M 2016

U.PORTO

INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR
UNIVERSIDADE DO PORTO

# IMAGING TELOMERES BY CRISPR/Cas9

JOÃO PEDRO CURADO AGRA AMORIM

**DISSERTAÇÃO DE MESTRADO APRESENTADA AO INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR DA UNIVERSIDADE DO PORTO EM ONCOLOGIA**

U.PORTO
INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR
UNIVERSIDADE DO PORTO

JOÃO PEDRO CURADO AGRA AMORIM

# IMAGING TELOMERES BY CRISPR/CAS9

Dissertação de Candidatura ao grau de Mestre em Oncologia – Especialização em Oncologia Molecular submetida ao Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto

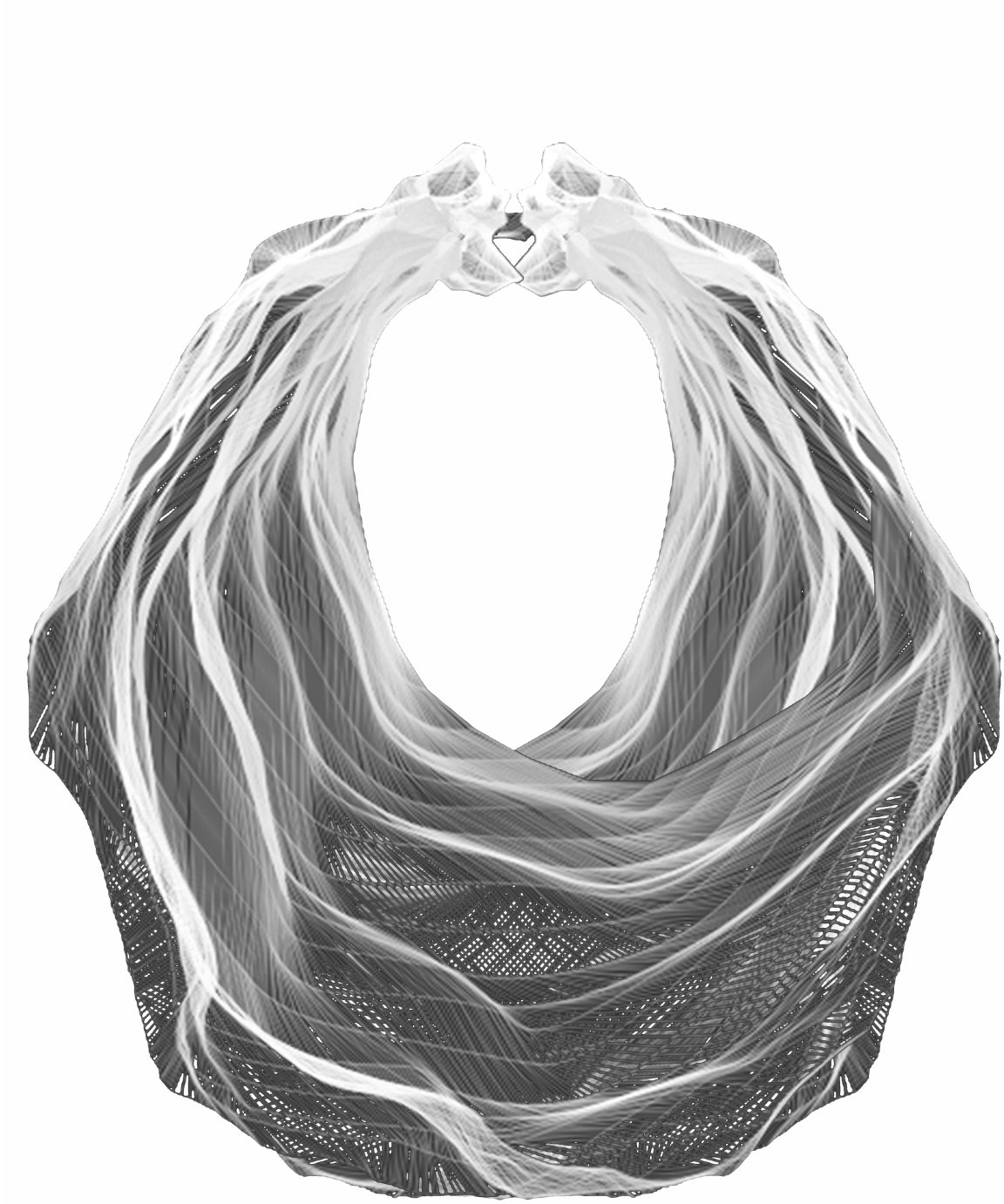
**Orientadora –** Professora Doutora Ana Paula Soares Dias Ferreira

**Categoria –** Professora Auxiliar na Faculdade de Medicina da Universidade do Porto, Coordenadora do Grupo Signaling and Metabolism no Instituto de Investigação e Inovação em Saúde/Instituto de Patologia e Imunologia Molecular da Universidade do Porto

**Afiliações –** Faculdade de Medicina da Universidade do Porto, Instituto de Investigação e Inovação em Saúde/Instituto de Patologia e Imunologia Molecular da Universidade do Porto.

25 ANOS A DESCOBRIR O CANCRO
ipatimup

INSTITUTO DE INVESTIGAÇÃO E INOVAÇÃO EM SAÚDE
UNIVERSIDADE DO PORTO



***OUROBOROS* by Jaime Manso**

**Acknowledgments**

First, I would like to express my gratitude to Drª Paula Soares and João Vinagre for the guidance and help in designing and executing this project. I also thank João for helping me overcome my difficulties in the laboratory material handling and scientific thinking and for often making me laugh.

I am deeply thankful to Raquel Lima, without whom I couldn't have obtained the most important results in this thesis. She was extremely supportive and her help was fundamental for me to learn about various techniques and for my evolution as an undergraduate research student. Also, her patience towards me was inexhaustible.

I thank Gustavo Santos for all his support. His companion and our conversations (regarding both science and other society's issues) were of most importance for my personal and professional well being. His help was also fundamental for the thinking and writing of the introduction of this dissertation.

Then, I would like to thank everyone in the Cancer Signaling and Metabolism group at I3S/IPATIMUP for their support, conversations, meal companion (especially Catarina Tavares and Luciana Bueno) and good work environment.

I place on record the importance of the various people and valences at I3S research institute for the completion of mine and other projects. Without, per example, the work of Paula Magalhães at CCGEN, I couldn't have completed my work.

Lastly, I would like to thank all my friends and family that supported me during this crucial stage of my life. Alexandra's support and companionship was invaluable. The support of my mother and the relaxation I felt when in her house in the middle of the forest were irreplaceable. As was the support of my friends in my hometown who I encountered mostly on weekends. The guys in my two musical projects were amazing in helping me maintain my focus by providing me with a mean to experience freedom and alienation from every day's life. My friend Filipe Nunes Vicente with whom I share the love for science, art and entropy (in the form of random conversations) was also a pillar for my growth in every aspect of life.

INDEX

**RESUMO**

As células neoplásicas humanas são capazes de atingir a imortalidade. Os mecanismos através dos quais isto é conseguido baseiam-se na capacidade destas células de contrariar o encurtamento dos telómeros. Estes são estruturas nucleoproteicas que protegem o término dos cromossomas e consistem em sequências repetitivas $(TTAGGG)_n$ ricas em guanina. A telomerase é uma enzima transcriptase reversa que contém tanto um componente catalítico proteico como um molde de RNA e é a principal responsável pelo alongamento dos telómeros e células somáticas adultas humanas reprimem normalmente a expressão desta enzima. Porém, aproximadamente 90% das células neoplásicas humanas voltam a expressá-la com o objectivo de manter o alongamento dos telómeros, sendo que mutações no promotor da telomerase são a causa genética conhecida mais frequente. Por outro lado 10 a 15% dos cancros não apresentam actividade telomerase detectável e uma parte desses atinge a imortalidade através de um mecanismo de elongação de telómeros independente da telomerase, o ALT (Alternative Lenghtening of Telomeres), que depende do mecanismo de reparação de DNA HR (Homologous Recombination). Mutações inactivadoras foram identificadas nas proteínas ATRX e DAXX e no complexo Histona H3.3 em várias células tumorais consideradas ALT-positivas, maioritariamente em tumores endócrinos do pâncreas, glioblastomas multiformes e oligodendrogliomas. Além disso, o fenótipo ALT é caracterizado pela presença de telómeros de tamanhos heterogéneos, APBs (ALT-associated promyelocytic leukaemia bodies) e recombinação telomérica com presença de repetições teloméricas extra-cromossomais (ECTRs).

O recentemente descoberto sistema de defesa bacteriano CRISPR/Cas9 foi rapidamente convertido numa ferramenta importante na área da engenharia genética, tornando-se imediatamente numa das mais vantajosas técnicas neste campo da Ciência. Consiste na utilização de uma endonuclease Cas9 e um RNA guia (sgRNA) específico para um gene ou loci genómico. Uma vez co-expressos, formam um complexo ribo-proteico que permite que a Cas9 ligue ao alvo genómico e o corte,resultando numa dupla quebra do DNA. Dependendo do mecanismo de reparação de DNA activado em resposta, a técnica permite fazer silenciamento (Knock-Out) de um gene ou edição do genoma. Recentemente, várias modificações à enzima Cas9 permitiram a extensão das aplicações da técnica. A construção de uma Cas9 desactivada – dCas9 – (através da mutação dos domínios catalíticos) que pode ser fundida com outras moléculas como factores de transcrição ou proteínas fluorescentes potenciou parte dessas novas aplicações. Neste trabalho explorámos a possibilidade de usar a dCas9 fundida com a proteína GFP e um sgRNA que guiasse este complexo para repetições teloméricas, com

o objectivo de visualizar e comparar o tamanho dos telómeros, numa linha celular de tumor neuroendócrino do pâncreas (Bon), *in vivo.* Para isso, produzimos lentivírus contendo o DNA dos vectores de interesse e infectámos as referidas células com os mesmos. Conseguimos assim visualizar a fluorescência característica da infeção com dCas9-EGFP mas, como a eficiência da transdução com os lentivírus foi aparentemente muito baixa, não foi possível fazer a seleção dessas células (para obter apenas aquelas que expressam estavelmente esse DNA). Por essa razão, foi também impossível infectar com sucesso as células com os sgRNAs e observar a pretendida fluorescência específica de telómeros. Num futuro próximo pretendemos, no entanto, optimizar o protocolo e atingir o objectivo de estabelecer uma linha celular expressando estavelmente dCas9-EGFP, produzindo assim a ferramenta pretendida.

**ABSTRACT**

Human cancer cells are capable of achieving immortality. The mechanisms by which they achieve it partially rely in being able to counteract the shortening of their telomeres, which are DNA-protein structures that protect chromosome ends and consist of arrays of guanine-rich (G-rich) repeats $(TTAGGG)_n$, Telomerase is a reverse transcriptase enzyme with both a protein catalytic component and a RNA template and is the main responsible for telomere elongation. Human adult somatic cells usually repress the telomerase enzyme expression. However, up to 90% of human cancers cells de-repress its expression to maintain telomere length, telomerase promoter mutations being the known most frequently found event. On the other hand, 10% to 15% of cancers do not present detectable telomerase activity and a part of those achieves immortalization via a telomerase-independent mechanism of telomere lengthening, the alternative lengthening of telomeres (ALT), which depends on a homologous recombination (HR) DNA-repair mechanism. Inactivating mutations were identified in the ATRX, DAXX and Histone H3.3 complex in ALT-positive tumor cells, mainly in pancreatic endocrine tumors, glioblastoma multiforme and oligodendrogliomas. Besides that, the ALT phenotype is characterized by the presence of heterogeneous telomere lengths, ALT-associated promyelocytic leukaemia bodies (APBs) and telomere recombination with the presence of extrachromosomal (linear and circular) telomeric repeats (ECTRs).

The recently discovered bacterial defence system CRISPR/Cas9 was readily converted into a tool in the area of genome engineering, turning immediately into one of the most advantageous techniques to be used in this field of science. It consists in the presence of a Cas9 endonuclease and a guide RNA (sgRNA) targeting a specific gene/genomic loci. Once co-expressed, they form a riboprotein complex allowing Cas9 to ligate the target and perform cleavage, resulting in a DNA Double Strand Break. Depending on the mechanisms of DNA repair that is then activated, this technique allows performing KO's or genome editing. More recently, various modifications to the Cas9 enzyme have extended its applications. The generation of a deactivated dCas9 (by mutationally inactivating its two nuclease domains) that could be fused to other molecules such as transcriptions factors or fluorescent proteins made it clear that it could be used for various purposes.. Altogether, this stressed the possibility of using dCas9 fused to the GFP protein and a sgRNA guiding to a telomeric repeat sequence to visualize and compare the telomeric lengths and their observable appearance in Bon pancreatic neuroendocrine tumor cells, *in vivo*. In this project, we produced lentivirus harbouring the necessary vectors and transduced the interest cells with those. We were able to confirm the characteristic fluorescence of the Tet3G/dCas9 infection but, as the transduction

efficiency was low, we could not select cells stably expressing the vectors. Therefore, it was impossible to successfully transduce the cells with the sgRNAs and produce the telomere-specific fluorescence that we aimed to see. In the near future, we intend to optimize this protocol until we reach our goal of establishing a clonal cell line stably expressing dCas9-EGFP.

**List of Figures**

**Supplementary Data**

**List of Tables**

**Abbreviations**

| | |
|---|---|
| ADD | ATRX-DNMT3-DNMT3L domain |
| ALT | Alternative Lengthening of Telomeres |
| APB | ALT-associated promyelocytic leukemia body |
| ATRX | Alpha thalassemia/mental retardation syndrome X-Linked protein |
| BIR | Break-induced repair |
| BLM | Bloom syndrome RecQ like helicase |
| Bon | Human carcinoid cell line |
| bp | Base pair |
| Cas1 | CRISPR-associated protein 1 |
| Cas2 | CRISPR-associated protein 2 |
| Cas3 | CRISPR-associated protein 3 |
| Cas9 | CRISPR-associated protein 9 |
| Cas10 | CRISPR-associated protein 10 |
| $CO_2$ | Carbon dioxide |
| CRISPR | Clustered regularly interspaced short palindromic repeats |
| crRNA | CRISPR RNA |
| DAXX | Death-domain associated protein |
| dCas9 | Deactivated CRISPR-associated protein 9 |
| dCas9-EGFP | Template for NLS-dCas9-NLS-EGFP fusion protein for CRISPR imaging |
| DNA | Deoxyribonucleic acid |
| DSB | Double strand break |
| dsDNA | Double strand DNA |
| DMEM | Dulbecco's Modified Eagle Medium |
| Dox | Doxycycline |

| | |
|---|---|
| DMEM/F-12 | Dulbecco's Modified Eagle Medium/Nutrient F-12 Ham |
| ECL | Enhanced Chemiluminescence substrate |
| ECTR | Extra-chromosomal telomeric repeat |
| EDTA | Ethylenediaminetetraacetic acid |
| EGFP | Enhanced green fluorescent protein |
| EST-1 | Even short telomeres 1 yeast line |
| FBS | Fetal bovine serum |
| GFP | Green fluorescent protein |
| G4 | G-quadruplex |
| Hek293FT | Highly transfectable clonal isolate derived from human embryonal kidney (HEK) cells transformed with the SV40 large T antigen |
| HGT | Horizontal gene transfer |
| HindIII | Restriction endonuclease isolated from *Haemophilus influenza* |
| HNH | HNH endonuclease |
| HR | Homologous Recombination |
| H3.3 | Histone 3 variant |
| H3K9m3 | Histone H3 trimethylated Lysine 9 |
| H3 K4m0 | Histone H3 unmodified Lysine 4 |
| Iap | Inhibitor of apoptosis |
| InDels | Small nucleotide insertions or deletions |
| KD | Knockdown |
| kDa | KiloDalton |
| KO | Knock-out |
| LB | Luria Broth, powder microbial growth medium |
| MRE11 | MRE11 homolog A, double strand break repair nuclease |
| MRN | MRE11-RAD50-NBS1 complex |

| | |
|---|---|
| MUS81 | Structure-specific endonuclease subunit |
| NBS1 | DNA repair and telomere maintenance protein nbs1 |
| NHEJ | Non-homologous end joining |
| NLS | Nuclear localization signal |
| OptiMEM | Reduced Serum Media |
| ORF | Open reading frame |
| PAM | Protospacer adjacent motif |
| pMD2.G | VSV-G envelope expressing plasmid |
| PML | Promyelocytic Leukemia |
| PHD | Plant homeodomain |
| psPAX2 | 2nd generation lentiviral packaging plasmid |
| NheI | Restriction endonuclease isolated from *Neisseria mucosa heidelbergensis* |
| qRT-PCR | Quantitative real-time polymerase chain reaction |
| RIPA | Radio-immunoprecipitation assay buffer |
| RAD50 | RAD50 double strand break repair protein |
| RAD52 | RAD52 homolog, DNA repair protein |
| rDNA | Ribossomal DNA |
| RNA | Ribonucleic acid |
| RNAi | RNA interference |
| RuvC | Crossover junction endodeoxyribonuclease |
| SDS/PAGE | Sodium dodecyl sulfate/Polyacrylamide gel electrophoresis |
| sgGAL4 | Human pSico-based U6 vector containing murine U6 promoter and sgRNA targeting GAL4 UAS promoter (negative control) |
| sgRNA | Short guide RNA |
| sgTelomere | Lentiviral vector that contains an optimized *S. pyogenes* sgRNA targeting human telomeres |

| | |
|---|---|
| SGTB | Simon Gonçalo Timoteo Buffer |
| shRNA | Short-hairpin RNA |
| SMC5/6 | Structural Maintenance of Chromosomes 5/6 complex |
| SP100 | Sp100 nuclear antigen |
| Snf2 | Sucrose non-fermenting |
| TALEN | Transcription activator-like effector nucleases |
| Tet3G | Tetracycline inducible expression systems (Tet-On 3G) transactivator vector |
| tracrRNA | Trans-activating crRNA |
| Tris-Hcl | Tris hydrochloride |
| T-SCE | Unequal telomeric sister chromatid exchange |
| U-2 OS | Human osteosarcoma cell line |
| ZFN | Zing finger nucleases |

# INTRODUCTION

## INTRODUCTION

Through life, human beings experience the effects of ageing, ultimately leading to death. However, even within mortal beings, there are cells that can achieve immortality: the cancer cells. The mechanisms by which these cells become immortal partially rely in being capable of counteracting the shortening of their telomeres, the universal biological clock.

### Telomeres and immortalization

In the first half of the 20$^{th}$ century, it was discovered that the chromosomes, localized in the cell nucleus, carry the genetic information (Thomas H. Morgan; Nobel Prize in Physiology or Medicine 1933). A decade later, Herman Muller and Barbara McClintock discovered that broken chromosomes were unstable, i.e. prone to rearrangements and/or fusions [1, 2]. However, chromosome ends seemed to be protected from such events. The special nature of chromosome ends were described and they were named telomeres (from the Greek words “telos” (end) and “meros” (part)) by Muller for the first time [1]. Currently, it is known that telomeres are DNA-protein structures that protect chromosome ends and consist of arrays of guanine-rich (G-rich) repeats $(TTAGGG)_n$, in vertebrates.

The discovery of DNA polymerases was of extreme importance in order to understand chromosome dynamics. They were demonstrated to be dependent on a primer for enzyme coupling in order to replicate DNA (Severo Ochoa and Arthur Kornberg; Nobel Prize in Physiology or Medicine 1959) and, in 1972, James Watson observed that it was theoretically impossible that both strands could be synthesized in these conditions [3]. This observation was denominated the end-replication problem (Figure 1). By this time, Leonard Hayflick had introduced two novel concepts in science (“The Hayflick Limit”): normal cells in culture are not immortal *and* there is a memory that allows cells to undergo a specific number of population doublings (what we call nowadays the biological clock) [4]. In 1973, Alexsei Olovnikov reported that newly formed DNA would have an incomplete strand at some point in time and made the connection with the facts reported by Hayflick, hypothesizing that the incomplete replication could be the basis for the chromosome shortening observed in cultured cells overtime, culminating in reduced viability [5].

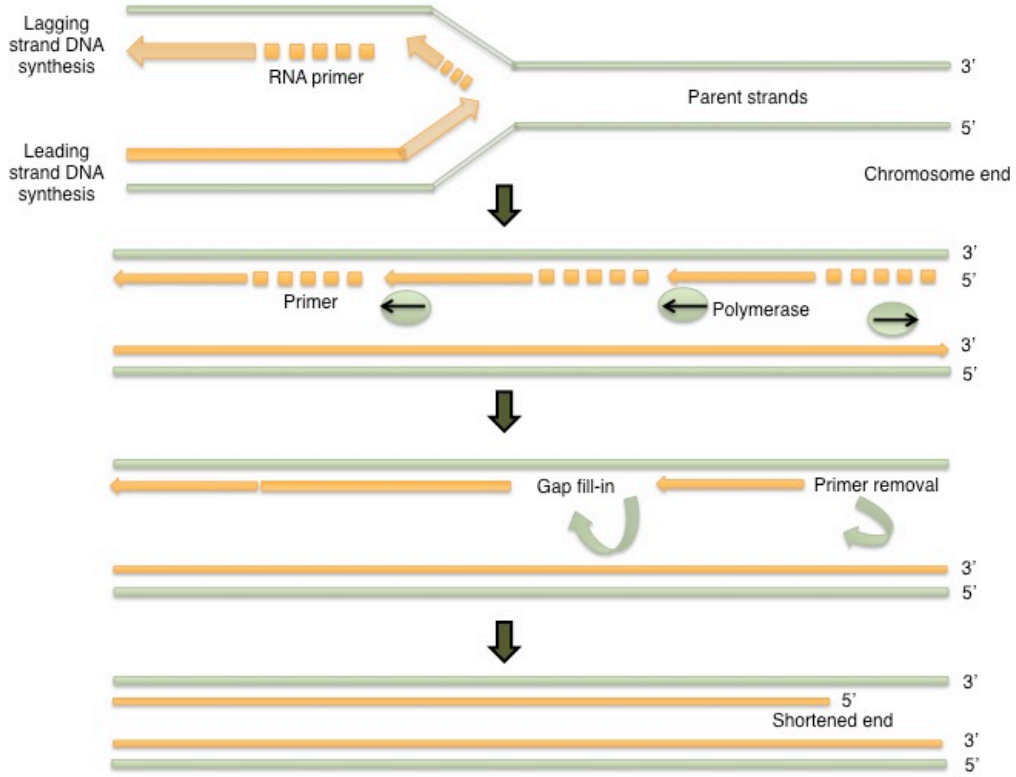


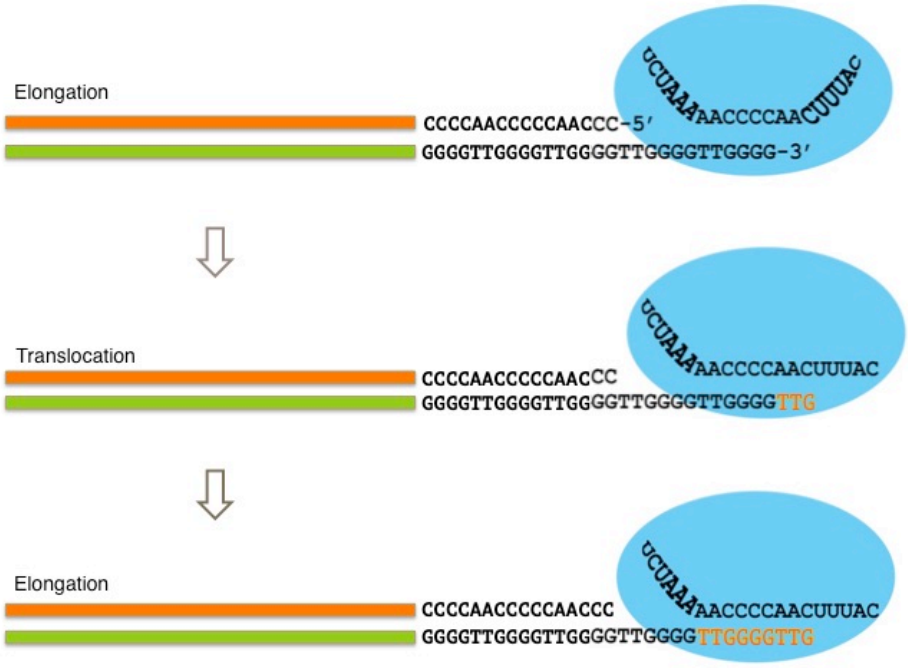

**Figure 1** – The end-replication problem. Copying of the lagging strand of a DNA molecule by the DNA polymerase enzyme occurs gradually. The process is dependent on an RNA primer that is degraded immediately after. The gap is filled and ligated to produce a new semi conserved DNA molecule. DNA polymerases are not able to fill the gap in the end of the chromosomes. This results in these becoming shorter upon each round of replication. Adapted from "The 2009 Nobel Prize in Physiology or Medicine – Advanced Information". *Nobelprize.org.* Nobel Media AB 2014. Web. 13 Jun 2016. http://www.nobelprize.org/nobel_prizes/medicine/laureates/2009/advanced.html

By the late 70s, Jack Szostak was working on the unravelling of the mechanisms occurring in the homologous recombination process, in yeast. He found that the ends of linear plasmid DNA molecules introduced into yeast cells were highly recombinogenic. This, as he assumed, would lead to rearrangements or interaction with homologous chromosomal DNA and integration. However, the observation was made that long-term maintenance was not possible [6]. These findings were in line with the early observations by Muller and McClintock (mentioned above) that broken chromosomes are highly unstable.

Elizabeth Blackburn and Jack Szostak attended the same Gordon conference in 1980, where they discussed the astonishing properties of telomeres and, finally, decided to collaborate. They investigated if adding *Tetrahymena* chromosome end repeat DNA sequences to linear yeast plasmids could stabilize the ends to allow replication. The most striking conclusion from these experiments was that the evolutionarily conversed functional properties of the *Tetrahymena* rDNA end sequences provided chromosomal stability. Also, there were functionally similar sequences in yeast. All this thus allowed to deduct that such end sequences corresponded to functional telomeres [7].

At this point, Carol Greider had joined Elizabeth Blackburn laboratory and they started investigating the mechanisms underlying the synthesis of these telomeric structures. Greider developed a biochemical assay using *Tetrahymena* extracts, where synthetic DNA oligonucleotides complementary to the telomeric repeats were used as a primer. Using this technique, she showed addition of TTGGGG repeats *in vitro*, leading to the notion that there was enzymatic activity present in these cell extracts [8]. In their following work, Greider and Blackburn finally presented the term telomerase and began demonstrating that this enzyme is a ribonucleoprotein dependent on both its protein and RNA components. About the RNA component, whose function was still not fully understood at the time, the two scientists wrote: *"It is tempting to speculate that the RNA component might be involved in determining the sequence of the telomeric repeats that are synthesized and/or the specific primer recognition. If the RNA of telomerase contains the sequence CCCCAA, this sequence could act as an internal guide sequence"*. As reported by Greider and Blackburn in the following year, the RNA component was indeed found to contain this template sequence. Also, the two scientists interestingly found that blocking of this motif availability compromised telomerase enzymatic activity [9]. The role of the CAACCCCAA RNA sequence as a template for telomere synthesis was confirmed when Blackburn's group showed that mutations in this sequence resulted in corresponding alterations in the DNA sequence being synthesized [10]. This data allowed to confirm the role of telomerase as a reverse transcriptase with a protein catalytic component as well as a RNA template for telomere elongation. A new model for the elongation of telomeres was thus presented by Greider and Blackburn (Figure 2): First, telomerase recognizes *Tetrahymena* chromosome ends containing a 13-base overhanging TTGGGG sequence, leading to hybridization between the complementary nucleotides. TTG nucleotides are then added stepwise and a translocation of the telomerase enzyme occurs as it is relocated to the 3'end of the TTGGGG strand such that the 3' TTG nucleotides are hybridized to the enymatic RNA component. Finally elongation occurs, copying the template sequence to complete the TTGGGGTTGG sequence. This

mechanisms explains how oligonucleotides with 3'ends terminating at any nucleotide with the




sequence TTGGGG are elongated to yield perfect tandem repeats of $(TTGGGG)_n$ [9].

**Figure 2** – A model for telomere elongation. Adapted from Greider and Blackburn, 1989 [9].

Human adult somatic cells usually repress the telomerase enzyme expression, although it continues to be expressed in proliferative cells (germ cells and tissue stem cells) [11]. Up to 90% of human cancers cells de-repress its expression to maintain telomere length (critical for tumor genomic stability and immortality) [11], telomerase promoter mutations being the most frequently found known event [12]. However, 10% to 15% of cancers do not present detectable telomerase activity. A subset of those (approximately 5%) achieves immortalization via a telomerase-independent mechanism of telomere lengthening, the so-called alternative lengthening of telomeres (ALT) that, as will be further addressed, depends on a homologous recombination (HR) DNA-repair mechanism to maintain telomere length [13].

Cancers that have a mesenchymal origin are reported to activate ALT more frequently, while epithelial cancers rely more frequently in telomerase reactivation/re-expression [13, 14]. As mesenchymal stem cells are known to express minimal or no detectable amounts of telomerase [15], this may predispose cells from this lineage to depend on ALT activation more frequently. Genome mapping by the somatic cell

hybridization technique had already demonstrated that ALT activation occurred through the loss of one or more repressor molecules present in normal somatic telomerase positive cells [16]. More recently, with the appearance of mass genome sequencing techniques, inactivating mutations were identified in the ATRX, DAXX and Histone H3.3 complex in ALT-positive tumor cells, mainly in pancreatic endocrine tumors [17], glioblastoma multiforme and oligodendrogliomas [18]. Moreover, it has been pointed that the appearance of the ALT phenotype can, in part, be a consequence of *ATRX* inactivating mutations [19].

**Alternative Lengthening of Telomeres**

In cancers where no detectable telomerase activity is found, telomere length is frequently assured by an homologous recombination (HR) mechanism, ALT [13]. The first observation of this phenomenon was a telomerase knockout model in yeast (even-short telomeres 1 - EST1), where it was discovered that this cells could maintain telomere length without telomerase activity. However, if a double knockout for RAD52 protein was performed, the cells did not survive [20]. RAD52 is a critical protein for the double strand break repair by HR and these experiments enlightened the central basis behind ALT mechanism. Moreover, studies using exogenous DNA integration into telomeric regions with tagged DNA were able to demonstrate that, in cells bearing the ALT phenotype, transference from one telomere to another occurs, subsequent to cell divisions [21]. It is important to refer that the ALT phenotype is characterized by the presence of: heterogeneous telomere lengths; ALT-associated promyelocytic leukaemia bodies (APBs), that differ from common promyelocytic leukaemia (PML) bodies found in other cell types by the inclusion of telomeric DNA and numerous specific recombination factors [22]; and telomere recombination with the presence of extrachromosomal (linear and circular) telomeric repeats (ECTRs) [23]. These features are consistent with hyperactive HR and the extrachromosomal telomeric DNA has been proposed to serve as a template for the extension of telomeres [24].

**Mechanisms for ALT**

Various proposed models describe how telomeres are maintained and extended without telomerase enzymatic activity (Figure 3) [25]. The first (Figure 3A) represents a model based in break-induced repair (BIR), a mechanism where DNA is synthetized away from a break site using HR donor template that, in this case, is telomeric DNA. The models 1 and 2 in Figure 3A differ in the timing of the lagging strand and, in both, there is loss of material in recipient telomeres. The model 3 represents a unidirectional replication

fork that will require a Holliday junction in the end of the process and, in this case, both donor and recipient telomeres will be semi-conserved [26]. Unequal telomeric sister chromatid exchange (T-SCE) has also been suggested as a mechanism for ALT (Figure 3B). This is supported by the fact that T-SCE is elevated in ALT+ cells as measured by CO-FISH [27]. Finally, the recurrent presence of linear and circular extrachromosomal telomere DNA (ECTRs) in ALT cells suggests that these molecules can be used as templates for HR-driven telomere elongation or simply undergo rolling-circle replication (Figure 3C) [25].

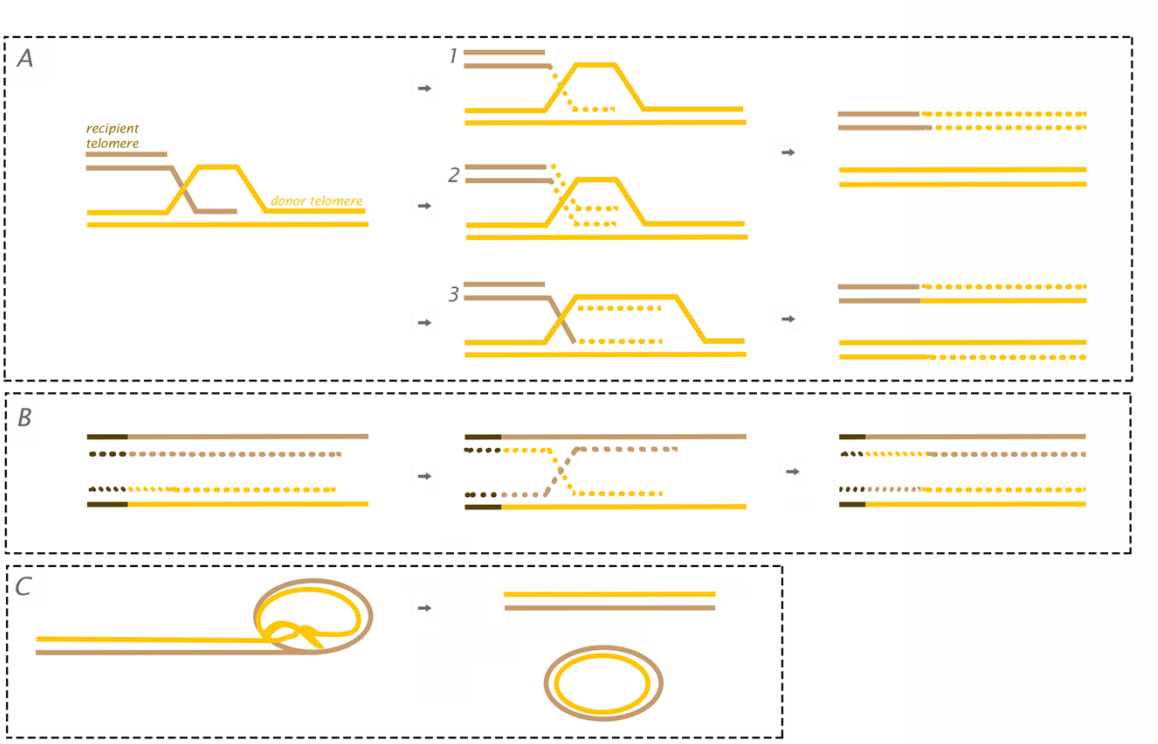



**Figure 3 -** Theoretical models for ALT mechanism. A) Break-induced replication loss from the donor telomere (1 and 2 without loss from the donor telomere). In model 3 there is a unidirectional replication fork with semi-conservative replication of the telomeres. B) Unequal telomeric sister chromatid exchange (T-SCE) model. C) Rolling circle and t-circle formation after t-loop resolution providing a linear double strand break for subsequent HR-mediated activation into the homologous templates. Adapted from Durant, ST 2012 [25].

### ATRX protein role in ALT

ATRX gene is localized in chromosome Xq21.1 and it encompasses 37 exons [28, 29]. This gene encodes a chromatin remodeler which is a 280 kDa protein that includes an unusual N-terminal plant homeodomain (PHD), designated the ATRX-DNMT3-DNMT3L (ADD) domain [30-32]. At the C terminus it presents seven helicase subdomains that confer ATPase activity and identify ATRX as a Snf2 family member of chromatin-

associated proteins [33]. As previously mentioned, it is known that ATRX, in collaboration with its partner DAXX, functions as a complex for the deposition of the histone variant H3.3 into telomeric and pericentromeric chromatin. DAXX is a highly specific histone chaperone that is able to discriminate H3.3 from the other major histone variants, whereas ATRX is involved in targeting DAXX to repetitive sequences enhancing histone deposition[34, 35].

Immunofluorescence studies have demonstrated that ATRX has a preference for binding within PML bodies and also to repetitive heterochromatic regions such as ribossomal DNA (rDNA), telomeric and pericentric DNA repeats [36-38]. The localization of ATRX to heterochromatin involves the interaction with the heterochromatin protein 1 (HP1) [39-41]. Also, targeting of ATRX to heterochromatin is dependent on its interaction with histone H3 trimethylated Lys9 (H3K9m3) and unmodified Lys4 (H3 K4m0) [42]. An hypothesis is that HP1 might serve as a protein scaffold, facilitating the recruitment of ATRX to the heterochromatin through binding of H3K9me3 via its N-terminal chromodomains [42, 43] (Figure 2). Still, this recruitment seems to be of extreme complexity and rely on multiple interactions between effector proteins and chromatin.

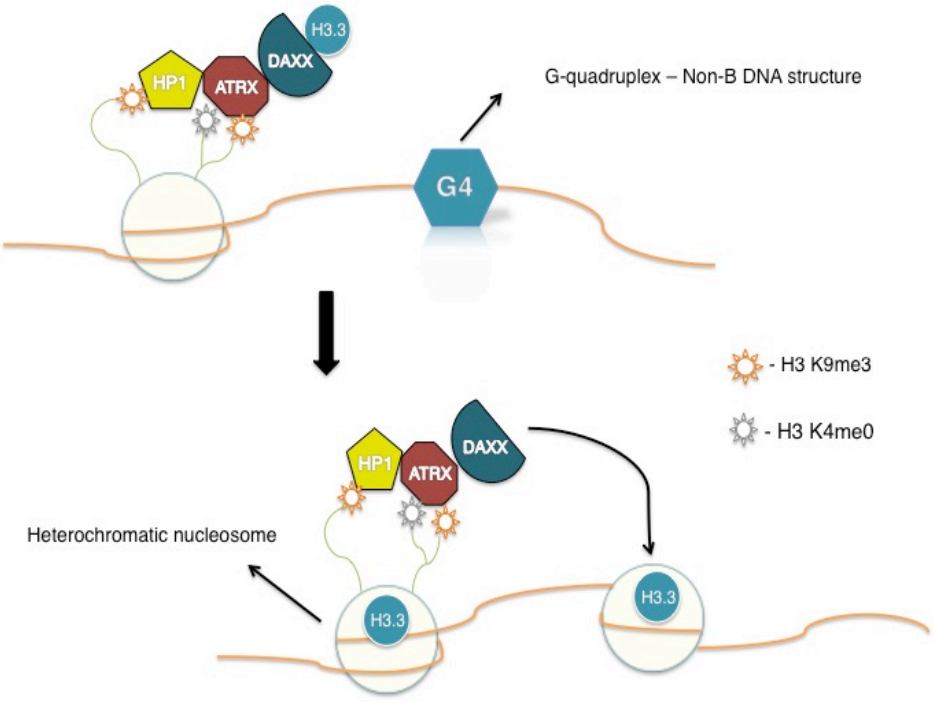



**Figure 4:** A proposed theoretical mechanism for ATRX/DAXX chromatin landscaping: ATRX binds to histone H3 at heterochromatin through interaction of its ADD domain with a H3 N-terminal tail, trimethylated at Lys9 (H3K9me3) and unmodified at Lys4 (H3K4me0). HP1 also recognizes H3K9me3, boosting this recruitment. Once in its target site, ATRX, in

combination with DAXX, facilitates the deposition of the histone variant H3.3. Adapted from Clynes et al. 2013 [43].

It is reported that heterochromatic regions are prone to DNA damage in the absence of the ATRX protein [44, 45]. This could mean that it plays a role in one of the two DNA DSB (Double Strand Break) repair mechanisms (Non Homologous End Joining – NHEJ in post mitotic cells; Homologous Recombination - HR in proliferative ones). However, exposing ATRX-KO cells to irradiation led to the suggestion that the protein is not required for repair of induced DSBs and, therefore, its role might be of protecting cells from replicative stress instead of repairing it [45, 46]. Interestingly, and in contrast with the latter, a more recent approach to this question by Koschmann, C. et al., brought new insights regarding the role of ATRX in DNA-repair pathways. The authors used an animal model of *ATRX*-deficient glioblastoma to uncover the impact of this protein loss in tumor proliferation and loss of genetic stability. They found out that the reduction of ATRX levels impaired NHEJ pathway, as transfection with a DNA-damaged GFP (designed to be restored by NHEJ) simultaneous to ATRX silencing reduced NHEJ function by 50% [47]. These findings suggest that a relative increase in HR in relation to NHEJ in *ATRX*-deficient gliomas may be supportive of ALT. Also, this data is explicative of the genetic instability, and altered DNA damage response in ALT cell lines [44, 46, 47], besides providing a molecular target that can be explored to design new targeted therapies regarding ALT-positive gliomas.

Telomeres composition of repetitive DNA with an high content of guanine and cytosine residues leads to the formation of secondary DNA structures, the G4-DNA [48]. These are known to form obstacles (like bulky DNA adducts) to multiple nuclear processes including DNA replication and transcription. A large proportion of ATRX target sites are predicted to adopt non-B DNA structures like these G-quadruplexes [49], leading to the concept that this protein might be responsible for aiding in the replication in presence of such structures or by preventing their formation, therefore facilitating replication at telomeres. A recent study suggests that ATRX does not appear to possess G4-DNA unwinding activity and that it might overcome these impediments indirectly, perhaps by favouring the histone H3.3 deposition to maintain DNA in a B-form conformation (Figure 2) or by promoting a fork bypass via template switching [50, 51]. Furthermore, a potential consequence of loss of ATRX function is an increased frequency of G4-DNA, causing DNA damage [52]. So, it is likely that the presence of such structures in telomeric DNA favours the ALT phenotype by presenting a barrier to the replication fork, causing fork stalling, collapse and subsequent restart by the homologous recombination DNA-repair mechanism [51].

Depletion of proteins or protein-complexes that are involved in HR and/or are required for replication fork restart like MRE11-RAD50-NBS1 (MRN complex), SMC5/6 complex, MUS81 or the BLM helicase suppress telomeric recombination in ALT cells [53, 54]. Particularly, the MRN complex functions includes DSB repair (via HR or NHEJ) and the restart of stalled replication forks [55] and, like ATRX, it co-localizes to telomeres during the S and G2 phases of the cell cycle [56]. Also, G-quadruplexes were found to be a preferred substrate for the MRX (MRN yeast homologue) complex-component MRE11 [57]. The association of MRN complex with PML bodies (constituting the APBs) seems to be a requirement for ALT activity and likely constitutes the sites of telomeric recombination [21]. This data shows that the proteins in the MRN complex might have an important role in the ALT mechanism as they interact with ATRX, PML bodies and telomeres.

Recently, it was demonstrated that ectopic expression of ATRX in an ALT-positive cell line significantly reduced replication fork stalling, likely limiting the substrate for HR and downregulating ALT [51]. Also, the same authors had already revealed that ATRX interacts with endogenous MRN complex components in a direct manner during DNA replication, suggesting that this interaction is likely to play a direct role in facilitating this process [50]. Moreover, the re-expression of ATRX in ALT cell lines resulted in a dramatic redistribution of the MRN complex away from telomeric DNA and PML nuclear bodies [51]. Overexpression of a PML nuclear body component, Sp100, also showed sequestration of MRN components away from APBs leading to suppression of ALT [53]. This suggests that ATRX expression in a cell with the ALT phenotype may limit ALT by redistributing MRN away from sites of telomeric recombination. The fact that G4-DNA are a substrate for MRE11 raises the possibility that MRN directly cleaves these structures triggering DBS formation and HR [51].

All the available data has allowed to propose that the lack of ATRX is linked to ALT by two main mechanisms, which can act independently or not: i) without this protein there is an increase in stalled replication forks due to G4-DNA structures that promote HR and ii) the redistribution of the MRN complex to APBs may also promote this DNA-repair mechanism.

### ATRX loss promotes ALT in mesenchymal cells

A recent study evaluated the effect of knocking down ATRX in fibroblasts and epithelial cells, to determine whether loss of this gene promoted the ALT mechanism [58]. They found out that ATRX depletion in epithelial cells does not promote ALT activation. However, in fibroblasts, ATRX knock down led to an increased ALT frequency [58].

Embryonic stem cells are known to have telomerase activity. However, mesenchymal stem cells express little or no detectable telomerase [59] and this may be the reason why these cells do not activate telomerase in a similar frequency to epithelial cells. This suggests that telomerase expression may be repressed by continuous chromatin-mediated repression at mesenchymal lineage. Another possible explanation may be that epithelial cells express proteins that mesenchymal cells do not and one or more of those might have a role in helping to overcome G4-DNA structures and/or chromatinization.

Although ATRX loss promotes ALT, it seems it is not enough to activate this alternative mechanism, even in fibroblast cell lines. In a recent study, knocking down ATRX in fibroblasts was capable of activating ALT in ten of twelve cultures [58]; one of the six controls activated ALT. This leads to the concept that loss of ATRX is not enough to activate ALT and the presence of ATRX is not enough to avoid ALT. This data suggests that there is a missing link in this mechanism, which can be other protein(s) or (epi-) genetic changes that can interact with ATRX.

**ATRX expression suppresses ALT**

It has been proposed that ATRX acts as a tumor suppressor. A study conducted in telomerase-negative U-2 OS ALT cells, which are also null for ATRX, revealed that re-expression of ATRX in these cells suppresses the ALT phenotype and, if that re-expression is switched off, it is acquired again. However, if G-quadruplexes are stabilized (using pyridostatin), ALT suppression by ATRX is prevented. This study also revealed that this suppression is dependent on the histone chaperone DAXX as re-expression of ATRX in ALT cells concomitant with DAXX KO was not capable of suppressing the ALT phenotype [51].

Another study, conducted in three different ALT mesenchymal cell lines that lack ATRX expression, showed that ATRX re-expression represses the ALT mechanism [58], even though the cells had to enter crisis before the ALT phenotype emerged. The explanation for this event, as referred before, might be the fact that ATRX facilitates replication and resolution of the G-quadruplex secondary structures, not allowing replication fork stalling, subsequent DSB formation and activation of the HR mechanism (the latter is a known hallmark of ALT). Another possible explanation is that ATRX helps in the formation of telomeric chromatin, not allowing the formation of DNA secondary structures like G-quadruplexes, therefore potentiating the normal mechanism of telomere replication, with no intervention by the HR-mediated repair of DNA.

Lastly, it is pertinent to say that, as the results of a recent study show, loss of ATRX also suppresses resolution of cohered telomeres as, in the absence of this protein,

the histone variant macroH2A1.1 binds to the poly(ADP-ribose) polymerase tankyrase 1, preventing it from localizing to telomeres and resolving cohesion (as cohesion resolution is the known function of this specific polymerase) [60]. The persistent cohesion that follows loss of ATRX thus promotes recombination between sister telomeres (T-SCE), a known ALT hallmark. This data suggests that this phenomenon indeed controls recombination in ALT cells and points to a role of ATRX in the ALT phenotype, in addition to its DAXX/H3.3 histone chaperone function. Even though this does not resolve the question of whether ATRX has an unmistakable role in activating ALT, the possibility is raised that the precondition of persistent telomere cohesion might favour activation of the ALT pathway over the upregulation of telomerase during tumorigenesis [60].

**Genome engineering**

The capacity to engineer biological systems and organisms or to specifically alter genomes holds extreme potential for application in the areas of basic science, medicine and biotechnology. Per example, KO experiments have been one of the most informative approaches to unsolved problems or questions in science. Also, genome engineering techniques have been discovered and ameliorated to a point humanity would never have imagined, decades ago. These have even been extended to other purposes than the originally thought: imaging of specific genomic regions through genome engineering of specific cells/organisms would have never been possible without the discovery of proteins such as GFP (or more recently, CRISPR associated protein 9 (Cas9)) that occur naturally as they are expressed by certain species. These findings were thus readily converted in tools that might help solve the intrinsic questions that follow consciousness.

There are now several known programmable sequence-specific endonucleases that permit precise editing of endogenous genomic loci in a broad range of species. The discovery of these enzymes led to the arising of a number of genome editing technologies in the recent years, including zinc-finger nucleases (ZFN) [61-64], transcription activator-like effector nucleases (TALEN) [64-68] and the RNA-guided Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) - Cas nuclease system [69]. The first two rely on a strategy of tethering endonuclease catalytic domains to modular DNA-binding proteins for inducing targeted DNA DSBs at specific genomic loci. By contrast, *S. Pyogenes* Cas9 (component of bacterial CRISPR-Cas Type II) is a nuclease guided by small RNAs through base pairing with target DNA. The later represents the most recently and advantageous designed technology. Amongst Cas9 advantages over ZFN and TALEN are: the rapid customization (the enzyme can be easily retargeted to new DNA sequences by using a new pair of oligos encoding the 20-nt guide sequence), higher

targeting efficiency (contrasting with TALENs, that cleave non-specifically, SpCas9 performs a blunt cut 3bp 5' of the protospacer adjacent motif (PAM) sequence and the multiplex genome editing possibility (by co-delivering a combination of small guided RNAs (sgRNAs) to the cells of interest) [70].

## Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR/Cas)

Bacteria and Archaea are frequently exposed to stresses such as infection from bacteriophages and other foreign genetic elements. The constant exposure to this exogenous DNA often results in Horizontal Gene Transfer (HGT), which can be mediated by transduction, conjugation and transformation of mobile elements such as plasmids [71]. This exposure eventually forced microbes to establish an array of defence mechanisms that allow the cell to recognize and distinguish "foreign" from "self" DNA and to survive exposure to these invasive elements. Occasionally, they even thrive in hostile and competitive environments. These mechanisms maintain genetic integrity; yet occasionally allow exogenous DNA uptake and conservation of genetic material advantageous for adaptation to the environment [72].

Bacteria possess multiple methods to regulate genetic flux and resist phage infection. These include the mutation or masking of cell surface receptors, restriction-modification, abortive infection and the Clustered Interspaced Short Palindromic Repeats (CRISPR) systems [71] which is a widespread mechanism that equips bacteria with a sequence-specific heritable acquired immunity against viruses and plasmids [73].

### History

In 1987, Ishino and his collaborators detected five 29 bp repeats with 32 bp spacers near the Inhibitor of apoptosis (*Iap*) gene of *Escherichia coli* [69]. This was the first time interspaced palindromic repeat sequences were observed. Following this, other reports appeared for *Mycobacterium tuberculosis* [74], *Haloferax spp* [75] and *Archaeoglobus fulgidus* [76]. Later, Mojica and co-workers determined structure and sequence similarity as well as phylogenetic distribution of those repetitive elements [77], showing that CRISPR arrays display a high degree of homology between phylogenetically distant species and a wide distribution in bacteria and archaea.

At first terminology was inconsistent but, to simplify, the name Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) was proposed by Ruud Jansen and collaborators [78]. The same authors identified the four genes closely located to the CRISPR array, which were termed CRISPR-associated (*Cas)* genes 1-4 [78], this being a crucial step in understanding the role of this system.

Surprisingly, in 2005, three independent studies showed that spacers matched sequences of extrachromosomal origin, including phages, prophages and plasmids [79-81]. In combination with detailed bioinformatics analyses of the Cas proteins, it was hypothesized that CRISPR/Cas systems provided an RNAi-like mechanism of resistance against invading genetic elements for bacteria and archaea [82, 83]. Direct evidence confirming this hypothesis was later provided by two studies; Barrangou *et al.* found that, after viral challenge, bacteria integrated new spacers derived from phage genomic sequences and, removal or addition of particular spacers, modified the phage-resistance phenotype of the cell [84]. Subsequently, Marraffini and Sontheimer showed that CRISPR/Cas systems can also prevent both conjugation and transformation of plasmids in *Staphylococcus epidermidis* [85].

## CRISPR components

CRISPR Arrays: Repeats, Spacers and Leader Sequence

A single genome can harbour more than one CRISPR array and these vary in size. Within arrays, repeat sequences alternate with spacers [69]. Repeats are typically identical in length and base sequence amongst themselves [77]. They are normally between 23 and 47 bp in length and those of the same subtype share a consensus sequence [86].

Spacer sequences are frequently unique in a genome. A large number of spacers match sequences originating from extrachromosomal sources (such as bacteriophages or plasmids) [79-81], thus conferring the sequence-specific immunity against those. Their length can vary throughout an array and lengths up to 72bp have been reported [87]. Surprisingly, in some organisms, spacers matching sequences within their own genome are found. Their function is unknown but it has been proposed that they are accidentally incorporated [88].

The leader sequence is the final CRISPR array component, localizing upstream of the first repeat [78]. This is an AT-rich (200-500bp long) sequence and it includes the promoter necessary for transcription [89-91].

Cas proteins

Located near the CRISPR array are the genes encoding Cas proteins [78]. These enzymes are necessary for the acquiring of new spacers and targeting of invading elements. CRISPR/Cas systems are generally classified into types I, II and III based on the phylogeny and presence of specific Cas proteins [86]. Cas1 and Cas2 are found

across the three types and Cas1 is considered the universal maker of these systems. Cas3, Cas9 and Cas10 are specific for types I, II and III, respectively [86].

### General mechanism

As referred, CRISPR represents a family of DNA repeats found in most archaeal (~90%) and bacterial (~40%) genomes [78, 92, 93]. It uses small non-coding RNAs for defence and functions together with Cas proteins. Its mechanism involves three phases (Figure 5): The first is the integration of short sequences from "foreign" genetic elements (known as spacers) into repetitive "self" genetic elements (known as CRISPR arrays). Then, these are transcribed and processed into small RNAs (crRNAs) by Cas proteins. Finally, a Cas protein complex together with crRNAs mediates targeting of the invading element, interfering with their nucleic acids [71].

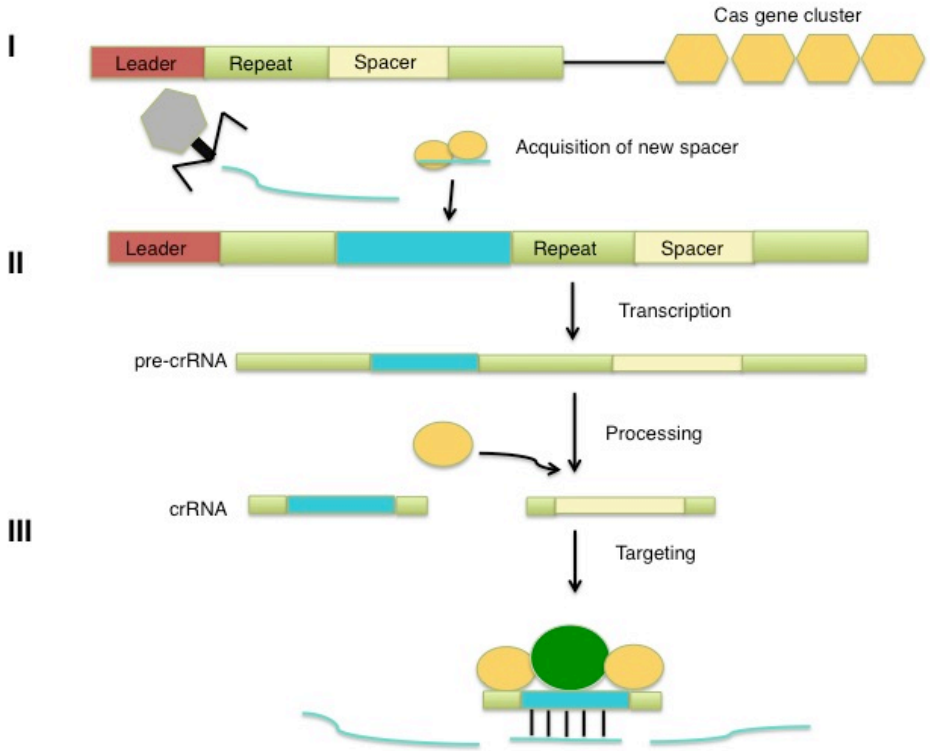



**Figure 5** - Overview of CRISPR/Cas adaptive immunity. **I – Adaptation:** CRISPR arrays are composed of short repeats sequences derived from foreign invaders (spacers). Upon infection by a foreign element, part of its genome is incorporated into the leader end of the CRISPR array (which is adjacent to a cluster of *Cas* genes); **II – crRNA generation**: The CRISPRs are transcribed into pre-crRNAs that are then processed into mature crRNAs; **III – Interference:** The Cas protein-crRNA complex binds and degrades the target nucleic acid. Adapted from Richter, C. et. al 2012 [71].

### Type II systems

Type II CRISPR/Cas was first discovered in *Streptococcus pyogenes* [94] and it is characterized by the presence of Cas9 (besides Cas1 and Cas2 ubiquitous proteins), a protein containing two nuclease domains (RuvC-like and HNH) [86]. In this type of system, a short trans-encoded transcript (tracrRNA – trans-activating crRNA) is necessary for processing of the pre-crRNA into mature crRNA [94]. Cas9, in complex with the tracrRNA-crRNA structure (that will be now on referred to as single-guide RNA (sgRNA)), binds the target dsDNA, creating a double strand break by cleaving the complementary and non-complementary strands with its HNH- and RuvC- like domains, respectively [95]. Garneau and co-workers observed *in vivo* cleavage of dsDNA at a position 3 bp from the PAM sequence within the protospacer (the sequence that gives rise to the spacer) [96]. This sequence (5'-NGG-3') seems to be of extreme importance in the affinity towards the target as cleavage of genomic sequences with mutated PAM is impaired in dsDNA [97]. Phages with PAM mutations have thus the ability to avoid targeting by type II systems. Moreover, complementarity between crRNA and target over a 13 bp stretch proximal to the PAM is required for interference with the nucleic acid, meaning that phages with mutations in this region of the protospacer might evade it [97]. The CRISPR/Cas9 genome engineering technique was thus derived from this specific type of prokaryotic defence system.

### CRISPR/Cas9 applications

CRISPR was originally employed to "knock-out" target genes [95, 98] in various cell types and organisms. For this it is necessary to induct co-expression of the Cas9 endonuclease and a sgRNA targeting a specific gene. The genomic target can be any ~20 nucleotide DNA sequence, provided that it is unmatchable (comparing to the rest of the genome) and is located immediately upstream of a PAM sequence. Once co-expressed, the Cas9 protein and the sgRNA form a riboprotein complex, allowing Cas9 to adopt an active DNA-binding conformation. This complex will then bind any genomic sequence with a PAM but the extent to which the sgRNA matches the target DNA determines whether the enzyme will cut. The end result of Cas9-mediated DNA cleavage is a DSB within the target DNA (~3-4 nucleotides upstream of the PAM sequence) that can then be repaired by one of two DNA-repair mechanisms (HR or NHEJ), depending on the presence or absence of a template sequence, respectively. The first is the most active repair mechanism, capable of rapidly repairing DSBs. In most cases, this mechanism gives rise to small nucleotide insertions or deletions (InDels) in the target DNA which result in in-frame amino acid deletions, insertions, or frameshift mutations. This leads to premature stop codons within the open reading frame (ORF) of the targeted gene and, ideally, result

in loss-of-function mutation of such gene. On the other hand, Homologous Recombination can be used to generate specific nucleotide changes (also known as “gene edits”), ranging from a single nucleotide change to large insertions. To achieve this, a DNA “repair template” containing the desired “edit” sequence must be delivered into the cell type of interest together with the sgRNA(s) and Cas9. The repair template must contain the desired edit as well as additional homologous sequence immediately upstream and downstream of the target (termed left & right homology arms) [99].

More recently, various modifications to the Cas9 enzyme have extended the applications of this technique: selectively activate or repress target genes [100], purify specific regions of DNA [101], image DNA in live cells using fluorescence microscopy [102] and even genome-wide screens [103]. Furthermore, the ease of generating sgRNAs makes CRISPR one of the most scalable genome editing technologies.

Imaging DNA in live cells

In 2013, a repurposed nuclease-deactivated Cas9 (dCas9) had been used by Gilbert and collaborators to regulate endogenous gene expression by controlling the RNA polymerase activity or modulating promoter accessibility when fused with transcription factors [100]. The generation of a deactivated Cas9 (by mutationally inactivating its two nuclease domains) that could be fused to other molecules such as transcriptions factors or the GFP protein brought new possibilities for the application of the system. Following this, the same authors sought that the CRISPR system could serve as a universal and flexible platform for the dynamic imaging of specific genomic elements in living mammalian cells [102]. This stressed the possibility, amongst others, of using dCas9 fused to the GFP protein and a sgRNA guiding to a telomeric repeat sequence to visualize and compare the telomeric lengths and observable appearance in living cancer cells.

### Limitations of CRISPR/Cas9

Each Cas9 ortholog has a unique PAM sequence. This does not totally limit the targeting range of this enzyme – in the human genome such target sites can be found on average every 9-12 bp [104, 105]. Another limitation of the CRISPR-Cas9 technique is the potential for off-target mutagenesis for a given guide sequence [105, 106]; however, they can be predicted and likely minimized by following general design guidelines [105], with the aid of bioinformatics.

### Ethical challenges

The simplicity of using the CRISPR/Cas9 system to engineer genomes in living organisms truly represents a new era in the fields of biology and genetics. However, several concerns regarding its ethics and biosafety have emerged in the past few years, especially since the publication of a paper reporting gene editing in human embryos [107]. Thenceforth, the editorial teams of both *Science* and *Nature* have written that the use of this technology for genome editing in human germline cells should be "addressed very seriously" and that "progressive policies should be developed/discussed". Human embryo editing is thus a controversial theme among scientists; some countries have even banned the use of the CRISPR/Cas9 technology in humans. This implies that a framework for open discourse on this issue (both within the scientific community and publicly, regarding scientific and ethical social concerns) is urgently needed. The safety and efficacy issues regarding this technology must be investigated and understood before attempts at human engineering for clinical testing. Considerations include the possibility of off-target alterations as well as on-target events that have unintended consequences [108]. These discussions should be initiated rather sooner than later as they are crucial in order to lead society's choices in this new era of biological/genetic engineering.

# AIMS OF THE STUDY

## AIMS OF THE STUDY

The aim of this study was to take advantage of the recently designed genome engineering technology CRISPR/Cas9 and optimize its application in the imaging of genomic loci in living human cells.

In particular, we aimed to produce a tool that allowed the visualization of telomeres through the targeting of a fluorescent protein complex (dCas9-EGFP) to telomeric repeats (by sgRNA) in Bon pancreatic neuroendocrine tumour cells.

# METHODOLOGY

# METHODOLOGY

## Material

### Plasmids

dCas9-EGFP, sgRNAs (sgTelomere and sgGAL4) and lentiviral (packaging and envelope-expressing) plasmids were commercially available (Addgene, Cambridge, USA). pSLQ1658-dCas9-EGFP (Addgene plasmid # 51023) and pSLQ1651-sgTelomere(F+E) (Addgene plasmid # 51024) were a gift from Bo Huang and Stanley Qi. pU6-sgGAL4-4 (Addgene plasmid # 46916) was a gift from Stanley Qi and Jonathan Weissman. pMD2.G (Addgene plasmid # 12259) and psPAX2 (Addgene plasmid # 12260) were a gift from Didier Trono. pCMV-Tet3G Vector (Clontech Cat. No. 631335, Tet-On 3G Transactivator) was obtained as part of the Tet-On® 3G Inducible Expression System (with mCherry) (Clontech – Takara Bio, Mountain View, USA).

### Cell lines

The two cell lines used in this work were available in the Ipatimup cell bank. Human embryonic kidney HEK293FT cells were used for lentiviral production by transient transfection. Bon cell line was used as it is originally derived from a human metastatic neuroendocrine pancreatic tumour [109] and represents a potential model for the study of the impact of ATRX depletion in the generation of the ALT phenotype. These cells were in three different conditions (WT, shRNA Control and shRNA ATRX – Knockdown (KD)). Control and KD were previously established in the lab.

## Methods

### Bacterial culture and transformation

Bacteria were plated in Luria Broth (LB) agar with ampicillin (100 $\mu$g/mL) overnight at 37ºC in an incubator. In the next morning, single colonies were isolated using a sterile pipette tip and inoculated in liquid LB media with ampicillin (100 $\mu$g/mL) (Sigma-Aldrich, St.Louis, USA). Cultures were incubated for 12-18h in an orbital shaker incubator (Biosan, Riga, Latvia), at 37ºC. LB media with ampicillin without any bacterial inoculation was used as negative control. After bacterial growth, 500 µL of the overnight cultures were added to 500µL of 50% glycerol (Sigma) in a cryovial to create glycerol stocks stored at -80ºC.

Clontech's plasmid DNA was received in storage buffer (10 mM Tris-HCl (pH 8.0), 1 mM EDTA (pH 8.0)). This plasmid DNA (0,5µL) was mixed with 50µL of *E. coli* competent cells in a microcentrifuge tube. The DNA was transformed in bacteria by the standard heat shock transformation protocol. Competent cells with plasmid DNA mixture were put on ice for 30 minutes and submitted to heat shock at 42ºC for 45 seconds in a thermomixer (Eppendorf, Hamburg, Germany). Following this, cells were put back on ice for 2 minutes. Competent cells (250µL) were then added to LB media and grown at 37ºC in an orbital shaking incubator for 45 minutes. After this, transformed cells were plated in LB agar plates with 100 $\mu$ g/mL ampicillin and incubated at 37ºC overnight. Following this, an overnight liquid culture was inoculated with a single colony and a glycerol stock was created, as described before.

All plasmids were posteriorly isolated and purified. First, bacterial cultures were grown from the glycerol stocks: a sterile pipette tip was used to inoculate the frozen bacteria into 100 mL of LB media that was then incubated at 37ºC with constant shaking for 12-16 hours. For isolation and purification, NZYMaxiprep (nzytech, Lisboa, Portugal) kit was used according to the manufacturer instructions.

**Electrophoresis and Sanger sequencing.**

Plasmid DNA size (kb) and sequence were confirmed by electrophoresis and Sanger sequencing, respectively. To linearize plasmids, 1 $\mu$ L (≅ 500ng) plasmid DNA was mixed with 15 $\mu$ L $H_2O$, 2 $\mu$ L reaction buffer and 1 $\mu$ L restriction enzyme Hind III (ThermoFisher, Waltham, USA) was used to cut dCas9, Tet3G and pMD2.G; NheI (ThermoFisher) for psPAX2, sgGAL4 and sgTelomere. A 1% agarose gel was prepared by mixing 0,8g GRS agarose (Grisp, Porto, Portugal) with 80mL SGTB 1x Agarose Electrophoresis Buffer (Grisp). Loading samples were prepared by mixing 1 $\mu$ L glicerol and bromophenol blue Loading Buffer + Gel Red (Biotium, Fremont, USA) with 2,5 $\mu$ L DNA. 1kb Plus DNA Ladder (Invitrogen, Carlsbad, USA) served as ladder marker. Samples were loaded in the gel, which was run at 120V by 40 minutes.

Sequencing reaction was performed with the ABI Prism BigDye Terminator Kit (Perkin-Elmer, Foster City, USA) and the fragments were run in an ABI prism 3100 Genetic Analyser (Perkin-Elmer). *M13rev* reverse primer (Sigma-Aldrich) , 5′-CAG GAA ACA GCT ATG ACC-3′ was used to sequence dCas9 and psPAX2 vectors. *CMVfw* forward primer (Eurogentec, Seraing, Belgium), 5'-GCT GTT TTG ACC TCC ATA GAA G-3' was used to sequence pMD2.G and Tet3G vectors. *pcDNA4_6_CMV* forward primer (Invitrogen), 5'-CATGAAGAATCTGCTTAGGG-3' was used to sequence sgGAL4 and sgTelomere vectors.

### Cell culture

HEK293FT cells were cultured in Dulbecco's Modified Eagle Medium High Glucose (Capricorn Scientific, Ebsdorfergrund, Germany) supplemented with 10% FBS (Capricorn), 1% Penicillin/Streptomycin (Capricorn) and 0,5% Fungizone (Capricorn). BON cells were cultured in Dulbecco's Modified Eagle Medium/Nutrient F-12 Ham (Capricorn) supplemented with 10% FBS (Capricorn), 1% Penicillin/Streptomycin (Capricorn) and 0,5% Fungizone (Capricorn). For culturing Bon shRNA CTRL and shRNA ATRX, the media was maintained with 2ηg/mL puromycin (InvivoGen, San Diego, USA), to perform continuous selection of the cells containing the shRNAs. All cells were maintained at 37ºC in a 5% $CO_2$ controlled environment. Both cell lines were tested for the presence of *Mycoplasma*. Upon reaching confluence, cells were trypsinized and seeded at the desired density in T25 flasks for growth, 6-well plates for transfection or 24-well plates for lentiviral transduction.

### Protein extraction and Western Blot

Cells were lysed for 15 min at 4ºC using Radioimmunoprecipitation Assay (RIPA) buffer (1% NP-40 in 150 mM NaCl, 50 mM Tris (pH 7.5), 2 mM EDTA) containing phosphatase and protease inhibitors. Cells were then collected by scraping and pelleted by centrifugation for 5 minutes at 14000 rpm and 4°C, to remove cell debris. The supernatant, which contained the fraction of cytoplasmic proteins, was recovered and stored at -20ºC. Proteins were quantified using a modified Bradford assay (Bio-rad, Hercules, USA), at 650 nm. Protein samples (40 µg) were separated in 10% SDS/ PAGE gels and transferred to Hybond ECL membrane (GE Healthcare, Little Chalfont, UK). Anti-ATRX rabbit (Sigma-Aldrich) was used as primary antibody. Goat anti-rabbit secondary antibody (Santa Cruz Biotechnology, Dallas, USA) was conjugated with peroxidase and visualized by ECL solution. Membranes were re-stained with a goat polyclonal anti-actin (Santa Cruz Biotechnology) antibody for protein loading control. All antibodies were used in 1:2000 dilutions.

CRISPR/Cas9 (imaging)

### Lentiviral production

HEK293FT cells were cultured in Dulbecco's Modified Eagle Medium High Glucose (Capricorn) supplemented with 10% Tet-System Approved FBS (Clontech), 1% Penicillin/Streptomycin (Capricorn) and 0,5% Fungizone (Capricorn) in 100mm tissue culture dishes until 70-80% confluence. The following mixtures were prepared, per plate,

and incubated for 15 minutes each: I. 8µg interest vector (dCas9, Tet3G, sgTelomere, sgGAL4) + 6µg lentiviral packaging plasmid (psPAX2) + 2µg VSV-G envelope expressing plasmid (pMD2.G) in 300µL OptiMEM (ThermoFisher) ; II. 16µL Lipofectamine 2000 (ThermoFisher) in 300µL OptiMEM. After this, both were mixed and incubated for 30 minutes. Meanwhile, media of HEK293FT cells was removed and 3mL of OptiMEM added. The plasmids were then added drop wise into the cells and this mixture was incubated for 3h, at which point 10mL of DMEM+10% Tet System Approved FBS were added. In the next morning, media was changed and cells were incubated at 32,5ºC for the rest of the protocol. Supernatants containing lentivirus were collected at 36, 48 and 72h post-transfection and kept at 4ºC. Supernatants were then filtered with a 0,45µm filter and 1mL of each was collected and stored at -80ºC. Virus were then concentrated using Amicon Ultra Centrifugal Filter Units (Merck Milipore, Billerica, USA) at 1600g for 30 minutes. Concentrated viruses were stored at -80ºC.

**Lentivirus titration by qRT-PCR**

Supernatants (150µL) were measured for lentivirus titer using Lenti-X qRT-PCR Titration Kit (Clontech). It is composed by a quick RNA purification step (Nucleospin RNA Virus Kit and respective protocol) before quantifying the number of lentiviral genome copies using qRT-PCR and SYBR® technologies. The protocol in the kit was executed according to manufacturer instructions.

**dCas9-EGFP/Tet-On 3G lentiviral transduction**

Bon WT cells stably co-expressing dCas9-EGFP and Tet-On 3G transactivator (Tet3G) were generated by lentiviral transduction. Approximately 10,000 Bon WT cells were plated into 3 wells of 24-well plates (20 to 30% confluence) 12-18h before infection. Each well contained 0,5mL complete growth medium (supplemented with 10% Tet System Approved FBS. Concentrated lentiviral stocks (dCas9-EGFP and Tet 3G) were thawed, combined at a 1:1 ratio and added to the cells. Transduction was done overnight at 37ºC and medium was supplemented with polybrene (8µg/mL) (Sigma). In the next morning, virus/polybrene-containing media was discarded and replaced with fresh complete medium (10% FBS). After 36h, medium was replaced again and doxycycline (Sigma) was added to one of the transduced wells at a concentration of 100ng/mL for inducing high expression of dCas9-EGFP. Cells were incubated for 12-24h, at which point, the ZOE fluorescent cell imager (Bio-Rad) was used to determine if dCas9-EGFP was properly localized in the cell's nucleus (as it contains NLS function). After being

confluent, cells were trypsinized and passaged to a 6-well plate to further visualization and to T25 flasks to freeze.

### sgRNA lentiviral infection

Approximately 4,000 cells were plated into a 8-well chambered coverglass 12-18h before infection. Each well contained 0,4-0,5mL DMEM-F12. Concentrated lentiviral stocks (sgTelomere and sgGAL4) were thawed and added to the cells. Transduction was done overnight at 37ºC and medium was supplemented with polybrene (8µg/mL) (Sigma). In the next morning, virus/polybrene-containing media was discarded and replaced with fresh complete medium. 48h post-transduction, the cells were observed with ZOE fluorescent cell imager.

# RESULTS

## RESULTS

### Purified plasmidic DNA quantification

Plasmid stocks were expanded and purified. DNA quantity in the resulting products was 697,3ηg/μL for sgTelomere; 529,4ηg/μL for pMD2.G; 453,5ηg/μL for Tet3G; 430,5ηg/μL for psPAX2; 812,4ηg/μL for sgGAL4 and 164ηg/μL for dCas9-EGFP (Table 1). All DNA was resuspended in 200-400μL of DNAse/RNase-free sterile $H_2O$ before quantification.

| Sample ID | ηg/μL | A260 | A280 | 260/280 | 260/230 |
|---|---|---|---|---|---|
| sgTelomere | 697,3 | 13,946 | 7,517 | 1,85 | 2,29 |
| pMD2.G | 529,4 | 10,588 | 5,759 | 1,84 | 2,33 |
| Tet3G | 453,5 | 9,070 | 4,860 | 1,87 | 2,29 |
| psPAX2 | 430,5 | 8,609 | 4,615 | 1,87 | 2,32 |
| sgGAL4 | 812,4 | 16,248 | 8,647 | 1,88 | 2,33 |
| dCas9-EGFP | 164 | 3,260 | 1,790 | 1,83 | 2,19 |

**Table 1** – Plasmid DNA quantification. Quantities in ηg/μL. A260/280 ratio accounts for assessment of nucleic acid purity. As these values are close to 2, it is inferable that the data is reliable and samples are not contaminated.

### Plasmid electrophoresis

After obtaining the purified plasmid DNA, it was linearized by digestion with one of two (HindIII and NheI) restriction endonucleases and ran in a 1% agarose gel to confirm vector size (bp). As expected, pMD2.G size was 5,8 Kb (well 2), sgTelomere was 8Kb (well 3), sgGAL4 was 8Kb (well 4), psPAX2 was 10Kb (well 5), Tet3G was 7,1Kb (well 6) and dCas9-EGFP was 11Kb long (well 7). The traces of DNA seen in the top of wells 2 and 3 are probably the result of an incomplete digestion by the restriction enzyme and/or the appearance of secondary structures of plasmidic DNA, creating extra bands (Figure 6).

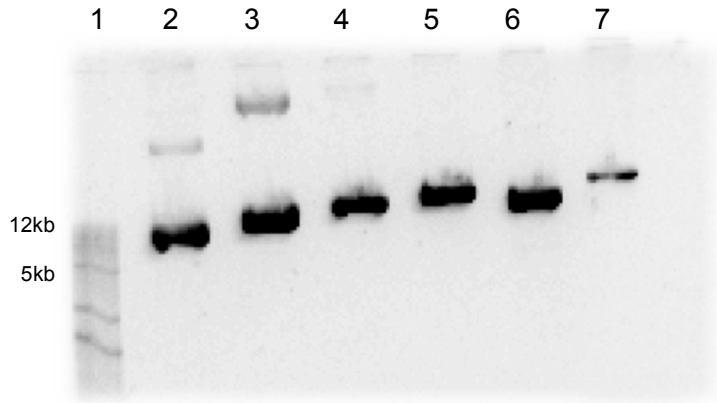



**Figure 6** – Plasmid electrophoresis gel. 1 – Marker, 2 – pMD2.G, 3 – sgTelomere, 4 – sgGAL4, 5 – psPAX2, 6 – Tet3G, 7 - dCas9-EGFP.

### Plasmid sequencing

The next step was to confirm vector sequence. Sequencing reaction was performed as described before and the plasmid DNA sequence was compared with the ones provided by the suppliers (Addgene, Clontech). As expected, all sequences matched up and no mutations were detected (Figure 7).

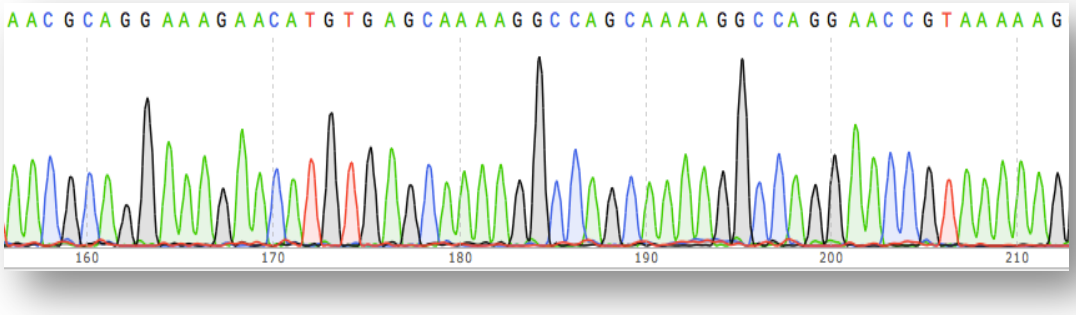



**Figure 7 –** An example of a part of a sequence (dCas9-EGFP) is shown. See other vector's partial sequences in Supplementary Figures.

### Transfection with Lipofectamine 2000

HEK293FT transfection with the packaging (psPAX2), envelope-expressing (pMD2.G) and interest (dCas9-EGFP, Tet3G, sgGAL4 or sgTelomere) vectors was performed in order to obtain four different types of lentiviral particles (corresponding to the four interest DNA vectors). The first time, only the exact necessary amount of lentivirus to transduce two wells of a 24-well plate containing Bon WT cells was produced. We infected these cells with only the Tet3G lentivirus and immediately tried to perform selection using

the G418 antibiotic (Sigma) (from which Tet3G confers resistance). All other lentivirus were stored at -80ºC. 15 days after starting selection all cells were dead, meaning that we hadn't been able to perform selection and we had lost all Tet3G lentivirus. For this reason, we decided to produce another batch of lentivirus by transfection. Once again, all products were stored at -80ºC. By this time, we had resolved that we would do (simultaneous) co-transduction of Bon cells with Tet3G and dCas9-EGFP lentivirus in order to be able to identify the infected cells by fluorescence microscopy (only in the presence of the two vectors the cells will produce nuclear fluorescence) instead of by selection.

**Lentiviral titration – Quantitative RT-PCR**

A small part of the Tet3G and dCas9-EGFP lentiviral samples obtained in the 2nd transfection experiment was used to perform quantitative Real Time Polymerase Chain Reaction (qRT-PCR), which was the chosen method to titrate the lentiviral samples. This provided us the information about the presence and quantity of the lentiviral particles in the two samples, comparing with the standard provided by the commercial kit. As observed in the amplification plot (Figure 9), we obtained the amplification curves for five standards (in grey) as well as for the two produced lentiviral samples (Tet3G in blue, dCas9-EGFP in green). The established threshold was 0,040395, which defined the $C_T$ values for the analysed samples.

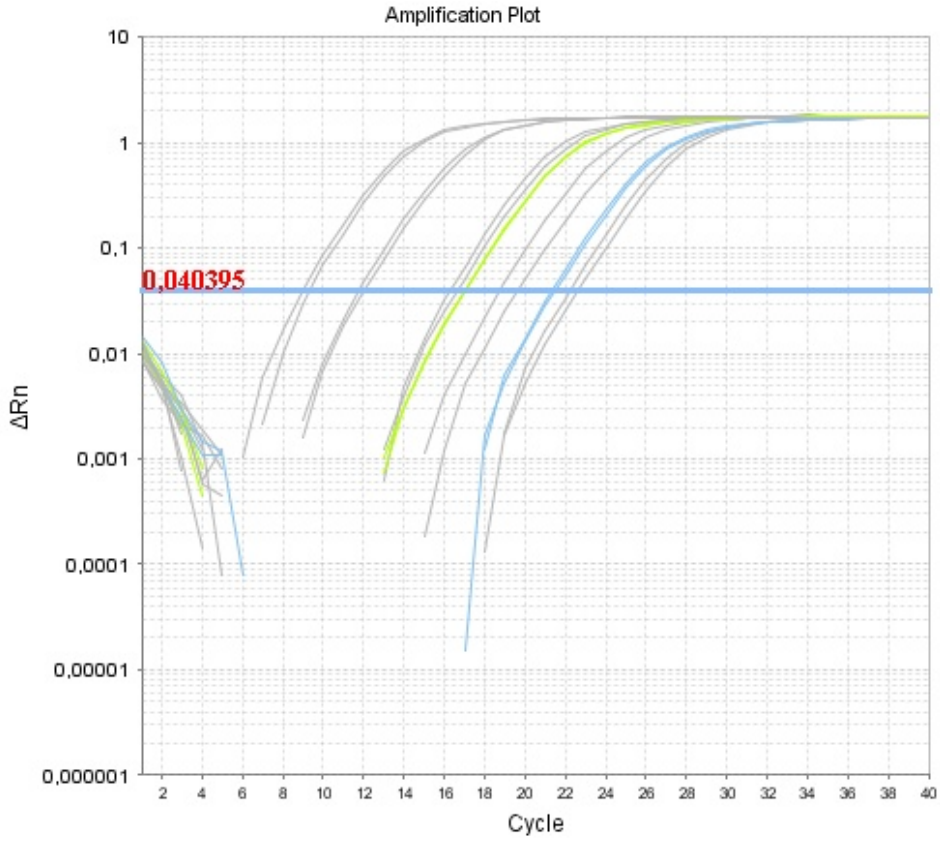


**Figure 9** – qRT-PCR amplification plot. In grey, the five standards used to produce the standard curve. In blue, the Tet3G sample. In green, the dCas9-EGFP sample. All samples were duplicated. Threshold was 0,040395.

We also obtained the standard curve and melt curve plots (Figure 10). The first provided us data about the efficacy of the reaction, which was 97,9% (meaning that it was a very efficient reaction, as this value should be between 90-110%). Also, the slope value (-3,374) shows a concordant result with the efficacy. The melt curve plots provided the information about the quality of the data. The fact that all melt curves (five standards and two samples) are in the same spectrum and always consist of a unique peak indicates that the data is reliable.

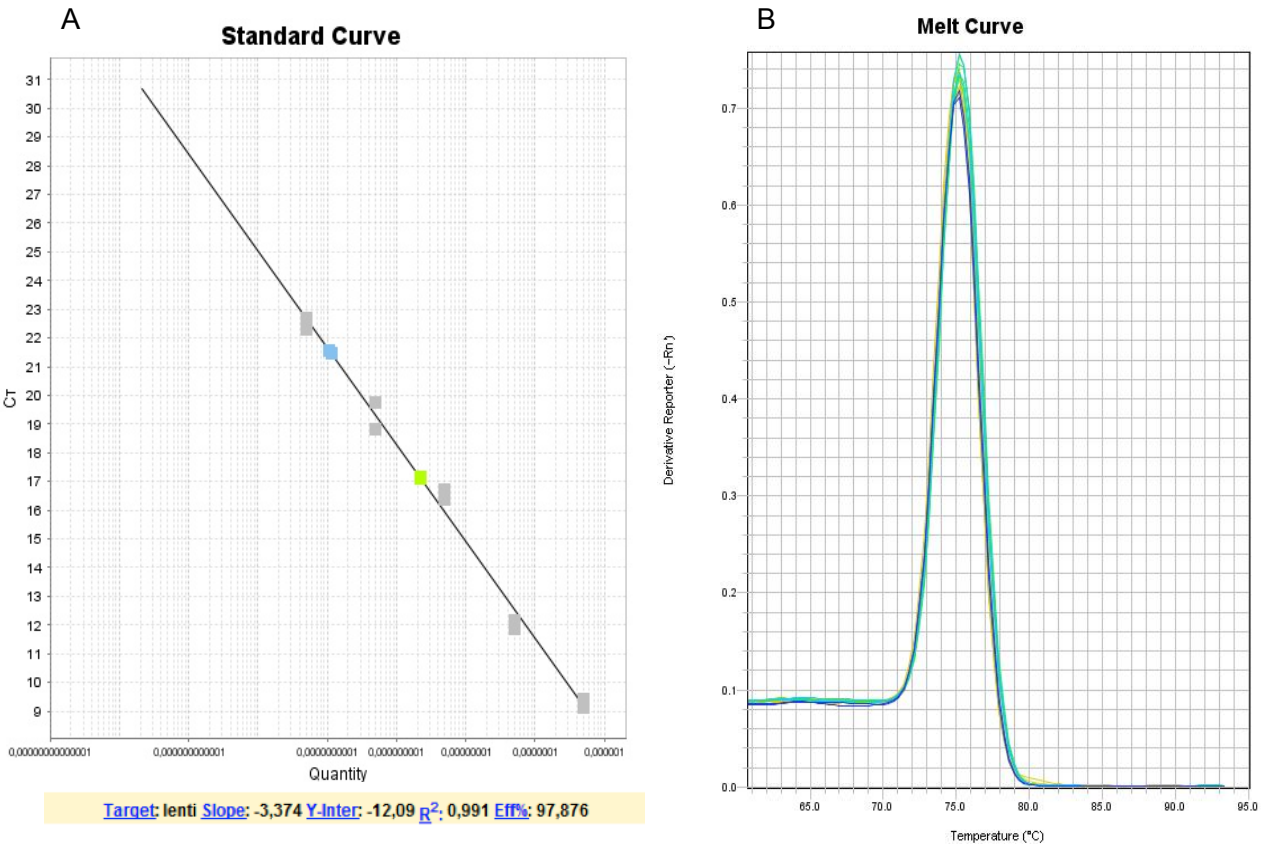



**Figure 10** – Lentiviral titration by quantitative RT-PCR. (A) Standard Curve. In grey, the five standards. In blue, the Tet3G sample. In green, the dCas9-EGFP sample. The x-axis corresponds to the quantity of lentivirus in the sample. The y-axis corresponds to the $C_T$ value of the various samples. Curve slope was -3,374. Efficacy was 97,9%; (B) Melt Curve. All seven samples (and duplicates) produced one single peak in similar temperature values.

Lastly, we obtained the data referring to the quantity (in units) of lentivirus in all samples and the cycle threshold ($C_T$) values. Standards 1, 2, 3, 4 and 5 were successive dilutions of the lentiviral standard provided by the kit, being the quantity of lentivirus in each 5E-07, 5E-08, 5E-09, 5E-010 and 5E-11 units, respectively. Its $C_T$ values means were 9,27; 11,99; 16,54; 19,28; 22,50, respectively. For our samples, Tet3G mean lentiviral quantity was 1,10E-10 units and mean $C_T$ was 21,51. For dCas9-EGFP, mean lentiviral quantity was 2,19E-09 units and mean $C_T$ was 17,13 (Table 2). Here, we confirmed what we could already observe in the amplification plot: dCas9-EGFP sample had 10-fold more lentiviral units than the Tet3G one.

| Well | Sample Name | Reporter | Cт | Cт Mean | Cт SD | Quantity | Quantity Mean | Quantity SD |
|---|---|---|---|---|---|---|---|---|
| **B4** | Std1 | SYBR | 9,107636452 | 9,274354935 | 0,235776216 | 5E-07 | | |
| **B5** | Std1 | SYBR | 9,441074371 | | | | | |
| **C4** | Std2 | SYBR | 11,84287167 | 11,9901638 | 0,208303213 | 5E-08 | | |
| **C5** | Std2 | SYBR | 12,13745689 | | | | | |
| **D4** | Std3 | SYBR | 16,38324928 | 16,54104614 | 0,223157108 | 5E-09 | | |
| **D5** | Std3 | SYBR | 16,69884109 | | | | | |
| **E4** | Std4 | SYBR | 18,81335449 | 19,27684021 | 0,65546912 | 5E-10 | | |
| **E5** | Std4 | SYBR | 19,74032784 | | | | | |
| **F4** | Std5 | SYBR | 22,7010994 | 22,50058174 | 0,28357479 | 5E-11 | | |
| **F5** | Std5 | SYBR | 22,30006409 | | | | | |
| | | | | | | | | |
| **B6** | Tet_1x | SYBR | 21,57520676 | 21,51176071 | 0,089724906 | 1,05144E-10 | 1,09899E-10 | 6,72542E-12 |
| **B7** | Tet_1x | SYBR | 21,44831657 | | | 1,14655E-10 | | |
| **B8** | dCas9_1x | SYBR | 17,11772919 | 17,1286087 | 0,015387309 | 2,20266E-09 | 2,18643E-09 | 2,29601E-11 |
| **B9** | dCas9_1x | SYBR | 17,13949013 | | | 2,17019E-09 | | |

**Table 2** - Quantity of lentiviral particles and $C_T$ values for the standards and Tet3G and dCas9-EGFP samples.

### Tet3G/dCas9-EGFP lentiviral co-transduction

We performed co-transduction of Bon WT cells with Tet3G and dCas9-EGFP lentivirus in 24-well plates, overnight. 48h later, a small percentage of the cells had nuclear fluorescence. As Tet3G codes for an activator of the TRE3G promoter present in the dCas9-EGFP plasmid and this activation only occurs in the presence of doxycycline (Tet-On inducible expression), we maintained transduced cells in two conditions: dCas9 (transduced cells in normal media) and dCas9+Dox (transduced cells in media containing 100ηg/mL Doxycycline). We took pictures of all wells using two conditions (Brightfield and Green fluorescence) of the ZOE fluorescent cell imager (Figure 11). Compared to the Control Bon WT cells (CTRL) where there is only random fluorescence in the media or in auto-fluorescing cells, the co-transduced cells (dCas9) exhibit points of green

fluorescence that seem to be specifically localized in the nucleus of the cells. These are in little number, which indicates that the lentivirus transduction might not have been of great efficacy. In the presence of doxycycline, there is a notorious difference in the number of fluorescent nucleus, indicating higher expression of the dCas9-EGFP protein complex.

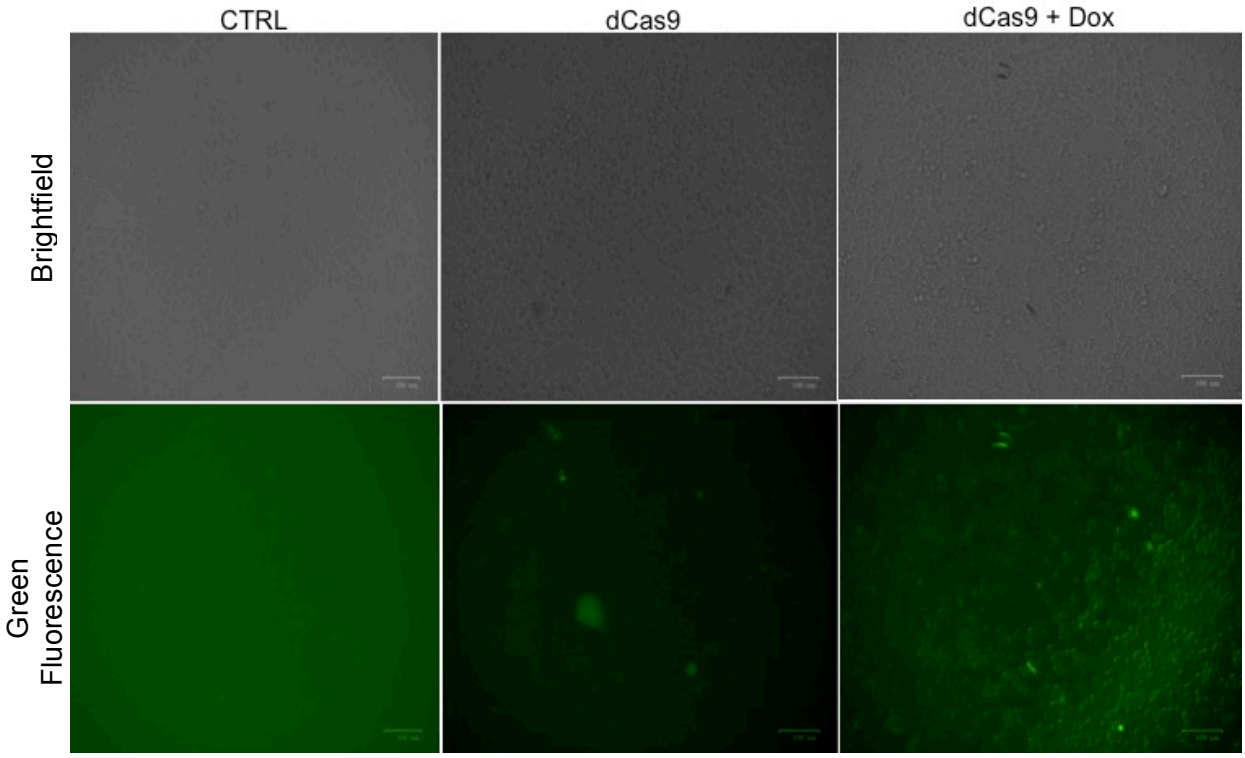



**Figure 11 –** Tet3G/dCas9-EGFP lentiviral co-transduced cells in a 24-well plate. Pictures taken in Brightfield and Green Fluorescence (488nm) conditions. CTRL are non-transduced Bon WT cells. dCas9 are transduced Bon WT cells in normal media. dCas9+Dox are transduced cells in media containing Doxycycline.

As we couldn't expand the cells because of a contamination, we recurred to our frozen lentivirus samples and repeated the experiment. We visualized and photographed fluorescence in 6-well plates (Figure 12). The same as in Figure 11 was observed as co-transduced cells exhibited green fluorescence, being this effect more evident in those cultured with doxycycline. This indicated that the second experiment had been successful, although the efficacy seemed to remain low.

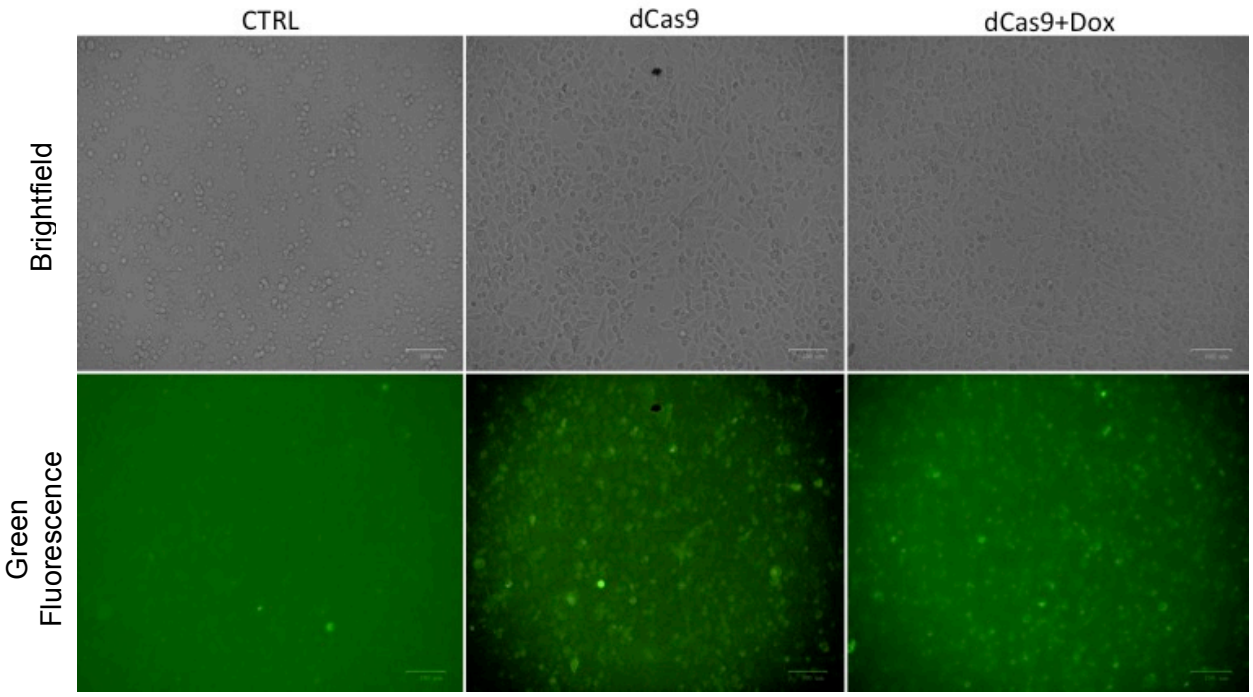



**Figure 12** – Tet3G/dCas9-EGFP lentiviral co-transduced cells expanded to a 6-well plate. Pictures taken in Brightfield and Green Fluorescence (488nm) conditions. CTRL are non-transduced Bon WT cells. dCas9 are co-transduced Bon WT cells in normal media. dCas9+Dox are co-transduced cells in media containing Doxycycline.

This time, we were able to expand transduced cells to T25 flasks, at which point, a part of those were frozen. Pictures were taken by merging the brightfield and green flourescent conditions, providing a better visualization of the fluorescence (Figure 13). The difference between control cells and co-transduced remained observable and was evident when doxycycline was added.

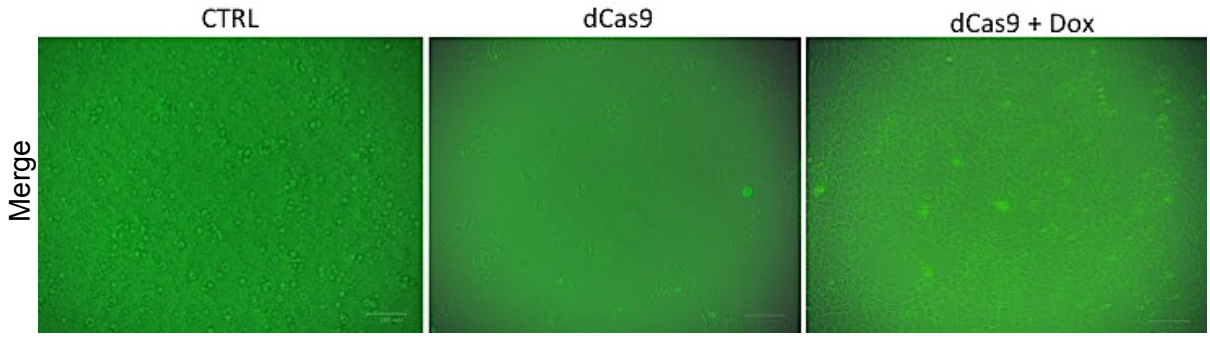



**Figure 13** - Tet3G/dCas9-EGFP lentiviral co-transduced cells seeded in T25 flasks. Pictures taken in Merge (Brightfield + Green Fluorescence) condition. CTRL are non-

transduced Bon WT cells. dCas9 are co-transduced Bon WT cells in normal media. dCas9+Dox are co-transduced cells in media containing Doxycycline.

We then tried to select clonal cells stably expressing Tet3G/dCas9 by two methods, simultaneously: serial dilutions and antibiotic resistance. In the various dilutions trials, we could not isolate a single fluorescent colony of cells originating doubtlessly from a unique cell. We tried to perform selection with media containing 500 ηg/mL of puromycin as the dCas9-EFDP vector confers resistance to this antibiotic but all cells died after 15-17 days.

We thus sought to experiment transducing the Bon cells with a different ratio of lentivirus, which could improve the efficacy of the process. The remaining stored amount of lentiviral samples allowed us to infect Bon WT cells with a 2:1 ratio of Tet3G/dCas9-EGFP lentivirus. As before, cells were expanded to T25 flasks, at which point we took pictures (Figure 14). Comparing to the previously tested 1:1 lentiviral ratio, there were no observable differences in the localization or intensity of the fluorescence. It remained low and in little quantity, again showing apparent low transduction efficacy.

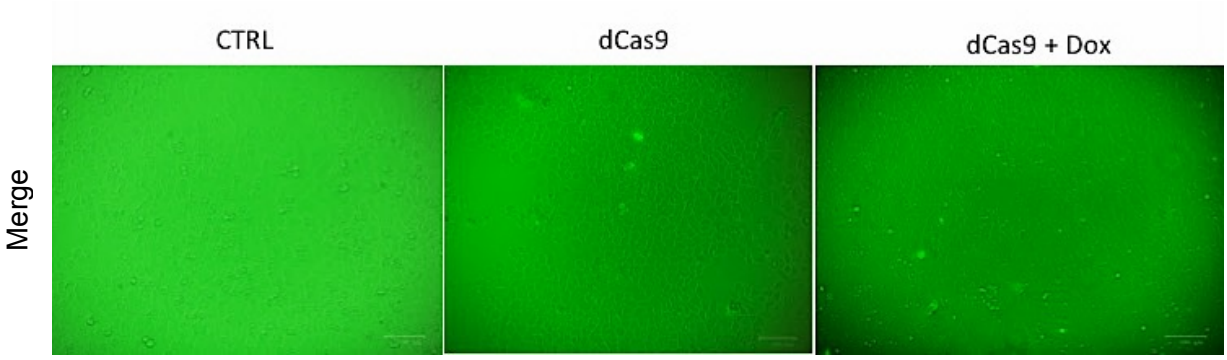



**Figure 14 -** Tet3G/dCas9-EGFP 2:1 lentiviral co-transduced cells seeded in T25 flasks. Pictures taken in Merge (Brightfield + Green Fluorescence) condition. CTRL are non-transduced Bon WT cells. dCas9 are co-transduced Bon WT cells in normal media. dCas9+Dox are co-transduced cells in media containing Doxycycline.

### sgRNA lentiviral infection

As we didn't have the time to try to select these cells, we decided to transduce all the non-selected dCas9/Tet3G previously infected cells that were being cultured (with both 1:1 and 2:1 lentiviral ratio) with the sgRNAs (sgTelomere and sgGAL4). The result was inconclusive as we were not able to identify points of fluorescence specific to the nucleus of the cells that would indicate the presence of telomeres.

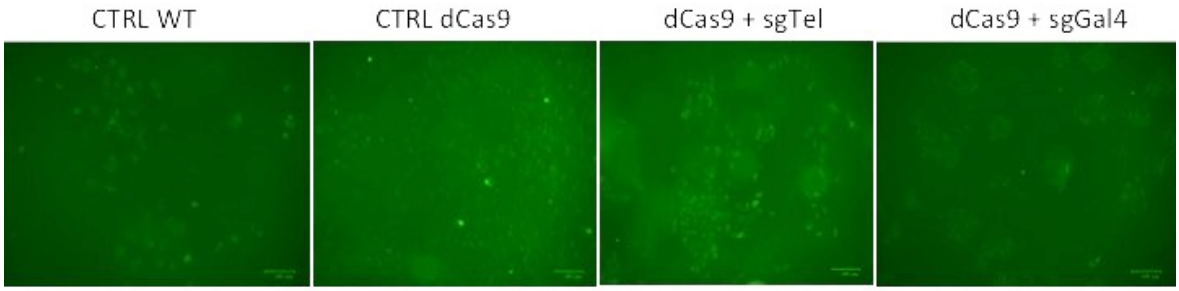



**Figure 15** - Tet3G/dCas9-EGFP/sgRNA lentiviral co-transduced cells seeded in 8-well plate. Pictures taken in Green Fluorescence condition. CTRL WT are non-transduced Bon WT cells. CTRL dCas9 are Tet3G/dCas9 transduced cells. dCas9 + sgTel are Tet3G/dCas9/sgTelomere transduced cells. dCas9 + sgGal4 are Tet3G/dCas9/sgGal4 transduced cells.

# DISCUSSION

## DISCUSSION

In this project, we aimed to establish a tool for the visualization of specific genomic loci (in this case telomeres) in living cells, without any kind of treatment or global DNA denaturation, by a CRISPR/Cas9 method where the Cas9 protein is deactivated (dCas9) and coupled to a Green Fluorescent Protein. This approach thus preserves the spatial relationships of the genetic elements, which is crucial for understanding the organization of the genome and its association with gene expression. We also wanted to compare the characteristics (type of fluorescence: its spatial localization, size, intensity and number of puncta) between different genetic conditions of the Bon cell line (WT, shRNA Control and shRNA ATRX).

Our approach to this aim was to produce lentivirus containing the necessary DNA for the Bon cells to express dCas9-EGFP and the interest sgRNAs and infect Bon cells with such lentivirus. Four types of lentivirus were generated, containing dCas9-EGFP (coding for a deactivated Cas9 protein coupled to and enhanced GFP), Tet3G (a vector coding for a protein that ligates dCas9 promoter and positively regulates its expression in the presence of doxycycline), sgTelomere (a single-guide RNA targeting telomeric repeats) and sgGAL4 (a single guide RNA with no known target in the human genome, to serve as control). Afterwards, we sought to establish a stable cell line expressing both Tet3G and dCas9 proteins, by selection, so that we could further infect these cells with the sgRNAs and observe the telomeric fluorescence.

Our initial results in the preparation of the vectors were as expected. The purified plasmid DNA was analysed for its size and sequence and this data was concordant with the information provided by the suppliers (Addgene and Clontech).

In this work it was not possible to attain our ultimately goal due to technical difficulties that affected some phases of the process. Regarding the CRISPR/Cas9 data, it is not conclusive but it opens the possibility of discussing the explanations for the observed events. As seen in Figure 13, Control (non-infected) cells are different from the infected ones. The fluorescence in the first ones is random and mostly located in what seem to be dying/dead cells or other particles in the media – autofluorescence. The Tet3G/dCas9 infected cells, on the other hand, exhibit blurs of fluorescence that are located in living cells, as we can see in the merged pictures (Figures 13 and 14), which is concordant with what is seen in the literature [102, 110, 111]. However, the fluorescent foci were in small number throughout the area of the flasks and were mostly located in cells that had a rounded format, contrasting with the adherent cells with a flattened format. This indicates that the infection process was successful but its efficacy low, drawback probably related to the processes of transfection and/or transduction.

In the tentative to quantify and normalize the transfection experiments we have done quantification of the lentiviral load by qRT-PCR. Nevertheless the transfection process was not fully controlled as the quantitative RT-PCR protocol only provided data about the presence of lentiviral-specific sequences in the samples and not about its functionality. This means that the qRT-PCR analysis does not give indication about the percentage of lentivirus containing the DNA inserts of interest. Also, we could not be sure if they were correctly assembled and capable of successfully infecting the cells. Although we found sufficient quantity of lentivirus for transduction in the two analysed samples, the percentage of functional ones was thus unknown. Also, importantly, the qRT-PCR data indicated that the Tet3G lentiviral solution was 10 times less concentrated than the dCas9, as seen in Table 2.

Figures 11, 12 and 13 indicate that the transduction process was not of great efficacy. As referred, the fluorescent foci seen in these photographs are in small quantity and appear to be weak indicating that, while some cells might have been successfully infected, these represent a small percentage. These results could be related to a number of factors apart from the fact that the lentiviral production by transient transfection could have not been functional, as referred previously. As the Tet3G lentiviral solution was 10-fold less concentrated comparing to dCas9 and as it is necessary that the two lentivirus are incorporated into the cells to produce fluorescence (the cells are co-transduced), the ratio of both lentivirus used in the transduction experiment was not intuitive and had to be optimized. We tried both 1:1 and 2:1 (Tet3G:dCas9) ratios (Figures 13 and 14, respectively) but no striking difference were observed between those two conditions. Nevertheless, a deeper analysis to this question should be done in future attempts to optimize this method in Bon cells and other ratios between the lentiviral solutions must be essayed. The dCas9 vector contains an inducible TRE3G promoter, meaning that the dCas9 protein expression is very strongly induced by the ligation of a Transactivator (Tet3G) protein to its promoter, in the presence of doxycycline. For this reason, we tried to stimulate the expression of dCas9 by adding doxycycline in two different concentrations: 100 ηg/mL (Figure 13) and 200 ηg/mL (data not shown). This experiment was also inconclusive, as there were no significant differences between the two. In fact, 100 ηg/mL should be more than enough to immensely stimulate fluorescence [102].

Throughout the entire experiment, it was only possible for us to use the ZOE Fluorescent Cell Imager in order to obtain the fluorescent cell imaging. This means our observations were limited by the characteristics of this specific microscope. As it only allowed us to observe cells in one magnification and its background signal was always high, this represents a great limitation to our work.

There is little information in the literature regarding the transduction of Bon cells with lentivirus. The intrinsic availability of these cells to transfection and transduction of lentivirus can also be an important factor to explain the low efficacy of the process. Although we didn't observe massive cell death upon infection, it might not be intuitive to successfully infect these cells. Existing experiments in this field of research (imaging genomic elements in living cells by CRISPR/Cas9) are mostly conducted in the RPE cell line [102, 110].

Afterwards, we tried to perform selection of infected Bon cells by the serial dilutions and antibiotic resistance methods. Both methods were unsuccessful, mostly likely due to the low efficacy of the transduction protocol. The little number of successfully infected cells producing fluorescence likely explains the fact that the cells didn't survive the antibiotic selection. This experiment was performed in 6-well plates, which cannot contain a large number of cells. For this reason, if there were cells surviving after the 15-17 days period of selection, they would be in very small number and probably couldn't form new colonies in order to regrow. Regarding the serial dilutions method, we weren't able to isolate, identify and plate a single fluorescent cell, as these cells were in such small percentage.

We lastly tried to infect the bulk population of cells, non-selected, with the sgRNAs, hoping that we could generate an example of what would be the targeted fluorescence to the telomeres in the nucleus, as seen in the literature [102]. Our result was inconclusive, as we could not see specific nuclear fluorescence in any of these cells. This is most likely explained by the fact that cells were plated in 8-well chambered coverglass, which also cannot contain a sufficient number of cells. As these cells were not stably expressing Tet3G/dCas9, the lack of fluorescence was not striking.

As our goals were not fully attained, we intend to continue trying to establish this tool by troubleshooting the whole process. There are various possible approaches to this problem in the future. We can try to directly transfect (instead of infecting with lentivirus) the cells with the Tet3G and dCas9 vectors to see if the cells successfully incorporate those vectors and produce a clear fluorescence signal that allows us to select them. Also, it might be promising to try to better control the processes of producing the lentivirus and infecting the cells in order to know if other ratios/quantities of lentivirus should be used.

# CONCLUSIONS

## CONCLUSIONS

We confirmed the capability of Bon cells to produce fluorescence by infecting them with lentivirus containing Tet3G and dCas9 vectors, even though the efficacy of the process was low. Following that, we tried to optimize this protocol in order to be able to produce telomeric-specific fluorescence. As this was unsuccessful, we will continue trying to establish this tool by troubleshooting the entire experiment and try again with a higher number of conditions and controls.

## REFERENCES


1. Muller, H., *The remaking of chromosomes.* Collecting net, 1938. **13**(198): p. 181-195.
2. McClintock, B., *The Stability of Broken Ends of Chromosomes in Zea Mays.* Genetics, 1941. **26**(2): p. 234-82.
3. Watson, J.D., *Origin of concatemeric T7 DNA.* Nat New Biol, 1972. **239**(94): p. 197-201.
4. Hayflick, L. and P.S. Moorhead, *The serial cultivation of human diploid cell strains.* Exp Cell Res, 1961. **25**: p. 585-621.
5. Olovnikov, A.M., *A theory of marginotomy. The incomplete copying of template margin in enzymic synthesis of polynucleotides and biological significance of the phenomenon.* J Theor Biol, 1973. **41**(1): p. 181-90.
6. Orr-Weaver, T.L., J.W. Szostak, and R.J. Rothstein, *Yeast transformation: a model system for the study of recombination.* Proc Natl Acad Sci U S A, 1981. **78**(10): p. 6354-8.
7. Szostak, J.W. and E.H. Blackburn, *Cloning yeast telomeres on linear plasmid vectors.* Cell, 1982. **29**(1): p. 245-55.
8. Greider, C.W. and E.H. Blackburn, *Identification of a specific telomere terminal transferase activity in Tetrahymena extracts.* Cell, 1985. **43**(2 Pt 1): p. 405-13.
9. Greider, C.W. and E.H. Blackburn, *A telomeric sequence in the RNA of Tetrahymena telomerase required for telomere repeat synthesis.* Nature, 1989. **337**(6205): p. 331-7.
10. Yu, G.L., et al., *In vivo alteration of telomere sequences and senescence caused by mutated Tetrahymena telomerase RNAs.* Nature, 1990. **344**(6262): p. 126-32.
11. Kim, N.W., et al., *Specific association of human telomerase activity with immortal cells and cancer.* Science, 1994. **266**(5193): p. 2011-5.
12. Vinagre, J., et al., *Telomerase promoter mutations in cancer: an emerging molecular biomarker?* Virchows Arch, 2014. **465**(2): p. 119-33.
13. Heaphy, C.M., et al., *Prevalence of the alternative lengthening of telomeres telomere maintenance mechanism in human cancer subtypes.* Am J Pathol, 2011. **179**(4): p. 1608-15.
14. Henson, J.D. and R.R. Reddel, *Assaying and investigating Alternative Lengthening of Telomeres activity in human cells and cancers.* FEBS Lett, 2010. **584**(17): p. 3800-11.
15. Serakinci, N., et al., *Telomerase promoter reprogramming and interaction with general transcription factors in the human mesenchymal stem cell.* Regenerative Medicine, 2006. **1**(1): p. 125-131.
16. Perrem, K., et al., *Repression of an alternative mechanism for lengthening of telomeres in somatic cell hybrids.* Oncogene, 1999. **18**(22): p. 3383-3390.
17. Jiao, Y., et al., *DAXX/ATRX, MEN1, and mTOR pathway genes are frequently altered in pancreatic neuroendocrine tumors.* Science, 2011. **331**(6021): p. 1199-203.
18. Schwartzentruber, J., et al., *Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma.* Nature, 2012. **482**(7384): p. 226-31.
19. Heaphy, C.M., et al., *Altered telomeres in tumors with ATRX and DAXX mutations.* Science, 2011. **333**(6041): p. 425.

20. Lundblad, V. and E.H. Blackburn, *An alternative pathway for yeast telomere maintenance rescues est1- senescence.* Cell, 1993. **73**(2): p. 347-60.
21. Dunham, M.A., et al., *Telomere maintenance by recombination in human cells.* Nature Genetics, 2000. **26**(4): p. 447-450.
22. Lallemand-Breitenbach, V. and H. de The, *PML nuclear bodies.* Cold Spring Harb Perspect Biol, 2010. **2**(5): p. a000661.
23. Cesare, A.J. and J.D. Griffith, *Telomeric DNA in ALT cells is characterized by free telomeric circles and heterogeneous t-loops.* Mol Cell Biol, 2004. **24**(22): p. 9948-57.
24. Henson, J.D., et al., *Alternative lengthening of telomeres in mammalian cells.* Oncogene, 2002. **21**(4): p. 598-610.
25. Durant, S.T., *Telomerase-independent paths to immortality in predictable cancer subtypes.* J Cancer, 2012. **3**: p. 67-82.
26. Royle, N.J., et al., *Telomere length maintenance--an ALTernative mechanism.* Cytogenet Genome Res, 2008. **122**(3-4): p. 281-91.
27. Bailey, S.M., M.A. Brenneman, and E.H. Goodwin, *Frequent recombination in telomeric DNA may extend the proliferative life of telomerase-negative cells.* Nucleic Acids Res, 2004. **32**(12): p. 3743-51.
28. Richard J. Gibbons, D.J.P., Laurent Villard, Douglas R. Higgs, *Mutations in a Putative Global Transcriptional Regulator Cause X-Linked Mental Retardation with a-Thalassemia (ATR-X Syndrome).* Cell, 1995. **80**: p. 837-845.
29. Picketts, D.J., et al., *ATRX encodes a novel member of the SNF2 family of proteins: Mutations point to a common mechanism underlying the ATR-X syndrome.* Human Molecular Genetics, 1996. **5**(12): p. 1899-1907.
30. Otani, J., et al., *Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX-DNMT3-DNMT3L domain.* EMBO Rep, 2009. **10**(11): p. 1235-41.
31. Ooi, S.K., et al., *DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA.* Nature, 2007. **448**(7154): p. 714-7.
32. Argentaro, A., et al., *Structural consequences of disease-causing mutations in the ATRX-DNMT3-DNMT3L (ADD) domain of the chromatin-associated protein ATRX.* Proc Natl Acad Sci U S A, 2007. **104**(29): p. 11939-44.
33. Flaus, A., et al., *Identification of multiple distinct Snf2 subfamilies with conserved structural motifs.* Nucleic Acids Res, 2006. **34**(10): p. 2887-905.
34. Drane, P., et al., *The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3.* Genes Dev, 2010. **24**(12): p. 1253-65.
35. Lewis, P.W., et al., *Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres.* Proc Natl Acad Sci U S A, 2010. **107**(32): p. 14075-80.
36. Xue, Y.T., et al., *The ATRX syndrome protein forms a chromatin-remodeling complex with Daxx and localizes in promyelocytic leukemia nuclear bodies.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(19): p. 10635-10640.
37. McDowell, T.L., et al., *Localization of a putative transcriptional regulator (ATRX) at pericentromeric heterochromatin and the short arms of acrocentric chromosomes.* Proc Natl Acad Sci U S A, 1999. **96**(24): p. 13983-8.

38. Gibbons, R.J., et al., *Mutations in ATRX, encoding a SWI/SNF-like protein, cause diverse changes in the pattern of DNA methylation.* Nat Genet, 2000. **24**(4): p. 368-71.
39. Bertrand Le Douarin, A.L.N., Jean-Marie Garnier, Hiroshi Ichinose, Fran9ois Jeanmougin, Regine Losson, PierreChambon, *A possible involvement of TIFIa and TIF1, in the epigenetic control of transcription by nuclear receptor.* EMBO Rep, 1996. **10**: p. 1235-1241.
40. Nan, X., et al., *Interaction between chromatin proteins MECP2 and ATRX is disrupted by mutations that cause inherited mental retardation.* Proc Natl Acad Sci U S A, 2007. **104**(8): p. 2709-14.
41. Nielsen, P.R., et al., *Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9.* Nature, 2002. **416**(6876): p. 103-7.
42. Eustermann, S., et al., *Combinatorial readout of histone H3 modifications specifies localization of ATRX to heterochromatin.* Nat Struct Mol Biol, 2011. **18**(7): p. 777-82.
43. Clynes, D., D.R. Higgs, and R.J. Gibbons, *The chromatin remodeller ATRX: a repeat offender in human disease.* Trends Biochem Sci, 2013. **38**(9): p. 461-6.
44. Lovejoy, C.A., et al., *Loss of ATRX, genome instability, and an altered DNA damage response are hallmarks of the alternative lengthening of telomeres pathway.* PLoS Genet, 2012. **8**(7): p. e1002772.
45. Watson, L.A., et al., *Atrx deficiency induces telomere dysfunction, endocrine defects, and reduced life span.* J Clin Invest, 2013. **123**(5): p. 2049-63.
46. Huh, M.S., et al., *Compromised genomic integrity impedes muscle growth after Atrx inactivation.* J Clin Invest, 2012. **122**(12): p. 4412-23.
47. Koschmann, C., et al., *ATRX loss promotes tumor growth and impairs nonhomologous end joining DNA repair in glioma.* Sci Transl Med, 2016. **8**(328): p. 328ra28.
48. Yong Wang, D.J.P., *Guanine residues in d(T2AG3) and d(T2G4) form parallel-stranded potassium cation stabilized G-quadruplexes with anti glycosidic torsion angles in solution.* Biochemistry, 1992. **31**(35): p. 8112-8119.
49. Law, M.J., et al., *ATR-X syndrome protein targets tandem repeats and influences allele-specific expression in a size-dependent manner.* Cell, 2010. **143**(3): p. 367-78.
50. Clynes, D., et al., *ATRX dysfunction induces replication defects in primary mouse cells.* PLoS One, 2014. **9**(3): p. e92915.
51. Clynes, D., et al., *Suppression of the alternative lengthening of telomere pathway by the chromatin remodelling factor ATRX.* Nat Commun, 2015. **6**: p. 7538.
52. Whitehouse, I. and T. Owen-Hughes, *ATRX: Put me on repeat.* Cell, 2010. **143**(3): p. 335-6.
53. Jiang, W.Q., et al., *Suppression of alternative lengthening of telomeres by Sp100-mediated sequestration of the MRE11/RAD50/NBS1 complex.* Mol Cell Biol, 2005. **25**(7): p. 2708-21.
54. Zeng, S., et al., *Telomere recombination requires the MUS81 endonuclease.* Nat Cell Biol, 2009. **11**(5): p. 616-23.
55. Robison, J.G., et al., *Replication protein A and the Mre11.Rad50.Nbs1 complex co-localize and interact at sites of stalled replication forks.* J Biol Chem, 2004. **279**(33): p. 34802-10.

56. Verdun, R.E. and J. Karlseder, *The DNA damage machinery and homologous recombination pathway act consecutively to protect human telomeres.* Cell, 2006. **127**(4): p. 709-20.
57. Ghosal, G. and K. Muniyappa, *Saccharomyces cerevisiae Mre11 is a high-affinity G4 DNA-binding protein and a G-rich DNA-specific endonuclease: implications for replication of telomeric DNA.* Nucleic Acids Res, 2005. **33**(15): p. 4692-703.
58. Napier, C.E., et al., *ATRX represses alternative lengthening of telomeres.* Oncotarget, 2015. **6**(18): p. 16543-58.
59. Zimmermann, S., et al., *Lack of telomerase activity in human mesenchymal stem cells.* Leukemia, 2003. **17**(6): p. 1146-9.
60. Ramamoorthy, M. and S. Smith, *Loss of ATRX Suppresses Resolution of Telomere Cohesion to Control Recombination in ALT Cancer Cells.* Cancer Cell, 2015. **28**(3): p. 357-69.
61. Porteus, M.H. and D. Baltimore, *Chimeric nucleases stimulate gene targeting in human cells.* Science, 2003. **300**(5620): p. 763.
62. Miller, J.C., et al., *An improved zinc-finger nuclease architecture for highly specific genome editing.* Nat Biotechnol, 2007. **25**(7): p. 778-85.
63. Sander, J.D., et al., *Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA).* Nat Methods, 2011. **8**(1): p. 67-9.
64. Wood, A.J., et al., *Targeted genome editing across species using ZFNs and TALENs.* Science, 2011. **333**(6040): p. 307.
65. Christian, M., et al., *Targeting DNA double-strand breaks with TAL effector nucleases.* Genetics, 2010. **186**(2): p. 757-61.
66. Zhang, F., et al., *Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription.* Nat Biotechnol, 2011. **29**(2): p. 149-53.
67. Hockemeyer, D., et al., *Genetic engineering of human pluripotent cells using TALE nucleases.* Nat Biotechnol, 2011. **29**(8): p. 731-4.
68. Sanjana, N.E., et al., *A transcription activator-like effector toolbox for genome engineering.* Nat Protoc, 2012. **7**(1): p. 171-92.
69. Ishino, Y., et al., *Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product.* J Bacteriol, 1987. **169**(12): p. 5429-33.
70. Ran, F.A., et al., *Genome engineering using the CRISPR-Cas9 system.* Nat Protoc, 2013. **8**(11): p. 2281-308.
71. Richter, C., J.T. Chang, and P.C. Fineran, *Function and regulation of clustered regularly interspaced short palindromic repeats (CRISPR) / CRISPR associated (Cas) systems.* Viruses, 2012. **4**(10): p. 2291-311.
72. Horvath, P. and R. Barrangou, *CRISPR/Cas, the immune system of bacteria and archaea.* Science, 2010. **327**(5962): p. 167-70.
73. Marraffini, L.A. and E.J. Sontheimer, *CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea.* Nat Rev Genet, 2010. **11**(3): p. 181-90.
74. Groenen, P.M., et al., *Nature of DNA polymorphism in the direct repeat cluster of Mycobacterium tuberculosis; application for strain differentiation by a novel typing method.* Mol Microbiol, 1993. **10**(5): p. 1057-65.
75. Mojica, F.J., et al., *Long stretches of short tandem repeats are present in the largest replicons of the Archaea Haloferax mediterranei and Haloferax volcanii and could be involved in replicon partitioning.* Mol Microbiol, 1995. **17**(1): p. 85-93.

76. Klenk, H.P., et al., *The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon Archaeoglobus fulgidus.* Nature, 1997. **390**(6658): p. 364-70.
77. Mojica, F.J., et al., *Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria.* Mol Microbiol, 2000. **36**(1): p. 244-6.
78. Jansen, R., et al., *Identification of genes that are associated with DNA repeats in prokaryotes.* Mol Microbiol, 2002. **43**(6): p. 1565-75.
79. Mojica, F.J., et al., *Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements.* J Mol Evol, 2005. **60**(2): p. 174-82.
80. Bolotin, A., et al., *Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin.* Microbiology, 2005. **151**(Pt 8): p. 2551-61.
81. Pourcel, C., G. Salvignol, and G. Vergnaud, *CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies.* Microbiology, 2005. **151**(Pt 3): p. 653-63.
82. Makarova, K.S., et al., *A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action.* Biol Direct, 2006. **1**: p. 7.
83. Haft, D.H., et al., *A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes.* PLoS Comput Biol, 2005. **1**(6): p. e60.
84. Barrangou, R., et al., *CRISPR provides acquired resistance against viruses in prokaryotes.* Science, 2007. **315**(5819): p. 1709-12.
85. Marraffini, L.A. and E.J. Sontheimer, *CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA.* Science, 2008. **322**(5909): p. 1843-5.
86. Makarova, K.S., et al., *Evolution and classification of the CRISPR-Cas systems.* Nat Rev Microbiol, 2011. **9**(6): p. 467-77.
87. Grissa, I., G. Vergnaud, and C. Pourcel, *The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats.* BMC Bioinformatics, 2007. **8**: p. 172.
88. Stern, A., et al., *Self-targeting by CRISPR: gene regulation or autoimmunity?* Trends Genet, 2010. **26**(8): p. 335-40.
89. Agari, Y., et al., *Transcription profile of Thermus thermophilus CRISPR systems after phage infection.* J Mol Biol, 2010. **395**(2): p. 270-81.
90. Pougach, K., et al., *Transcription, processing and function of CRISPR cassettes in Escherichia coli.* Mol Microbiol, 2010. **77**(6): p. 1367-79.
91. Pul, U., et al., *Identification and characterization of E. coli CRISPR-cas promoters and their silencing by H-NS.* Mol Microbiol, 2010. **75**(6): p. 1495-512.
92. van der Oost, J., et al., *CRISPR-based adaptive and heritable immunity in prokaryotes.* Trends Biochem Sci, 2009. **34**(8): p. 401-7.
93. Sorek, R., V. Kunin, and P. Hugenholtz, *CRISPR--a widespread system that provides acquired resistance against phages in bacteria and archaea.* Nat Rev Microbiol, 2008. **6**(3): p. 181-6.
94. Deltcheva, E., et al., *CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III.* Nature, 2011. **471**(7340): p. 602-7.

95. Jinek, M., et al., *A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.* Science, 2012. **337**(6096): p. 816-21.
96. Garneau, J.E., et al., *The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA.* Nature, 2010. **468**(7320): p. 67-71.
97. Deveau, H., et al., *Phage response to CRISPR-encoded resistance in Streptococcus thermophilus.* J Bacteriol, 2008. **190**(4): p. 1390-400.
98. Mali, P., et al., *RNA-guided human genome engineering via Cas9.* Science, 2013. **339**(6121): p. 823-6.
99. *CRISPR/Cas9 Guide*. Available from: https://www.addgene.org/crispr/guide/.
100. Gilbert, L.A., et al., *CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes.* Cell, 2013. **154**(2): p. 442-51.
101. Fujii, H. and T. Fujita, *Isolation of Specific Genomic Regions and Identification of Their Associated Molecules by Engineered DNA-Binding Molecule-Mediated Chromatin Immunoprecipitation (enChIP) Using the CRISPR System and TAL Proteins.* Int J Mol Sci, 2015. **16**(9): p. 21802-12.
102. Chen, B., et al., *Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system.* Cell, 2013. **155**(7): p. 1479-91.
103. Chen, S., et al., *Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis.* Cell, 2015. **160**(6): p. 1246-60.
104. Cong, L., et al., *Multiplex genome engineering using CRISPR/Cas systems.* Science, 2013. **339**(6121): p. 819-23.
105. Hsu, P.D., et al., *DNA targeting specificity of RNA-guided Cas9 nucleases.* Nat Biotechnol, 2013. **31**(9): p. 827-32.
106. Fu, Y., et al., *High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells.* Nat Biotechnol, 2013. **31**(9): p. 822-6.
107. Liang, P., et al., *CRISPR/Cas9-mediated gene editing in human tripronuclear zygotes.* Protein Cell, 2015. **6**(5): p. 363-72.
108. Baltimore, D., et al., *Biotechnology. A prudent path forward for genomic engineering and germline gene modification.* Science, 2015. **348**(6230): p. 36-8.
109. Evers, B.M., et al., *The human carcinoid cell line, BON. A model system for the study of carcinoid tumors.* Ann N Y Acad Sci, 1994. **733**: p. 393-406.
110. Chen, B. and B. Huang, *Imaging genomic elements in living cells using CRISPR/Cas9.* Methods Enzymol, 2014. **546**: p. 337-54.
111. Chen, B., et al., *Expanding the CRISPR imaging toolset with Staphylococcus aureus Cas9 for simultaneous imaging of multiple genomic loci.* Nucleic Acids Res, 2016. **44**(8): p. e75.

## SUPPLEMENTARY DATA

## SUPPLEMENTARY FIGURES

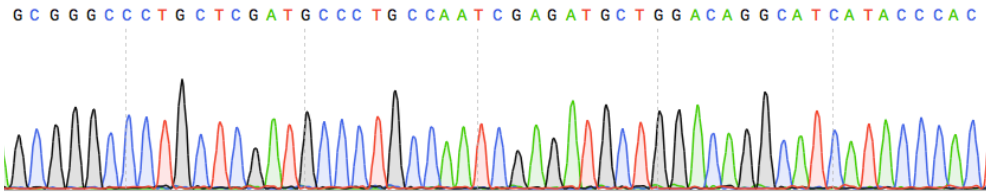


**Figure 1 –** Partial sequence of Tet3G vector.

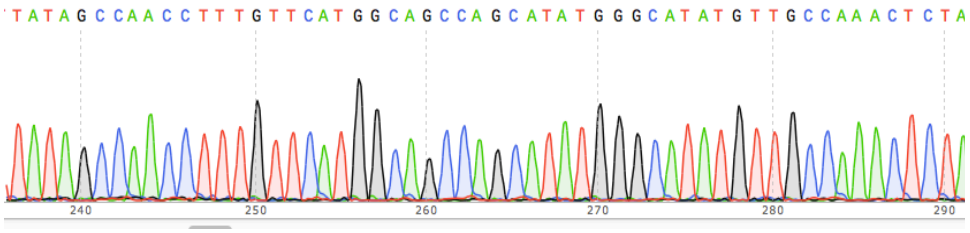



**Figure 2 –** Partial sequence of psPAX2 vector

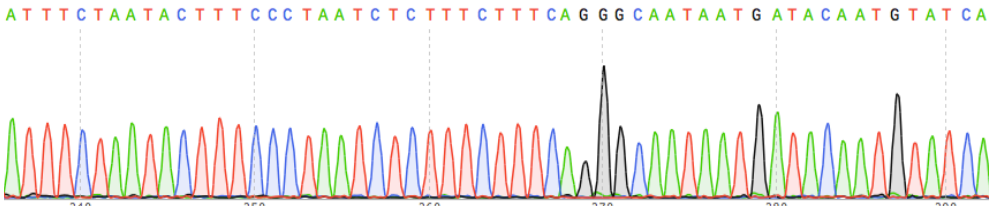


**Figure 3 –** Partial sequence of pMD2.G vector.

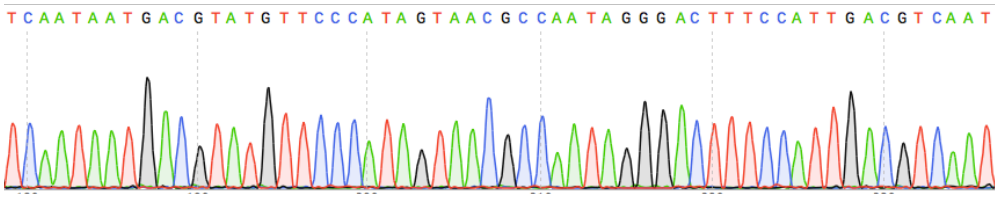


**Figure 4 –** Partial sequence of SgTelomere vector.

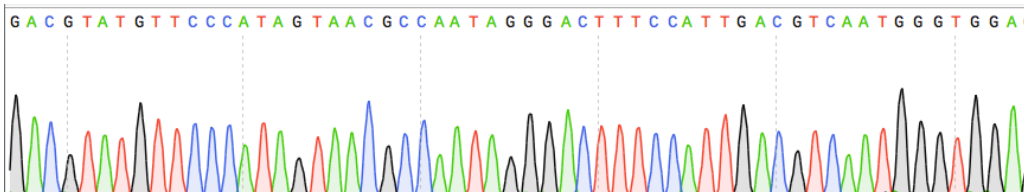


**Figure 5 –** Partial sequence of sgGAL4 vector.

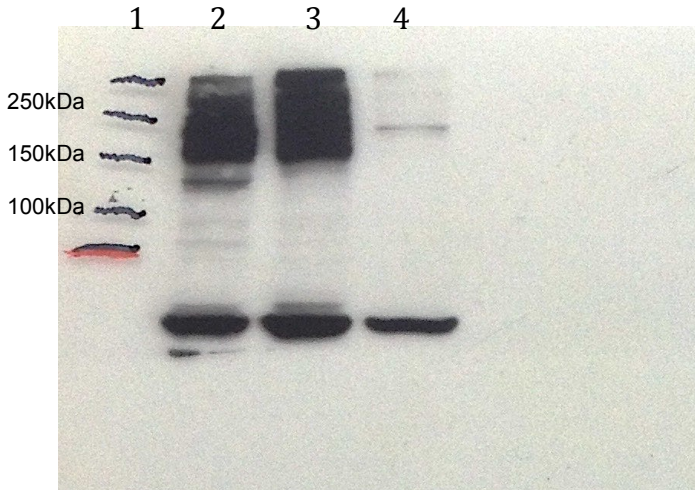



**Figure 6 –** Western Blot for the ATRX protein in Bon cells. 1 – Marker, 2 - Bon WT, 3 – Bon shRNA CTRL, 4 – Bon shRNA ATRX. As our secondary objective was to visualize and compare telomeric lengths in various Bon cells with different genetic backgrounds (WT, shRNA CTRL and shRNA ATRX), we performed Western Blot to confirm the KD of the ATRX protein in the latter type of cell. Indeed, the resulting membrane showed us that the KD was successfully established and ready for further experimenting as the bands in the membrane at ≈280kDa (the ATRX protein size) are much weaker in the shRNA ATRX well, compared to WT and shRNA CTRL.


**UNIVERSIDADE DO PORTO**