

de Almeida Marques, Catarina (2015) *Initiation of nuclear DNA replication in Trypanosoma brucei and Leishmania.* – Appendices. PhD thesis.

http://theses.gla.ac.uk/6518/

Copyright and moral rights for this thesis are retained by the author

A copy can be downloaded for personal non-commercial research or study

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Glasgow Theses Service http://theses.gla.ac.uk/ theses@gla.ac.uk Accompanying material from the thesis entitled "Initiation of Nuclear DNA Replication in *Trypanosoma brucei* and *Leishmania*", submitted by Catarina de Almeida Marques in fulfilment of the requirements for the Degree of Doctor in Philosophy.

Wellcome Trust Centre for Molecular Parasitology; Institute of Infections, Immunity and Inflammation; College of Medical, Veterinary and Life Sciences

University of Glasgow

April 2015

Table of Contents

Table of Contents	2
List of Tables	3
List of Figures	4
7 Appendices	6
7.1 Conservation of the putative origin recognition complex factors within	
the kinetoplastid group	7
7.2 Protein sequence Alignments	9
7.2.1 TbORC1/CDC6 alignment with other organisms' Orc1 subunits	9
7.2.2 TbORC1B alignment with model eukaryotes' Cdc6 proteins1	3
7.2.3 TbORC4 alignment against Orc4 subunits of model eukaryotes1	7
7.2.4 Tb7980 alignment against Orc5 subunits of model eukaryotes20	C
7.2.5 Tb3120 alignment against model eukaryotes Orc2 subunits2	2
7.3 RNAi in PCF cells	7
7.3.1 Extra TbORC1/CDC6 and Tb3120 RNAi clones	7
7.3.2 Flow cytometry analysis of the effects on cell cycle resultant from	
gene-targeted RNAi induction29	9
7.4 Cloning the constructs	8
7.4.1 Confirmation enzymatic digestion of the plasmids for endogenous tag	ž
of Lister 427 cell lines	8
7.4.2 Plasmid map of the parental construct using for gene deletion4	6
7.5 Cell cycle mRNA levels of TbORC1/CDC6 interacting factors	6
7.6 TbORC1/CDC6 and other factors subcellular localisation	7
7.7 Gel Filtration of TbORC1/CDC6 12 myc	2
7.8 Mapping Origins of Replication in BSF cells	3
7.9 Origin and non-origins size in T. brucei, L. major, and L. mexicana 5	6
List of References	2

List of Tables

Table 7-1. T. brucei distance in Kbp between the first two genes within	the SSRs,
with or no origin activity detected by MFA-seq	
Table 7-2. L. major distance in Kbp between the first two genes within	the SSRs,
with or no origin activity detected by MFA-seq	
Table 7-3. L. mexicana distance in Kbp between the first two genes with	hin the
SSRs, with or no origin activity detected by MFA-seq	

List of Figures

Figure 7.1. The putative ORC factors are syntenic between different
Figure 7.2. Alignment of TbORC1/CDC6 with Orc1 subunits of a range of model
organisms
Figure 7.3. Alignment of TbORC1B with Cdc6 proteins of a range of model
eukaryotes
Figure 7.4. IbORC4 alignment with Orc4 subunits of a range of eukaryotes 18
Figure 7.5. Ib/980 alignment with Orc5 subunits of model eukaryotes
Figure 7.6. Alignment of 1b3120 with Orc2 subunits from model eukaryotes23
Figure 7.7. Effect of TbORC1/CDC6 depletion by induction of specific gene
Figure 7.8 Effect of Th2120 deplotion by induction of specific gone targeted
PNAi over time
Figure 7.9, 2013 cell line, 12 h, 24 h and 48 h time points 20
Figure 7.10, 2013 cell line, 72 h, 96 h and 120 h time points. $\dots \dots \dots$
Figure 7.11, 2013 cell line, $1/1$ h and 168 h time points
Figure 7.12 ThOPC1/CDC6 PNAi Clb cell line 12 h 24 h and 48 h time points 31
Figure 7.13 ThORC1/CDC6 RNAi Clb cell line, 72 h, 96 h and 120 h time points.
31
Figure 7.14. TbORC1/CDC6 RNAi Clb cell line. 144 h time points
Figure 7.15. TbORC4 RNAi Cla cell line. 12 h. 24 h and 48 h time points 32
Figure 7.16. TbORC4 RNAi Cla cell line, 72 h, 96 h and 120 h time points 33
Figure 7.17. TbORC4 RNAi Cla cell line, 144 h time points
Figure 7.18. TbORC4 RNAi Clb cell line, 12 h, 24 h and 48 h time points 34
Figure 7.19. TbORC4 RNAi Clb cell line, 72 h, 96 h and 120 h time points 34
Figure 7.20. TbORC4 RNAi Clb cell line, 144 h time points
Figure 7.21. Tb3120 RNAi Clb cell line, 24 h, 48 h and 72 h time points35
Figure 7.22. Tb3120 RNAi Clb cell line, 96 h and 120 h time points
Figure 7.23. Tb3120 RNAi Clb cell line, 144 h and 168 h time points
Figure 7.24. TbORC1B RNAi cell line, 6 h, 12 h and 24 h time points
Figure 7.25. TbORC1B RNAi cell line, 48 h, 72 h and 96 h time points
Figure 7.26. TbORC1B RNAi cell line, 120 h and 144 h time points
Figure 7.27. Alignment of the DNA sequences of the TbORC1/CDC6 gene from the
reference genomes of the TREU 927 and Lister 427 strains of T. brucei 39
Figure 7.28. Alignment of the DNA sequences of the TbORC1B gene from the
reference genomes of the TREU 927 and Lister 427 strains of 1. brucei 40
Figure 7.29. Alignment of the DNA sequences of the TDURC4 gene from the
For the T
rigure 7.30. Alignment of the DNA sequences of the 107960 gene from the
Figure 7 31 Alignment of the DNA sequences of the Tb3120 gene from the
reference genomes of the TRFU 927 and Lister 427 strains of T. brucei 43
Figure 7.32. (continue next page, together with description)
Figure 7.33. Enzymatic digestion of the plasmids used for endogenous tagging of
proteins with N- or C-terminal 12-myc tag, to be used for the transfection of
Lister 427 cell lines
Figure 7.34. Original plasmid used for the generation of the constructs used for
gene deletion described in Chapter 3
Figure 7.35. mRNA levels of TbORC1/CDC6, TbORC1B, TbORC4, Tb7980, Tb3120
and Tb1120 throughout the cell cycle46

7 Appendices

7.1 Conservation of the putative origin recognition complex factors within the kinetoplastid group



Figure 7.1. The putative ORC factors are syntenic between different kinetoplastid species. Each panel (A-F) represents a 10 Kbp window of the parasite's genome (top row shows the *T. brucei* chromosome as well as the chromosome coordinates). The *T. brucei* (Tb927) gene of interest is highlighted in yellow, and the respective orthologues in *L. major* (Lmaj), *L. mexicana* (Lmex) and *T. cruzi* (TcNEL) are shown in the "syntenic sequences and genes" section. A) shows the orthologues of TbORC1/CDC6; B) orthologues of TbORC1B; C) orthologues of TbORC4; D) orthologues of Tb7980; E) orthologues of Tb3120; and F) orthologues of Tb1120. Images retrieved from TriTrypDB version 8.0. *(continued below)*





Figure 7.1. (continued).

7.2 Protein sequence Alignments

7.2.1 TbORC1/CDC6 alignment with other organisms' Orc1 subunits

The protein sequences for TbORC1/CDC6 (Tb927.11.7216), as well as for the Trypanosoma cruzi (TcCLB.508239.10) and Leishmania major orthologues (LmjF.28.0030) were retrieved from the TriTrypDB database (http://tritrypdb.org/tritrypdb/). Sequences of the Orc1 subunits of all the other herein represented organisms were obtained from the NCBI Protein Database (http://www.ncbi.nlm.nih.gov/protein), although the protein identification numbering refers to styles used by different databases such as GenBank, InterPro, and NCBI Reference Sequence. Orc1 subunits of model eukaryotes such as human, *Homo sapiens* (HmOrc1, AAC50325.1), domestic mouse, Mus musculus (MmOrc1, NP_035145.2), fruit fly, Drosophila melanogaster (DmOrc1, NP 477303.1), model plant, Arabidopsis thaliana (AtOrc1, NP_567440.1), and the baking yeast, Saccharomyces cerevisiae (ScOrc1, NP_013646.1), were used for the alignment. In addition, the sequence for the archaea Aeropyrum pernix (ApOrc1, BAA79440.2) Orc1/Cdc6 protein was also used, as this was the sequence used in the first published work on TbORC1/CDC6 and TcORC1/CDC6 factors (Godoy et al., 2009). Alignment was performed and graphically represented using CLC genomics, version 7.5.1 (QIAGEN Aarhus A/S), with a gap open cost of 10.0, a gap extension cost of 1.0, an end gap cost of "as any other", and the very accurate (slow) option for alignment. Results are depicted in Figure 7.2. Regions of high conservation were used to deduce potential motif regions and thus produce the domain and motif schematic representation of TbORC1/CDC6, shown and discussed in Chapter 3, Figure 3.1.



Figure 7.2. Alignment of TbORC1/CDC6 with Orc1 subunits of a range of model organisms. Conservation is depicted by the colour gradient, and sequences are ordered by similarity. Hs, *H. sapiens*; Mm, *M. musculus*; Dm, *D. melanogaster*; At, *A. thaliana*; Ap, *A. pernix*; Sc, *S. cerevisiae*; Tc, *T. cruzi*; Tb, *T. brucei*; Lm, *L. major*. (----) underlines the WHD domain. *(continued below)*





100

Conservation

%

'n





100

Conservation

%

C



Figure 7.2. (continued)

7.2.2 TbORC1B alignment with model eukaryotes' Cdc6 proteins

Protein sequences were retrieved from the appropriate databases, as described in the previous section. For simplicity, the *T. cruzi* and *L. major* orthologues of TbORC1B (Tb927.9.2030) were here represented as TcORC1B (TcCLB.507939.14) and LmORC1B (LmjF.26.2210), respectively. The following Cdc6 protein sequences from model eukaryotes were used for the alignment: *M. musculus* (MmCdc6, NP_035929.1), *H. sapiens* (HsCdc6, NP_001245.1), *D. melanogaster* (DmCdc6, AAF50387.1), *A. thaliana* (AtCdc6, AEC08293.1), and *S. cerevisiae* (ScCdc6, CAA89490.1). In addition, the Cdc6 protein of *Clonorchis sinensis* (CsCdc6, GAA29103.2) was also used, as this was the first hit obtained when using TbORC1B as a query in blastp (see Chapter 3, section 3.1.2). The alignment was performed as described in the previous section, and the results are represented in Figure 7.3. Regions of high conservation were used to deduce potential motif regions and thus produce the domain and motif schematic representation of TbORC1B, shown and discussed in Chapter 3, Figure 3.1.



Figure 7.3. Alignment of TbORC1B with Cdc6 proteins of a range of model eukaryotes. Conservation is depicted by the colour gradient, and sequences are ordered by similarity. Mm, *M. musculus*; Hs, *H. sapiens*; Dm, *D. melanogaster*; Cs, *Clonorchis sinensis*; At, *A. thaliana*; Sc, *S. cerevisiae*; Tb, *T. brucei*; Tc, *T. cruzi*; Lm, *L. major*. (----) underlines the WHD domain.(*continued below*)





100

Conservation

8

'n





Figure 7.3. (continued).

7.2.3 TbORC4 alignment against Orc4 subunits of model eukaryotes

Protein sequences were retrieved from the appropriate databases, as described in section 7.2.1. For simplicity, the *T. cruzi* and *L. major* orthologues of TbORC4 (Tb927.10.13380) were here represented as TcORC4 (TcCLB.506357.20) and LmORC4 (LmjF.18.0720), respectively. The following Orc4 protein sequences from model eukaryotes were used for the alignment: *D. melanogaster* (DmOrc4, AAD39473.1), *H. sapiens* (HsOrc4, NP_001177808.1), *M. musculus* (MmOrc4, CAA76188.1), *A. thaliana* (AtOrc4, AEC05404.1), and *S. cerevisiae* (ScOrc4, AAB68149.1). The Orc4 subunit from *Musca domestica* (MdOrc4, XP_005179312.1) was also used in the alignment, as this was one of the top hits in the blastp analysis performed using TbORC4 as query (see Chapter 3, section 3.1.2). The alignment was performed as described in section 7.2.1, and the results are represented in Figure 7.4. Regions of high conservation were used to infer potential motif regions and thus generate the domain and motif schematic representation of TbORC4, shown and discussed in Chapter 3, Figure 3.1.



Figure 7.4. TbORC4 alignment with Orc4 subunits of a range of eukaryotes. Conservation is depicted by the colour gradient, and sequences are ordered by similarity. Dm, *D. melanogaster*; Md, *Musca domestica*; Hs, *H. sapiens*; Mm, *M. musculus*; At, *A. thaliana*; Sc, *S. cerevisiae*; Tc, *T. cruzi*; Tb, *T. brucei*; Lm, *L. major*. (----) underlines the AAA+ ATPase domain. (continued below)





0

Figure 7.4. (continued).



Figure 7.4. (continued).

7.2.4 Tb7980 alignment against Orc5 subunits of model eukaryotes

Protein sequences were retrieved from the appropriate databases, as described in section 7.2.1. *T. cruzi* and *L. major* orthologues of Tb7980 (Tb927.10.7980) were here represented as Tc7980 (TcCLB.506247.280) and Lm7980 (LmjF.36.6700), respectively. The following Orc5 protein sequences from model eukaryotes were used for the alignment: *H. sapiens* (HsOrc5, NP_002544.1), *M. musculus* (MmOrc5, NP_036089.1), *D. melanogaster* (DmOrc5, NP_477132.1), *A. thaliana* (AtOrc5, NP_194720.2), and *S. cerevisiae* (ScOrc5, CAA65483.1). The alignment was performed as described in section 7.2.1, and the results are represented in Figure 7.5. Regions of high conservation were used to infer potential motif regions and thus generate the domain and motif schematic representation of Tb7980, shown and discussed in Chapter 3, Figure 3.1.

100

Conservation

%



Figure 7.5. Tb7980 alignment with Orc5 subunits of model eukaryotes.

Conservation is depicted by the colour gradient. Hs, H. sapiens; Mm, M. musculus; Dm, D. melanogaster; At, A. thaliana; Sc, S. cerevisiae; Tb, T. brucei; Tc, T. cruzi; Lm, L. major. (continued below)





7.2.5 Tb3120 alignment against model eukaryotes Orc2 subunits

Protein sequences were retrieved from the appropriate databases, as described in section 7.2.1. *T. cruzi* and *L. major* orthologues of Tb3120 (Tb927.9.4530) were here represented as Tc3120 (TcCLB.511585.90) and Lm3120 (LmjF.01.0660), respectively. The following Orc2 protein sequences from model eukaryotes were used for the alignment: *H. sapiens* (HsOrc2, NP_006181.1), *M. musculus* (MmOrc2, Q60862.1), *D. melanogaster* (DmOrc2, AAF99606.1), *A. thaliana* (AtOrc2, AEC09416.1), and *S. cerevisiae* (ScOrc2, CAA85003.1). The alignment was performed as described in section 7.2.1, and the results are represented in Figure 7.6. Regions of high conservation were used to infer potential motif regions and thus generate the domain and motif schematic representation of Tb3120, shown and discussed in Chapter 3, Figure 3.1.



Figure 7.6. Alignment of Tb3120 with Orc2 subunits from model eukaryotes.

Conservation is depicted by the colour gradient. Hs, *H. sapiens*; Mm, *M. musculus*; Dm, *D. melanogaster*; At, *A. thaliana*; Sc, *S. cerevisiae*; Tb, *T. brucei*; Tc, *T. cruzi*; Lm, *L. major*. Potential Walker A and Walker B motifs are shown within the dashed green boxes. *(continued below)*

HsOrc2 (NP_006181.1) - - RDVQESL - KNGSATG - - - - GGNKVYSFQNRKHSEKMAKL 109 MmOrc2 (Q60862.1) - - RNIQESL - GNGSAKD - - - - GRNKVYSFQQRKHPEEMTKL 109 DmOrc2 (AAF99606.1) AEDQEEESIEESENAARP - - - - AAKDLHLIQSEYNVAGTSMF 132 AtOrc2 (AEC09416.1) 8 ScOrc2 (CAA85003.1) RKNT SPDPALKPKTPSKAP- - - - - RKRGRPRKIQEELTDRIKKD 95 Tb3120 (Tb927.9.4530) WTCEGPRPPRRLTRGQGA SYDVDVVQRYYTMKEQREKRAQSQAPL 172 TC3120 (TCCLB.511585.90) WT CEGPRPARRLORGAGA SYDMDVVORYYTMKEOREKRAL SQAPL 167 Lm3120 (LmjF.01.0660) WCCEGPAPAARQHSGGSLVYDTANLHKYHFMRAQREKRRSSTEPL 445 Conservation _______ _____ HsOrc2 (NP_006181.1) A SELAKT POK SV SFSLKND PEIT INV POS SKGH SA SDK VO PKNND 154 MmOrc2 (060862.1) ALELAKT SGKKDPLD- SNDPEITKNIAQK SKGHST SEKAPLVNNN 153 DmOrc2 (AAF99606.1) GFN---TPKKRDAMALAA-LNATPCTPKTPKTPRLGVKTPDTKRK 173 125 Tb3120 (Tb927.9.4530) TAQFLGT------EPSP--TE--FGGNIGVPPE-AS-Tc3120 (TcCLB.511585.90) IAHFHAA-----EGPP--THVAFSSVVGEEKDDVI-197 195 Lm3120 (LmjF.01.0660) LARVLAVFPATLGGAGTRADE SAPCVTQSPLQSRRRPHGDASSG 490 Conservation _____ - ------170 169 DmOrc2 (AAF99606.1) K SMDQPKTPAHVRTRVK-----190 AtOrc2 (AEC09416.1) 8 ScOrc2 (CAA85003.1) KQVMEKTGIKEKREREK------142 Tb3120 (Tb927.9.4530) RTGVLGRRPF-----Tc3120 (TcCLB.511585.90) GNGPLNRRPF-----207 205 Lm3120 (LmjF.01.0660) KSAVRGRDPLPSASSVHGRRLTLSRLDARANAAASLRDGAALPHL 535 Conservation DmOrc2 (AAF99606.1) KQIA-----KIVADSDEDF--SGDESDFRPSDEESSSSSSSDA 227 AtOrc2 (AEC09416.1) 17 ScOrc2 (CAA85003.1) IQVATTTYEDNVTPQTDDNFVSNSPEPPEPATPSKKSLTTNHDFT 187 - CEEGCSFSCVVQNPRLEQSLVWCGVNRQLRES---HP----- CVEGGSFVFSVQNPRMEVSCVWCGVDRRLSEA---HP----Tb3120 (Tb927.9.4530) 241 Tc3120 (TcCLB.511585.90) 239 Lm3120 (LmjF.01.0660) DRCEEGGSYVFSVWNPRMEVQLVWCAVNREAREGAPALHPHHYQR 580 Conservation HSOR2 (NP 006181 1) GVACE----HEEDTNAVIESOK LOAONBVV SAPVGKETPSKBMKB 233 MmOrc2 (060862.1) EATKD---- EEEDTNVARLSQKSQGQNRLLPAPVSKETLPKKKKR 232 DmOrc2 (AAF99606.1) GNSSD----NDAADDEPKTPSRAR---RAIVVPVLPKTPSAARLR 265 AtOrc2 (AEC09416.1) 17 ScOrc2 (CAA85003.1) SPLKQIIMNNLKEYKDSTSPGKLTLSRNFTPTPV----PKNKKLY 228 Lm3120 (LmjF.01.0660) SSVGH - - - - RGHGVLPPSASVTPLFMWRCPPSLPGLPARPFRV 619 Conservation ___ 255 254 295 28 256 Tb3120 (Tb927.9.4530) HIPYGSYGIHKIHAVALYVTEDIYRRATARSGSPRPVVKSPGGRR 316 TC3120 (TCCLB.511585.90) RVPYGVYGLHNVHALALYLTNDIFRKATTKL - QLTTADTLTVEE Lm3120 (LmjF.01.0660) TVPYGLYAFANTHALALYEASDIFRKVTTQLAA - - AASGGDGAH 312 661 Conservation

 HsOrc2 (NP_006181.1)
 ----- DRTLQKLKRAKL-DQQTLRNLL-----

 MmOrc2 (Q60862.1)
 ---- DRTLQRLRRARV-DQKTLHNLL-----

 DmOrc2 (AAF99606.1)
 ---- DHTLDRLKNPRL-AADRVFSLL-----
 276 275 316 AtOrc2 (AEC09416.1) 50 ScOrc2 (CAA85003.1) 277 Tb3120 (Tb927.9.4530) MEQGH PQRDGQLMAVKEL PRNTVRPLLVVDVLHRTAT SD PVRNLL 361 Tc3120 (TcCLB.511585.90) NEEPPPWRRGQLRTGKALPRHAVRPLLVVDVLHRIAT SDPVWDLL 357 Lm3120 (LmjF.01.0660) HAQPHPR---QLSSVKVSPAATMRALLIVHVMHHVATT-PV----698 Conservation 299 298 -----SEIKT SAEHEGSINA IMEEYRSY FP------DmOrc2 (AAF99606.1) 341 -----STIEMKHSKE--ISELMSDYKTMYS-------AtOrc2 (AEC09416.1) 73 ScOrc2 (CAA85003.1) -----SNFFNENFQKRPRQKLFEIQKKMFP--------302 Tb3120 (Tb927.9.4530) PTQWLPEH- MRRPPSRICSPRRENSITETAGCV- - - GSGVTKGN Tc3120 (TcCLB.511585.90) SVEACATYVLNAAVGPLKYQKGCQGVTEEKATV- - - ETHAVEET 401 398 - - - - PLRLGRALPPPLPDPT SQHHHLKEQRPMPQASSSHVARGT Lm3120 (LmiF.01.0660) 738 Conservation HsOrc2 (NP 006181.1) 299 MmOrc2 (Q60862.1) 298 DmOrc2 (AAF99606.1) 341 AtOrc2 (AEC09416.1) 73 Conservation



Figure 7.6. (continued).





100

Conservation

%

0





100

'n

% Conservation

7.3 RNAi in PCF cells

These results refer to the results shown in Chapter 3, section 3.3.2.

7.3.1 Extra TbORC1/CDC6 and Tb3120 RNAi clones

Here are shown the results for two clones, TbORC1/CDC6 RNAi Cla (Figure 7.7) and Tb3120 RNAi Cla (Figure 7.8) that revert the RNAi phenotype after 72 h and 168 h post-induction, respectively.



Figure 7.7. Effect of TbORC1/CDC6 depletion by induction of specific gene targeted RNAi over time.

A) Growth curves of un-induced (Tet -) and tetracycline-RNAi induced (Tet +) cell cultures over five days. Cell concentration was assessed every 24 h, and plotted on a Log₁₀ y-axis graph. The individual points represent the mean from two independent experiments, while the error bars depict the standard error of the mean (SEM). The red arrow pinpoints a 1:10 dilution of both Tet – and Tet + cultures. B) Quantification of cells in the different cell cycle stages throughout the course of five days of RNAi induction, based on the nuclear and kinetoplast configuration of the cells stained with DAPI. A minimum of 150 cells was counted per time point and experimental group (Tet - and Tet +), and percentages of each cell type (1N1K, 1N2K, 2N2K, 0N1K, and others) were calculated relative to the total amount of cells analysed. Only one experiment is shown. C) Percentage of EdU positive cells in the Tet - culture from the same time point. A minimum of 150 cells were analysed per time point and group (Tet - and Tet +). Only one experiment is shown.



Figure 7.8. Effect of Tb3120 depletion by induction of specific gene targeted RNAi over time. A) Growth curves of un-induced (Tet -) and tetracycline-RNAi induced (Tet +) cell cultures over eight days. Cell concentration was assessed every 24 h, and plotted on a Log_{10} y-axis graph. The individual points represent the mean from two independent experiments, while the error bars depict the standard error of the mean (SEM). The red arrow pinpoints a 1:10 dilution of both Tet – and Tet + cultures. B) Quantification of cells in the different cell cycle stages throughout the course of five

days of RNAi induction, based on the nuclear and kinetoplast configuration of the cells stained with DAPI. A minimum of 150 cells was counted per time point and experimental group (Tet - and Tet +), and percentages of each cell type (1N1K, 1N2K, 2N2K, 0N1K, and others) were calculated relative to the total amount of cells analysed. Only one experiment is shown. C) Percentage of EdU positive cells in the Tet + samples relative to the number of EdU positive cells in the Tet - culture from the same time point. A minimum of 150 cells were analysed per time point and group (Tet - and Tet +). Only one experiment is shown.

7.3.2 Flow cytometry analysis of the effects on cell cycle resultant from gene-targeted RNAi induction

Here are shown the individual DNA content histograms obtained from flow cytometry analysis of the RNAi-induced cell lines, and used to generate the graphs in Chapter 3, Figure 3.11.





Top row, scatter plots of the side scatter (SSC) x forward scatter (FSC), showing the cells in the population by internal complexity (SSC) and size (FSC). The gate is shown as a rectangle enclosing the whole cellular population; the percentage of cells within is shown. These cells are shown in the row below, displayed in scatter plots of FL2-W (width) x FL2-A (area), as FL2 is the detector that detects the propidium iodide fluorescent signal. The gate is shown as a rectangle, and the percentage of cells within is shown. The graphs below represent the cells within that gate but as histograms of the FL2-A data. Here, gates are shown for the G1 and G2 populations, as well as a sub-G1 population labelled as zoids. The percentages for each are shown, and these were used to generate the graph in Chapter 3, Figure 3.11. Using the back-gating function of FlowJo, the "zoid" population was back-gated to the SSCxFSC plot, and the cells are shown in orange. Bottom graphs show the overlap of the histograms of the Tet – (full blue line) and Tet + (dashed blue line) samples for each time point.



Figure 7.10. 2913 cell line, 72 h, 96 h and 120 h time points. Description as in Figure 7.9.



Figure 7.11. 2913 cell line, 144 h and 168 h time points. Description as in Figure 7.9.



Figure 7.12. TbORC1/CDC6 RNAi Clb cell line, 12 h, 24 h and 48 h time points. Description as in Figure 7.9.



Figure 7.13. TbORC1/CDC6 RNAi Clb cell line, 72 h, 96 h and 120 h time points. Description as in Figure 7.9.



Figure 7.14. TbORC1/CDC6 RNAi Clb cell line, 144 h time points. Description as in Figure 7.9.



Figure 7.15. TbORC4 RNAi Cla cell line, 12 h, 24 h and 48 h time points. Description as in Figure 7.9.



Figure 7.16. TbORC4 RNAi Cla cell line, 72 h, 96 h and 120 h time points. Description as in Figure 7.9.



Figure 7.17. TbORC4 RNAi Cla cell line, 144 h time points. Description as in Figure 7.9.



Figure 7.18. TbORC4 RNAi Clb cell line, 12 h, 24 h and 48 h time points. Description as in Figure 7.9.



Figure 7.19. TbORC4 RNAi Clb cell line, 72 h, 96 h and 120 h time points. Description as in Figure 7.9.



Figure 7.20. TbORC4 RNAi Clb cell line, 144 h time points. Description as in Figure 7.9.



Figure 7.21. Tb3120 RNAi Clb cell line, 24 h, 48 h and 72 h time points. Description as in Figure 7.9.



Figure 7.22. Tb3120 RNAi Clb cell line, 96 h and 120 h time points. Description as in Figure 7.9.



Figure 7.23. Tb3120 RNAi Clb cell line, 144 h and 168 h time points. Description as in Figure 7.9.



Figure 7.24. TbORC1B RNAi cell line, 6 h, 12 h and 24 h time points. Description as in Figure 7.9.



Figure 7.25. TbORC1B RNAi cell line, 48 h, 72 h and 96 h time points. Description as in Figure 7.9.



Figure 7.26. TbORC1B RNAi cell line, 120 h and 144 h time points. Description as in Figure 7.9.

7.4 Cloning the constructs

7.4.1 Confirmation enzymatic digestion of the plasmids for endogenous tag of Lister 427 cell lines

When comparing the available sequences for the strains TREU 927 and Lister 427 of *T. brucei* (TriTrypDB, version 8.1), it was evident that there were some minor nucleotide differences between the two (Figure 7.27, Figure 7.28, Figure 7.29, Figure 7.30, Figure 7.31, and Figure 7.32). While in the case of both TbORC1/CDC6 and TbORC1B the regions used for cloning did not include any nucleotide difference that resulted in an amino acid change in the resultant proteins, the same was not observed for TbORC4, Tb7980, Tb3120 and Tb1120. To avoid any differences between strains, gDNA from TREU 927 and Lister 427 was used as a template to produce the constructs used for tagging the proteins of interest in parasites of the respective two strains. While the plasmids generated for the tagging of cells from the strain TREU 927 are shown in Chapter 3, section 3.2.1.1, the plasmids used for the tagging of cells from the tagging of cells from the Lister 427 strain are here shown in Figure 7.33.

TREU 927 TbORC1/CDC6	GCATGGAACTGGCAGCTGATGCGAAAACGAGAAGTCGACTGGACA	45
Lister 427 TbORC1/CDC6	GCATGGAACTGGCAGCTGATGCGAAAACGAGAAGTCGACTGGACA	45
TREU 927 TbORC1/CDC6	TAACGAAACGTCTCGTCTTCGAACCATATAGTCTGCCTGAGCTCA	90
Lister 427 TbORC1/CDC6	TAACGAAACGTCTCGTCTTCGAACCATATAGTCTGCCTGAGCTCA	90
TREU 927 TbORC1/CDC6	AGGAAATTATTCTTCGAAGAGT <mark>©</mark> AGTCACATCAAACCTACGTTAT	135
Lister 427 TbORC1/CDC6	AGGAAATTATTCTTCGAAGAGT <mark>G</mark> AGTCACATCAAACCTACGTTAT	135
TREU 927 TbORC1/CDC6	TTGC <mark>A</mark> GAGAAAGCAATCAACTATCTATGTAACCAAACTGCATCGC	180
Lister 427 TbORC1/CDC6	TTGC <mark>G</mark> GAGAAAGCAATCAACTATCTATGTAACCAAACTGCATCGC	180
TREU 927 TbORC1/CDC6	ACTACGGGGATGTGAGACGTCTTTTACAATCTGCTTCCTCCGCCA	225
Lister 427 TbORC1/CDC6	ACTACGGGGATGTGAGACGTCTTTTACAATCTGCTTCCTCCGCCA	225
TREU 927 TbORC1/CDC6	TTTGCGGTCTCATGATGAGAATAGAGGAGGGTTACAAGTTGCCGG	270
Lister 427 TbORC1/CDC6	TTTGCGGTCTCATGATGAGAATAGAGGAGGGTTACAAGTTGCCGG	270
TREU 927 TbORC1/CDC6	AGAAGCATGATGGCTTGCTAACTGTAAAGGATGTTCACTCAGTTG	315
Lister 427 TbORC1/CDC6	AGAAGCATGATGGCTTGCTAACTGTAAAGGATGTTCACTCAGTTG	315
TREU 927 TbORC1/CDC6	TTCGCCAAATATTTCACGATCGCTTTGTTGAGTTTATCCAAACTA	360
Lister 427 TbORC1/CDC6	TTCGCCAAATATTTCACGATCGCTTTGTTGAGTTTATCCAAACTA	360
TREU 927 TbORC1/CDC6	TTCGTCTTCCCGTAGTGTTTATCAGCGTTGCTGTCATTGCAGTAG	405
Lister 427 TbORC1/CDC6	TTCGTCTTCCCGTAGTGTTTATCAGCGTTGCTGTCATTGCAGTAG	405
TREU 927 TbORC1/CDC6	AGACAGCAAGGCTTTTTCGAGCGAACTGCGAGGACAGCCGACTAC	450
Lister 427 TbORC1/CDC6	AGACAGCAAGGCTTTTTCGAGCGAACTGCGAGGACAGCCGACTAC	450
TREU 927 TbORC1/CDC6	CCATAGATAGCTTGTTTACGGCAACGAAGAGAGCCCAAGAGCGTT	495
Lister 427 TbORC1/CDC6	CCATAGATAGCTTGTTTACGGCAACGAAGAGAGCCCAAGAGCGTT	495
TREU 927 TbORC1/CDC6	TTGGCTCAGTTTTTGCAGACCTACATGCCGTCACTTTGAACTATG	540
Lister 427 TbORC1/CDC6	TTGGCTCAGTTTTTGCAGACCTACATGCCGTCACTTTGAACTATG	540
TREU 927 TbORC1/CDC6	GGGCGTACCTAGAAATAGTAGAGATGCTGCGGGAGGTAGCACTGA	585
Lister 427 TbORC1/CDC6	GGGCGTACCTAGAAATAGTAGAGATGCTGCGGGAGGTAGCACTGA	585
TREU 927 TbORC1/CDC6	TTGACGTTTCGGTAGGTGAAGAGCGCATTCCCGTCAAAACGGTTC	630
Lister 427 TbORC1/CDC6	TTGACGTTTCGGTAGGTGAAGAGCGCATTCCCGTCAAAACGGTTC	630
TREU 927 TbORC1/CDC6	AGTCACTACTCGAGGGCCACTGAGAGGGCACACGCGTCAATGCTAC	675
Lister 427 TbORC1/CDC6	AGTCACTACTCGAGGCCACTGAGAGGGCACACGCGTCAATGCTAC	675
TREU 927 TbORC1/CDC6	AACCATTCCAAACAGTCGTCGATGCATGCAAGCTCCACGATGACT	720
Lister 427 TbORC1/CDC6	AACCATTCCAAACAGTCGTCGATGCATGCAAGCTCCACGATGACT	720
TREU 927 TbORC1/CDC6	TTGGTACGGGGATATGCCCACTGTTTTCGATATAG 755	

Lister 427 TbORC1/CDC6 TTGGTACGGGGATATGCCCACTGTTTTCGATATAG 755

Figure 7.27. Alignment of the DNA sequences of the TbORC1/CDC6 gene from the reference genomes of the TREU 927 and Lister 427 strains of *T. brucei*.

Only the region of the gene used for the cloning is represented. None of the nucleotide differences resulted in amino acid changes.

TREU 927 TbORC1B	ACAACGAGACAGTCAAATCGCAAAGGTGAGCATCGGTAGCACCTC 4	5
Lister 427 TbORC1B	ACAACGAGACAGTCAAATCGCAAAGGTGAGCATCGGTAGCACCTC 4	5
TREU 927 TbORC1B	CGCTGTTGGGCGGCG <mark>G</mark> CAAAACAAACGGCCATTGGGTCTTCGTT 90	0
Lister 427 TbORC1B	CGCTGTTGGGCGGCGACAAAACAAACGGCCATTGGGTCTTCGTT 9	0
TREU 927 TbORC1B	CGGCTCATACTCCGCGGCAGAGCCGGGTCCAGTAGTTCCCGGCAC 1	35
Lister 427 TbORC1B	CGGCTCATACTCCGCGGCAGAGCCGGGTCCAGTAGTTCCCGGCAC 1	35
TREU 927 TbORC1B	CTTCAC <mark>G</mark> ACGCGGGTTGTGCACCGCCTGTATACAGCACTAATGGG 1	80
Lister 427 TbORC1B	CTTCACAACGCGGGTTGTGCACCGCCTGTATACAGCACTAATGGG 1	80
TREU 927 TbORC1B	CCAACAAAGATTCCCGTCAATGAATGCGGCAGGGATTTCGTCGGC 2	25
Lister 427 TbORC1B	CCAACAAAGATTCCCGTCAATGAATGCGGCAGGGATTTCGTCGGC 2	25
TREU 927 TbORC1B	CATCGATGGTTTTGCGGACATTGGCATCATCTCCCGCCCACAACG 2	70
Lister 427 TbORC1B	CATCGATGGTTTTGCGGACATTGGCATCATCTCCCGCCCACAACG 2	70
TREU 927 TbORC1B	CCGTGGGAATGAAGAAGTTTTCTCCTTTAATGGAACCTGGACACT	15
Lister 427 TbORC1B	CCGTGGGAATGAAGAAGTTTTCTCCTTTAATGGAACCTGGACACT 3.	15
TREU 927 TbORC1B	AGAGTCCATGCAAGCCGCTCTCACAGCGCGCGGGGGGGGG	60
Lister 427 TbORC1B	AGAGTCCATGCAAGCCGCTCTCACAGCGCGCGGGGGGGGG	60
TREU 927 TbORC1B	ACAGGAACGAGTGGATTGTGGACTGGATTCTGCAGAGAATCGTTT 4	05
Lister 427 TbORC1B	ACAGGAACGAGTAGATTGTGGACTGGATTCTGCAGAGAATCGTTT 4	05
TREU 927 TbORC1B	TGAAGAAGTGCTGCGGGAACTCAAGGGCATTTTATCCCTGTGA 448	
Lister 427 TbORC1B	TGAAGAAGTGCTGCGGGAACTCAAGGGCATTTTATCCCTGTGA 448	

Figure 7.28. Alignment of the DNA sequences of the TbORC1B gene from the reference genomes of the TREU 927 and Lister 427 strains of *T. brucei*. Only the region of the gene used for the cloning is represented. None of the nucleotide differences resulted in amino acid changes.

TREU 927 TbORC4	CGTTTCTGCTGTCTTTGGGGAAGTGTGTTCAGGGGAAATTCCTCTC	46
Lister 427 TbORC4	CGTTTCTGCTGTCTTTGGGGAAGTGTGTTCAGGGGAAATTCCTCTC	46
TREU 927 TbORC4	CTTAGCAGTGCCAGTTGTGGTAAACTTTTGTCGTGGTTTGAGAGAA	92
Lister 427 TbORC4	CTTAGCAGTGCCAGTTGTGGTAAACTTTTGTCGTGGTTTGAGAGAA	92
TREU 927 TbORC4	CTTCG <mark>C</mark> CGAAACTGAAAGTTGCGCAAAATATCCGTAGTGCCGTCTT	138
Lister 427 TbORC4	CTTCGGCGAAACTGAAAGTTGCGCAAAATATCCGTAGTGCCGTCTT	138
TREU 927 TbORC4	TAGTGTTGCGCCAGAGGCAGTGAAGTCATTGTGGCGTGGTGAATTT	184
Lister 427 TbORC4	TAGTGTTGCGCCAGAGGCAGTGAAGTCATTGTGGCGTGGTGAATTT	184
TREU 927 TbORC4	ACCAATGCGAATCGGAAAAGCGCAGGGTCGGTGAAACCTCTCAGGA	230
Lister 427 TbORC4	ACCAATGCGAATCGGAAAAGCGCAGGGTCGGTGAAACCTCTCAGGA	230
TREU 927 TbORC4	ATTACGCGCTTTTAGTTGGTGATATGCTTTCCGAATGCAAACTCGT	276
Lister 427 TbORC4	ATTACGCGCTTTTAGTTGATGATATGCTTTCCGAATGCAAACTCGT	276
TREU 927 TbORC4	TGAACTGGGTTACTGTACCCGGGAAATGTTTTTACTTCTTACGTAC	322
Lister 427 TbORC4	TGAACTGGGTTACTGTACCCGGGAAATGTTTTTACTTCTTACGTAC	322
TREU 927 TbORC4	GTATACCTCCGCCATGAGGCTGGTGTAGTGCGTACCGTCGTTGACT	368
Lister 427 TbORC4	GTATACCTCCGCCATGAAGCTGGTGTAGTGCGTACCGTCGTTGACT	368
TREU 927 TbORC4	TGCTGGAGGATGTTGCCTCGTCCATGGGAACACATGCTGCTGCTGC	414
Lister 427 TbORC4	TGCTGGAGGATGTTGCCTCGTCCATGGGAACACATGCTGCTGCTGC	414
TREU 927 TbORC4	ACTTGACCGTGCAGCATTCACCGCAGCTGTCGGGTTGCTTAACCGT	460
Lister 427 TbORC4	ACTTGACCGTGCAGCATTCACCGCAGCTGTCGGGTTGCTTAACCGT	460
TREU 927 TbORC4	TGGCGAATTGTTCGTGTCGGCGGTCGAGATGGCTCGACGGCAGTTT	506
Lister 427 TbORC4	TGGCGAATTGTTCGTGTCGGCGGTCGAGATGGCTCGACGGCAGTTT	506
TREU 927 TbORC4	CGTTGAGGGGTAGCCCAGCCAGACTTCGAGAGTTTCTTCAGGAGGT	552
Lister 427 TbORC4	CGTTGAGGGGTAGCCCAGCCAGACTTCGAGAGTTTCTTCAGGAGGT	552
TREU 927 TbORC4	GTTACATCGCTCAGAGTACTGCA <mark>A</mark> CGAAACCCTCGGTTTGGATACC	598
Lister 427 TbORC4	GTTACATCGCTCAGAGTACTGCAGCGAAACCCTCGGTTTGGATACC	598
TREU 927 TbORC4	AAGGAGGTGGCGCGATTACGCAGCCTCGTGTGA 631	

Lister 427 TbORC4 AAGGAGGTGGCGCGATTGCGCAGCCTCGTGTGA 631

Figure 7.29. Alignment of the DNA sequences of the TbORC4 gene from the reference genomes of the TREU 927 and Lister 427 strains of *T. brucei*.

Only the region of the gene used for the cloning is represented. From the five nucleotide differences, only three result in amino acid changes.

Lister 427 Tb7980	gacatgccgtgacgaactcaggcggaaatgggaggcttg <mark>c</mark> ggttc	45
TREU 927 Tb7980	gacatgccgtgacgaactcaggcggaaatgggaggcttg <mark>t</mark> ggttc	45
Lister 427 Tb7980	gtggttggtgtttggtaa <mark>a</mark> gaacgtggaagtcacgatagttttcg	90
TREU 927 Tb7980	gtggttggtgtttggtaa <mark>g</mark> gaacgtggaagtcacgatagttttcg	90
Lister 427 Tb7980	acgtgcgtacgcgcataatatgtggctttactatgtgttttagca	135
TREU 927 Tb7980	acgtgcgtacgcgcataatatgtggctttactatgtgttttagca	135
Lister 427 Tb7980	cgaatgttttccctacccttcgtgtttgacattttctgttctcat	180
TREU 927 Tb7980	cgaatgttttccctacccttcgtgtttgacattttctgttctcat	180
Lister 427 Tb7980	tcacaatagcgctgttccccctcccctgtgcacttcagtgg <mark>g</mark> tgg	225
TREU 927 Tb7980	tcacaatagcgctgttccccctcccctgtgcacttcagtgg <mark>t</mark> tgg	225
Lister 427 Tb7980	gcatcaccaaggtagagcactgtgcccctcccctcccct	270
TREU 927 Tb7980	gcatcaccaaggtagagcactgtgcccctcccctcccct	270
Lister 427 Tb7980	actgtagcaattacgtacgagcgttctgtttccccgtcgtatggc	315
TREU 927 Tb7980	actgtagcaattacgtacgagcgttctgtttccccgtcgtatggc	315
Lister 427 Tb7980	agcccaaacaccacgcaaggaggt tggagtggacctgcgggatga	360
TREU 927 Tb7980	agcccaaacaccacgcaaggaggt tggagtggacctgcgggatga	360
Lister 427 Tb7980	attagacgcagacgtggagtgtggctccggttacatctccatctt	405
TREU 927 Tb7980	attagacgcagacgtggagtgtggctccggttacatctccatctt	405
Lister 427 Tb7980	cggaccaccgggtagcggcaaaacgactatgctcctgcagtacct	450
TREU 927 Tb7980	cggaccaccgggtagcggcaaaacgactatgctcctgcagtacct	450
Lister 427 Tb7980	cacttcgcgggagcatcacgtacgtaatatgaagcggtctttctt	495
TREU 927 Tb7980	cacttcgcgggagcatcacgtacgtaatatgaagcggtctttctt	495
Lister 427 Tb7980	gctagagtatgttactggcagggcgctagctggtgactccctgcg	540
TREU 927 Tb7980	gctagagtatgttactggcagggcgctagctggtgactccctgcg	540
Lister 427 Tb7980	g c g g c t g g c g t g g c g t c t t t t	585
TREU 927 Tb7980	gcggctggcgtggcgtcttttgccgcaaactaaacggcggcggac	585
Lister 427 Tb7980	agaatgttcacatcttcagtttggcttgctggtacgacagtggct	630
TREU 927 Tb7980	agaatgttcacatcttcagtttggcttgctggtacgacagtggct	630
Lister 427 Tb7980	tgacaatgcagtggagggggggggggggggggggggggg	
TREU 927 Tb7980	tgacaatgcagtggaggggagtgagcttcacttcgtcgtg 670	

Figure 7.30. Alignment of the DNA sequences of the Tb7980 gene from the reference genomes of the TREU 927 and Lister 427 strains of *T. brucei*.

Only the region of the gene used for the cloning is represented. Because Tb7980 was tagged in the N-terminal, both the 5' region immediately before the gene's start codon (dark red), and the first 357 bp of the gene (red), were used for cloning. Although there were no differences between the sequences of the gene, the 5' region used presented three nucleotide variances. To avoid interference with this region, plasmids were generated using gDNA from both strains.

TREU 927 Tb3120	AGTGCATGGTATAGACGAACTCGACCCACCACTTCTGGTAGAGCTT	46
Lister 427 Tb3120	AGTGCATGGCATAGACGAACTCGACCCACCACTTCTGGTAGAGCTT	46
TREU 927 Tb3120	CAGAACATTGCACGTGACCATCCCAATCGTGTAATGTTGCTATGCT	92
Lister 427 Tb3120	CAGAACATTGCACGTGACCATCCCAATCGTGTAATGTTGCTGTGCT	92
TREU 927 Tb3120	CATTCGACGACCCAAACTGGGCCATGTCAAACAGTGCGGCGCAGTT	138
Lister 427 Tb3120	CATTCGACGATCCAAACTGGGCCATGTCAAACAGTGCGGCGCAGTT	138
TREU 927 Tb3120	GGAGCCGTTTCGATTGGCGTATGTACACCTCCGCTCGATGTTGCT	184
Lister 427 Tb3120	GGAGCCGTTTCGATTGGCGTATGTACACCTCCGCTCGATGTTGCT	184
TREU 927 Tb3120	CCACGGGTGCATGAAATGGCATGTGTTAAAAGTCTCACGTTGCTCA	230
Lister 427 Tb3120	CCACGGGTGCATGAAATGGCATGTGTTAAAAGTCTCACGTTGCTCA	230
TREU 927 Tb3120	CAGACCTTGAGGCAGCAGCAGCAGGGGGGAAAAGGTTTGGTCATCA	276
Lister 427 Tb3120	CAGACCTTGAGGCAGCAGCAGCAGGGGGGAAAAGGTTTGGTCATCA	276
TREU 927 Tb3120	CGGGTTAAGAGGAAGTCTTGGCCCTGGGACCTCTTACCACTTCAA	322
Lister 427 Tb3120	CGGGTTAAGAGGAAGTCTTGGCCCTGGGACCTCTTACCACTTCAA	322
TREU 927 Tb3120	GACACCATTAGACGTATACTTTTTAG <mark>T</mark> CTTCCCGCCACGTTTACTG	368
Lister 427 Tb3120	GACACCATTAGACGTATACTTTTTAGCCTTCCCGCCACGTTTACTG	368
TREU 927 Tb3120	ATGTCTTACGATGCATGATTGAGCGACAGGAAGCGTCGGGCGAAAA	414
Lister 427 Tb3120	ATGTCTTACGATGCATGATTGAGCGACAGGAAGCGTCGGGCGAAAA	414
TREU 927 Tb3120	TGTCTTCGTTCCCATGAG <mark>T</mark> CTCCACCAGCAACACTTCGACGATCGA	460
Lister 427 Tb3120	TGTCTTCGTTCCCATGAGCCTCCACCAGCAACACTTCGACGATCGA	460
TREU 927 Tb3120	GGAATGATGATTTCAGTGGGCCGCCTCAGAGCGATTGAACGGGAAC	506
Lister 427 Tb3120	GGAATGATGATTTCAGTGGGCCGCCTCAGAGCGATTGAACGGGAAC	506
TREU 927 Tb3120	TAACATCCAATCGGCTGGCTGTGTTTGATGCAGCAGAAAATAAAT	552
Lister 427 Tb3120	TAACATCCAATCGGCTGGCTGTGTTTGATGCAGCAGAAAATAAAT	552
TREU 927 Tb3120	GATGATTCCTCAACACAAAAAACTGCTGCGGGTGTTGGAGGAAGTC	598
Lister 427 Tb3120	GATGATTCCTCAACACAAAAACTGCTGCGGGTGTTGGAGGAAGTC	598
TREU 927 Tb3120	GCGG <mark>G</mark> ACAGAGACAAAACACTCGGTCAAACGGTGGAGCTCCAGTGG	644
Lister 427 Tb3120	GCGGAACAGAGACAAAACACTCGGTCAAACGGTGGAGCTCCAGTGG	644
TREU 927 Tb3120	AGGCATAG 652	
Lister 427 Tb3120	AGGCATAA 652	

Figure 7.31. Alignment of the DNA sequences of the Tb3120 gene from the reference genomes of the TREU 927 and Lister 427 strains of *T. brucei*.

Only the region of the gene used for the cloning is represented. From the seven nucleotide differences, only one result in amino acid changes.

TREU 927 Tb1120	CTTCTGTTGCTTTCTGCGAGAGGCCCTTCGCCGTTTCCCAGTGAT	45
Lister 427 Tb1120	CTTCTGTTGCTTTCTGCGAGAGGCCCTTCGCCGTTTCCCAGTGAT	45
TREU 927 Tb1120	ACTGCCGCCGTTGTTGCTTCGGCAGTTGCAGTCGCTCTGGGCGTT	90
Lister 427 Tb1120	ACTGCCGCCGTTGTTGCTTCGGCAGTTGCAGTCGCTCTGGGCGTT	90
TREU 927 Tb1120	TCGACACCAACTTTCGGATGTCGTGGTGGGTCTCCATTCACTGCT	135
Lister 427 Tb1120	TCGACACCAACTTTCGGATGTCGTGGTGGGTCTCCATTCACTGCT	135
TREU 927 Tb1120	CTCTCCCTTCGCACTATCCTCCTCG <mark>A</mark> TGCATGCCACAATGAGTGA	180
Lister 427 Tb1120	CTCTCCCTTCGCACTATCCTCCTCG <mark>G</mark> TGCATGCCACAATGAGTGA	180
TREU 927 Tb1120	GGCTAATGATGTCACAACTTCTCCAGCAAATGTGGTGCAGGGGAC	225
Lister 427 Tb1120	GGCTAATGATGTCACAACTTCTCCAGCAAATGTGGTGCAGGGGAC	225
TREU 927 Tb1120	ATCGAGAGGCTACCATGAGAATATG <mark>A</mark> CACAGCTTTCTAATAGCAT	270
Lister 427 Tb1120	ATCGAGAGGCTACCATGAGAATATG <mark>G</mark> CACAGCTTTCTAATAGCAT	270
TREU 927 Tb1120	CGAGCTTCTGAACATCCTCTTCGCTGCCGCTTGTGAATGGGCAGA	315
Lister 427 Tb1120	CGAGCTTCTGAACATCCTCTTCGCTGCCGCTTGTGAATGGGCAGA	315
TREU 927 Tb1120	TGGTCATCTATTGAATGAGGCTGTTGCGTTCGC <mark>T</mark> GTACTGTACGA	360
Lister 427 Tb1120	TGGTCATCTATTGAATGAGGCTGTTGCGTTCGCCCGTACTGTACGA	360
TREU 927 Tb1120	AGACCTTGTTTTCGGTAGGCTGACGCGGCTGAA <mark>A</mark> CGTATACCCGG	405
Lister 427 Tb1120	AGACCTTGTTTTCGGTAGGCTGACGCGGCTGAA <mark>G</mark> CGTATACCCGG	405
TREU 927 Tb1120	TTTGACCGCGTATGTGGAGCGCTTATCGCCGAAGTCGGTGGGGCC	450
Lister 427 Tb1120	TTTGACCGCGTATGTGGAGCGCTTATCGCCGAAGTCGGTGGGGGCC	450
TREU 927 Tb1120	TACCGATTTGTCCAAGGCACTGACCAGCTGCGTTCCGGTGTTCGT	495
Lister 427 Tb1120	TACCGATTTGTCCAAGGCACTGACCAGCTGCGTTCCGGTGTTCGT	495
TREU 927 Tb1120	GCCTAACGCATTAACAGGCTCTGGGGCGTATATTCGCGACGAACT	540
Lister 427 Tb1120	GCCTAACGCATTAACAGGCTCTGGGGCGTATATTCGCGACGAACT	540
TREU 927 Tb1120	TTCTCAGGAACAACAGTACGTAATGAGAAGGCCCCTTTCTTACC	585
Lister 427 Tb1120	TTCTCAGGAACAACAGTACGTAATGAGAAGGCCCCTTTCTTT	585
TREU 927 Tb1120	ACCAATTATGGCCGTACCCGGCGCAAAGGTGTCGCTGGCAACAGA	630
Lister 427 Tb1120	ACCAATTATGGCCGTACCCGGCGCAAAGGTGTCGCTGGCAACAGA	630
TREU 927 Tb1120	GTTGCTCCGCTCCACCCTTTTGGCCGTGCTCCCACCGCACGACTC	675
Lister 427 Tb1120	GTTGCTCCGCTCCACCCTTTTGGCCGTGCTCCCACCGCACGACTC	675
TREU 927 Tb1120	GTTGGAGGAAGTGCAGAAGGCTGTAAGAGCCTTTCCGTGTGCCGG	720
Lister 427 Tb1120	GTTGGAGGAAGTGCAGAAGGCTGTAAGAGCCTTTCCGTGTGCCGG	720
TREU 927 Tb1120	CGGTAAATGTGATGAGGGTACTCCAAGTGAGGCTCTGCGGCATGG	765
Lister 427 Tb1120	CGGTAAATGTGATGAGGGTACTCCAAGTGAGGCTCTGCGGCATGG	765

Figure 7.32. (continue next page, together with description)

TREU 927 Tb1120	AGACACGTACTGTGGTGTGACGTACAATAACTTTTTCAGTCCCCT	810
Lister 427 Tb1120	©GACACGTACTGTGGTGTGACGTACAATAACTTTTTCAGTCCCCT	810
TREU 927 Tb1120	AATTCCTGACTCGGTGCGAGTATTGCACCTTCTCACCTCCCATGC	855
Lister 427 Tb1120	AATTCCTGACTCGGTGCGAGTATTGCACCTTCTCACCTCCCATGC	855
TREU 927 Tb1120	GATGGCAAGCTCTAACG <mark>T</mark> TAAGCAGCAATTTGTGCCACTTTCTCT	900
Lister 427 Tb1120	GATGGCAAGCTCTAACGCTAAGCAGCAATTTGTGCCACTTTCTCT	900
TREU 927 Tb1120	AATACAGCGAATTTGTCAGCTTTCGGATGAATCGCTGATACGATC	945
Lister 427 Tb1120	AATACAGCGAATTTGTCAGCTTTCGGATGAATCGCTGATACGATC	945
TREU 927 Tb1120	GTTAGTGGAACTGCAATTGACGGGAATGGCGACCGTTAACATGCG	990
Lister 427 Tb1120	GTTAGTGGAACTGCAATTGACGGGAATGGCGACCGTTAACATGCG	990
TREU 927 Tb1120	AGAGTTTAAAGCGAGGTCCTCCTTGCTTGCGCTAGCTGA 1030	
Lister 427 Tb1120	AGAGTTTAAAGCGAGGTCCTCCTTGCTTGCGCCCTAGCTGA 1030	

Figure 7.12. Alignment of the DNA sequences of the Tb1120 gene from the available genomes of the TREU 927 and Lister 427 strains of *T. brucei*.

Only the region of the gene used for the cloning is represented. From the seven nucleotide differences, only four result in amino acid changes.



Figure 7.33. Enzymatic digestion of the plasmids used for endogenous tagging of proteins with N- or C-terminal 12-myc tag, to be used for the transfection of Lister 427 cell lines. Description as in Chapter 3, Figure 3.13.





Figure 7.34. Original plasmid used for the generation of the constructs used for gene deletion described in Chapter 3.

7.5 Cell cycle mRNA levels of TbORC1/CDC6 interacting factors



Figure 7.35. mRNA levels of TbORC1/CDC6, TbORC1B, TbORC4, Tb7980, Tb3120 and Tb1120 throughout the cell cycle.

Levels of mRNA of TbORC1/CDC6, TbORC1B, TbORC4, Tb7980, Tb3120 and Tb1120 throughout the cell cycle of PCF cells; the graphs were exported from TriTrypDB version 8.0, and represent the data originated by (Archer *et al.*, 2011).

7.6 TbORC1/CDC6 and other factors subcellular localisation

	DAPI 5 µm myc	5 µm	5 μm
	AH		
	DIC		and
	927 wt	TbORC1/CDC6 -/12myc	TbORC1/CDC6 -/12myc TbORC1B 6HA
anti-myc	+	+	+
anti-HA	+	+	-
Anti-mouse A596	+	+	+

Figure 7.36. Immunofluorescence of TbORC1/CDC6^{12myc} and TbORC1B^{6HA}, TbORC4^{6HA}, Tb7980^{6HA}, Tb3120^{6HA}, and Tb1120^{6HA} – Controls used.

Top panel row shows the staining of the cells with DAPI. Panel row below shows detection with the anti-myc antiserum, and the next one detection with anti-HA and anti-mouse AlexaFluor® 594 conjugate antisera. Lowest panel row shows the cells outline by DIC. PCF 927 wt cells were stained for both myc and HA detection. Images show that no clear non-specific signal is detected for both tags' labelling. TbORC1/CDC6 -/12myc cells were stained for both myc and HA detection. Images show that no clear non-specific signal is detected. Images show that no clear non-specific signal is detected for both tags' labelling. TbORC1/CDC6 -/12myc cells were stained for both myc and HA detection. Images show that no clear non-specific signal is detected for the HA tag detection, demonstrating that there is no interference of the signal emitted from the anti-myc antisera into the TRITC filter set (used for the detection of AlexaFluor® 594 fluorophore). TbORC1/CDC6 -/12myc TbORC1B 6HA cells were incubated with the anti-mouse AlexaFluor® 594 antiserum alone, followed by the incubation with anti-myc. Images show that the secondary antibody alone is not detecting any signal in the TRITC filter channel, and that a posterior incubation of the sample with the anti-myc antibody, raised in mouse like the anti-HA antibody, leads to no cross-detection of the myc-tagged protein.



Figure 7.37. Intensity plots of DAPI signal throughout the cell cycle.

Intensity of the DAPI signal, represented (dots) as the mean of pixel intensity within the region of interest (ROI, of 21 x 21 pixels) enclosing the each cell nucleus. At least 125 cells were analysed per cell line ($n \ge 125$). The median of the values is represented, with the error bars depicting the interquartile range. Statistic significance was assessed through analysis using the Kruskal-Wallis non-parametric test. (*) p-value < 0.05; (**) p-value < 0.01; (***) p-value < 0.001; (****) p-value < 0.0001.





Intensity of the myc signal, represented (dots) as the mean of pixel intensity within the region of interest (ROI, of 21 x 21 pixels) enclosing the each cell nucleus. At least 125 cells were analysed per cell line ($n \ge 125$). The median of the values is represented, with the error bars depicting the interquartile range. Statistic significance was assessed through analysis using the Kruskal-Wallis non-parametric test. (*) p-value < 0.05; (**) p-value < 0.01; (***) p-value < 0.001; (****) p-value < 0.001; (****) p-value < 0.001.



Figure 7.39. Intensity plots of EdU signal throughout the cell cycle.

Intensity of the EdU signal, represented (dots) as the mean of pixel intensity within the region of interest (ROI, of 21 x 21 pixels) enclosing the each cell nucleus. Dots in red represent cells with EdU signal with enough intensity to be perceived by eye. At least 125 cells were analysed per cell line ($n \ge 125$). The median of the values is represented, with the error bars depicting the interquartile range. Statistic significance was assessed through analysis using the Kruskal-Wallis non-parametric test. (*) p-value < 0.05; (**) p-value < 0.01; (***) p-value < 0.001; (****) p-value < 0.001; (****) p-value < 0.001.

2N2

2N2K



Figure 7.40. Intensity plots of EdU signal throughout the cell cycle.

Results from the experiment performed independently from the one shown in Figure 7.37 to Figure 7.39, and shown in Chapter 3 (DAPI and myc signals). Intensity of the EdU signal, represented (dots) as the mean of pixel intensity within the region of interest (ROI, of 21 x 21 pixels) enclosing the each cell nucleus. Dots in red represent cells with EdU signal with enough intensity to be perceived by eye. At least 125 cells were analysed per cell line ($n \ge 125$). The median of the values is represented, with the error bars depicting the interquartile range. Statistic significance was assessed through analysis using the Kruskal-Wallis non-parametric test. (*) p-value < 0.05; (**) p-value < 0.01; (***) p-value < 0.001.





Figure 7.41. Chromatogram resultant from the gel filtration of TbORC1/CDC6 12myc. TbORC1/CDC6 -/12myc cell line was lysed and separated into fractions of 1 ml by gel filtration, which are shown in the x axis (eluted volume in ml). The red sections and numbers represent the wells in the plate in which the fraction was eluted to. The y axis represents the UV (ultraviolet) absorbance values. The samples eluted from 43 ml (1D12) to 84 ml (1G5) were selected for western blot analysis, as shown in Chapter 3, section 3.6.3.

7.8 Mapping Origins of Replication in BSF cells

The script used to map the origins of replication in BSF and PCF cells to the Lister 427 genome (Chapter 4) is shown below. The script was conceived and designed by Dr Nicholas J. Dickens, based on the methodology used previously to map the origins in PCF cells strain TREU 927 (Tiengwe *et al.*, 2012). The same script was also used to map the origins in *L. major* and *L. mexicana*, shown in Chapter 5.

First step - check quality of the sequencing results
retrieved from the sequencing system. The data is obtained as
fastaq files, which are analysed using the FastQC software
package, a quality control tool for high throughput
sequencing data
(http://www.bioinformatics.babraham.ac.uk/projects/fastqc/).
This tool opens a graphical interface in which the sequences
are uploaded and analysed. The results are shown in a graphic
and quality score. A threshold was set at a quality score of
20, therefore the sequences were trimmed (below), to exclude
those poor quality reads.

Second step - trim the reads, to exclude poor quality ones, with a quality score lower than 20. In addition, the reads obtained from the sequencing system contain the sequences of the adapters used for the library preparation and sequencing processes, and these have to be removed. For these purposes, the fastq-mcf tool (https://code.google.com/p/eautils/wiki/FastqMcf) was used.

"dir" refers to the directory in which the files are stored. The next line defines "dir", so that it can be later used to call files from that specific folder using just \$dir.

dir=/wtcmpdata/WTCMP/McCulloch/DNA/Trypanosoma/brucei

-q allows to define the quality threshold, which is 20. # -w window-size for quality trimming. # adapters.fa refers to the file that has the adapter sequences. # -o output file. # only the files for BSF early S and G2 phases are shown. All other samples were processed identically.

fastq-mcf -q 20 -w 5 adapters.fa \$dir/H9B89ADXX_BSF427-EarleyS_GTCCGC_L001_R1_001.fastq \$dir/H9B89ADXX_BSF427-EarleyS_GTCCGC_L001_R2_001.fastq -o BSF427-earlyStrimmed.fwd.fastq -o BSF427-earlyS-trimmed.rev.fastq

fastq-mcf -q 20 -w 5 adapters.fa \$dir/H9B89ADXX_BSF427-G21M_TAGCTT_L001_R1_001.fastq \$dir/H9B89ADXX_BSF427-G21M_TAGCTT_L001_R2_001.fastq -o BSF427-G2-trimmed.fwd.fastq -o BSF427-G2-trimmed.rev.fastq

this generated the files containing the trimmed sequences.

Third step - align the trimmed reads to the reference genome. The reference Lister 427 genome used was retrieved from TriTrypDB v8.0 (file name TriTrypDB-8.0_TbruceiLister427.fa). The reference genome needs to be indexed using the Bowtie2 v2.2.0 tool, and the reads were aligned to the reference genome using the very sensitive local -k1 mode.

folder containing the reference genome file

cp /wtcmpdata/Genomes/EuPathDB/TriTrypDB-8.0_TbruceiLister427.fa

index of the reference genome, generating a new file bowtie427Ref.

bowtie2-build TriTrypDB-8.0 TbruceiLister427.fa bowtie427Ref

alignment of the BSF early S and G2 phases trimmed data to the indexed reference genome. All other data was processed identically.

bowtie2 -p 4 --very-sensitive-local -x bowtie427Ref -1
BSF427-earlyS-trimmed.fwd.fastq -2 BSF427-earlyStrimmed.rev.fastq 1> BSF427-EarlyS-aligment.sam 2> BSF427EarlyS-aligment.log

bowtie2 -p 4 --very-sensitive-local -x bowtie427Ref -1
BSF427-G2-trimmed.fwd.fastq -2 BSF427-G2-trimmed.rev.fastq 1>
BSF427-G2M-aligment.sam 2> BSF427-G2M-aligment.log

Fourth step - the aligned reads are compared using a simplified version of the method described in (Tiengwe et al., 2012). The reads are binned in 2.5 Kbp sections along each chromosome, and the number of reads per bin are then used to calculate the ratios between the early S (or late S) and G2 samples, scaled for the total size of the read library (reads per 2.5 Kbp per million reads mapped). The output is then used to create the graphical representations of the chromosomes (in this case, using GraphPad, version 6.0). All steps are done using the samtools software package.

the bin size is set up to 2.5 Kbp

binzise=2500

defining the reference genome to use.

refFile=/wtcmpdata/Genomes/EuPathDB/TriTrypDB-8.0 TbruceiLister427.fa

need to index the fasta file containing the reference
genome.

samtools faidx \$refFile

need to generate a .bed file out of the reference genome. Use a script generated by Dr Nicholas J. Dickens. It then bins the reference genome into 2500 bp intervals.

fai2bed="/homes/nd48m/old_Projects/2014/03_March/mfaseq/cfg/fai2bed.pl"

\$fai2bed \$refFile.fai 2500 > ref.bed

now need to index the data files resultant from the alignment. Again, only the BSF early S and G2 are shown, the remaining were processed identically.

samtools index BSF427-EarlyS-aligment.sam

samtools index BSF427-G2M-aligment.sam

the coverage of the genome with the reads from each sample
is then assessed and the output retrieved in the form of a
.bed file.

coverageBed _abam \$BSF427-EarlyS-aligment.sam _b ref.bed >
BSF427vs427_EarlyS.2500.bed

coverageBed _abam \$BSF427-G2M-aligment.sam _b ref.bed >
BSF427vs427 G2M.2500.bed

the ratios between the coverage of the early S (or late S) phase and G2 phase samples are calculated.

for simplicity, labels are given to the bed files.

bed1=BSF427vs427_EarlyS.2500.bed

bed2=BSF427vs427_G2M.2500.bed

output file with the final results to then be used to create the graphs.

out=2040816 BSF427vs427 EarlyS v g2.2500.bedgraph

get the total number of aligned reads for each bam file total1=`awk 'BEGIN{sum=0}{ sum+=\$4} END {print sum}' \$bed1` total2=`awk 'BEGIN{sum=0}{ sum+=\$4} END {print sum}' \$bed2`

create a bedgraph of the files

paste \$bed1 \$bed2 | awk 'BEGIN{FS="\t"; OFS="\t";scale='\$total2'/'\$total1'}{if(\$11==0){next;}}{print \$1,\$2,\$3,scale * \$4/\$11}' > \$out

7.9 Origin and non-origins size in *T. brucei*, *L. major*, and *L. mexicana*

Table 7-1. *T. brucei* distance in Kbp between the first two genes within the SSRs, with or no origin activity detected by MFA-seq..

		Origins Non-orig			Non-origins			
Chr	Gene left	Gene right	SSR type	origin (kb)	Gene left	Gene right	SSR type	non- origin
1	unclea	r, it's at the cent	tromere		Tb927.1.2080	Tb927.1.200	dss	5.068
	Tb927.2.5710	Tb927.2.5720	h-t	1.047	Tb927.2.2590	Tb927.2.2650	dss	8.063
2	Tb927.2.1600	Tb927.2.1680	dss	7.987	Tb927.2.5080	Tb927.2.5120	dss	6.795
2					Tb927.2.3330	Tb927.2.3340	CSS	1.377
					Tb927.2.5360	Tb927.2.5410	CSS	6.485
	unclea	r, it's at the cent	tromere		Tb927.3.580	Tb927.3.590	dss	18.339
	Tb927.3.4390	Tb927.3.4400	h-t	10.14 5	Tb927.3.860	Tb927.3.860	dss	4.247
3				Í	Tb927.3.1030	Tb927.3.1040	h-t	4.965
				I	Tb927.3.2260	Tb927.3.2270	dss	2.434
				i	Tb927.3.4890	Tb927.3.4900	dss	2.447
	Tb927.4.1190	Tb927.4.1210	h-t	11.08 2	Tb927.4.2080	Tb927.4.2090	dss	5.132
4	Tb927.4.3740	Tb927.4.3760	h-t	36.35 2	Tb927.4.5390	Tb927.4.5400	dss	10.255
	Tb927.4.4660	Tb927.4.4670	h-t	8.668	Tb927.4.5080	Tb927.4.5090	h-t	1.899
	unclea	r, it's at the cent	tromere		Tb927.5.1580	Tb927.5.1590	dss	6.642
5	Tb927.5.2150	Tb927.5.2160	dss	10.70 5	Tb927.5.2900	Tb927.5.2910	dss	6.688
	Tb927.5.4530	Tb927.5.4540	dss	8.14	Tb927.5.3500	Tb927.5.3510	dss	2.199
	unclea	r, it's at the cent	tromere		Tb927.6.750	Tb927.6.760	dss	5.965
6	Tb927.6.4580	Tb927.6.4590	dss	7.145	Tb927.6.1290	Tb927.6.1300	h-t	9.803
					Tb927.6.2150	Tb927.6.2160	dss	4.216
	unclea	r, it's at the cent	tromere		Tb927.6.5200	Tb927.6.5210	CSS	4.774
	Tb927.7.920	Tb927.7.930	dss	7.236	Tb927.7.1360	Tb927.7.1370	h-t	3.452
7					Tb927.7.1940	Tb927.7.1950	CSS	1.154
1					Tb927.7.2730	Tb927.7.2740	dss	5.79
					Tb927.7.3470	Tb927.7.3480	h-t	3.65
		-			Tb927.7.3880	Tb927.7.3890	dss	5.56
	Tb927.8.1370	Tb927.8.1380	dss	0.414	Tb927.8.1940	Tb927.8.1950	dss	2.414
	Tb927.8.2880	Tb927.8.2890	dss	1.076	Tb927.8.1650	Tb927.8.1970	CSS	8.651
	Tb927.8.3930	Tb927.8.3940	h-t	4.781	Tb927.8.3520	Tb927.8.3530	CSS	6.603
8	Tb927.8.4890	Tb927.8.4900	dss	7.252	Tb927.8.4760	Tb927.8.4770	h-t	5.271
	Tb927.8.6560	Tb927.8.6570	CSS	6.468	Tb927.8.5430	Tb927.8.5440	dss	4.763
	Tb927.8.7740	Tb927.8.7750	h-t	12.05	Tb927.8.5920	Tb927.8.5930	h-t	0.922
					Tb927.8.6920	Tb927.8.6930	dss	4.422
	Tb927.9.3130	Tb927.9.3180	dss	4.10	Tb927.9.1960	Tb927.9.1970	h-t	5.87
	Tb927.9.7280	Tb927.9.7290	CSS	2.59	Tb927.9.4900	Tb927.9.4910	h-t	0.54
9	Tb927.9.11150	Tb927.9.11220	dss	6.33	Tb927.9.9870	Tb927.9.9940	h-t	5.41
	Tb927.9.14510	Tb927.9.14530	dss	3.54	Tb927.9.8880	Tb927.9.8950	h-t	9.64
					Tb927.9.13150	Tb927.9.13160	CSS	2.12
	Tb927.9.3040	1b927.10.3060	h-t	4.06	Tb927.10.2450	1b927.10.2460	dss	1.36
	ID927.10.4960	10927.10.4970	dss	4.77	1 b927.10.4180	1 by 27.10.4190	dss	3.34
	1b927.10.6420	1b927.10.6430	dss	5.40	Tb927.10.8340	Tb927.10.8350	dss	2.64
10	грудини приставляется приставляется годинальные приставляется пристав приставляется приставлеется приставляется пристав пристав приставляется пристав	rb927.10.7640	dss	0.57	Tb927.10.11270	Tb927.10.11280	dss	4.30
	ID927.10.9670	ID927.10.9590	n-t	1.86	10927.10.12550	10927.10.12570	n-t	3.49
	16927.10.10850	16927.10.10870	h-t	12.03	16927.10.13550	10927.10.13560	h-t	5.25
				ļ	Tb927.10.14000	Tb927.10.14010	h-t	1.39
					Tb927.10.14840	Tb927.10.14860	dss	8.65

Table 7-1. (continue).

	Origins						Non-origins		
Chr	Gene left	Gene right	SSR type	origin (kb)		Gene left	Gene right	SSR type	non- origin
	Tb927.10.720	Tb927.10.730	dss	6.24		Tb927.11.3220	Tb927.11.3230	dss	8.53
	Tb927.11.1920	Tb927.11.1930	dss	5.11		Tb927.11.3560	Tb927.11.3570	h-t	4.51
	Tb927.11.4760	Tb927.11.4780	h-t	10.25		Tb927.11.6300	Tb927.11.6310	h-t	0.13
	Tb927.11.6720	Tb927.11.6730	h-t	1.01		Tb927.11.7213	Tb927.11.7214	dss	4.85
	Tb927.11.7970	Tb927.11.7980	dss	1.93		Tb927.11.8840	Tb927.11.20730	h-t	9.28
	Tb927.11.9810	Tb927.11.9820	dss	7.06		Tb927.11.11400	Tb927.11.11420	h-t	6.38
	Tb927.11.11060	Tb927.11.11080	h-t	8.90		Tb927.11.11980	Tb927.11.12020	h-t	5.80
11						Tb927.11.13710	Tb927.11.13720	dss	2.23
						Tb927.11.15380	Tb927.11.15390	h-t	5.897
						Tb927.11.16170	Tb927.11.16180	dss	3.941
						Tb927.11.6250	Tb927.11.14630	CSS	1.873
						Tb927.11.12710	Tb927.11.12720	CSS	1.389
						Tb927.11.5920	Tb927.11.5940	CSS	2.749
						Tb927.11.4120	Tb927.11.4130	CSS	1.598
						Tb927.11.2400	Tb927.11.2410	CSS	11.902

Table 7-2. *L. major* distance in Kbp between the first two genes within the SSRs, with or no origin activity detected by MFA-seq.

	Origins				Non-origins				
Chr	Gene left	Gene right	SSR type	origin (kb)	Gene left	Gene right	SSR type	non- origin	
1		unclear			LmjF.01.0315	LmjF.01.0320	dss	0.973	
2	LmjF.02.0570	LmjF.02.SLRNA.0010	dss	8.342		unclear			
2	LmjF.03.0670	LmjF.03.0690	CSS	13.304	LmjF.03.0010	LmjF.03.0020	dss	1.296	
5					LmjF.03.0970	LmjF.03.0980	h-t	0.683	
4	LmjF.04.0380	LmjF.04.0390	CSS	9.03	LmjF.04.0625	LmjF.04.0630	h-t	0.678	
5	LmjF.05.1040	LmjF.05.1050	dss	1.453	LmjF.05.0450	LmjF.05.0460	h-t	0.893	
	LmjF.06.0360	LmjF.06.0370	dss	5.747	LmjF.06.0560	LmjF.06.0570	h-t	1.351	
6					LmjF.06.1250	LmjF.06.1260	CSS	2.786	
					LmjF.06.1290	LmjF.06.1300	dss	4.79	
7	LmjF.07.0470	LmjF.07.0475	dss	7.646	LmjF.07.0010	LmjF.07.0020	dss	5.105	
·					LmjF.07.0802	LmjF.07.0805	h-t	1.032	
8	LmjF.08.1090	LmjF.08.1101	dss	11.792	LmjF.08.0860	LmjF.08.0870	CSS	3.724	
9	LmjF.09.0690	LmjF.09.0700	CSS	6.475	LmjF.09.1000	LmjF.09.1010	dss	3.464	
	LmjF.10.0600	LmjF.10.0610	h-t	6.454	LmjF.10.1227	LmjF.10.1228	dss	1.189	
10					LmjF.10.0030	LmjF.10.0040	dss	9.307	
					LmjF.10.0510	LmjF.10.0520	CSS	4.084	
11	LmjF.11.0475	LmjF.11.0480	h-t	9.055	LmjF.11.0920	LmjF.11.0930	h-t	3.031	
12	LmjF.12.0510	LmjF.12.0520	dss	10.541	LmjF.12.0400	LmjF.12.0405	CSS	2.353	
					LmjF.12.0010	LmjF.12.0020	h-t	0.416	
	LmjF.13.0450	LmjF.13.0460	dss	5.679	LmjF.13.0700	LmjF.13.0710	CSS	1.286	
13					LmjF.13.1370	LmjF.13.1380	h-t	1.025	
					LmjF.13.1680	LmjF.13.1690	dss	1.847	
14	LmjF.14.0470	LmjF.14.0480	CSS	9.477	LmjF.14.1050	LmjF.14.1060	dss	2.186	
15	LmjF.15.0740	LmjF.15.0750	CSS	8.166	LmjF.15.0223	LmjF.15.0225	dss	1.054	
					LmjF.15.1560	LmjF.15.1570	dss	4.443	
16	LmjF.16.0920	LmjF.16.0930	dss	5.697	LmjF.16.1130	LmjF.16.1140	CSS	0.949	
					LmjF.16.1520	LmjF.16.1530	dss	2.923	
17	LmjF.17.0733	LmjF.17.0790	h-t	6.894	LmjF.17.0860	LmjF.17.0870	dss	1.637	
					LmjF.17.0340	LmjF.17.0350	h-t	2.032	
18	LmjF.18.1050	LmjF.18.1060	h-t	6.521	LmjF.18.0560	LmjF.18.0570	dss	1.826	
19	LmjF.19.1420	LmjF.19.1430	h-t	9.272	LmjF.19.0980	LmjF.19.0985	h-t	5.121	
					LmjF.19.0220	LmjF.19.0230	dss	7.524	

Table 7-2. (continue).

	Origins										
Chr	Gene left	Gene right	SSR type	origin (kb)	Ge						
	LmjF.20.1175	LmjF.20.1180	h-t	7.807	LmjF						
20					LmjF						
	Lm:F 21 0720	Lm:E 21 072E	dee	7 46 5	LmjF						
	LIIIJF.21.0720	LIIIJF.21.0/25	ass	7.405	LIIIJF						
21					LmiF						
					ImiF						
	LmiF.22.1480	LmiF.22.1490	dss	6.176	LmiF						
22		,			LmjF						
					LmjF						
23	LmjF.23.1155	LmjF.23.1160	dss	5.259	LmjF						
23					LmjF						
	LmjF.24.1300	LmjF.24.1310	h-t	4.69	LmjF						
. .					LmjF						
24					LmjF						
					LmjF						
	lmiE 25 1460	lmiE 25 1470	h t	1 216	LmiE						
25	Liiiji .23.1400	LIIIJI .23.1470	11-1	1.210	LmiF						
25					LmiF						
	LmjF.26.1665	LmjF.26.1670	h-t	5.82	LmjF						
26	-				LmjF						
					LmjF						
27	LmjF.27.2337	LmjF.27.rRNA.01	dss	7.629	LmjF						
21					LmjF						
	LmjF.28.2100	LmjF.28.2110	dss	4.37	LmjF						
28					LmjF						
					LmjF						
	lmiE 20 0885	lmiE 20 0800	dee	6 124	LmjF						
29	LIIIJI .29.000J	LIIIJI .27.0070	uss	0.134	LmiF						
27					LmiF						
	LmiF.30.0710	LmiF.30.0720	dss	5.046	LmiF						
30	-				LmjF						
					LmjF						
	LmjF.31.1640	LmjF.31.1650	h-t	5.505	LmjF						
					LmjF						
31					LmjF						
					LmjF						
	lmiE 22 2095	LmiE 22 2000	dee	7 04 9	LmjF						
32	LIIIJF.32.2903	LIIIJF.32.2990	uss	7.000	Liijr						
52					ImiF						
	LmiF.33.1610	LmiF.33.1620	h-t	6.358	LmiF						
33	,	,	-		LmjF						
					LmjF						
	LmjF.34.0690	LmjF.34.0700	dss	3.402	LmjF						
34					LmjF						
					LmjF						
	LmjF.35.1190	LmjF.35.1200	h-t	4.675	LmjF						
					LmjF						
35					LMJF						
J					l miF						
					LmiF						
					LmiF						

Non-origins										
Gene left	Gene right	SSR type	non- origin							
LmjF.20.0840	LmjF.20.0850	h-t	0.715							
LmjF.20.0260	LmjF.20.0625	dss	1.32							
LmjF.20.1450	LmjF.20.1460	CSS	1.498							
LmjF.21.0015	LmjF.21.0020	dss	1 924							
LmjF.21.0520	LmjF.21.0530	CSS	1.024							
LilijF.21.1090	LilijF.21.1100	CSS	1.005							
LmjF 22 0790	LmiF 22 0800	dss	1.005							
I miF 22 1220	LmiF 22 1230	CSS	0.938							
LmiF.22.0010	LmiF.22.0020	h-t	5.944							
LmjF.23.1630	LmjF.23.1640	CSS	3.237							
LmjF.23.0010	LmjF.23.0020	h-t	1.736							
LmjF.24.0650	LmjF.24.0660	h-t	1.012							
LmjF.24.1900	LmjF.24.1905	h-t	0.518							
LmjF.24.2330	LmjF.24.2340	dss	1.239							
LmjF.24.1640	LmjF.24.1650	CSS	4.483							
LmjF.24.0010	LmjF.24.0020	h-t	0.683							
LmjF.25.0715	LmjF.25.0720	dss	2.555							
LmjF.25.1080	LmjF.25.1090	CSS	2.532							
LmjF.25.2160	LmjF.25.2170	dss	1.363							
LmjF.26.1015	LmjF.26.1020	dss	1.475							
LmjF.26.0760	LmjF.26.0770	h-t	0.811							
LmjF.26.2270	LmjF.26.2280	h-t	1./34							
LmjF.27.1265	LmjF.27.12/0	dss	1.816							
LmjF.27.0290	LmjF.27.0300	dss	3.030							
$L_{111}F.20.0770$	LilijF.20.0705	ass	0.622							
Linji .28.0330	Linji .28.0300	C35	8 18							
LmiF 28 2690	LmiF 28 2700	C55	0.10							
LmiF29.1800	LmiF.29.1810	h-t	2,489							
LmiF.29.2350	LmiF.29.2360	dss	1.331							
LmjF.29.1440	LmjF.29.1450	CSS	5.27							
LmjF.30.1730	LmjF.30.1740	h-t	0.84							
LmjF.30.2090	LmjF.30.2100	CSS	0.938							
LmjF.30.3235	LmjF.30.3240	dss	3.884							
LmjF.31.0570	LmjF.31.0580	h-t	4.409							
LmjF.31.1230	LmjF.31.1240	h-t	3.127							
LmjF.31.1990	LmjF.31.2000	h-t	6.972							
LmjF.31.2715	LmjF.31.2720	h-t	1.024							
LmjF.31.3160	LmjF.31.3170	dss	1.679							
LmjF.32.2155	LmjF.32.2160	h-t	1.434							
LmjF.32.13/0	LmjF.32.1380	CSS	4.191							
LmjF.32.0480	LmjF.32.0490	dss	1.907							
LmjF.33.1/90	LmjF.33.1800	dss	1.163							
LmjF.33.0600	LmjF.33.0605	ass b t	3.000							
Liliji .33.0290	Liliji .33.0295	h t	1 /28							
Liliji .34.1240	Linji .34.1243	dee	6 258							
LmiF 34 3520	LmiF 34 3530	h-t	1.515							
LmiF.35.0180	LmiF.35.0190	dss	2.323							
LmjF.35.1750	LmjF.35.1755	dss	3.028							
LmjF.35.1470	LmjF.35.1480	CSS	6.276							
LmjF.35.2590	LmjF.35.2600	CSS	0.814							
LmjF.35.2130	LmjF.35.2140	h-t	4.043							
LmjF.35.3470	LmjF.35.3480	h-t	3.373							
LmjF.35.3910	LmjF.35.3920	dss	1.508							

Table 7-2. (continue).

	Origins				Non-origins				
Chr	Gene left	Gene right	SSR type	origin (kb)	Gene left	Gene right	SSR type	non- origin	
	LmjF.36.2720	LmjF.36.2730	h-t	6.401	LmjF.36.0535	LmjF.35.0537	dss	1.089	
					LmjF.36.1955	LmjF.36.1960	dss	1.073	
					LmjF.36.3660	LmjF.36.3670	dss	0.823	
24					LmjF.36.4220	LmjF.36.4230	h-t	1.037	
30					LmjF.36.5365	LmjF.36.5370	h-t	3.97	
					LmjF.36.6350	LmjF.36.6360	h-t	2.112	
					LmjF.36.1350	LmjF.36.1360	CSS	6.289	
					LmjF.36.4880	LmjF.36.4890	CSS	1.442	

Table 7-3. *L. mexicana* distance in Kbp between the first two genes within the SSRs, with or no origin activity detected by MFA-seq.

		Origins					Non-origins		
Chr	Gene left	Gene right	Туре	origin (kb)		Gene left	Gene right	Туре	non- origin
1		unclear				LmxM.01.0315	LmxM.01.0320	dss	0.969
2	LmxM.02.0570	LmxM.02.0580	dss	9.91			unclear		
3	LmxM.03.0670	LmxM.03.0690	CSS	12.691		LmxM.03.0010	LmxM.03.0020	dss	1.388
						LmxM.03.0970	LmxM.03.0980	h-t	0.703
4	LmxM.04.0380	LmxM.04.0390	CSS	8.378		LmxM.04.0625	LmxM.04.0630	h-t	0.706
5	LmxM.05.1040	LmxM.05.1050	dss	1.461		LmxM.05.0450	LmxM.05.0460	h-t	0.897
	LmxM.06.0360	LmxM.06.0370	dss	4.647		LmxM.06.0560	LmxM.06.0570	h-t	1.385
6						LmxM.06.1250	LmxM.06.1250	CSS	2.537
						the dss doe	es not exist in L m	ex chr6	1
7	LmxM.07.0470	LmxM.07.0475	dss	5.695		LmxM.07.0010	LmxM.07.0020	dss	5.033
Ľ			-			LmxM.07.0802	LmxM.07.0805	h-t	0.981
8	LmxM.08.1090	LmxM.08.1091	dss	11.346		LmxM.08.0860	LmxM.08.0870	CSS	3.343
9	LmxM.09.0690	LmxM.09.0700	CSS	6.591		LmxM.09.1000	LmxM.09.1010	dss	3.66
	LmxM.10.0600	LmxM.10.6010	h-t	5.629		LmxM.10.1227	LmxM.10.1228	dss	1.178
10						LmxM.10.0030	LmxM.10.0040	dss	9.675
			-			LmxM.10.0510	LmxM.10.0520	CSS	3.857
11	LmxM.11.0475	LmxM.11.0480	h-t	8.557		LmxM.11.0920	LmxM.11.0930	h-t	3.127
12	LmxM.12.0510	LmxM12.0520	dss	8.776		LmxM.12.0400	LmxM.12.0405	CSS	2.878
			-		╽┟	LmxM.12.0010	LmxM.12.0020	h-t	0.399
	LmxM.13.0450	LmxM.13.0460	dss	5.072		LmxM.13.0700	LmxM.13.0710	CSS	1.265
13						LmxM.13.1370	LmxM.13.1380	h-t	1.012
						LmxM.13.1680	LmxM.13.1690	dss	1.833
14	LmxM.14.0470	LmxM.14.0480	CSS	8.147	╎╎	LmxM.14.1050	LmxM.14.1060	dss	2.212
15	LmxM.15.0740	LmxM.15.0750	CSS	6.619		LmxM.15.0223	LmxM.15.0225	dss	1.049
					╎╎	LmxM.15.1560	LmxM.15.1570	dss	4.47
16	LmxM.16.0920	LmxM.16.0930	dss	4.634		LmxM.16.1130	LmxM.16.1140	CSS	0.954
					╎╎	LmxM.16.1520	LmxM.16.1530	dss	2.869
17	LmxM.17.0733	LmxM.17.0790	h-t	5.666		LmxM.17.0860	LmxM.17.0870	dss	1.59
					╎╎	LmxM.17.0340	LmxM.17.0350	h-t	2.048
18	LmxM.18.1050	LmxM.18.1060	h-t	4.814	╎╎	LmxM.18.0560	LmxM.18.0570	dss	1.904
19	LmxM.19.1420	LmxM.19.1430	h-t	6.663		LmxM.19.0980	LmxM.19.0985	h-t	5.007
					╽┟	LmxM.19.0220	LmxM.19.0230	dss	7.342
	LmxM.20.1175	LmxM.20.1180	h-t	6.661		LmxM.20.0840	LmxM.20.0850	h-t	0.694
20						LmxM.20.0260	LmxM.20.0625	dss	1.313
					╽┟	LmxM.20.1450	LmxM.20.1460	CSS	1.467
	LmxM.21.0720	LxmM.21.0725	dss	6.302		LmxM.21.0015	LmxM.21.0020	dss	1.043
21						LmxM.21.0520	LmxM.21.0530	CSS	1.765
<u> </u>						LmxM.21.1090	LmxM.21.1100	CSS	1.727
					۱L	LmxM.21.1720	LmxM.21.1730	dss	1.609

Table 7-3. (continue).

	Origins					Non-origins				
Chr	Gene left	Gene right	Туре	origin (kb)		Gene left	Gene right	Туре	non- origin	
	LmxM.22.1480	LmxM.22.1490	dss	5.98] [LmxM.22.0790	LmxM.22.0800	dss	2.112	
22						LmxM.22.1220	LmxM.22.1230	CSS	0.842	
						LmxM.22.0010	LmxM.22.0020	h-t	5.832	
23	LmxM.23.1155	LmxM.23.1160	dss	4.65		LmxM.23.1630	LmxM.23.1640	CSS	3.121	
						LmxM.23.0010	LmxM.23.0020	h-t	1.611	
	LmxM.24.1300	LmxM.24.1310	h-t	4.157		LmxM.24.0650	LmxM.24.0660	h-t	0.938	
						LmxM.24.1900	LmxM.24.1905	h-t	0.507	
24						LmxM.24.2330	LmxM.24.2340	dss	1.225	
						LmxM.24.1640	LmxM.24.1650	CSS	4.23	
			L .	4 202	┥┝	LmxM.24.0010	LmxM.24.0020	n-t	0.678	
25	LmxM.25.1460	LmxM.25.14/0	n-t	1.202		LmxM.25.0/15	LmxM.25.0/20	dss	2.517	
25						LmxM.25.1080	LmxM.25.1090	CSS	2.333	
	L muth 26 466	L	b 4	E (12		LmxM.25.2160	LmxM.25.2170	ass	1.357	
24	LMXM.20.1000	LMXM.20.1070	n-t	5.012		LmxM.26.1015	LmxM.26.1020		1.404	
20						LmxM.26.0760	LmxM.20.0770	n-t	1 924	
		myth 27.rDNA.rfam				LIIIX/M.20.2270	LIIIXM.20.2200	n-t	1.030	
27	LmxM.27.2337 ^L s	can:982402-983034	dss	14.154		LmxM.27.1265	LmxM.27.1270	dss	1.77	
						LmxM.27.0290	LmxM.27.0300	dss	3.648	
	LmxM.28.2100	LmxM.28.2110	dss	3.946		LmxM.28.0770	LmxM.28.0785	dss	1.097	
28						LmxM.28.0350	LmxM.28.0360	CSS	0.701	
						LmxM.28.1570	LmxM.28.1580	CSS	8.415	
						LmxM.28.2690	LmxM.28.2700	CSS	0.758	
						LmxM.08_29.0890	LmxM.08_29.0885	dss	3.628	
8	-	-	-	-		LmxM.08_29.1810	LmxM.08_29.1800	h-t	2.495	
						LmxM.08_29.1440	LmxM.08_29.1450	CSS	5.352	
	Lmv44 20 0710	Lmv44 20 0720	dee			LmxM.08_29.2360	LmxM.08_29.2330	dss b.t	1.495	
20	LIIIX/M.29.0710	LIIIXM.29.0720	ass	3.719		$L_{111X}M.29.1730$	LIIIXM.29.1740	11-L	0.000	
29						LIIIX/W.29.2090	LIIIXM.29.2100	CSS dcc	2 826	
	1 mxM 30 1640	LmxM 30 1650	h-t	1 103		LIIIX/M.29.3233	LmxM 30 0571	uss h-t	5.030	
	LIIIXM.30.1040	LIIXM. 30. 1030	11-0	4.105		LmxM 30 1230	LmxM 30 1240	h-t	2 952	
30						LmxM 30 1990	LmxM 30 2000	h-t	67	
50						LmxM 30 2715	LmxM 30 2720	h-t	0.991	
						LmxM 30 3160	LmxM 30 3170	dss	1 681	
	LmxM.31.2985	LmxM.31.2990	dss	6.307	1	LmxM.31.2155	LmxM.31.2160	h-t	1.435	
31						LmxM.31.1370	LmxM.31.1380	CSS	3.042	
						LmxM.31.0480	LmxM.31.0490	dss	1.899	
	LmxM.32.1610	LmxM.32.1620	h-t	5.452	11	LmxM.32.1790	LmxM.32.1800	dss	1.143	
32					H	LmxM.32.0600	LmxM.32.0605	dss	2.807	
					H	LmxM.32.0290	LmxM.32.0295	h-t	3.939	
	LmxM.33.0690	LmxM.33.0700	dss	3.464	11	LmxM.33.1240	LmxM.33.1245	h-t	1.427	
33					lí	LmxM.33.2530	LmxM.33.2540	dss	6.211	
						LmxM.33.3520	LmxM.33.3530	h-t	1.668	
	LmxM.34.1190	LmxM.34.1200	h-t	3.981	ן ן	LmxM.34.0180	LmxM.34.0190	dss	2.426	
						LmxM.34.1750	LmxM.34.1755	dss	2.936	
						LmxM.34.1470	LmxM.34.1480	CSS	5.334	
34						LmxM.34.2590	LmxM.34.2610	CSS	1.003	
						LmxM.34.2130	LmxM.34.2140	h-t	4.037	
						LmxM.34.3470	LmxM.34.3480	h-t	3.148	
						LmxM.34.3910	LmxM.34.3920	dss	1.547	

Table 7-3. (continue).

		Origins				Non-origins		
Chr	Gene left	Gene right	Туре	origin (kb)	Gene left	Gene right	Туре	non- origin
					LmxM.36.2720	LmxM.36.2730	h-t	3.549
					LmxM.36.0535	LmxM.35.0537	dss	1.101
					LmxM.36.1955	LmxM.36.1960	dss	1.079
					LmxM.36.3660	LmxM.36.3670	dss	0.662
20	-	-	-	-	LmxM.36.4220	LmxM.36.4230	h-t	1.089
					LmxM.36.5365	LmxM.36.5370	h-t	3.122
					LmxM.36.6350	LmxM.36.6360	h-t	2.098
					LmxM.36.1350	LmxM.36.1360	CSS	5.217
					LmxM.36.4870	LmxM.36.4890	CSS	1.394

List of References

- Archer, S. K., Inchaustegui, D., Queiroz, R. & Clayton, C. (2011). The cell cycle regulated transcriptome of *Trypanosoma brucei*. *PLoS One*, **6**, e18425.
- Godoy, P. D., Nogueira-Junior, L. A., Paes, L. S., Cornejo, A., Martins, R. M., Silber, A. M., Schenkman, S. & Elias, M. C. (2009). Trypanosome prereplication machinery contains a single functional orc1/cdc6 protein, which is typical of archaea. *Eukaryot Cell*, 8, 1592-603.
- Tiengwe, C., Marcello, L., Farr, H., Dickens, N., Kelly, S., Swiderski, M., Vaughan, D., Gull, K., Barry, J. D., Bell, S. D. & McCulloch, R. (2012). Genome-wide analysis reveals extensive functional interaction between DNA replication initiation and transcription in the genome of *Trypanosoma brucei*. *Cell Rep*, 2, 185-97.