNA polymerases transcribe telomeric RNAs [22]. Recently, polyadenylated TERRA transcripts were identified in different developmental stages of *Trypanosoma brucei* and some *Leishmania* species [23,24]. Also, in *T. brucei*, increased TERRA transcription was detected in cells depleted of TbRAP1. TbRAP1 is a homolog of the yeast and mammalian RAP1 (Repressor Activator Protein 1) that associates with the parasite telomeric protein TbTRF (Telomeric Repeat Factor) and is essential for variant surface glycoproteins (VSG) silencing and cell viability [23].

In model organisms, TERRA transcripts are involved in telomere length regulation and replication, is essential for telomeres maintenance, DNA damage response at telomeres, and for the telomeric chromatin assembly [26,27]. TERRA can associate with hnRNPA1, which can also bind telomeres, and both can regulate telomere extension by telomerase in a three-state model [28]. But, it is still unclear which mechanisms regulate the abundance of TERRA transcripts since it is highly dependent on several factors, such as cell cycle, cell developmental stage, telomere length, and different stress conditions [26,27,29].

TERRA can form telomeric R-loops (TERRA R-loops), and there is evidence to show that telomeric R-loops' formation in mammals, yeast, and trypanosomes is influenced by the amount of TERRA transcripts, telomere size, and cell conditions [20-27]. R-loops are triple-stranded structures composed of RNA-DNA hybrids formed within the DNA double helix. R-loops formation causes the displacement of a single-stranded DNA generally found in G-C-rich genomic regions in different organisms, such as telomeres [30–34]. Recently, it was shown that the formation of TERRA R-loops *in vitro* is catalyzed by the recombinase RAD51, which physically interacts with TERRA and can mediate TERRA-telomere strand invasion [35]. It is already known that human somatic cells maintained under normal conditions transcribe low TERRA levels and form R-loops [36–39]. In contrast, in abnormal conditions, such as during disease, high levels of TERRA and telomeric R-loops are detected [40–44]. In yeast, TERRA and TERRA R-loops accumulate at short telomeres,

activating a DNA damage response that promotes homology-directed repair to avoid premature senescence [34]. Therefore, TERRA regulates the telomeric chromatin structure and telomere elongation [37,38,40].

The present work shows the identification of TERRA transcripts in *Leishmania major*, using SL-RNA-Seq libraries obtained from the three parasite life stages and *in silico* mining based on known TERRA loci features described in other organisms. Validation of these results by northern blot, RNA-FISH, and RT-qPCR analyses confirmed the existence of polyadenylated TERRA transcripts originating from some, but not all chromosome ends in each parasite life stage, supporting the SL-RNA-Seq results. The results also showed that, in *Leishmania*, TERRA is processed through *trans*-splicing and polyadenylation coupled reactions. Northern blot assays demonstrated the presence of TERRA transcripts, showing different expression levels depending on the parasite life stage and *in vitro* passage. The number of TERRA transcripts was higher in the parasite's infective forms (metacyclics>amastigotes) than in newly-transformed procyclic promastigotes and in parasites from continuous *in vitro* passages. We had also noted differences in the telomeric restriction fragment (TRF) profiles during parasite development. We speculate that it is probably due to differences in telomeric base J levels. In trypanosomatids, base J is considered an RNA polymerase II transcription terminator and an epigenetic marker, suggesting that it could be implicated with TERRA transcription regulation in *Leishmania major*. TERRA R-loops were also detected in all parasite life stages, mainly in the infective forms, which are usually exposed to a hostile environment in the mammalian host. Our results show that *Leishmania* TERRA shares conserved features with other eukaryotes and advances new specific functions in a parasite of great medical importance. It is the first time TERRA transcript is thoroughly assessed through reductionist and high-throughput corroborative assays in *Leishmania*, paving the way for further applied research using TERRA as a potential target to fight against leishmaniasis. The impact of these findings on parasite telomere biology is discussed.

2. Material and Methods

2.1. Parasite cultures and isolation of amastigotes and metacyclic promastigotes

In the present study, we used *Leishmania major* strain (MHOM/IL/1980/FRIEDLIN) from Oswaldo Cruz Institute collection, confirming that we worked with a genetically homogenous population. For all assays, parasites in the promastigote form were cultivated in exponential phase at 26 °C in 1X M199 medium pH7.3 (Cultilab), supplemented with 10% (v/v) heat-inactivated fetal calf serum (Cultilab), 25 mM HEPES and 1%

(v/v) antibiotic/antimycotic solution (Cultilab). Parasite cultures showed no contamination with Mycoplasma using the MycoFluor™ Mycoplasma Detection Kit (Molecular Probes).

In most assays, we used parasite life stages obtained from the same developmental cycle. Therefore, newly *in vitro*-transformed promastigotes, here named promastigotes P1 (passage 1) or procyclic promastigotes, were differentiated from amastigotes extracted from mice footpad lesions [45] after inoculation into M199 medium (see above) at 26 °C for 24h. Metacyclics M1 were selected from stationary phase promastigotes P1 cultures using agglutination with peanut lectin [10].

Promastigotes passages P2 to P24 represent the continuous cultivation in the exponential growth (every four days) of promastigotes P1. Metacyclics M2-M24 were obtained from stationary phase cultures (day ten of culture) of promastigotes passages P2-P24. To calculate parasites proliferation rate, we considered an *L. major* promastigote to take about 10h to complete one cell cycle [46]. Growth curves constructed using different promastigotes passages did not detect differences in cell growth profiles.

2.2. Mapping of the subtelomeric CSB (Conserved Sequence Box) motifs in *L. major* genome

We developed and ran an ad-hoc PERL script (<https://github.com/eltonjrv/bioinfo.scripts/blob/master/pattern-position.pl>) to map subtelomeric CSB (Conserved Sequence Box) motifs [47] onto the *L. major* Friedlin genome. We considered either forward or reverse orientation CSB motif matches, by running the script twice like the following:

```
$ perl pattern-position.pl TriTrypDB-38_LmajorFriedlin_Genome.fasta
GTACAGT.\{1,51\}GGAGAGGGTGT >CSBs-position-fwr.tab
$ perl pattern-position.pl TriTrypDB-38_LmajorFriedlin_Genome.fasta
ACACCCTCTCC.\{1,51\}ACTGTAC >CSBs-position-rev.tab.
```

The regular expression “CSB motif” representation on the command lines above was based on Fu & Barker (1998) [47]. Either “no matches” or matches within chromosomes ends-only (less than 5 kb from the extremities) were reported for each chromosome. Once we obtained the CSB genomic coordinates, we were able to load them into the Artemis genome browser (PMID: 22199388) [48] and assign them to their respective putative TERRA SL signal, using an arbitrary non-CDS SL-surrounding region of 2 kb.

2.3. Prediction of putative subtelomeric CpG islands

To predict CpG islands throughout the whole *L. major* genome, we ran `cpgi130.pl` [49] with the following parameters: `GCC=65`, `OE=0.65`, and `LENGTH=400`. The `cpgi130.pl` per-chromosome output files were then parsed through an ad-hoc PERL script (<https://github.com/eltonjrv/bioinfo.scripts/blob/master/cpg2gff.pl>) that generates a gff file suitable for visual inspection within subtelomeric regions using the Artemis genome browser [48]. Within the putative CpG island, we identified a putative methylated *Ava*I restriction site at the subtelomeres of *L. major* Chr10R in all three parasite life stages.

To validate the in-silico results, we treated *L. major* genomic DNA obtained from the three parasite life stages with bisulfite (see section 2.11 for the protocol to obtain *L. major* genomic DNA) using the EZ DNA Methylation Gold Kit (Zymo Research), followed by amplification by nested PCR of an 857 bp fragment that contains the putative subtelomeric CpG island. TOPO TA[®] plasmid (Invitrogen) was used to clone the PCR products and transform *E. coli* DH5- α . Plasmids DNA of ten colonies of each parasite life stage was automated sequenced and then aligned with ClustalW multiple sequences alignment tool through the MEGA software [50], using the putative reference CpG island sequence as the query.

2.4. Splice acceptor sites (SAS) mining within subtelomeric regions

An ad-hoc PERL script (<https://github.com/eltonjrv/bioinfo.scripts/blob/master/mining-subtel-polyY.pl>) was developed to search for SAS (Splice Acceptor Sites) right downstream of the last subtelomeric protein-coding gene, as an indication of eventual splice acceptor sites for TERRA. SAS are generally composed of a polypyrimidine tract followed by an AG dinucleotide (PolyY tract + AG) [51,52].

2.5. RNA-Seq analysis of putative TERRA transcripts

Independent Splice Leader (SL) RNA-Seq (SL-RNA-Seq) libraries from *L. major* developmental stages (amastigotes, promastigotes procyclics, and metacyclics) were constructed according to Cuypers et al. (2017) [53] and used to search for TERRA transcripts at each parasite chromosome end termini.

All three SL-RNA-Seq libraries passed through a quality control assessment with FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>), followed by alignment against the *L. major* Friedlin genome (TriTrypDB v38*) with bowtie2 (PMID: 22388286) [54], using the following parameters: `--very-sensitive-local -N 1`. Bowtie2-generated bam files were loaded into the Artemis genome browser (PMID:

22199388) [48], which allowed us to compute RPKM values [55] for putative TERRA transcripts, using RPKM ≥ 1 as a threshold for expression. As the transcript length metric for the RPKM calculation, we arbitrarily established 500 nt surrounding the L- or R-most non-CDS SL signal positioned 5' upstream of the subtelomeric region in each chromosome end. A heatmap of putative TERRA transcripts expression was plotted within the R environment (version 3.5.2) using the 'heatmap.2' function from the gplots library. Myler's Lab provided those SL-RNA-Seq libraries to the TriTrypDB web resource (https://tritrypdb.org/tritrypdb/app/record/dataset/DS_8bc463a882).

2.6. RNA isolation

Total RNA was obtained from approximately 5×10^6 cells from the three *L. major* life stages (promastigotes, metacyclics, and amastigotes) using Trizol (Invitrogen). Each 20 μg of total RNA was treated three times with 2 μl DNase I *Amplification* Grade (Invitrogen). Each sample's amount and purity were estimated by measuring $\text{OD}_{260 \text{ nm}}$ in a spectrophotometer Epoch (BioTek). RNA samples were submitted to PCR amplification using primers that specifically amplify the *L. major* alfa-tubulin gene (GenBank Acc# AL359777) to check for genomic DNA contamination (data not shown).

2.7. Northern blot analysis

Total RNA (5 μg each) from *L. major* parasites originated from the same developmental cycle (amastigotes, promastigotes P1, and metacyclics M1) and promastigotes and metacyclics from continuous *in vitro* passages (respectively, P2, P4, and P6, and, M2, M4, and M6) were used to perform the Northern blot analyses. As controls, RNAs were treated with RNase A (10 $\mu\text{g}/\mu\text{l}$). Before detecting telomeric transcripts, RNAs were treated 3-5 times with DNase I followed by PCR amplification to ensure samples were not contaminated with DNA (data not shown). RNA samples were fractionated onto 12-15% PAGE gels containing 10M urea in 1X TBE due to the limited amount of samples after DNaseI treatment. RNA samples were transferred to nylon membranes and hybridized with DIG-labeled TELC and with DIG-labeled TELG used as probes (Suppl. Table 1). The reactions were revealed using an anti-DIG-AP (Roche) and CPD Star (Roche).

2.8. RT-qPCR

1 For the RT-qPCR reactions, oligodT was also used for the first strand cDNAs synthesis using the
2
3 Improm II cDNA synthesis kit (Promega). RT-qPCR assays were done in triplicates with 1µL of cDNA and
4
5 10µM of each forward and reverse primers specific for the subtelomeres (Suppl. Table 1) of tested
6
7 chromosomes (Chr04L, Chr09L, Chr10R, Chr20L, and Chr29R), using *PowerUP SYBR Green mix (Applied*
8
9 *Biosystems)*. Amplifications were performed in *QuantStudio 12K* (Life Technologies). Dissociation curves
10
11 were included in all amplification runs, and Ct (*threshold cycle*) values were calculated. The gene encoding
12
13 the 19S proteasome non-ATPase subunit 8 (RPN8) (LmF.32.0390) was used as an internal control based
14
15 on Inbar et al. (2017) [56]. The relative gene quantification was calculated using the *Delta Ct* method [57],
16
17 obtained through the difference in threshold cycle between the target and the reference gene (RPN8). P-
18
19 values were obtained using a two-tailed non-parametric Wilcoxon Signed Rank Test, and $P < 0,05$ was
20
21 assumed as statistically significant. Statistical analyses were performed using GraphPad Prism (version 5.0).
22
23
24
25
26

2.9. RNA-FISH

27
28 Exponentially growing promastigotes (1×10^6) and metacyclics (1×10^6) were fixed with 4% formalin for
29
30 5 min, followed by several washes in 1X PBS and attached to glass coverslips coated with 0.1% poly-L-lysine
31
32 (Sigma). Subsequently, cells were dehydrated through an ethanol series (70%, 90%, and 100%) for 5 min
33
34 each at 10 °C and air dry. A PNA FITC-labeled telomeric DNA oligo probe (CCCTAA)₃ (PANAGENE) was
35
36 used. For the hybridization step, cells were incubated 3h at 56°C in the dark with 10 µM of the probe in the
37
38 presence of 1X hybridization buffer (70% formamide, 20 mMTris-HCl, pH 7.0, and 1% BSA). After the
39
40 hybridization step, cells were washed twice at 39 °C with 50% deionized formamide and 50 mMTris-HCl, pH
41
42 7.6, washed twice at 39 °C with 50 mMTris-HCl, pH 7.6, and finally washed with 50 mMTris-HCl, pH 7.6 at
43
44 room temperature. DNA in the nucleus and kinetoplast were stained with Vectashield® mounting medium
45
46 DAPI (Vector Labs). As a control, fixed cells were treated with RNase A (10µg/µl) for 1h at 37°C before
47
48 hybridization. PNA FITC-labeled telomeric DNA oligo probe (CCCTAA)₃ was also used to hybridize telomeric
49
50 DNA using a modified telomeric FISH protocol [58]. Three biological replicates of each parasite life stage and
51
52 about 100 images of each slide were analyzed with a Nikon 80i fluorescence microscope and captured with
53
54 a digital camera (Nikon). When necessary, images were superimposed (merge images) using NIS elements
55
56 software (v. Ar 3.10).
57
58
59
60
61
62
63
64
65

2.10. Cloning and sequencing of *TERRA* transcripts

1 The terminal end of the Chr29R arm was cloned using a PCR-based strategy to confirm that the
 2 transcripts amplified from the chromosomes' terminal ends were *TERRA*. Clones were obtained using cDNA
 3 amplified with oligo dT for the first strand synthesis, and for the second strand synthesis, we used the Chr29R
 4 specific primer (oligonucleotide SubChr29 F). The cDNA amplicons were amplified using *Platinum® Taq* DNA
 5 Polymerase (Invitrogen), followed by fractionation in a 1% ethidium bromide agarose gel and cloning into the
 6 PCR 4-TOPO vector (Invitrogen). Recombinant plasmid DNA was used to transform *E. coli* DH5-alpha and
 7 purified using PureLink 96 HQ Mini Plasmid DNA Purification Kit (Invitrogen). Purified plasmid DNA was
 8 automated sequenced, and the nucleotide sequence was analyzed using CLC Sequence Viewer 7.6 and
 9 Blastn (<http://www.ncbi.nlm.nih.gov/blast>).

2.11. Extraction of genomic DNA and telomeric Southern blotting analysis

10 Parasite cells were harvested (1.0×10^8 cells), washed three times in sterile phosphate-buffered
 11 saline, and lysed in the presence of 10 $\mu\text{g}/\mu\text{l}$ proteinase K overnight at 56°C. Total genomic DNA was
 12 obtained using both phenol: chloroform extraction [59] and DNeasy Blood and Tissue kit (Qiagen). The DNA
 13 samples were resuspended in 10mMTris-HCl, 1mM EDTA pH 8,0, and subsequently stored at 4°C.

14 Genomic DNA was digested with 10U *AfaI* (Thermo Scientific, Waltham, MA, USA) at 37°C overnight
 15 to liberate chromosome end termini [60]. DNA fragments were fractionated onto a 0.8% agarose gel and
 16 transferred to nylon membranes. Southern blots were hybridized using a DIG-labeled telomeric probe (DIG-
 17 TELC) (Suppl. Table 1). The hybridization signals were developed by chemiluminescence after incubating
 18 the membranes with an anti-DIG serum (Roche) covalently coupled to alkaline phosphate, followed by
 19 incubation with CPD-Star (Roche).

2.12. DRIP: DNA-RNA hybrid immunoprecipitation followed by qPCR

20 Approximately 2×10^7 cells of *L. major* amastigotes, log-phase promastigotes (P2, P4, and P6), and
 21 metacyclics (M2, M4, and M6) were used to obtain formaldehyde cross-linked chromatin as described before
 22 [61]. Ten percent of the chromatin was used for immunoprecipitation, and 10% served as input and qPCR
 23 reference. In summary, cross-linked chromatin was immunoprecipitated with the S9.6 antibody (final
 24 concentration 32 $\mu\text{g}/\text{ml}$) using the Immunoprecipitation kit Dynabeads Protein A (Life Technologies),
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65

according to the manufacturer instructions. DNA was cleaned with QIAquick PCR Purification Kit (Qiagen).

As the control, samples were pre-treated with 50 U recombinant RNase H (New England BioLabs) for 2.5 h at 37 °C, followed by immunoprecipitation with the S9.6 antibody.

DNA obtained from immunoprecipitated chromatin was subjected to qPCR reactions using specific combinations of TERRA primers (Suppl. Table 1) originated from the ends of Chr10R, Chr20L, and Chr29R. Reactions were performed in 40 cycles using the *PowerUP SYBR Green mix (Applied Biosystems)*. Ct (*threshold cycle*) values were used to analyze the percentage of TERRA R-loops in each chromosome end and were calculated as described [62], using the formula: $100 * 2^{(adjusted\ input - Ct(IP))}$, with $adjusted\ input = Ct(input) - \log_2(20)$.

Results and Discussion

3.1. Northern blot confirms the existence of TERRA transcripts originated from different *L. major* chromosome ends

We performed an *in-silico* mining based on the conserved features associated with telomeric transcription loci in other organisms to search for telomeric transcripts in *L. major* genome. We first searched all *L. major* Friedlin chromosome ends for the presence of subtelomeric CSB (Conserved Sequence Box) motifs using an ad-hoc PERL script [47,51,60,63–65] (see Methods section). This analysis showed that not all the 72 chromosome ends from the *L. major* genome contain subtelomeric CSB elements. They were found at the termini of the left (L) arm from 20 chromosomes and the termini of the right (R) arm from 25 chromosomes (Table 1).

We then mapped SL (splice leader) signals, positioned 5' upstream of all subtelomeric regions, using data obtained from SL-RNA-Seq libraries, since SL signals at that position would indicate telomere transcription. SL are tri-methyl guanosine 5'-capped 39 nt-long conserved markers of RNA maturation in trypanosomatids. SL-RNA-seq libraries are sequenced based on the SL sequence's presence at the 5' end of each transcript [53]. Here is important to remind that trypanosomatids' genes are organized in directional clusters that are polycistronically transcribed by RNA polymerase II in pre-mRNAs (applying for all mRNAs and some ncRNAs). Pre-mRNAs are then quickly processed in the nucleus through both *trans*-splicing and polyadenylation coupled reactions [47,52,66]. During *trans*-splicing, an SL sequence derived from the SL

RNA is added to the 5' end of all mature mRNAs, where the SL acceptor sites (SAS) are generally composed of a polypyrimidine tract followed by an AG dinucleotide located within the intercistronic region of the pre-mRNA molecule in the polycistron. The addition of an SL signal is coupled with the addition of a poly-A tail at the 3' end of the upstream gene in the polycistron [51,65]. An ad-hoc PERL script was developed to search for SAS at all *L. major* chromosome ends. The results showed SAS positioned right downstream of the last subtelomeric protein-coding gene, in all chromosome ends (data not shown) regardless of the presence of previously detected SL signals.

We further detected subtelomeric SL signals originated only from the C-rich telomeric strand at some, but not all parasite chromosome ends (explained in depth below, and shown in Table 2 and Suppl Fig.1A).

To validate this result, we used Northern blot analyses. In Figs. 1A-C, the assays were done with total RNA obtained from amastigotes, promastigotes P1 or procyclic promastigotes [56], and metacyclics M1. The results showed that only G-rich transcripts were detected in the amastigote and metacyclic (M1) stages (smeared signal in Fig. 1A). In contrast, no transcripts were detected in newly *in vitro*-transformed procyclic promastigotes (P1) (Fig. 1A).

Also, no hybridization signal was detected when RNA samples were pre-treated with RNase A (Fig. 1B), and any C-rich RNA transcripts were detected (Fig. 1D). Fig. 1C contains the quantitative analysis of the results presented in Fig. 1A.

Northern blots from total RNA obtained from promastigotes P2, P4, P6, and metacyclics M2, M4, and M6 were also performed, and the results showed G-rich transcripts in procyclic promastigotes P2-P6 (Fig. 2A). In metacyclics M2, a more diffuse and stronger signal can be detected (Fig. 2C). However, the intensity of the hybridization signal diminished in M4 and M6 (Fig. 2C). As a control, RNA samples were pre-treated with RNase A (Figs. 2B and 2D). Fig. 2E contains the quantitative analysis of the results presented in Figs. 2A and 2C.

These findings strongly suggest that the abundance of G-rich transcripts varies with the parasite life stage and within the same parasite life stage, depending on the passage.

According to the Northern blots shown in Figs. 1 and 2, there is a direct relationship between the totality of G-rich transcripts, the parasite life stage, and the continuous *in vitro* passages. To assess TERRA's eventual transcriptional regulation, we searched and found putative CpG islands at the subtelomeric regions using a reputable computational tool [49]. The results were validated using bisulfite sequencing of one of the

candidates at Chr10R subtelomeres, found in all three parasite life stages. The bisulfite sequence was negative for the presence of methylated cytosines, and thus, did not succeed in confirming subtelomeric promoters (Suppl. Figs. 1A-B, summarize this finding). This result was not a surprise since *Leishmania* sp., as other trypanosomatids, do not present canonical promoters for RNA polymerase II, implying an absence of gene expression regulation at the transcriptional level [52,65]. Thus, the Northern blots' results are more likely to reflect the recently described global genetic reprogramming occurring during parasite development inside the insect vector. After the insect blood meal, the ingested amastigotes differentiate into procyclic promastigotes (P1), transforming into nectomonads and further into leptomonads and metacyclics [56]. Nectomonads represent axenic promastigotes with multiple passages. Their transcriptome analyses revealed changes consistent with cell cycle arrest and the upregulation of genes associated with starvation and stress. At this stage, nectomonads approach metabolic similarities to metacyclics, its subsequent stage [4,56]. Inbar et al. (2017) [56] also showed that, although nectomonads and metacyclics transcriptomes are not so different, they both present unique signatures. Also, metacyclics obtained in axenic cultures *in vitro* did not show more significant differences in gene expression profiles than the metacyclics obtained *in vivo*. However, they both show changes compatible with pre-adaptation to the mammalian host's intracellular environment, making their transcriptome profile more similar to amastigotes [56].

3.2. In silico analysis of independent SL-RNA-Seq libraries indicates that G-rich RNAs are transcribed from the C-rich subtelomeric/telomeric strand

The studies from SL-RNA-Seq libraries (see ref. [53] and Methods section for details) revealed that, from a total of 72 chromosomes' ends (*L. major* has 36 chromosomes), SL-signals were detected only on the C-strand of 28 L arms and 27 R arms. SL-signals were found upstream of both subtelomeric and telomeric repeats, which indicates TERRA transcription. No transcripts originating from the telomeric G-rich strand (C-rich RNAs) [21,67] were detected in *L. major* by the adopted approach.

Analysis of potential TERRA transcription through SL-RNA-Seq libraries showed that amastigotes displayed 55/72 ends with SL signals positioned upstream of the subtelomeric sequences (28 at the L arms and 27 at the R arms). Procyclic promastigotes showed 38/72 chromosome ends with SL signals upstream of the subtelomeric sequences (17 at the L arms and 21 at the R arms). And metacyclics displayed 50/72 chromosome ends with SL signals upstream of the subtelomeric sequences (23 at the L arms and 27 at the R arms). As summarized in Table 2, regardless of the parasite life stage, TERRA transcripts were not

originated from the L arms of chromosomes 4, 8, 9, 11, 12, 13, and 24 nor the R arms of chromosomes 3, 4, 8, 9, 13, 16, 20, 22 and 27. Thus, chromosomes 4, 8, 9, and 13 are the only ones in *L. major* that do not express TERRA. Similarly, in humans, it was previously shown that TERRA transcripts are originated from most but not all chromosomes (36 different chromosome termini out of 46) [68].

We also checked for the number of SL signals (SL-RNA-Seq reads' stack at specific positions of the genome) localized upstream of the subtelomeric region in each chromosome end of the three parasites' life stages. We identified 55 subtelomeric SL signal at the chromosome ends of amastigotes, from which 23 had only one SL signal (14 at the L arm and 9 at the R arm). In procyclic promastigotes, from 38 chromosome ends containing SL signals, 17 had only one SL signal (9 at the L arm and 8 at the R arm). Finally, in metacyclics, 19 out of 50 subtelomeric SL signal-containing chromosome ends had only one SL signal (10 at the L arm and 9 at the R arm). All other chromosome ends that are transcribed, regardless of the parasite life stage, showed more than one SL signal localized upstream of the subtelomeric sequence (Suppl. Fig. 2A). For example, in Chr2R, there are two SL signals, one located 1,161 bp and the other located 616 bp from the first telomeric repeat, which could indicate that there might be other transcripts originating from the C-strand chromosome ends (Suppl. Fig. 2B). Similarly, fission yeast, besides TERRA, also expresses α ARRET, a G-rich lncRNA transcribed from more internal subtelomeric regions [21].

3.3. TERRA transcripts show different expression patterns across *L. major* developmental stages

RPKM (Reads Per Kilobase per Million mapped reads) normalization metric was used to evaluate possible TERRA expression differences throughout the *L. major* developmental cycle. This evaluation was based on the data obtained from the three SL-RNA-Seq libraries (AMA, PRO, and META) and by setting an arbitrary and fixed transcript length of 500 nt for all subtelomeric SL signals (Suppl. Table 2) (see Methods section for more details).

We compared RPKM of 72 *L. major* chromosome end termini obtained from each parasite life stage (Suppl. Table 3). Amastigotes and procyclic promastigotes share one telomeric read originating from Chr35L. In contrast, amastigotes and metacyclics share two telomeric reads originating from Chr21L and Chr27L, and four telomeric reads originating from Chr15R, Chr17R, Chr23R, and Chr33R (Suppl. Table 3). Curiously, procyclics and metacyclic promastigotes do not share any telomeric reads. These results suggest that amastigotes and metacyclics share a similar expression profile of some TERRA transcripts. It was already shown that when *L. major* amastigotes and metacyclics are residing in their natural hosts, they present

identical protein-coding gene expression patterns [56,69], corroborating the results shown herein for TERRA transcription.

We also compared TERRA-associated RPKMs from each chromosome end termini among all three *L. major* life stages. We observed that regardless of the chromosome arm, TERRA transcripts are more abundant in metacyclics than amastigotes and procyclics, following in that order (i.e., TERRA Chr29R) (Fig. 3). Our results also indicated differences in the same chromosome end's transcription level depending on the parasite life stage (i.e., comparing the RPKM of TERRA Chr29R in metacyclic and procyclic, as well as the northern blots results presented in Figs. 1 and 2). Differences in the number of TERRA transcripts originated from the same chromosome were also described in human cell lines [27,29,68]. A non-hierarchical clustering of TERRA expression, ordered by chromosome end terminus in each parasite life stage, illustrates these results. It shows that metacyclics and amastigotes present higher TERRA transcription levels than procyclic promastigotes (Fig. 3).

3.4. In *L. major* TERRA transcripts are localized in many nuclear foci in metacyclic forms

RNA-FISH allows the detection of RNA molecules on fixed cells and, thus, it was our method of choice for visualizing the intranuclear distribution of TERRA [18]. In this assay, we used *L. major* promastigotes P2 and metacyclics M2. Parasites were hybridized *in situ* using a PNA FITC-labeled probe containing three C-rich telomeric repeats. As shown in Fig. 4A, it was impossible to detect TERRA in promastigote cells through RNA-FISH. In contrast, all metacyclics showed hybridization signals distributed throughout DAPI-stained nuclei, with some more intense regions standing out (Fig. 4B). As a probe control, we used the hybridization of telomeric DNA (Figs. 4A-B, top panels). Hybridization signals were not detected in cells treated with RNase A, strongly suggesting that the identified nuclear foci in metacyclic cells correspond to TERRA. In agreement with this result, mammals' TERRA transcripts appear as foci distributed in the nucleoplasm and associated with telomeres throughout the cell cycle [18,70].

3.5. TERRA transcripts are processed by trans-splicing and are polyadenylated in all *L. major* life stages

TERRA transcripts were individually analyzed by RT-qPCR using primers specific for some individual chromosome end termini with unique sequences to confirm the SL-RNA-Seq findings. To choose which chromosome end terminus would be assessed, we relied on different features, such as *i*) the presence or

absence of the subtelomeric CSB elements, *ii*) the presence or absence of an SL signal upstream of the subtelomeric sequence, and *iii*) the possibility to find DNA sequences that were unique for a specific chromosome end and not shared with any other genomic loci. We started selecting the chromosome ends according to the presence/absence of both CSB sequences and SL signals by performing BLASTn alignments [71]. BLAST results (Table 1, Table 2) helped us classify chromosome ends into four different types, which presented exclusive features shared among the three parasite life stages (Suppl. Table 4, Suppl. Table 5, and Suppl. Table 6).

RT-qPCR reactions were done using oligo dT for the first strand synthesis and cDNA obtained from the three *L. major* life stages and parasites from continuous *in vitro* passages (promastigotes P2, P4, P6, P12, and P24, and metacyclics M2, M4, M6, M12, and M24). The assays were done to test for the presence of polyadenylated mature TERRA transcripts originated from Chr04L, Chr09L, Chr10R, Chr20L, and Chr29R. The results showed that any TERRA transcripts were detected from Chr04L and Chr09L in any analyzed sample (data not shown). In contrast, TERRA transcripts were amplified from Chr10R, Chr20L, and Chr29R in all three parasite stages (Table 3). However, TERRA transcripts from Chr20L and Chr29R were more abundant in amastigotes and promastigotes, respectively. TERRA transcripts were also amplified from Chr10R and Chr29R in metacyclics and from Chr20L only in metacyclics M2. TERRA originated from Chr29R, followed by the ones that originated from the terminus of Crh10R showed to be the most abundant mainly in metacyclics (Table 3). These results corroborate the SL-RNA-Seq data (Table 2 and Fig. 3) and confirm that mature TERRA lncRNAs are transcribed and processed from subtelomeric-telomeric regions of some *L. major* chromosomes. We also observed developmental stage-associated differences in TERRA expression profiles among chromosome ends, and we could only detect transcription from chromosome ends containing an SL signal right upstream of the subtelomeric sequence. Here is worth reminding that polyadenylation and *trans*-splicing are coupled reactions occurring at the first steps of RNA processing in trypanosomatids. We believe that, as already described in model eukaryotes, these transcripts are likely transcribed by RNA polymerase II [19–21,23–25,68]. These results strongly suggest that canonical early-stage RNA processing in trypanosomatids also occurs at subtelomeric-telomeric regions from some *L. major* chromosomes, responsible for the maturation of TERRA transcripts rather than a byproduct of a transcriptional readthrough event.

Northern blots using total RNAs obtained from promastigotes P2, P4, and P6 and metacyclics M2, M4, and M6 hybridized with a probe from Chr29R arm, showed that the mature TERRA transcribed from Chr29R is about 1,000 nt long and is expressed from both parasite life stages (Fig. 5A-B). In both metacyclics and promastigotes, we can also see a transcript $\geq 1,500$ nt-long hybridized with the Chr29R probe, which possibly corresponds to a precursor of TERRA Chr29R RNA (Fig. 5B).

Thus, these results corroborate the Northern blot and the SL-RNA-Seq results confirming that TERRA transcripts are more abundant in *L. major* metacyclics and amastigotes than in procyclic promastigotes (Figs 1-3).

We further cloned and sequenced TERRA transcripts originating from Chr29R obtained from promastigotes cDNA to ensure that the amplicons obtained were indeed TERRA. The nucleotide sequence analysis using BLASTn showed that the amplicon was 100% identical to Chr29R and did not share similarities with any other chromosome locus in the parasite genome (Suppl. Fig. 3). Similar results were obtained for TERRA Chr10R and TERRA Chr20L (data not shown).

3.7. Variation in the Telomeric Restriction Fragment profiles during parasite development may be directly involved with TERRA transcription regulation in *Leishmania major*

In model organisms, TERRA is engaged with various cellular processes, including telomere maintenance, regulation of telomere length and telomerase activity, and heterochromatin deposition [18,26,27,36]. In yeast, it was also shown that the abundance of TERRA transcripts is inversely proportional to telomere length, meaning that cells presenting short telomeres accumulate TERRA transcripts and R-loops, activating local DNA damage response (DDR) [20,27,36–38]. Thus, we decided to check whether this is a conserved feature also shared by *Leishmania* sp. *Leishmania* telomeres are formed by TTAGGG repeated sequences, and at the subtelomeric region, we can find the CSB elements [47]. *Afa* I restriction sites are located within the conserved CSB elements and are used to digest subtelomeric DNA [60] and to obtain Telomere Restriction Fragment (TRF) profiles (Fig. 6A). We analyzed the TRF profiles of *L. major* life stages that originated from a single developmental cycle and promastigotes and metacyclics from continuous *in vitro* passages (Figs. 6B and 6C).

Here is worth recalling that *L. major* telomeres are enriched of base J (β -D-glucosyl-hydroxymethyl uracil), a modification of thymines first described in *Trypanosoma brucei* and later found in other

kinetoplastids, including *Leishmania*. In *Leishmania*, 98% of all base J throughout the genome are at
 1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65

subtelomeres/telomeres, compared to other trypanosomatids, some of which present only 1% of base J at telomeres [72–74]. More important, it was demonstrated by others [71] and by Genest et al. [74] that the presence of high amounts of base J at *L. major* promastigotes chromosome ends inhibits DNA cleavage by frequently cutting restriction enzymes that would digest most of the subtelomeric repeats. Thus, the predominance of the hybridized telomeric repeat-containing fragments in the upper part of the Southern blots shown in Figs. 6B and C can explain the promastigotes TRF profiles. Also, there is no description of *L. major* amastigotes and metacyclics telomere length profiles in the literature. Therefore, this is the first report of TRF profiles from all *L. major* developmental stages. As shown in Fig. 6B, it seems that amastigotes may also present subtelomeric base J since most *AfaI* subtelomeric sites were not cleaved. The non- or partially digested subtelomeric DNA was detected in DNA fragments ranging from ~3.6 kb – >8.5 kb (here named base J-containing telomeric fragments). The TRF profile of amastigotes also showed faint hybridization signals in short DNA fragments ranging from <0.35 kb - \geq 1.0 kb (here named base J-free telomeric fragments). The opposite seems true for metacyclics M1 (Fig. 6B) and less for metacyclics M2 and M6 (Fig. 6C). In M1, telomere hybridization is mainly detected in short DNA fragments ranging from <0.35 kb - \geq 1.0 kb (Fig. 6B), suggesting that most of the subtelomeric DNA are free of modified base J. In contrast, the TRFs of metacyclic M2 and M6 showed an increased amount of non-digested subtelomeric DNA (here named base J-containing telomeric fragments) than M1 (Fig. 6C). Also, no visible differences were detected in the TRF profiles of promastigotes (P2 and P6) compared to the respective promastigotes P1 (Fig. 6C). Together, these results strongly suggest that only metacyclics, among the three developmental stages assessed herein, possibly present less amount of base J at subtelomeres.

Intriguingly, in 1997, van Leeuwen et al. [75] showed that in *T. brucei*, base J is present in silent VSG (variant surface glycoprotein, the main parasite virulent factor) expression sites. Still, it is absent from active VSG expression sites, suggesting that base J may play a role in transcription silencing in *T. brucei*. More recently, it was also shown that in *Leishmania*, base J prevents transcriptional readthrough, suggesting that it could serve as a recruitment signal to localize RNA polymerase II termination factors or a mark for a repressive chromatin structure [76]. Thus, in our opinion, the detected differences in TRF profiles combined with the noticed variations in TERRA expression during *L. major* developmental cycle (>metacyclics>amastigotes>promastigotes) and in parasites from continuous *in vitro* passages, argues in

favor of Hazelbaker & Buratowski [77] hypothesis, which expatiates on-base J functioning as i) a non-protein transcription terminator, ii) a recruitment signal to localize an unknown termination factor or iii) an epigenetic mark. We are currently working on proving how base J's presence affects TERRA expression in different *Leishmania* species.

3.8. TERRA R-loops are formed at some *L. major* chromosome ends

To verify whether in *Leishmania* TERRA can form R-loops, we performed RNA–DNA hybrid immunoprecipitation (DRIP) followed by qPCR. TERRA R-loops were confirmed by pre-treating the samples with RNase H, followed by immunoprecipitation with the S9.6 antibody.

The results showed that TERRA Chr10R, TERRA Chr20L, and TERRA Chr29R formed R-loops in amastigotes, while only TERRA Chr29R formed R-loops in promastigotes and metacyclics (Fig. 7 and Suppl. Table 7). Curiously, the Chr29R end terminus appears to transcribe high amounts of TERRA in all parasite life stages, mainly in metacyclics, according to the SL-RNA-Seq and RT-qPCR results (Figs. 3 and Table 3). In agreement, in *T. brucei*, R-loops accumulate in the VSG's subtelomeric expression sites of the infective bloodstream forms [78].

It is likely that high amounts of TERRA and TERRA R-loops, principally in *L. major* infective forms, would be part of a mechanism that helps parasites maintain telomere integrity by avoiding DNA damage generated from the hostile oxidative environment inside the mammalian host. We have already observed that acute oxidative stress induces telomere shortening in *Leishmania* promastigotes [79]. This observation argues in favor of TERRA exerting a protective role at *Leishmania* sp. telomeres, as demonstrated before in other eukaryotes [38,80]. Another hypothesis worth being tested is if TERRA R-loops' formation would be associated with replication-transcription conflicts recently described in trypanosomes [81]. If R-loops are the cause or the consequence of replication-transcription conflicts is a matter of debate that requires further experiments.

4. Conclusions

TERRA's conserved features in a divergent eukaryote, such as *Leishmania* spp., may shed light on the biological importance of the telomeric chromatin and the contributions of telomeric ncRNAs to maintain telomere homeostasis during evolution. Although it has to be proven, our results also suggest that base J at

parasite telomeres controls TERRA transcription during *Leishmania major* development and *in vitro* passages. Understand how important is the balance between TERRA transcription and TERRA R-loops for telomere maintenance and parasite survival may be detrimental in the development of antiparasitic therapies.

5. Acknowledgments

The authors want to acknowledge Mr. Mark Ewusi Shiburah for proofreading and English edition and Dr. Marcelo Santos da Silva for critical reading the manuscript.

Funding

This work was supported by São Paulo Research Foundation (FAPESP, Fundação de Amparo a Pesquisa do Estado de São Paulo, Brazil, <http://www.fapesp.br/>), grant 2018/04375-2 to MINC and grant 2018/23302-6 to SG. The National Council for Scientific and Technological Development, CNPq (grant 405581/2018-1), <http://www.cnpq.br/> also supports SG. Morea, E.G.O is a postdoctoral fellow from FAPESP (grant 2019/11496-3) and, MINC is a PQ-CNPq fellow (grant 302433/2019-8). The funders had no role in study design, data collection, analysis, decision to publish, or manuscript preparation.

Conflicts of Interest: The authors declare no conflicts of interest.

7. References

- [1] D.A. Johnston, M.L. Blaxter, W.M. Degraeve, J. Foster, A.C. Ivens, S.E. Melville, Genomics and the biology of parasites, *BioEssays*. 21 (1999) 131–147. [https://doi.org/10.1002/\(SICI\)1521-1878\(199902\)21:2<131::AID-BIES7>3.0.CO;2-I](https://doi.org/10.1002/(SICI)1521-1878(199902)21:2<131::AID-BIES7>3.0.CO;2-I).
- [2] K. Gull, The biology of kinetoplastid parasites: insights and challenges from genomics and post-genomics., *Int. J. Parasitol.* 31 (2001) 443–52. [https://doi.org/10.1016/s0020-7519\(01\)00154-0](https://doi.org/10.1016/s0020-7519(01)00154-0).
- [3] S. Antinori, L. Schifanella, M. Corbellino, Leishmaniasis: new insights from an old and neglected disease, *Eur. J. Clin. Microbiol. Infect. Dis.* 31 (2012) 109–118. <https://doi.org/10.1007/s10096-011-1276-0>.
- [4] S.M. Gossage, M.E. Rogers, P.A. Bates, Two separate growth phases during the development of *Leishmania* in sand flies: implications for understanding the life cycle., *Int. J. Parasitol.* 33 (2003) 1027–34. [https://doi.org/10.1016/s0020-7519\(03\)00142-5](https://doi.org/10.1016/s0020-7519(03)00142-5).
- [5] R. Killick-Kendrick, D.H. Molyneux, R.W. Ashford, Ultrastructural observations on the attachment of *Leishmania* in the sandfly., *Trans. R. Soc. Trop. Med. Hyg.* 68 (1974) 269. PMID: 4417943.
- [6] D.P. Neves, A.L. de Melo, O. Genaro, P.M. Linardi, *Parasitología humana*, 10th ed., Atheneu, São Paulo, 2000.
- [7] L. Rey, *Parasitologia*, Guanabara Koogan, Rio de Janeiro, 2001.
- [8] D. Sacks, S. Kamhawi, Molecular Aspects of Parasite-Vector and Vector-Host Interactions in Leishmaniasis, *Annu. Rev. Microbiol.* 55 (2001) 453–483. <https://doi.org/10.1146/annurev.micro.55.1.453>.

- [9] D.L. Sacks, P. V Perkins, Identification of an infective stage of *Leishmania* promastigotes., *Science*. 223 (1984) 1417–9. <https://doi.org/10.1126/science.6701528>.
- [10] R. da Silva, D.L. Sacks, Metacyclogenesis is a major determinant of *Leishmania* promastigote virulence and attenuation., *Infect. Immun.* 55 (1987) 2802–6. <https://doi.org/10.1128/IAI.55.11.2802-2806.1987>.
- [11] Leishmaniasis, WHO, World Health Organization. https://www.who.int/health-topics/leishmaniasis#tab=tab_1, 2020 (accessed 11 August 2020).
- [12] E.H. Blackburn, J.G. Gall, A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in *Tetrahymena*., *J. Mol. Biol.* 120 (1978) 33–53. [https://doi.org/10.1016/0022-2836\(78\)90294-2](https://doi.org/10.1016/0022-2836(78)90294-2).
- [13] F.E. Pryde, H.C. Gorham, E.J. Louis, Chromosome ends: all the same under their caps., *Curr. Opin. Genet. Dev.* [https://doi.org/10.1016/s0959-437x\(97\)80046-9](https://doi.org/10.1016/s0959-437x(97)80046-9).
- [14] S.R.W.L. Chan, E.H. Blackburn, Telomeres and telomerase, *Philos. Trans. R. Soc. B Biol. Sci.* 359 (2004) 109–122. <https://doi.org/10.1098/rstb.2003.1370>.
- [15] E.H. Blackburn, Cell biology: Shaggy mouse tales, *Nature*. 436 (2005) 922–923. <https://doi.org/10.1038/436922a>.
- [16] E.R. Henderson, E.H. Blackburn, An overhanging 3' terminus is a conserved feature of telomeres., *Mol. Cell. Biol.* 9 (1989) 345–8. <https://doi.org/10.1128/mcb.9.1.345>.
- [17] W.E. Wright, V.M. Tesmer, K.E. Huffman, S.D. Levene, J.W. Shay, Normal human chromosomes have long G-rich telomeric overhangs at one end., *Genes Dev.* 11 (1997) 2801–9. <https://doi.org/10.1101/GAD.11.21.2801>.
- [18] C.M. Azzalin, P. Reichenbach, L. Khoriatuli, E. Giulotto, J. Lingner, Telomeric Repeat Containing RNA and RNA Surveillance Factors at Mammalian Chromosome Ends, *Science* 318 (2007) 798–801. <https://doi.org/10.1126/science.1147182>.
- [19] S. Schoeftner, M.A. Blasco, Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II, *Nat. Cell Biol.* 10 (2008) 228–236. <https://doi.org/10.1038/ncb1685>.
- [20] B. Luke, A. Panza, S. Redon, N. Iglesias, Z. Li, J. Lingner, The Rat1p 5' to 3' Exonuclease Degrades Telomeric Repeat-Containing RNA and Promotes Telomere Elongation in *Saccharomyces cerevisiae*, *Mol. Cell.* 32 (2008) 465–477. <https://doi.org/10.1016/j.molcel.2008.10.019>.
- [21] A. Bah, H. Wischnewski, V. Shchepachev, C.M. Azzalin, The telomeric transcriptome of *Schizosaccharomyces pombe*, *Nucleic Acids Res.* 40 (2012) 2995–3005. <https://doi.org/10.1093/nar/gkr1153>.
- [22] G. Rudenko, L.H. Van der Ploeg, Transcription of telomere repeats in protozoa., *EMBO J.* 8 (1989) 2633–8. PMID: 2511008.
- [23] V. Nanavaty, R. Sandhu, S.E. Jehi, U.M. Pandya, B. Li, *Trypanosoma brucei* RAP1 maintains telomere and subtelomere integrity by suppressing TERRA and telomeric RNA:DNA hybrids., *Nucleic Acids Res.* 45 (2017) 5785–5796. <https://doi.org/10.1093/nar/gkx184>.
- [24] J.D. Damasceno, G. LA Silva, C. Tschudi, L.R. Tosi, Evidence for regulated expression of Telomeric Repeat-containing RNAs (TERRA) in parasitic trypanosomatids, *Mem. Inst. Oswaldo Cruz.* 112 (2017) 572–576. <https://doi.org/10.1590/0074-02760170054>.
- [25] A. Porro, S. Feuerhahn, P. Reichenbach, J. Lingner, Molecular Dissection of Telomeric Repeat-Containing RNA Biogenesis Unveils the Presence of Distinct and Multiple Regulatory Pathways, *Mol. Cell. Biol.* 30 (2010) 4808–4817. <https://doi.org/10.1128/MCB.00460-10>.
- [26] C.M. Azzalin, J. Lingner, Telomere functions grounding on TERRA firma, *Trends Cell Biol.* 25 (2015) 29–36. <https://doi.org/10.1016/j.tcb.2014.08.007>.
- [27] J.J. Montero, I. López de Silanes, O. Graña, M.A. Blasco, M. Serrano, Telomeric RNAs are essential to maintain telomeres, *Nat. Commun.* 7 (2016) 12534. <https://doi.org/10.1038/ncomms12534>.
- [28] S. Redon, I. Zemp, J. Lingner, A three-state model for the regulation of telomerase by TERRA and hnRNPA1, *Nucleic Acids Res.* 41 (2013) 9117–9128. <https://doi.org/10.1093/nar/gkt695>.
- [29] R. Arora, C.M.C. Brun, C.M. Azzalin, TERRA: Long Noncoding RNA at Eukaryotic Telomeres, in: Springer Berlin Heidelberg, 2011: pp. 65–94. https://doi.org/10.1007/978-3-642-16502-3_4.
- [30] P.A. Ginno, P.L. Lott, H.C. Christensen, I. Korf, F. Chédin, R-Loop Formation Is a Distinctive Characteristic of Unmethylated Human CpG Island Promoters, *Mol. Cell.* 45 (2012) 814–825. <https://doi.org/10.1016/j.molcel.2012.01.017>.
- [31] M.E. Reaban, J. Lebowitz, J.A. Griffin, Transcription induces the formation of a stable RNA-DNA hybrid in the immunoglobulin α switch region, *J. Biol. Chem.* 269 (1994) 21850–21857. PMID: 8063829.
- [32] P.A. Ginno, Y.W. Lim, P.L. Lott, I. Korf, F. Chédin, GC skew at the 59 and 39 ends of human genes links R-loop formation to epigenetic regulation and transcription termination, *Genome Res.* 23 (2013)

1590–1600. <https://doi.org/10.1101/gr.158436.113>.

- [33] L.A. Sanz, S.R. Hartono, Y.W. Lim, S. Steyaert, A. Rajpurkar, P.A. Ginno, X. Xu, F. Chédin, Prevalent, Dynamic, and Conserved R-Loop Structures Associate with Specific Epigenomic Signatures in Mammals, *Mol. Cell.* 63 (2016) 167–178. <https://doi.org/10.1016/j.molcel.2016.05.032>.
- [34] S. Toubiana, S. Selig, DNA:RNA hybrids at telomeres - when it is better to be out of the (R) loop, *FEBS J.* 285 (2018) 2552–2566. <https://doi.org/10.1111/febs.14464>.
- [35] M. Feretzaki, M. Pospisilova, R. Valador Fernandes, T. Lunardi, L. Krejci, J. Lingner, RAD51-dependent recruitment of TERRA lncRNA to telomeres through R-loops, *Nature.* (2020) 1–6. <https://doi.org/10.1038/s41586-020-2815-6>.
- [36] V. Pfeiffer, J. Lingner, D. Demarini, N. Shah, A. Wach, TERRA Promotes Telomere Shortening through Exonuclease 1–Mediated Resection of Chromosome Ends, *PLoS Genet.* 8 (2012) e1002747. <https://doi.org/10.1371/journal.pgen.1002747>.
- [37] E. Cusanelli, C.A.P. Romero, P. Chartrand, Telomeric Noncoding RNA TERRA Is Induced by Telomere Shortening to Nucleate Telomerase Molecules at Short Telomeres, *Mol. Cell.* 51 (2013) 780–791. <https://doi.org/10.1016/j.molcel.2013.08.029>.
- [38] M. Graf, D. Bonetti, A. Lockhart, K. Serhal, V. Kellner, A. Maicher, P. Jolivet, M.T. Teixeira, B. Luke, Telomere Length Determines TERRA and R-Loop Regulation through the Cell Cycle, *Cell.* 170 (2017) 72–85.e14. <https://doi.org/10.1016/j.cell.2017.06.006>.
- [39] S. Sagie, S. Toubiana, S.R. Hartono, H. Katzir, A. Tzur-Gilat, S. Havazelet, C. Francastel, G. Velasco, F. Chédin, S. Selig, Telomeres in ICF syndrome cells are vulnerable to DNA damage due to elevated DNA:RNA hybrids, *Nat. Commun.* 8 (2017). <https://doi.org/10.1038/ncomms14015>.
- [40] R. Arora, Y. Lee, H. Wischnewski, C.M. Brun, T. Schwarz, C.M. Azzalin, RNaseH1 regulates TERRA-telomeric DNA hybrids and telomere maintenance in ALT tumour cells., *Nat. Commun.* 5 (2014) 5220. <https://doi.org/10.1038/ncomms6220>.
- [41] H. Episkopou, I. Draskovic, A. Van Beneden, G. Tilman, M. Mattiussi, M. Gobin, N. Arnoult, A. Londoño-Vallejo, A. Decottignies, Alternative Lengthening of Telomeres is characterized by reduced compaction of telomeric chromatin, *Nucleic Acids Res.* 42 (2014) 4391–4405. <https://doi.org/10.1093/nar/gku114>.
- [42] C.A. Lovejoy, W. Li, S. Reisenweber, S. Thongthip, J. Bruno, T. de Lange, S. De, J.H.J. Petrini, P.A. Sung, M. Jasin, J. Rosenbluh, Y. Zwang, B.A. Weir, C. Hatton, E. Ivanova, L. Macconail, M. Hanna, W.C. Hahn, N.F. Lue, R.R. Reddel, Y. Jiao, K. Kinzler, B. Vogelstein, N. Papadopoulos, A.K. Meeker, Loss of ATRX, genome instability, and an altered DNA damage response are hallmarks of the alternative lengthening of Telomeres pathway, *PLoS Genet.* 8 (2012). <https://doi.org/10.1371/journal.pgen.1002772>.
- [43] Z. Deng, A.E. Campbell, P.M. Lieberman, TERRA, CpG methylation and telomere heterochromatin: Lessons from ICF syndrome cells, *Cell Cycle.* 9 (2010) 69–74. <https://doi.org/10.4161/cc.9.1.10358>.
- [44] S. Yehezkel, Y. Segev, E. Viegas-Péquignot, K. Skorecki, S. Selig, Hypomethylation of subtelomeric regions in ICF syndrome is associated with abnormally short telomeres and enhanced transcription from telomeric regions, *Hum. Mol. Genet.* 17 (2008) 2776–2789. <https://doi.org/10.1093/hmg/ddn177>.
- [45] C.L. Barbiéri, A.I. Doine, E. Freymuller, Lysosomal depletion in macrophages from spleen and foot lesions of Leishmania-infected hamster, *Exp. Parasitol.* 71 (1990) 218–228. [https://doi.org/10.1016/0014-4894\(90\)90024-7](https://doi.org/10.1016/0014-4894(90)90024-7).
- [46] A. Ambit, K.L. Woods, B. Cull, G.H. Coombs, J.C. Mottram, Morphological events during the cell cycle of leishmania major, *Eukaryot. Cell.* 10 (2011) 1429–1438. <https://doi.org/10.1128/EC.05118-11>.
- [47] G. Fu, D.C. Barker, Rapid cloning of telomere-associated sequence using primer-tagged amplification., *Biotechniques.* 24 (1998) 386–90. <https://doi.org/10.2144/98243bm10>.
- [48] T. Carver, S.R. Harris, M. Berriman, J. Parkhill, J.A. McQuillan, Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data., *Bioinformatics.* 28 (2012) 464–9. <https://doi.org/10.1093/bioinformatics/btr703>.
- [49] D. Takai, P.A. Jones, Comprehensive analysis of CpG islands in human chromosomes 21 and 22, *Proc. Natl. Acad. Sci. U. S. A.* 99 (2002) 3740–3745. <https://doi.org/10.1073/pnas.052410099>.
- [50] S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms, (n.d.). <https://doi.org/10.1093/molbev/msy096>.
- [51] S.I. Miller, S.M. Landfear, D.F. Wirth, Cloning and characterization of a Leishmania gene encoding a RNA spliced leader sequence., *Nucleic Acids Res.* 14 (1986) 7341–60. <https://doi.org/10.1093/nar/14.18.7341>.
- [52] S.M.R. Teixeira, B.M. Valente, Mechanisms Controlling Gene Expression in Trypanosomatids, in: M.S. da Silva, M.I.N. Cano (Eds.), *Frontiers in Parasitology (Volume 1) Molecular Cell Biology of Pathogenic*

Trypanosomatids, Bentham Science Publishers, Sharjah, UAE, 2017, pp. 261-290.

- [53] B. Cuypers, M.A. Domagalska, P. Meysman, G. de Muylder, M. Vanaerschot, H. Imamura, F. Dumetz, T.W. Verdonckt, P.J. Myler, G. Ramasamy, K. Laukens, J.-C. Dujardin, Multiplexed Spliced-Leader Sequencing: A high-throughput, selective method for RNA-seq in Trypanosomatids, *Sci. Rep.* 7 (2017) 3725. <https://doi.org/10.1038/s41598-017-03987-0>.
- [54] B. Langmead, S.L. Salzberg, Fast gapped-read alignment with Bowtie 2, *Nat. Methods.* 9 (2012) 357–359. <https://doi.org/10.1038/nmeth.1923>.
- [55] A. Mortazavi, B.A. Williams, K. McCue, L. Schaeffer, B. Wold, Mapping and quantifying mammalian transcriptomes by RNA-Seq, *Nat. Methods.* 5 (2008) 621–628. <https://doi.org/10.1038/nmeth.1226>.
- [56] E. Inbar, V.K. Hughitt, L.A.L. Dillon, K. Ghosh, N.M. El-Sayed, D.L. Sacks, The Transcriptome of *Leishmania major* Developmental Stages in Their Natural Sand Fly Vector, *MBio.* 8 (2017) e00029-17. <https://doi.org/10.1128/mBio.00029-17>.
- [57] M.W. Pfaffl, Quantification strategies in real-time PCR, in: S. A. Bustin (Ed.), *AZ of quantitative PCR*, International University Line, La Jolla, CA, 2004, pp. 89-113.
- [58] R.D.C.V. Da Silveira, M.S. Da Silva, V.S. Nunes, A.M. Perez, M.I.N. Cano, The natural absence of RPA1N domain did not impair *Leishmania amazonensis* RPA-1 participation in DNA damage response and telomere protection., *Parasitology.* 140 (2013) 547–559. <https://doi.org/10.1017/S0031182012002028>.
- [59] E. Medina-Acosta, G.A. Cross, Rapid isolation of DNA from trypanosomatid protozoa using a simple “mini-prep” procedure., *Mol. Biochem. Parasitol.* 59 (1993) 327–9. [https://doi.org/10.1016/0166-6851\(93\)90231-l](https://doi.org/10.1016/0166-6851(93)90231-l).
- [60] F.F. Conte, M.I.N. Cano, Genomic organization of telomeric and subtelomeric sequences of *Leishmania (Leishmania) amazonensis*, *Int. J. Parasitol.* 35 (2005) 1435–1443. <https://doi.org/10.1016/j.ijpara.2005.05.011>.
- [61] C.B.B. Lira, M.A. Giardini, J.L.S. Neto, F.F. Conte, M.I.N. Cano, Telomere biology of trypanosomatids: beginning to answer some questions, *Trends Parasitol.* 23 (2007) 357–362. <https://doi.org/10.1016/j.pt.2007.06.005>.
- [62] B. Balk, A. Maicher, M. Dees, J. Klermund, S. Luke-Glaser, K. Bender, B. Luke, Telomeric RNA-DNA hybrids affect telomere-length dynamics and senescence., *Nat. Struct. Mol. Biol.* 20 (2013) 1199–205. <https://doi.org/10.1038/nsmb.2662>.
- [63] K.J. Livak, T.D. Schmittgen, Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2– $\Delta\Delta$ CT Method, *Methods.* 25 (2001) 402–408. <https://doi.org/10.1006/meth.2001.1262>.
- [64] M.A. Giardini, M.F. Fernández, C.B.B. Lira, M.I.N. Cano, *Leishmania amazonensis*: Partial purification and study of the biochemical properties of the telomerase reverse transcriptase activity from promastigote-stage, *Exp. Parasitol.* 127 (2011) 243–248. <https://doi.org/10.1016/j.exppara.2010.08.001>.
- [65] X.H. Liang, A. Haritan, S. Uliel, S. Michaeli, trans and cis splicing in trypanosomatids: Mechanism, factors, and regulation, *Eukaryot. Cell.* 2 (2003) 830–840. <https://doi.org/10.1128/EC.2.5.830-840.2003>.
- [66] J.L. Reis-Cunha, H.O. Valdivia, D. Castanheira, Bartholomeu, Trypanosomatid Genome Organization and Ploidy, in: M.S. da Silva, M.I.N. Cano (Eds.), *Frontiers in Parasitology (Volume 1) Molecular Cell Biology of Pathogenic Trypanosomatids*, Bentham Science Publishers, Sharjah, UAE, 2017, pp. 61-103.
- [67] J. Greenwood, J.P. Cooper, Non-coding telomeric and subtelomeric transcripts are differentially regulated by telomeric and heterochromatin assembly factors in fission yeast, *Nucleic Acids Res.* 40 (2012) 2956–2963. <https://doi.org/10.1093/nar/gkr1155>.
- [68] S.G. Nergadze, B.O. Farnung, H. Wischnewski, L. Khorauli, V. Vitelli, R. Chawla, E. Giulotto, C.M. Azzalin, CpG-island promoters drive transcription of human telomeres, *RNA.* 15 (2009) 2186–2194. <https://doi.org/10.1261/rna.1748309>.
- [69] P.D.C. Ruy, N.M. Monteiro-Teles, R.D. Miserani Magalhães, F. Freitas-Castro, L. Dias, T.P. Aquino Defina, E.J. Rosas De Vasconcelos, P.J. Myler, A. Kaysel Cruz, Comparative transcriptomics in *Leishmania braziliensis*: disclosing differential gene expression of coding and putative noncoding RNAs across developmental stages, *RNA Biol.* 16 (2019) 639–660. <https://doi.org/10.1080/15476286.2019.1574161>.
- [70] H.P. Chu, C. Cifuentes-Rojas, B. Kesner, E. Aeby, H. goo Lee, C. Wei, H.J. Oh, M. Boukhali, W. Haas, J.T. Lee, TERRA RNA Antagonizes ATRX and Protects Telomeres, *Cell.* 170 (2017) 86-101.e16. <https://doi.org/10.1016/j.cell.2017.06.017>.
- [71] S.F. Altschul, T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST

and PSI-BLAST: A new generation of protein database search programs, *Nucleic Acids Res.* 25 (1997) 3389–3402. <https://doi.org/10.1093/nar/25.17.3389>.

- [72] J.H. Gommers-Ampt, F. Van Leeuwen, A.L.J. de Beer, J.F.G. Vliegthart, M. Dizdaroglu, J.A. Kowalak, P.F. Crain, P. Borst, β -d-glucosyl-hydroxymethyluracil: A novel modified base present in the DNA of the parasitic protozoan *T. brucei*, *Cell.* 75 (1993) 1129–1136. [https://doi.org/10.1016/0092-8674\(93\)90322-H](https://doi.org/10.1016/0092-8674(93)90322-H).
- [73] F. Van Leeuwen, M.C. Taylor, A. Mondragon, H. Moreau, W. Gibson, R. Kieft, P. Borst, β -d-glucosyl-hydroxymethyluracil is a conserved DNA modification in kinetoplastid protozoans and is abundant in their telomeres, *Proc. Natl. Acad. Sci. U. S. A.* 95 (1998) 2366–2371. <https://doi.org/10.1073/pnas.95.5.2366>.
- [74] P.A. Genest, B. ter Riet, T. Cijssouw, H.G.A.M. van Luenen, P. Borst, Telomeric localization of the modified DNA base J in the genome of the protozoan parasite *Leishmania*, *Nucleic Acids Res.* 35 (2007) 2116–2124. <https://doi.org/10.1093/nar/gkm050>.
- [75] F. Van Leeuwen, E.R. Wijsman, R. Kieft, G.A. Van Der Marel, J.H. Van Boom, P. Borst, Localization of the modified base J in telomeric VSG gene expression sites of *Trypanosoma brucei*, *Genes Dev.* 11 (1997) 3232–3241. <https://doi.org/10.1101/gad.11.23.3232>.
- [76] H.G.A.M. Van Luenen, C. Farris, S. Jan, P.A. Genest, P. Tripathi, A. Velds, R.M. Kerkhoven, M. Nieuwland, A. Haydock, G. Ramasamy, S. Vainio, T. Heidebrecht, A. Perrakis, L. Pagie, B. Van Steensel, P.J. Myler, P. Borst, Glucosylated hydroxymethyluracil, DNA base J, prevents transcriptional readthrough in *Leishmania*, *Cell.* 150 (2012) 909–921. <https://doi.org/10.1016/j.cell.2012.07.030>.
- [77] D.Z. Hazelbaker, S. Buratowski, Base J: Blocking RNA Polymerase's Way, *Curr. Biol.* 20 (2012) 960–62. <https://doi.org/10.1016/j.cub.2012.10.010>.
- [78] E. Briggs, K. Crouch, L. Lemgruber, C. Lapsley, R. McCulloch, Ribonuclease H1-targeted R-loops in surface antigen gene expression sites can direct trypanosome immune evasion, *PLoS Genet.* 14 (2018) e1007729. <https://doi.org/10.1371/journal.pgen.1007729>.
- [79] M.S. da Silva, M. Segatto, R.S. Pavani, F. Gutierrez-Rodrigues, V. da S. Bispo, M.H.G. de Medeiros, R.T. Calado, M.C. Elias, M.I.N. Cano, Consequences of acute oxidative stress in *Leishmania amazonensis*: From telomere shortening to the selection of the fittest parasites, *Biochim. Biophys. Acta - Mol. Cell Res.* 1864 (2017) 138–150. <https://doi.org/10.1016/j.bbamcr.2016.11.001>.
- [80] H.P. Chu, C. Cifuentes-Rojas, B. Kesner, E. Aeby, H. goo Lee, C. Wei, H.J. Oh, M. Boukhali, W. Haas, J.T. Lee, TERRA RNA Antagonizes ATRX and Protects Telomeres, *Cell.* 170 (2017) 86-101.e16. <https://doi.org/10.1016/j.cell.2017.06.017>.
- [81] M.S. da Silva, G.R. Cayres-Silva, M.O. Vitarelli, P.A. Marin, P.M. Hiraiwa, C.B. Araújo, B.B. Scholl, A.R. Ávila, R. McCulloch, M.S. Reis, M.C. Elias, Transcription activity contributes to the firing of non-constitutive origins in African trypanosomes helping to maintain robustness in S-phase duration, *Sci. Rep.* 9 (2019) 1–19. <https://doi.org/10.1038/s41598-019-54366-w>.

1 **Figure Legend**

2
3

4
5 **Figure 1. Detection of telomeric C-strand transcripts in total RNA from *L. major* amastigotes and**
6
7 **metacyclics.** A) Northern blot hybridized with DIG-labeled TELC probe (on the right) identified telomeric C-
8
9 strand transcripts in amastigotes (A, lane 1), early-transformed promastigotes (P1, lane 2), and in metacyclics
10
11 (M1, lane 3). The ethidium bromide-stained gel is on the left. The smeared signals in lanes 1 and 3 represent
12
13 transcripts of different sizes and abundance. B) total RNA used in A) were pre-treated with RNase A. C) The
14
15 data plotted on the graph represent the average \pm SD of three independent Northern blot assays shown in
16
17 A). D) RNA samples used in A were hybridized with a DIG-labeled TELG probe to identify RNAs transcribed
18
19 from the G-rich telomeric strand. The ethidium bromide-stained gel (on the left) is shown as the loading
20
21 control. MW, molecular weight marker used in A) 1kb DNA plus (Invitrogen), and in B) and D), molecular
22
23 weight marker DIG-VI (Roche).
24
25
26
27
28

29 **Figure 2. The abundance of transcripts originating from C-strand varies with *L. major* life stage and**
30
31 **continuous *in vitro* passages.** A) Total RNA (ethidium bromide-stained gel, on the left) was hybridized with
32
33 a DIG-labeled TELC probe (northern blot on the right) to identify transcripts from the C-rich telomeric strand
34
35 in promastigotes P2 (lane 1), P4 (lane 2), and P6 (lane 3). B) total RNA used in A) was pre-treated with
36
37 RNase A. C) Total RNA (ethidium bromide-stained gel on the left) was hybridized with a DIG-labeled TELC
38
39 probe (northern blot on the right) to identify transcripts from the C-rich telomeric strand in metacyclics M2
40
41 (lane 1), M4 (lane 2), and M6 (lane 3). D) total RNA used in C) was pre-treated with RNase A. E) The data
42
43 plotted on the graph represent the average \pm SD of three independent Northern blot assays shown in A) and
44
45 C). MW, molecular weight marker 1kb DNA plus (Invitrogen) used in the assays presented in A) and C); in B)
46
47 and D), molecular weight marker DIG-VI (Roche). EtBr, ethidium bromide.
48
49
50
51
52
53

54 **Figure 3. RPKM was calculated from each chromosome end termini of all three *Leishmania* life stages.**
55
56 TERRA expression level was obtained in RPKM, followed by a log₂ transformation of individual transcripts
57
58 originating from each chromosome end after analyses of the three SL-Seq libraries constructed from different
59
60 *L. major* life stages (PRO-promastigote, AMA-amastigote, and META-metacyclic). Expression of *L. major*
61
62 TERRA transcripts was plotted within the R environment (version 3.45.2) using the 'heatmap.2' function from
63
64
65

the 'gplots' library. The transcripts were ordered by chromosome using no hierarchical clustering.

Figure 4. TERRA is abundantly found in the nucleus of *L. major* metacyclic forms. In A and B, RNA-FISH (RNA) was performed in non-denaturing conditions using a PNA FITC-labeled telomeric C strand oligo probe. As the probe control, telomeric FISH (DNA) was done using the same PNA FITC-labeled telomeric C strand oligo probe. Vectashield mounting with DAPI was used to stain DNA in the nucleus (N) and kinetoplast (k). As the reaction control, fixed cells were treated with RNase A (RNase A) before hybridization. Images were acquired using a Nikon 80i fluorescence microscope and were superimposed using NIS elements software (v. Ar 3.10). Enlarged images of individual cells were displayed (bottom left squares). Bars 2 μ m.

Figure 5. Detection of TERRA Chr29R transcripts in promastigotes and metacyclics from different *in vitro* passages. Northern blots using total RNAs obtained from promastigotes P2, P4, and P6 and metacyclics M2, M4, and M6 were hybridized with a probe from Chr29R arm. The arrows in A) and B) point to the bands of mature TERRA Chr29R mRNA and the asterisks (*) the putative precursors of TERRA Chr29R. The data plotted on the graph represent the average \pm SD of three independent Northern blot assays shown in A) and B).

Figure 6. Telomere Restriction Fragment (TRF) analysis throughout the *L. major* developmental cycle and in parasites from continuous *in vitro* passages. A) A diagram showing a representation of the *L. major* chromosome end termini. TAS, telomere-associated sequences, containing the subtelomeric CSB elements. Black arrows represent the presence of subtelomeric *Afa*I restriction sites within the conserved CSB elements. The *L. major* telomeric repeat is represented as TTAGGG sequences, and the G-overhang corresponds to the G-rich single-strand protrusion at the chromosome end termini [45,57]. B) Southern blot of genomic DNA (1 μ g each) obtained from amastigotes (A), promastigotes (P1), and metacyclics (M1) originated from a single developmental cycle. C) Southern blot of genomic DNA (1 μ g each) obtained from promastigotes P2 and P6 and metacyclics M2 and M6. DNA samples in A and B were digested with *Afa* I, fractionated onto a 0.8% agarose gel, and hybridized with the DIG-TELC probe. The assays were developed by chemiluminescence using anti-DIG serum and CPD-Star (Roche). EtBr, ethidium bromide-stained gels in B and C were used as the load control. Molecular weight markers DIG-VII (Roche). In B) and C), base J-containing and base J-free telomeric fragments are signaled.

1 **Figure 7. Detection of telomeric R-loops in *L. major* life stages and parasites from continuous *in vitro***
2 **passages.** TERRA R-loops from amastigotes, promastigote (P2, P4, and P6), and metacyclics (M2, M4, and
3
4 P6) were immunoprecipitated with S9.6 antibody. Immunoprecipitated DNA from RNA-DNA hybrids was
5
6 amplified from the ends of Chr10R, Chr20L, and Chr29R arms using qPCR. The percentage of
7
8 immunoprecipitated RNA-DNA hybrids in each analyzed sample, relative to input (telomeric DNA), was
9
10 plotted on the graph.
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Table 1. *In silico* analysis of the 72 *L. major* chromosome end termini (R and L arms)

Chromosome arms	Chromosomes with CSB elements
L	4; 6; 10; 11; 14; 17; 18; 19; 20; 22; 23; 25; 26; 28; 30; 31; 32; 33; 34; 36
R	5; 6; 7; 9; 10; 12; 13; 14; 15; 16; 19; 20; 21; 23; 25; 26; 27; 28; 29; 30; 32; 33; 34; 35; 36

**L. major* genome contains 36 linear chromosomes (72 chromosome end termini)

Table 2. Chromosome end termini of all three *L. major* life stages that present SL signals located upstream of the subtelomeric regions.

Parasite life stage	Chromosome	
	L arm	R arm
Amastigote	1;2;3;5;7;10;14;15;16;17;18;19;20;21;22;23;25;26;27;28;29;30;31;32;33;34;35;36	1;2;5;6;7;10;11;12;14;15;17;18;19;21;23;24;25;26;28;29;30;31;32;33;34;35;36
Promastigote	1;3;7;10;15;16;19;20;22;26;27;30;31;32;33;34;35	1;2;5;6;7;10;12;14;18;19;21;24;25;26;29;30;31;32;33;34;35
Metacyclic	1;2;3;6;7;10;15;16;18;19;20;21;22;26;27;29;30;31;32;33;34;35;36	1;2;5;6;7;10;11;12;14;15;17;18;19;21;23;24;25;26;28;29;30;31;32;33;34;35;36

Table 3. Relative expression of TERRA in *L. major* life stages using RT-qPCR

LIFE STAGE	DATA	CHROMOSOME		
		Chr10R	Chr20L	Chr29R
Amastigote	Minimum	4,311	1,347	4,562
	Median	4,520	2,360	4,831
	Maximum	6,887	4,676	5,516
	Std. Error	0,5224	0,5819	0,1534
	P value (two tailed)	0,0345	0,0313	0,0313
	Significant (alpha=0.05)	Yes	Yes	Yes
Promastigote P2	Minimum	5,148	1,428	4,031
	Median	8,357	6,350	6,184
	Maximum	9,327	8,633	7,157
	Std. Error	0,7986	1,195	0,5022
	P value (two tailed)	0,0345	0,0313	0,0313
	Significant (alpha=0.05)	Yes	Yes	Yes
Promastigote P4	Minimum	6,025	4,596	3,754
	Median	7,260	5,448	4,694
	Maximum	7,444	6,291	6,409
	Std. Error	0,2818	0,3036	0,3870
	P value (two tailed)	0,0345	0,0313	0,0313
	Significant (alpha=0.05)	Yes	Yes	Yes
Promastigote P6	Minimum	5,228	2,122	3,076
	Median	7,266	4,318	4,697
	Maximum	7,274	6,458	5,936
	Std. Error	0,4305	1,231	0,7495
	P value (two tailed)	0,0345	0,1250	0,1250
	Significant (alpha=0.05)	Yes	No	No
Promastigote P12	Minimum	5,226	1,493	4,438
	Median	5,226	6,114	5,027
	Maximum	5,226	9,671	6,427
	Std. Error	0,0	2,001	0,3564
	P value (two tailed)	0,3458	0,1250	0,0313
	Significant (alpha=0.05)	No	No	Yes
Promastigote P24	Minimum	7,633	5,503	4,000
	Median	7,938	6,797	4,353
	Maximum	8,919	9,030	4,757
	Std. Error	0,2454	0,7457	0,09934
	P value (two tailed)	0,0345	0,1250	0,0313
	Significant (alpha=0.05)	Yes	No	Yes
Metacyclic P2	Minimum	0,4360	5,6300	-4,305
	Median	1,571	6,22	-1,787
	Maximum	2,705	8,3460	1,362
	Std. Error	0,6550	0,6117	1,026
	P value (two tailed)	0,0947	0,125	0,1563
	Significant (alpha=0.05)	No	No	No
Metacyclic P4	Minimum	2,242	Undetermined*	-2,341
	Median	3,451	Undetermined*	-2,205
	Maximum	4,659	Undetermined*	-0,7000
	Std. Error	0,6977	Undetermined*	0,3888
	P value (two tailed)	0,0947	Undetermined*	0,1250
	Significant (alpha=0.05)	No	Undetermined*	No
Metacyclic P6	Minimum	2,173	Undetermined*	-3,352
	Median	3,830	Undetermined*	-2,056
	Maximum	5,486	Undetermined*	0,01108
	Std. Error	0,9564	Undetermined*	0,8673
	P value (two tailed)	0,0947	Undetermined*	0,2500
	Significant (alpha=0.05)	No	Undetermined*	No

1		Minimum	2,251	Undetermined*	-1,577
2	Metacyclic P12	Median	5,285	Undetermined*	-1,120
3		Maximum	8,319	Undetermined*	-0,7050
4		Std. Error	1,752	Undetermined*	0,2238
5		P value (two tailed)	0,0947	Undetermined*	0,1250
6		Significant (alpha=0.05)	No	Undetermined*	No
7		Minimum	5,341	Undetermined*	0,4030
8	Metacyclic P24	Median	5,769	Undetermined*	1,290
9		Maximum	6,197	Undetermined*	2,166
10		Std. Error	0,2471	Undetermined*	0,3755
11		P value (two tailed)	0,0947	Undetermined*	0,1250
12		Significant (alpha=0.05)	No	Undetermined*	No

*Undetermined means that it was not possible to detect any amplification

13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

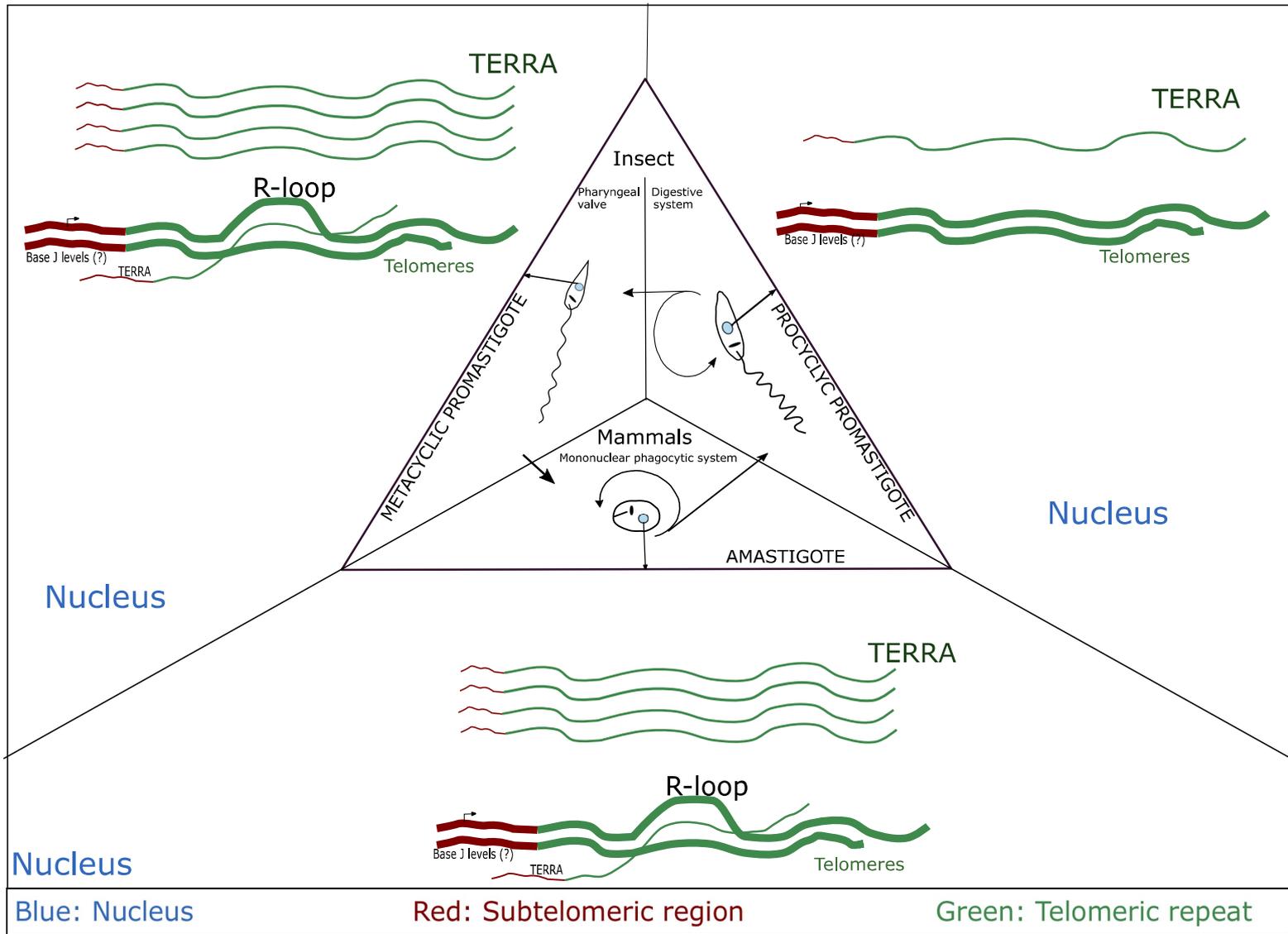


Fig. 1

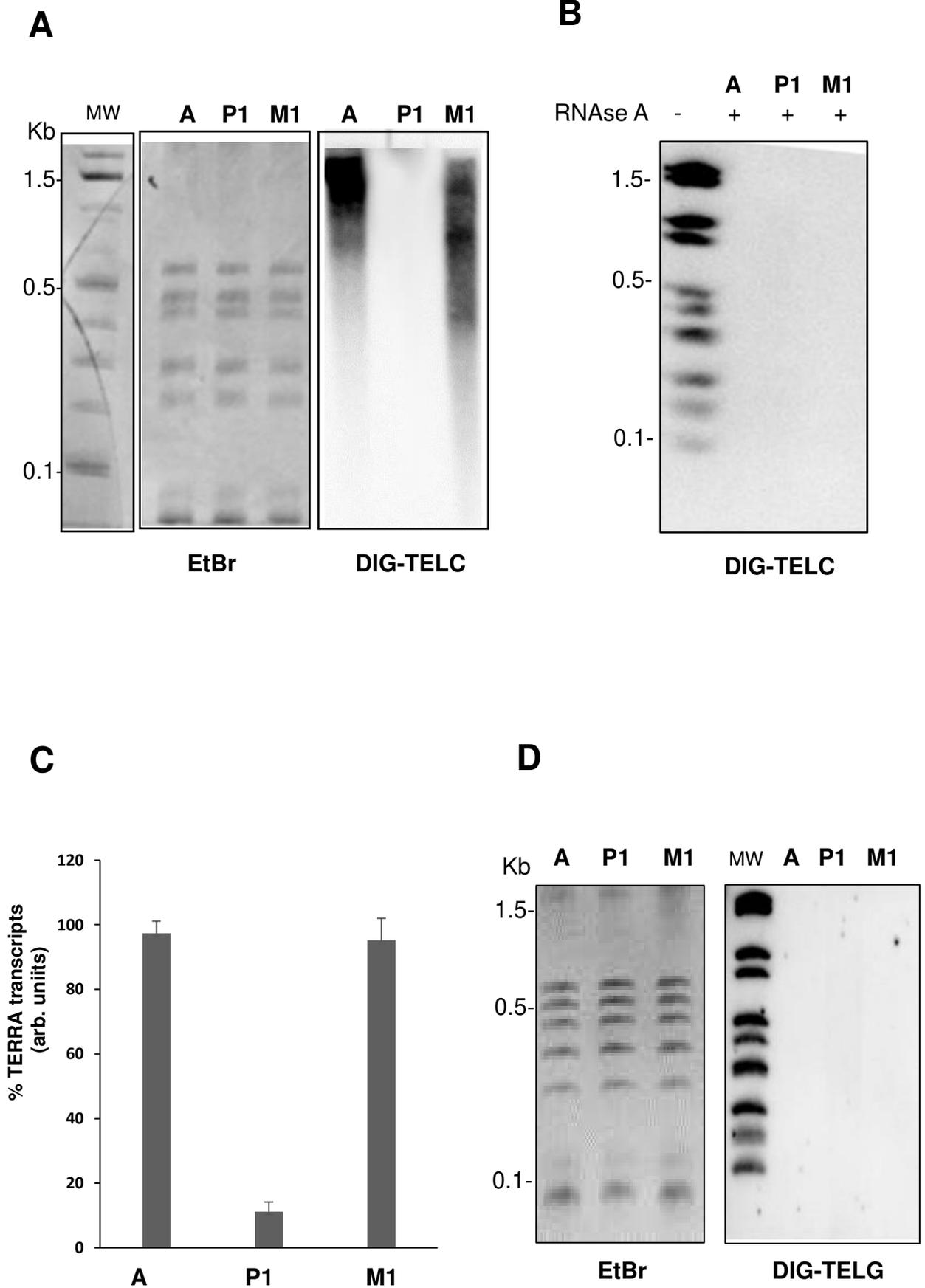


Fig. 2

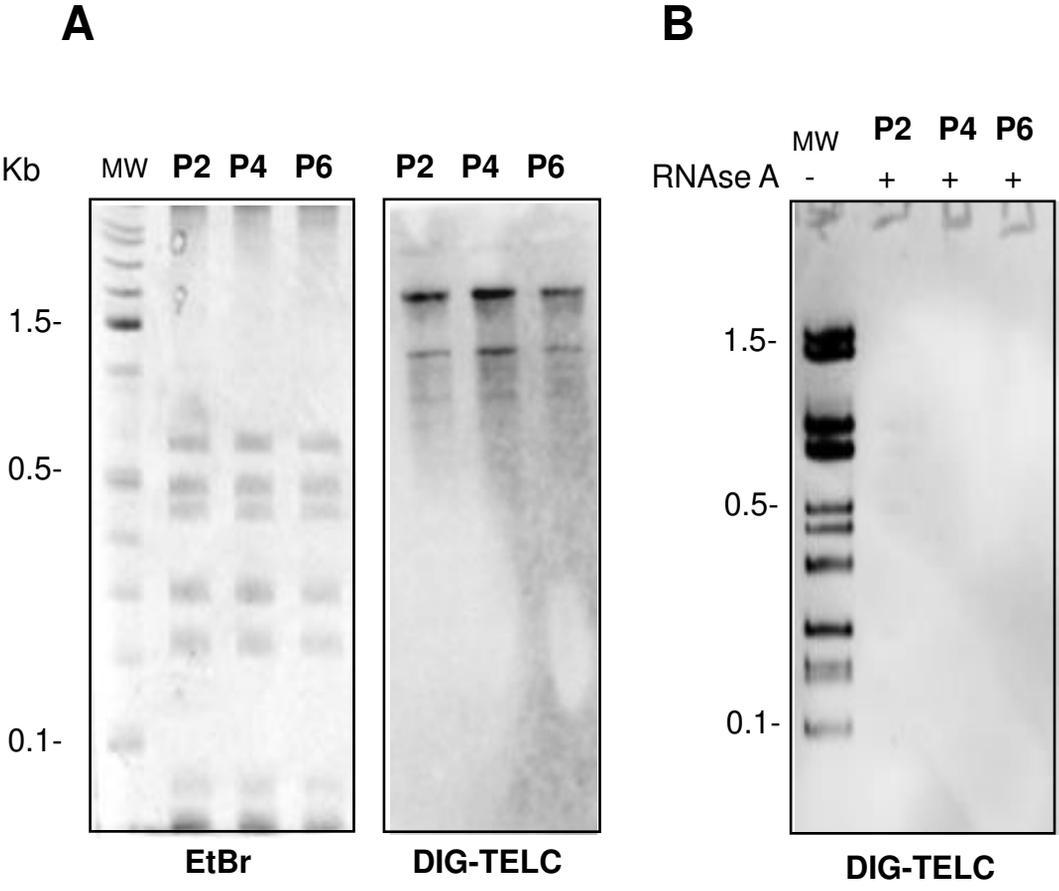
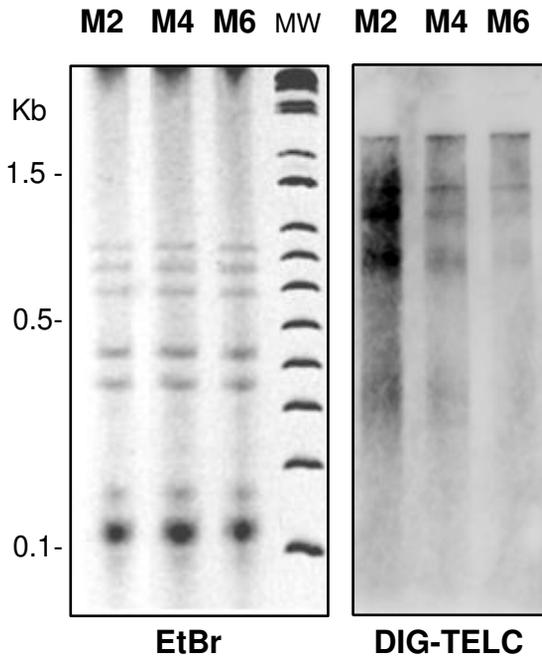
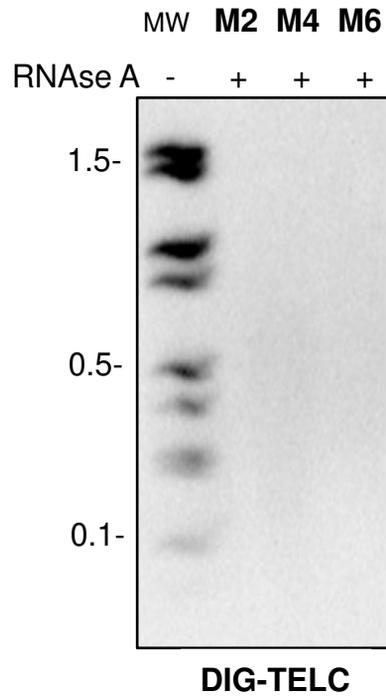


Fig. 2

C



D



E

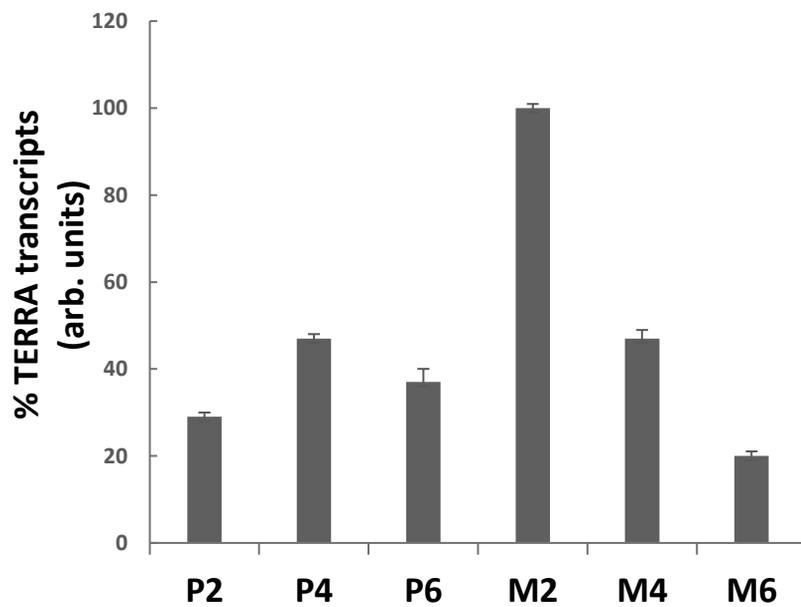


Fig. 3

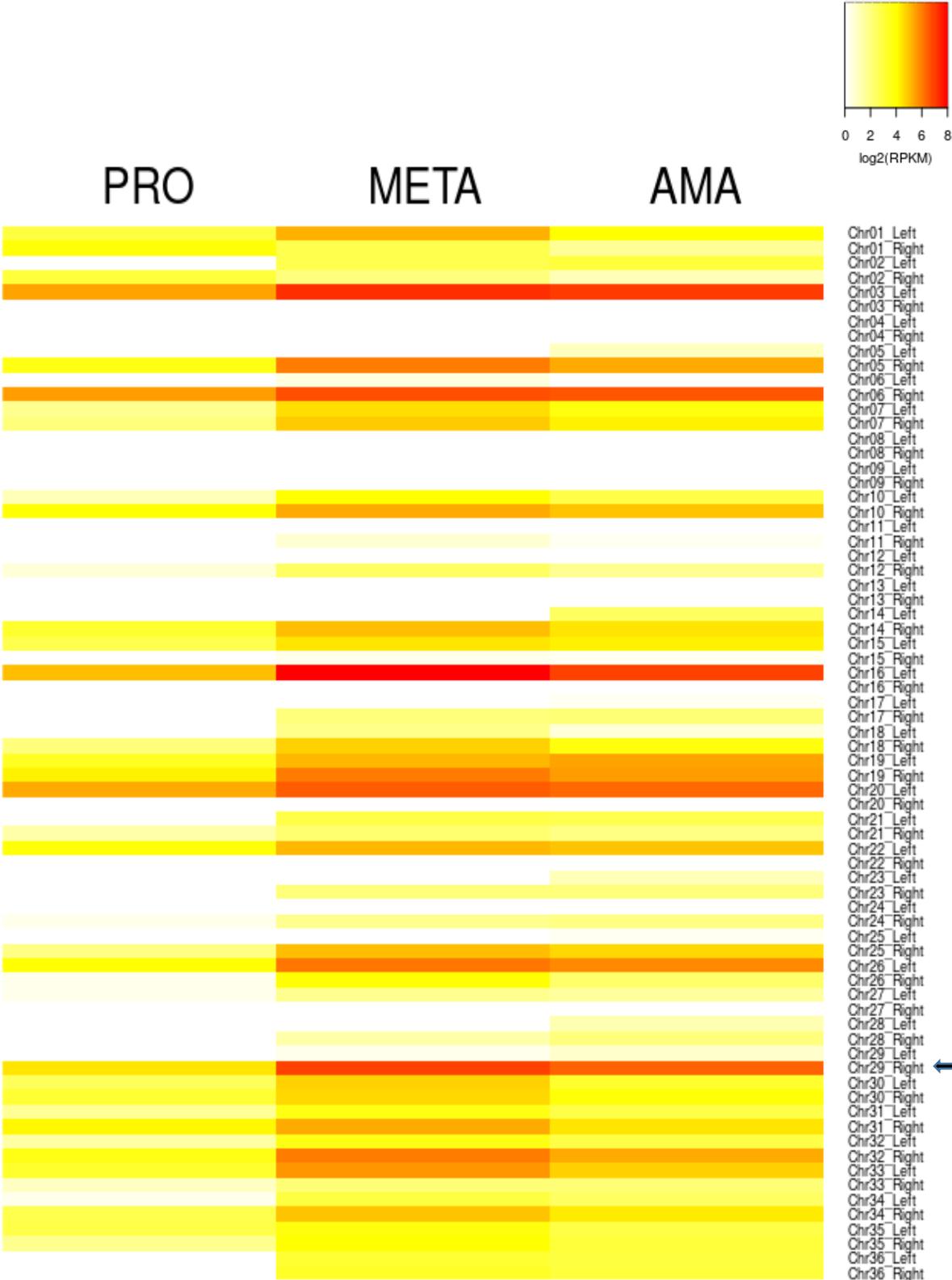


Fig. 4

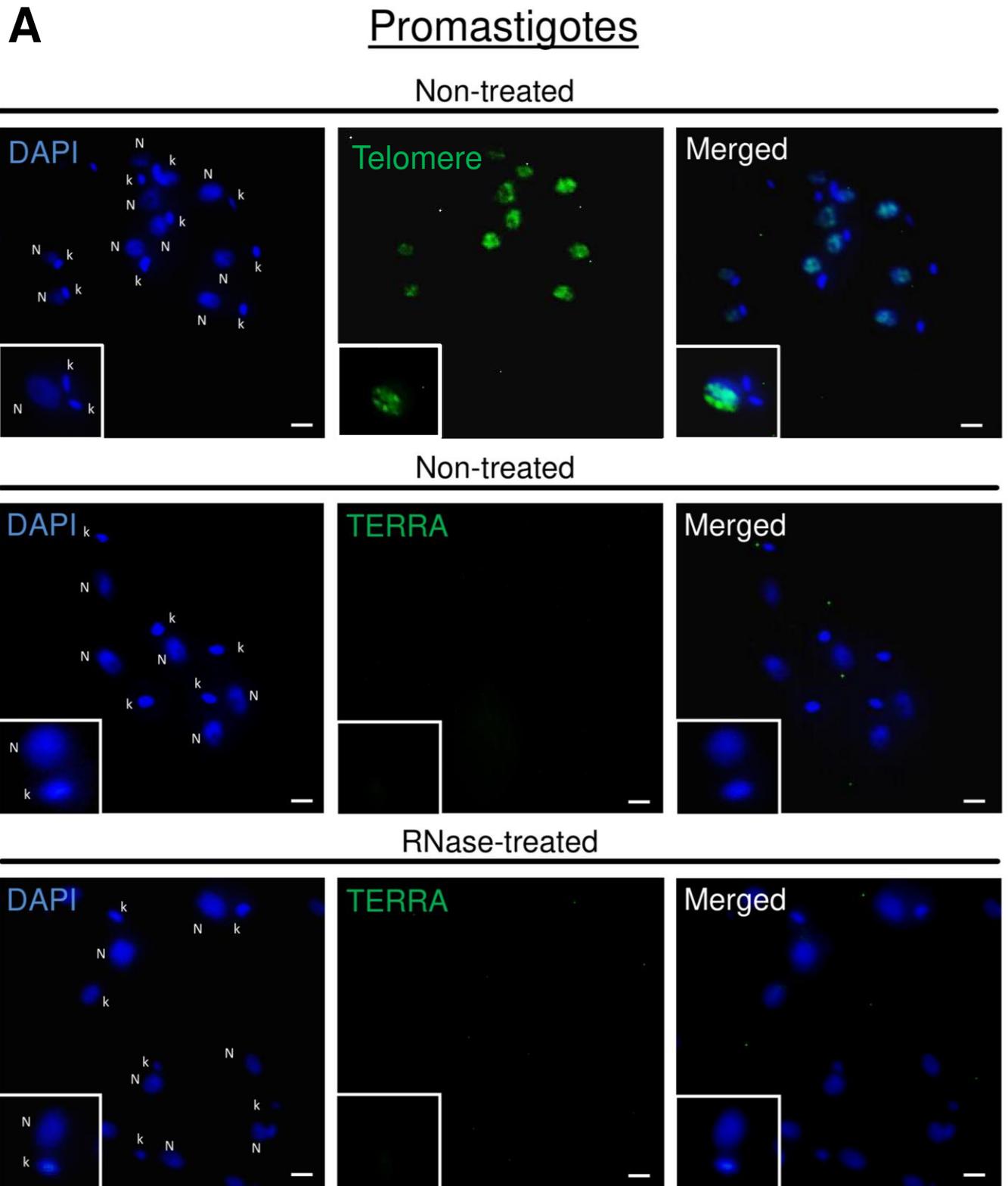
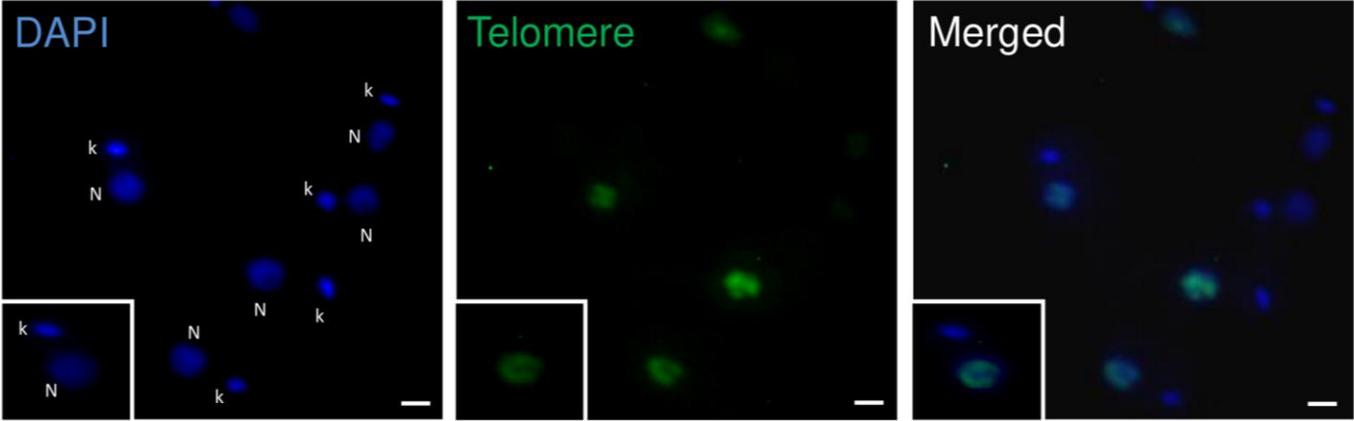


Fig. 4

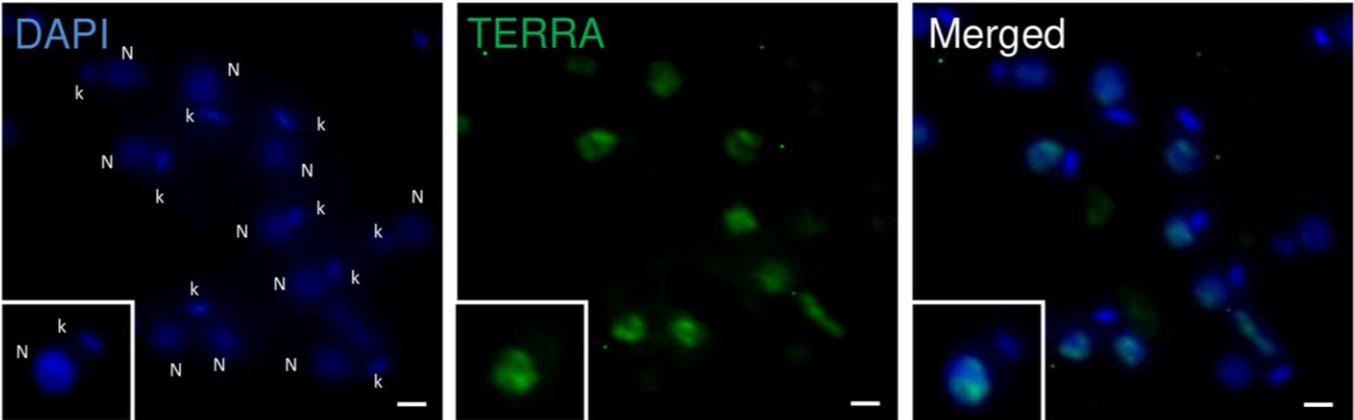
B

Metacyclics

Non-treated



Non-treated



RNase-treated

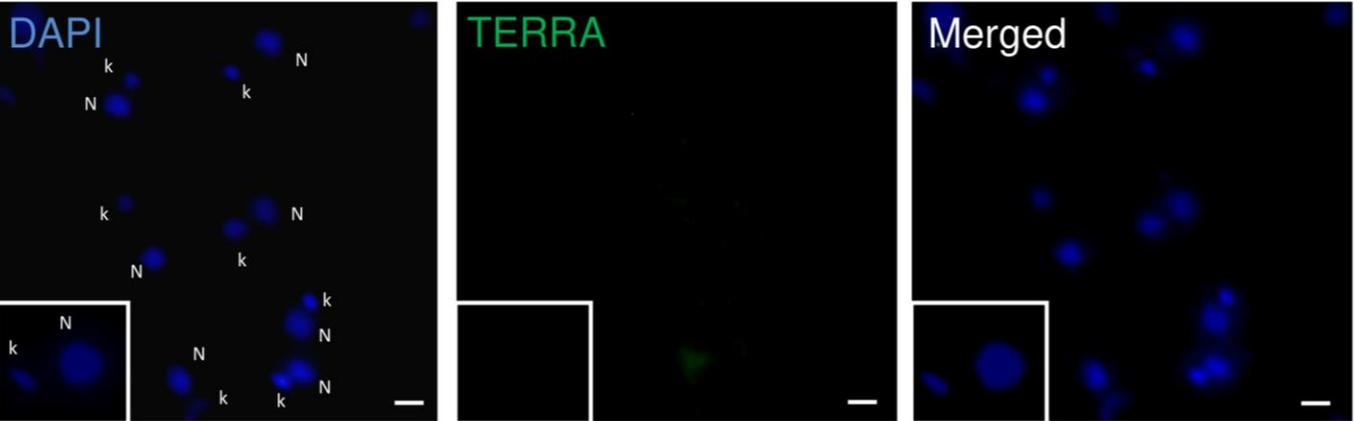


Fig. 5

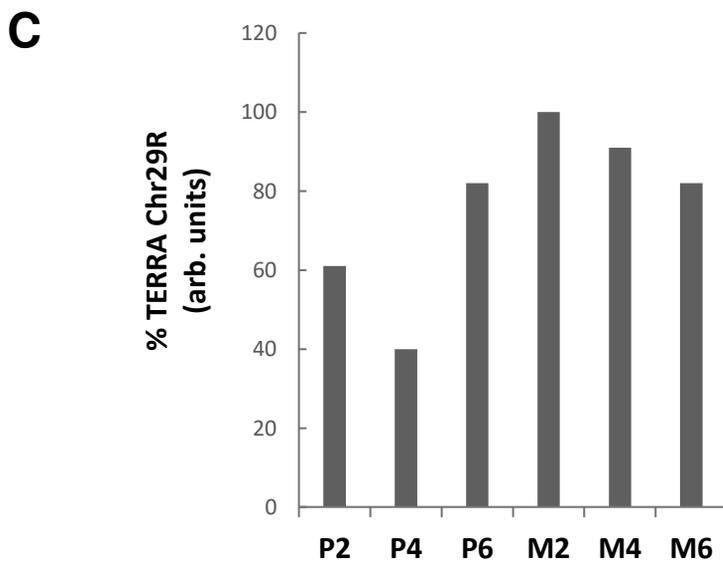
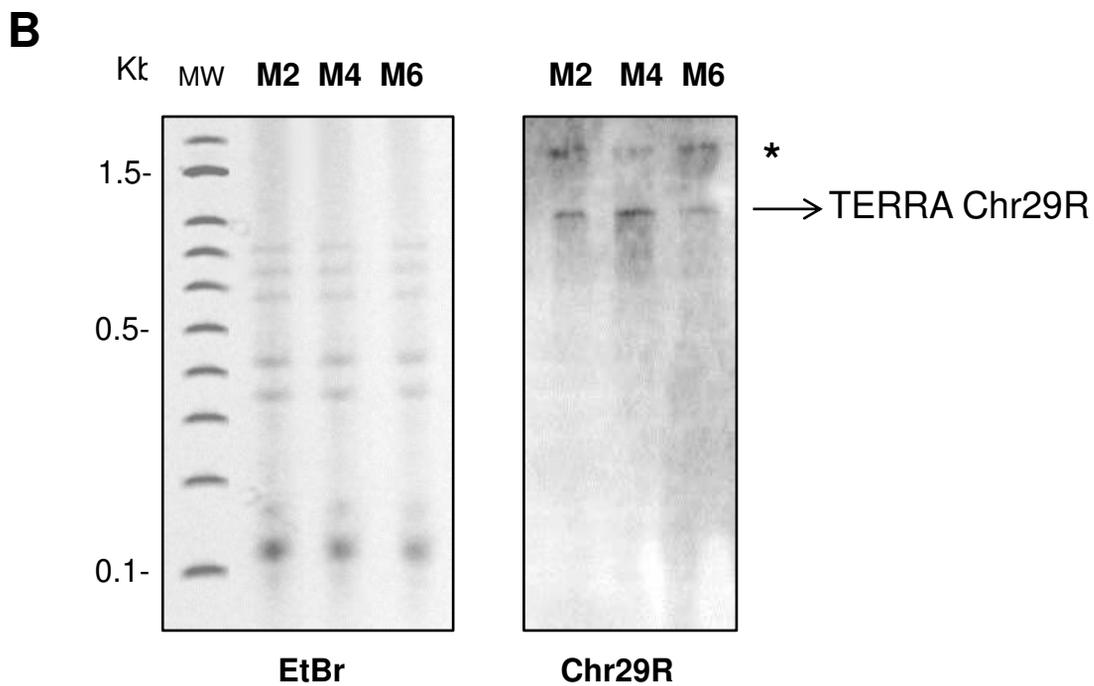
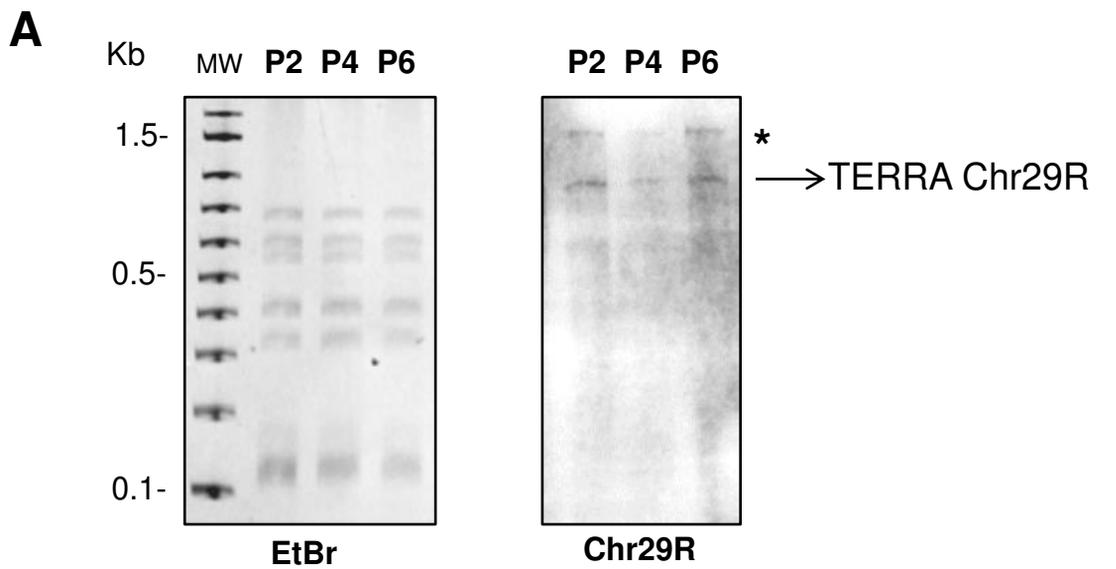
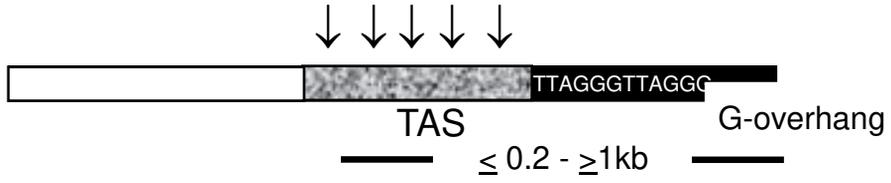


Fig. 6

A



B

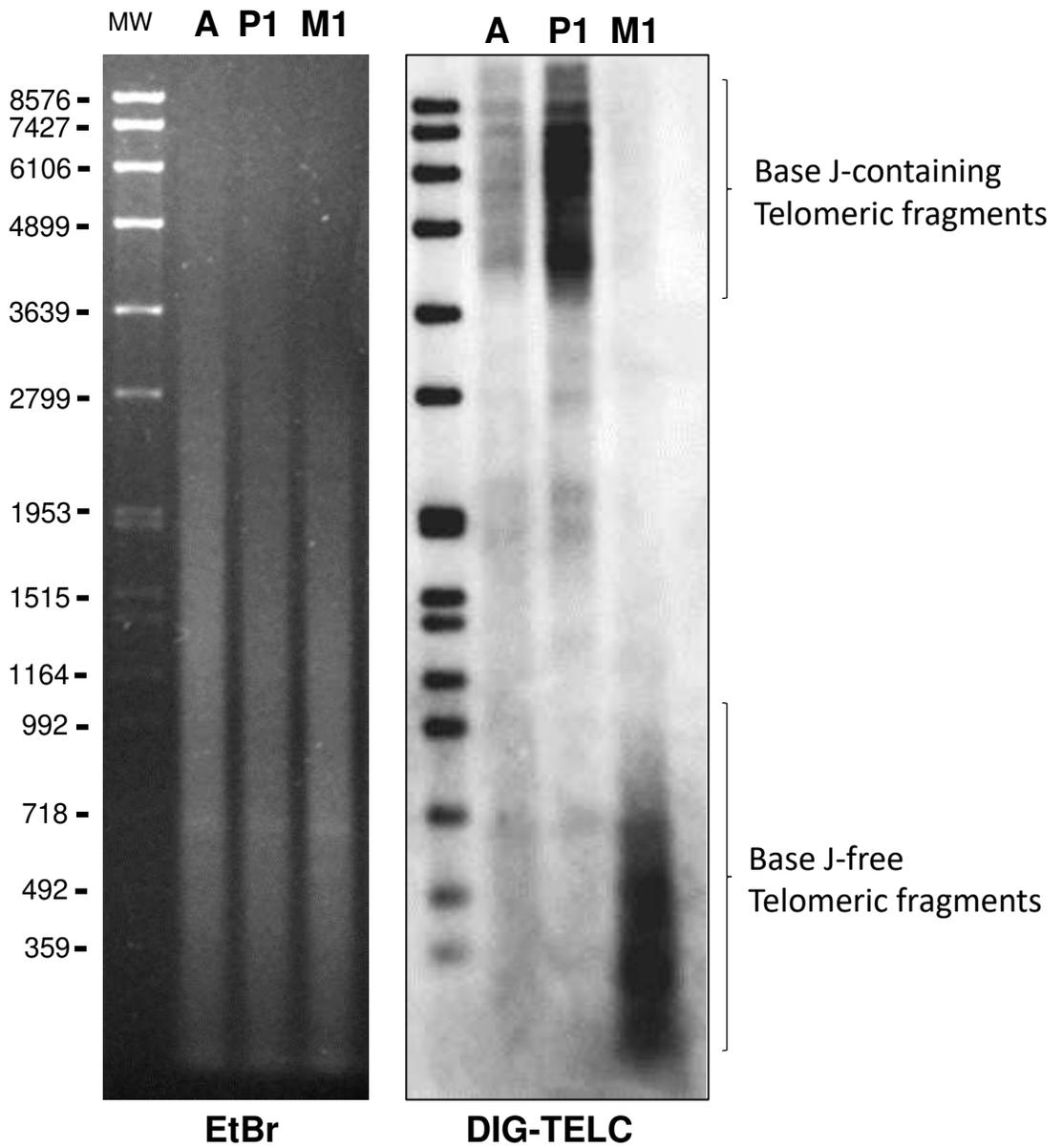


Fig. 6

C

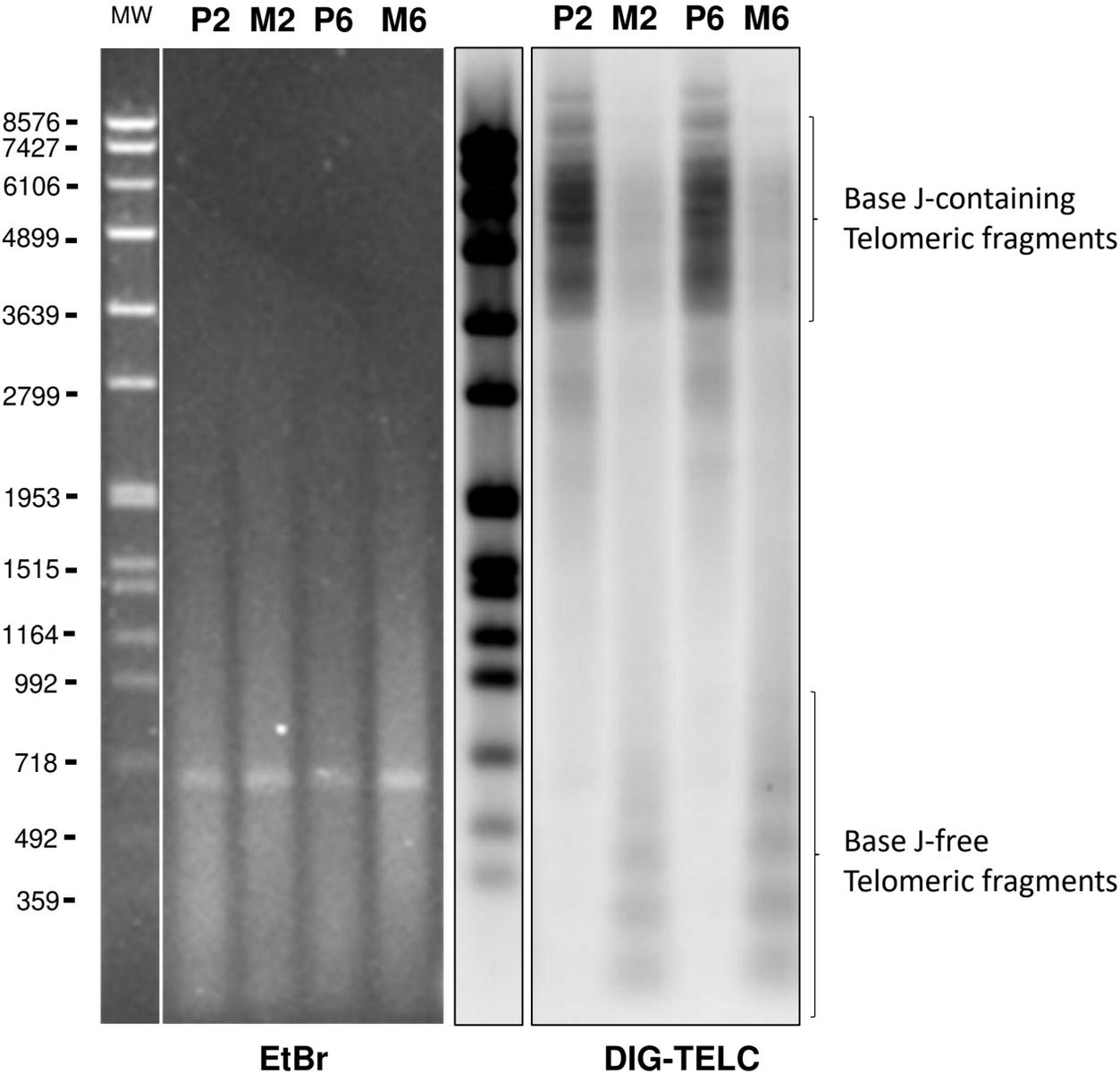
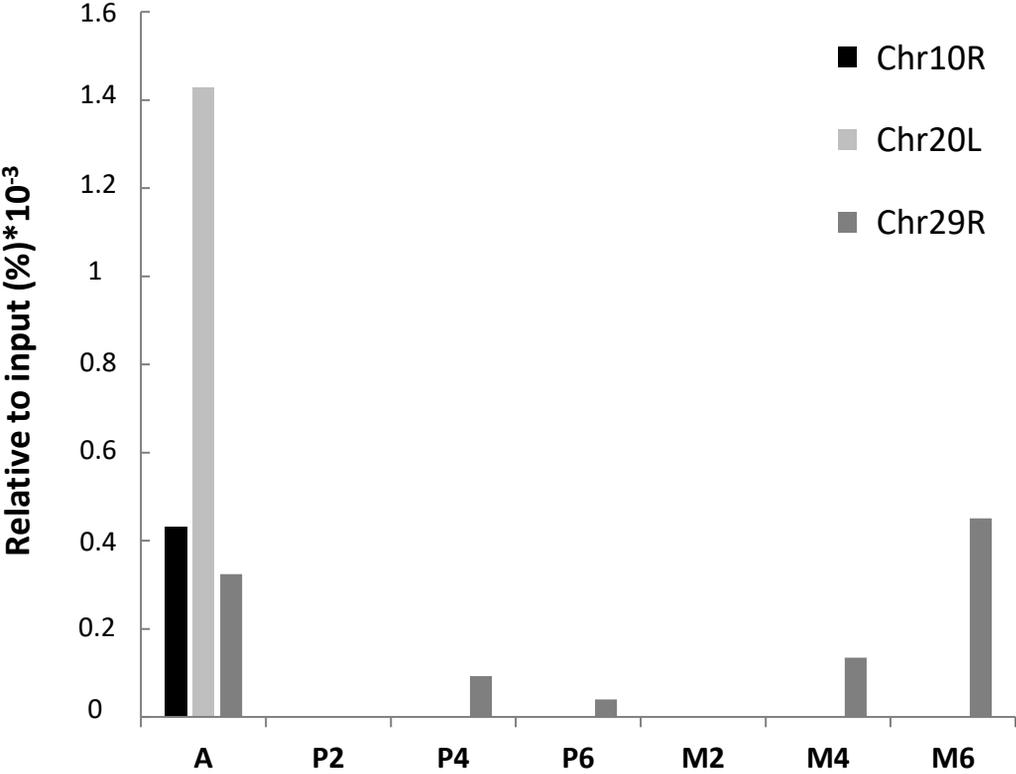


Fig. 7



Author Contributions

Edna Gicela Ortiz Morea: Investigation, Methodology, Validation, Writing- Original draft preparation

Elton José Rosas de Vasconcelos: Data Curation, Validation, Software, Writing-Original draft preparation, Writing- Reviewing and Editing

Cristiane Alves de Santis: Investigation, Methodology, Validation

Selma Giorgio: Resources, Visualization, Writing- Reviewing and Editing

Peter Myler: Resources, Visualization, Writing- Reviewing and Editing

Helio Langoni: Resources, Visualization, Writing- Reviewing and Editing

Claus Maria Azzalin: Conceptualization, Supervision, Writing- Reviewing and Editing

Maria Isabel Nogueira Cano: Conceptualization, Funding acquisition, Project administration, Supervision, Writing- Reviewing and Editing

Article: Exploring TERRA during *Leishmania major* developmental cycle and continuous *in vitro* passages



[Click here to access/download](#)

Supplementary Material

Moreal et al_IJBM_2020 Suppl Material.docx

