UNIVERSITY OF SÃO PAULO

RIBEIRÃO PRETO MEDICAL SCHOOL

POST GRADUATION PROGRAM OF CELLULAR AND MOLECULAR
BIOLOGY AND PATHOGENIC BIOAGENTS

Characterisation of the role of the ATR kinase homolog in *L. major* response to replication stress.

GABRIEL LAMAK ALMEIDA DA SILVA

RIBEIRÃO PRETO

2022

GABRIEL LAMAK ALMEIDA DA SILVA

Characterisation of the role of the ATR kinase homolog

in *L. major* response to replication stress.

Thesis presented to the postgraduate program of Cellular and Molecular Biology from Ribeirão Preto Medical School at University of São Paulo, as a requirement to obtain the title of doctor in Science.

Specific Area: Cellular and Molecular Biology

Supervisor: Prof. Dr. Luiz Ricardo Orsini Tosi

RIBEIRÃO PRETO

2022

---

da Silva, Gabriel Lamak Almeida

Characterisation of the role of the ATR kinase homolog

in L. major response to replication stress, 2022.

140 p.: il.; 30 cm

Phd Thesis, presented at Ribeirão Preto Medical School/USP.

Field: Cellular and Molecular Biology.

Supervisor: Tosi, Luiz Ricardo Orsini

1. ATR kinase   2. DNA damage response   3. Leishmania
4. DNA Replication   5. Genome Maintanence

Name: Gabriel Lamak Almeida da Silva

Title of the work: Characterisation of the role of the ATR kinase homolog

in L. major response to replication stress.

> Thesis presented to the postgraduate department of Cellular and Molecular Biology and Pathogenic Bioagents from Ribeirão Preto Medical School at University of São Paulo, as a requirement to obtain the title of doctor in Science.
>
> Specific Area: Cellular and Molecular Biology

Approved in: ___/___/____

Examiners

Prof. Dr. Luiz Ricardo Orisini Tosi

Institution: Ribeirão Preto Medical School – University of São Paulo

Judgement: _____ Signature: _____

Prof.(a) Dra. Leticia Fröhlich Archangelo

Institution: Ribeirão Preto Medical School – University of São Paulo

Judgement: _____ Signature: _____

Prof.(a) Dra. Julia Pinheiro Chagas Cunha

Institution: Butantã Institute – São Paulo

Judgement: _____ Signature: _____

Prof. Dr. José Renato Rosa Cussiol

Institution: Biochemistry Department – Federal University of São Paulo

Judgement: _____ Signature: _____

Dedication

Dedico este trabalho a memória das minhas tias: Alexandrina, Antônia, Mariquinha e todos os familiares que faleceram; que junto à minha mãe, não mediram esforços para garantir que eu e meus irmãos tivéssemos "estudo".

Dedico este trabalho também a todos os professores brasileiros, em especial: aos do Externato Tico e Teco, CDS, UFRN e USP; por todos os dias superarem as adversidades, dar oportunidade e mudança na vida de muitos alunos brasileiros.

*The dedication is written in Portuguese
to address some people that cannot read in English.

Acknowledgment

Ao(s) Deus(es) e Nossa senhora Sant'ana de Caicó que de alguma forma, possam ter me guiado durante toda minha trajetória.

Ao Prof. Dr. Luiz Ricardo Orsini Tosi, pela amizade, confiança, orientação e ter acreditado em mim. Obrigado pela oportunidade e paciência que teve para que eu pudesse desenvolver minhas habilidades cientificas durante esses anos. Eu não me imagino sendo o profissional que eu sou, se não tivesse as discussões, conversas e ensinamentos diários. Sou eternamente grato.

A todos os colegas que estão ou passaram pelo Laboratório Molecular de *Leishmanias*: Renato, Jeziel, Elaine, Ricardo, Stela, Matheus e Jennifer. Por todos os ensinamentos e colaboração durante esses anos.

A todos os professores do Departamento de Biologia Molecular e Celular da FMRP-USP que estiveram comigo desde o início do meu doutorado; muito obrigado pelos ensinamentos e discussões científicas. Em especial a Prof.(a) Angela Cruz pela amizade e exemplo, muito obrigado.

Aos funcionários do Departamento de Biologia Molecular e Celular da FMRP-USP em especial: Tânia, Viviane, Bete, Roberta, Fabiana (farmácia), Silvinha, Gabriella, Camila, Lúcia e Ricardo; pela amizade e suporte durante os experimentos.

To Prof. Richard McCulloch, Dr. Jeziel Damasceno, all from McCulloch's lab to host me during my internship at University of Glasgow (2020-2021); thank you for guide me during my last PhD year during such challenging times.

A todos amigos que fiz durante meu doutorado em especial aos do grupo do cocô: Natalia Melquie, Natalia Barbosa, Marjorie, Lorenzon, Pedro Gobira, Ítalo, Rafael Ricci, Rafael "spider" e Juliana. Aos do grupo dos Emocionados: Amanda, Camila, Lucão, Isadora e Letícia. Aos amigos do futebol: Murilo, Tiozão, Flávio Protásio, e todos os outro pernas-de-pau. Aos amigos e

*The acknowledgment is written in Portuguese
Or in English to direct address the target person or people.   **VII**

 funcionários do CEFER-USP: Felipe, Abel, Leonardo, dentre outros. Foram muitos momentos de alegria e descontração que compartilhei com todos e com certeza foram importantes para superar dificuldades e a distância da família.

Aos amigos da república bicho-solto: Vinícius Detoni, Warrison (leitim), Renan de Carvalho, João e Israel. Obrigado pela irmandade durante meu doutorado, vocês estarão sempre no meu coração.

To all friends in Glasgow: Marco, Hérve, Vicky, Dan, Félix and all from the institute. You guys showed me the best feature of Glasgow, thanks.

Special Acknowledgment

A minha mãe, Ana Maria de Almeida, aos meus irmãos, José Lucas Almeida da Silva e Matheus Almeida da Silva; pelo amor incondicional apesar da distância, incentivo e confiança. Obrigado pelo apoio mesmo em momentos difíceis.

A minha namorada e companheira, Olivia Kemp, pelo amor, carinho, confiança e apoio durante momentos difíceis. Obrigado por sempre estar do meu lado, te amo.

To David and Susan Kemp, Cameron and Pamela Mcgrail for all support and friendship during my first time in Glasgow.

*The acknowledgment is written in Portuguese
Or in English to direct address the target person or people.

Title: Characterisation of the role of the ATR kinase homolog in L. major response to replication stress.

Abstract

Eukaryotic cells have evolved mechanisms to maintain and replicate their genomes, whose integrity and transmission are constantly challenged by DNA damage and replication impediments. The protein kinase Ataxia-Telangiectasia and Rad3-related (ATR), a member of the phosphatidylinositol 3-kinase-like family, ensures genome maintenance and stability, considered as a master regulator of the eukaryotic response to DNA injuries. In this PhD thesis, I investigate the conservation and functional relevance of the ATR homolog in the DNA metabolism of *Leishmania major*, a protozoan parasite with a remarkably plastic genome. Herein, CRISPR/cas9 genome editing was used to generate a Myc-tagged ATR cell line (^{myc}ATR), a heterozygous C-terminal deletion ATR cell line and a tagged at the N-terminal of RPA1 (ATR$^{\Delta C+/-mycRPA}$), and a Myc-tagged C-terminal knockout ATR (^{myc}ATR$^{\Delta C}$). Our findings show that ATR localisation depends upon the C-terminal region and is not limited to the nuclear compartment, but also located to the kinetoplast and in another extra nuclear localisation. The ATR$^{\Delta C+/-mycRPA}$ cells accumulate single-stranded DNA at the kinetoplast and at the nuclear under or not hydroxyurea treatment. In contrast, we observe abolishment of the possible S-G2/M checkpoint in ^{myc}ATR$^{\Delta C}$ cells, leading to aberrant and increased DNA content. However, the complete loss of ATR C-terminal domain did not hinder the replication initiation program. In fact, both cells, ^{myc}ATR and ^{myc}ATR$^{\Delta C}$ cells, use alternative replication sites to maintain the internal chromosomes replication. ATR C-terminal domain is also important to guarantee the replication of the sub-telomeric regions under replicative stress. Finally, complete loss of ATR C-terminal region leads to increased DNA damage and we detect an increased number of aberrant cells accumulating after replication stress exposure. Altogether, the results suggest that the *L. major* ATR kinase is important during cell cycle control and genome maintenance under replication stress, preventing the accumulation of ssDNA at the nucleus and kinetoplast; as well protecting the replication at sub-telomeric regions.

Key words: ATR kinase; DNA damage response; *Leishmania*; DNA Replication; Genome Maintanence.

Título: Caracterização do papel do homologo da quinase ATR na resposta ao estresse replicativo em *L. major*.

Resumo

As células eucarióticas desenvolveram mecanismos para manter e replicar seu genoma cuja integridade e transmissão são constantemente desafiadas por danos ao DNA e impedimentos da replicação. A proteína quinase Ataxia-Telangiectasia and Rad3-related (ATR), um membro da família phosphatidylinositol 3-kinase-like, garante a manutenção e a estabilidade do genoma, e é um dos principais reguladores da reposta a danos no DNA de eucariotos. Nesta tese de doutorado, eu investiguei a conservação e a relevância funcional do homologo da ATR no metabolismo do DNA de *Leishmania major*, um parasito protozoário com um singular genoma plástico. Usando o sistema de CRISPR/cas9 para edição do genoma nós conseguimos gerar uma linhagem com a quinase ATR fusionada a uma tag de Myc (^{myc}ATR); uma linhagem heterozigota para a deleção da região C-terminal da quinase com uma tagging de myc no N-terminal da proteína RPA1 (ATR$^{\Delta C+/-mycRPA}$); e uma linhagem que expressa uma tag de Myc e homozigótica para a deleção da região C-terminal da ATR (^{myc}ATR$^{\Delta C}$). Nossos resultados mostram que a localização da quinase ATR é dependente da região C-terminal e não é só limitada ao núcleo, mas também localizada no kinetoplasto eoutra localização extranuclear. As células ATR$^{\Delta C+/-mycRPA}$ acumulam DNA de fita simples no kinetoplasto e no núcleo na presença ou não de hidroxiureia. Entretanto, o S-G2/M checkpoint parece ser perdido nas células ^{myc}ATR$^{\Delta C}$, levando a um acúmulo de células aberrantes e um aumento do conteúdo de DNA. A perda total da região C-terminal da ATR não afetou o programa de iniciação da replicação. De fato, ambas linhagens, ^{myc}ATR and ^{myc}ATR$^{\Delta C}$, usam sítios alternativos de replicação para manter a replicação nas regiões internas do cromossomo. A região C-terminal da ATR também é importante para garantir a replicação das regiões sub-teloméricas sob estresse replicativo. Por fim, a perda total da região C-terminal da ATR quinase leva a um aumento do marcador de dano no DNA yH2A e de células aberrantes após estresse replicativo. Portanto, os resultados sugerem que a ATR de *L. major* é importante para o controle do ciclo celular, bem como para a manutenção do genoma sob estresse replicativo, prevenindo o acúmulo de fita simples de DNA no núcleo e no kinetoplasto, com também protegendo a replicação das regiões sub-teloméricas.

Palavras-chave: quinase ATR, resposta ao dano no DNA, *Leishmania*, Replicação do DNA, Manutenção do genoma.

Index

Figure List

Figure 1 – Leishmaniasis Clinical Manifestation.

Figure 2 - Status of endemicity leishmaniasis reported in the world in 2020.

Figure 3 - *Leishmania* Life Cycle.

Figure 4 – Current models of circular and linear amplification in *Leishmania*.

Figure 5 – PIKK family kinases: Recruitment and Activation in Response to DNA Damage.

Figure 6 - Predicted domain locations in Trypanosomatid PIKKs compared with their human homologs.

Figure 7 - The PIKK-driven DDR pathways in Trypanosomatid parasites and the canonical Eukaryotic pathways.

Figure 8 - Workflow of MFA-seq analysis.

Figure 9 - Deficient ATR cells accumulate ssDNA and RPA signal in comparison to control cells under hydroxyurea treatment.

Figure 10 - *L. major* ATR kinase localises to nuclear and extranuclear compartments.

Figure 11 - *L. major* ATR kinase kinetoplast location duplicates in cells synthetising DNA.

Figure 12 - *L. major* ATR kinase subcellular location depends on the kinase C-terminal.

Figure 13 - ^{myc}ATR partially co-localises with DNA synthesis.

Figure 14 - Loss of ATR C-terminal leads to increased DNA content and nuclear area.

Figure 15 - ^{myc}ATR$^{\Delta C}$ cells show an altered cell cycle progression upon and after replication stress.

Figure 16 - ^{myc}ATR$^{\Delta C}$ cells show an increase in cells with abnormal DNA content after HU treatment.

Figure 17 - Loss of ATR C-terminal decreases the protein levels of the kinase and correlates with increased yH2A levels after HU treatment.

Figure 18 - ^{myc}ATR$^{\Delta C+/-}$ cells do not show cell cycle defects or an increasement of yH2A levels in comparison with ^{myc}ATR cells.

Table List*

*The tables are at the supplementary section Item 7

Abbreviation List

WHO: World Health Organization

VL: Visceral Leishmaniasis

PKDL: Post-kala-azar dermal Leishmaniasis

CL: Cutaneous Leishmaniasis

Mb: Megabase

PTU: polycistronic transcription unit

Pol II: RNA polymerase II

AcH3: acetylated histone H3

DHFR-TS: dihydrofolate reductase-thymidylate synthase

MTX: methotrexate (an antifolate compound)

CNV: copy number variation

IR: inverted repeats

DSBs: double strand breaks

ssDNA: single strand DNA

DDR: DNA damage response

PIKKs: phosphoinositide 3-kinase (PI3K)-related kinases

ATM: ataxia-telangiectasia mutated

DNA-PK: DNA-dependent protein kinase

ATR: Ataxia telangiectasia and Rad3-related protein

CHK2: checkpoint kinase protein 2

CHK1: checkpoint kinase protein 1

MRN: MRE11-RAD50-NBS1 complex

NHEJ: non-homologous end joining

HR: homologous recombination

H2AX: histone H2AX

HU: hydroxyurea

RPA: Replication Protein A (a heterotrimeric ssDNA binding protein complex)

dNTP: Deoxyribonucleotide triphosphate

yH2A: phosphorylated H2A (the trypanosomatid equivalent of H2AX)

VSG: Variant Surface Glycoprotein

sgRNA: single guide RNA

MFA-seq: Marker Frequency Analysis followed by next generation sequencing

Kb: kilobase

Introduction

1. Introduction

1.1 Leishmania and Leishmaniasis

Leishmaniasis are a group of diseases caused by the protozoan parasites *Leishmania*. According to World Health Organization (WHO), more than 20 *Leishmania* species cause zoonotic diseases in around 70 animal species including rodents, dogs, and humans (Burza, Croft, Boelaert, 2019). *Leishmania* is transmitted to the mammalian vertebrate host through bites of infected female sandflies, of which over 90 species have been identified as disease transmitters, according to World Health Organization, 2020 (Maroli et al., 2013). Leishmaniasis continues to be a major health and economic problem affecting around 35 million people under vulnerable conditions in the Americas, Africa, and East Asia (WHO, 2017); being included in the group of diseases called Neglected Tropical Diseases (NTDs) (WHO, 2021). No vaccines are currently available for human usage and current treatment, mostly based on repurposed drugs normally used to control other diseases, has severe side effects and coincides with increased reports of field drug resistance. Alternatively, increasing vaccination of animals using the currently available vaccines such as Leishmune[@], CaniLeish[@], and Leish-Tec could help control and prevent Leishmaniasis (Moafi et al., 2019).

There are four main forms of the disease: Visceral Leishmaniasis (VL, also known as Kala-azar), which is characterized by irregular bouts of fever, weight loss, enlargement of the spleen and liver, and anaemia. Post-kala-azar dermal Leishmaniasis (PKDL), which is a cutaneous form derived from VL, and Cutaneous Leishmaniasis (CL) are forms that cause skin lesions, mainly ulcers (Figure 1). The Mucosal or Mucocutaneous Leishmaniasis (ML) leads to considerable destruction of mucous membranes of the nose, mouth, and throat (Figure 1). Whilst CL is the most common form of the disease, VL is the most serious, being fatal in more than 95% of undiagnosed or untreated cases (WHO, 2020).

Figure 1 – Leishmaniasis Clinical Manifestation.
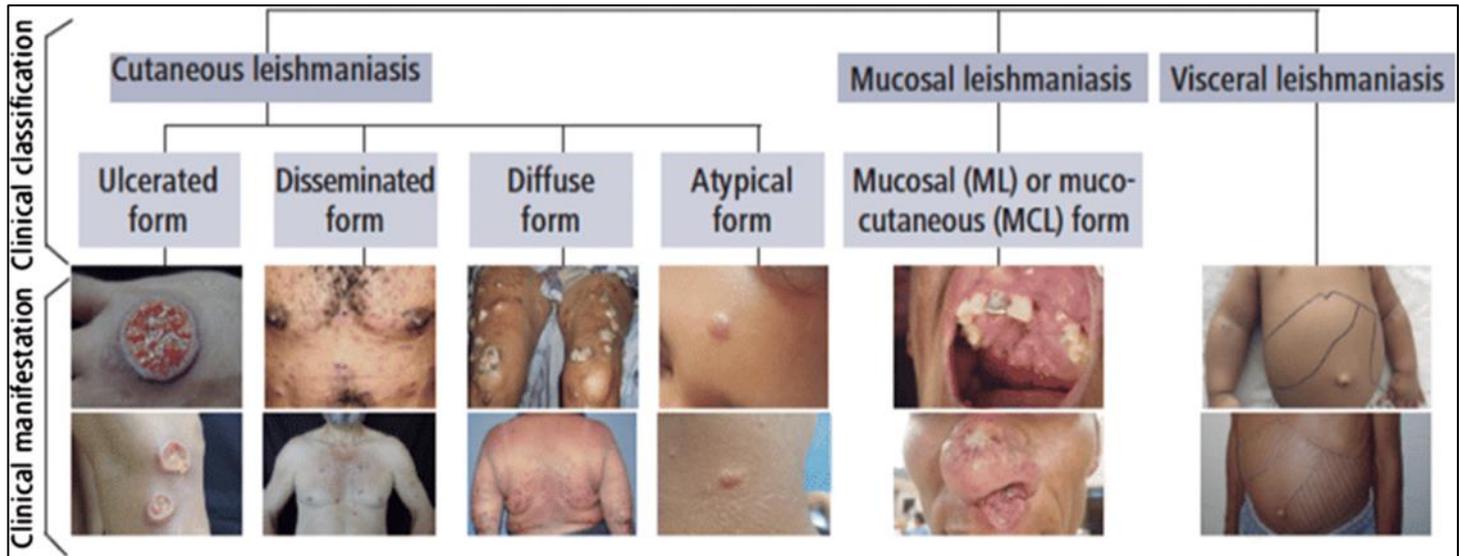
Figure 1 – Leishmaniasis Clinical Manifestation. Clinical classification for leishmaniasis diagnosis in Americas (Reproduced from: Mann et al., 2021).

1.1.1 Epidemiology

Leishmaniasis is widely distributed and considered endemic in 49% of the countries and territories that report to the WHO. From 2015 to 2020 more than 1 million new cases of CL and more than 90,000 new cases of VL were reported, with 208,357 new cases of CL and 12,838 new cases of VL reported in 2020 alone (WHO, 2021). The last report released by the WHO revealed a higher number of CL cases in Middle Eastern countries and in the Americas, especially Brazil and Colombia (Figure 2 - Map *A*), with each of them reporting more than 6,000 new cases (WHO, 2021). In Latin America, Brazil represents more than 96% of the VL cases (Figure 2 - Map *B*) (Maia-Elkhoury et al., 2021), with most cases concentrated at border regions where there is high deforestation activity and urbanization (de Oliveira, Moreno, 2021). In fact, several studies showed the presence of the sandfly species *Lutzomyia cruzi* (Oliveira et al., 2016) and several dogs (the main reservoir) diagnosed with leishmaniasis in those regions (Lopez et al., 2021).

In addition, increased risk of developing leishmaniasis and its severe forms correlates with poor housing and sanitary conditions, favouring sandfly breeding and increasing their access to humans. Families in endemic areas with diets lacking protein-energy, iron, vitamin A and zinc have increased risk of developing more severe

symptoms of the disease (Lima Maciel et al., 2014). Finally, climate change can+ strongly impact the vectors ecology and, consequently, reproduction and the development of *Leishmania* promastigotes in sandflies (González et al., 2010; WHO, 2020).

Figure 2 - Status of endemicity leishmaniasis reported in the world in 2020.

A - Status of endemicity of cutaneous leishmaniasis (CL) worldwide, 2020

Countries reporting imported CL cases – *Pays notifiant des cas importés de LC*

Brazil – Brésil: 291
Lebanon – Liban: 239
France: 170
United Arab Emirates – Émirats arabes unis: 41
Saudi Arabia – Arabie saoudite: 32

Belgium – Belgique: 27
Ecuador – Équateur: 22
Egypt – Égypte: 19
Qatar: 15

Colombia – Colombie: 7
Argentina – Argentine: 6
Suriname: 4
Kuwait – Koweït: 3

Belarus – Biélarussie: 1
Russian Federation – Fédération de Russie: 1
Venezuela (Bolivarian Republic of) – Venezuela (République bolivarienne du): 1

Number of new CL cases reported in 2020 – *Nombre de nouveaux cas de LC notifiés en 2020*

- 0 cases reported
- <100
- 100–999
- 1000–4999
- ≥5000
- No autochtonous cases reported – *Pas de cas autochtone notifié*
- No data – *Aucune donnée*
- Not applicable – *Non applicable*

B - Status of endemicity of visceral leishmaniasis (VL) worldwide, 2020
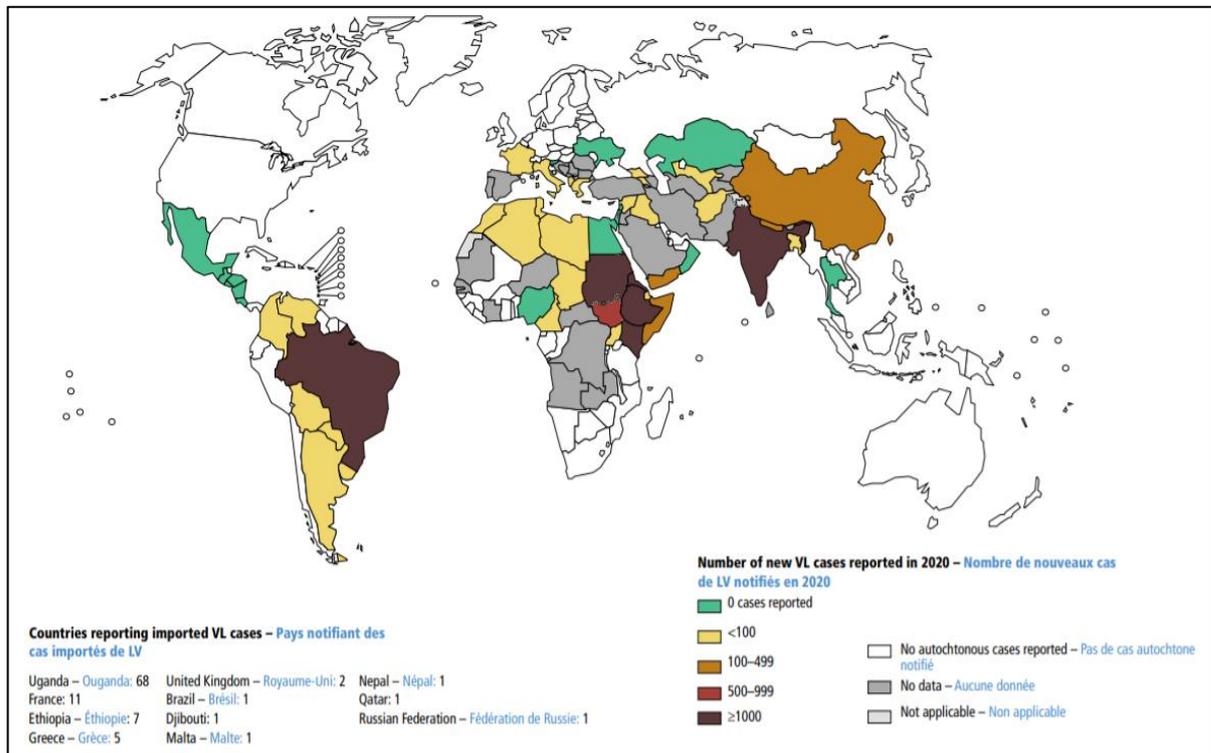
Figure 1 - Status of endemicity leishmaniasis reported in the world in 2020. The distribution of the Cutaneous Leishmaniasis (CL) around the world is show on the Map A, while the distribution of Visceral Leishmaniasis (VL) around the world is shown on Map B. (Reproduced from: Ruiz-Postigo , Jain et al. 2021)

1.1.2 Life Cycle

Leishmania parasites have a complex digenetic life cycle alternating between promastigotes, a flagellated form in the invertebrate vector (sandfly), and amastigotes, in the vertebrate host (mammalian) (Harhay, Oliaro, 2011). During their blood meals, the infected female sandflies regurgitate the parasite metacyclic form, characterized by the longer flagellum in comparison to the parasite body. At the insect bite, promastigotes are taken by phagocytic cells and differentiate into the round amastigotes form in which the flagellum is short and immotile (Kaye, Scott, 2011).

Leishmania amastigotes are well adapted to multiply in infected host cells. Following rounds of replication, the parasite burden eventually overwhelms the host cell, rupturing the cell membrane and allowing the parasites to infect other cells. The free amastigotes are then phagocytosed by new macrophages allowing the infection to reach different tissues. The life cycle is complete when sandflies become infected during a bloodmeal from an infected vertebrate host (Kaye, Scott, 2011). In the sandfly

gut, the amastigotes differentiate into procyclic promastigotes (short-flagellated and weakly motile form), stimulated by the internal vector environment changes such as decrease of temperature and increase of pH (Dostálová, Volf, 2012). The procyclic promastigotes multiply in the sandfly gut and, after 48 hours in order to establish the infection, they differentiate into strongly motile and long forms called nectomonad promastigotes. The nectomonad promastigotes may migrate to the sandfly anterior midgut and differentiate into a shorter parasite form called leptomonad promastigotes (Dostálová, Volf, 2012). This form is responsible for producing an important factor for the *Leishmania* transmission process called promastigote secretory gel (Rogers, Chance, Bates, 2002). The differentiation of leptomonad promastigotes into metacyclic prosmastigotes, with a smaller body and long flagellum responsible for faster motility, eventually occurs in the proximity of the vector proboscis, being able to infect another vertebrate (Figure 3) (Dostálová, Volf, 2012).

Figure 3 - *Leishmania* Life Cycle
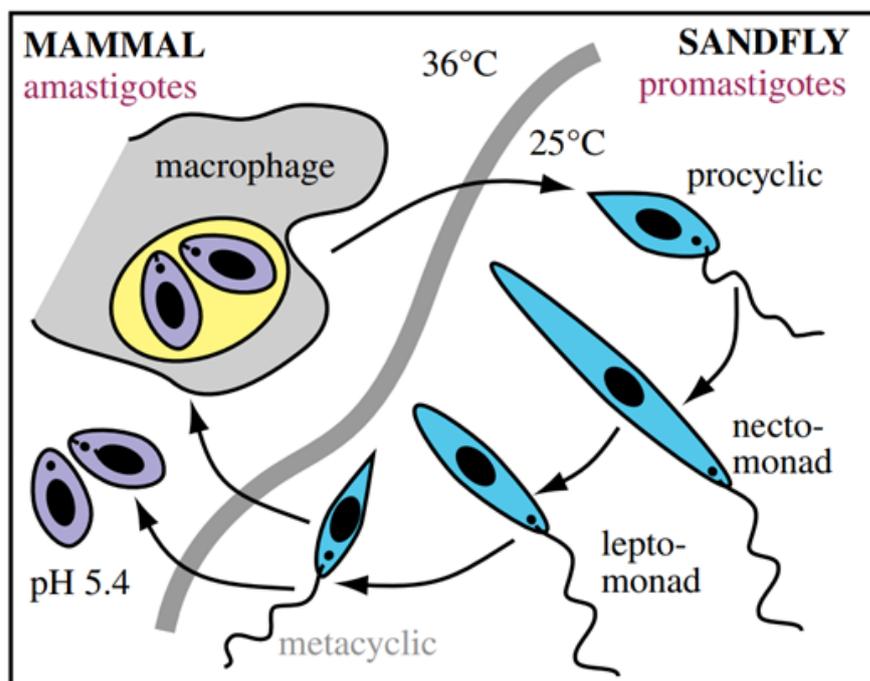
Figure 3 - *Leishmania* Life Cycle. In an infected female sandfly, the *Leishmania* Procyclic Promastigotes proliferate in the midgut and migrate to the sandfly proboscis; there they differentiate into Metacyclic Promastigotes. During the sandfly blood meal, the Metacyclic promastigotes invade the mammalian blood system and they are phagocytosing by macrophages and neutrophiles. Inside the macrophages the *Leishmania* Metacyclic promastigotes differentiate into Amastigotes and begin to multiply. Eventually the macrophages are disrupted, releasing the amastigotes allowing infection of other cells. When a female sandfly feeds from an infected vertebrate host, the insect absorbs the amastigotes from the blood and within the sandfly midgut the Leishmania differentiates again into Procyclic Promastigotes, starting the cycle again. (Reproduced from: Clayton, 2019)

1.2 *Leishmania* Genome Organisation.

The *Leishmania* genome is approximately 32Mb in size, typically distributed across 34 or 36 chromosomes, depending upon the *Leishmania* species (Ivens et al., 2005). The genome contains at least 8300 encoded genes (Ivens et al., 2005). Amongst species, for instance *L. major, L. infantum, L. braziliensis* conservation of the genome can reach 99% (Peacock et al. 2007). Despite the high synteny, chromosome size polymorphism is a typical feature of the Leishmania genome (Pagès et al., 1989). Notably, structural chromosome modifications are found in *L. mexicana* which show a fusion of chromosomes 8 and 29 with chromosomes 20 and 36; while in *L. braziliensis* the fusion happened between chromosomes 20 and 34 (Peacock et al. 2007).

Leishmania parasite belong to the Trypanosomatidae family in the Kinetoplastida class. In contrast to most Eukaryotes, trypanosomatids organise their genes into polycistronic transcription unit (PTU). Genes encoded in a PTU are not functionally related. PTUs are transcribed by RNA polymerase (Pol) II. In these parasites, transcription initiation does not appear to correlate with defined promotor sequences, but instead, initiates at sites associated with histone modifications, such as the acetylation of histone H3 (AcH3) (Martínez-Calvillo et al. 2003; Thomas et al., 2009). The transcription termination at the end of PTU seems to correlates with the location of a glycosylated T base on the DNA sequence, termed base-J (van Luenen et al., 2012). The polycistronic pre-mRNA is processed by *trans*-splicing and polyadenylation to generate mature mRNA transcripts. This is achieved through the through the addition of a 39 nucleotides RNA sequence, termed the splice leader (SL) at the 5`-end, and a polyadenine (Poly-A) tail at the 3'-end of each individual mRNA (Haile, Papadopoulou, 2008; Matthews et al., 1994).

To overcome an apparent lack of classical transcription regulation, *Leishmania* has evolved other mechanisms to control mRNA abundance and/or availability by post-transcription regulation and gene dosage variation (Iantorno et al., 2017; Rochette et al., 2008). As evidence of post-transcriptional regulation, many transcripts showed altered and regulated levels upon protein synthesis inhibition (Clayton, Shapira, 2007). Furthermore, the 3'UTR regions of genes also seem to play an important role in mRNA stabilization and their translation rate (Haile, Papadopoulou, 2008; Müller,

Padmanabhan, Papadopoulou, 2010). Recently, it was demonstrated that the presence of a 16-mer motif in the 3' untranslated region of VSG genes is important to guarantee the m^6A residues at the poli(A) tail of the actively expressed VSG transcripts in *T. brucei*, protecting the mRNA from degradation (Viegas et al., 2022).

For *Leishmania*, the intrinsic plasticity of its genome may offer a means of gene expression regulation through alterations to result in gene and chromosome copy numbers (i.e copy number variation; CNV) (Laffitte et al., 2016), gene amplification or ploidy mosaicism (Ubeda et al., 2014). Those features have been linked with *Leishmania* environment adaptation, drug resistance and changes in tissue tropism (Alvar et al., 2012; Finkel, Holbrook, 2000). Fast-growing eukaryotic cells, such fungi and cancer cells, show the same genome adaptation that occurs in *Leishmania* (Gordon, Resio, Pellman, 2012; Shor, Perlin, 2015).

1.3 Copy Number Variation and the proposed models for CNV generation in *Leishmania*.

The first studies that shed light on the mechanism of extrachromosomal gene amplification in *Leishmania* used multiple rounds of selection in increased drug concentration (Haimeur, Ouellette, 1998; Beverley et al. 1984). The products of these events usually present different structures, mainly circular and/or linear amplicons, which can be differentiated by Pulse Field Gel Electrophoresis, based on the differences in topology between linear or circular DNA molecules (Gómez, doctoral dissertation, 2017).

A bioinformatic screens identified that direct and inverted repeated sequence elements are generally non coding and highly conserved among *Leishmania sp.*, with a length between 0.2 and 2.5 Kb and widely distributed in the *Leishmania* genome (Ubeda et al., 2014). According to the current model for CNV generation, these sequences rearrange during DNA replication by homologous recombination (Laffitte et al., 2016). It has been suggested that direct repeats generate circular extrachromosomal DNAs by recombining repeats at the same chromatid; or tandem duplication of the locus by unequal crossing at sister chromatids (Figure 4A) (Laffitte et al., 2016). Classically, these amplification events have been studied by exploiting the knowledge that dihydrofolate reductase-thymidylate synthase gene (*DHFR-TS*)

becomes amplified on extrachromosomal circles in *Leishmania* cells serially passaged in the antifolate methotrexate. It has been suggested that HR between direct sequences produces amplicons containing DHFR leading to MTX resistance (Berveley et al. 1984).

Inverted repeats (IR), in contrast to DRs, have been implicated in the formation of linear amplicons (Figure 4B) (Laffitte et al., 2014; Ubeda et al., 2014). It was shown previously that the DNA repair nuclease MRE11 is required for the ability of *L. infantum* to generate linear amplicons (Laffitte et al., 2014). MRE11, when complexed with interacts with RAD50 and NBS1, form the MRN complex, which is an important complex for DNA repair via HR or NHEJ (Stracker, Petrini, 2011). Current data suggests that during DNA replication, the presence of IRs can cause hairpins to form, leading to single strand breaks or double strand breaks in the DNA, which are key targets for processing by the MRN complex (Laffitte et al., 2014). After resection of the damaged strand by MRE11 nuclease, the IR sequences anneal to generate a platform for DNA synthesis orientated towards the telomeric regions (Figure 4B) (Ubeda et al., 2014). In fact, sequence-based studies have shown an increased read depth in different *Leishmania* isolates around the chromosome ends, a region known by the presence of repeat elements, after selection for fast *in vitro* growth, revealing a possible nonclassical telomere maintenance mechanism which is associated with the intrinsic strain-specific genetic factors under environmental selection which could allow fitness gain (Bussotti et al. 2018).

Figure 4 – Current models of circular and linear amplification in *Leishmania*.
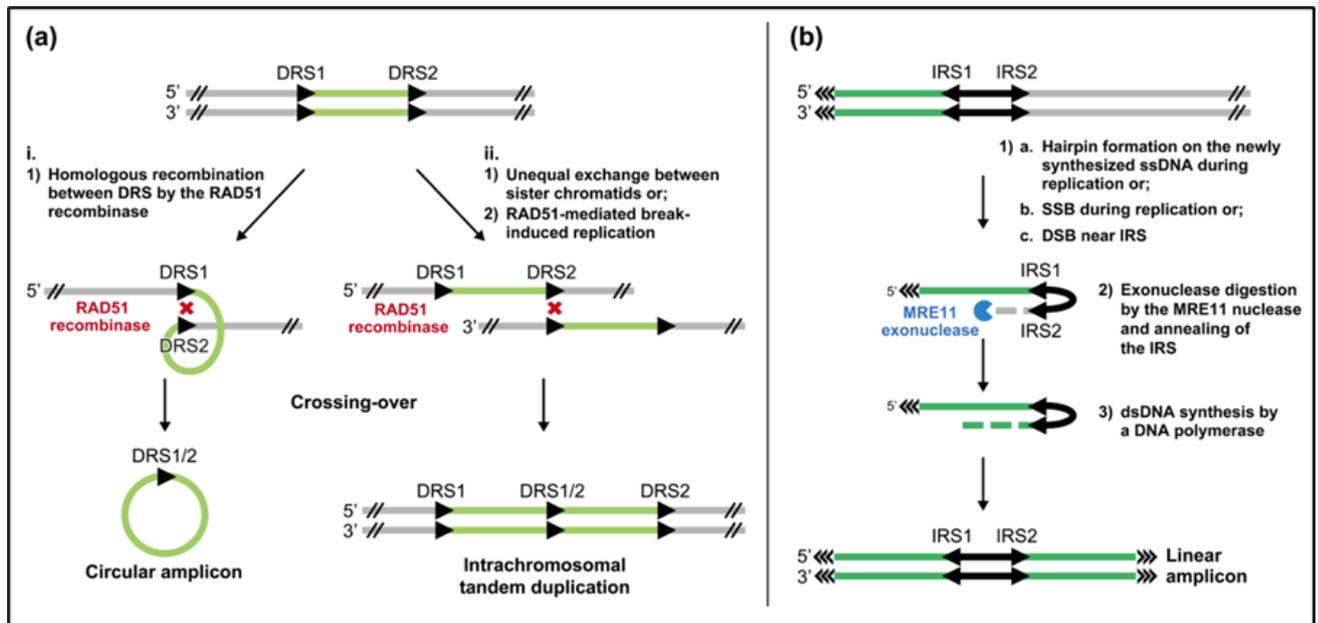
Figure 4 – Models of genomic amplification in *Leishmania*. A) Model of the formation of extrachromosomal circular amplicon (i) or intrachromosomal tandem duplication, both depend on homologous recombination mediated by RAD51. **B)** Model of the formation of extrachromosomal linear amplification by the pairing of repeated sequences during the replication process. (Reproduced from Laffitte et al., 2016)

Leishmania cells are largely diploid but aneuploidy is common. Moreover, a variable chromosomal ploidy among individual cells generates intra-strain heterogeneity which is known as chromosome mosaicism (Sterkers et al., 2012). Aneuploidy and chromosome mosaicism could arise from a number of events including unusual replication dynamics of the parasite, hybridisation via cell-cell fusions and/or defective chromosome segregation (Sterkers et al., 2011). Altogether, the available information suggests that the generation of the *Leishmania* genome plasticity is multifactorial. Hence, a better understanding of genome maintenance strategies could uncover a vital characteristic of the parasite biology.

1.4 Genomic Stability in Eukaryotes.

DNA is challenged daily and throughout the cell cycle by numerous sources of stress which may compromise DNA integrity. For instance, chemicals that cross-link or chemically modify DNA bases. All of which can physically damage DNA generating single and double strand breaks (SSBs; DSBs). Examples include ultraviolet light (UV)

or ionizing radiation. In addition to exogenous sources, endogenous processes like DNA replication can leave the DNA molecule more susceptible to events that cause injuries, such as, DNA-base mismatches and accumulation of single strand DNA (ssDNA); a state of perturbed DNA replication is known as replication stress (Ciccia, Elledge, 2010; Jackson, Bartek, 2009). The cellular response to replication stress is discussed in more detail in section x.

To maintain genome integrity and guarantee the transmission of the correct DNA information to the next generation, cells have evolved robust and organised mechanisms to deal with genomic injuries called the DNA damage response (DDR). Activation of the DDR is responsible for detecting damage and signalling the cell cycle to pause. DDR also is involved in recruiting repair proteins or, in case of extensive damage, to promote the elimination of cells by apoptosis (Blackford, Jackson, 2017).

A key early step of the DDR is the recognition of damage to the DNA by a group of sensor proteins which are responsible to recruit and active signalling proteins (Ciccia, Elledge, 2010; Jackson, Bartek, 2009). The majority of signalling proteins are kinases, which are responsible for the phosphorylation of downstream effective proteins involved in different cell processes that are important for guaranteeing genomic stability. In humans, six PIKK (phosphoinositide 3-kinase (PI3K)-related kinases) family members have been identify: ataxia-telangiectasia mutated (ATM), DNA dependent protein kinase catalytic subunit (DNA-PKcs), Ataxia-Telangiectasia and Rad3-related (ATR), mammalian target of rapamycin (mTOR), suppressor with morphological effect on genitalia family member (SMG1) and transformation/transcription-associated protein (TRAAP); in other eukaryotic cells there are at least one of the six PIKKs family members has been identified (Cimprich, Cortez et al., 2008; Mcmahon et al., 1998; Smith, Jackson et al., 1999; Yamashita et al., 2001; Yang et al., 2013). The first three PIKK family members are signalling proteins of DDR, where ATM and DNA-PKcs both respond primarily to DSBs (Figure 5); and ATR, which is required for an appropriate response to tackle replication stress (Figure 5) (Abraham, 1998; Blackford, Jackson, 2017).

Figure 5 – PIKK family kinases: Recruitment and Activation in Response to DNA Damage.
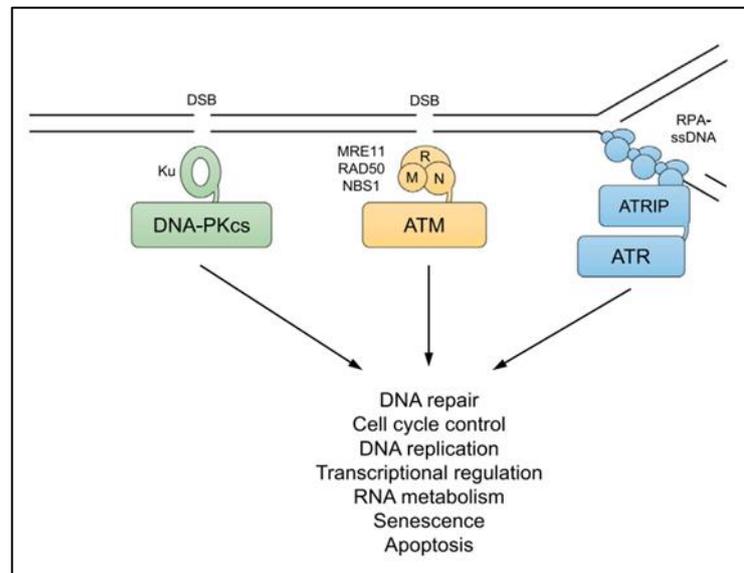
Figure 5 - PIKK family kinases: Recruitment and Activation in Response to DNA Damage. In the presence of DSBs ATM is recruited and activated by MRE-11-RAD50-NBS1 (MRN complex) and DNA-PK is recruited and activated by Ku DSB-bound. ATR is recruited to RPA-coated ssDNA with its partner ATRIP (From: Blackford, Jackson, 2017).

As part of the same family, all three kinases share similar structural characteristics and biochemical characteristics. For example, their kinase domains, present at the C-terminal, are flanked by the FAT and FATC domains (the last one present only in ATR and ATM kinases), which are regulators of the kinase activity (Bosotti, Isacchi, Sonnhammer, 2000). In human cells, the ATM kinase FAT domain interacts and inhibit the kinase domains in absence of DNA damage (Pellegrini et al., 2006). However, upon DNA damage, the phosphorylation of Ser1981 in the FAT domain disrupts the inhibitory interaction (Pellegrini et al., 2006). It is noteworthy that the equivalent phosphorylation in mouse cells is not necessary for ATM activation (Pellegrini et al., 2006). In ATR kinase, the FAT phosphorylation at Thr1989 is required for kinase activation by TopBp1, one of the ATR activators (see below) (Liu et al., 2011); in addition, the FATC domain plays an indispensable role due to its basal kinase activity (Mordes, Cortez, 2008).

1.5 ATM and DNA-PKcs activation during DSBs.

In eukaryote models, both ATM and DNA-PKcs are recruited to the chromatin in response to DSBs. ATM binds to MRE11-RAD50-NBS1 (MRN) complex which acts to unwind the DNA strands and performs end-resection (Carney et al.,1998). Once actived by MRN complex, ATM phosphorylates important downstream substrates involved in halting cell cycle progression, such as p53 (a tumour suppressor protein) and the checkpoint kinase protein 2 (CHK2) (Melchionna et al., 2000; Shiloh, Ziv, 2013). The range of ATM activity spans across two key DNA repair pathways, NHEJ and HR, involved in preserving the integrity of the DNA strands (Riballo et al., 2004).

DNA-PK can also become activated in the presence of DBSs. DNA-PK is recruited to damaged sites by the Ku heterodimer (formed by the Ku70 and Ku80 proteins), which leads to DNA-PK autophosphorylation (Walker, Corpina, Goldberg, 2001). To promote the DNA repair in this scenario, the DNA ends are resected by nucleases and the DNA gap filled by DNA polymerases mainly Pol mu (Pol μ) and Pol epsilon (Pol ε), which act in a template independent manner (Pannuzio, Watanabe, Lieber, 2018). Lastly, DNA ligase IV, in conjunction with the x-ray repair cross-complementing protein 4 (XRCC4) and the XRCC4-like factor (XLF), seals the break (Grawunder et al., 1997; Li et al., 1995).

Several reports showed that both ATM and DNA-PK have overlapping functions during DSB repair as both kinases activate in response to DSBs (Britton, Coates, Jackson, 2015; Jiang et al., 2015). During the NHEJ repair, the MRN-dependent activity of ATM may trigger the phosphorylation of DNA-PK to promote the recruitment and activation of the endonuclease Artemis (Moshous et al., 2011). In fact, studies using high-throughput single-molecule microscopy demonstrate that the MRN complex gains access to occluded DNA ends by removing Ku or other DNA adducts via an Mre11-dependent nucleolytic reaction (Myler et al., 2017; Riballo et al. 2004). The adaptor protein PAXX (Paralog of XRCC4 and XLF), has been showed to be essential to stabilize the ATM activity during NHEJ (Balmus et al. 2016; Zha et al., 2011). In addition, ATM substrates on chromatin involved in DNA-end protection and bridging, such as H2AX and 53BP1, also share functional redundancy with XLF in NHEJ (Kumar, Oksenych, 2014b). ATM also promotes HR, a high-fidelity repair pathway. HR requires DSB-end processing to produce tracts of RAD51-nucleofilament, invade the

sister chromatid that serves as a template (Kowalczykowski, 2015). The DNA-end resection is stimulated by MRN complex and ATM activity, that leads to Ku removal from the DSB-ends at collapsed replication forks to allow strand invasion (Chanut et al., 2016).

1.6 ATR kinase pathway.

Though activated by a myriad of genotoxic stressors, the ATR kinase also sits at the apex of the replication stress response (RSR). The RSR, as mentioned, tackles the effects of stalled or slowed DNA polymerase progression (Blackford, Jackson, 2017). The result of which leads to the formation of a common DNA structure at the replication fork that ATR recognizes (Saldivar, Cortez, Cimprich, 2017); ssDNA. ssDNA forms when deregulation between the helicase activity and the replication fork progression (Zou, Elledge, 2003) arises. Canonically, the ATR pathway triggers by the recognition of ssDNA bound by a heterotrimeric complex called RPA. In human cells RPA is formed by proteins RPA70, RPA32 and RPA14. The RPA complex coats the ssDNA, forming an ssDNA-RPA nucleofilament, and protecting the ssDNA from degradation by nucleases (Zou, Elledge, 2003). This RPA-ssDNA association serve as platform for the recruitment of ATR-interacting protein (ATRIP) which it is necessary for the recruitment of ATR to stressed sites (Cortez et al., 2001).

Following recruitment, two factors were found to participate in ATR activation: TOPBP1 (Topoisomerase II-binding protein 1), which interacts with the FAT domain in both ATR and ATRIP through its ATR-activation domain (Mordes et al., 2008), is recruited to sites containing a 5'-ended ssDNA-dsDNA junction (Kumagai et al., 2006). The ssDNA-RPA nucleofilament also acts as a loading platform for the 9-1-1 checkpoint complex (Rad9-Rad1-Hus1). The 9-1-1 clamp acts in the stabilization of TOPBP1 and, consequently stimulates ATR-ATRIP activation (Delacroix et al., 2007).

The formation of ssDNA-RPA nucleofilaments has a pivotal participation in DNA replication and repair functioning not only as a major sensor to elicit the DDR following the genotoxic stress, but also as a tool for protecting DNA replication forks against breakage (Toledo et al., 2013). Under replication stress, RPA activates the ATR pathway by two possible routes in vertebrates: 1) by directly recruiting and activating ATR-ATRIP, as described above, or 2) by the recruitment and interaction to second

ATR activator, the Ewing tumor-associated antigen 1 (ETAA1) that may act on stressed replication forks (Bass et al., 2016). In the canonical ATR pathway activation, the N-terminal of RPA70 subunit is responsible for the recruitment of ATR-ATRIP to the damage site, it also acts a facilitator of the Rad17/RFC to mediate loading of the 9-1-1 clamp at the ssDNA-dsDNA junction (Day, Oliver and Pearl, 2022, ***preprint***).

Once activated at stressed replication forks, ATR orchestrates a multifaceted response that protects the integrity of the fork and ultimately the genome. Instrumental to this response is the downstream effector kinase, CHK1, which is activated by phosphorylation by ATR (de Oliveira et al., 2015). Together, ATR and CHK1 function to arrest the cell cycle, suppress replication origin firing, stabilize replication forks, and promote fork repair and replication restart. By coupling cell cycle arrest, fork stabilization and restart, ATR and CHK1 act to ensure that cells do not enter mitosis when replication is perturbed (Saldivar, Cortez and Cimprich, 2017).

1.6.1 Cell cycle arrest

One of the main functions of the activation of ATR is cell cycle arrest. When DNA damage arises in S phase of the cell cycle, activation of ATR leads to ATR phosphorylating the checkpoint protein 1 (CHK1). ATR activation by CHK1 is mediate by the adaptor protein Claspin, which recruits CHK1 to the damage site (Kumagai and Dunphy, 2000). Upon activation, CHK1 phosphorylates and inactivates the cell division cycle 25 (CDC25) phosphatase, which removes the inhibitory phosphorylation of CDK2 and CDK1, stopping the progression of the cell cycle. The regulation of CDC25 by ATR-CHK1 is evidenced in studies that inhibit the activity of ATR or CHK1 (Karlsson-Rosenthal and Millar, 2006).

1.6.2 Regulation of Origin Firing

DNA replication origin firing is orchestrated spatially and temporally through a variety of processes involving cell cycle progression control, chromatin structure and nuclear dynamics. Therefore, the replication firing, for a particular chromosome or region, will occur at different moments of the S phase depending upon the cell type and physiological environment (Fragkos et al., 2015). ATR is a negative regulator of the origin firing and prevents inappropriate activation of origin firing even in unperturbed cells (Shechter, Costanzo, Gautier, 2004). This function of ATR is

important to guarantee sufficient supply of DNA precursors and replication factors for optimal fork progression.

Under replication stress, ATR has a role limiting the DNA replication. In this case ATR blocks the recruitment of CDC45 to the MCM2-7 (minichromosome maintenance 2-7) complex, which is the helicase complex responsible for unwinding DNA at the replication fork (Kamani, Dutta, 2011). A similar mechanism occurs during the intra S checkpoint where ATR phosphorylates the histone methyltransferase myeloid/lymphoid or mixed-lineage leukemia (MLL) that promote the stabilization of histone modifications on the chromatin blocking the loading of CDC45 at the replication forks (Liu et al., 2010; Zegerman, Diffley, 2010). The regulation of origin firing is a key aspect of the ATR-dependent checkpoint, which supresses origin firing globally, but allow dormant origin to be fired locally. Such mechanism guarantees a supported replication DNA at specific regions and prevent problematic replication widespread across the genome (Ge, Blow 2010).

1.6.3   ATR maintains replication fork stability and restart.

The ability of stalled polymerases to restart DNA synthesis after removal or bypass of a block in replication elongation is a *bona fide* indication of replication fork stabilization. Unstable forks may lead to replication fork collapse, a process that can generate DSBs (Buisson et al. 2017; Nelseen et al. 2015). ATR also participates in the stabilization of stressed replication forks. An interesting example is the protection effect of ATR during fork reversal (Sogo, Lopes, Foiani, 2002). Upon replication blockage, replication fork reversal is a protective mechanism where a conversion of typical three-way junction at the replication fork turns into a four-way junction. This process is mediated by SMARCAL1 (Couch et al., 2013; Liao et al., 2018). ATR is essential to prevent the nuclease-mediated cleavage of replication fork at late stages of replication reversal (Ragland et al., 2013). *In vitro*, ATR phosphorylates SMARCAL1 supressing fork reversal, protecting the DNA from the excessive DNA cleavage that processes the four-way junctions of the reversed replication fork (Couch et al., 2013). A direct activity of ATR and its downstream targets also work protecting the stalled forks to collapse inhibiting endonucleases as MUS81 (Franchitto et al., 2008).

Another way that ATR maintains replication fork stability is by modulating replisome function. Under replication stress, ATR phosphorylates many replisome

components including DNA polymerases, the MCM helicase, the RPA complex and Claspin (Kumagai, Dunphy, 2000). In human cells, ATR recruits FANCD2 to the replication fork, more specifically to be associate with MCM complex. FANCD2 slows the progression of the DNA polymerase, and presumably the helicase, minimizing the formation of ssDNA (Lossaint et al. 2013). Furthermore, ATR inhibition combined with replication stress cause the depletion of the pool of RPA, leading cells to replication catastrophe. This phenomenon is particularly detrimental for replication forks emanating from aberrantly fired origins (Toledo et al. 2013). Besides its role in replication fork stabilization, ATR seems to be actively involved in fork restart. It has been shown that ATR phosphorylates two of trans-lesion polymerases, REV1 and Pol η, suggesting that the ATR might control lesion bypass (Göhler et al., 2011). ATR acts to ensure fork restart by regulating template switching, fork reversal and homologous recombination (Göhler et al., 2011). For that, ATR phosphorylates RAD51-dependent factors that act in the recruitment of RAD51 to stalled forks and DSBs (Ahlskog et al., 2016).

1.6.4 Ensuring dNTP availability

ATR plays a key role in regulating dNTP levels. Studies in budding yeast showed that the loss of Mec1 (ATR in yeast) could be rescued by increasing the activity of ribonucleotide reductase (RnR), which is the rate limiting enzyme in dNTP production (Zhao, Muller, Rothstein, 1998). Similar results were found in humans where ATR regulates the factor E2F1-dependent expression of ribonucleoside-diphosphate reductase subunit M2 (RRM2), which is particularly important during early stages of S phase (Buisson et al. 2015).

1.7 Our current understanding of DDR signalling kinases in Trypanosomatids.

Trypanosomatids diverged earlier from more classical model eukaryotes. Thus, several unorthodox or divergent aspects of the DDR and DNA repair have emerged. For instance, current data supports a lack of the classical NHEJ pathway involved in DSBs repair in trypanosomatids seems to prioritise the use of regions of micro-homology, utilising microhomology mediated end joining (MMEJ) or single-strand annealing (SSA) at a higher frequency when compared to other eukaryotes (Laffitte et al., 2016). A summary of the DNA repair mechanisms is described on the supplementary 9.1.

The predicted trypanosomatid ATR kinase present identities of ~25% when compared to the human homologs (Figure 6). The enzyme also conserved around ~70% of an α-solenoid-like domain, characteristic from PIKK family members. For instance, the PRD domain, which in humans is involved in the kinase activation by TopBP1 interaction, appears to be lacking from the ATR kinases of all kinetoplastid sequences (Mordes et al., 2008). Also, until now there are no studies describing functional roles of the putative TopBP1 identified in Genois (2014); as well its possible interactions with the kinetoplastid ATR. This might indicate that a different effector and/or interaction site is involved in the activation of ATR in these protozoa.

The knowledge about ATR functions is more advanced in *T. brucei*. Using RNAi, depletion of ATR in mammalian infective *T. brucei* leads to an accumulation of replicative cells followed by a growth arrest, increase of genotoxic markers such as yH2A levels, RAD51 and RPA foci (Black et al., 2020; Jones et al., 2014); all characteristics of ATR canonical implications in DDR. In another study using the insect stage of *T. brucei*, showed that the ATR kinase is involves signalling the DNA damage and DNA repair in response to ionizing radiation (Marin et al., 2020). Interesting, ATR is also important for the switch of Surface Variant Protein (VSG) that allow the bloodstream *T. brucei* to evade the host immune system (Black et al., 2020), directing correlating the kinase activity with a particular feature of this parasite biology.

Figure 6 - Predicted domain locations in Trypanosomatid PIKKs compared with their human homologs.
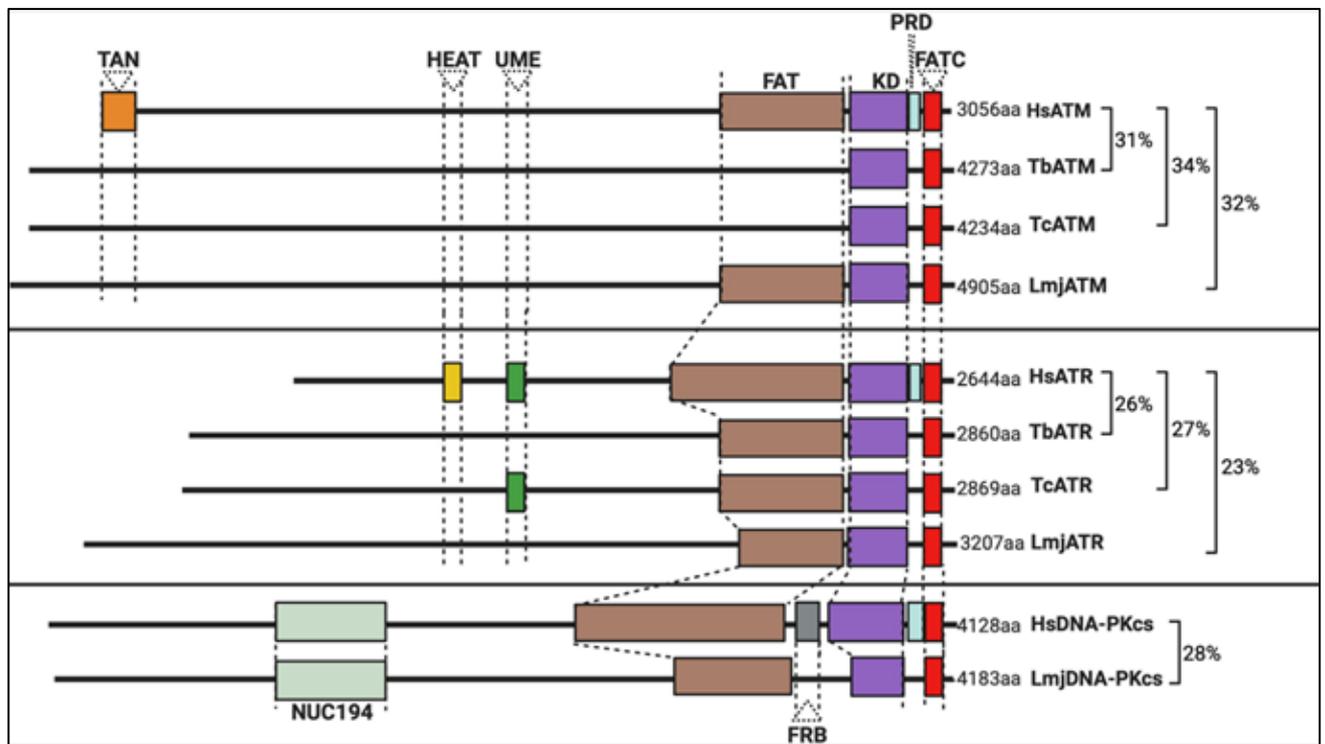
Figure 6 - Predicted domain locations in Trypanosomatid PIKKs compared with their human homologs. Schematic illustration of the predicted domain locations in Trypanosomatid PIKKs compared with their human homologs. Putative domains were identified using Pfam (http://pfam.xfam.org), Prosite (https://prosite.expasy.org), and Interpro (https://www.ebi.ac.uk/interpro/). Sequence similarity was determined using BLAST (Altschul et al., 1997), and all sequences from trypanosomatids are compared to the corresponding human kinase sequence. Gene IDs, the percentage identity, and the E value for each sequence are as follows: HsATM (AAB65827.1), TbATM (TbATM427_020008900; 31.47%, 2e-99), TcATM (TcCLB.509395.20; 33.95%, 7e-108), and LmjATM (LmjF.02.0120; 31.73%, 9e-94). HsATR (NP_001175.2), TbATR (Tb427_110165100; 26.14%, 7e-119), TcATR (TcBrA4_0103840; 27.17%, 2e-199), and LmjATR (LmjF.32.1460; 23.45%, 2e-97). HsDNA-PKcs (NP_008835.5), LmjDNA-PKcs (LmjF.36.2940; 27.58%, 4e-30). (From: Silva et al., 2021).

The first studies that try to investigate ATR kinase in Leishmania were using a selective inhibitor of the ATR kinase in humans VE-821 (Charrier et al., 2011; Reaper et al., 2011). The promastigotes cells VE-821 treated cells showed a small decrease of proliferation which is exacerbated when the cells are also treat with H_2O_2, which could suggest that ATR might play a role during oxidative stress (da Silva et al., 2018). Recently using the same approach, inhibition of ATR in a LmjEXO1 overexpression cell line results in the formation of extended single-stranded DNA (da Silva et al., 2022). However, there is no validate data confirming that the ATR inhibitors available to human ATR are specific to the *Leishmania* version. Our laboratory has been dedicating

to characterize the ATR pathway factors, specifically the 9-1-1 complex, which its subunits showed canonical roles in replication response pathway (Damasceno et al. 2016; Damasceno et al. 2013); and divergent functions as individual subunits, as HUS1 which exists in its monomeric form, and its depletion revealed its possible involvement in genome plasticity (Damasceno et al. 2018). As demonstrated at the Figure 7, many questions remain without answers about DDR in Trypanosomatids; so far there is no evidence of the presence of CHK1 and p53 which are essential upstream targets of ATR activity. In addition, there is no reported data specifically targeting the *Leishmania* ATR version.

Figure 7 - The PIKK-driven DDR pathways in Trypanosomatid parasites and the canonical Eukaryotic pathways.
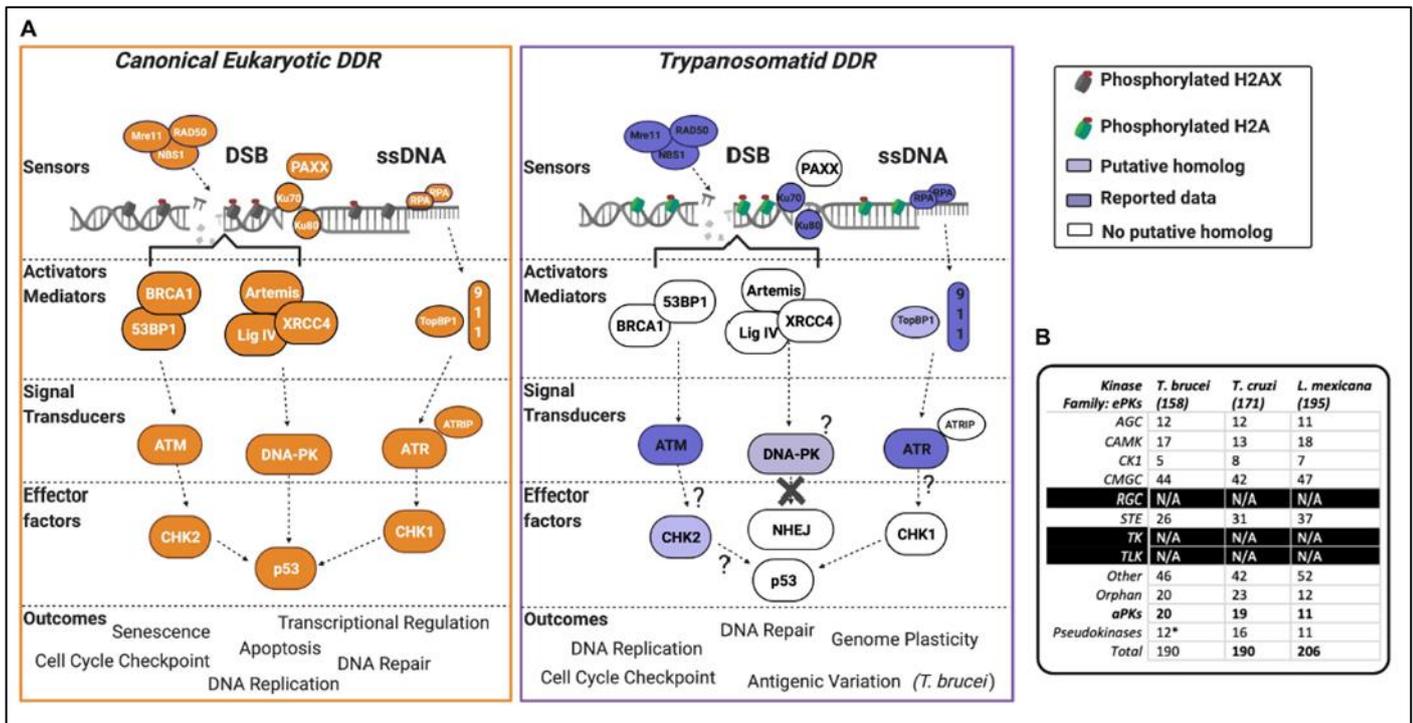
Figure 7 - The PIKK-driven DDR pathways in Trypanosomatid parasites and the canonical Eukaryotic pathways. A) A schematic illustration of a simplified eukaryotic DDR pathway (left) compared to known or predicted components of the trypanosomatid DDR pathway (right). Dark shaded factors indicate that functional characterization has been performed in one or more organisms. Light shading indicates limited data availability. White indicates no data are available or the factor is not present in the genome, as further illustrated by question marks. For more intricate details on eukaryotic DDR factors and pathways, we encourage the reader to refer to recent reviews (Alexander and Orr-Weaver, 2016; Blackford and Jackson, 2017; Wright et al., 2018; Sun et al., 2020; Zhao et al., 2020; Ghosh and Raghavan, 2021). DSB, double-stranded break; DDR, DNA damage response; ssDNA, single-stranded DNA. (B) Summary table of PKs and their associated families in *T. brucei*, *T. cruzi*, and *L. mexicana*. Data collated from Parsons et al. (2005), Jones et al. (2014), and Baker et al. (2021). (*) = the pseudokinases in T. brucei are included among the counts for the other families and their respective numbers have not been adjusted to remove pseudokinase family members. N/A = no kinases have been identified as members of these kinase families. (From: Silva et al., 2021).

Considering the scarcity of studies detailing ATR kinase functions in genome maintenance in *Leishmania major*, we chose to investigate whether the canonical features of the ATR activity are conserved in this parasite. To this end, we generated ATR-deficient *L. major* cell lines and an endogenously tagged ATR cell line, to allow us to characterise the function of ATR in *L. major* promastigote cells. Using a combination of assays, we measured the accumulation of ssDNA, investigated RPA relocation, ATR subcellular localization and cell cycle progression dynamics. In addition, we set a Marker Frequency Analysis (MFA-seq) to investigate how the *Leishmania* replication program is affected in an ATR-deficient manner.

Methods

2. Methods

2.1 Parasite Culture

Cell lines were derived from *L. major* strain LT252 (MHOM/IR/1983/IR). Promastigotes were cultured at 26°C in M199 or HOMEN medium supplemented with 10% heat-inactivated fetal bovine serum and 1% penicillin/streptomycin. Transfections were performed with exponentially growing cells with Amaxa Nucleofactor II, using the pre-set program X-001 (Burkard, Jutzi, Roditti, 2011). Transfectants were selected by limiting dilution in 96-well plates in the presence of the appropriate antibiotic. Cas9/T7-expressing cells were selected with 20 µg.mL^{-1} hygromycin. Cells expressing the gene of interest (GOI) were selected in 20 µg.mL^{-1} hygromycin, 10 µg.mL^{-1} puromycin and 8 µg.mL^{-1} neomycin (G418).

2.2 DNA constructs and cell line generation.

To use CRISP/Cas9 genome editing in LT252, a parental cell line was established in which Cas9 and T7 polymerase were inserted into and expressed from the tubulin array (Figure sup. 9.2). WT cells were transfected with plasmid pTB007, previously digested with PacI, to generate the Cas9/T7-expressing cell line. Correct integration of Cas9/T7-encoding cassette was confirmed by PCR. This strategy is previously described in Beneke (2017).

The donor fragment and the sgRNAs for the transfection were generated by PCR as described previously (ref). Briefly, the primers were produced using the gene ID of ATR, LmjF.32.1460, on Leishgedit.net as described in Beneke (2017). For each fragment amplified (sgRNA and donor fragment) five PCR reactions were performed as follows using Phusion polymerase (98°C/1min, 98°C/30s, 60°C/30s, 72°C/45s or 1:45min, 72°C/5min, 35 cycles). All PCR reactions were precipitated and pooled as follows: DNA was extracted using Phenol/Chloroform (in equal volume), centrifuged at max. velocity for 5min and the upper phase collected. The DNA was then precipitated in ETOH 100% (cold): 250ul; 3M Sodium acetate: 10ul; Glycoblue: 3ul overnight at-20 °C: 16h. After, the DNA was centrifuged as before for 20min, the upper phase collected

and then washed in ETOH 70% (cold): 500ul by centrifugation as before for 20 min. The pellet was air dried then resuspended in 20ul of MilliQ water. For each transfection, 10 µl of the precipitated DNA was used. For recombination transfection or double tagging, $5x10^5$ *L. major* Cas9 cells were resuspend in 100 µl of 1× Tb-BSF buffer and pulse using the X-100 program in the Amaxa Nucleofactor IIb (Lonza) (Burkard, Jutzi, Roditi, 2011). After transfection, cells were resuspended in 10 mL of fresh media with 20% of FBS and recovered for at least 12 hours, then the appropriate selection drug was added, and the cells plated on a 96-well plate on the dilution of 1:6 – 1:72 – 1:864 for ~ 5 weeks.

2.3 Western blotting

Whole cell extracts ($3\text{-}4x10^6$ cell/mL^{-1}) were prepared by centrifuge (1400g/5min), wash with PBS and resuspend in NuPAGE LDS Sample Buffer (ThermoFisher) with 10% protein inhibitors (Thermo Fisher A32953) and 2% β-mercaptoethanol or with the sample buffer (300mM Tris-HCl ph 6,8, 50% Glycerol, 10% SDS, 5% β-mercaptoethanol). Samples were boiled prior to loading either in NuPAGE LDS Sample Buffer (ThermoFisher) or with the sample buffer (300 mM Tris-HCl ph 6.8, 50% Glycerol, 10% SDS, 5% β-mercaptoethanol). After boiling the samples, extracts were resolved on 4–12% gradient Bis-Tris Protein Gels (ThermoFisher) or homemade gels (30% Acrylamide mix, 1,5 M Tris pH 8.8, 10% SDS, 10% ammonium persulfate, TEMED); then transferred to Polyvinylidene difluoride (PVDF) membranes (GE Life Sciences). Before probing for specific proteins, membranes were blocked with 10% (w/v) non-fat dry milk in phosphate-buffered saline or tris-buffered salina supplemented with 0.05% Tween-20 (PBS-T or TBS-T). Primary antibody incubation was performed for 2 h at room temperature with PBS-T or TBS-T supplemented with 5% non-fat dry milk. Membranes were washed with PBS-T or TBST-T and then incubated with secondary antibodies in the same conditions as the primary antibodies. For HRP-conjugated secondary antibodies, ECL Prime Western Blotting Detection Reagent (GE Life Sciences) was used for band detection as visualized with Hyperfilm ECL (GE Life Sciences). For IR Dye-conjugates secondary antibodies, the Odyssey Imaging Systems (Li-COR Biosciences) was used for band detection and visualization. We used ImageJ software to quantify the bands and the protein levels were normalised to the levels of the loading control.

da Silva, Gabriel Lamak Almeida *Methods*

2.4 Antibodies

Generation of affinity-purified antibodies against γH2A (1:1000) from rabbit serum was previously described (Glover and Horn, 2012). Commercial primary antibodies used are as follows: mouse anti-myc (1:5000, Merck Millipore), anti-EF1α (1: 40 000, Merck Millipore) and anti-BrdU clone B44 (1: 500, BD Bioscience). Commercial secondary antibodies used are as follows: goat anti-Rabbit IgG HRP-conjugated (ThermoFisher), goat anti-Rabbit IgG HRP-conjugated (ThermoFisher), goat anti-Rabbit IgG Alexa Fluor 488-conjugated (ThermoFisher), goat anti-Rabbit IgG IRDye 800CW-conjugated (Li-COR Biosciences) and goat anti-Mouse IgG IRDye 680CW-conjugated (Li-COR Biosciences).

2.5 IdU incorporation (S phase detection and ssDNA detection).

Cells were incubated with 150 mM IdU for 30 min (S phase detection) or for 16 hours followed by HU treatment (ssDNA detection); and then fixed at -20°C with a mixture (7:3) of ethanol and 1x PBS for at least 16 hr. Next, cells were rinsed with washing buffer (1x PBS supplemented with 1% BSA) and the DNA denatured for 30 min with 2N HCL, followed by neutralisation with phosphate buffer (0.2 M Na2HPO4, 0.2 M KH2PO4, pH 7.4). Detection of incorporated IdU was performed with anti-BrdU antibody (diluted in washing buffer supplemented with 0.2% Tween-20) for 1 hour at room temperature. After washing, cells were incubated with anti-mouse secondary antibody conjugated with Alexa Fluor 488 (diluted in washing buffer supplemented with 0.2% Tween-20) for 1 hour at room temperature and then washed. Finally, cells were stained with 1xPBS supplemented with 10 mg.mL^{-1} Propidium Iodide (PI) and 10 mg.mL^{-1} RNAse A and filtered through a 35 mm nylon mesh. FACSCelesta (BD Biosciences) was used for data acquisition and FlowJo software for data analysis. Negative control (omission of anti-BrdU antibody during IdU detection step) was included in each experiment and used to draw gates to discriminate positive and negative events.

2.6 EdU incorporation

Cells were incubated for 30 min with 10 mM of EdU (Click-iT; Thermo Scientific) during exponential growth (no treat) and after chronic and acute HU treatment, then fixed with 3.7%paraformaldehyde for 15 min, adhered into poly-L-lysine coated slides, then permeabilized with 0.5% TritonX100 for 20 min. After washing with PBS supplemented with 3% BSA, cells were subjected to a Click-iT reaction, as per the manufacturer's instructions. Subcellular localisation of the endogenously tagged ATR kinase was probed for using anti-myc antibody, followed by secondary anti-mouse 488 antibody. DNA was stained with Hoechst 33342. Images were acquired with a DV microscope (Leica) and processed with ImageJ software (reference). Quantification of the fluorescence was performed using the ROI Manager tool in ImageJ software. Bright and contrast were adjusted for improved image representation.

2.7 Fluorescent activated cell sorting (FACS) and genomic DNA extraction.

FACS and DNA extraction were performed as previously described (Damasceno et al., 2020). Briefly, exponentially growing cells were collected by centrifugation and washed in 1 × PBS supplemented with 5 mM EDTA. Cells were fixed at -20°C in a mixture (7:3) of methanol and 1 × PBS. Prior to sorting, fixed cells were collected by centrifugation and washed once in 1 × PBS supplemented with 5 mM EDTA, re-suspended in 1 × PBS containing 5 mM EDTA, 10 µg/ml PI and 10 µg/ml RNase A, and passed through a 35 µm nylon mesh and read in a FACSAria I cell sorter (BD Biosciences). The genomic DNA was extracted using Blood and Tissue DNA extraction kit (Qiagen), by omitting the lysis step. Genomic DNA from non-sorted (exponentially growing or stationary) cells was also extracted using Blood and Tissue DNA extraction kit (Qiagen), following the manufacture instructions.

2.8 Marker Frequency Analysis (MFA-seq)

The library preparation and the sequence were done by Eurofins genomics (Highlights of NGSelect DNA). After alignment and filtering, the MFA-seq analysis was performed using a previously in house established pipeline (Marques et al., 2015), with

modifications. The workflow detailing all the analysis is represented in Figure 8. The sequence files were uploaded at the Galaxy server and the sequence adaptors trimmed using the tool Trimmomatic and the reads mapped using the reference Genome of *L. major* Friendlin available on Tritrypdb (Table 7.1 and 7.2). Reads were binned in 1.0 kb windows along chromosomes and the number of reads in each bin was used to calculate the ratio between each cell cycle stage (or exponentially growing cells) versus stationary cells, scaled for the total library size. Bins with less than 100 reads in either sample were discarded, as were reads with a quality score below 20. To reduce noise due to collapsed regions, bins with a ratio above 2.8 were discarded (Table 7.3). Bins overlapping with other problematic mapping regions, as previously described (Lombraña et al., 2016), were also removed (Table 7.4). Due to strain differences, further aberrant mapping regions (Table 7.5) were excluded from samples. Due to the prevalence of mosaic aneuploidy in *Leishmania* (add reference here to back this statement up), ratio values were converted into Z scores using in-house R-scripts wrote by Dario Beraldi from University of Glasgow in Rstudio (Script 8.1); the data were plotted in 5 kb sliding windows for each individual chromosome, and MFA-seq profiles were represented in a graphical form using Gviz (Hahne, Ivanek, 2016). ChIP data for AcH3, base J and KKT1 were previously published (Garcia-Silva et al., 2017, Thomas et al., 2009; van Luenen et al., 2012;). The bedgraph files were uploaded into Galaxy server and convert into Bigwig format (Table 7.6), the matrix was generated using computeMatrix tool where the files were plotting according with the region of interest. For the metaplots, we use the deeptools plotHeatmap (Ramírez et al., 2014).

To calculate the Copy Number Variation (CNV), the coverage was measure from each BAM file, in which reads were binned in 2.0 Kb windows along chromosomes, excluding reads with quality below 30 and ignoring missing data (Table 7.7). Each file was downloaded and the CNV calculate using in-house R-scripts wrote by Dr. Dario Beraldi from University of Glasgow in Rstudio (Script 8.2). SNP's and insertions and deletions (INDELS) were examined separately. Briefly, every BAM file was filtered using the FreeBayes tool (Blankenberg et al., 2014) excluding the reads with less than 20 of map quality (Table 7.8). SNPs or INDELS were measure individually from data acquired with the VCF filter tool (Table 7.9), to examine only new occurrences of SNPs or INDELS across the conditions on Figure 25A. VCF-VCF intersect (Table 7.10) was

used to quantify these occurrences and then the numbers were compared with the size of each chromosome (Table 7.11).

Figure 8 – Workflow of MFA-seq analysis.
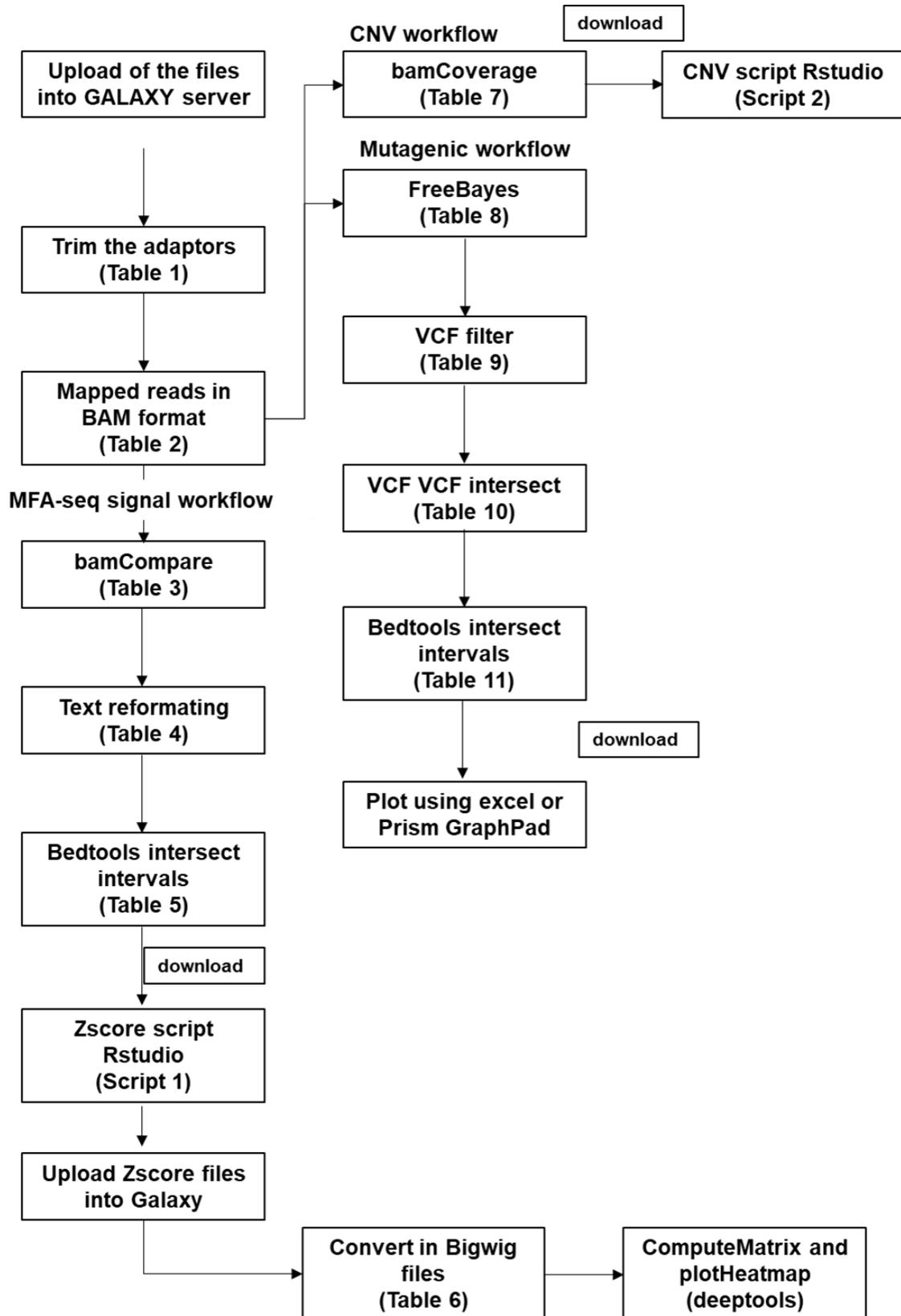
Figure 8 – Workflow of MFA signal analysis. Workflow showing how we processed and analyse the Whole Genome Sequence to obtain the MFA-seq signal, Mutagenic counts and CNV, using Galaxy server and Rstudio.

Results

3. Results

3.1 Chapter 1

In this chapter we describe how we generated several *L. major* ATR mutants and provide characterization of these mutants, which we use to test for conserved features of the ATR pathway in *Leishmania* and specific findings that might be associated with parasite cell biology.

3.1.1 Deficient ATR cells accumulate ssDNA and RPA relative to control cells under hydroxyurea treatment.

The primary substrates of ATR are elongated tracks of ssDNA coated by the RPA complex, forming foci which could be observed by immunofluorescence (IFA) (see below), to protecting the DNA strand and stimulate the ATR recruitment (Vassin et al., 2009). Without ATR, ssDNA and RPA foci accumulate in human cells under conditions of replication stress (Barr et al., 2003; Toledo et a., 2013). To ask if these replication stress associated features also accumulate in *L. major*, we generated a cell line deficient for the ATR kinase. As previously mentioned, the kinase domain of ATR and other key activation domains are present within the C-terminal portion of the protein. The function of these domains has been intensively explored in human and yeast cells (Bosotti, Isacchi, Sonnhammer, 2000; Mordes et al., 2008; Rao et al., 2018; Zhou et al., 2013). The sequence analysis of the *L. major* putative homolog for ATR (LmjF.32.1460) revealed a 23% homology with the human version of the kinase (da Silva et al., 2020). A classical approach to gene replacement of the entire ATR coding sequence was deemed to challenging due either to the predicted size of the ATR locus (~9.6 Kb) or to the fact that ATR may be an essential gene. Therefore, we instead aimed to truncate the locus by targeting the C-terminal region of the gene, which encodes relevant domains for ATR kinase function. This approach would increase the efficiency for mutant recovery and allow study of the cell expressing a kinase-defective ATR. To truncate *L. major* ATR, we replaced one allele of the kinase C-terminal using CRISPR-Cas9 system as originally developed to use in *Leishmania mexicana* by

Gluenz's lab (Beneke et al., 2017) and adapted for use in our laboratory. We used CRISPR/Cas9 combine with a fragment resulted from the digestion of a vector generated in our laboratory by cloning two regions of ATR locus with 500 bp length each for homologue recombination, flanking a select marker and six copies of hemagglutinin (HA) that it will replace the ATR C-terminal domains (Figure 9A and Supplementary Figure 9.3). To test for alterations to the behaviours of the RPA complex, we first tagged both alleles of RPA1, one of the complex components, using the same CRISPR-Cas9 system, and the generation of the donor fragments was made as described in Beneke (2016) (Figure 9A). The homozygotes $_{myc}$RPA cells were submitted to another round of transfection to replace at least one copy of ATR C-terminal using the *in-house* fragment. The modification to the endogenous locus in both cases was confirmed by PCR (Figure 9B) and the cell line was named ATR$^{\Delta C+/-}$/$_{myc}$RPA cells. We performed a western blotting to confirm the expression of the HA tag at the C-terminal of the truncated ATR allele, however did not seems to be expressed (Supplementary Figure 9.3).

To measure ssDNA accumulation, ATR$^{\Delta C+/-}$/ RPA^{myc} cells in exponential growth were incubated with the thymidine analogue IdU for 16 hours, using a cell line expressing only the tagged-RPA as a control. After the IdU incorporation, the cultures were washed and cultivated both in the absence and presence of 5mM of hydroxyurea (HU). As for other organisms, HU may cause replication stress in Leishmania by the inhibition of the Ribonucleotide Reductase (RNR), involved in the production of deoxyribonucleotides (dNTP) (Koç et al., 2004). Samples were collected at two- and five-hours post-HU treatment and prepared for IFA. ssDNA accumulation and RPA intensity were analysed using confocal microscopy. RPA subcellular localisation was detected by indirect immunofluorescence using anti-myc antibody and the ssDNA was detected with anti-BrdU antibody which recognizes the incorporated IdU in single strand conformation (Figure 9C). The results showed that in fact the loss of one ATR C-terminal copy leads to an increase of ~33% of the mean of nuclear ssDNA intensity in comparison with the control in untreated cells (13000 A.U. in $_{myc}$RPA cells and 19000 A.U. in ATR$^{\Delta C+/-}$ /$_{myc}$RPA cells (Figure 9E and 9I). Furthermore, we also observed a positive ssDNA signal in ATR deficient cells at the kinetoplast DNA (kDNA) in the shape of two dots in untreated cells (Figure 9C). It is known that the replication of (kDNA) starts at the extreme edges of the kinetoplast (Shapiro, Englund, 1995). The

quantification of the ssDNA intensity at the kinetoplast showed a significant difference between ATR deficient cells and the control in untreated cells, however, the intensity progressively decreases when the cells are in presence of HU (Figure 9D). In contrast, we observed an opposite trend when we measured the nuclear ssDNA intensity; the intensity of ssDNA in untreated cells significantly increases with the HU treatment in ATR deficient cells compared with the control (Figure 9E).

Next, we investigated the response of the RPA complex in *Leishmania* in the absence of one allele of the C-terminus of ATR. In humans the RPA complex is distributed throughout the nucleus and relocates into foci in the presence of DNA stressors (Barr et al., 2003). Firstly, RPA is located at the nuclear compartment, therefore we measured the RPA intensity at the nucleus since we observed an increase of ssDNA in ATR deficient cells. The analysis showed a significant increase in RPA intensity in ATR deficient cells in comparison with the control in untreated and treated cells (Figure 9F). Next, we counted the number of RPA foci. The number of foci in ATR deficient cells was significantly higher, with an average of 2 foci per nucleus, in comparison with the control cells, with an average of 1.5 foci per nucleus, after addition of HU (Figure 9G).

Lastly, we wanted to observe where in the cell cycle the accumulation of ssDNA was occurring. Here, we treated cells with two different concentrations of HU. These concentrations form the basis for most subsequent experiments performed in this thesis: Under 0.5 mM HU for 20 hours, replication is slowed in *L. major* ('Chronic stress') (Damasceno et al., 2013); and under 5 mM HU for 8 hours, the cells synchronize at the G1/S boundary of the cell cycle ('Acute stress') (Damasceno et al., 2013). Here, we analysed the cells by flow cytometry and the DNA stained with propidium iodide (PI) to examine ssDNA across the cell cycle. To define the positive events, every time point included a negative, un-stained control, in which the anti-BrdU antibody was omitted to allow for subtraction of background signal using FlowJo software. In agreement with the IFA data, the deficient ATR cells showed a higher accumulation of positive events (ssDNA) in all conditions compare with the control cells. With significant values during 8 hours of chronic treatment where the majority of cells are at S phase, and 4 hours of acute treatment where the cells are synchronizing at G1/S boundary; the positive events value in these time points in ATR$^{\Delta C+/-}$/$_{myc}$RPA cells are 15 and 16 times higher while the control cells showed an increase of 2 and

1.3 times higher, respectively, both relative to the untreated control value. (Figure 9H and 9I).

Figure 9 - Deficient ATR cells accumulate ssDNA and RPA signal in comparison to control cells under hydroxyurea treatment.
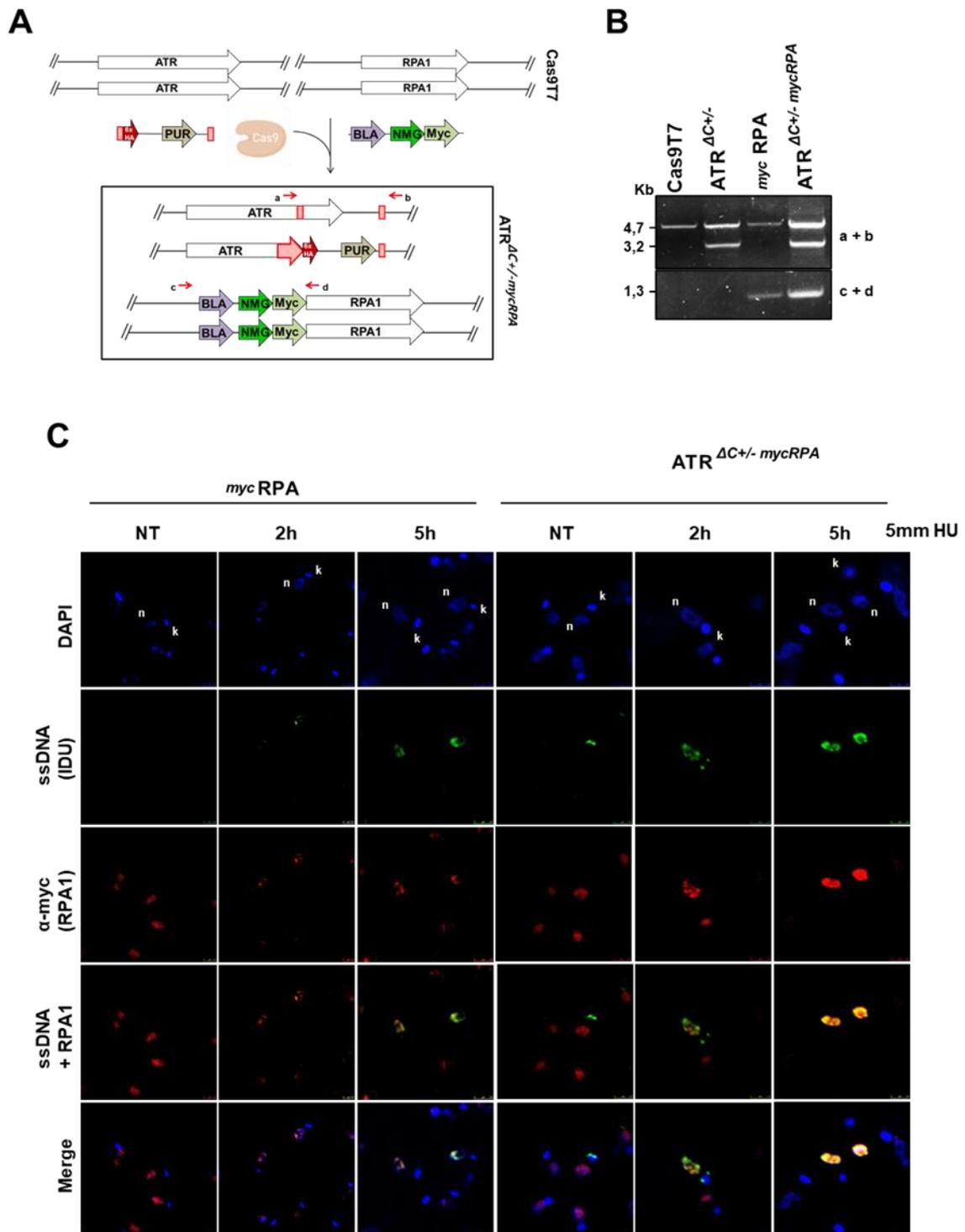
Figure 9 - Deficient ATR cells showed an increase of ssDNA and RPA signal in comparison to wildtype cells under hydroxyurea treatment. A) Schematic representation showing a cell line expressing Cas9 which was used to delete one copy of ATR C-terminal region; the fragment used to replace the targeted ATR C-terminal was generated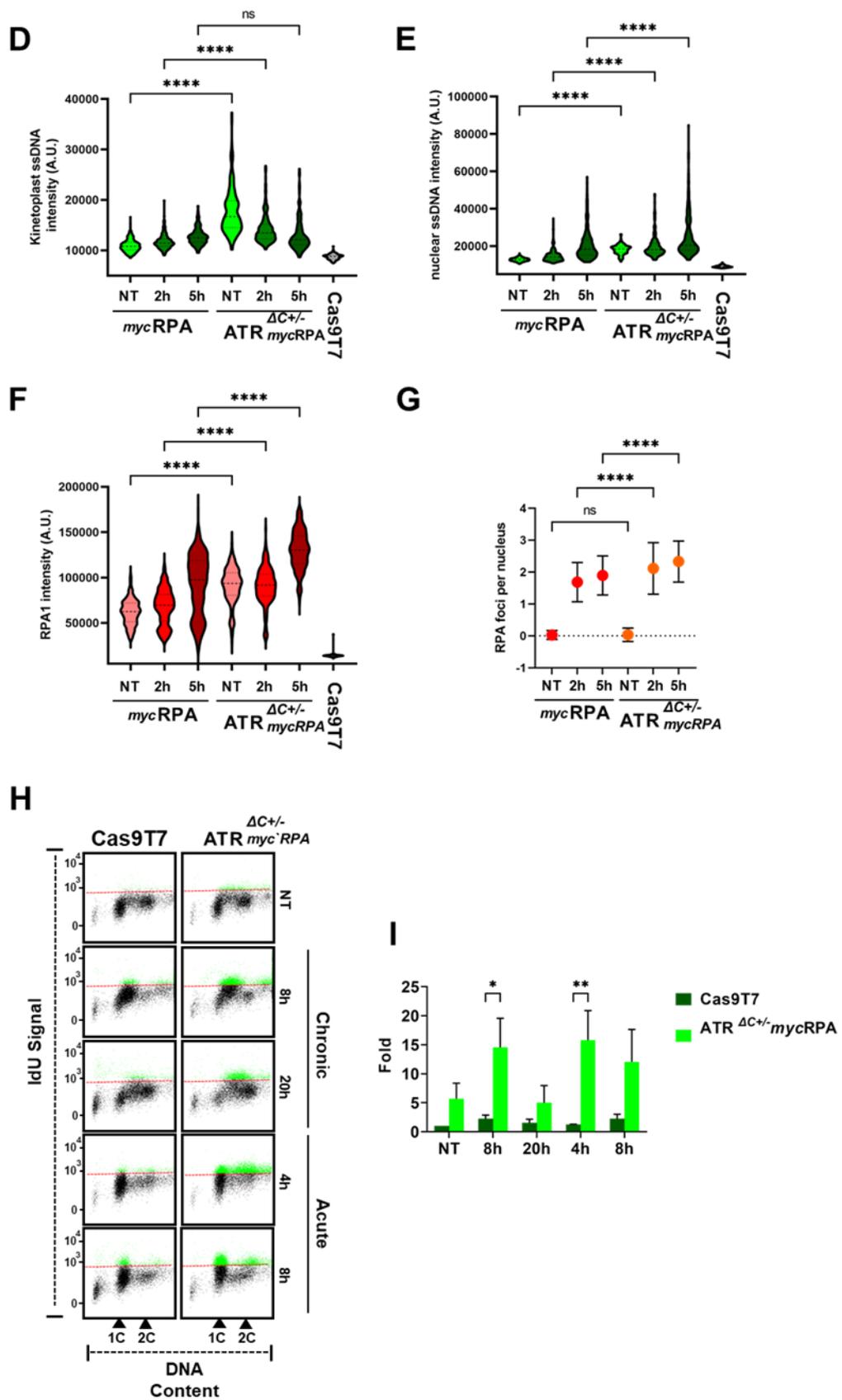 by cloning in a PGEM vector: 500pb upstream of the target break region and 500pb of the ATR locus UTR for the homologous recombination (red dashed regions), between were cloned six copies of the tag hemagglutinin (6xHA) and a gene that confers resistance to Puromycin (PUR). The fragment is obtained by digesting the plasmid with NdeI and PacI. The myc tag in N-terminal of RPA1gene was generated using the system described in Beneke (2017). **B)** PCR analysis indicating the deletion in one of the alleles of the ATR$^{ΔC+/-}$ cells using the set of primers (a + b), and the set of primers to identify the tag at RPA1 gene using the set of primers (c + d) shown in (A). **C)** Representative images acquired in SP6 Leica confocal microscope from one stack of Z images of RPA^{myc} cells and ATR$^{ΔC+/-}$ / RPA^{myc} cells untreated (NT) or treated with 5mM of HU for 2 and 5 hours. Cells were fixed with 3% formaldehyde and the cells stained with anti-BrdU (ssDNA) in green, anti-myc (RPA1) in red, and the genomic DNA stained with DAPI in blue. Bar scale = 2.5 um; the bright and contrast were adjusted. **D)** Quantification of the ssDNA signal around the kinetoplast DNA. An area was drawn around 150 kinetoplasts based on the DAPI stain and the ssDNA was measured using ImageJ software. Error bars ± SD; n = 2 experiments, * P<0,05; ** P<0,005; *** P<0,00. t-test unpaired. **E)** Quantification of the ssDNA signal around the nucleus. An area was drawn around 200 nuclei based on the DAPI stain and the ssDNA was measured using ImageJ software. Error bars ± SD; n = 2 experiments, * P<0,05; ** P<0,005; *** P<0,001. Unpaired t-test. **F)** Quantification of the RPA signal around the nucleus. An area was drawn around 200 nuclei based on the DAPI stain and the myc signal was measured using ImageJ software. Error bars ± SD; n = 2 experiments, * P<0,05; ** P<0,005; *** P<0,001. Unpaired t-test. **G)** Quantification of RPA foci. The foci were visually counted in at least 100 nuclei of each time point. Error bars ± SD; n = 2 experiments, * P<0,05; ** P<0,005; *** P<0,001. Unpaired t-test. **H)** Cells were incubated ~16 hours with IdU and then exposed to chronic HU treatment (8 and 20 hours) or acute HU treatment (4, 8 hours) and processed for IdU detection. The probing of anti-BrdU was under native conditions; a negative control (without antibody) of each sample was used it to retrieve the background, the DNA was stained using propidium iodate (IP) and the samples analysed by FACS. **I)** Quantification of IdU fluorescence detected under native conditions. Signal from individual cells was determined using ImageJ software; the increase of ssDNA was measured by dividing the number of positive events in each sample by the number of positive events in the untreated control cells (fold); Error bars ± SD; n = 2 experiments, * P<0,05; ** P<0,005. unpaired t-test.

3.1.2 *L. major* ATR kinase localises to the nucleus and kinetoplast.

The location of a protein provides valuable information on its function. ATR from other eukaryotes, including *T. brucei*, are localised to the nucleus (Barr et al., 2003; Black et al., 2020; Keegan et al., 1996). To investigate *L. major* ATR localization, we generated a cell line expressing a tagged version of ATR. To do this, we used the genome editing tool CRISPR-Cas9 based system as adapted to *L. major* described in Beneke (2017), which amplify the donor fragments and the single guide RNA by PCR. As we did to generated the $_{myc}$RPA cells, we aim to insert a fluorescent protein mNeonGreen and three copies of the epitope myc at the N-terminus of ATR (Figure 10A). Upon selection on puromycin, modification of the endogenous locus was confirmed by PCR (Figure 10B) and western blotting (showed later at Figure 12C) and the generated cell line named ^{myc}ATR.

da Silva, Gabriel Lamak Almeida   **Results**

To observe the subcellular localisation of ATR, we performed live cell microscopy using the ^{myc}ATR cells, and control untagged cells. First, exponential growth cells were incubated with Hoescht for 10 minutes to stain the DNA, washed, and resuspended in 500uL of fresh M199 medium. Three microliters of the culture were carefully placed on a microscope slide and spread using the coverslip. We used Cas9T7 cells as a negative control and the cells viability (flagellar movement) and mNG signal were analysed on a SP6 Leica confocal microscope. Fluorescent signal of mNG showed that ATR localises within the nuclear compartment and at the kinetoplast. It is noteworthy that ATR potentially localises within a region close which seems to be the flagella pocket (Figure 10C).

Figure 10 - *L. major* ATR kinase localises to nuclear and extranuclear compartments.
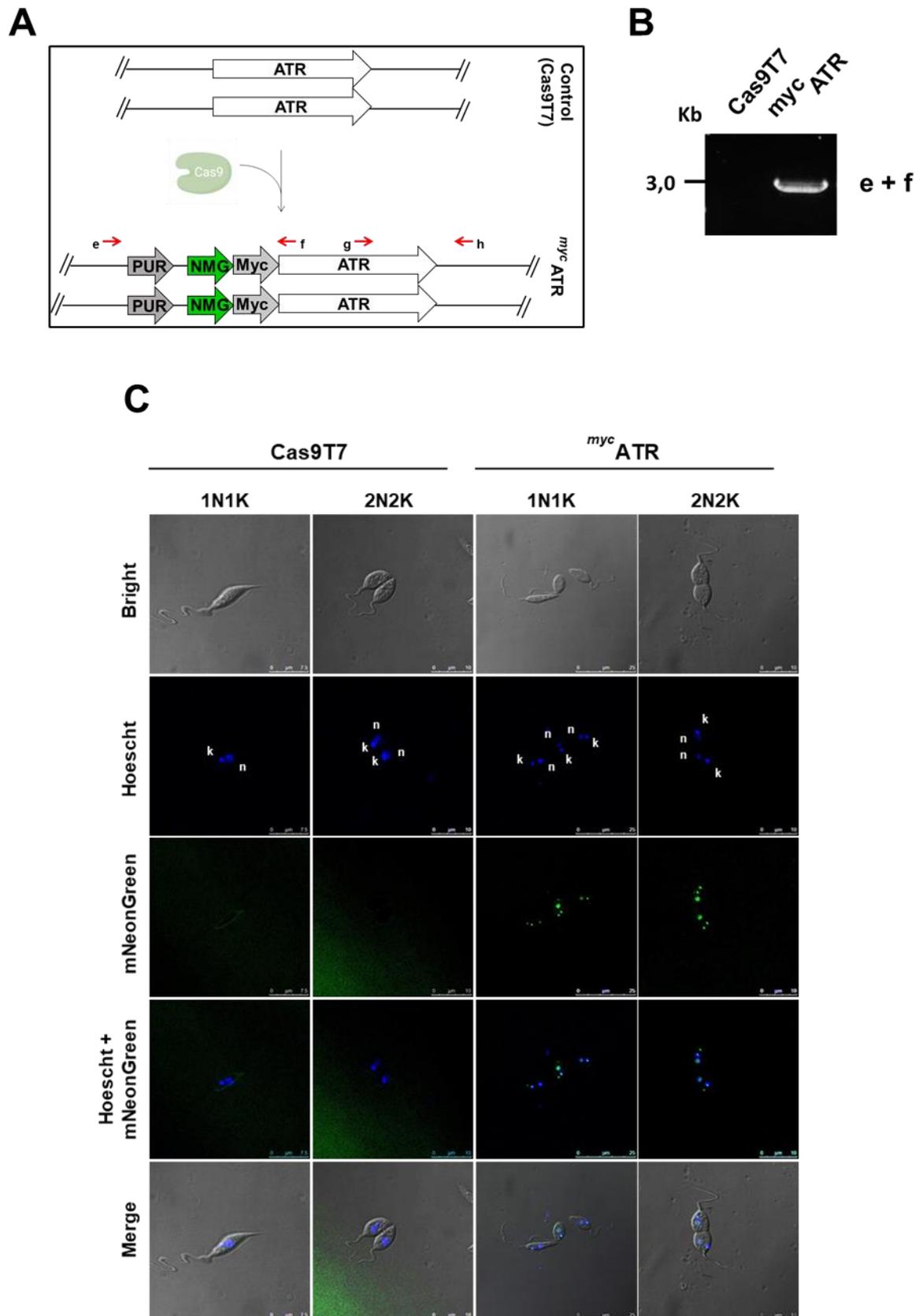
Figure 10 - L. major ATR kinase locates at nuclear, kinetoplast and flagella tip. A) Schematic representation showing a cell line expressing Cas9 which was used to add a tag (NeonGreen + 3xmyc) N'terminally on LmJATR kinase (LmjF.32.1460) generating ^{myc}ATR cell line. **B)** PCR analysis from ^{myc}ATR cell line indicating the addition of the tagging in two alleles of both cell lines (e + f). **C)** Representative images of the sub-cellular localization of mNeongreen signal which is expressed N'terminally in live ^{myc}ATR cell; scale bar = 7.5, 10, 25 µm. Images were captured on a SP6 microscope (Leica). Nuclear (n) and Kinetoplast (k) DNA shown in blue, were stained with 1x M199 and Hoescht solution for 10 min/ 27°C followed by image acquisition on the microscope; the mNeongreen signal is in green.

3.1.3 The kinetoplast localisation of *L. major* ATR kinase changes in cells synthetising kDNA.

The previous results showed that ATR localises within the nucleus and in proximity to the kinetoplast. The kinetoplast DNA (kDNA) is composed of two types of DNA rings called minicircles and maxicircles; and its replication begins with the assemble of several proteins in two antipodal sites flanking the kDNA disk and positioned 180^θ apart (Figure 11A) (Lui et al., 2005). Our results in the section prior revealed deficiency of ATR combined with replication stress also causes ssDNA accumulation at these compartments suggesting ATR may operate to resolve ssDNA regions at these sites during the DNA replication process. To ask if ATR may be involved during DNA replication of the nuclear and/or kinetoplast DNA replication, we incubated ^{myc}ATR cells and Cas9T7 cells with a 30-minute pulse of the thymidine analogue EdU. The cells were fixed and the EdU detectd by Click-IT chemistry (see methods section). The subcellular localisation of ATR was also examined by indirect IFA using α-myc. Images were captured on a Deconvolution microscope (Figure 11A). With the EdU pulse we can quantify the number of cells that are synthetising DNA as they appear positive for EdU signal. Both nuclear and kDNA incorporate EdU during DNA synthesis, and here we opted to focus specifically on replication within the kinetoplast. Here, the EdU signal appears as two dots at the kinetoplast extremities which correspond to the locations of the kinetoplast replication machinery (Shapiro, Englund, 2015) (Figure 11A). The percentage of cells synthetising kDNA were similar between the two cell lines (Figure 11B).

In cells with positive EdU signal, ATR kinetoplast signal appears with one dot or two dots shape, where the proportion of each conformation is around 50% of the total of ^{myc}ATR EdU positive cells (Figure 11C). However, in cells with no EdU signal only a

single dot is detected at the kinetoplast (Figure 11D). These results suggest the changes on subcellular location of ATR kinase at the kDNA correlates with active kinetoplast DNA synthesis.

Figure 11 – *L. major* ATR kinase kinetoplast location duplicates in cells synthetising DNA.
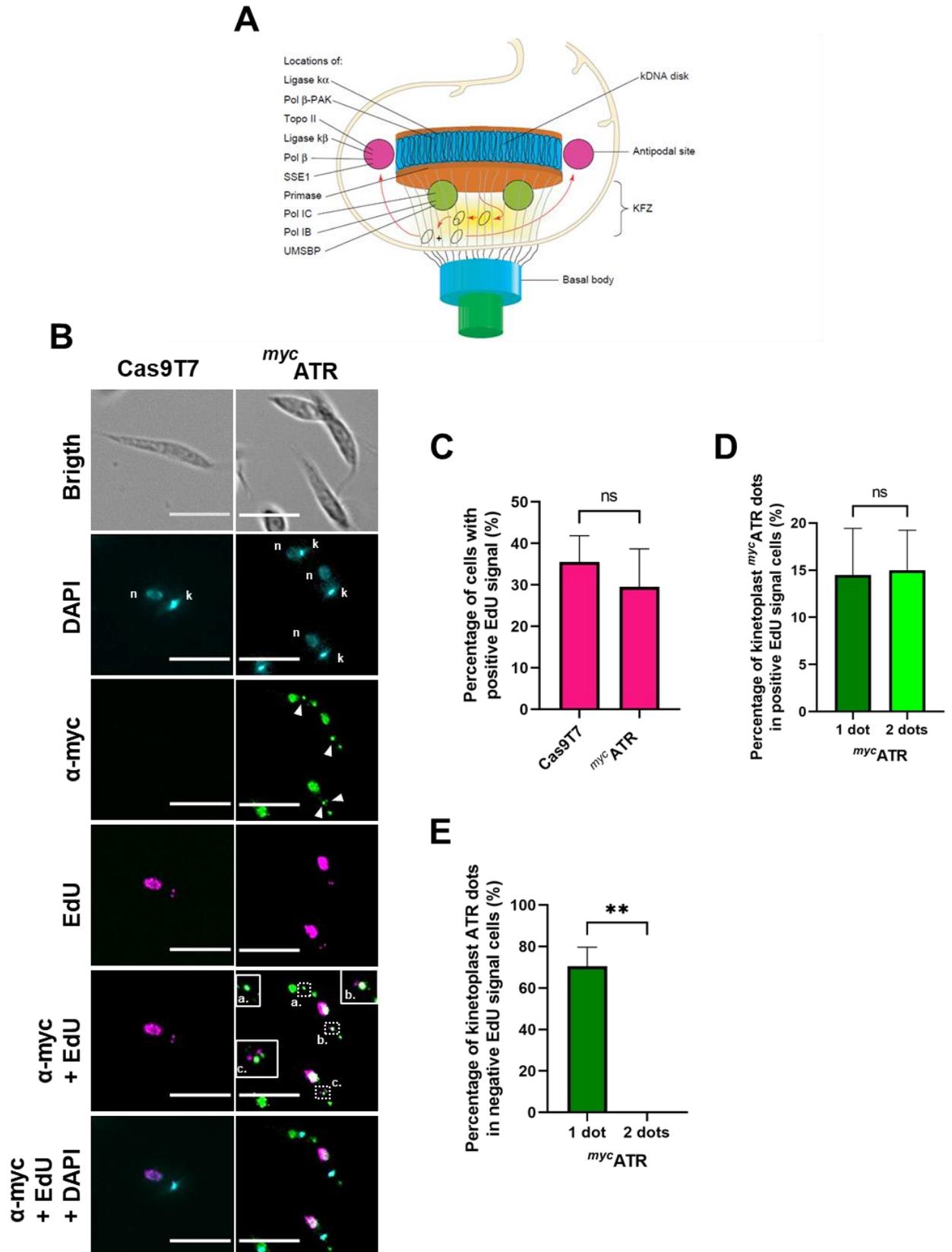
Figure 11 - *L.* **major** **ATR kinase kinetoplast location duplicates in cells synthetising DNA. A)** Schematic representation of kDNA-replication model based in studies in *Crithidea fasciculata* and *Trypanosoma brucei* (Figure from Liu et al., 2005). **B)** Representative images showing Cas9T7 and ^{myc}ATR cells that incorporated the thymidine analog EdU in untreated condition and exposed to EdU pulse for another 30 min. The images were acquired on a DV microscope (Leica) and the channels are: ATR (anti-myc) in green, the EdU in magenta and the DNA was stained with DAPI and shown in cyan; head arrow show the ATR dots; dashed square showed an amplified region for better visualization (a. single ATR dot negative EdU, b. single ATR dot positive EdU, c. double ATR dot positive EdU). bar scale = 8 µm. **C)** Quantification of the percentage of positive EdU signal. Error bars ± SD, n = 2 (100 cells/experiment). Unpaired t-test. **D)** Quantification of percentage of kinetoplast ^{myc}ATR dots in positive EdU signal cells. Error bars ± SD, n = 2 (100 cells/experiment). Unpaired t-test. **E)** Quantification of percentage of kinetoplast ^{myc}ATR dots in negative EdU signal cells. Error bars ± SD, n = 2 (100 cells/experiment). ** P<0,005, Unpaired t-test.

3.1.4 *L. major* ATR kinase subcellular location depends on the protein C-terminal domain.

Next, we aimed to understand if the C-terminal region of ATR was important for the localisation pattern we describe above. To answer this question, we deleted the C-terminal region using the CRISPR-Cas9 system in our ^{myc}ATR cell line. Specifically, we designed a guide RNA which would result in the deletion of the predicted kinase FAT and FATC domains (from ~6220 to 9624 pb), as in Figure 9A. The deleted region was replaced by the selected marker Neomycin (G418) and cells selected in HOMEM media (Figure 12A). The truncation to the ATR locus was confirmed by PCR (Figure 12B) and the population of cells resistant to the selectable marker named ^{myc}ATR$^{\Delta C}$ cells. We also verified that the N-terminal tag was expressed on the truncated version of ATR. To check the levels and the size of both modified proteins, ^{myc}ATR and ^{myc}ATR$^{\Delta C}$ cells, we performed a western blot analysis probing with anti-myc antiserum to detect the myc tag. Based on its predicted amino acid sequence, the molecular mass of *L. major* ATR is approximately 351 kilodaltons (Kda; Tritryp.db). Western analysis revealed an estimated 420 Kda tagged protein in ^{myc}ATR cells; and an estimated 238 Kda protein in the ^{myc}ATR$^{\Delta C}$ cell line (Figure 12C). It is noteworthy that the truncation of the ATR C-terminal region affected the levels of the protein (discussed below).

Due to the suspected low levels of the truncated ATR protein, the visualisation of the mNG in ^{myc}ATR$^{\Delta C}$ cells for live cells analysis proved challenging. To overcome this issue to investigate the location of ^{myc}ATR$^{\Delta C}$, we formaldehyde fixed cells and probed for the myc tag using anti-myc antibodies (α-myc) by indirect IFA. Images were

captured by structural illuminated super resolution microscopy (SR-SIM) on an Elyra microscope (Zeiss). Consistent with what was observed in live cells, ^{myc}ATR accumulated within the nuclear compartment, proximal to the kinetoplast and nearby the site of the flagella pocket (Figure 12D). The analysis of ATR$^{\Delta C}$ cells revealed a different localization pattern. ATR signal was distributed throughout the cell body, likely within the cytoplasm (Figure 12D). These results suggests that the truncation of the C-terminal domain impaired the localization of ATR to the nucleus and the kDNA compared with cells expressing full-length ATR.

Figure 12 – *L. major* **ATR kinase subcellular location depends on the kinase C-terminal.**
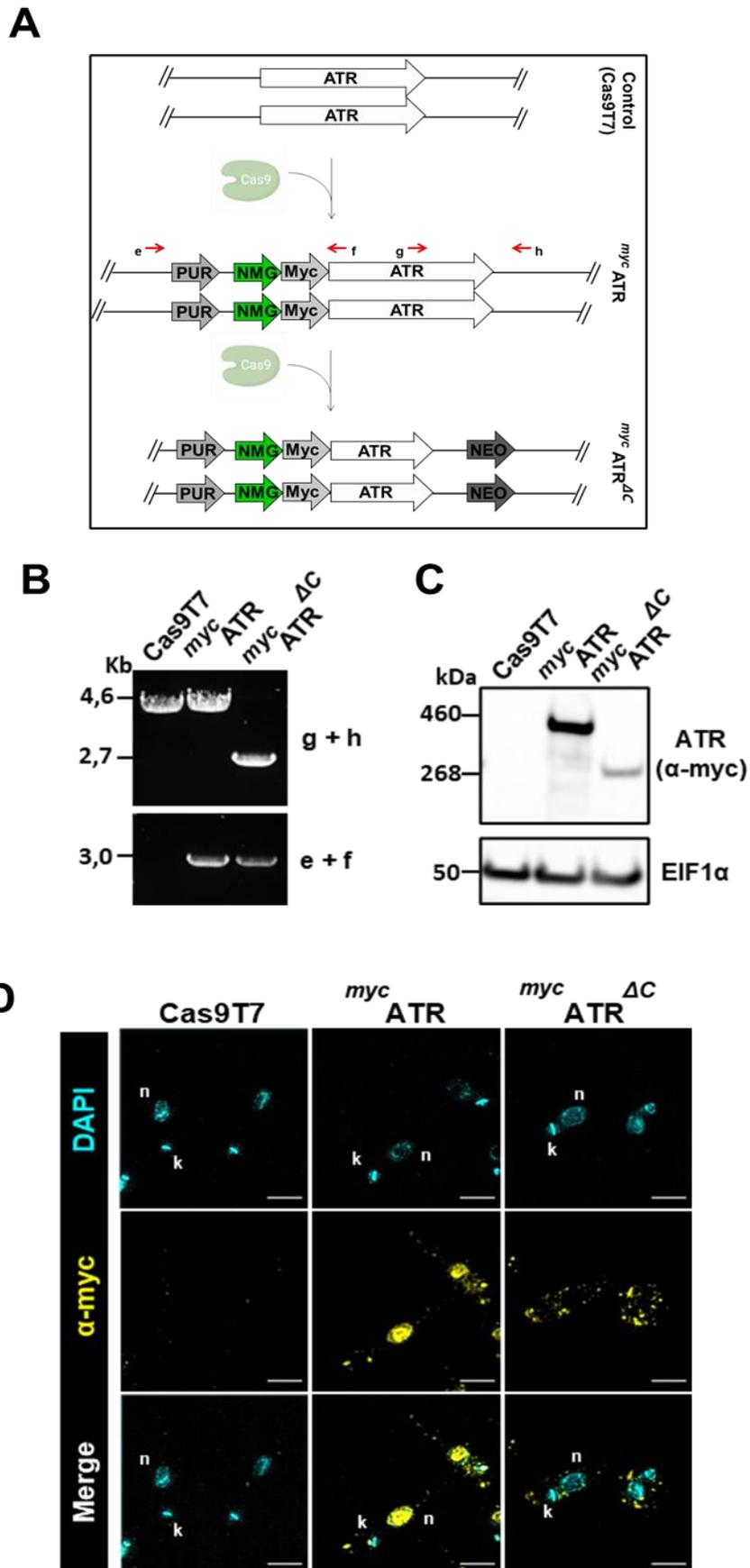
Figure 12 - *L. major* ATR kinase subcellular location depends on the kinase C-terminal. A) Schematic representation showing a cell line expressing Cas9 which was used to add a tag (3xmyc + NeonGreen) N'terminally on LmJATR kinase (LmjF.32.1460) generating ^{myc}ATR cell line. The predicted kinase domains present on the C-terminal region, indicated by a black arrow (~6220 – 9624 pb), of ^{myc}ATR cell line were replaced by a selected marker (NEO), generating ^{myc}ATR$^{\Delta C}$ cell line; red arrow represents the primers used to amplify those alterations. **B)** PCR analysis from ^{myc}ATR and ^{myc}ATR$^{\Delta C}$ cell lines indicating the addition of the tagging in two alleles of both cell lines (e + f) and the deletion of the C-terminal region (g + h) in both alleles of ^{myc}ATR$^{\Delta C}$ cells. **C)** Western analysis of whole cell extract of indicated cell lines in exponential growth; extracts were probed with anti-myc antiserum and anti-EIF1α were used as a loading control (predicted protein sizes are indicated, kDa). **D)** Representative images of the sub-cellular localization of myc signal which is expressed N-terminally in ^{myc}ATR and ^{myc}ATR$^{\Delta C}$ cells; scale bar = 2.6 µm. Images were captured on an Elyra super resolution microscope (Zeiss). Nuclear (n) and Kinetoplast (k) DNA are shown in cyan; and the myc signal in yellow.

Next, we investigated if ^{myc}ATR co-localise with DNA synthesis at the nuclear compartment, since the canonical function of ATR as a safeguard of replication is described. We pulsed ^{myc}ATR and ^{myc}ATR$^{\Delta C}$ cells (as a control since ATR lost its location at the nuclear compartment) with 30 min of EdU, the cells fixed and probed with anti-myc to detect endogenously tagged ATR. Images were acquired as before (Figure 13A).. The results showed that 50% of the ATR signal in ^{myc}ATR cells co-localise with EdU signal (Figure 13B); On the other hand, the ATR signal in ^{myc}ATR$^{\Delta C}$ cells and the EdU signal showed a poor correlation 9% (Figure 13B). This result showed that ATR, when present at the nuclear compartment, co-localise with DNA synthesis.

Figure 13 - ^{myc}ATR partially co-localises with DNA synthesis.
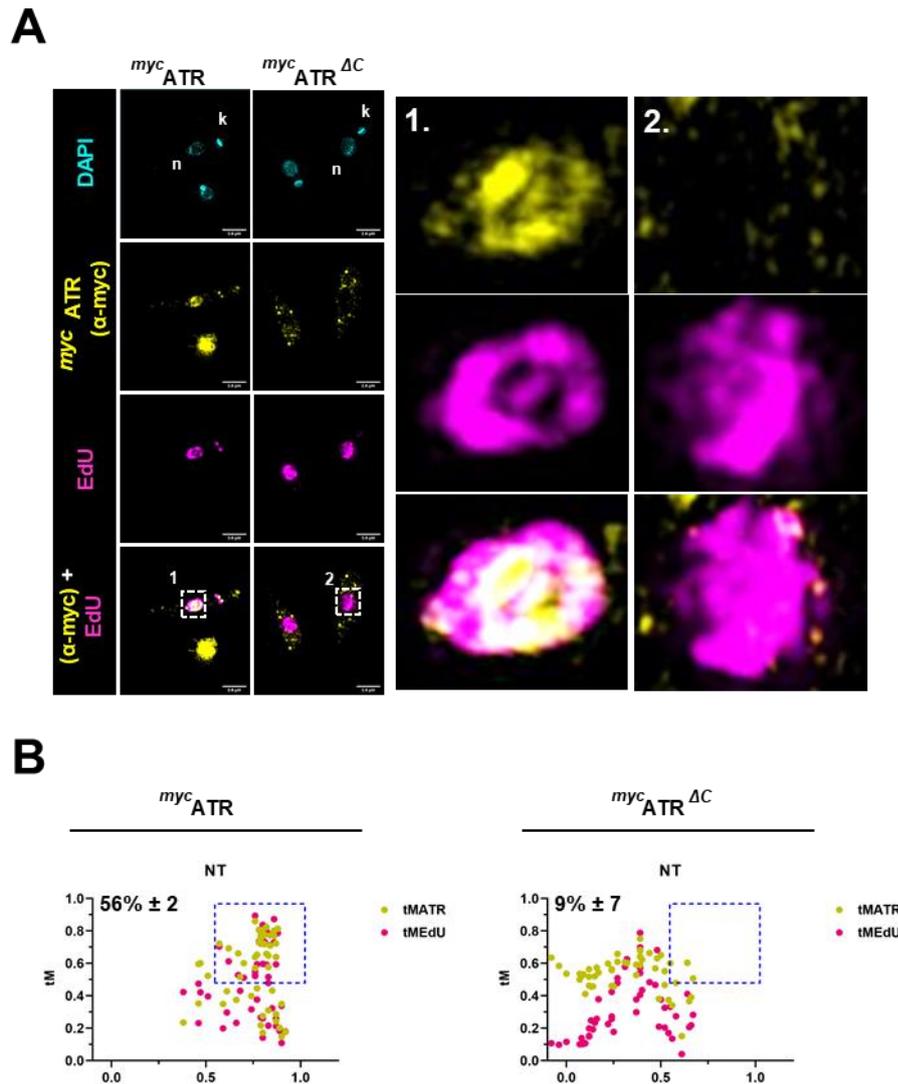
Figure 13 - ^{myc}ATR partially co-localises with DNA synthesis. A) Representative images of ^{myc}ATR cells and ^{myc}ATR^{ΔC} cells in untreated condition showing the co-localization of the endogenous ATR tagging and EdU at nuclear compartment. The cells were fixed and prepared to myc and EdU detection and the images acquired on Elyra microscope (Zeiss), scale bar = 8 µm. (B) Quantitative analysis of co-localization between endogenous tagged ATR with EdU in ^{myc}ATR and ^{myc}ATR^{ΔC} cells in untreated and submit to EdU pulse for 30 min. Co-localization between ATR and EdU in the nucleus of all samples was quantified in a single fluorescence slide using Coloc 2 Fiji plugin. Two factors were taking into account to determine the relevant co-localization events; Pearson's R Value (special correlation between two signals) and the Thresholded Mander's (tMEdU, proportion of EdU signal overlapping with ATR (myc); tMATR, proportion of ATR (myc) signal overlapping with EdU). Blue box indicates the nucleus where co-localization where above 0.5 for all parameters; mean ± SD, n = 2 experiments.

3.1.5 Loss of ATR C-terminal leads to increased DNA content and nuclear area.

Considering that ATR is a key factor in the regulation of the cell cycle profile, we used FACS analysis to characterize cell cycle progression in the ^{myc}ATR and ^{myc}ATR$^{\Delta C}$ cell lines. Surprisingly, we found ^{myc}ATR$^{\Delta C}$ cells contained an overal increased DNA content when compared to ^{myc}ATR cells (Figure 14A). The ^{myc}ATR$^{\Delta C}$ population and ^{myc}ATR$^{\Delta C}$ individual clones, showed an increase of at least ~50% content of DNA when compared to ^{myc}ATR cells (Figure 14A). Except for the increase in DNA content, the cell cycle profile of ^{myc}ATR$^{\Delta C}$ cells was comparable to that of ^{myc}ATR cells. This profile was consistent among individual clones and the selected population of ^{myc}ATR$^{\Delta C}$ cells, suggesting that the population is homogeneous in respect to this feature (Figure 14A).

Given the increased the DNA content, we asked if this effect also correlated with alterations to nuclear size. Therefore, we decided to measure the nuclear area based on DAPI stained microscopy images of ^{myc}ATR and ^{myc}ATR$^{\Delta C}$ cells (Figure 14B). The result showed a significant increase of the nuclear area in ^{myc}ATR$^{\Delta C}$ cells (mean of ~2900 A.U.) when compared to ^{myc}ATR cells (mean of ~1740 A.U.) (Figure 14C); which could be a consequence of the increased DNA content observed in these cells.

Figure 14 - Loss of ATR C-terminal leads to increased DNA content and nuclear area.
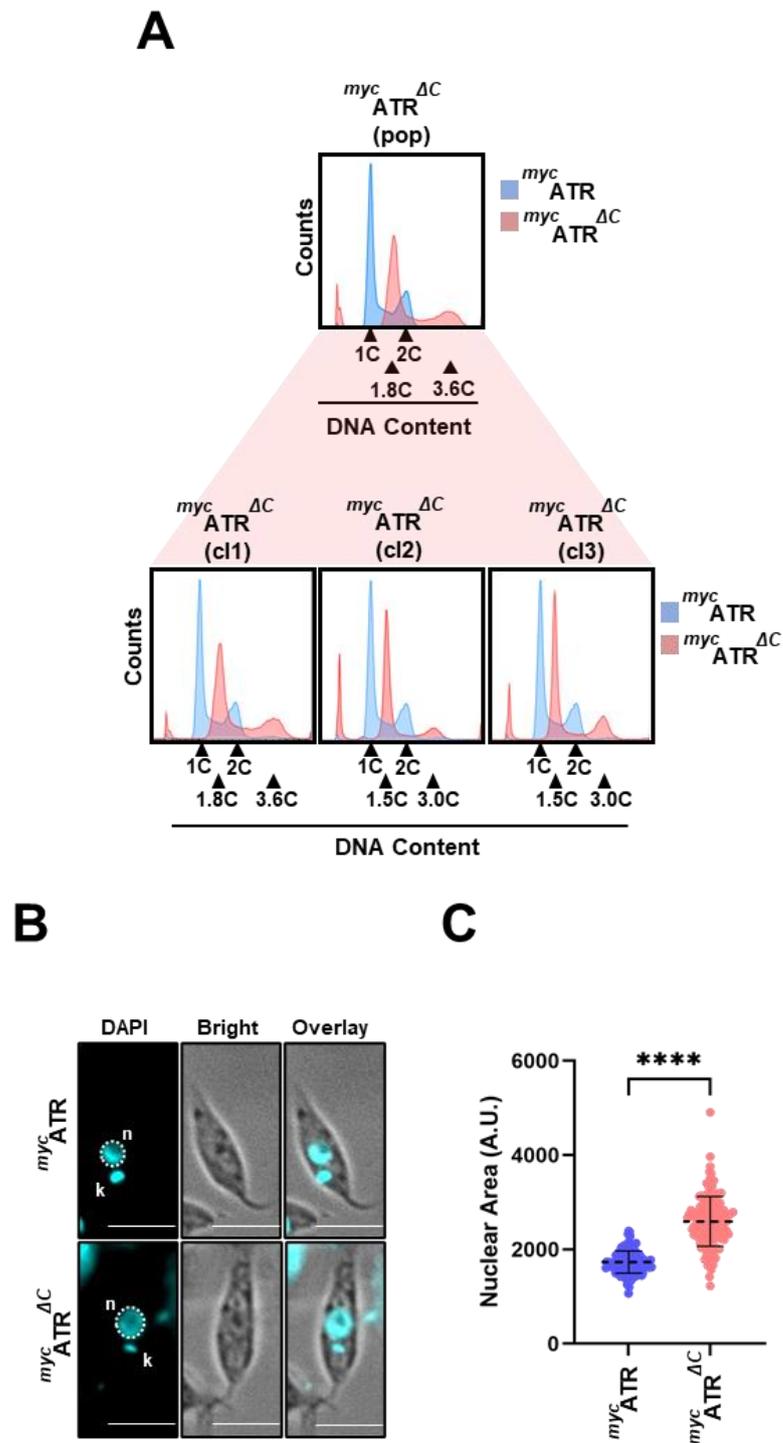
Figure 14 - Loss of ATR C-terminal leads to an increased DNA content and nuclear area. A) DNA content analysis by FACS of ^{myc}ATR and ^{myc}ATR$^{\Delta C}$ cells (population, Cl1, Cl2, Cl3), in exponential growth; DNA was stained with Propidium iodide (PI). 1C and 2C indicate single and double DNA content for ^{myc}ATR cells, respectively; and 1.8C and 3.6C indicate single and double DNA content for ^{myc}ATR$^{\Delta C}$ cells (population and Cl1), respectively; and 1.5C and 3.0C indicate single and double DNA content for ^{myc}ATR$^{\Delta C}$ cells (Cl2 and Cl3), respectively. The coral shadow represents that the clones are originated from the selected ^{myc}ATR$^{\Delta C}$ population. **B)** Representative images of nuclear area based on DAPI staining (dashed circle) in ^{myc}ATR and ^{myc}ATR cells, scale bar = 4.7 µm. Images were captured on a DV microscope (Leica). Nuclear (n) and Kinetoplast (k) DNA shown in cyan. **C)** The nuclear areas were measured using ImageJ software and are shown as Arbitrary Units values (A.U.). Error bars ± SD; n=3 experiments (>150 cells counted/experiment). (****) p-value<0.0001, unpaired t-test.

3.1.6 ^{myc}ATR$^{\Delta C}$ cells show an altered cell cycle progression upon and after replication stress.

We used FACS analysis to investigate the cell cycle profile progression in ^{myc}ATR and ^{myc}ATR$^{\Delta C}$ cells during and after chronic or acute replication stress induced by the addition of HU; intermediate time points were collected during each treatment to monitor cell cycle progression (Figure 15A).

Under chronic treatment, we could observe that the ^{myc}ATR cells slowly progressed throughout the cell cycle, and at 8 hours of treatment the majority of cells are at the S phase (~70%), completing the cell cycle after 20 hours. However, when we examined the profile of ^{myc}ATR$^{\Delta C}$ cells this progression was compromised, since the cells did not seem to performance a similar progression to S phase (~35%) as observed with the control. At 8 hours treatment, a considerable proportion of cells still in G1 or early S stage (~35%); and, at 20 hours, ^{myc}ATR$^{\Delta C}$ the cells were still distributed along the S phase (~30%). Moreover, it is noticeable the considerable increase in the sub-G1 population in ^{myc}ATR$^{\Delta C}$ cells when exposed to replication stress cells during 8 and 20 hours of treatment (~12% and ~22%, respectively) (Figure 15C, top). After the removal of the HU, the cell cycle progression of ^{myc}ATR cells continue to show a normal profile without major alteration, but, in other hand, the cell cycle profile of ^{myc}ATR$^{\Delta C}$ cells showed unchanged even after the HU removal with high proportion of Sub G1 population cells (~20%). (Figure 15C bottom)

As shown in previous work (Damasceno et al., 2016), under acute HU treatment *L. major* wildtype cells synchronise at the G1-S boundary. This is the profile we observed when ^{myc}ATR cells are exposed to this treatment for 8 hours (Figure 15B).

After removal of HU, ^{myc}ATR cells begin to progress synchronously as expected, re-establishing the normal profile after 8 hours (Figure 15B). However, the cell cycle profile in ^{myc}ATR$^{\Delta C}$ cells also synchronise upon acute treatment but, as observed after chronic treatment, these mutated ATR cells begin to accumulate sub-G1 cells after HU removal (~42%) (Figure 15B and 15D). These data indicated that a possible G1/S checkpoint inoperative in ^{myc}ATR$^{\Delta C}$ cells.

Figure 15 - ^{myc}ATR$^{\Delta C}$ cells show an altered cell cycle progression upon and after replication stress.
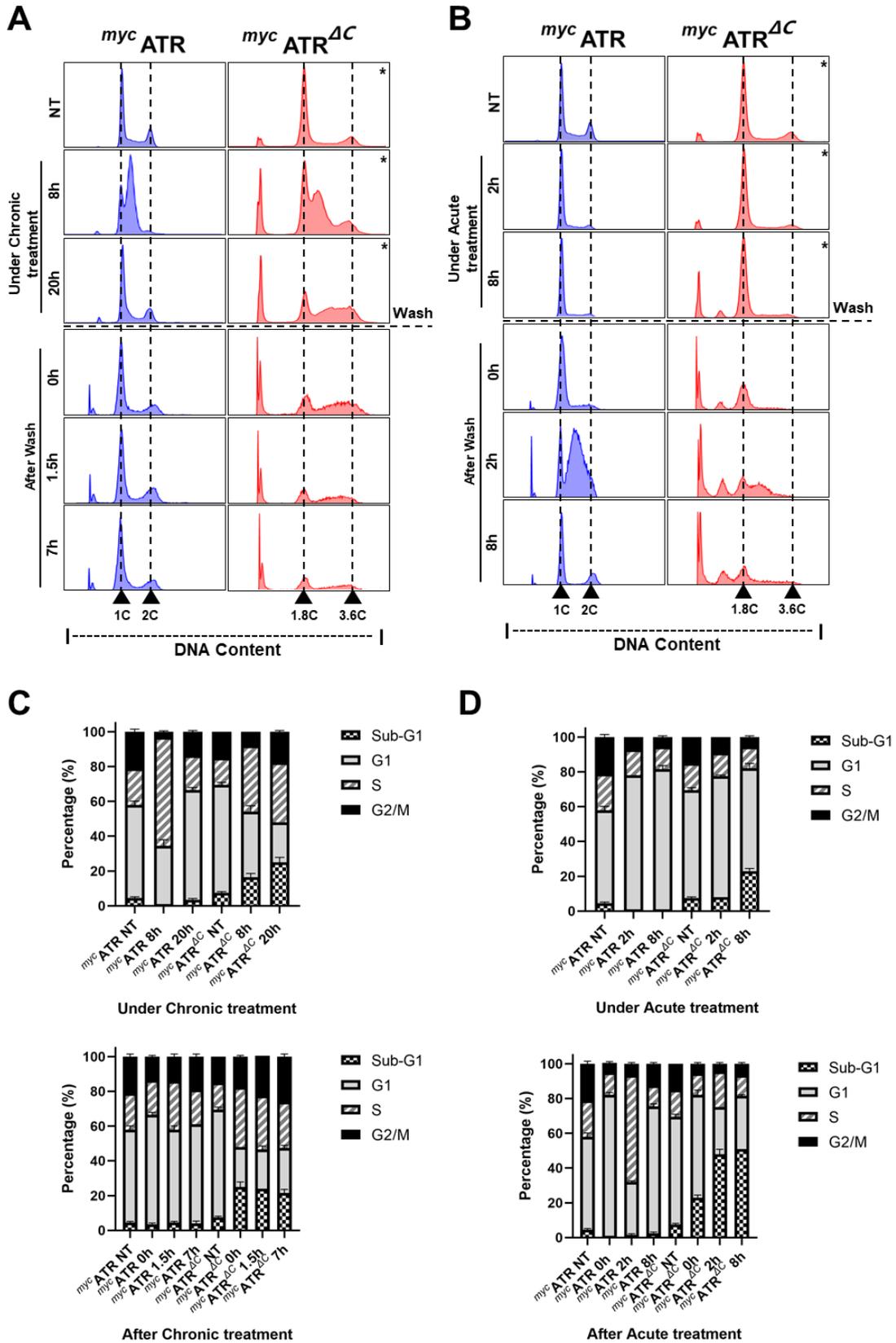
Figure 15 - ^{myc}**ATR**^{ΔC} **cells showed an altered cell cycle progression upon and after replication stress. A)** ^{myc}ATR and ^{myc}ATR^{ΔC} cells were left untreated (NT) or treated with 0.5mM of Hydroxyurea (HU) for 8 and 20 hours (Chronic). In an independent experiment the cells were treated with 0.5mM of Hydroxyurea (HU) for 20 hours, washed and resuspended in fresh media, and time points collected in 0, 1.5, 7 hours after the treatment. Cells were fixed with methanol and the DNA stained with propidium iodate, followed by flow cytometer analysis. **B)** ^{myc}ATR and ^{myc}ATR^{ΔC} cells were left untreated (NT) or treated with 5mM of Hydroxyurea (HU) for 2 and 8 hours (Acute), In an independent experiment the cells were treated with 5mM of Hydroxyurea (HU) for 8 hours, washed and resuspended in fresh media, and time points collected in 0, 2, 8 hours after the treatment. Cells were fixed with methanol and the DNA stained with propidium iodate, followed by flow cytometer analysis**. C)** Percentage of cells in each cell cycle stage in each time point of the (A). The gates based on the DNA content were used to sort the cells into Sub-G1, G1, S and G2/M phases; 1C and 2C indicate single and double DNA content in ^{myc}ATR cells, respectively; and 1.8C and 3.6C indicate single and double DNA content in ^{myc}ATR^{ΔC} cells; 30,000 cells of each time point were analysed; Error bars ± SD; n=2 experiments. t-test (paired). **D)** Percentage of cells in each cell cycle stage in each time point of the (B). The gates based on the DNA content were used to sort the cells into Sub-G1, G1, S and G2/M phases; 1C and 2C indicate single and double DNA content in ^{myc}ATR cells, respectively; and 1.8C and 3.6C indicate single and double DNA content in ^{myc}ATR^{ΔC} cells; 30,000 cells of each time point were analysed; Error bars ± SD; n=2 experiments. t-test (paired). *FACS from the samples that were also used on the MFA-seq (see below).

The progression through the cell cycle in trypanosomatids can also be investigated by analysing the proportion of nucleus (n) to kinetoplast (k) DNA via DAPI staining in microscopy images. Cell cycle progression can be represented through the distinct patterns of 1n1k, 1n2k and 2n2k cells (Ambit et al., 2011). Hence, we fixed cells as per the conditions described for FAC and performed a cell cycle profile analysis by counting the proportion of "n" and "k" in ATR^{WT} and ATR^{ΔC} cells (Figure 16A). The result showed that, for unperturbed cells, the proportion of cells with abnormal DNA content (0n1k/1n0k or 2n1k) was undetectable for ATR^{WT} cells and ~5% for ATR^{ΔC} cells. Upon replication stress, ATR^{WT} cells presented 5% and 2% of abnormal cells after chronic and acute stress, respectively. The effect of the treatment was more evident among ATR^{ΔC} cells where the increase of abnormal cells was ~12% after chronic treatment and ~10% after acute treatment (Figure 16B). Altogether, these data suggested that ATR plays a relevant role in genome homeostasis in these cells.

Figure 16 - ^{myc}ATR^{ΔC} cells show an increase in cells with abnormal DNA content after HU treatment.
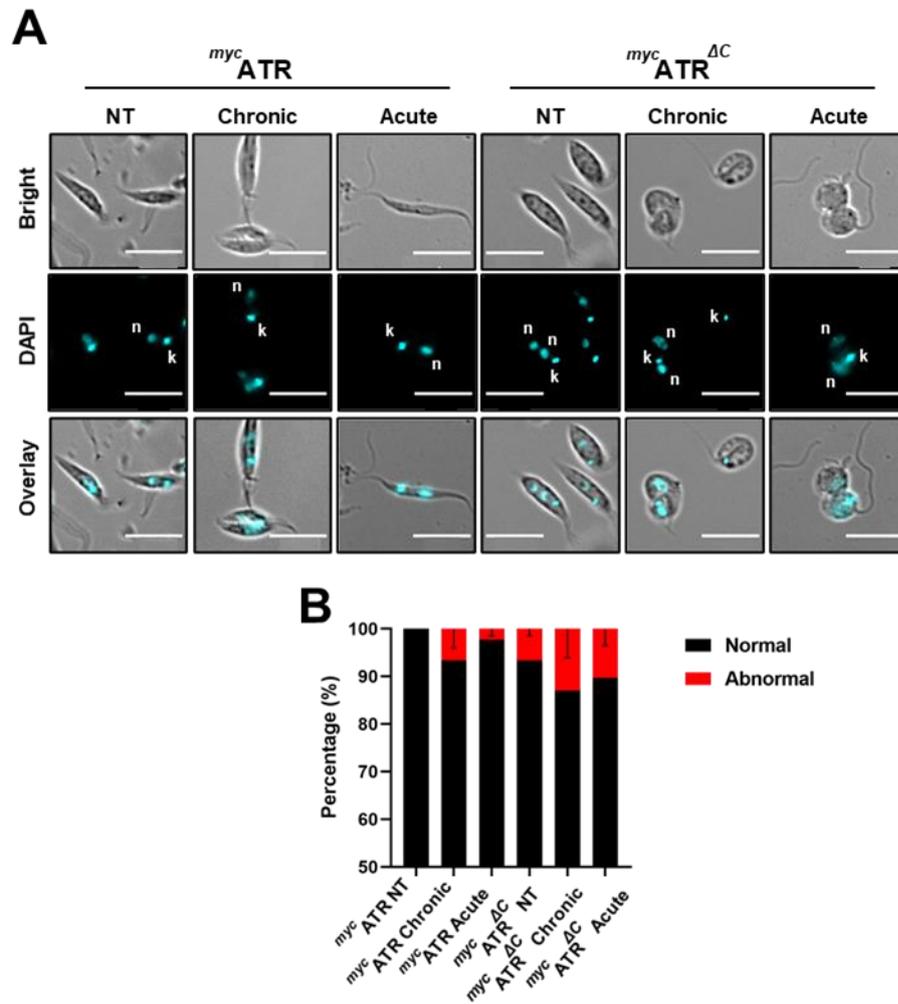
Figure 16 - ^{myc}ATR^{ΔC} cells showed an increase in cells with abnormal DNA content after HU treatment. A) Representative images of ^{myc}ATR and ^{myc}ATR^{ΔC} cells untreated (NT) or treated with 0.5mM of HU for 20 hours (Chronic) or with 5mM of HU for 8 hours (Acute); scale bar = 8µm. Images were captured on a DV microscope (Leica). Nuclear (n) and Kinetoplast (k) DNA were stained with DAPI and shown in cyan. **B)** Percentage of DNA content populations based on the proportion of Nucleus/Kinetoplast (n/k) in ^{myc}ATR and ^{myc}ATR^{ΔC} cells untreated (NT) or treated with 0.5mM of HU for 20 hours (Chronic) or with 5mM of HU for 8 hours (Acute). The populations were classified according to the proportion of n/k in Normal (1n1k, 1n2n, 2n2k) and Abnormal (0n1k, 1n0k, 2n1k). Error bars ± SD, n = 3 (>100 cells counted/experiment).

3.1.7 Loss of ATR C-terminal decreases the levels of the kinase and correlates with increased yH2A levels, a marker of DNA damage, after HU treatment.

We next asked if ATR levels in ^{myc}ATR and ^{myc}ATR^{ΔC} cells changed after chronic and acute HU treatment. Cells were treated or harvested in the absence of chronic or acute HU treatment, after which the cells were washed and resuspended in fresh HOMEM media devoid of HU. Whole protein extracts were prepared from cells collected at 0, 8 and 20 hours after chronic treatment and 0, 4 and 8 hours after acute treatment. Western blotting analysis using anti-myc antibodies showed that the levels of ATR in ^{myc}ATR and ^{myc}ATR^{ΔC} cells were not affected upon treatement (Figure 17A). However, as the result in Figure 12C indicates, the level of ATR is significantly reduced in ^{myc}ATR^{ΔC} cells.

Next, we asked if truncation of ATR affected the protein levels of putative downstream factors, including under different treatments of HU. The same conditions as above were used, excluding the 4 hour time point after acute treatment. The 9-1-1 complex has been described in *L. major* by our laboratory and it seems to have divergent and conserved functions befitting with the canonical ATR pathway described in other eukaryotes (Damasceno et al., 2013; Damasceno et al., 2016; Nunes et al., 2013). The western blot was probed to examine protein levels of two 9-1-1 components, Hus1 and Rad9. The results show that both Hus1 and Rad9 levels do not change after HU treatment or in response to the deletion of ATR C-terminal (Figure 17B). We also tested for accumulation of yH2A, a marker of DNA damage in trypanosomatids. These results show a significant increase in yH2A levels in ^{myc}ATR^{ΔC} cells in untreated and after chronic HU treatment in comparison with ^{myc}ATR cells (Figure 17B Right). Surprisingly, after acute treatment the levels of yH2A were found to be significantly higher in ^{myc}ATR cells than in ^{myc}ATR^{ΔC} cells after 8 hours after the treatment which could indicate H2A phosphorylation (Figure 17B Left).

Figure 17 - Loss of ATR C-terminal decreases the protein levels of the kinase and correlates with increased yH2A levels after HU treatment.
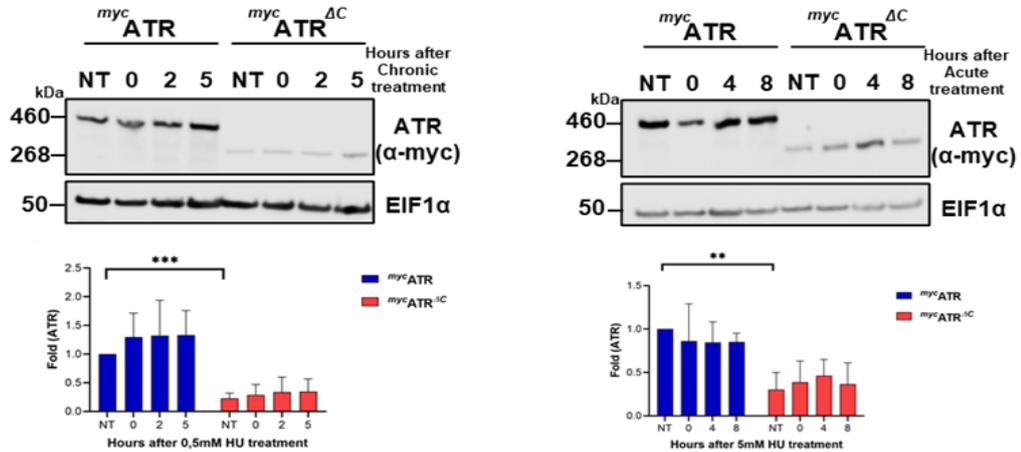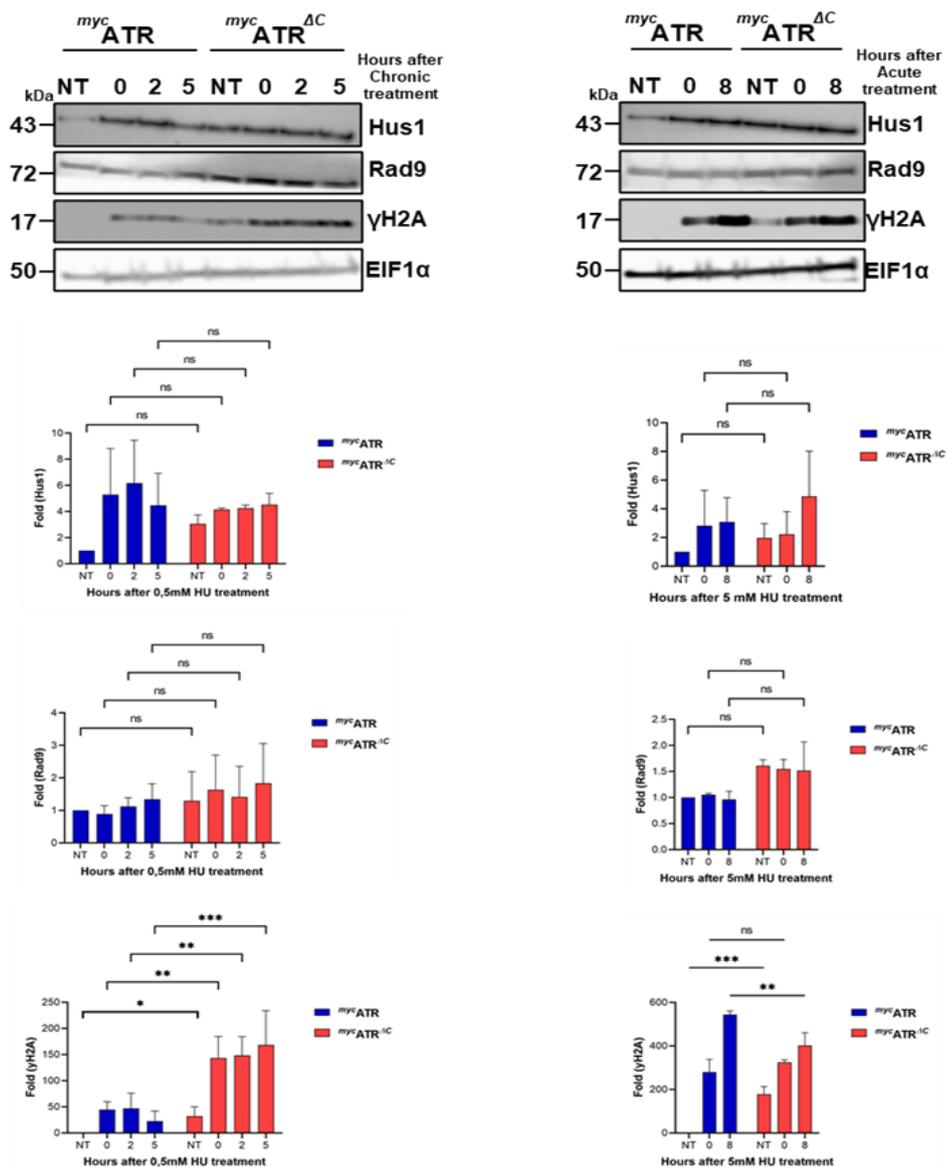
Figure 17 - Loss of ATR C-terminal decreases the protein levels of the kinase and correlates with increased yH2A levels after HU treatment. A) Western analysis of whole cell extract of each time points in (A) after chronic and acute HU treatment; extracts were probe to anti-myc antiserum (ATR) and anti-EIF1α was used as a loading control (predicted protein sizes are indicated, kDa). The quantification of the signal was measured at ImageJ software; Error bars ± SD; n = 3 experiments. * P<0,05; ** P<0,005; *** P<0,001. Unpaired t-test **B)** Western analysis of whole cell extract of each time points in (A) after chronic and acute HU treatment; extracts were probe to anti-hus1, anti-rad9 and anti-γH2A antiserum; an anti-EIF1α was used as a loading control (predicted protein sizes are indicated, kDa). The quantification of the signal was measured at ImageJ software; Error bars ± SD; n = 3 experiments. * P<0,05; ** P<0,005; *** P<0,001. Unpaired t-test.

3.1.8 ^{myc}ATR$^{ΔC+/-}$ cells do not show cell cycle defects or an increasement of yH2A levels in comparison with ^{myc}ATR cells.

We next asked if partial loss of the ATR C-terminal region causes the phenotypes shown in ^{myc}ATR^{ΔC} cells. We used the same strategy designed to generate ^{myc}ATR^{ΔC} cells to try to select a cell line where only one allele of the ATR C-terminal region was deleted using the same background cells (^{myc}ATR) to avoid any possible bias of the tagging presence on the N-terminal of the kinase could case (Figure 18A). The modification of one allele of the ATR locus was confirmed by PCR using the genomic DNA from selected population and from one clone (Figure 18B). The new heterozygote cell line was called ^{myc}ATR$^{ΔC+/-}$.

To examine for proliferative defects, we seeded ^{myc}ATR$^{ΔC+/-}$ cells at a concentration of $5x10^5$ cell/mL and cell density was measured every 24 hours for 5 days by counting. The ^{myc}ATR$^{ΔC+/-}$ cells did not show growth defects, in the population and/or the clone, in comparison with ^{myc}ATR cells (Figure 18C). Next, FACS analysis was performed to assess cell cycle progression in ^{myc}ATR$^{ΔC+/-}$ cells, relative to controls. Both, heterozygote cell lines and ^{myc}ATR cells were left untreated or treated with chronic or acute HU treatment; the cultures were washed and resuspended in fresh media and samples from different time points were collected and analysed by flow cytometry. The result showed that ^{myc}ATR$^{ΔC+/-}$ cells have a similar cell cycle profile in comparison with ^{myc}ATR cells in both conditions (Figure 18D). In addition, the heterozygote cells do not show an increase in DNA content or accumulation of a sub-G1 population, as observed in ^{myc}ATR^{ΔC} cells (Figure 18D). Lastly, we measured the levels of yH2A after chronic or acute treatment. The experiment was set as the one in section 3.1.7. The western blot analysis showed that in untreated cells and after each

da Silva, Gabriel Lamak Almeida   *Results*

HU treatments the levels of yH2A were not significantly different in ^{myc}ATR$^{\Delta C+/-}$ cells whem compared to ^{myc}ATR cells. However, the yH2A levels in ^{myc}ATR$^{\Delta C+/-}$ cells was similar to that observed in ^{myc}ATR$^{\Delta C}$ cells (Figure 18E).

Figure 18 - ^{myc}ATR$^{ΔC+/-}$ cells do not show cell cycle defects or an increasement of yH2A levels in comparison with ^{myc}ATR cells.
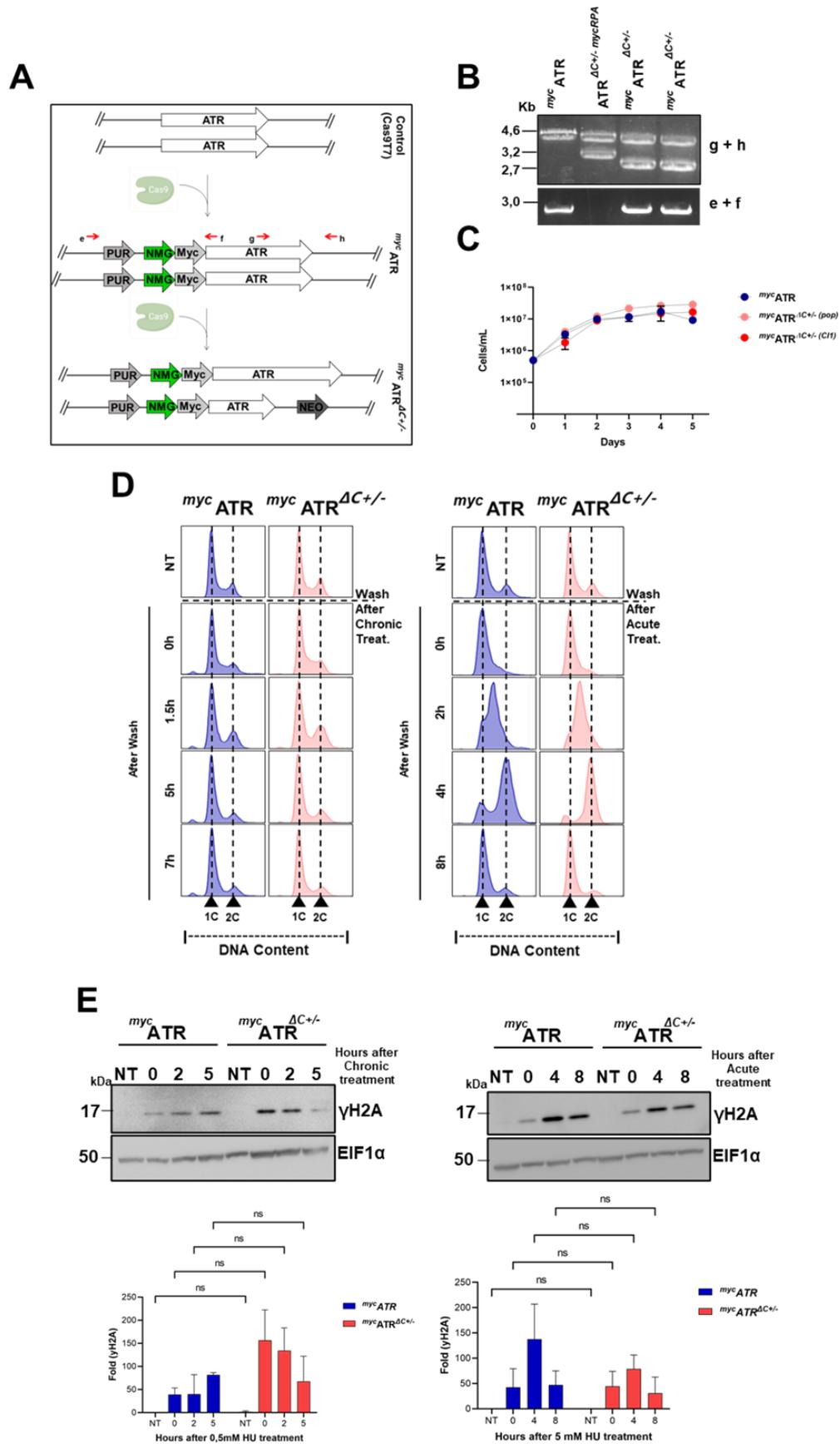
Figure 18 - ^{myc}ATR$^{ΔC+/-}$ cells do not show cell cycle defects or an increasement of yH2A levels in comparison with ^{myc}ATR cells. A) Schematic representation showing a cell line expressing Cas9 which was used to add a tag (3xmyc + NeonGreen) N'terminally on LmJATR kinase (LmjF.32.1460) generating ^{myc}ATR cell line. The predicted kinase domains present on the C-terminal region, indicated by a black arrow (~6220 – 9624 pb), of ^{myc}ATR cell line were replaced in one allele by a selected marker (NEO), generating ^{myc}ATR$^{ΔC+/-}$ cell line; red arrow represents the primers used to amplify those alterations. **B)** PCR analysis from ^{myc}ATR and ^{myc}ATR$^{ΔC+/-}$ cell lines indicating the addition of the tagging in two alleles of both cell lines (e + f) and the deletion of the C-terminal region (g + h) in one allele of ^{myc}ATR$^{ΔC+/-}$ cells, the ATR$^{ΔC+/- mycRPA}$ was use as a control cell line for the C-terminal deletion, showing that the fragment used for the replacement was distinct between the two strategies. **C)** Growth curve of the indicated cell lines in either M199 medium; cells were seeded at ~1x10^5 cells.ml^{-1} at day 0; growth profile was evaluated after cells were kept in culture for 5 days and the cell density was assessed every 24h, and error bars depict standard deviation from two replicate experiments. **D) Left.** ^{myc}ATR and ^{myc}ATR$^{ΔC+/-}$ cells were left untreated (NT) or treated with 0.5mM of Hydroxyurea (HU) for 20 hours, washed and resuspended in fresh media, and time points collected in 0, 1.5, 7 hours after the treatment. Cells were fixed with methanol and the DNA stained with propidium iodate, followed by flow cytometer analysis. **Right.** ^{myc}ATR and ^{myc}ATR$^{ΔC+/-}$ cells were left untreated (NT) or treated with 5mM of Hydroxyurea (HU) for 8 hours, washed and resuspended in fresh media, and time points collected in 0, 2, 8 hours after the treatment. Cells were fixed with methanol and the DNA stained with propidium iodate, followed by flow cytometer analysis. The cell cycle profile was generated using FlowJo software. **E)** Western analysis of whole cell extract of each time points in ^{myc}ATR and ^{myc}ATR$^{ΔC+/-}$ cells after chronic and acute HU treatment; extracts were probe to anti-γH2A antiserum and an anti-EIF1α was used as a loading control (predicted protein sizes are indicated, kDa). The quantification of the signal was measured at ImageJ software; Error bars ± SD; n = 3 experiments. * P<0,05; ** P<0,005; *** P<0,001. Unpaired t-test.

3.2 Chapter 2

In this chapter, we focused our investigations into the roles of ATR relating to *Leishmania* specific biology, particularly the unorthodox dynamics of replication, specifically examining how ATR kinase influences the dynamics around the mapped origin points. Understanding the replication profile with or without full length ATR, in the presence or absence of replication stress may give key insights relating to the phenotypes we have observed in ^{myc}ATR^{ΔC} cells.

3.2.1 ^{myc}ATR^{ΔC} cells show proliferative defects associated with a delayed progression to stationary phase.

^{myc}ATR^{ΔC} cells contain increased quantities of DNA, relative to controls. To further characterize ^{myc}ATR^{ΔC} cells we investigated its growth profile. ^{myc}ATR^{ΔC} and control cells were seeded at a concentration of 3×10^5 cells/ml and the cell density was counted every 24 hours for 6 days. No differences between the control cell lines Cas9T7 cells and ^{myc}ATR were observed, with both cell lines reaching stationary phase at the third day of the growth curve (Figure 19A). In contrast, the ^{myc}ATR^{ΔC} cell line presented a slower growth profile when compared to control cells reaching stationary phase 48 hours later, on the fifth day of the growth curve (Figure 19A).

Based on these findings, we investigated the proportion of cells synthesizing DNA at each point of the growth curve. Cells were subjected to a 30 minutes pulse of the thymidine analogue IdU every 48 hours for seven days and the proportion of cells that incorporated the analogue was quantified using FACS analysis. As shown in Figure 19B, the ^{myc}ATR cell line presented a peak in the proportion of cells synthesizing DNA at day 1 (~29%), which quickly decreased at day 3 (~6%), when those cells reached stationary phase and DNA synthesis was halted as expected. The ATR^{ΔC} cell line also presented a peak of replicating cells at day 1 (~26%). Notably, and different from ^{myc}ATR cells, the majority of these cells seemed to be concentrated at early-S phase of the cell cycle. Moreover, there was a considerable proportion of cells synthesizing DNA up to day 5 of the growth curve (~4%). For instance, at day 3, the proportion of cells replicating their DNA was more than twice (~14%) of that observed for ^{myc}ATR cells (~6%). These results could in part explain the altered dynamics of DNA replication observed during the growth analysis of ^{myc}ATR^{ΔC} cells (Figure 19B).

Figure 19 - ^{myc}ATR^{ΔC} cells show growth defect associated with delayed progression to stationary phase.
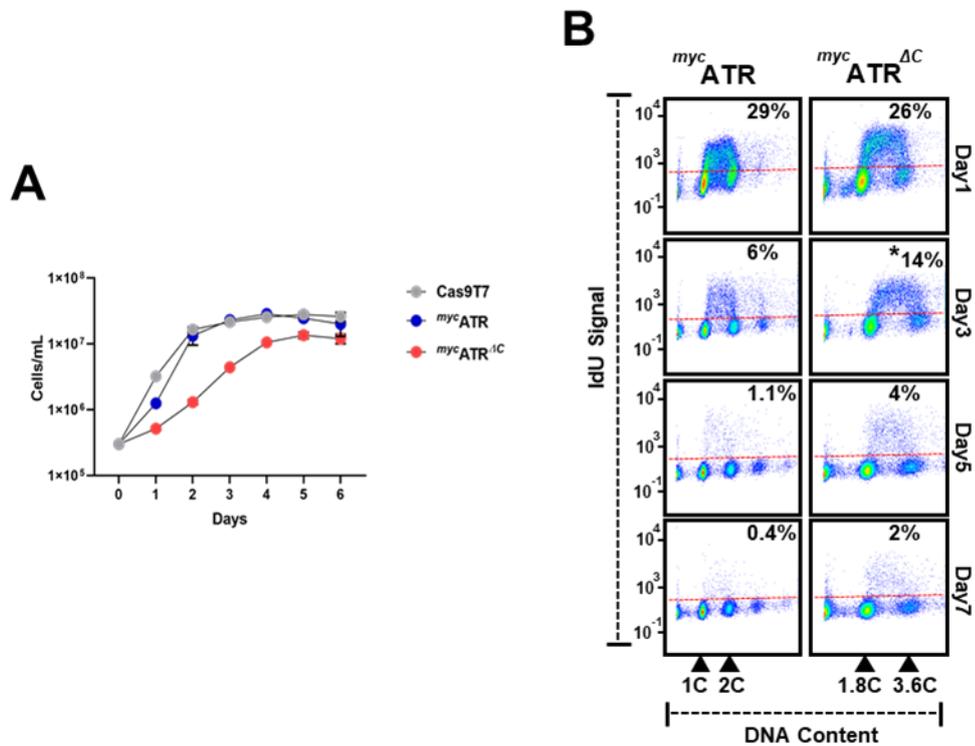
Figure 19 - ^{myc}ATR^{ΔC} cells showed a growth defect which is associated with a longer replication period. **A)** Growth curve of the indicated cell lines in either HOMEM medium; cells were seeded at ~3x10^5 cells.ml^{-1} at day 0; growth profile was evaluated after cells were kept in culture for 6 days and the cell density was assessed every 24h, and error bars depict standard deviation from two replicate experiments. **B)** Representative pseudocolor plots from a flow cytometry analysis to detect DNA synthesis in the indicated cell lines; ^{myc}ATRand ^{myc}ATR^{ΔC} cells were seeded at ~5x10^5 cells.ml^{-1} at day 0; at the indicated time points an aliquot of each cell line was incubated with IdU for 30 min and IdU detected under denaturing conditions; 30,000 cells were analyzed per sample; 1C and 2C indicate single and double DNA content for ^{myc}ATR cells, respectively; and 1.8C and 3.6C indicate single and double DNA content for ^{myc}ATR^{ΔC} cells, respectively; dashed red lines indicate the threshold used to discriminate negative from IdU-positive events; inset numbers indicate total percentage of IdU-positive events on the whole population. n=2 experiment, * P < 0,05.

3.2.2 Loss of *L. major* ATR C-terminal does not affect the replication initiation program.

The study published by Marques (2015) using Marker Frequency Analysis coupled with deep sequencing (MFA-seq) showed that *Leishmania* cells have a distinct DNA replication program where sequences enriched from cell sorted (FACS) in S phase cells relative to G2 cells have a single MFA-seq peak per chromosome, suggesting a single origin per molecule which diverge from the multiple origins mapped in the chromosomes of *Trypanosoma brucei* and *Trypanosoma cruzi,* other kinetoplastid members (de Araujo et al., 2020; Tiengwe et al., 2012). This unusual replication scenario is shown to be insufficient to complete the replication of larger chromosomes, for example (Marques et al., 2015). Later on, a further characterization of the *Leishmania* replication program, this time, comparing the MFA-seq signal enrichment of *L. major* cells in exponential growth (replicating) with cells in stationary phase (non-replicating) showed that the replication program emerges from two sites: the consistent single region (Marques et al. 2015) in early S; and also from the sub-telomers in late S and G1 phases (Damasceno et al., 2020), which could explain how *Leishmania* complete the genome duplication. In this work, we observed that ^{myc}ATR^{ΔC} cells showed a growth defect which is associated with a delay in reaching the stationary phase. So, using the MFA-seq protocol described in Damasceno (2020), we set out to investigate the effect of ATR C-terminal deletion on the *L. major* DNA replication program.

To map the origins of replication, the MFA-seq was performed (two replicates) by calculating the sequences enrichment obtained from the relative amount of DNA in ^{myc}ATR and ^{myc}ATR^{ΔC} cells collected at day 1 and day 3 of the growth curve (exponentially growing) versus Cas9T7 cells in stationary phase (non-replicating) which would act as our normalisation control for the MFA-seq(Figure 20A and Figure sup.9.2). To have clear detection of the MFA-seq signal, the data was analysed by converting the ratio values, which is the average values resulted from the ratio between exponential/stationary alignment to a given reference base, in Z score, that is normalised to the ratio values by how many standard deviations (score) are above or below the mean (score) in each chromosome (see in Methods); this step is important in this case of an aneuploid genome like *Leishmania*. The meta-analysis comparing MFA-seq signal profile across all chromosomes was organized in clusters based on

the sequence enrichment distribution along each chromosome and the data showed the majority of the ^{myc}ATR cells are replicating the internal region (single peak in each cluster) on each chromosome at day 1, while, at day 3 when the cells are close to the stationary phase, the replication concentrates in the sub-telomeric regions in all three clusters (Figure 20B). These findings are in agreement with prior findings (Damasceno et al., 2020). The meta-analyses of ^{myc}ATR$^{\Delta C}$ cells, at day 1 and day 3, showed that the MFA-seq signal profile in each cluster is consistent with the ^{myc}ATR cells in the first day of exponential growing (Figure 20B). Thus, these data indicate that the loss of the ATR C-terminal may not affect the activation of early S origins.

The chromatin composition and organization seem to be very important to drive the replication initiation in *Leishmania* and *T. brucei* where all the origin's location correspond with switch strand regions (SSRs) (Marques et al., 2015; Tiengwe et al., 2012). Also, it has been reported that the internal replication region in early S location overlaps with the concomitant position of kinetoplastid kinetochore protein 1 (KKT1), the transcription initiation marker acetylated histone H3 (AcH3) and a transcription termination marker ß-D-glucosyl-hydroxymethylur-acil (base J), while the sub-telomeric replication just with AcH3 (Damasceno et al., 2020). Next, we look to the MFA-seq signal around those regions and the meta-analysis of all chromosomes, in ^{myc}ATR and ^{myc}ATR$^{\Delta C}$ cells, showed, in agreement with the literature, a strong correlation between the single MFA-seq signal on each chromosome and the location of KKT1, AcH3 and base J, during the exponential growth (day1); the amplitude of the signal, which represents the proportion of cells that have an enriched sequences mapped at these regions, diminished as the cells progress to stationary phase (day3) (Figure 20C). We also looked to the MFA-seq of the transcription-related markers (AcH3 and base J) in SSRs where they are found individually or in association or in SSRs where none of those makers are present (Non Origin). The meta-analysis of the SSRs-Non Origin, AcH3/base J, and AcH3 located at the internal regions of the chromosomes showed peaks with small amplitudes which indicates a low replication activity around these locations in both ^{myc}ATR and ^{myc}ATR$^{\Delta C}$ cells during exponential growth (Figure 20C). However, in both cell lines, a sustainable MFA-seq signal was found around the AcH3 location at the chromosomes ends, during exponential growth and it increases when the cells are reaching stationary phase (Figure 20C). Altogether, the results demonstrated that the replication program previously reported in

Leishmania (Damasceno et al., 2020; Marques et al., 2015) is preserved and any extra replication region is activated in ^{myc}ATR$^{\Delta C}$ cells.

Figure 20 - Loss of *L. major* ATR C-terminal does not affect the replication initiation program.
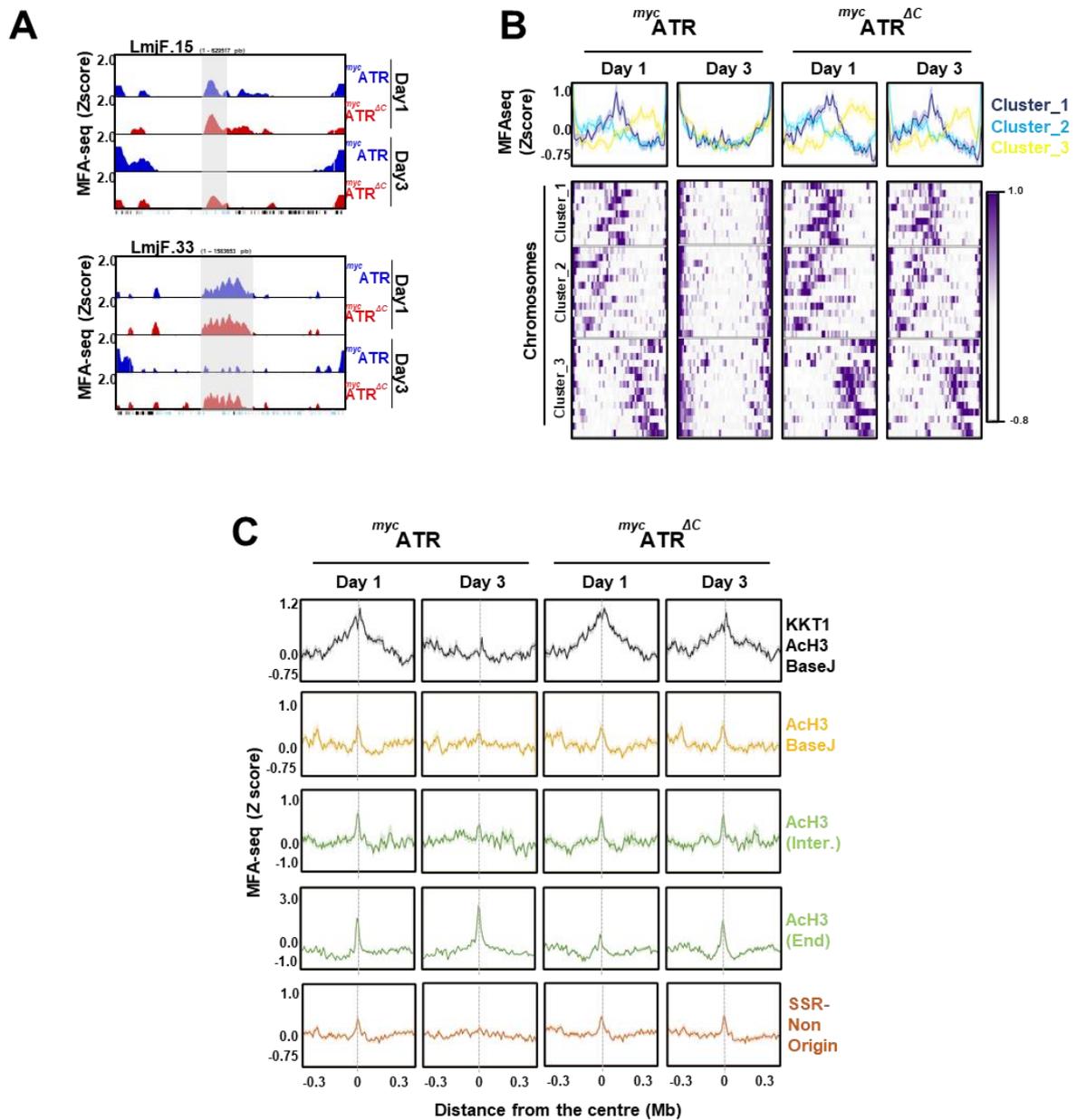
Figure 20 - Loss of L. major ATR C-terminal does not affect the replication initiation program. A) Graphs show the distribution of sites of DNA replication initiation across two complete chromosomes in the indicated box, in each case the ^{myc}ATR and ^{myc}ATR^{ΔC} cells were seeded at ~5x105 cells.ml-1 (day 0) and growth in a fresh HOMEM media, the DNA collected at day 1 and day 3 as indicated. MFA-seq signal is represented by Z-scores across the chromosomes and calculated by comparing read depth coverage of a DNA from each time point relative to stationary Cas9T7 cells; the bottom track of each chromosome displays coding sequences, with genes transcribed from right to left in black, and from left to right in light blue, the light gray box indicates the Main Origin region. **B)** MFA-seq profile in ^{myc}ATR and ^{myc}ATR^{ΔC} cells at day 1 and day3; panels at the top represent global MFA-seq signal for the whole genome; colourmaps indicate MFA-seq profile for individual chromosomes (each row). Chromosomes were grouped in three cluster based on k-meaning, using deepTools. **C)** Metaplots of global MFA-seq signal, in ^{myc}ATR and ^{myc}ATR^{ΔC} cells at day 1 and day3, respectively, relative to stationary Cas9T7 ± 0.3 Mb from the centre of regions containing the indicated combination of chromatin marks; light-coloured areas around the lines indicate the standard error of the mean (SEM).

3.2.3 Loss of *L. major* ATR C-terminal is not associated with alternative origin firing during replication stress.

In other Eukaryotes, the ATR kinase is an important regulator of origin firing during DNA replication stress. Next, we decided to understand the dynamics of replication in ^{myc}ATR and ^{myc}ATR^{ΔC} cells treated with or in the absence of chronic HU (8 and 20 hours) or acute HU (2 and 8 hours). As with the item 3.2.2, we first performed a meta-analysis clustering each chromosome based on the sequence enrichment distribution. The MFA-seq signal profiles under HU conditions showed an increase of the peak's amplitude around a single region on each cluster chromosome at 8 hours of chronic treatment in both cells (Figure 21B). After 20 hours, we observed a substantial decrease of the signal amplitude in ^{myc}ATR cells, but not as pronounced in ^{myc}ATR^{ΔC} cells (Figure 21A). A similar effect was observed upon acute HU treatment, a single narrow peak is forming at 2 and 8 hours treatment in each cluster in both cells which is befitting with a synchronisation profile at the G1/S boundary (Figure 21B).

Figure 21 - Loss of *L. major* ATR C-terminal does not affect new origin firing during replication stress.

Figure 21 - Loss of L. major ATR C-terminal seems to not interfere in new origin firing during replication stress. MFA-seq profile in ^{myc}ATR and ^{myc}ATR$^{\Delta C}$ cells untreated or treated with 0,5mM of HU for 8 or 20 hours (Chronic) **(A)**; or treated with 5mM of HU for 2 and 8 hours (Acute) **(B)** relative to stationary Cas9T7; panels at the top represent global MFA-seq signal for the whole genome; colourmaps indicate MFA-seq profile for individual chromosomes (each row). Chromosomes were grouped in three cluster based on k-meaning, using deepTools.

To detail if the changes on the MFA-seq signal amplitude in ^{myc}ATR and ^{myc}ATR$^{\Delta C}$ cells during the different HU treatments is because of an increase in replication in a particular replication region, we next plot the MFA-seq signal correlating with the specific chromatin markers combinations (KKT1, AcH3, base J, and SSRs Non Origin). The meta-analysis showed that not only around the main internal replication region (KKT1, AcH3 and base J overlap signal), but also around mapped regions with overlap of AcH3 and base J, AcH3 in internal regions and SSRs-Non Origin, the MFA-seq profile in both cell lines showed the same pattern as described at the Figure 21 A and B, an increase in the peak's amplitude at 8 hours followed by a decrease at 20 hours, under chronic HU treatment (Figure 22A); and the formation of a narrow peak during the cell synchronization under acute HU treatment (Figure 22B). Notice, that, as explained at Item 3.2.2, the amplitude of the MFA-seq signal in each time point of the different treatment reflects the proportion of cells in S phase observed at Figure 15.. The data suggest that; first, Leishmania can use other regions related with transcription units (SSRs) as replication sites, but not as frequent as the main internal region (KKT1, AcH3 and base J overlap); second, the deletion of the C-terminal region from *L. major* ATR does not have a different replication program or cause firing of new origins across internal regions of the chromosomes under replication stress in comparison with the ATR full length cells. However, upon 2 hours of acute treatment, the MFA-seq profile of ^{myc}ATR$^{\Delta C}$ cells are comparable with that of control cells subjected to 8 hours of acute treatment (Figure 22B). This could suggest that the ^{myc}ATR$^{\Delta C}$ cells are more sensitive to acute replication stress.

da Silva, Gabriel Lamak Almeida   *Results*

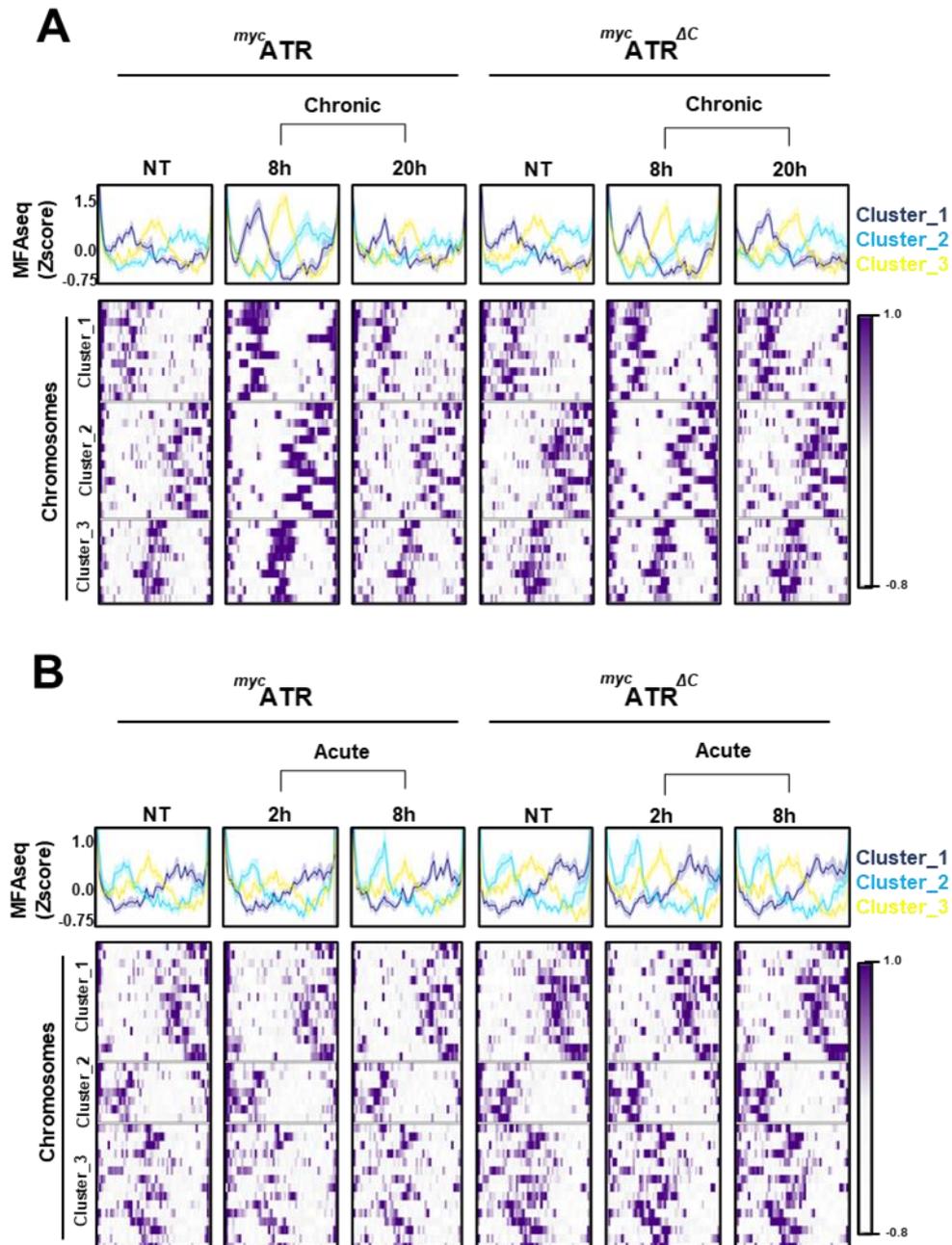**Figure 22 - Loss of *L. major* ATR C-terminal seems to not interfere in new origin firing during replication stress.**
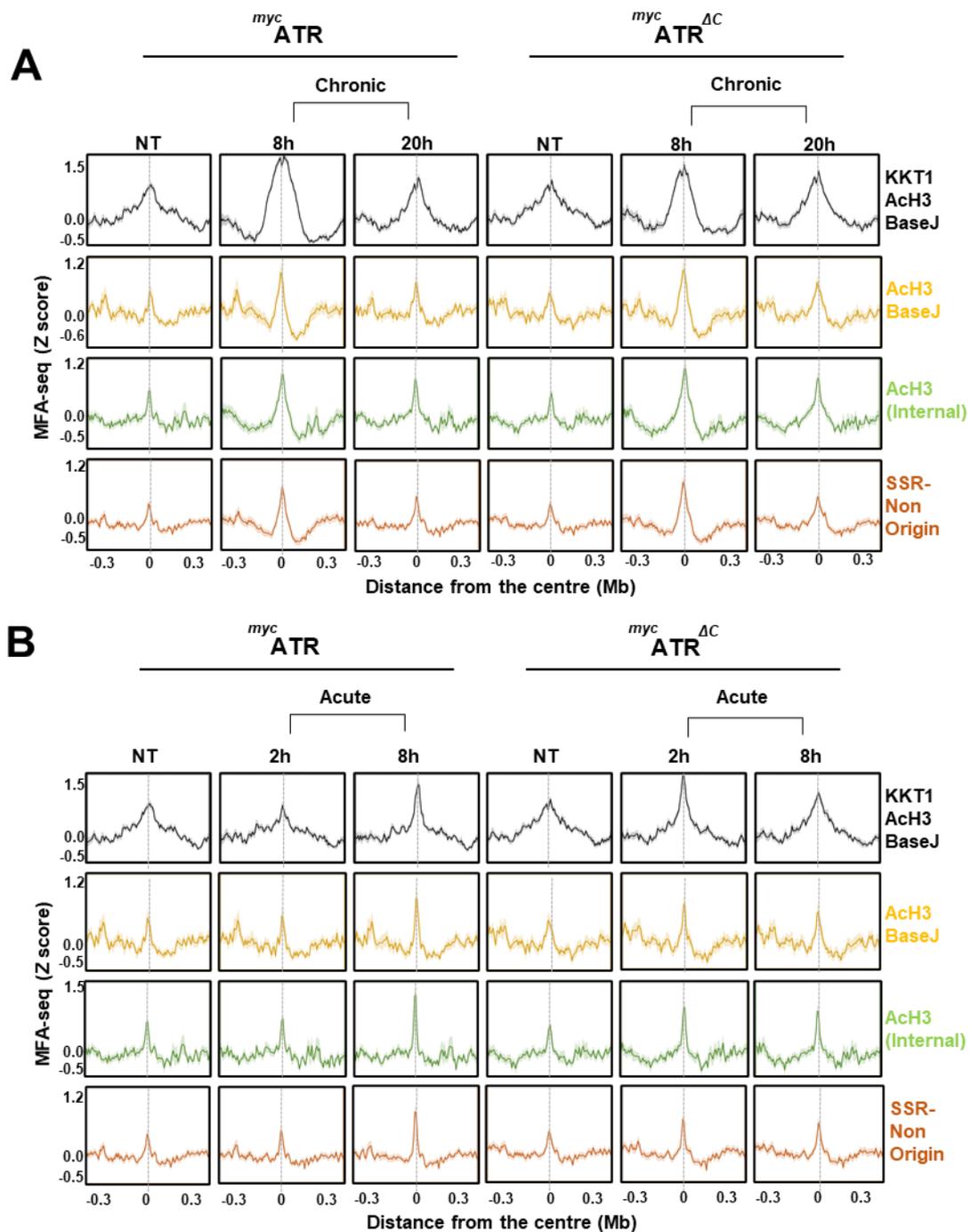
Figure 22 - Loss of L. major ATR C-terminal seems to not interfere in new origin firing during replication stress. Metaplots of global MFA-seq signal, in ^{myc}ATR and ^{myc}ATR^{ΔC} cells untreated or treated with 0,5mM of HU for 8 or 20 hours (Chronic) (A); or treated with 5mM of HU for 2 and 8 hours (Acute) (B) relative to stationary Cas9T7 ± 0.3 Mb from the center of regions containing the indicated combination of chromatin marks; light-colored areas around the lines indicate the standard error of the mean (SEM).

3.2.4 Loss of *L. major* C-terminal affects the sub-telomeric replication in large chromosomes.

The analysis of MFA-seq normalized reads along each chromosome revealed an absence of signal at the chromosome end farthest from the main origin in a few large chromosomes (24, 25, 29, 36) in ^{myc}ATR$^{\Delta C}$ cells (Figure 23A). Also, the signal enrichment at these regions in ^{myc}ATR cells correlates with the mapped signal of AcH3 (Figure 23A) at the very end of the chromosome. The depletion of MFA-seq signals were not due to chromosome deletion at these regions since the read coverage in ^{myc}ATR$^{\Delta C}$ cells was similar to that found in ^{myc}ATR cells (Figure 23A, Bottom). Therefore, the absence of MFA-seq signal observed suggests a constitutive replication defect across these regions.

The meta-analysis of the MFA-seq signals across all chromosome ends in ^{myc}ATR and ^{myc}ATR$^{\Delta C}$ cells shows that the decrease of MFA-seq signal in some chromosome ends is extended to all sub-telomeric regions on the mutant cells under 20 hours of chronic treatment and 2 hours of acute treatment compared with the MFA-seq signal in ^{myc}ATR cells (Figure 23B). We performed a meta-analysis correlating the MFA-seq signal and the AcH3 localised within the sub-telomeric region (Figure 23C). The results showed that the decrease of replication at the very end of the chromosomes overlaps with the position of this chromatin marker and reinforces the hypothesis that the presence of Ach3 at the sub-telomere is used to drive the initiation of replication.

Figure 23 - Loss of *L. major* C-terminal affects sub-telomeric replication in large chromosomes.
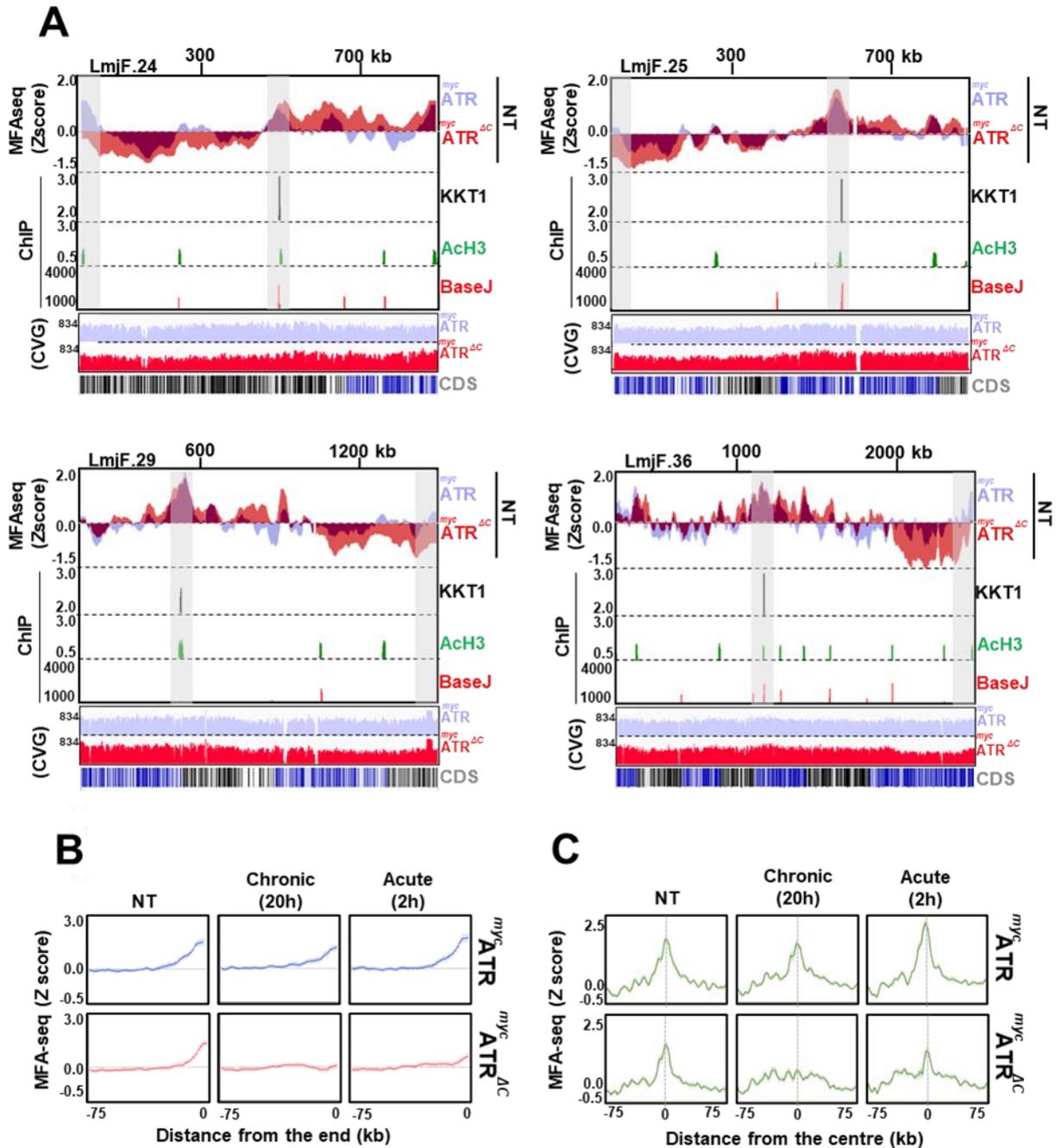
Figure 23 - Loss of L. major C-terminal affects the sub-telomeric replication in large chromosomes. A) Representative region of Chromosome 24, 25, 29 and 36 of ^{myc}ATR and ^{myc}ATR$^{\Delta C}$ cells untreated showing MFA-seq signal compared with the position of AcH3, Base J and KKT1 enriched sites (Main Origin) and in AcH3 position at one telomeric region; the bottom track indicates annotated CDSs (right to left in blue, and from left to right in black); coverage tracks (CVG/Bottom) were generated with deepTools, using bamCoverage tool; the track in each chromosome bottom displays coding sequences. **C)** Metaplots of MFA-seq signal across 75 kb of sequence from all chromosomes ends in ^{myc}ATR and ^{myc}ATR$^{\Delta C}$ cells untreated or treat with 0,5mM of HU for 20 hours (Chronic) or with 5mM of HU for 2h (Acute). **D)** Metaplots of MFA-seq signal around 75 kb of sequence from all AcH3 locate at telomeric regions in ^{myc}ATR and ^{myc}ATR$^{\Delta C}$ cells untreated or treated with 0,5mM of HU for 20 hours (Chronic) or with 5mM of HU for 2h (Acute).

3.2.5 ^{myc}ATR$^{\Delta C}$ cells do not have an alteration in copy number variation.

Considering the typical genome plasticity of *Leishmania* and the observed increase in DNA content in ^{myc}ATR$^{\Delta C}$ cells (Figure 14), we asked if the increase in DNA content correlated with alterations to the number of chromosomes. To do this we determined the ploidy of each chromosome in ^{myc}ATR and ^{myc}ATR$^{\Delta C}$ cells, treated or not, correlating the coverage of each sequence with the reference genome from the database (Tritrypdb.org). Although the cell cycle progression analysis showed an increase of ~50% of DNA content in ^{myc}ATR$^{\Delta C}$ cells; the results from the copy number variation analysis did not reveal any significant increase in chromosome number of ^{myc}ATR$^{\Delta C}$ cells that could explain their higher DNA content (Figure 24). In fact, we noticed a decrease in one copy of chromosome 31 from the five copies observed in ^{myc}ATR cells (Figure 24). A plausible explanation for not detecting the massive increase in ^{myc}ATR$^{\Delta C}$ cells DNA content in this analysis is the possibility of the extra amount of DNA to be evenly distributed along the genome. Such alteration would escape the detection since this analysis used a whole genome sequencing approach. An appropriate spike-in control would have to be used to verify the possibility of an evenly distributed increase in DNA content.

Figure 24 – The copy number variation analysis does not explain the increase of DNA content in ^{myc}ATR$^{\Delta C}$ cells.
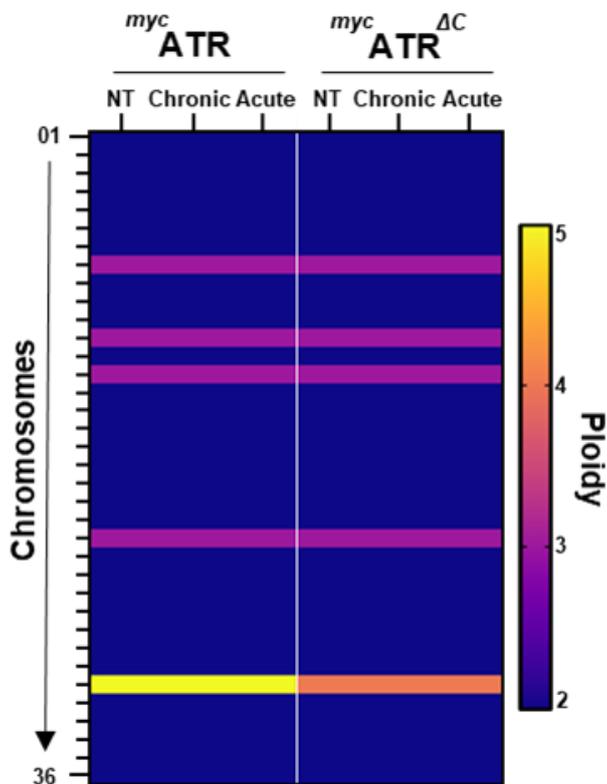
Figure 24 - ^{myc}ATR$^{\Delta C}$ cells do not have an alteration in copy number variation. Heatmap showing the ploidy of each chromosome of ^{myc}ATR and ^{myc}ATR$^{\Delta C}$ cells untreated or treated with 0.5mM of HU for 20 hours (Chronic) or 5mM of HU for 8 hours (Acute). The values were generated comparing each chromosome coverage which were generated with deepTools, using bamCoverage tool.

3.2.6 Acute HU treatment correlates with enhanced mutagenesis in ^{myc}ATR$^{\Delta C}$ cells.

Next, in order to determine if the loss of ATR C-terminal causes genome instability, we used short-read Illumina sequencing of ^{myc}ATR and ^{myc}ATR$^{\Delta C}$ cells DNA content throughout their growth curve, or comparing untreated and chronic or acute HU-treated cells. We measured the number of single nucleotide polymorphisms (SNPs), insertions (INS) and deletions (DEL). Sequences were aligned and compared with reference genome (see in Methods). To use SNPs and INDELs as indicators of enhanced variability, SNPs, INS, DEL that were common to the two time points compared (either day 1/day 3; non-treated/chronic HU-treated; or non-treated/acute HU-treated) were excluded (Figure 25A). The total number of each marker in each chromosome was then normalized by the chromosome size.

The result showed no significant change in the number of SNPs under any condition in both cell lines. While ^{myc}ATR cells showed an increase in insertions and deletions after three days of growth and an increase in deletions upon chronic treatment (Figure 25B) the ^{myc}ATR$^{\Delta C}$ cells showed a substantial increase in the insertion and deletions mostly upon the acute treatment. It is noteworthy that the accumulation of these variability markers seems to be concentrated in small and median size chromosomes in both cell lines. (Figure 25B).

Figure 25 - Acute HU treatment correlates with enhanced mutagenesis in ^{myc}ATR^{ΔC} cells.

Figure 25 - Acute HU treatment correlates with enhanced mutagenesis in ^{myc}ATR^{AC} cells. A) SNPs, Insertions (INS) and Deletions (DEL) relative to the reference genome were identified; events common to the two conditions were discarded; events exclusively found in Day3, 20h and 8h were considered to the following analysis. **B)** Heatmap representing the density of SNPs, INSs, DELs (/Kb) in Day3, 20h, 8h of each condition in (A); numbers at the top of each row indicate the Pearson correlation between SNPs, INSs, DELs and chromosome size; where correlation is significant, it is indicated by * P<0,05; ** P<0,005; *** P<0,001.

<u>Discussion</u>

4. Discussion

Safeguarding replication is vital to guarantee genome fidelity. For Leishmania, whose genome is unusually plastic, the study of DNA damage response seems to be vital to understand the dichotomy question of how the balance between genome integrity and variability works. Our findings uncover many key aspects of ATR kinase functions in Leishmania major, collecting evidence of its function under replication stress or lack of. The use of different cell lines bearing an altered ATR locus (^{myc}ATR; ATRΔC$^{+/-mycRPA}$ and ^{myc}ATR^{ΔC}) revealed different aspects of ATR function in *L. major*: (i) ATR localization is not restricted to the nuclear compartment and is also expressed in the kinetoplast and in a third location near or at the flagellar pocket; (ii) Proper ATR location is dependent on its C-terminal region; (iii) At the kinetoplast, ATR seems to be involved in the DNA replication process; (iv) Deletion of ATR C-terminal substantially affects DNA metabolism causing increase in DNA content, accumulation of DNA damage marker and increase in the proportion of aberrant cells; (v) Upon replication stress, the proper replication of sub-telomeric sequences is dependent on an intact ATR, and (vi) The dynamics of cell cycle progression in cells exposed to replication stress is also strongly dependent on a full-length ATR. Altogether, the results showed that ATR kinase is important to control the cell cycle and the maintenance of replication during replication stress in *L. major*.

ATR kinase is the pinnacle of DNA damage response during replication stress and its activity connects with many pathways, as in its overlap with ATM activity during the exonucleolytic resection of DBSs during S-G2 phase (Jazayeri et al., 2006). The importance of ATR to the many housekeeping cell processes is suggested by the failed attempts to generate a viable ATR knockout mouse (de Klein et al., 2000) and, in humans, mutation in ATR is associated with Seckel syndrome (Lavin et al., 2008). The studies that describe ATR functions in humans and yeast cells use approaches based on ATR inhibitors (Ramkumar et al., 2021), localized motif mutation or disruption of the RNA processing (Murga et al., 2009; Tannous et al., 2021), or knockdown using the RNAi system (Postigo et al., 2017). There are few studies discussing ATR functions in trypanosomatids, with most of them exploring the fact that T. brucei has a functional RNAi system (Black et al., 2020; Marin et al., 2020 and Zhou et al., 2019). Upon the

adaptation of the Cas9 editing system for Leishmania (Beneke et al., 2017), the successful deletion of ATR in Leishmania mexicana has been described (Baker et al., 2021) However, its function during DNA replication stress response has not yet been explored in other Leishmania species. The strategy to delete the predicted functional domains of the *L. major* ATR kinase allowed us to generate stable mutants, partially or with a complete absence of those domains, allowing the study of the ATR function on this parasite.

ATR participates in the activation of checkpoints in the presence of DNA damage, delaying the cell cycle in S or G2 phases (Eykelenboom et al., 2013; Ruiz et al., 2016). Our findings showed the ^{myc}ATR$^{\Delta C+/-}$ cells have a similar cell cycle profile in comparison with the control cell after HU treatment, but in ^{myc}ATR$^{\Delta C}$ cells, under the same conditions we observed an increase of sub-G1 DNA content population suggesting that ATR is required at the end of replication and the entrance into G2/M phase or even during mitosis. In unperturbed human cells, inhibition of ATR kinase during S phase causes premature mitosis by early phosphorylation of FOXM1 which drives premature cyclin B expression (Saldivar et al., 2018). The same study describes an ATR-related checkpoint where ETAA1 is the primary ATR activator during S/G2 transition (Saldivar et al., 2018). Knockdown of ATR kinase in T. brucei has been linked with defects during checkpoints activation (Marín et al., 2020), and it might be the cause of sub-G1 population accumulation in ^{myc}ATR$^{\Delta C}$ cells in *L. major*; however, the mechanisms of checkpoint activation could be different in Trypanosomatids, since they seem to lack ETAA1, for example. ATR kinase is also linked to have an important role during mitosis by activating Aurora B through CHK1, in response to R-loop formation at the centromeres, which prevent chromosome mis-segregation (Kabeche et al., 2018). So far, our results are insufficient to distinguish which checkpoint(s) are activated by ATR kinase in *L. major*, however they do strongly suggest that ATR kinase prevents cells to undergo mitosis before the duplication of the DNA is completed, since the complete loss of the kinase C-terminal generates daughter cells with reduced DNA content when compared to their parental cell.

The control of the cell cycle by ATR could be more extensive in *L. major* since the kinase is located at the nucleus, kinetoplast and a third location suggestive of the flagellar pocket, with all three locations dependent on the C-terminal region of the kinase. Trypanosomatids cell cycle is timed and ordered being characterized by three

sub-cycles that involve the duplication and segregation of the nucleus, kinetoplast and cytoskeletal structures (Wheeler, Gull, Sunter, 2019). In *Leishmania mexicana* the duplication of kinetoplast and nucleus DNA is near-synchronous and followed by the elongation of the new flagellum (Wheeler, Gluenz, Gull, 2011). We showed that ATR is equally distributed among its cellular compartments during cell division of ^{myc}ATR cells, and the deletion of its C-terminal region leads to an increase of aberrant cells: 0N1K (zoids), 1N0K and 2N1K; irrespective of hydroxyurea treatment. Several studies in *T. brucei*, showed that failure to progress through S phase or mitosis does not prevent kinetoplast and flagellum duplication or cytokinesis, generating zoids or 1N1K cells; also, failure to duplicate kinetoplast DNA does not prevent flagellum duplication or cytokinesis, generating zoids; whereas, disruption of basal body (flagellar structure) duplication does not prevent kinetoplast S phase or mitosis, with the outcome being cells with 2N1K (Hammarton et al., 2003; Hu et al., 2015; Tiengwe, Marques, McCulloch, 2014; Wang, Paul, 2001). Important findings that also support the hypothesis of failures during the cell cycle, are that ^{myc}ATR$^{\Delta C}$ cells have an increased DNA content and their culture takes longer to reach the stationary phase. Altogether we can formulate at least two explanations: the cytokinesis occurs while the ^{myc}ATR$^{\Delta C}$ cells are still replicating their DNA and/or the absence of ATR activity at the different compartments disrupt the cell cycle organization.

Our findings raise questions about *L. major* ATR function at its different sub-cellular locations. Indeed, little is known about ATR functions outside of the nucleus. In human cells, UV radiation causes ATR conformation changes from trans-isomer (Nuclear) to cis-isomer (Mitochondrial) through the phosphorylation of Pin1 (peptidylprolyl cis/trans isomerase NIMA-interacting 1), an enzyme that alters proline peptidyl bonds causing alteration in protein structure. This modification blocks the recruitment of Bax (proapoptotic Bcl2–associated X) to mitochondria and prevents apoptosis. This ATR action does not depend on its association to ATRIP (Hilton et al., 2015). Our findings showed that ATR locates at the kinetoplast, a dense structure composed by kinetoplast DNA (kDNA) within the mitochondrial matrix (de Souza et al., 2010). Given the peculiar kDNA organization, composed of thousands of topologically interlocked DNA circles, we speculate that, similar to the canonical function at the nucleus, ATR possibly orchestrates the faithful replication of each of the minicircles and maxicircles during kDNA replication. Supporting this hypothesis, our data showed

that loss of one ATR C-terminal copy leads to the accumulation of ssDNA in the kinetoplast. Also, the kinetoplast ATR location duplicates during kDNA replication. Considering the machinery that replicates the kDNA minicircles and maxicircles (Liu et al., 2005), our findings are consistent with a functional involvement of ATR in the DNA replication that takes place in the kinetoplast. Since *T. brucei* or *L. mexicana* share with *L. major* the same mitochondrial organization, it is unclear why ATR was not observed in these kinetoplastids (Black et al., 2020 and Baker et al., 2021). To this date, there is no report of involvement of ATR, or any other DDR kinase, in the flagellum biogenesis or cytokinesis. However, some kinases have different locations depending on the cell cycle stage in *L. mexicana* (Baker et al., 2021). In any case, a possible interaction between ATR and the kinases that regulate the flagellum biosynthesis and cytokinesis cannot be discarded.

Leishmania has a distinct DNA replication program in which the genome duplication starts in early S phase, from a single region in every chromosome, and a new replication initiation site starts in late S phase, at sub-telomeric regions, and continues towards G2/M and G1 phase (Damasceno et al., 2020; Marques et al., 2015). In addition, DNA replication initiation in *Leishmania* correlates with transcription initiation and termination markers (Marques et al., 2015). In unperturbed human cells, ATR is essential to modulate the DNA replication program by balancing the firing of DNA replication origins. This is achieved through the stabilization of Rif1-directed PP1 interaction which regulates and counteracts the activity of DDK dependent phosphorylation on MCMs. (Alver et al., 2017; Moiseeva et al., 2019). Our findings suggest a limited role for ATR in the DNA replication program of *L. major*. Upon deletion of the ATR C-terminal domain, the same frequent replication sites were found suggesting that the deleted ATR domains do not cause significant firing of alternative or dormant origins of replication. Considering that the molecular mechanism of DNA synthesis initiation and elongation in trypanosomatids are not fully understood, the possibility of distinct pathways controlling DNA replication in *Leishmania* cannot be ruled out. The participation of alternative regulation events could have pivotal influence in the replication process of an organism that relies on DNA amplification to control and allow dosage of gene expression (Laffitte et al., 2016a). It is noteworthy that mutated ATR cells have an increase in DNA content that is not restricted to specific

chromosomes, but distributes across the entire genome, indicating that the capacity to cope with different gene dosage is still intact upon ATR inactivation.

Under replication stress, the human ATR version supresses origin firing globally and allows local dormant origin firing to the stressed site, a crucial function to support the completion of DNA replication. (Ge and Blow, 2010). In *Leishmania*, SNS-seq (Small Nascent strand-seq) can detect multiple initiation sites across the genome (Lombraña et al., 2016). On the other hand, mapping of MFA-seq signals mostly detected a single early-S main origin besides the replication of sub-telomeres (Damasceno et al., 2020). In this work we analyzed MFA-seq signals in cells exposed to 0.5mM hydroxyurea. This treatment slows down replication but is not enough to hold cells at the G1/S transition. Thus, we were able to detect replication sites used under replication stress and correlated them with transcription initiation and termination markers across the genome. This revealed that Leishmania can use regions involved in transcription control as replication sites. These findings support the hypothesis that SSRs associated with or without markers of transcription initiation and/or termination may also participate in the replication process (Lombraña et al., 2016; Marques et al., 2015). Those background replication sites could help the replication progression during replication stress and their activity is reduced after the resumption of normal replication rate. It is noteworthy that, ATR is not necessary for the activation of the background origins because the same MFA-seq signal profile was found both in wildtype and in mutant *L. major* ATR.

This work provided evidence that ATR influences the late S phase replication at the sub-telomeres in *L. major*. Two hypotheses have been proposed to explain the emergence of sub-telomeric replication: an extension of replication forks from the chromosome-central initiation region and/or a replication initiation event that could be dependent on DNA repair processes (Damasceno et al., 2020). In mammalian and yeast cells, ATR kinase regulates not only the pool of dNTPs (Buisson et al. 2015 and Zhao et al. 1998), but also, RPA exhaustion (Toledo et al., 2013) and phosphorylation of HR factors (Petermann et al., 2010) upon replication stress. We observed the absence of MFA-seq signal around the sub-telomeric region in some large chromosomes in the mutated ATR cell line. This could be a result of deregulation in replication substrates and/or the factors that prevent fork collapse which compromise the late stages of the S phase, including the activation of the late replication sites

(Buisson et al., 2015; D'angiolella et al., 2012; Técher et al., 2016). Another explanation could be related with a non-canonical replication initiation at *Leishmania* sub-telomeres involving a possible clash between transcription and replication machinery. In fact, we can observe the presence of transcription initiation and termination elements flaking the MFA-seq negative signal region. However, it is not clear why this was observed specifically in some large chromosomes. This might be related, as in other eukaryotes, to a combination of several factors which include how chromosomes are organized within the nucleus, the timing of origin, the distribution of origins along the chromosome and the mechanism of replication activation (Koren et al., 2014; Schübeler et al., 2002).

Indeed, a strong indication of the hypothesis of *Leishmania* sub-telomeric regions' replication initiation is linked with DNA damage response is the fact that the MFA-seq signal from chromosome ends is globally affected in mutated ATR cells under chronic and acute replication stress, and this decrease in signal correlates with the AcH3 sites. The replication at chromosome ends in *Leishmania* is also affected in cells deficient of RAD9 and HUS1, two components of the 9-1-1 complex (Damasceno et al., 2020). Although the components of DNA replication initiation have not been mapped in *Leishmania*, the replication at the single replication region coincides with the position of centromere marker (KKT), which is also positionally conserved in *T. brucei*, and with ORC localization in early S phase (Lombraña et al., 2016; Marques et al., 2015; Tiengwe et al., 2012). On the other hand, the replication at the sub-telomeric region only correlates with AcH3 in *Leishmania* (Damasceno et al., 2020). Also, telomeres of Leishmania are relevant regions for genome maintenance (Lovejoy et al., 2012), and are particularly susceptible to copy number variation (Bussotti et al., 2018). These features could promote the activation of DNA repair activity at these locations, and through an unorthodox pathway, drive the replication machinery.

In mammalian and yeast cells, ATR has also been described at participating in the restart of stalled forks that may be generated in response to DNA injuries (Zeman and Crimpich, 2014). Although we have not probed this function of ATR in Leishmania, cells with mutated ATR are unable to recover the normal cell cycle progression upon chronic or acute replication under stress. Instead, these cells accumulate yH2A and aberrant DNA content. In fact, the ^{myc}ATR$^{\Delta C}$ cells seem to be unable to hold the G1/S synchronization under acute hydroxyurea treatment accumulating insertion and

deletions, that might add obstacles to replication resumption after hydroxyurea removal. Also, we cannot rule out that ATR might be directly involved in the fork restart. Many processes have been proposed to promote fork restart. These involve repriming ahead stalled fork by PrimPol; a DNA damage tolerance pathway that allows continuous DNA synthesis by bypassing DNA lesions; fork reversal processes and template switching using the undamaged chromatid (Buisson et al., 2017; Nelseen et al., 2015). In this context, it has been reported that ATR phosphorylates two of the trans-lesion polymerases, REV1 and Pol η (Göhler et al., 2011); also, the ATR kinase substrates are important to the recruitment of RAD51 which acts during template switching, fork reversal and homologous recombination (Murphy et al. 2014; Vassin et al., 2009). Despite the fact that the ATR mutated cells did not show a differential number of single nucleotide polymorphism (SNPs) in comparing the cells treated with hydroxyurea and non-treated cells, this does not disprove ATR involvement in *Leishmania* genome variability; therefore, we must consider that these variety markers require time to be established among the population of cells.

Overall, this work showed that ATR kinase in *Leishmania major* conserved canonical features related with genome maintenance and integrity, since mutated ATR cells showed an increase of DNA content; cells with abnormal proportion of N/K; accumulation of ssDNA, increase of mutagenic markers and an affected cell cycle profile under or after replicative stress. More important, ATR kinase is closely linked with some of the parasite biology features; been locate at the kinetoplast, preventing ssDNA accumulation during kDNA replication, and flagellar pocket. *Leishmania* has a particular DNA replication program where ATR kinase seems to be involved in the sub-telomeric replication, during late S phase, in regions with the presence of the transcription marker AcH3; supporting the hypothesis DNA damage response components could be involve in the sub-telomeric DNA replication. For the future, it will be important to identify other protein that interact with ATR during the replication stress response; identify the proteins that are upregulated in mutated ATR cells which could indicate compensatory pathways; identify where on the genome ATR might binds to correlate with the MFA-seq findings; and finally, track and sequence the direction and location of the replication forks in mutated ATR cells.

Conclusion

5. Conclusion.

In summary, this work describes a detailed characterization of the ATR kinase in *Leishmania major;* uncovering important aspects of its functions in response to replicative stress and contributing to a better understanding of the parasite biology. The main findings described in this work are:

- *L. major* ATR kinase localizes to three sub-cellular compartments: nucleus, kinetoplast and possibly the flagella pocket. ATR localization is dependent on its C-terminal region.
- At the kinetoplast, ATR seems to be involved in kDNA replication process.
- Loss of one ATR C-terminal copy leads to accumulation of ssDNA at the kDNA and the nucleus, which it enhanced under hydroxyurea treatment.
- Loss of both ATR C-terminal copies leads to increased DNA content, abnormal proportion of nucleus/kinetoplast, and defective cell cycle progression.
- Knockout ATR C-terminal cells not only accumulate the DNA damage marker yH2A and are unable to restore normal cell cycle profile, but also accumulate mutagenic marker; upon hydroxyurea treatment.
- Knockout of ATR C-terminal does not affect the general replication program of the genome. However, some large chromosomes showed a defective replication signal of sub-telomeric sequences.
- Replication sites at chromosome ends overlap with the transcription marker AcH3 and is affected in knockout ATR C-terminal cells exposed to replication stress.

<u>References</u>

6. References.

1. ABRAHAM, Robert T. Mammalian target of rapamycin: immunosuppressive drugs uncover a novel pathway of cytokine receptor signaling. Current opinion in immunology, v. 10, n. 3, p. 330-336, 1998.

2. AHLSKOG, Johanna K. et al. ATM/ATR-mediated phosphorylation of PALB 2 promotes RAD 51 function. EMBO reports, v. 17, n. 5, p. 671-681, 2016.

3. ALVAR, Jorge et al. Leishmaniasis worldwide and global estimates of its incidence. PloS one, v. 7, n. 5, p. e35671, 2012.

4. ALVER, Robert C. et al. Reversal of DDK-mediated MCM phosphorylation by Rif1-PP1 regulates replication initiation and replisome stability independently of ATR/Chk1. Cell reports, v. 18, n. 10, p. 2508-2520, 2017.

5. AMBIT, Audrey et al. Morphological events during the cell cycle of Leishmania major. Eukaryotic cell, v. 10, n. 11, p. 1429-1438, 2011.

6. BAKER, N. et al. Systematic functional analysis of Leishmania protein kinases identifies regulators of differentiation or survival. Nature communications, v. 12, n. 1, p. 1-15, 2021.

7. BALMUS, Gabriel et al. Synthetic lethality between PAXX and XLF in mammalian development. Genes & development, v. 30, n. 19, p. 2152-2157, 2016.

8. BARR, Sharon M. et al. ATR kinase activity regulates the intranuclear translocation of ATR and RPA following ionizing radiation. Current biology, v. 13, n. 12, p. 1047-1051, 2003.

9. BASS, Thomas E. et al. ETAA1 acts at stalled replication forks to maintain genome integrity. Nature cell biology, v. 18, n. 11, p. 1185-1195, 2016.

10. BENEKE, Tom et al. A CRISPR Cas9 high-throughput genome editing toolkit for kinetoplastids. Royal Society open science, v. 4, n. 5, p. 170095, 2017.

11. BEVERLEY, Stephen M. et al. Unstable DNA amplifications in methotrexate resistant Leishmania consist of extrachromosomal circles which relocalize during stabilization. Cell, v. 38, n. 2, p. 431-439, 1984.

12. BLACK, Jennifer Ann et al. Trypanosoma brucei ATR links DNA damage signaling during antigenic variation with regulation of RNA polymerase I-transcribed surface antigens. Cell reports, v. 30, n. 3, p. 836-851. e5, 2020.

13. BLACKFORD, Andrew N.; JACKSON, Stephen P. ATM, ATR, and DNA-PK: the trinity at the heart of the DNA damage response. Molecular cell, v. 66, n. 6, p. 801-817, 2017.

14. BLANKENBERG, Daniel et al. Dissemination of scientific software with Galaxy ToolShed. Genome biology, v. 15, n. 2, p. 1-3, 2014.

15. BOSOTTI, Roberta; ISACCHI, Antonella; SONNHAMMER, Erik LL. FAT: a novel domain in PIK-related kinases. Trends in biochemical sciences, v. 25, n. 5, p. 225-227, 2000.

16. BRITTON, Sébastien; COATES, Julia; JACKSON, Stephen P. A new method for high-resolution imaging of Ku foci to decipher mechanisms of DNA double-strand break repair. Journal of Cell Biology, v. 202, n. 3, p. 579-595, 2013.

17. BUISSON, Rémi et al. Distinct but concerted roles of ATR, DNA-PK, and Chk1 in countering replication stress during S phase. Molecular cell, v. 59, n. 6, p. 1011-1024, 2015.

18. BURKARD, Gabriela Schumann; JUTZI, Pascal; RODITI, Isabel. Genome-wide RNAi screens in bloodstream form trypanosomes identify drug transporters. Molecular and biochemical parasitology, v. 175, n. 1, p. 91-94, 2011.

19. BURZA, Sakib; CROFT, Simon L.; BOELAERT, Marleen. Leishmaniasis–Authors' reply. The Lancet, v. 393, n. 10174, p. 872-873, 2019.

20. BUSSOTTI, Giovanni et al. Leishmania genome dynamics during environmental adaptation reveal strain-specific differences in gene copy number variation, karyotype instability, and telomeric amplification. MBio, v. 9, n. 6, p. e01399-18, 2018.

21. CARNEY, James P. et al. The hMre11/hRad50 protein complex and Nijmegen breakage syndrome: linkage of double-strand break repair to the cellular DNA damage response. Cell, v. 93, n. 3, p. 477-486, 1998.

22. CHANUT, Pauline et al. Coordinated nuclease activities counteract Ku at single-ended DNA double-strand breaks. Nature communications, v. 7, n. 1, p. 1-12, 2016.

23. CICCIA, Alberto; ELLEDGE, Stephen J. The DNA damage response: making it safe to play with knives. Molecular cell, v. 40, n. 2, p. 179-204, 2010.

24. CIMPRICH, Karlene A.; CORTEZ, David. ATR: an essential regulator of genome integrity. Nature reviews Molecular cell biology, v. 9, n. 8, p. 616-627, 2008.

25. CLAYTON, Christine; SHAPIRA, Michal. Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. Molecular and biochemical parasitology, v. 156, n. 2, p. 93-101, 2007.

26. CODERRE, Jeffrey A. et al. Overproduction of a bifunctional thymidylate synthetase-dihydrofolate reductase and DNA amplification in methotrexate-resistant Leishmania tropica. Proceedings of the National Academy of Sciences, v. 80, n. 8, p. 2132-2136, 1983.

27. CORTEZ, David et al. ATR and ATRIP: partners in checkpoint signaling. Science, v. 294, n. 5547, p. 1713-1716, 2001.

28. COUCH, Frank B. et al. ATR phosphorylates SMARCAL1 to prevent replication fork collapse. Genes & development, v. 27, n. 14, p. 1610-1623, 2013.

29. CRUZ, Angela; BEVERLEY, Stephen M. Gene replacement in parasitic protozoa. Nature, v. 348, n. 6297, p. 171-173, 1990.

30. DA SILVA, Raíssa Bernardes et al. Selective human inhibitors of ATR and ATM render Leishmania major promastigotes sensitive to oxidative damage. PloS one, v. 13, n. 9, p. e0205033, 2018.

31. DA SILVA, Raíssa Bernardes et al. Specific Human ATR and ATM Inhibitors Modulate Single Strand DNA Formation in Leishmania major Exposed to Oxidative Agent. Frontiers in cellular and infection microbiology, p. 1302, 2022.

32. DAMASCENO, Jeziel D. et al. Conditional genome engineering reveals canonical and divergent roles for the Hus1 component of the 9–1–1 complex in the maintenance of the plastic genome of Leishmania. Nucleic Acids Research, v. 46, n. 22, p. 11835-11846, 2018.

33. DAMASCENO, Jeziel D. et al. Functional compartmentalization of Rad9 and Hus1 reveals diverse assembly of the 9-1-1 complex components during the DNA damage response in Leishmania. Molecular microbiology, v. 101, n. 6, p. 1054-1068, 2016.

34. DAMASCENO, Jeziel D.; NUNES, Vinicius S.; TOSI, Luiz RO. LmHus 1 is required for the DNA damage response in L eishmania major and forms a complex with an unusual Rad 9 homologue. Molecular microbiology, v. 90, n. 5, p. 1074-1087, 2013.

35. DAMASCENO, Jeziel Dener et al. Genome duplication in Leishmania major relies on persistent subtelomeric DNA replication. Elife, v. 9, p. e58030, 2020.

36. D'ANGIOLELLA, Vincenzo et al. Cyclin F-mediated degradation of ribonucleotide reductase M2 controls genome integrity and DNA repair. Cell, v. 149, n. 5, p. 1023-1034, 2012.

37. DE ARAUJO, Christiane Bezerra et al. Replication origin location might contribute to genetic variability in Trypanosoma cruzi. BMC genomics, v. 21, n. 1, p. 1-16, 2020.

38. DE KLEIN, Annelies et al. Targeted disruption of the cell-cycle checkpoint gene ATR leads to early embryonic lethality in mice. Current biology, v. 10, n. 8, p. 479-482, 2000.

39. DE OLIVEIRA, Francisco Meirelles Bastos et al. Phosphoproteomics reveals distinct modes of Mec1/ATR signaling during DNA replication. Molecular cell, v. 57, n. 6, p. 1124-1132, 2015.

40. DE OLIVEIRA, Lara Camila; MOREIRA, Neide Martins. Epidemiological aspects of visceral leishmaniasis in Brazil and in international border regions. Research, Society and Development, v. 10, n. 12, p. e549101220684-e549101220684, 2021.

41. DE SOUZA, Wanderley; DE CARVALHO, Tecia Maria Ulisses; BARRIAS, Emile Santos. Ultrastructure of Trypanosoma cruzi and its interaction with host cells. In: American Trypanosomiasis. Elsevier, 2010. p. 393-432.

42. DELACROIX, Sinny et al. The Rad9–Hus1–Rad1 (9–1–1) clamp activates checkpoint signaling via TopBP1. Genes & development, v. 21, n. 12, p. 1472-1477, 2007.

43. DOSTÁLOVÁ, Anna; VOLF, Petr. Leishmania development in sand flies: parasite-vector interactions overview. Parasites & vectors, v. 5, n. 1, p. 1-12, 2012.

44. DUNCAN, Samuel M. et al. Conditional gene deletion with DiCre demonstrates an essential role for CRK3 in L eishmania mexicana cell cycle regulation. Molecular microbiology, v. 100, n. 6, p. 931-944, 2016.

45. EYKELENBOOM, John Kenneth et al. ATR activates the SM checkpoint during unperturbed growth to ensure sufficient replication prior to mitotic onset. Cell reports, v. 5, n. 4, p. 1095-1107, 2013.

46. FALCÃO DE OLIVEIRA, Everton et al. Monthly distribution of phlebotomine sand flies, and biotic and abiotic factors related to their abundance, in an urban area to which visceral leishmaniasis is endemic in Corumbá, Brazil. PLoS One, v. 11, n. 10, p. e0165155, 2016.

47. FINKEL, Toren; HOLBROOK, Nikki J. Oxidants, oxidative stress and the biology of ageing. nature, v. 408, n. 6809, p. 239-247, 2000.

48. FRAGKOS, Michalis et al. DNA replication origin activation in space and time. Nature reviews Molecular cell biology, v. 16, n. 6, p. 360-374, 2015.

49. FRANCHITTO, Annapaola et al. Replication fork stalling in WRN-deficient cells is overcome by prompt activation of a MUS81-dependent pathway. The Journal of cell biology, v. 183, n. 2, p. 241-252, 2008.

50. GE, Xin Quan; BLOW, J. Julian. Chk1 inhibits replication factory activation but allows dormant origin firing in existing factories. Journal of Cell biology, v. 191, n. 7, p. 1285-1297, 2010.

51. GENOIS, Marie-Michelle et al. DNA repair pathways in trypanosomatids: from DNA repair to drug resistance. Microbiology and Molecular Biology Reviews, v. 78, n. 1, p. 40-73, 2014.

52. GLOVER, Lucy; HORN, David. Trypanosomal histone γH2A and the DNA damage response. Molecular and biochemical parasitology, v. 183, n. 1, p. 78-83, 2012.

53. GÖHLER, Thomas et al. ATR-mediated phosphorylation of DNA polymerase η is needed for efficient recovery from UV damage. Journal of Cell Biology, v. 192, n. 2, p. 219-227, 2011.

54. GÓMEZ, Ricardo Obonaga. A deficiência das proteínas de checkpoint HUS1 e RAD9 promove a variação do número de cópias no genoma de Leishmania major. Tese de Doutorado. Universidade de São Paulo.

55. GONZÁLEZ, Camila et al. Climate change and risk of leishmaniasis in North America: predictions from ecological niche models of vector and reservoir species. PLoS neglected tropical diseases, v. 4, n. 1, p. e585, 2010.

56. GORDON, David J.; RESIO, Benjamin; PELLMAN, David. Causes and consequences of aneuploidy in cancer. Nature Reviews Genetics, v. 13, n. 3, p. 189-203, 2012.

57. GRAWUNDER, Ulf et al. Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells. Nature, v. 388, n. 6641, p. 492-495, 1997.

58. HAILE, Simon; DUPE, Aurelien; PAPADOPOULOU, Barbara. Deadenylation-independent stage-specific mRNA degradation in Leishmania. Nucleic acids research, v. 36, n. 5, p. 1634-1644, 2008.

59. HAIMEUR, Anass; OUELLETTE, Marc. Gene amplification in Leishmania tarentolae selected for resistance to sodium stibogluconate. Antimicrobial agents and chemotherapy, v. 42, n. 7, p. 1689-1694, 1998.

60. HAMMARTON, Tansy C. et al. Stage-specific differences in cell cycle control in Trypanosoma brucei revealed by RNA interference of a mitotic cyclin. Journal of Biological Chemistry, v. 278, n. 25, p. 22877-22886, 2003.

61. HARHAY, Michael O. et al. Urban parasitology: visceral leishmaniasis in Brazil. Trends in parasitology, v. 27, n. 9, p. 403-409, 2011.

62. HILTON, Benjamin A. et al. ATR plays a direct antiapoptotic role at mitochondria, which is regulated by prolyl isomerase Pin1. Molecular cell, v. 60, n. 1, p. 35-46, 2015.

63. HU, Huiqing et al. The centriole cartwheel protein SAS-6 in Trypanosoma brucei is required for probasal body biogenesis and flagellum assembly. Eukaryotic Cell, v. 14, n. 9, p. 898-907, 2015.

64. IANTORNO, Stefano A. et al. Gene expression in Leishmania is regulated predominantly by gene dosage. MBio, v. 8, n. 5, p. e01393-17, 2017.

65. IVENS, Alasdair C. et al. The genome of the kinetoplastid parasite, Leishmania major. Science, v. 309, n. 5733, p. 436-442, 2005.

66. JACKSON, Stephen P.; BARTEK, Jiri. The DNA-damage response in human biology and disease. Nature, v. 461, n. 7267, p. 1071-1078, 2009.

67. JAZAYERI, Ali et al. ATM-and cell cycle-dependent regulation of ATR in response to DNA double-strand breaks. Nature cell biology, v. 8, n. 1, p. 37-45, 2006.

68. JIANG, Wenxia et al. Differential phosphorylation of DNA-PKcs regulates the interplay between end-processing and end-ligation during nonhomologous end-joining. Molecular cell, v. 58, n. 1, p. 172-185, 2015.

69. KABECHE, Lilian et al. A mitosis-specific and R loop–driven ATR pathway promotes faithful chromosome segregation. Science, v. 359, n. 6371, p. 108-114, 2018.

70. KARLSSON-ROSENTHAL, Christina; MILLAR, Jonathan BA. Cdc25: mechanisms of checkpoint inhibition and recovery. Trends in cell biology, v. 16, n. 6, p. 285-292, 2006.

71. KARNANI, Neerja; DUTTA, Anindya. The effect of the intra-S-phase checkpoint on origins of replication in human cells. Genes & development, v. 25, n. 6, p. 621-633, 2011.

72. KAYE, Paul; SCOTT, Phillip. Leishmaniasis: complexity at the host–pathogen interface. Nature reviews microbiology, v. 9, n. 8, p. 604-615, 2011.

73. KEEGAN, Kathleen S. et al. The Atr and Atm protein kinases associate with different sites along meiotically pairing chromosomes. Genes & development, v. 10, n. 19, p. 2423-2437, 1996.

74. kineto

75. KOÇ, Ahmet et al. Hydroxyurea arrests DNA replication by a mechanism that preserves basal dNTP pools. Journal of Biological Chemistry, v. 279, n. 1, p. 223-230, 2004.

76. KOREN, Amnon et al. Genetic variation in human DNA replication timing. Cell, v. 159, n. 5, p. 1015-1026, 2014.

77. KOWALCZYKOWSKI, Stephen C. An overview of the molecular mechanisms of recombinational DNA repair. Cold Spring Harbor Perspectives in Biology, v. 7, n. 11, p. a016410, 2015.

78. KUMAGAI, Akiko et al. TopBP1 activates the ATR-ATRIP complex. Cell, v. 124, n. 5, p. 943-955, 2006.

79. KUMAGAI, Akiko; DUNPHY, William G. Claspin, a novel protein required for the activation of Chk1 during a DNA replication checkpoint response in Xenopus egg extracts. Molecular cell, v. 6, n. 4, p. 839-849, 2000.

80. KUMAR, Vipul; ALT, Frederick W.; OKSENYCH, Valentyn. Functional overlaps between XLF and the ATM-dependent DNA double strand break response. DNA repair, v. 16, p. 11-22, 2014.

81. LAFFITTE, Marie-Claude N. et al. Plasticity of the Leishmania genome leading to gene copy number variations and drug resistance. F1000Research, v. 5, 2016.

82. LAVIN, Martin F. Ataxia-telangiectasia: from a rare disorder to a paradigm for cell signalling and cancer. Nature reviews Molecular cell biology, v. 9, n. 10, p. 759-769, 2008.

83. LI, Zhiying et al. The XRCC4 gene encodes a novel protein involved in DNA double-strand break repair and V (D) J recombination. Cell, v. 83, n. 7, p. 1079-1089, 1995.

84. LIAO, Hongwei et al. Mechanisms for stalled replication fork stabilization: new targets for synthetic lethality strategies in cancer treatments. EMBO reports, v. 19, n. 9, p. e46263, 2018.

85. LIMA MACIEL, Bruna L. et al. Dual immune modulatory effect of vitamin A in human visceral leishmaniasis. PLoS One, v. 9, n. 9, p. e107564, 2014.

86. LIU, Beiyu et al. Fellowship of the rings: the replication of kinetoplast DNA. Trends in parasitology, v. 21, n. 8, p. 363-369, 2005.

87. LIU, Han et al. Phosphorylation of MLL by ATR is required for execution of mammalian S-phase checkpoint. Nature, v. 467, n. 7313, p. 343-346, 2010.

88. LIU, Shizhou et al. ATR autophosphorylation as a molecular switch for checkpoint activation. Molecular cell, v. 43, n. 2, p. 192-202, 2011.

89. LOMBRAÑA, Rodrigo et al. Transcriptionally driven DNA replication program of the human parasite Leishmania major. Cell reports, v. 16, n. 6, p. 1774-1786, 2016.

90. LORDELO, Peresi. EPIDEMIOLOGIA DA LEISHMANIOSE VISCERAL NO MUNICÍPIO DE TRÊS LAGOAS, MATO GROSSO DO SUL, UMA NOVA REGIÃO ENDÊMICA NO BRASIL. 2021.

91. LOSSAINT, Gérald et al. FANCD2 binds MCM proteins and controls replisome function upon activation of s phase checkpoint signaling. Molecular cell, v. 51, n. 5, p. 678-690, 2013.

92. LOVEJOY, Courtney A. et al. Loss of ATRX, genome instability, and an altered DNA damage response are hallmarks of the alternative lengthening of telomeres pathway. PLoS genetics, v. 8, n. 7, p. e1002772, 2012.

93. MAIA-ELKHOURY, Ana Nilce Silveira et al. ORGANIZAÇÃO PAN-AMERICANA DA SAÚDE: LEISHMANIOSES: INFORME EPIDEMIOLÓGICO NAS AMÉRICAS. WASHINGTON: ORGANIZAÇÃO PAN-AMERICANA DA SAÚDE. 2021.

94. MARIN, Paula Andrea et al. ATR kinase is a crucial player mediating the DNA damage response in Trypanosoma brucei. Frontiers in cell and developmental biology, v. 8, p. 1651, 2020.

95. MAROLI, M. et al. Phlebotomine sandflies and the spreading of leishmaniases and other diseases of public health concern. Medical and veterinary entomology, v. 27, n. 2, p. 123-147, 2013.

96. MARQUES, Catarina A. et al. Genome-wide mapping reveals single-origin chromosome replication in Leishmania, a eukaryotic microbe. Genome biology, v. 16, n. 1, p. 1-12, 2015.

97. MARTÍNEZ-CALVILLO, Santiago et al. Transcription initiation and termination on Leishmania major chromosome 3. Eukaryotic cell, v. 3, n. 2, p. 506-517, 2004.

98. MATTHEWS, Keith R.; TSCHUDI, Christian; ULLU, Elisabetta. A common pyrimidine-rich motif governs trans-splicing and polyadenylation of tubulin polycistronic pre-mRNA in trypanosomes. Genes & development, v. 8, n. 4, p. 491-501, 1994.

99. MCMAHON, Steven B. et al. The novel ATM-related protein TRRAP is an essential cofactor for the c-Myc and E2F oncoproteins. Cell, v. 94, n. 3, p. 363-374, 1998.

100. MELCHIONNA, Roberta et al. Threonine 68 is required for radiation-induced phosphorylation and activation of Cds1. Nature cell biology, v. 2, n. 10, p. 762-765, 2000.

101. MOAFI, Mohammad et al. Leishmania vaccines entered in clinical trials: a review of literature. International Journal of Preventive Medicine, v. 10, 2019.

102. MOISEEVA, Tatiana N. et al. An ATR and CHK1 kinase signaling mechanism that limits origin firing during unperturbed DNA replication. Proceedings of the National Academy of Sciences, v. 116, n. 27, p. 13374-13383, 2019.

103. MORDES, Daniel A. et al. TopBP1 activates ATR through ATRIP and a PIKK regulatory domain. Genes & development, v. 22, n. 11, p. 1478-1489, 2008.

104. MORDES, Daniel A.; CORTEZ, David. Activation of ATR and related PIKKs. Cell cycle, v. 7, n. 18, p. 2809-2812, 2008.

105. MOSHOUS, Despina et al. Artemis, a novel DNA double-strand break repair/V (D) J recombination protein, is mutated in human severe combined immune deficiency. Cell, v. 105, n. 2, p. 177-186, 2001.

106. MÜLLER, Carolin A. et al. Capturing the dynamics of genome replication on individual ultra-long nanopore sequence reads. Nature methods, v. 16, n. 5, p. 429-436, 2019.

107. MÜLLER, Michaela; PADMANABHAN, Prasad K.; PAPADOPOULOU, Barbara. Selective inactivation of SIDER2 retroposon-mediated mRNA decay contributes to stage-and species-specific gene expression in Leishmania. Molecular microbiology, v. 77, n. 2, p. 471-491, 2010.

108. MURGA, Matilde et al. A mouse model of ATR-Seckel shows embryonic replicative stress and accelerated aging. Nature genetics, v. 41, n. 8, p. 891-898, 2009.

109. MURPHY, Anar K. et al. Phosphorylated RPA recruits PALB2 to stalled DNA replication forks to facilitate fork recovery. Journal of Cell Biology, v. 206, n. 4, p. 493-507, 2014.

110. MYLER, Logan R. et al. Single-molecule imaging reveals how Mre11-Rad50-Nbs1 initiates DNA break repair. Molecular cell, v. 67, n. 5, p. 891-898. e4, 2017.

111. NEELSEN, Kai J.; LOPES, Massimo. Replication fork reversal in eukaryotes: from dead end to dynamic response. Nature reviews Molecular cell biology, v. 16, n. 4, p. 207-220, 2015.

112. NUNES, Vinicius S. et al. The Hus1 homologue of Leishmania major encodes a nuclear protein that participates in DNA damage response. Molecular and biochemical parasitology, v. 177, n. 1, p. 65-69, 2011.

113. PAGÈS, Michel et al. Chromosome size and number polymorphisms in Leishmania infantum suggest amplification/deletion and possible genetic exchange. Molecular and biochemical parasitology, v. 36, n. 2, p. 161-168, 1989.

114. PANNUNZIO, N. R., Watanabe, G., & Lieber, M. R. (2018). Nonhomologous DNA end-joining for repair of DNA double-strand breaks. Journal of Biological Chemistry, 293(27), 10512-10523.

115. PEACOCK, Christopher S. et al. Comparative genomic analysis of three Leishmania species that cause diverse human disease. Nature genetics, v. 39, n. 7, p. 839-847, 2007.

116. PELLEGRINI, Manuela et al. Autophosphorylation at serine 1987 is dispensable for murine Atm activation in vivo. Nature, v. 443, n. 7108, p. 222-225, 2006.

117. PEREZ-MORGA, David L.; ENGLUND, Paul T. The attachment of minicircles to kinetoplast DNA networks during replication. Cell, v. 74, n. 4, p. 703-711, 1993.

118. PETERMANN, Eva et al. Hydroxyurea-stalled replication forks become progressively inactivated and require two different RAD51-mediated pathways for restart and repair. Molecular cell, v. 37, n. 4, p. 492-502, 2010.

119. POSTIGO, Antonio et al. Cytoplasmic ATR activation promotes vaccinia virus genome replication. Cell reports, v. 19, n. 5, p. 1022-1032, 2017.

120. RAGLAND, Ryan L. et al. RNF4 and PLK1 are required for replication fork collapse in ATR-deficient cells. Genes & development, v. 27, n. 20, p. 2259-2273, 2013.

121. RAMÍREZ, Fidel et al. deepTools: a flexible platform for exploring deep-sequencing data. Nucleic acids research, v. 42, n. W1, p. W187-W191, 2014.

122. RAMKUMAR, Kavya et al. AXL Inhibition Induces DNA Damage and Replication Stress in Non–Small Cell Lung Cancer Cells and Promotes Sensitivity to ATR Inhibitors. Molecular Cancer Research, v. 19, n. 3, p. 485-497, 2021.

123. RAO, Qinhui et al. Cryo-EM structure of human ATR-ATRIP complex. Cell research, v. 28, n. 2, p. 143-156, 2018.

124. RIBALLO, Enriqueta et al. A pathway of double-strand break rejoining dependent upon ATM, Artemis, and proteins locating to γ-H2AX foci. Molecular cell, v. 16, n. 5, p. 715-724, 2004.

125. ROCHETTE, Annie et al. Genome-wide gene expression profiling analysis of Leishmania major and Leishmania infantum developmental stages reveals substantial differences between the two species. BMC genomics, v. 9, n. 1, p. 1-26, 2008.

126. ROGERS, Matthew E.; CHANCE, M. L.; BATES, P. A. The role of promastigote secretory gel in the origin and transmission of the infective stage of Leishmania mexicana by the sandfly Lutzomyia longipalpis. Parasitology, v. 124, n. 5, p. 495-507, 2002.

127. ROYO, Hélène et al. ATR acts stage specifically to regulate multiple aspects of mammalian meiotic silencing. Genes & development, v. 27, n. 13, p. 1484-1494, 2013.

128. RUIZ, Sergio et al. A genome-wide CRISPR screen identifies CDC25A as a determinant of sensitivity to ATR inhibitors. Molecular cell, v. 62, n. 2, p. 307-313, 2016.

129. RUIZ-POSTIGO, Jose Antonio et al. Global leishmaniasis surveillance: 2019-2020, a baseline for the 2030 roadmap/Surveillance mondiale de la leishmaniose: 2019-2020, une periode de reference pour la feuille de route a l'horizon 2030. Weekly Epidemiological Record, v. 96, n. 35, p. 401-420, 2021.

130. SALDIVAR, Joshua C. et al. An intrinsic S/G2 checkpoint enforced by ATR. Science, v. 361, n. 6404, p. 806-810, 2018.

131. SALDIVAR, Joshua C. et al. Science, v. 361, n. 6404, p. 806-810, 2018.

132. SALDIVAR, Joshua C.; CORTEZ, David; CIMPRICH, Karlene A. The essential kinase ATR: ensuring faithful duplication of a challenging genome. Nature reviews Molecular cell biology, v. 18, n. 10, p. 622-636, 2017.

133. SCHÜBELER, Dirk et al. Genome-wide DNA replication profile for Drosophila melanogaster: a link between transcription and replication timing. Nature genetics, v. 32, n. 3, p. 438-442, 2002.

134. SEIDEL, Jeffrey J.; ANDERSON, Carol M.; BLACKBURN, Elizabeth H. A novel Tel1/ATM N-terminal motif, TAN, is essential for telomere length maintenance and a DNA damage response. Molecular and cellular biology, v. 28, n. 18, p. 5736-5746, 2008.

135. SERPELONI, Mariana et al. UAP56 is a conserved crucial component of a divergent mRNA export pathway in Toxoplasma gondii. Molecular microbiology, v. 102, n. 4, p. 672-689, 2016.

136. SHAPIRO, Theresa A.; ENGLUND, Paul T. The structure and replication of kinetoplast DNA. Annual review of microbiology, v. 49, n. 1, p. 117-143, 1995.

137. SHECHTER, David; COSTANZO, Vincenzo; GAUTIER, Jean. ATR and ATM regulate the timing of DNA replication origin firing. Nature cell biology, v. 6, n. 7, p. 648-655, 2004.

138. SHILOH, Yosef; ZIV, Yael. The ATM protein kinase: regulating the cellular response to genotoxic stress, and more. Nature reviews Molecular cell biology, v. 14, n. 4, p. 197-210, 2013.

139. SHOR, Erika; PERLIN, David S. Coping with stress and the emergence of multidrug resistance in fungi. PLoS pathogens, v. 11, n. 3, p. e1004668, 2015.

140. SMITH, Graeme CM; JACKSON, Stephen P. The DNA-dependent protein kinase. Genes & development, v. 13, n. 8, p. 916-934, 1999.

141. SOGO, José M.; LOPES, Massimo; FOIANI, Marco. Fork reversal and ssDNA accumulation at stalled replication forks owing to checkpoint defects. Science, v. 297, n. 5581, p. 599-602, 2002.

142. STERKERS, Yvon et al. FISH analysis reveals aneuploidy and continual generation of chromosomal mosaicism in Leishmania major. Cellular microbiology, v. 13, n. 2, p. 274-283, 2011.

143. STERKERS, Yvon et al. Novel insights into genome plasticity in Eukaryotes: mosaic aneuploidy in Leishmania. Molecular microbiology, v. 86, n. 1, p. 15-23, 2012.

144. STORTZ, Jennifer A. et al. Genome-wide and protein kinase-focused RNAi screens reveal conserved and novel damage response pathways in Trypanosoma brucei. PLoS pathogens, v. 13, n. 7, p. e1006477, 2017.

145. STRACKER, Travis H.; PETRINI, John HJ. The MRE11 complex: starting from the ends. Nature reviews Molecular cell biology, v. 12, n. 2, p. 90-103, 2011.

146. TANNOUS, Elias A. et al. Mechanism of auto-inhibition and activation of Mec1ATR checkpoint kinase. Nature structural & molecular biology, v. 28, n. 1, p. 50-61, 2021.

147. TÉCHER, Hervé et al. Signaling from Mus81-Eme2-dependent DNA damage elicited by Chk1 deficiency modulates replication fork speed and origin usage. Cell reports, v. 14, n. 5, p. 1114-1127, 2016.

148. TIENGWE, Calvin et al. Genome-wide analysis reveals extensive functional interaction between DNA replication initiation and transcription in the genome of Trypanosoma brucei. Cell reports, v. 2, n. 1, p. 185-197, 2012.

149. TIENGWE, Calvin; MARQUES, Catarina A.; MCCULLOCH, Richard. Nuclear DNA replication initiation in kinetoplastid parasites: new insights into an ancient process. Trends in parasitology, v. 30, n. 1, p. 27-36, 2014.

150. TOLEDO, Luis Ignacio et al. ATR prohibits replication catastrophe by preventing global exhaustion of RPA. Cell, v. 155, n. 5, p. 1088-1103, 2013.

151. UBEDA, Jean-Michel et al. Genome-wide stochastic adaptive DNA amplification at direct and inverted DNA repeats in the parasite Leishmania. PLoS biology, v. 12, n. 5, p. e1001868, 2014.

152. VAN DER PLOEG, Lex HT et al. Antigenic variation in Trypanosoma brucei analyzed by electrophoretic separation of chromosome-sized DNA molecules. Cell, v. 37, n. 1, p. 77-84, 1984.

153. VASSIN, Vitaly M. et al. Human RPA phosphorylation by ATR stimulates DNA synthesis and prevents ssDNA accumulation during DNA-replication stress. Journal of cell science, v. 122, n. 22, p. 4070-4080, 2009.

154. VIEGAS, Idalio J. et al. N6-methyladenosine in poly (A) tails stabilize VSG transcripts. Nature, v. 604, n. 7905, p. 362-370, 2022.

155. VU, Giang TH et al. Repair of site-specific DNA double-strand breaks in barley occurs via diverse pathways primarily involving the sister chromatid. The Plant Cell, v. 26, n. 5, p. 2156-2167, 2014.

156. WALKER, John R.; CORPINA, Richard A.; GOLDBERG, Jonathan. Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. Nature, v. 412, n. 6847, p. 607-614, 2001.

157. WANG, Zefeng; ENGLUND, Paul T. RNA interference of a trypanosome topoisomerase II causes progressive loss of mitochondrial DNA. The EMBO journal, v. 20, n. 17, p. 4674-4683, 2001.

158. WHEELER, Richard J.; GLUENZ, Eva; GULL, Keith. The cell cycle of Leishmania: morphogenetic events and their implications for parasite biology. Molecular microbiology, v. 79, n. 3, p. 647-662, 2011.

159. WHEELER, Richard J.; GULL, Keith; SUNTER, Jack D. Coordination of the cell cycle in trypanosomes. Annual review of microbiology, v. 73, p. 133-154, 2019.

160. WORLD HEALTH ORGANIZATION et al. Ending the neglect to attain the sustainable development goals: a road map for neglected tropical diseases 2021–2030. 2020.

161. YAMASHITA, Akio et al. Human SMG-1, a novel phosphatidylinositol 3-kinase-related protein kinase, associates with components of the mRNA surveillance complex and is involved in the regulation of nonsense-mediated mRNA decay. Genes & development, v. 15, n. 17, p. 2215-2228, 2001.

162. YANG, Haijuan et al. mTOR kinase structure, mechanism and regulation. Nature, v. 497, n. 7448, p. 217-223, 2013.

163. ZEGERMAN, Philip; DIFFLEY, John FX. Checkpoint-dependent inhibition of DNA replication initiation by Sld3 and Dbf4 phosphorylation. Nature, v. 467, n. 7314, p. 474-478, 2010.

164. ZEMAN, Michelle K.; CIMPRICH, Karlene A. Causes and consequences of replication stress. Nature cell biology, v. 16, n. 1, p. 2-9, 2014.

165. ZHA, Shan et al. ATM damage response and XLF repair factor are functionally redundant in joining DNA breaks. Nature, v. 469, n. 7329, p. 250-254, 2011.

166. ZHANG, Wen-Wei; MATLASHEWSKI, Greg. CRISPR-Cas9-mediated genome editing in Leishmania donovani. MBio, v. 6, n. 4, p. e00861-15, 2015.

167. ZHAO, Xiaolan; MULLER, Eric GD; ROTHSTEIN, Rodney. A suppressor of two essential checkpoint genes identifies a novel protein that negatively affects dNTP pools. Molecular cell, v. 2, n. 3, p. 329-340, 1998.

168. ZHAO, Xiaolan; MULLER, Eric GD; ROTHSTEIN, Rodney. A suppressor of two essential checkpoint genes identifies a novel protein that negatively affects dNTP pools. Molecular cell, v. 2, n. 3, p. 329-340, 1998.

169. ZHOU, Qing et al. A kinetochore-based ATM/ATR-independent DNA damage checkpoint maintains genomic integrity in trypanosomes. Nucleic acids research, v. 47, n. 15, p. 7973-7988, 2019.

170. ZHOU, Zhong-Wei et al. An essential function for the ATR-activation-domain (AAD) of TopBP1 in mouse development and cellular senescence. PLoS genetics, v. 9, n. 8, p. e1003702, 2013.

171. ZOU, Lee; LIU, Dou; ELLEDGE, Stephen J. Replication protein A-mediated recruitment and activation of Rad17 complexes. Proceedings of the National Academy of Sciences, v. 100, n. 24, p. 13827-13832, 2003.

Supplementary Data

7. Tables

7.1 Trimmomatic

Table 7.1 – Trim the adaptors of the sequenced data

Trimmomatic	Coordanates
Single-end or paired-end reads?	pair_of_files
Input FASTQ file (R1/first of pair)	file1.fastaq.gz
Input FASTQ file (R2/second of pair)	file2.fastaq.gz
Perform initial ILLUMINACLIP step?	yes
Select standard adapter sequences or provide custom?	standard
Adapter sequences to use	TruSeq3 (paired-ended, for MiSeq and HiSeq)
Maximum mismatch count which will still allow a full match to be performed	2
How accurate the match between the two 'adapter ligated' reads must be for PE palindrome read alignment	30
How accurate the match between any adapter etc. sequence must be against a read	10
Minimum length of adapter that needs to be detected (PE specific/palindrome mode)	8
Always keep both reads (PE specific/palindrome mode)?	True
Select Trimmomatic operation to perform	SLIDINGWINDOW
Number of bases to average across	4
Average quality required	20
Output trimlog file?	False
Output trimmomatic log messages?	False

Galaxy server

7.2 Mapped reads (BAM format)

Table 7.2 – Mapped reads (BAM format)

Map with BWA-MEM on (mapped reads in BAM format)	Coordanates
Will you select a reference genome from your history or use a built-in index?	history
Use the following dataset as the reference sequence	TriTrypDB-50_LmajorFriedlin_Genome.fasta (TriTrypdb.org)
Single or Paired-end reads	paired
Select first set of reads	file 1 (R1 paired)
Select second set of reads	file 2 (R2 paired)
Enter mean, standard deviation, max, and min for insert lengths.	150
Set read groups information?	do_not_set
Select analysis mode	illumina

Galaxy server

7.3 Compare with Stationary cells

Table 7.3 – bamCompare

bamCompare	Coordenates
First file	file 1
Control file	file 2 (Stationary Cas9)
Bin size in bases	1000
Method to use for scaling the largest sample to the smallest	readCount
How to compare the two files	ratios
Pseudocount	1
Compute an exact scaling factor	False
Coverage file format	bedgraph
Region of the genome to limit the operation to	Empty.
Show advanced options	yes
Smooth values using the following length (in bases)	Not available
Extend reads to the given average fragment size.	yes
Ignore duplicates	true
Center regions with respect to the fragment length	true
Minimum mapping quality	0
Include reads based on the SAM flag	Not available
Minimum fragment length for inclusion.	0
Maximum fragment length for inclusion.	0
Ignore missing data?	False
Skip bins of no coverage	No
regions that should be excluded for calculating the scaling factor	Empty.

Galaxy server

7.4 Text reformating

Table 7.4 – Text reformating

Text reformatting	Coordenates
File to process	**file (bamCompare)**
AWK Program	**{if($4<2.8)print$0}**

Galaxy server

7.5 bedtools Intersect intervals

Table 7.5 - bedtools Intersect intervals

bedtools Intersect intervals	Coordenates
File A to intersect with B	text reformating file
Combined or separate output files	iterate
File B to intersect with A	ShadeAreas.bed
Calculation based on strandedness?	Overlaps on either strand
What should be written to the output file?	Nothing selected.
Treat split/spliced BAM or BED12 entries as distinct BED intervals when computing coverage.	False
Required overlap	default
Report only those alignments that **do not** overlap with file(s) B	True
Write the original A entry _once_ if _any_ overlaps found in B.	False
For each entry in A, report the number of overlaps with B.	False
When using BAM input (-abam), write output as BED instead of BAM.	False
For coordinate sorted input file the more efficient sweeping algorithm is enabled.	False
Print the header from the A file prior to results	False

Galaxy server

7.6 Convert in Bigwig format

Table 7.6 – Convert in Bigwig format

Wig/BedGraph-to-bigWig	Coordentes
Convert	file_SMOOTH_bedgraph
Converter settings to use	preset

Galaxy server

7.7 bamCoverage

Table 7.7 – bamCoverage

bamCoverage	Coordenates
BAM/CRAM file	file_BWA (BAM format)
Bin size in bases	2000
Scaling/Normalization method	no
Coverage file format	bedgraph
Compute an exact scaling factor	False
Region of the genome to limit the operation to	Empty.
Show advanced options	yes
Scaling factor	1.0
Smooth values using the following length (in bases)	Not available.
Regions that should be excluded for normalization	Empty.
Ignore missing data?	True
Extend reads to the given average fragment size.	yes
Ignore duplicates	True
Center regions with respect to the fragment length	True
Minimum mapping quality	30
Include reads based on the SAM flag	Not available.
Exclude reads based on the SAM flag	Not available.
Minimum fragment length for inclusion.	0
Maximum fragment length for inclusion.	0
Determine nucleosome positions from MNase-seq data	False
Offset inside each alignment to use for the signal location.	Empty.
Only include reads originating from fragments from the forward or reverse strand.	no
Blacklisted regions in BED/GTF format	ShadeAreas.bed

Galaxy server

7.8 FreeBayes

Table 7.8 – FreeBayes

FreeBayes	Coordenates
Choose the source for the reference genome	history
Run in batch mode?	individual
BAM dataset	file_BWA (BAM format)
Use the following dataset as the reference sequence	LmajorFriedlin_Genome.fasta (Trytripdb.or)
Limit variant calling to a set of regions?	do_not_limit
Read coverage	set
Require at least this coverage to process a site	0
Downsample per-sample coverage.	0
Skip processing of.	0
Choose parameter selection level	full
Turn off left-alignment of indels	False
Input filters	set
Include duplicate-marked alignments in the analysis	False
Exclude alignments.	20
Exclude alleles.	20
Consider any allele.	0
Consider any allele in which.	0
Exclude reads with more than this number of separate gaps	1000
Use stringent input base and mapping quality filters	False
Require at least this fraction.	0.05
Require at least this count.	2
Require at least this sum.	0
Require at least this count.	1

Galaxy server

7.9 VCF filter

Table 7.9 – VCF filter

VCF filter	Coordenates
VCF dataset to filter	file 1_FreeBayes
Select the filter type	Info filter (-f)
Specify filtering value	QUAL > 29
Select the filter type	Info filter (-f)
Specify filtering value	TYPE = snp or ins or del
Filter entire records, not just alleles	False
Tag vcf records as positively filtered with this tag, print all records	False
Tag vcf records as negatively filtered with this tag, print all records	False
Append the existing filter tag, don't just replace it	False
Apply --tag-pass on a per-allele basis, adds or sets the corresponding INFO field tag	False
Inverts the filter, e.g. grep -v	False
Use logical OR instead of AND to combine filters	False
Specify a region on which to target the filtering	Empty.

Galaxy server

7.10   VCF-VCF intersect

Table 7.10 – VCF-VCF intersect

VCF-VCFintersect	Coordenates
The first VCF dataset	file 2_VCFfilter
The second VCF dataset	file 1_VCFfilte
Choose the source for the reference genome	history
Using reference file	TriTrypDB-50_LmajorFriedlin_Genome.fasta (Trytripdb.org)
Union or intersection	Intersect
Invert selection?	True
compare records up to this many bp away (window size)	0
output whole loci when one alternate allele matches	False
Advanced controls	no

Galaxy server

7.11   bedtools Intersect intervals

Table 7.11 - bedtools Intersect intervals

bedtools Intersect intervals	coordenates
File A to intersect with B	01_LeishmaniaChromossomes.bed.txt
Combined or separate output files	iterate
File B to intersect with A	file_VCF_VCFintersect
Calculation based on strandedness?	Overlaps on either strand
What should be written to the output file?	Nothing selected.
Treat split/spliced BAM or BED12 entries as distinct BED intervals when computing coverage.	False
Required overlap	default
Report only those alignments that **do not** overlap with file(s) B	False
Write the original A entry _once_ if _any_ overlaps found in B.	False
For each entry in A, report the number of overlaps with B.	True
When using BAM input (-abam), write output as BED instead of BAM.	False
For coordinate sorted input file the more efficient sweeping algorithm is enabled.	False
Print the header from the A file prior to results	False

Galaxy server

8. Scripts

8.1 Zscore

```
library(zoo)
library(data.table)
library(readr)
library(tibble)



FILE_LIST <- Sys.glob("pathway"')
FILE_LIST
for(fin in FILE_LIST){
  bdg <- fread(fin, header = F, col.names=c('chrom', 'start', 'end', 'score'))
  bdg[, zscore := (score - mean(score))/sd(score), by= chrom]
  bdg[, mascore := rollmean(zscore, 5, fill= 'extend'), by= chrom]
  bdg$zscore <- NULL
  bdg$score <- NULL
  sample_name <- sub('\\.bedgraph$', '', basename(fin))
  suffix <- '_SMOOTH.bedgraph'
  outdir <- dirname(fin)
  outfile <- file.path(outdir, paste0(sample_name, suffix))
  cat(sprintf('Writing parsed %s to %s\n', fin, outfile))
  write.table(bdg, outfile, sep="\t", row.names = FALSE, col.names = FALSE,
quote = FALSE)
}
```

Script 8.1. Script to convert the MFA-seq signal bedgraph from ratio to Zscore using Rsudio.

8.2 CNV

```r
library(zoo)

library(data.table)

library(readr)

library(tibble)

FILE_LIST <- Sys.glob('*.bedgraph')

for(fin in FILE_LIST){

bdg1 <- fread(fin, header = F, col.names=c('chrom', 'start', 'end', 'score'))

bdg2 <-bdg1[,CCN :=(score/median(score))*2]

bdg3 <-aggregate(CCN ~ chrom, bdg2, median)

sample_name <- sub('\\.bedgraph$', '', basename(fin))

suffix <- '_PLOIDY.bedgraph'

outdir <- dirname(fin)

outfile <- file.path(outdir, paste0(sample_name, suffix))

cat(sprintf('Writing parsed %s to %s\n', fin, outfile))

write.table(bdg3, outfile, sep="\t", row.names = FALSE, col.names = FALSE,
quote = FALSE)}
```

Script 8.2. Script to calculate the CNV from bamCoverage file using Rsudio.

9. Supplementary Figures

9.1 Scheme of DNA repair pathways.

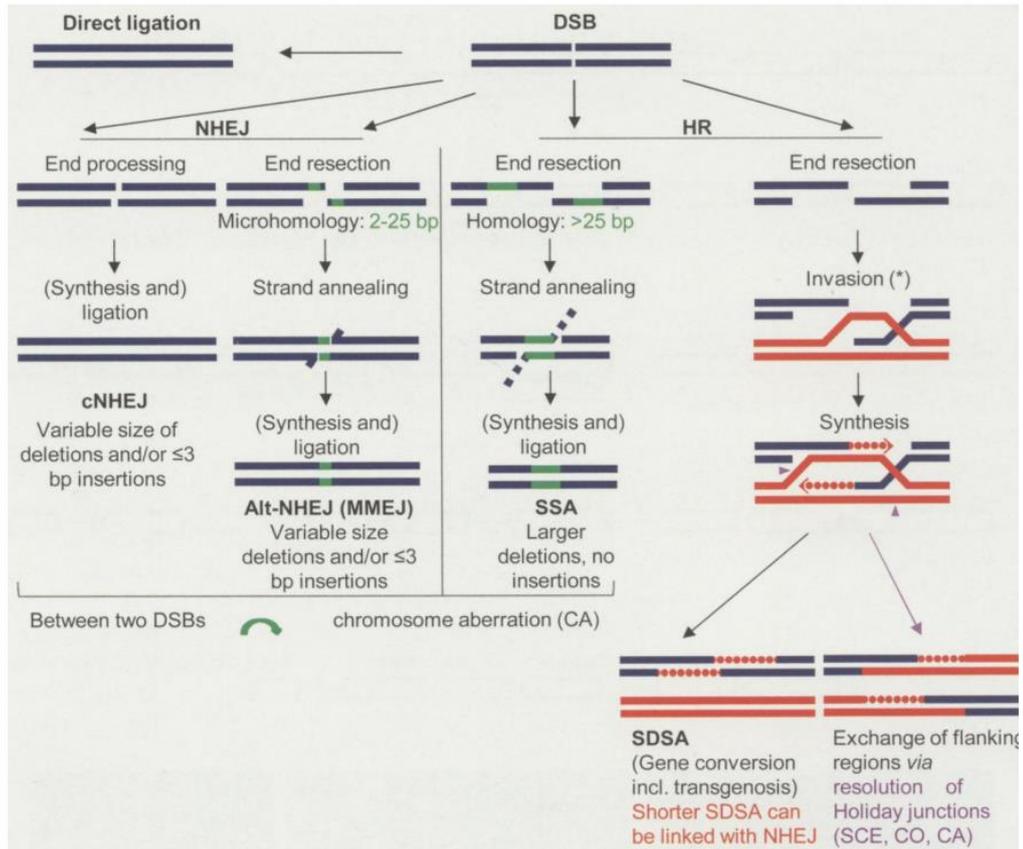**Figure 9.1 Scheme of DNA repair pathways.**

Figure 9.1. Schematic representation of different pathways to repair a DNA damage (From: Vu et al., 2014).

9.2 Western blotting to confirm the Cas9T7 linage.

Figure 9.2 Western blotting to confirm the Cas9T7 linage.
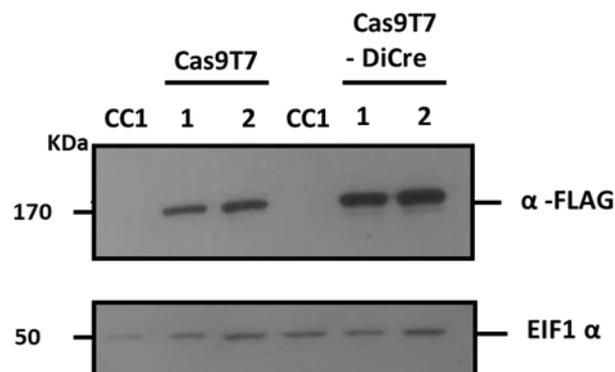
Figure 9.2. Western analysis of whole cell extract from *L.major* CC1 (control) and clones obtained from the transfection of *L.major* CC1 and *L.major* Dicre cells to express the constructor Cas9T7 (Beneke et al., 2017). The Cas9 endonuclease is tagged with a FLAG epitope. The detection was made using anti-FLAG and anti-EIF1α as a loading control.

9.3 Confirmation of locus alteration in ATR$^{ΔC+/- \, mycRPA}$ cell line.

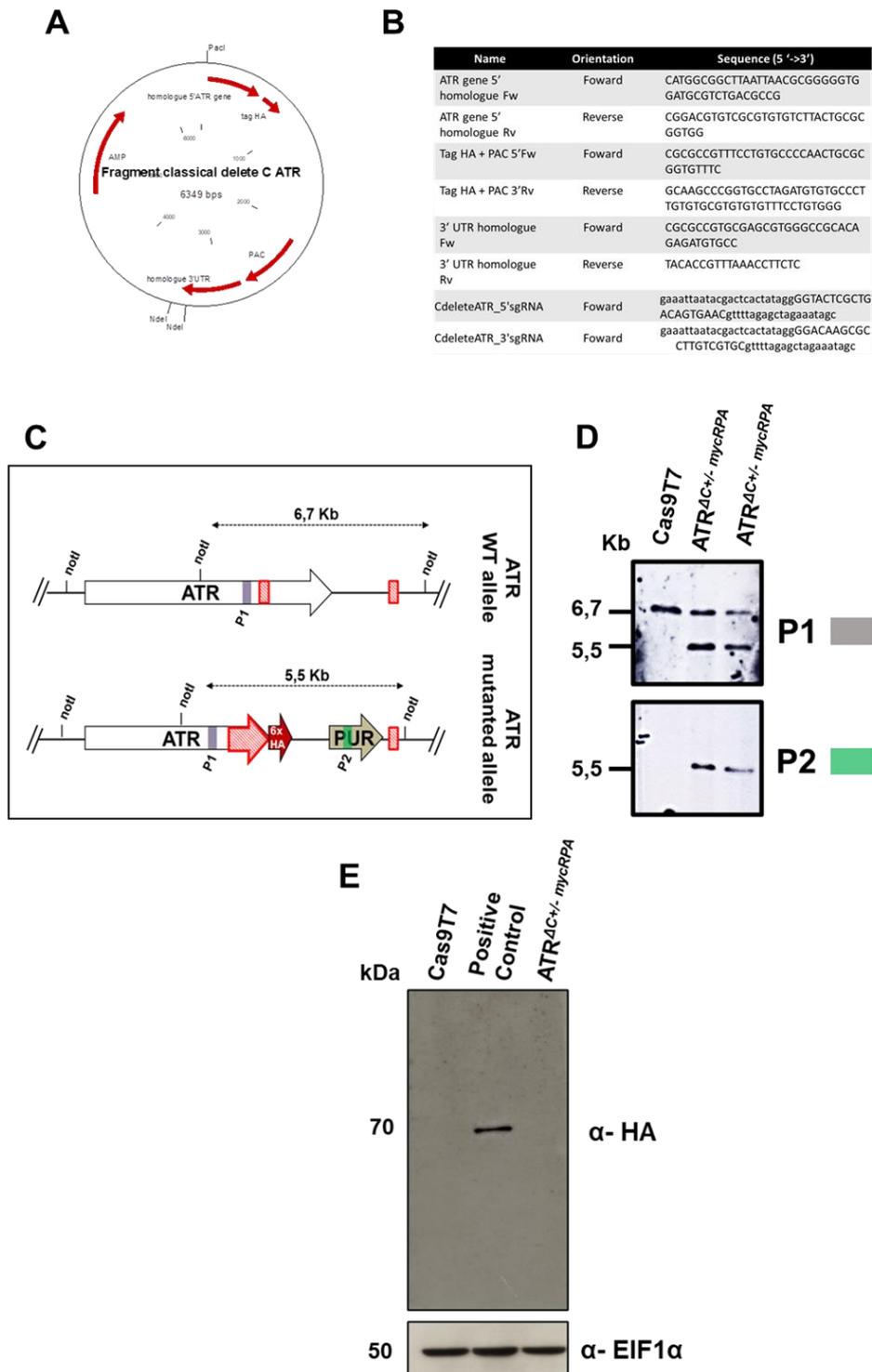**Figure 9.3 – Confirmation of locus alteration in ATR$^{ΔC+/- \, mycRPA}$ cell line.**

Figure 9.3. A) Schematic representation of the plasmid where it was extracting the fragment by using the restriction enzymes NdeI and PAcI, to delete ATR C-terminal; the fragment is composed by: two homolog recombination regions with around 500 pb size, six sequence that encode the tagging protein hemagglutinin (HA), a select marker puromycin (PAC). **B)** Table with the primers used to generate this plasmid mentioned in **(A)**. **C)** Schematic representation of the Southern Blotting strategy to identify the mutated ATR allele. Two probes were used: the first one target the middle region of ATR gene (gray square) and the second the selected marker gene (green square). The DNA of Cas9T7 cells (control) and ATR$^{ΔC+/-}$ cells were digested with the restriction enzyme NotI. **D)** Southern Blotting analysis showing that the ATR$^{ΔC+/-}$ cells have one wildtype ATR allele (6,7 kb) and one mutated ATR allele (5,5 kb). **E)** Western analysis of whole cell extract of indicated cell lines in exponential growth; extracts were probe with anti-HA antiserum and anti-EIF1α was used as a loading control (predicted protein sizes are indicated, kDa).

9.4 Genome-wide mapping of replication initiation in ^{myc}ATR and ^{myc}ATR^{ΔC} cells.

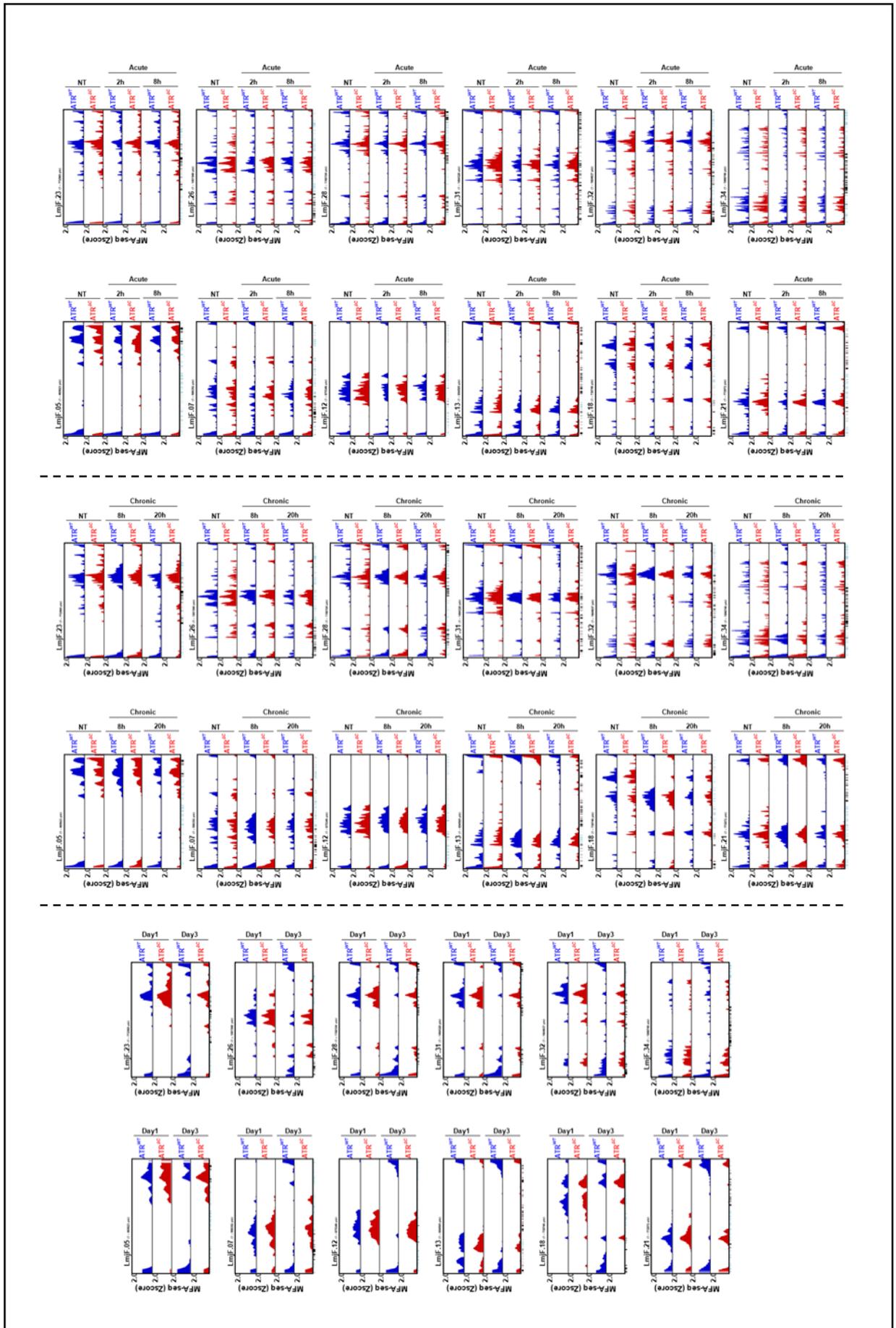## Figure 9.4 - Genome-wide mapping of replication initiation in ^{myc}ATR and ^{myc}ATR^{ΔC} cells.

Figure 9.4. Genome-wide mapping of replication initiation in ^{myc}ATR and ^{myc}ATR$^{\Delta C}$ cells in two times of growth curve or upon Chronic or Acute HU treatment. Graphs show the distribution of replication initiation across indicated chromosomes in the indicated cell lines; the bottom group show the DNA synthesis profile in different chromosomes at day 1 and day 3; the middle group show DNA synthesis profile in different chromosomes untreated or under chronic HU treatment, 0.5mM of HU for 8 and 20 hours; the top group show the DNA synthesis profile in different chromosomes untreated or under acute HU treatment, 5mM of HU for 2 and 8 hours. MFA-seq is represented by Z-scores across the chromosomes, calculated by comparing read depth coverage of each cell line DNA in the indicated time points relative to stationary Cas9T7 cells; the track in each chromosome bottom displays coding sequences, with genes transcribed from right to left in black, and from left to right in light blue.

<u>Attachments</u>

10. Attachments

10.1 First Author paper

10.1.1 (SILVA, Gabriel LA et al., Frontiers in Cell and Developmental Biology, 2021).

"Unpicking the Roles of DNA Damage Protein Kinases in Trypanosomatids". Published: 06 August 2021.

Abstract

To preserve genome integrity when faced with DNA lesions, cells activate and coordinate a multitude of DNA repair pathways to ensure timely error correction or tolerance, collectively called the DNA damage response (DDR). These interconnecting damage response pathways are molecular signal relays, with protein kinases (PKs) at the pinnacle. Focused efforts in model eukaryotes have revealed intricate aspects of DNA repair PK function, including how they direct DDR pathways and how repair reactions connect to wider cellular processes, including DNA replication and transcription. The Kinetoplastidae, including many parasites like Trypanosoma spp. and Leishmania spp. (causative agents of debilitating, neglected tropical infections), exhibit peculiarities in several core biological processes, including the predominance of multigenic transcription and the streamlining or repurposing of DNA repair pathways, such as the loss of non-homologous end joining and novel operation of nucleotide excision repair (NER). Very recent studies have implicated ATR and ATM kinases in the DDR of kinetoplastid parasites, whereas DNA-dependent protein kinase (DNAPKcs) displays uncertain conservation, questioning what functions it fulfills. The wide range of genetic manipulation approaches in these organisms presents an opportunity to investigate DNA repair kinase roles in kinetoplastids and to ask if further kinases are involved. Furthermore, the availability of kinase inhibitory compounds, targeting numerous eukaryotic PKs, could allow us to test the suitability of DNA repair PKs as novel chemotherapeutic targets. Here, we will review recent advances in the study of trypanosomatid DNA repair kinases.

10.2   Co-author papers:

10.2.1   (DAMASCENO, Jeziel D. et al. Memórias do Instituto Oswaldo Cruz. 2017)

"Evidence for regulated expression of Telomeric Repeat-containing RNAs (TERRA) in parasitic trypanosomatids". Published: 8 August 2017

Abstract

The Telomeric Repeat-containing RNAs (TERRA) participate in the homeostasis of telomeres in higher eukaryotes. Here, we investigated the expression of TERRA in Leishmania spp. and Trypanosoma brucei and found evidences for its expression as a specific RNA class. The trypanosomatid TERRA are heterogeneous in size and partially polyadenylated. The levels of TERRA transcripts appear to be modulated through the life cycle in both trypanosomatids investigated, suggesting that TERRA play a stage-specific role in the life cycle of these early-branching eukaryotes.

10.2.2 (SANTOS, Renato ERS et al. Molecular and biochemical parasitology. 2017)

"A DiCre recombinase-based system for inducible expression in Leishmania major". Published: September 2017.

Abstract

Here we present the establishment of an inducible system based on the dimerizable Cre recombinase (DiCre) for controlled gene expression in the protozoan parasite Leishmania. Rapamycin-induced DiCre activation promoted efficient flipping and expression of gene products in a time and dose-dependent manner. The DiCre flipping activity induced the expression of target genes from both integrated and episomal contexts broadening the applicability of the system. We validated the system by inducing the expression of both full length and truncated forms of the checkpoint protein Rad9, which revealed that the highly divergent C-terminal domain of Rad9 is necessary for proper subcellular localization. Thus, by establishing the DiCre-based inducible system we have created and validated a robust new tool for assessing gene function in Leishmania.

10.2.3 (DAMASCENO, Jeziel D. et al. Nucleic Acids Research. 2018)

"Conditional genome engineering reveals canonical and divergent roles for the Hus1 component of the 9–1–1 complex in the maintenance of the plastic genome of Leishmania". Published: 31 October 2018.

Abstract

Leishmania species are protozoan parasites whose remarkably plastic genome limits the establishment of effective genetic manipulation and leishmaniasis treatment. The strategies used by Leishmania to maintain its genome while allowing variability are not fully understood. Here, we used DiCre-mediated conditional gene deletion to show that HUS1, a component of the 9–1–1 (RAD9-RAD1-HUS1) complex, is essential and is required for a G2/M checkpoint. By analyzing genome-wide instability in HUS1 ablated cells, HUS1 is shown to have a conserved role, by which it preserves genome stability and also a divergent role, by which it promotes genome variability. These roles of HUS1 are related to distinct patterns of formation and resolution of single-stranded DNA and γH2A, throughout the cell cycle. Our findings suggest that Leishmania 9–1–1 subunit have evolved to co-opt canonical genomic maintenance and genomic variation functions. Hence, this study reveals a pivotal function of HUS1 in balancing genome stability and transmission in Leishmania. These findings may be relevant to understanding the evolution of genome maintenance and plasticity in other pathogens and eukaryotes.

11.   Awards

2021: "Best Oral Presentation at Cellular and Molecular Biology Symposium/ FMRP