



















# Contents

<b>Abstract</b> .....	<b>ii</b>
<b>Keywords</b> .....	<b>iii</b>
<b>Résumé</b> .....	<b>iv</b>
<b>Mots-clés</b> .....	<b>vi</b>
<b>List of abbreviations</b> .....	<b>vii</b>
<b>Chapter 1. Introduction</b> .....	<b>11</b>
1.1 Telomere structure and function.....	11
1.2 Telomere maintenance mechanisms .....	16
1.3 Telomere transcription .....	18
1.4 TERRA's role in telomere structure .....	21
1.5 TERRA regulation of telomere length and the ICF syndrome .....	24
1.5 Thesis outline .....	28
<b>Chapter 2. Quantification of TERRA in different cell lines</b> .....	<b>29</b>
2.1 Introduction.....	29
2.2 Results.....	31
Characterization of TERRA expression in HCT116 DKO cells.....	31
TERRA and <i>WASH</i> transcripts behave differently in DKO cells.....	36
ZNF148 is a repressor of TERRA transcription .....	37
2.3 Discussion and Perspectives .....	40
2.4 Project contributions.....	43
<b>Chapter 3. Role of TERRA in telomere length and chromatin structure</b> .....	<b>44</b>
3.1 Introduction.....	44
3.1.1 Telomere length dynamics .....	44

3.1.2 CRISPR interference and gene activation .....	46
3.2 Results .....	50
TERRA expression and telomere length do not directly correlate in DKO cells.....	50
Telomere length dynamics are changed in DKO cells.....	53
DKO cells fail to downregulate TERRA in S-phase.....	56
TERRA levels are not increased in DNMT3B KO cells .....	58
Tentative downregulation of TERRA using CRISPR interference .....	60
TERRA upregulation using CRISPR-based gene activation .....	63
Long-term TERRA upregulation does not cause telomere shortening <i>in cis</i> .....	66
Changes in telomeric and subtelomeric chromatin caused by TERRA .....	71
3.3 Discussion and perspectives .....	75
<b>Materials and Methods .....</b>	<b>83</b>
Cell culture and transfection .....	83
Plasmids and gRNAs.....	83
Lentiviral vector production and titration.....	83
Retroviral vector production and transduction .....	84
Western blot .....	85
Telomere restriction fragment analysis (TRF) .....	85
Single telomere length analysis (STELA) .....	86
Cell cycle sorting.....	87
RT-qPCR.....	87
TERRA absolute quantification .....	88
Northern blot.....	88

IF and TERRA FISH.....	89
DNMT3B CRISPR/Cas9 deletion .....	90
siRNA depletion.....	90
ChIP-dot blot and ChIP-qPCR.....	91
<b>References.....</b>	<b>97</b>
<b>Acknowledgements .....</b>	<b>108</b>



# Chapter 1. Introduction

## 1.1 Telomere structure and function

Telomeres are complex structures present at the end of every linear chromosome. They consist of DNA, RNA and associated proteins that must interact in harmony to correctly shape and protect chromosome ends (Palm and de Lange, 2008).

In mammals, the telomeric DNA typically comprise 2 to 50 kb of double-stranded 5'-TTAGGG-3' repeats. Mouse cells show telomeres of up to 50 kb, while human telomeres lie between 5 and 15 kb (Palm and de Lange, 2008). This long double-stranded sequence ends in a G-rich 3' overhang of 50 to 500 nucleotides in length. This overhang is generated by a tightly regulated process. Telomeric DNA replication generates a blunt-ended chromatid (leading strand) and another chromatid with a short overhang corresponding to removed initiator RNA primer (lagging strand). Initially, a short resection of the leading strand is performed by the Apollo nuclease (Lam et al., 2010; Wu et al., 2010). Further resection of both leading and lagging is performed by exonuclease I (ExoI), which results in over-resected telomere ends (Wu et al., 2012). This is corrected by the CST (CTC1-STN1-TEN) complex, possibly by engaging polymerase  $\alpha$ -primase in the fill-in reaction (Dai et al., 2010; Wu et al., 2012, 2012). The single-stranded overhang is able to invade the preceding double-stranded telomeric sequence and form a protective structure at chromosome ends. This structure was first observed using electron microscopy and was referred to as T-loop (Griffith et al., 1999). The 3' overhang is also essential for the activity of telomerase. The RNA component of telomerase (hTR) needs to anneal to the overhang for the catalytic subunit (hTERT) to add *de novo* telomeric repeats to the overhang.

Upstream of TTAGGG repeats are chromosomal specific sequences called the subtelomeres. Subtelomeric sequences are partially conserved segmental

duplications and often present telomeric-like repeats (Riethman et al., 2005). About two thirds of the assembled subtelomeres present high-density CpG-like islands.

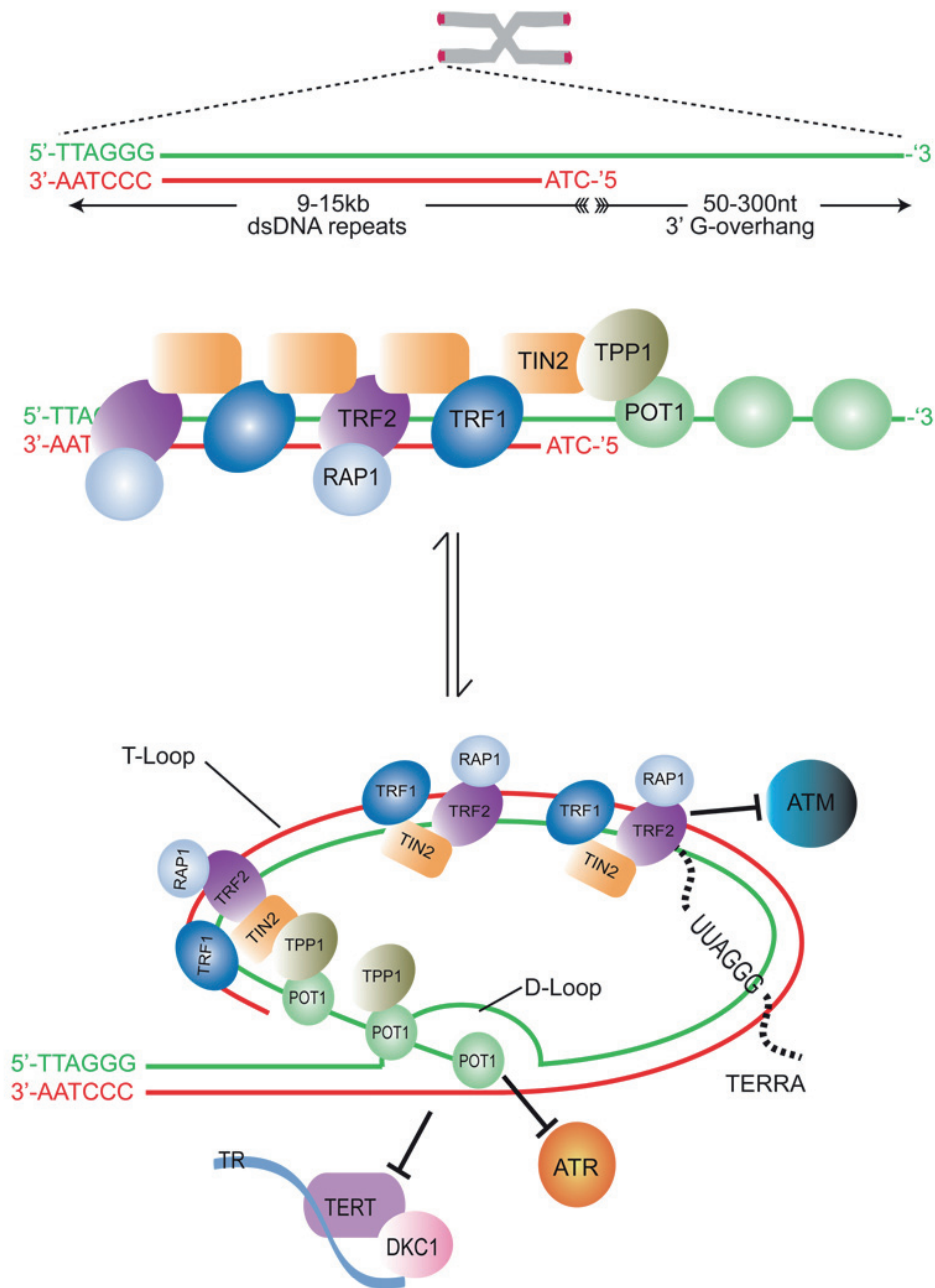
Telomeres and subtelomeres are heterochromatic regions and were believed to be devoid of transcriptional activity. However, in 2007 the telomeric-repeating containing RNA (TERRA) was discovered in human and mouse cells (Azzalin et al., 2007). TERRA was also identified in many other species, demonstrating its conserved nature (Bah et al., 2012; Luke et al., 2008; Rudenko and Van der Ploeg, 1989; Vrbsky et al., 2010). TERRA is now considered an integral part of telomeres, together with the telomeric DNA and associated proteins.

Telomeres are essential to maintain genomic integrity. With each cell cycle, telomeres inevitably erode due to the mechanisms generating the 3' overhang (Lingner et al., 1995) and the inability of DNA polymerase to complete the replication of the lagging telomeric strand (Watson, 1972). Telomeres which do not bear a sequence long enough for the binding of sufficient telomeric proteins are considered critically short. Insufficient amounts of the proteins TRF2 and POT1 (introduced below) can lead to ATM and ATR activation, respectively, and unleash DNA damage responses (DDR) at telomeres (Denchi and de Lange, 2007). Unresolved DDR and persistent ATM activation eventually promote p53-dependent senescence onset (Deng et al., 2008). Replicative senescence limits the number of divisions that a cell can make during its life (Hayflick's limit) and plays an important role in limiting tumor appearance and progression (Campisi, 2005).

Unlike mammalian cells, the telomeric DNA sequence of *S. cerevisiae* is more heterogenous, comprising  $300 \pm 75$  bp of different combinations of  $C_{1-3}A/TG_{1-3}$ . The dsDNA is covered by the double-strand binding protein RAP1, its binding partners Rif1 and Rif2, and other proteins. Budding yeast telomeres also finish in a 3' overhang and are transcribed into TERRA (Wellinger and Zakian, 2012). This overhang is bound by the Cdc13 complex (Cdc13, Stn1 and Ten1 proteins), which caps the single stranded telomeric DNA and recruits telomerase (Wu and Zakian, 2011).

In human telomeres, a core complex of 6 proteins, known as shelterin, abundantly binds to and protects the telomere DNA (Palm and de Lange, 2008). Two proteins, TRF1 and TRF2 (Telomere repeat factor 1 and 2) directly bind the double stranded TTAGGG repeats (5'-YTAGGGTTR-3') as homodimers through their Myb domains (Court et al., 2005). TRF1 promotes proper telomeric DNA replication and helps to control telomere length (Loayza and de Lange, 2003; Sfeir et al., 2009). TRF2 plays an important role in the establishment and protection of the T-loop. Through its basic domain, TRF2 is able to promote strand invasion and stabilize the three-way T-loop junction, protecting the T-loop from cleavage by Holliday Junction resolvases (Schmutz et al., 2017). TRF2 is essential for the repression of ATM signaling at telomeres. TRF2 could potentially inhibit ATM activation, since it was shown to co-immunoprecipitate with this kinase and to inhibit its autophosphorylation, which happens early in the DDR cascade (Karlseder et al., 2004). Another model suggests that this is achieved indirectly by the formation of the T-loop, where the telomeric DNA end is hidden from DNA damage detectors (Doksani et al., 2013; Griffith et al., 1999; Schmutz et al., 2017). More recently, the dimerization domain of TRF2 (TRFH) was shown to promote T-loop formation and to prevent ATM activation in MEFs. T-loop linearization correlated with telomeric DDR activation, but was not caused by the DDR, suggesting that the T-loop conformation plays a central role in telomere protection (Van Ly et al., 2018). In the absence of TRF2, ATM signaling is activated, which leads to ligase IV-dependent end-to-end chromosome fusions (Denchi and de Lange, 2007; Smogorzewska et al., 2002).

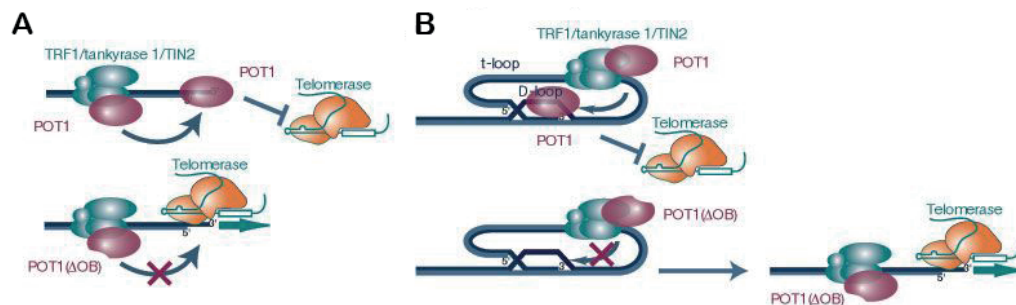
Another protein that binds telomeric DNA directly is POT1 (Protection of telomeres 1). POT1 recognizes single-stranded 5'-(T)TAGGGTTAG-3' sequences and is abundantly present in the telomeric 3' overhang (Loayza et al., 2004). By binding the overhang, POT1 inhibits ATR activation and NHEJ at telomeres (Denchi and de Lange, 2007). POT1 is present at telomeres as a heterodimer with TPP1. The N-terminus of TPP1's OB fold domain recruits telomerase and, together with POT1, stimulates telomerase activity *in vitro* (Grill et al., 2018; Wang et al., 2007). However, the POT1 protein alone inhibits telomerase activity,



**Figure 1: Telomere structure, shelterins and T-loop.** Human telomeres consist of up to 15 kb of double stranded TTAGGG repeats ending in a 3' G-rich overhang. The telomeric DNA is bound by a core complex of 6 proteins, referred to as shelterin (TRF1, TRF2, TIN2, TPP1, POT1 and RAP1). A number of other known and unknown proteins participate in the protection and maintenance of the telomere (not shown). The single stranded overhang is able to loop back into the double stranded sequence, forming the T-loop and a D-loop. POT1 inhibits ATR activation and telomere elongation by telomerase (TR, TERT and DKC complex) by binding and hiding the single stranded telomeric DNA and/or by stabilizing the D-loop. TRF2 is able to directly bind TERRA and protect telomeres against ATM-dependent DDR. Figure from O'Sullivan and Karlseder, 2010.



possibly by competing with the enzyme for binding to its substrate (Kelleher et al., 2005). *In vivo*, overexpression of a POT1 mutant lacking its OB domain (POT1 $\Delta$ OB) leads to rapid increase in telomere length (Loayza and de Lange, 2003). This mutant is still able to localize to telomeres, but fails to bind the 3' overhang. Two models were proposed to explain this behavior. By binding the 3' overhang, POT1 could directly prevent the binding of telomerase to the overhang and hamper telomere elongation. Alternatively, POT1 could stabilize the D-loop formed at the base of the T-loop and prevent telomere elongation by sequestering the 3' overhang (Figure 2) (Loayza and de Lange, 2003).



**Figure 2: Model proposing the mechanisms by which POT1 helps to control telomere length. (A)** POT1, recruited to telomeres via the TRF1 complex (TRF1, tankyrase and TIN2) binds the 3' overhang and hampers the binding of telomerase to the overhang. **(B)** POT1 helps to stabilize the D-loop at the base of the T-loop. This structure hinders telomerase binding to 3' overhang. In the absence of POT1's OB domain, the telomere end may become linear and allow for telomerase-dependent telomere elongation (image and model from Loayza and de Lange, 2003).

TIN2 is the protein that stabilizes the entire shelterin complex. It binds TRF1 and TRF2 homodimers and recruits TPP1 and POT1 to the telomeres. TIN2 knockout induces phenotypes similar to those seen when TPP1 and POT1 are absent, namely activation of ATM and ATR signaling, excessive 3' overhang formation, chromatid and chromosome fusions, as well as disproportionate loading of RPA to the 3' overhang (Takai et al., 2011). RPA and POT1 have very similar affinity for the telomeric ssDNA, however RPA is in 1,000-fold excess over POT1. POT1 is

probably only able to out-compete RPA and prevent ATR activation because it is physically tethered to the telomeres via TPP1 and TIN2 (Takai et al., 2011).

The last component of the shelterin complex is RAP1. Human cells and mice knockout for RAP1 are viable and do not show any major telomeric dysfunction. Overexpression of RAP1 leads to telomere elongation in HTC75 cells and its C-terminus domain is responsible for its recruitment to telomeres via TRF2 binding (Li et al., 2000). RAP1 and Ku70 double knockout in mouse embryonic fibroblasts (MEFs) leads to an increase in telomere recombination (Sfeir et al., 2010). RAP1 also seems to cooperate with TRF2 to avoid homology directed repair at telomeres and the generation of telomere-free chromosome fusions (Rai et al., 2016).

Together, these six proteins form the most abundant complex present at telomeres. They perform essential tasks related to telomere protection, replication, length maintenance and structure. However, they also rely on many other players to achieve their goals properly. The list of non-shelterin proteins associated with the telomeres grows continuously. Methods such as QTIP, PiCh and BioID have revealed new proteins important in the telomere biology and makes us realize that we are still only scratching the surface of such complex structures (Déjardin and Kingston, 2009; Garcia-Exposito et al., 2016; Grolimund et al., 2013; Majerská et al., 2017).

## 1.2 Telomere maintenance mechanisms

The loss of telomeric sequence that happens with each cycle of DNA replication has an important role in limiting cellular lifespan. Therefore, tumor cells must find a way to circumvent this problem.

The vast majority of tumors **reactivate telomerase expression**. Telomerase is a ribonucleoprotein complex consisting at its core of a catalytic (hTERT) and an RNA subunit (hTR). Telomerase localizes to Cajal bodies and is recruited to telomeres in S-phase via interaction with the TEL (TPP1 glutamate and leucine-

rich) patch in the OB-fold domain of TPP1. hTR binds to the telomeric 3' overhang and serves as a primer and a template for the addition of new telomeric repeats by hTERT (Schmidt and Cech, 2015). Telomerase is mostly active in fast dividing cells, such as germ line and stem cells, while hTERT expression is suppressed in human somatic cells. Recurrent promoter mutations have been shown to create new binding sites for transcription factors and increase hTERT expression in cancer cells (Horn et al., 2013; Huang et al., 2013).

**Alternative lengthening of telomeres (ALT)** is used by a minority of tumors to maintain telomere length. It is preferred by certain tumors of mesenchymal origin, such as sarcomas (leiomyosarcoma, chondrosarcoma and undifferentiated pleomorphic sarcoma) and astrocytomas (diffuse and anaplastic astrocytomas) (Heaphy et al., 2011a). ALT uses recombination to elongate telomeres. The product of such telomeric recombinations generates C-circles, whose presence is often used to determine if a tumor is ALT-positive (Henson et al., 2009). Mutations in ATRX and DAXX, proteins forming a chromatin remodeling complex, are commonly found in ALT tumors and cell lines (Heaphy et al., 2011b; Lovejoy et al., 2012). Loss of ATRX function is associated with high levels of genome instability and increased telomere length (Lovejoy et al., 2012). Although associated with ALT, ATRX mutations are not sufficient for ALT establishment, since these mutations are also occasionally found in telomerase-positive tumors (Lee et al., 2018).

Another feature of ALT cells is increased telomere transcription. Subtelomeric CpG islands, believed to drive the transcription of the telomeric long non-coding RNA TERRA, are less methylated in ALT than in telomerase-positive cell lines, and this phenomenon is associated with increased TERRA expression (Arora et al., 2014; Nergadze et al., 2009). In the absence of telomerase, TERRA can promote maintenance of telomere length by telomeric recombination, delaying replicative senescence onset and preserving genomic stability (Arora et al., 2014; Graf et al., 2017).

Although genetically very different, these telomere maintenance mechanisms are not mutually exclusive. One melanoma has been recently shown to possess both

telomerase promoter mutations and ALT phenotypes simultaneously (Lee et al., 2018). Other 8 tumors are suspected to behave similarly.

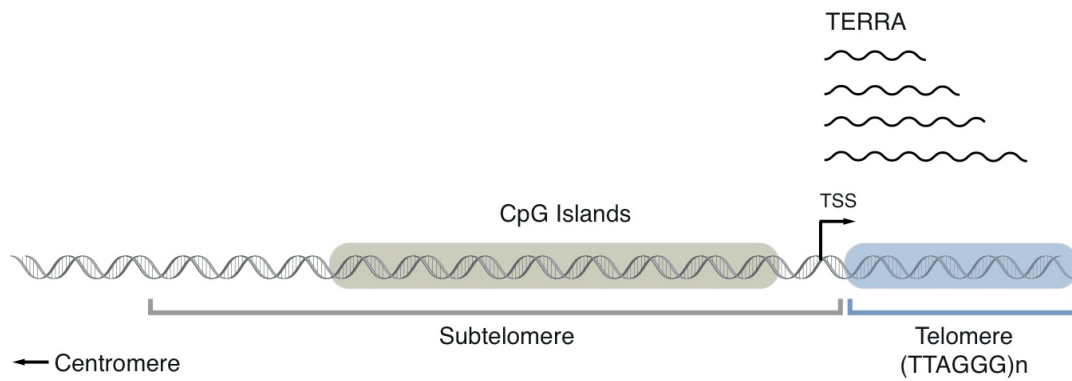
### 1.3 Telomere transcription

Due to the heterochromatic state of chromosome ends, telomeres were initially believed to be transcriptionally inactive. With the discovery of human lncRNA TERRA, it became clear that most eukaryotes transcribe their telomeres (Azzalin et al., 2007; Bah et al., 2012; Luke et al., 2008; Rudenko and Van der Ploeg, 1989; Schoeftner and Blasco, 2008; Solovei et al., 1994; Vrbsky et al., 2010).

TERRA molecules are capped by 7-methylguanosine and are rarely (<10%) polyadenylated. Non-polyadenylated TERRA has a 3h half-life and partially associates with telomeres, while polyA TERRA is more stable and vastly nucleoplasmatic (Porro et al., 2010, 2014). TERRA has been reported to interact with several telomere binding proteins, such as TRF1, TRF2, HP1 and some ORC subunits (Deng et al., 2009).

TERRA transcription starts at the subtelomeres and continues towards the end of the chromosome (Figure 3). Its length varies from 100 to 9,000 bases. The origin of this variation comes from both the distance between the transcription start site (TSS) to the telomeres and from how much the transcription progresses into the telomeric repeats (Porro et al., 2014).

In most eukaryotic cells, telomere transcription has only been characterized as proceeding towards the chromosome ends. *Schizosaccharomyces pombe* and *S. cerevisiae* were also shown to express ARRET, an RNA complementary to the subtelomeric sequence of TERRA, but lacking telomeric repeats (Bah et al., 2012; Luke et al., 2008). In addition, *S. pombe* expresses an RNA similar to TERRA, but lacking telomeric repeats ( $\alpha$ ARRET), and ARIA, a C-rich telomeric DNA devoid of subtelomeric sequences (Bah et al., 2012).



**Figure 3:** TERRA transcription often starts in the subtelomeric sequences and continues towards the chromosome end. Subtelomeric CpG islands, including the 61-29-37 bp repeats, are thought to act as TERRA promoters. The presence of unique subtelomeric sequences in the 5' end of TERRA allows us to determine from which chromosome TERRA comes from. Consequently, we are currently unable to identify and quantify TERRAs whose transcription starts within the telomeric repeats.

RNA polymerase II is responsible for TERRA transcription (Azzalin et al., 2007; Schoeftner and Blasco, 2008). The exact location of TERRA's TSS is difficult to define, but the few mapped TSSs lay close to the subtelomeric CpG islands (Porro et al., 2014). Many of these islands consist of variations of repetitive sequences known as 61-29-37 bp repeats, which are thought to act as TERRA promoters (Nergadze et al., 2009).

Up to 20 subtelomeric CpG have been predicted (Nergadze et al., 2009), but little is known about their role in TERRA transcription and regulation. It is also currently unknown if the presence of a CpG island is a determinant of TERRA transcription.

Participating in the regulation of TERRA transcription is DNMT3B, which methylate subtelomeric CpG islands, thereby inhibiting TERRA transcription (Yehezkel et al., 2008). Methylation of CpG islands is thought to silence transcription either by directly inhibiting the binding of methylation-sensitive transcription factors, or by recruiting methyl-CpG binding domain (MBD) proteins, such as Kaiso, MeCP2 and MBD1/2/3/4. These MBDs are able to interact with several chromatin modifiers, such as histone methyltransferases, histone deacetylases and HP1, to silence the nearby chromatin (Bogdanović and Veenstra, 2009; Deaton and Bird, 2011).

While subtelomeric CpG island methylation inhibits TERRA transcription, CTCF and Rad21 proteins, which bind in the proximity of these islands, are found to increase telomere transcription (Deng et al., 2012).

TERRA expression is cell cycle regulated. The highest TERRA levels are found in G1, while levels decrease throughout S-phase and reach minimal values in the transition between late-S and G2 phases. During G2, TERRA starts to accumulate again, leading to maximum levels in G1 (Porro et al., 2010). It is not yet addressed if the decreased TERRA levels are due to decreased transcription or increased degradation.

*In vitro* experiments suggest that changes in TERRA levels during the cell cycle are important to allow the binding of POT1 to the 3' overhang after DNA replication (Flynn et al., 2011). In S phase, RPA is loaded onto single stranded telomeric DNA. According to this model, POT1 may not be able to displace RPA after replication and requires assistance of hnRNPs. In the beginning of S-phase, TERRA would bind hnRNPA1 and prevent it from displacing RPA from the telomeres. Upon decrease in TERRA levels in late S, hnRNPA1 would strip RPA out of the telomere and give an opportunity for POT1 binding (Flynn et al., 2011). An alternative model suggests that the local tethering of POT1 to the telomeres via TPP1 and TIN2 ensures local POT1 concentrations high enough to induce the RPA-POT1 switch without aid from TERRA (Takai et al., 2011).

While the cell cycle regulation of TERRA appears to be important for telomerase-positive cells, the number of TERRA foci is unchanged in U2OS and HUO9 ALT cell lines during S-phase. Cell cycle regulation of TERRA is dependent on a protein frequently mutated in ALT cells, the chromatin remodeler ATRX. ATRX depletion in telomerase-positive cells leads to stable TERRA levels in S-phase and persistent loading of RPA to telomeric DNA after replication, structure that can potentially assist telomere recombination in ALT cells (Flynn et al., 2015).

## 1.4 TERRA's role in telomere structure

Several heterochromatin marks, such as H3K9me3, H3K27me3 and H4K20me3, are commonly found in telomeres and subtelomeres of mouse and human cells (Blasco, 2007). Although originally thought to be highly heterochromatic, a recent report claims that they are not (Cubiles et al., 2018). Using ChIP-seq, the authors demonstrate that, compared to Satellites II and III regions, telomeres harbour less H3K9me3 and H3K27me3, while having more open chromatin marks, such as H4K20me1, H3K27ac, H3K4me3 and H3K36me3. TERRA is believed to contribute to heterochromatin formation at telomeres due to its ability to directly bind SUV39H1 and enhance the deposition of H3K9me3 marks at telomeres upon telomere uncapping (Porro et al., 2014).

Clones with decreased TERRA expression have less telomeric H3K9me3, H4K20me3 and H3K27me3. Concomitantly, they also show less PCR2 at the telomeres, which could lead to the reduced heterochromatin marks (Montero et al., 2018). In TERRA-reduced clones or when siRNAs against TERRA are used, telomere-free chromosome ends and other chromosome anomalies are more frequent (Chu et al., 2017; Deng et al., 2009; Montero et al., 2016). This indicates a role for TERRA in the normal telomere structure, perhaps serving as a scaffold for binding or stabilization of different proteins. However, these results need careful interpretation. High levels of UUAGGG-containing siRNAs could compete with TERRA and telomeres for binding proteins, as well as bind to the telomeric 3' overhang, interfering with the telomere structure. Additionally, the TERRA-reduced clones were created by inducing double-strand breaks in the 20q and potentially other subtelomeres (Montero et al., 2016), which could by themselves generate such telomeric anomalies.

More recently, TERRA was also shown to compete with the chromatin remodeler ATRX not only for binding at telomeres, but also in other loci in the mouse genome. Using TERRA anti-sense capture probes and deep sequencing, more than 4,000 TERRA binding sites were identified. Only a fraction of them were telomeric or subtelomeric sites, demonstrating the ability of TERRA to bind chromatin throughout the genome. The same probes were used to capture

proteins interacting with TERRA. 134 proteins were identified, including the chromatin remodeller ATRX. Analysis of previously published ATRX ChIP-seq data revealed that more than 600 peaks are shared between this protein and TERRA. The expression of genes nearby these common peaks, such as *Mid1*, *Erdr1*, *Fyco1*, *Lphn2* and *Nfib* were downregulated by TERRA depletion by LNA gapmers, but upregulated upon ATRX depletion, indicating a functional antagonism between them. Moreover, TERRA was shown to outcompete telomeric DNA for ATRX binding *in vitro*, and TERRA knockdown increased the localization of ATRX to telomeres *in vivo*. It is hypothesized that the same mechanism is used by TERRA to influence gene expression in non-telomeric targets (Chu et al., 2017).

TERRA transcripts are capable of invading the double-stranded telomeric DNA and form TERRA:DNA hybrids, or R-loops. These structures represent a challenge for transcription and replication machinery and pose a threat to genomic stability (Aguilera and García-Muse, 2012). Because of that, cells developed several mechanisms to regulate R-loops, including RNase H1 and H2, which degrade the RNA in the R-loop, and the Pif1 helicase, able to unwind remaining hybrids (Arora et al., 2014; Boule and Zakian, 2007; Paeschke et al., 2013).

At telomeres, R-loops can also cause problems to the displaced telomeric strand. G-rich single-strands can form G-quadruplexes, which need to be removed by helicases, such as BLM and WRN (Drosopoulos et al., 2015), before replication can proceed. On the other hand, ALT cells seem to profit from the instability caused by R-loops at telomeres, which promotes telomere length maintenance by recombination. Arora and colleagues showed that ALT telomeres accumulate more R-loops and RNaseH1 than telomeres maintained by telomerase. They demonstrated that the levels of R-loops need to be strictly controlled to preserve telomere integrity in these cells. Whereas RNaseH1 overexpression leads to accumulation of telomere-free chromosome ends, probably due to impaired HR at telomeres, RNaseH1 depletion leads to replication stress, exaggerated telomere recombination (as measured by C-circle quantification) and sudden telomere loss (Arora et al., 2014). The authors propose that TERRA:DNA hybrids



cause replication stress, leading to the loading of phosphorylated RPA at telomeres, which in turn could recruit RNaseH1 to remove excessive hybrids and maintain optimal levels of R-loops at telomeres.

Using *Saccharomyces cerevisiae* strains, Graf and colleagues also showed the importance of regulating telomeric R-loops (Graf et al., 2017). They induced sudden telomere shortening by activating the FLP1 recombinase in a single telomere and observed rapid increase in TERRA expression and R-loop accumulation *in cis*. The removal of the R-loops by RNaseH1 (*RNH1*) overexpression resulted in decreased loading of the recombinogenic protein Rad51 at this telomere and premature senescence. They also show that short telomeres, from cells grown for 60 PDs in the absence of telomerase, recruit less RNA exonuclease Rat1, which leads to higher TERRA levels. On the other hand, R-loop accumulation at short telomeres seems to be due to decreased recruitment of RNaseH2 (*RNH201*) by Rif2, as Rif2 deletion leads to less RNaseH2 and more R-loops at telomeres. The authors propose that short and long telomeres manage the amounts of R-loops differently. Long telomeres recruit enough Rif2, Rat1 and RNaseH2 to control TERRA levels and remove R-loops before the arrival of the replication fork, whereas short telomeres fail to do so. When the replication machinery meets R-loops, replication stress and DSBs may follow, potentially triggering homologous recombination and telomere elongation (Aguilera and Gómez-González, 2008; Graf et al., 2017; Hamperl and Cimprich, 2016). In line with this hypothesis, a recent report shows that R-loops occurring upon DSBs in highly transcribed genes can recruit Rad52, which in turn promotes R-loop processing and homology directed repair in non-telomeric loci (Yasuhara et al., 2018). Whether this mechanism plays a role in ALT or uncapped telomeres remains to be investigated.

## 1.5 TERRA regulation of telomere length and the ICF syndrome

The role of TERRA in telomere length regulation is told by a long and conflicting list of publications, where studies *in vitro*, in mice, in yeast and in human cells often disagree. Perhaps one of the most striking evidences of a relationship between TERRA expression and telomere length in humans comes from analysis of ICF patients' cells. ICF (Immunodeficiency, centromere instability and facial anomalies syndrome) is a rare autosomal recessive disease, with only 50 cases reported since the 1970s (Ehrlich et al., 2006). Patients suffer from severe immunodeficiency and recurrent infections, which leads to a poor life expectancy (Hagleitner et al., 2007). Immunoglobulin supplementation improves the course the disease (Hagleitner et al., 2007). Although bone marrow and hematopoietic stem cell transplantation have been performed successfully in ICF infants, they are not routinely indicated (Gössling et al., 2017). The link between TERRA expression and telomere length was suggested when primary cells of ICF patients were shown to have very short telomeres and abnormally high TERRA expression (Yehezkel et al., 2008).

The most common mutation found in ICF patients falls within the coding sequence of the *DNMT3b* gene (classified as ICF1 variant) (Ehrlich et al., 2006). The remaining patients without *DNMT3b* mutations could have mutations in promoters or transcription factors involved in *DNMT3b* regulation or in other genes (Ehrlich et al., 2006). More recently, *ZBTB24* mutations have been found in eight ICF2 patients (Chouery et al., 2012; de Greef et al., 2011).

DNMT3B is the main *de novo* methyltransferase in human and murine cells. Together with DNMT3A, it establishes methylation patterns at an early stage of embryogenesis (Okano et al., 1999). In mice, *DNMT3b* deletion is embryonically lethal, while lack of DNMT3A shows postnatal lethality (Li et al., 1992; Okano et al., 1999). Disruption of human *DNMT3a* and *DNMT3b* genes results in viable cell lines (Liao et al., 2015).

As a probable result of impaired *de novo* methylation during embryogenesis, ICF1 patients present hypomethylation of certain genomic sequences, including satellites 1, 2 and 3 (Jeanpierre et al., 1993) and subtelomeric CpG islands (Yehezkel et al., 2008). Since subtelomeric hypomethylation is associated with increased TERRA transcription, it is no surprise that ICF1 patients have extremely high TERRA expression (Nergadze et al., 2009; Yehezkel et al., 2008). Additionally, ICF1 patients have very short telomeres, which is likely responsible for the outcomes of the disease (Yehezkel et al., 2008). The relationship between telomere length, subtelomeric methylation and TERRA transcription is further strengthened by the observation that patients with ICF types 2, 3 and 4 have normal telomere length, methylation and TERRA levels (Toubiana et al., 2018).

The creation of induced pluripotent stem cells (iPSCs) from three ICF1 patients helped to bring some insight into the mechanisms of the disease (Sagie et al., 2014). Introduction of c-MYC, OCT4, SOX2 and KLF4 in primary fibroblasts led to telomerase upregulation, telomere elongation and allowed cells to overcome the premature senescence phenotype observed in the fibroblasts. However, subtelomeric CpG islands remained hypomethylated and abnormally high TERRA expression was sustained. This suggests that, in the presence of increased telomerase activity, telomeres can be elongated in spite of their high transcription levels. The situation was reversed upon re-differentiation of patient iPSCs: telomerase activity is reduced, telomeres shorten and premature senescence is triggered. The authors speculate that expression of telomerase during embryonic development may support ICF syndrome embryos to develop and reach birth, at which point differentiated cells will have short telomeres.

Similarly to the phenotype of ICF1 cells, *DNMT1* and *DNMT3b* knockout human colorectal carcinoma cells (HCT116 DKO) also show subtelomeric hypomethylation, elevated TERRA expression and short telomeres (Farnung et al., 2010; Nergadze et al., 2009), providing a more accessible model for the syndrome. The successful generation of the initial *DNMT1* KO cell line (Rhee et al., 2000) was surprising, since this condition was thought to be lethal in mice (Li et al., 1992). *DNMT3b* was soon after knocked out in this *DNMT1* KO clone (Rhee et al., 2002), creating the HCT116 DKO line we use today. Only several years later

it became clear that this cell line expresses a truncated version of DNMT1 resulting from alternative splicing (Egger et al., 2006; Spada et al., 2007). In fact, this cell line expresses a truncated form of the DNMT1 protein, which is able to maintain up to 50% of the WT methylation levels (Egger et al., 2006). Nowadays this cell line is commercialized as *DNMT1* ( $\Delta$ exons3-5/ $\Delta$ exons3-5) *DNMT3B* (-/-), but is still referred to as DKO in publications and in the present thesis.

Analysis of subtelomeric methylation revealed only a minor decrease in the single KO cell lines, while virtually no methylation remained in the DKO cells (Nergadze et al., 2009). *DNMT3b* KO cells were predicted to show increased TERRA levels because of the phenotype seen in ICF patients, however TERRA expression was only increased in the DKO. To add another level of complexity, DNMT3B overexpression in ICF1 fibroblasts failed to rescue subtelomeric methylation. Partial rescue was only seen when this protein was expressed together with *DNA methyltransferase 3-like (DNMT3L)*, which stimulates DNMT3A and DNMT3B activity and is normally expressed with DNMT3B during early embryonic development (Yehezkel et al., 2013).

The question remains about why ICF1 and DKO cells have short telomeres and if these the short telomeres are due to high TERRA expression. It seems unlikely that increased TERRA is a consequence of short telomeres. Although telomere shortening causes a mild increase in TERRA expression in *Saccharomyces cerevisiae* (Cusanelli et al., 2013), no evidence shows that this is the case in human cells, since telomere shortening caused by telomerase inhibition does not increase TERRA expression (Farnung et al., 2012). In parallel, elongation of telomeres by overexpression of telomerase leads to longer TERRA molecules, but causes only a 50% reduction on its transcription (Arnoult et al., 2012; Smirnova et al., 2013; Van Beneden et al., 2013).

Two studies have shown that TERRA inhibits telomerase *in vitro* (Redon et al., 2010; Schoeftner and Blasco, 2008). In this context, the reduction in TERRA levels in late S-phase was suggested to alleviate its inhibitory effect over telomerase, thereby allowing for telomere elongation (Chu et al., 2017; Porro et al., 2010). In *S. cerevisiae*, TERRA overexpression stimulates Exonuclease 1

activity at telomeres and thereby induces telomere shortening (Pfeiffer and Lingner, 2012). Furthermore, accumulation of TERRA R-loops in cells lacking the yeast Rat1 RNA exonuclease can serve as an obstacle for telomere maintenance and cause telomere shortening (Luke et al., 2008). The same appears to be true for human cells. ICF patient cells show increased formation of R-loops and more DNA damage markers at telomeres, both of which are reduced with the overexpression of RNase H1 (Sagie et al., 2017). Unresolved damage caused by TERRA:DNA hybrids can potentially contribute to telomere dysfunction in the disease.

Arguing against TERRA-dependent telomere shortening are two models described in *S. cerevisiae*. In one of them, TERRA was shown to be transcribed from short telomeres and then to co-localize with the *TLC1* telomerase RNA and its telomere of origin (Cusanelli et al., 2013). The authors speculate that TERRA produced specifically from short telomeres recruits telomerase *in cis* to induce the preferential elongation of short chromosome ends. More recently, Graf and colleagues showed that TERRA R-loops accumulate during S-phase at critically short telomeres, where they help to elicit DDR and HDR to elongate these telomeres (Graf et al., 2017). The authors suggest that accumulation of R-loops would not happen at long telomeres, because those harbour sufficiently long telomeric repeats to recruit enough Rap1, Rif2 and RNaseH2 molecules. RNaseH2 would be responsible for degrading TERRA and R-loops before the arrival of the replication fork, preventing DDR activation. By eliciting HDR exclusively at critically short telomeres, cells can avoid premature senescence and inadequate DDR activation and HDR at long telomeres. It should be pointed out that both studies were conducted in yeast and that the effect of TERRA in mammalian telomeres may differ significantly.

In an attempt to induce high TERRA transcription and observe changes in telomere length, an artificial transcriptionally inducible telomere (tiTEL) was inserted at a chromosome end of HeLa cells (Farnung et al., 2012). Upon doxycycline addition, transcriptionally inducible TERRA (tiTERRA) increased 10 to 20-fold, but this did not affect telomere elongation *in cis* (Farnung et al., 2012). It is important to notice that these experiments were performed upon

overexpression of telomerase and could have developed differently when limited to the endogenous expression of the enzyme. The tiTEL system also failed to induce telomere length changes in *S.pombe* (Moravec et al., 2016). However, some telomere elongation was seen with concomitant tiTERRA expression and trichostatin A treatment, which enhances TERRA expression and decreases heterochromatin formation.

tiTEL experiments and the fact that telomeres from DKO cells can be normally elongated by telomerase (Farnung et al., 2012) argue against a TERRA-dependent telomerase inhibition *in vivo* in human cells. Clearly there is still a lot we do not understand in the complex relationship between TERRA, telomerase and telomere length.

## 1.5 Thesis outline

The aim of this thesis is to better understand TERRA transcription and how it affects telomere maintenance and structure. Chapter 2 is fruit of a collaboration with Marianna Feretzaki and Julien Delafontaine. We describe the characterization of TERRA transcription in HeLa, U2OS, HLF, HCT116 WT and DKO cells, perform an siRNA screening for new TERRA transcription factors and validate one of them by CRISPR/Cas9 knockout.

Chapter 3 describes a quest to comprehend why HCT116 DKO cells have such short telomeres. Starting by investigating the telomere length maintenance dynamics and the cell cycle control of TERRA, and passing by attempts to create new DKO clones, we arrive at the use of CRISPR interference and activation systems to manipulate endogenous TERRA expression. With this system, we show that overexpression of the 10q and 13q TERRAs does not lead to telomere shortening *in cis*, but increases H4K20 and H3K9 trimethylation. Although there is a general increase in heterochromatin marks in the telomeres, it seems that the effects in subtelomeric sequences is exclusively *in cis*.

# Chapter 2. Quantification of TERRA in different cell lines

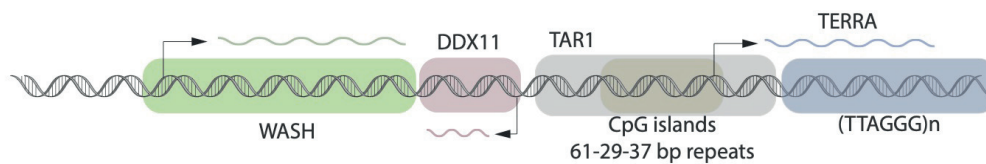
## 2.1 Introduction

The subtelomeres are generally considered to be the 500 kb of DNA sequence preceding the start of the TTAGGG-repeats. The first subtelomeric sequences were cloned in 1990 by Brown and colleagues (Brown et al., 1990). They succeeded in cloning 3.4 kb and 2.1 kb of the XqYq and 10q subtelomeres, respectively, into yeast artificial chromosomes. In these sequences, named TelBam3.4 and TelSau2.0, they identified shared conserved repetitive elements forming CpG islands of various sizes. The more centromere-proximal repeat is a tandem of 61-base-pairs (bp), which is followed by a 29-bp tandem repeat and a 37-bp tandem repeat, localized close to the telomeric repeats. The subtelomeric region containing these repeats is referred to as 61-29-37 bp repeats, and it is found not exclusively in the 10q and XqYq subtelomeres, but also in the 1p, 2p, 3q, 4p, 5p, 6p, 8p, 9p, 9q, 11p, 12p, 15q, 16p, 17p, 19p, 20p and 21q (Nergadze et al., 2009).

Telomere-associated repeat (TAR1) sequences, composed partially of degenerated telomeric repeats, were also identified in the 2kb near to the TTAGGG repeats of nearly all sequenced chromosome ends. This sequence varies considerably in size among the different subtelomeres (Riethman, 2008). TAR1 sequences often contain the 61-29-37 bp repeats and can be partially transcribed as the subtelomeric part of TERRA.

Also embedded in the subtelomeric sequences are long coding and non-coding RNAs from the *WASH* and *DDX11* families. *DDX11* family genes are present in the 1p, 2q, 3q, 6p, 9p, 9q, 11p, 12p, 15q, 16p, 17p, 19p, 20p, 20q, Xq/Yq, Xp. These genes are transcribed towards the centromere some of them are potentially translated (Costa et al., 2009).

Genes from a subclass of the Wiscott-Aldrich Syndrome Protein (WASP) family, the *WASH* genes, are also present in the subtelomeres. These genes are transcribed towards the chromosome end and finish roughly 5 kb away from the telomeric repeats (Riethman, 2008). *WASH* genes have undergone extensive duplication and degeneration, and can be found in at least 16 chromosome ends. In humans, they encode a class of WASP proteins that co-localizes with and promotes nucleation of actin (Linardopoulou et al., 2007). Montero and colleagues proposed that the *WASH* transcripts proceed until the end of the chromosome and contain telomeric repeats, thereby being wrongfully detected as TERRA (Montero et al., 2016). They suggest that the only real TERRA, which is not derived from other subtelomeric transcripts, is the 20q TERRA. We show here that *WASH* transcripts do not behave as TERRA molecules in HCT116 DKO cells and can therefore not represent the majority of TERRA.



**Figure 4: Representation of the main repetitive elements present at the subtelomeres.** *WASH* and *TERRA* RNAs are transcribed in the same orientation, towards the chromosome end. *DDX11* transcripts are generated in the opposite direction. The elements in this figure are not in scale.

Nergadze and colleagues were the first to propose that subtelomeric CpG islands could act as *TERRA* promoters. They cloned the 61-29-37 bp repeats upstream an eGFP gene and observed that the 29-37 repeats were sufficient to drive eGFP expression. These repeats are easily recognized in the XqYq subtelomeres, while many variants can be found in other chromosome ends. The hypomethylation of the 61-29-37 repeats was linked to high *TERRA* expression, as DKO cells seemed to have less than 5% of methylation left in these specific sequences (Nergadze et al., 2009).



Although the 61 bp repeat is not required for the promoter activity of the subtelomeric repeats, it plays a role in the regulation of transcription. These repeats are also often the binding site for CTCF and Rad21, whose depletion leads to decreased TERRA expression (Deng et al., 2012). Apart from CTCF and Rad21, NRF1 and HSF1 have been implicated in activating TERRA transcription (Diman et al., 2016; Koskas et al., 2017). On the other hand, ATRX and Snail1 seem to inhibit TERRA transcription (Chu et al., 2017; Mazzolini et al., 2018).

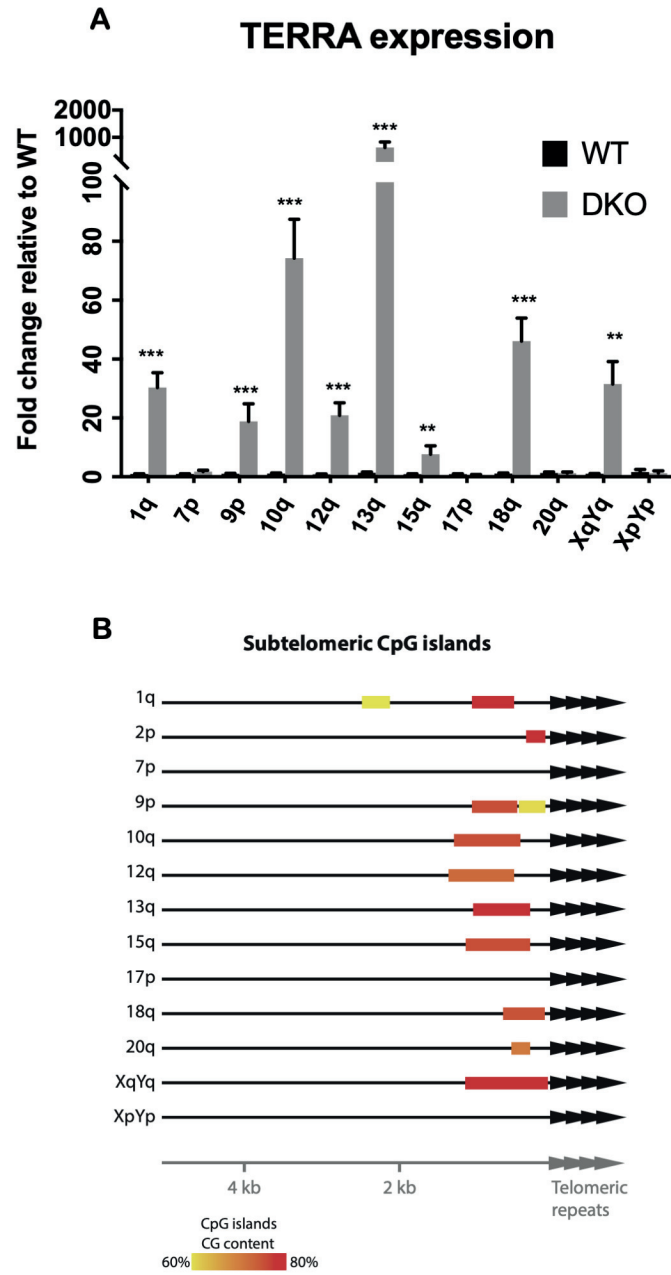
Bioinformatic analysis of subtelomeric sequences predicts a large number of binding sites for several transcription factors (Porro et al., 2014). Complemented with an additional analysis, we validated some of the predicted TERRA transcription factors and performed a detailed characterization of TERRA transcription in different cell lines.

## **2.2 Results**

### **Characterization of TERRA expression in HCT116 DKO cells**

Even though HCT116 DKO cells are known to overexpress TERRA, the extent of the increase was never measured in detail in the different chromosome ends. One of the reasons behind this is the challenge of measuring individual TERRA molecules by RT-qPCR, due to the sequence similarity between different subtelomeres. To circumvent this problem, we carefully designed and/or validated the TERRA qPCR primers used in our experiments by isolating the RT-qPCR products. These products were then subcloned into a plasmid and sequenced. Unless stated otherwise, all primers used in the experiments below were validated to amplify only the desired product.

With a set of specific TERRA primers on our hands, we measured the increase in TERRA expression in DKO cells. When compared to WT cells, most TERRAs are highly overexpressed in DKO (Figure 5A). However, the 7p, 17p, 20q and XpYp TERRA species seem to be unaffected by the lack of subtelomeric methylation.



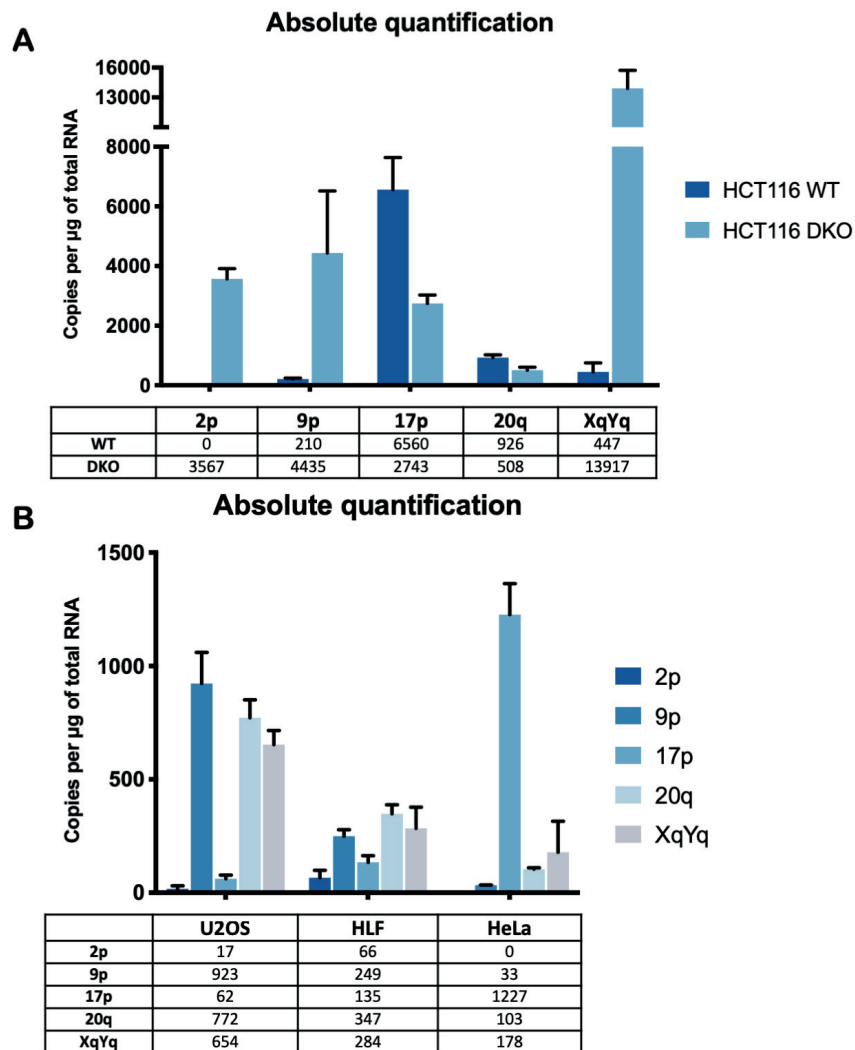
**Figure 5: Characterization of TERRA expression in HCT116 DKO cells. (A)** Relative quantification of TERRA expression in HCT116 WT and DKO cells. WT TERRA levels were adjusted to 1. **(B)** Representation of position and CG content of subtelomeric CpG islands in the indicated chromosome ends. \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ ,  $n = 3$ .

To elucidate the mechanism behind this phenotype, we looked at the composition of the first 15 kb of subtelomeric sequences. Our goal was to identify CpG islands potentially involved in promoting TERRA transcription. Any sequence longer than 200 bp with more than 60% GC content was considered a CpG island. As seen in Figure 5B, about 75% of the chromosome ends analyzed contain a CpG island, similarly to previously published data (Koskas et al., 2017). The length and the GC content of these CpG islands greatly varied among the different subtelomeres. Furthermore, the chromosome ends displaying no CpG islands are exactly those whose TERRA expression is insensitive to lack of methylation. The 20q TERRA, however, seems to be an exception. Its subtelomere shows a very short CpG island, very similar in length to the 2p, whose expression is greatly increased in DKO cells (Figure 6A). Only the absolute quantification of 2p TERRA is shown since the extremely low expression of this species in WT cells impedes a relative quantification. The main difference between the 2p and the 20q CpG islands seems to be their GC content. While the 2p island has 81% CG content, the 20q barely reaches 70% (Figure 5B). These data suggest that the presence of a subtelomeric CpG island predicts TERRA overexpression in DKO cells. For very short CpG islands, this prediction seems to also depend on their CG content.

Subtelomeric CpG islands are thought to be important promoters of TERRA transcription. We therefore tested if those subtelomeres lacking CpG islands express any TERRA at all and how their transcription levels compare to CpG-containing ones. To do so, Marianna Feretzaki developed an absolute quantification method for TERRA. We subcloned the subtelomeres of interest into pcDNA4 vectors. The final amount of each plasmid was carefully measured using the Qubit™ dsDNA High Sensitivity Assay Kit and a serial dilution from 1 to  $10^8$  copies of plasmid was used to create a standard qPCR curve. The standard curve was used to quantify absolute TERRA levels in five cell lines. For this analysis, we included not only HCT116 WT and DKO cells, but also another telomerase-positive cell line (HeLa), a primary line (human lung fibroblasts, HLF) and cells that maintain their telomeres by ALT (U2OS). This allows us to gain an overview of TERRA expression pattern across different cellular

backgrounds. In our calculations, we assume that the reverse transcription reaction is 100% efficient, which is probably not the case. Therefore, the results of the plasmid-based quantification are expressed as copies per  $\mu\text{g}$  of total RNA.

The results obtained by absolute quantification corroborate those stemming from the relative quantification (Figure 6A). It also became clear that subtelomeres without CpG islands are still able to produce TERRA (17p), and their transcripts appear to be abundant in HeLa and HCT116 WT cells.



**Figure 6: Absolute quantification of TERRA in different cell lines.** Absolute quantification was performed by qPCR using a standard curve prepared with plasmids containing subtelomeric sequences of the indicated telomeres. The quantification was done in HCT116 WT and DKO cells (A), in U2OS, HLF passage 14 and HeLa (B).

HeLa and HCT116 cells, both telomerase-positive, showed a very similar pattern of TERRA expression (Figure 6B). 17p was by far the highest expressed TERRA, followed by 20q or XqYq. In contrast, U2OS cells showed high expression of XqYq and 20q TERRAs, while relatively few copies of 17p were detected. Primary HLF cells present a more equilibrated profile, where the differences between the highest and lowest expressed TERRAs are not so pronounced. In all DNMT1/3B proficient cell lines, the 2p TERRA was the one with the lowest absolute expression in our panel.

To achieve an even more realistic quantification of TERRA, we quantified a few molecules using standard curves done with TERRA RNA. For that, cloned TERRA was *in vitro* transcribed, quantified, serially diluted and reverse transcribed before the qPCR. By doing so, we can take the limitation of the reverse transcription into account when calculating the number of TERRA molecules per cell (Table 1).

Table 1: Number of TERRA molecules per cell

Cell Line	10q	15q	20q
<b>U2OS</b>	164 ± 41	63 ± 9	4 ± 1
<b>HLF</b>	0.8 ± 0.1	8 ± 1	1 ± 1
<b>HeLa</b>	2 ± 1	5 ± 1	1 ± 1
<b>HCT116 WT</b>	2 ± 1	22 ± 4	8 ± 2
<b>HCT116 DKO</b>	954 ± 65	298 ± 12	4 ± 1

Due to different preparation of standard curves, the results from Table 1 and Figure 6 cannot be directly compared, but the relative amount of TERRA expression among the cell lines remains reasonably constant. In both experiments, 20q expression is about twice as much in HCT116 WT cells as in DKO cells. This specific TERRA molecule was previously suggested to be the most abundant TERRA in U2OS cells. A knockout of the subtelomeric 20q sequence seemed to produce a TERRA KO cell line (Montero et al., 2016). Here we use two methods to show that the 20q is not the most abundant and certainly not the only TERRA molecule produced by different cell lines.

## TERRA and WASH transcripts behave differently in DKO cells

In 2016, Montero and colleagues claimed that there is only one *bona fide* TERRA molecule and that all other TERRA molecules detected by Northern blot and RT-qPCR are in fact *WASH*, *DDX11* and *TAR1* transcripts (Montero et al., 2016). It is expected that some *TAR1* sequences are present in TERRA molecules, since they often comprise the 61-29-37 bp repeats and/or are adjacent to the TTAGGG repeats. The *DDX11* transcripts are unlikely to be misclassified as TERRA because they are transcribed in the opposite direction. However, if the transcription of *WASH* is not properly terminated, it could indeed advance towards the telomeric repeats. Assuming that all TERRAs, except the 20q, are in reality *WASH* transcripts, then these RNAs must simulate the behaviour previously associated to TERRA, including their upregulation in HCT116 DKO cells (Nergadze et al., 2009).

To verify if *WASH* transcripts are upregulated in DKO cells, we designed 3 primer pairs to amplify *WASH* by RT-qPCR. To avoid any bias, random primers were used in the reverse transcription reaction. In agreement with our previous experiments (Figure 5) more TERRA transcripts were detected in HCT116 DKO cells, with exception of the 20q TERRA. The expression of *WASH*, as measured by any primer pair, was unchanged in this cell line (Table 2). Therefore, the changes in TERRA expression seen by us and others in RT-qPCR or Northern blots cannot be explained by changes in *WASH* expression, arguing that these RNAs are unrelated and possess their own promoters.

Table 2. Fold change in the expression of the listed transcripts. Expression in HCT116 WT was adjusted to 1. N = 1.

Transcript	HCT116 WT	HCT116 DKO
9p TERRA	1	9.6
12q TERRA	1	31.3
18p TERRA	1	27.7
20q TERRA	1	0.52
WASH_1	1	1.1
WASH_2	1	1.0
WASH_3	1	0.9

## ZNF148 is a repressor of TERRA transcription

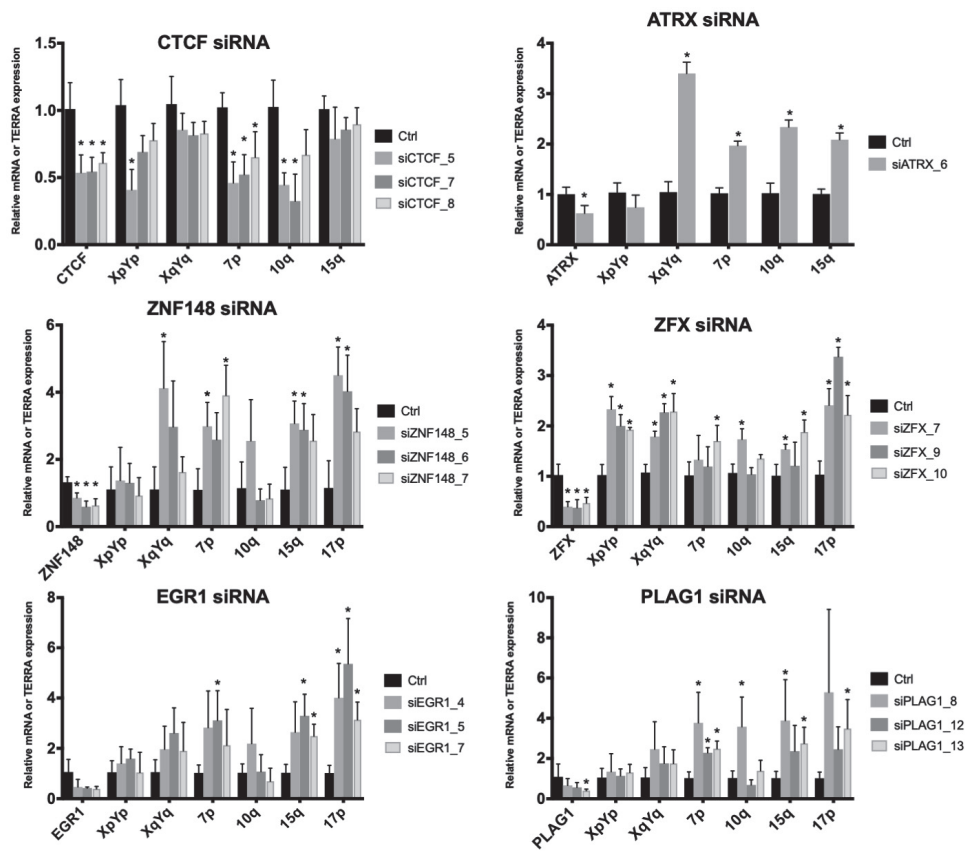
Currently, there are only a few transcription factors known to influence TERRA transcription, including CTCF, NRF1 and HSF1. However, TERRA expression is complex, responds to changes in telomeric DNA and protein composition and varies according to cell line and chromosome end, suggesting that many transcription factors may be potentially involved in its transcriptional regulation. In addition, previous analysis of the 2 kb surrounding predicted TERRA TSSs (Porro et al., 2014) revealed that subtelomeric sequences are rich in binding motifs for the most diverse transcription factors. Therefore, we set out to study which of these factors actually contribute to TERRA expression.

Firstly, in a collaboration with Marianna Feretzaki and Julien Delafontaine, we performed a motif search analysis (FIMO database) to complement what was previously published (Porro et al., 2014). The main difference between the two approaches is that, while Porro and colleagues screened the 2 kb around TERRA's predicted 5' end, we scanned the 2 kb sequence preceding the start of the telomeric repeats (Table 3). Marianna performed an siRNA screening with all transcription factors whose motifs were discovered in the bioinformatic analysis. CTCF and ATRX were used as positive controls. We were unable to efficiently knockdown SP1, EWSR1, SMAD4, PEBP1, MAZ, SP2 and ZNF263. AP2A1, KLF15 and ZNF281 depletions did not change TERRA expression (data not shown). However, ZNF148, ZFX, EGR1 and PLAG1 knockdowns induced TERRA expression (Figure 7), suggesting that they act as TERRA repressors.

Table 3: Motifs of transcriptional regulatory elements found in the different subtelomeres by FIMO.

TF	CpG-positive chromosome ends						CpG-negative chromosome ends				Sum
	9p	10q	15q	16p	19p	XqYq	7p	17p	20q	XpYp	
<b>CTCF</b>	25	8	19	16	24	29	15	8	18	8	170
<b>PLAG1</b>	11	8	16	12	8	13	18	1	19	60	166
<b>EGR1</b>	8	8	4	7	9	7	6	8	4	2	63
<b>SP1</b>	110	49	98	119	136	165	86	58	89	289	1199
<b>ZFX</b>	79	56	69	56	59	68	56	8	36	56	543
<b>ZNF148</b>	32	45	19	26	35	24	26	10	44	48	309
<b>EWSR1</b>	5	4	3	8	6	8	2	4	6	2	48
<b>AP2A1</b>	8	8	7	3	1	7	6	8	2	4	54

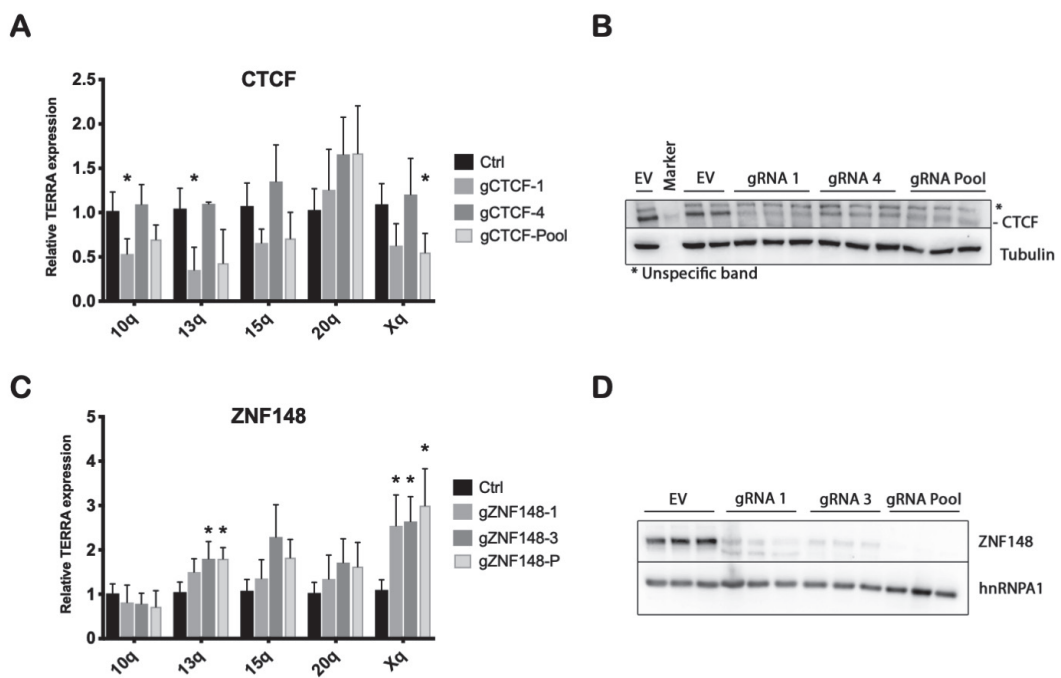
<b>SMAD4</b>	3	7	8	9	6	2	2	4	5	10	56
<b>CACYBP</b>	28	29	12	8	14	12	19	16	6	9	153
<b>KLF15</b>	8	4	3	14	5	15	5	6	22	22	104
<b>PEBP1</b>	4	6	3	1	2	7	6	6	1	9	45
<b>MAZ</b>	32	25	29	35	29	19	28	19	22	90	328
<b>SP2</b>	66	98	48	56	49	69	59	35	36	189	705
<b>ZNF263</b>	35	25	19	20	18	33	19	36	7	36	248
<b>ZNF281</b>	145	89	116	143	151	143	189	36	80	357	1449



**Figure 7: ZNF148, ZFX, EGR1 and PLAG1 are new factors affecting TERRA transcription.** HeLa cells were transfected with single siRNAs against transcription factors predicted to bind to subtelomeric sequences. \* $p < 0.05$ ,  $n = 3$ .



To confirm some of these results, we designed gRNAs to knockout ZNF148 in HeLa cells using CRISPR/Cas9. gRNAs against CTCF were used as positive control and caused a decrease in TERRA levels (Figure 8A-B). The gRNA 4 against CTCF had little effect in the protein level and therefore did not affect TERRA expression. ZNF148 KO with gRNA 1, 3 or with a pool of 5 gRNAs caused significant increase in 13q and Xq TERRAs (Figure 8C-D). Both siRNA and gRNA approaches suggest that ZNF148 acts as a negative regulator of TERRA expression.



**Figure 8: ZNF148 knockout increases TERRA transcription.** HeLa cells were transfected with 1, 2 or a pool of 5 gRNAs against CTCF and ZNF148. Transfected cells were selected for 3 days with puromycin. RNA from 1-3 million cells was used for RT-qPCR (A and C) and 100,000 cells were used for verification of protein levels (B and D). \* $p < 0.05$ ,  $n = 3$ .

## 2.3 Discussion and Perspectives

The increased TERRA expression phenotype in DKO cells was known since 2009 (Nergadze et al., 2009). However, the extent of this increase was never fully measured. Here we quantify TERRA overexpression in 12 chromosome ends. We found that subtelomeres lacking CpG islands do not overexpress TERRA in the DKO condition. It seems that the presence of a large, high density subtelomeric CpG island predicts that a TERRA transcript will be upregulated in DKO cells.

Without the “classical” TERRA promoters, do telomeres from 7p, 17p and XpYp express TERRA at all? By using relative and absolute TERRA quantification methods, we show that they do. But how do the levels of these transcripts compare to CpG-driven TERRAs? And what drives their transcription?

Estimating the number of TERRA molecules (or that of any RNA) inside cells has always been challenging. FISH staining shows several foci, which could be single TERRA molecules or an aggregation of the lncRNA. If the former is correct, then the method is biased towards the identification of the longest TERRAs, while short ones are lost in the background. Moreover, FISH protocols require a pre-extraction step, in which TERRA molecules that are not directly or indirectly attached to the chromatin are lost. Northern blots are similarly biased towards the detection of longer TERRAs, and the diversity of this lncRNA’s length makes it challenging to calculate a copy number.

The current method of choice to calculate the abundance of any RNA is RT-qPCR. However, we were so far only able to do this calculation in a comparative manner: relative to another RNA (e. g. GAPDH) and/or relative to another cell line or treatment. The reasons for that include different primer efficiencies, software-dependent calculation of cycle threshold values and qPCR method. To circumvent these limitations, we calculated the number of TERRA molecules using a method similar to the one used to calculate the number of hTR molecules (Xi and Cech, 2014). By preparing a standard qPCR curve with plasmids containing 2p, 9p, 17p, 20q or XqYq subtelomeres, we are able to eliminate most of the variables that prevent a reliable absolute quantification. The only variable

still present is the reverse transcription reaction. To account for that, we reverse transcribed 10q, 15q and 20q TERRAs containing approximately 540 bp of telomeric repeats in the same way we would treat our samples. The qPCR curve for absolute quantification was then prepared with the reverse-transcribed TERRA cDNA.

Another possible weakness of both methods is the RNA isolation method itself. We cannot guarantee that all cells were properly lysed and that no TERRA RNA was lost during the purification process, suggesting that the number in Table 1 are underestimated.

Nonetheless, both approaches clearly state that the 20q telomere is not the only one transcribing TERRA, contradicting recent reports (Montero et al., 2016, 2018). In fact, that is not even the most abundant TERRA species in U2OS cells. Surprisingly, in both telomerase-positive cell lines (HCT116 and HeLa), the 17p TERRA is the most expressed one in our plasmid-based absolute quantification. Its subtelomere is devoid of CpG islands, which were thought to be strong TERRA promoters. Therefore, in the 17p telomere, CpG islands are not a requirement for TERRA transcription, suggesting that CpG-negative subtelomeres can potentially express high levels of TERRA.

Montero and colleagues also argue that most non-20q TERRA molecules are actually *WASH* transcripts (Montero et al., 2016). We showed by RT-qPCR that these transcripts behave differently than TERRA. Unlike TERRA, *WASH* RNA is not increased in DKO cells, suggesting that they do not share the same promoter and are likely unrelated RNAs. Interestingly, the 20q TERRA is also not increased in DKO cells, as determined by relative and absolute quantifications. This TERRA can therefore not justify the increased TERRA levels in DKO cells, which is used as a positive control in the same publication. For this increase to be visible in Northern blots and RNA dot blots, many other non-20q TERRA molecules must be overexpressed. We confirmed here that the 1q, 2p, 9p, 10q, 12q, 13q, 15q, 18q and XqYq TERRAs behave as expected in DKO cells (Figure 5 and Figure 6). Furthermore, the same publication showed that *WASH* and TERRA transcripts have strikingly different patterns in Northern blots (Montero et al., 2016).

We next asked which transcription factors drive TERRA transcription in CpG-positive and negative telomeres. To answer that, the 2 kb preceding the start of TTAGGG repeats were scanned for the presence of binding motifs for any transcription factor. Both the current and a previous analysis (Porro et al., 2014) failed to identify transcription factors binding exclusively to CpG-negative chromosome ends. However, we uncovered several candidates to new TERRA transcription factors. siRNA depletion of ZFX, EGR1, PLAG1 and ZNF148 suggests that they have a repressor effect on TERRA transcription.

ZFX contains 13 zinc finger domains, as well as transcriptional activation and DNA binding domains. Its transcriptional activity is required for the self-renewal of mouse and human hematopoietic stem cells (Galan-Caridad et al., 2007; Harel et al., 2012). However, ZFX has not yet been linked to telomere biology. Early growth response 1 (EGR1) is responsible for the activation of genes involved in cell growth and angiogenesis, and its expression has been shown to increase upon hTERT overexpression (Park et al., 2016). In orthotopic xenograft models, depletion of EGR1 has been associated with diminished tumor invasion capabilities (Park et al., 2016). Conversely, other studies suggest that this protein acts as a tumor suppressor (Krones-Herzig et al., 2005; Mohamad et al., 2018). Although the hTERT promoter contains a EGR1 consensus motif, there is conflicting evidence about the effects of EGR1 in hTERT expression (Akutagawa et al., 2008; Jacob et al., 2016). The *PLAG1* zinc finger gene is consistently rearranged in salivary gland pleomorphic adenoma, and its upregulation has been shown to promote the development of human hepatoblastoma (Aström et al., 1999; Zatkova et al., 2004). Similarly to ZFX, a role for PLAG1 in telomere biology has not yet been demonstrated.

ZNF148 involvement in TERRA transcription regulation was further confirmed by CRISPR/Cas9-mediated deletion. This transcription factor is part of the Kruppel family of zinc finger DNA binding proteins and can function as an activator or repressor (Passantino et al., 1998). ZNF148, also known as ZBP-89, is often upregulated in tumors, where it can inhibit cell growth and induce apoptosis (Zhang et al., 2010). Decreased ZNF148 expression is also correlated with poor prognosis in colorectal cancer (Gao et al., 2013). In contrast, ZNF148

knockdown decreases *TERT* expression and telomerase activity, which is accompanied by telomere shortening, as measured by qPCR (Fang et al., 2017). No involvement of ZNF148 in TERRA expression was previously reported. Repression of TERRA expression by ZNF148 could potentially help to hamper the maintenance of telomere length in tumors by ALT.

Since ZFX, EGR1, PLAG1 and ZNF148 are transcription factors which influence the transcription of numerous genes, we cannot currently confirm that they act directly in TERRA promoters to repress its expression. However, the high number of consensus binding sites for these transcription factors in the subtelomeric sequences suggests that this effect could be direct. CHIP-qPCR for the endogenous protein may help us to confirm this hypothesis.

## 2.4 Project contributions

Marianna Feretzaki, Patricia Renck Nunes and Joachim Lingner designed the experiments.

MF and PRN designed and validated primers for TERRA quantification by RT-qPCR.

MF designed, performed and analyzed the siRNA screen (Figure 7); cloned plasmids for *in vitro* TERRA transcription; optimized and performed the *in vitro* transcription of TERRA; and implemented and executed the RT-qPCR for absolute quantification of TERRA (Figure 6).

PRN performed the analysis of subtelomeric sequences, identification of CpG islands and relative quantification of TERRA (Figure 5); cell culture and RNA isolation of all cell lines used for absolute quantification of TERRA; validation of ZNF148 antibody; design and cloning of ZNF148 and CTCF gRNAs; preparation of figures.

Julien Delafontaine: performed the FIMO (Find Individual Motif Occurrences) analysis.



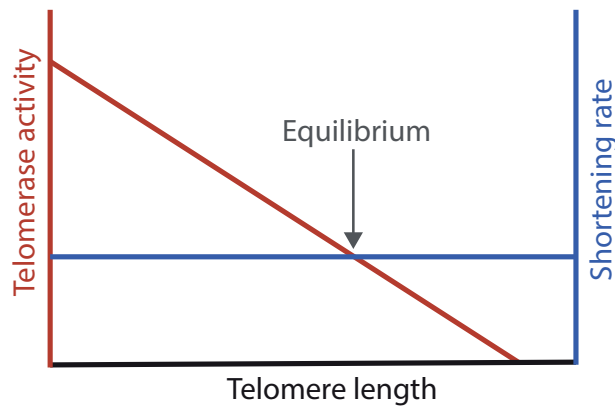
# Chapter 3. Role of TERRA in telomere length and chromatin structure

## 3.1 Introduction

### 3.1.1 Telomere length dynamics

Telomere length is determined by the equilibrium between telomere shortening and elongation. Telomere shortening is dependent on the end replication problem and on the resection caused during the formation of the 3' overhang (Lingner et al., 1995; Watson, 1972). The telomere shortening rate is approximately 100-200 bp per population doubling (PD) in human cells and varies slightly with cell strain and length of the 3' overhang (Huffman et al., 2000).

Experiments performed in *Saccharomyces cerevisiae* show that, while the shortening rate is constant and independent of telomere length, telomere elongation rate depends on telomere length (Marcand et al., 1999). The telomere length reaches its equilibrium when the telomere shortening rate equals the elongation rate (Figure 9). Marcand and colleagues showed that the telomere elongation rate is very high when telomeres are shorter than their equilibrium, and diminishes when telomere length exceeds this equilibrium (Marcand et al., 1999). By having different telomere elongation rates, different cells achieve different telomere lengths in the equilibrium state (Colgin et al., 2003). It is important to notice that this was not yet clearly demonstrated in mammalian cells.



**Figure 9: Model of telomere length regulation.** The telomere shortening rate is independent of telomere length. Telomerase activity varies according to the cell and increases proportionally with telomere shortening. Likewise, telomerase activity reduces upon telomere elongation. When the telomere shortening and elongation rates are equal, the cell reaches the equilibrium, with stable telomere length and telomerase activity.

In telomerase-positive cells, the telomere elongation rate is dictated by telomerase. Telomerase activity and processivity depend on a number of factors, including, but not limited to, hTERT and hTR expression, mutations in telomerase components, accessibility to the 3' overhang and redox state of the telomere (Ahmed and Lingner, 2018; Loayza and de Lange, 2003; Prescott and Blackburn, 1997).

Telomere shortening rate is best assessed upon telomerase knockout. Inhibitors of telomerase, such as BIBR1532 or GRN163L, can also be used. Calculating the telomere elongation rate is slightly more challenging. Overexpression of POT1 $\Delta$ OB is perhaps the preferable approach, since it allows telomerase to access the 3' overhang without interfering with the enzyme itself. However, this measurement is dependent on the levels of POT1 $\Delta$ OB expression, because the endogenous POT1 is still available. When comparing the elongation rate among different cell lines, it is imperative that they all have the same POT1 $\Delta$ OB protein levels.

As aforementioned, HCT116 DKO are not a true *DNMT1* KO. These cells translate a hypomorphic *DNMT1* allele into a truncated protein, which contain an intact catalytic domain and maintains up to 50% of the WT methylation levels (Egger et



al., 2006). However, it is a fact that they have high TERRA expression and short telomeres. Therefore, we took this opportunity to gain some insight on why their telomeres are so short. Here, we employ BIBR1532 treatment and POT1 $\Delta$ OB expression to try to understand the mechanism responsible for the short telomeres in DKO cells.

### **3.1.2 CRISPR interference and gene activation**

TERRA expression is complex, as it arises from several or all chromosome ends, and has no defined transcription start and end sites. Thus, manipulating TERRA expression has been a major challenge since the discovery of this lncRNA. TERRA knockdown has been initially achieved with siRNAs (Deng et al., 2009). This caused several defects in telomere structure and changes in telomeric chromatin. These effects could have been caused by TERRA depletion, but also by the presence of siRNAs, that could potentially disrupt the canonical maintenance of the telomere. The introduction of high levels of double stranded TERRA-like sequences could compete with endogenous TERRA and telomere binding proteins, leading to telomere uncapping and aberrations, and making it difficult to interpret the results. More recently, a similar phenotype was seen when TERRA was depleted by LNA gapmers (Chu et al., 2017). However, both LNA gapmers and siRNAs could potentially bind to the telomeric 3' overhang and interfere with normal telomere biology.

In 2016, Montero and colleagues tried to create a TERRA KO U2OS cell line by using CRISPR/Cas9 to delete 8 kb of the 20q subtelomere. One clone had a 4-fold reduction in TERRA levels, while the other two clones showed only a 50% reduction in TERRA expression. However, the Northern blot and qPCR analysis were highly inconsistent, implying that the KO method was probably affecting more than one subtelomere. More recently, the same approach was used to create another five "TERRA KO" clones, which again retained 50 to 80% of TERRA expression (Montero et al., 2018). The results of both papers need to be interpreted cautiously, not only because a 5-fold variation in TERRA levels is

common among WT clones, but also because we have shown that the 20q is not the most abundant TERRA in U2OS cells (Figure 6B). Additionally, the extend of the phenotypes observed in these clones do not correlate with their TERRA reduction. Due to the high similarity between subtelomeres, the gRNAs used for the creation of these clones likely targeted multiple chromosome ends. Induction of DSBs at subtelomeres could also cause sudden telomere loss and induce DDR leading to apoptosis or senescence. One evidence that this may have happened is the fact that the authors were unable to recover viable clones from HCT116 and HeLa cells, and only very few U2OS clones survived. It is also unknown if these survivors did not acquire additional mutations that enabled them to cope with the loss of subtelomeric sequences. This deletion could also have further consequences in the telomeric structure, destroy other subtelomeric transcripts and eliminate the binding sites for proteins involved in telomere architecture and protection. Therefore, it remains unclear if the observed phenotypes, which include severe telomeric anomalies, are indeed due to repressed 20q TERRA expression.

TERRA upregulation is perhaps even more challenging. The addition of high copy TTAGGG-containing vectors for TERRA overexpression meet the same competition problems with the endogenous telomeric sequence as the siRNAs. So far, the most elegant approach to increase TERRA expression are tiTELS, where a modest 10 to 20-fold increase in TERRA is observed (Farnung et al., 2012).

In this chapter, I describe how we used gRNA-guided gene activation to overexpress endogenous TERRA in HCT116 and HeLa cells. The Cas9 is a nuclease part of the clustered regularly interspaced short palindromic repeats (CRISPR) system, which functions as an adaptative immune system in *Streptococcus pyogenes* and other microbes. This nuclease is guided by an RNA (guiding RNA or gRNA) to the target locus and cleaves the dsDNA (Jinek et al., 2012). Repair of this double stranded break by NHEJ can be used to create knockouts, while HR in the presence of an exogenous repair template creates knockins and targeted gene modifications (Ran et al., 2013). Due to its ease of use, CRISPR/Cas9 rapidly became the method of choice for generation of knockouts and knockins.

To cleave both DNA strands, the Cas9 has two endonuclease domains, RuvC and HNH. Mutations in important residues of these domains (D10A and H840A) generates a catalytically inactive enzyme (dead/deactivated Cas9 or dCas9) (Qi et al., 2013). The dCas9 can still bind to the DNA and be used as a vehicle to bring virtually any other fused protein to a gRNA-targeted locus.

The first derivatives of the dCas9 were transcription manipulation systems. The dCas9 itself and dCas9-KRAB (Krüppel associated box) proteins can be used to repress gene expression (CRISPR interference or CRISPRi). The KRAB fusion protein does not consistently induce more repression than the dCas9 alone, suggesting that the mechanical effect (prevention of RNA polymerase binding, for example) plays an important role. Therefore, the trick to an efficient repression seems to be the position of the gRNA, and this varies vastly according to the target (Gilbert et al., 2013; Qi et al., 2013). Other tools, such as dCas9-DNMT3A and dCas9-LSD1, are also available, but do not improve significantly the repression effect and are mostly used for epigenetic studies (Kearns et al., 2015; Vojta et al., 2016).

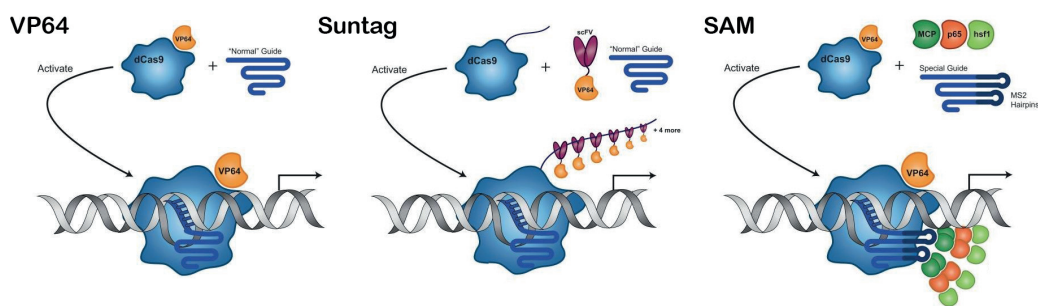
Rapidly following CRISPR interference was CRISPR activation. The dCas9 was fused to four copies of the transcription activator VP16 (VP64) and used to induce expression of exogenous constructs and endogenous human genes (Gilbert et al., 2013; Perez-Pinera et al., 2013). Just like for the CRISPRi, the position and combination of gRNAs is crucial for a high overexpression (Maeder et al., 2013; Perez-Pinera et al., 2013).

After the launch of dCas9-VP64, another 8 dCas9-based activation systems were published: dCas9-VP160, VP64-dCas9-BFP-VP64, dCas9-Tet1-CD, dCas9-p300, dCas9-Suntag, dCas9-VP64-p65-Rta, SAM and RNA scaffold systems based on MS2 and PP7 loops (reviewed in Chavez et al., 2016).

From all these options, the synergistic activation mediator (SAM) stands out for its efficiency and consistency. In this system, two MS2 loops are inserted in the gRNA sequence and can be bound by a MS2 coating protein fused to the HSF1 transactivation domain and the NF- $\kappa$ B transactivating subunit p65. In combination, gRNA-MS2-p65-HSF1 and dCas9-VP64 deliver five activation

domains to the target site and create the ideal conditions for gene expression (Koner mann et al., 2014).

Performing closely to SAM is the dCas9-Suntag. The dCas9 is fused to 10 copies of the GCN4 peptide, which is tightly bound by a single-chain variable fragment (scFv) fused to a VP64 (scFv-VP64) (Tanenbaum et al., 2014). The Suntag can theoretically deliver up to ten VP64 transcription activators and often performs better than dCas9-VP64 alone.



**Figure 10: dCas9 activation systems: VP64, Suntag and SAM.** See text for more details. Figure modified from Chavez et al., 2016.

Even with all these systems available, there is still no consensus about how to design the gRNAs for dCas9 targeting. The current approach is based on trial and error and the success of this method is very much target-based. While some genes will be easily 10,000 to 100,000-fold overexpressed, others will fiercely resist a 10 fold upregulation (Chavez et al., 2016).

Apart from gene repression and activation, dCas9 variants are being used widely due to their versatility. Single molecule live cell imaging, ChIP-sequencing, CRISPR-display and locus-specific proteomics are just some examples (Schmidtman et al., 2016; Shechner et al., 2015; Tanenbaum et al., 2014; Tsui et al., 2018).

In this chapter, we use the classical dCas9-VP64 to activate the expression of TERRA in HCT116 and HeLa cells and assess its impact on the subtelomere and

telomere structure. We also evaluate the telomere length dynamics in HCT116 WT and DKO cells to better understand the cause of the short telomere phenotype in the DKO.

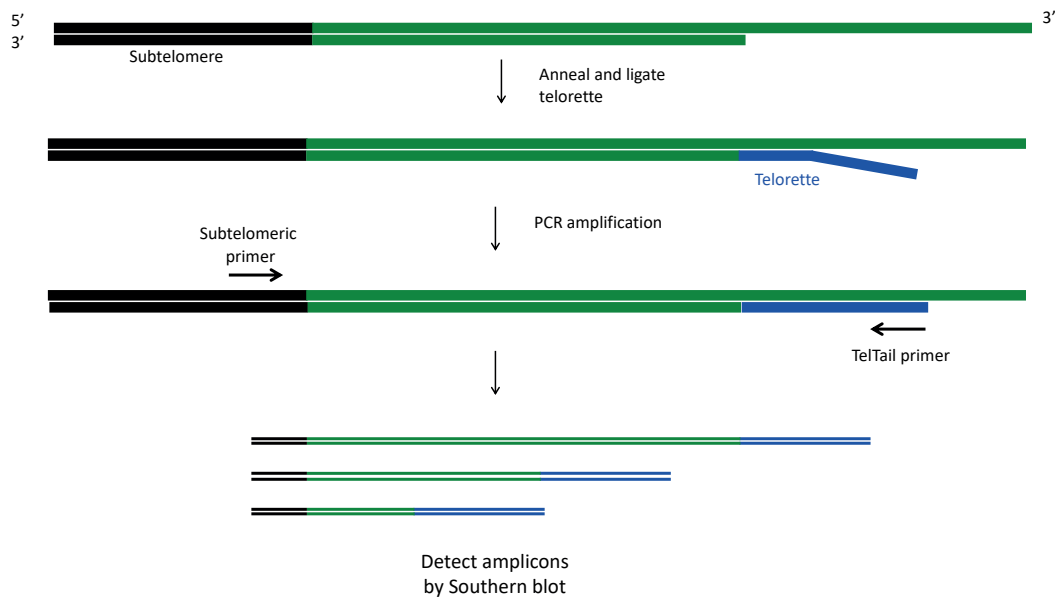
## 3.2 Results

### **TERRA expression and telomere length do not directly correlate in DKO cells**

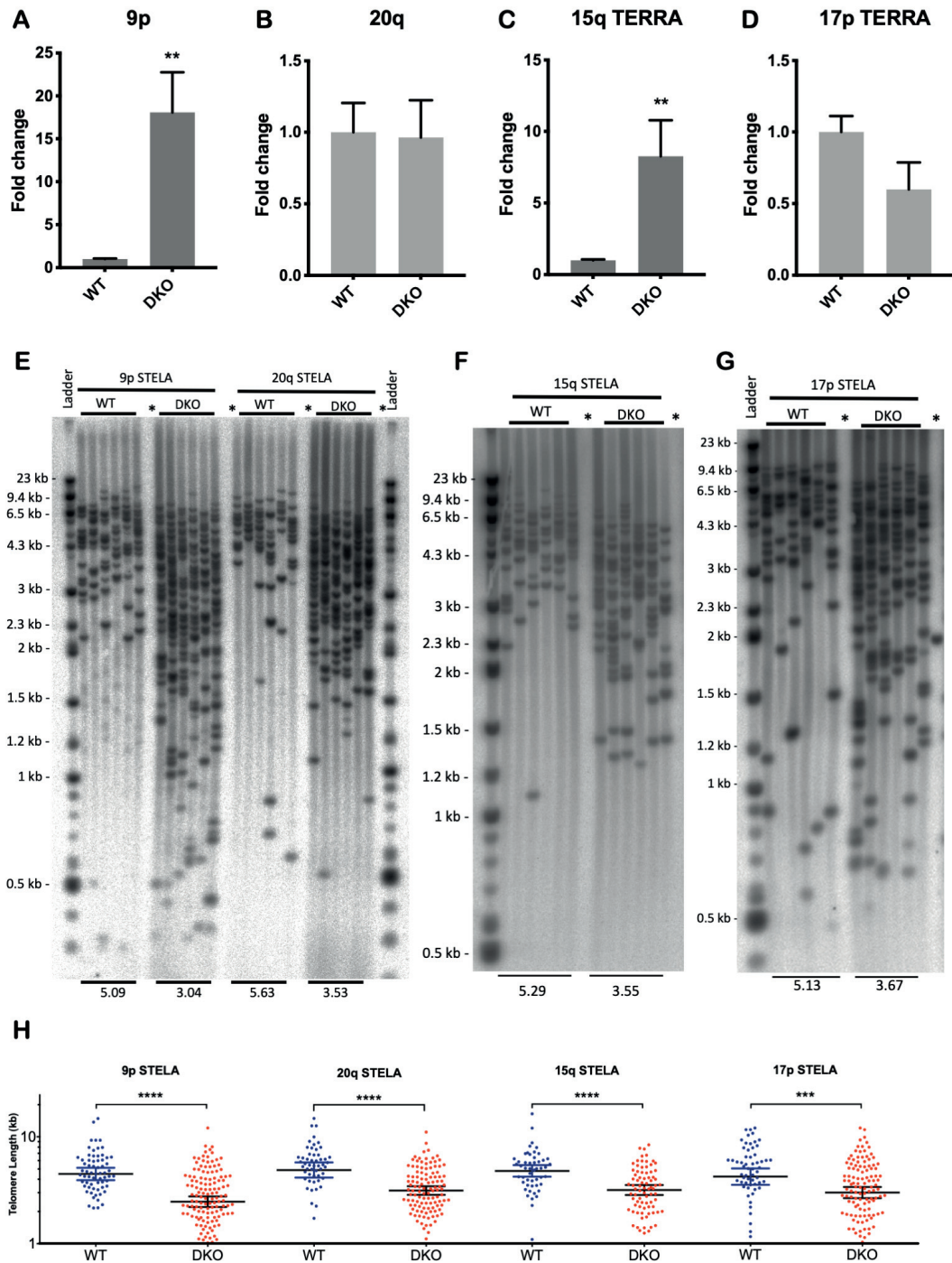
In *S. cerevisiae*, TERRA levels increase with telomere shortening (Cusanelli et al., 2013; Graf et al., 2017). In telomerase-positive cells, the upregulated TERRA may help telomere elongation by bringing telomerase to the shortest telomeres (Cusanelli et al., 2013). Another model suggests that short telomeres fail to degrade TERRA R-loops before the passage of the replication fork, leading to DDR, Rad51 loading and possibly to telomere elongation via HDR (Graf et al., 2017). In contrast, inducing TERRA expression in *S. cerevisiae*'s telomere 1L leads to Exo1-dependent telomere shortening *in cis* (Pfeiffer and Lingner, 2012). These observations could not yet be translated to human cells. Therefore, to understand if there would be a direct correlation between TERRA expression and telomere length in human cells, we studied both aspects in HCT116 DKO cells.

In chapter 2, we showed for the first time that some TERRA molecules are expressed at similar levels in WT and DKO cells (Figure 5). Since TERRA may regulate telomere length, we asked whether telomeres without TERRA overexpression also show the short telomere phenotype in the DKO. To measure the length of individual telomeres, we used the Single TELOmere Length Analysis (STELA). STELA was originally developed by Duncan Baird and consists of a long-range small pool PCR with forward primers specific to the telomere of interest (Baird et al., 2003). The genomic DNA is digested with EcoRI, followed by overnight ligation with telorette 3 adaptor. The small pool PCR is performed with a primer complementary to the adaptor (TelTail primer) and a subtelomeric

specific primer. We validated specific primers for 9p, 10q, 17p and 20q STELAs. The primer designed for 10q STELA may amplify 13q products in 12.5% of the cases, as measured by sequencing of short STELA products (data not shown). The primer for 17p STELA is able to bind the 7p subtelomere, but PCR amplification is not possible due the position of the EcoRI site in this subtelomere. The primer used for 15q STELA was previously described (Farnung et al., 2012).



**Figure 11: Schematic representation of STELA.** A primer complementary to the 3' overhang is ligated to the C-rich telomeric strand. The telorette contains a unique sequence in its 5' end, which is then recognized by a reverse primer in the PCR reaction (TelTail). The forward primer binds to the subtelomere sequence and allows for the amplification of a specific telomere. Amplified telomeres are detected by Southern blot with a telomeric probe.



**Figure 12: TERRA expression and telomere length do not correlate in HCT116 DKO cells.** 9p (A), 20q (B), 15q (C) and 17p (D) TERRA expression was measured by RT-qPCR in HCT116 WT and DKO cells. TERRA levels were normalized to GAPDH. N=3. STELA was performed to measure telomere length of 9p, 20q (E), 15q (F) and 17p (G) telomeres. Each lane corresponds to an independent small pool PCR reaction done with genomic DNA of the indicated cell line. Negative controls were done without telomere ligation (lanes marked with an asterisk). (H) Scattered plot (geometric mean  $\pm$  95% confidence interval) representing the same results as in E, F and G. Each band is represented by one dot. \*\* $p$ <0.01.

To verify if there is any correlation between TERRA expression and telomere length, we compared the telomere length of two telomeres overexpressing TERRA (9p and 15q) to the length of those with TERRA expression similar to the WT (20q and 17p). Independently of their level of TERRA expression (Figure 12A-D), 9p, 15q, 17p and 20q telomeres were significantly shorter in DKO cells compared to WT cells (Figure 12E-H). This suggests that, if TERRA is responsible for the short telomere phenotype, this effect is not performed exclusively *in cis*.

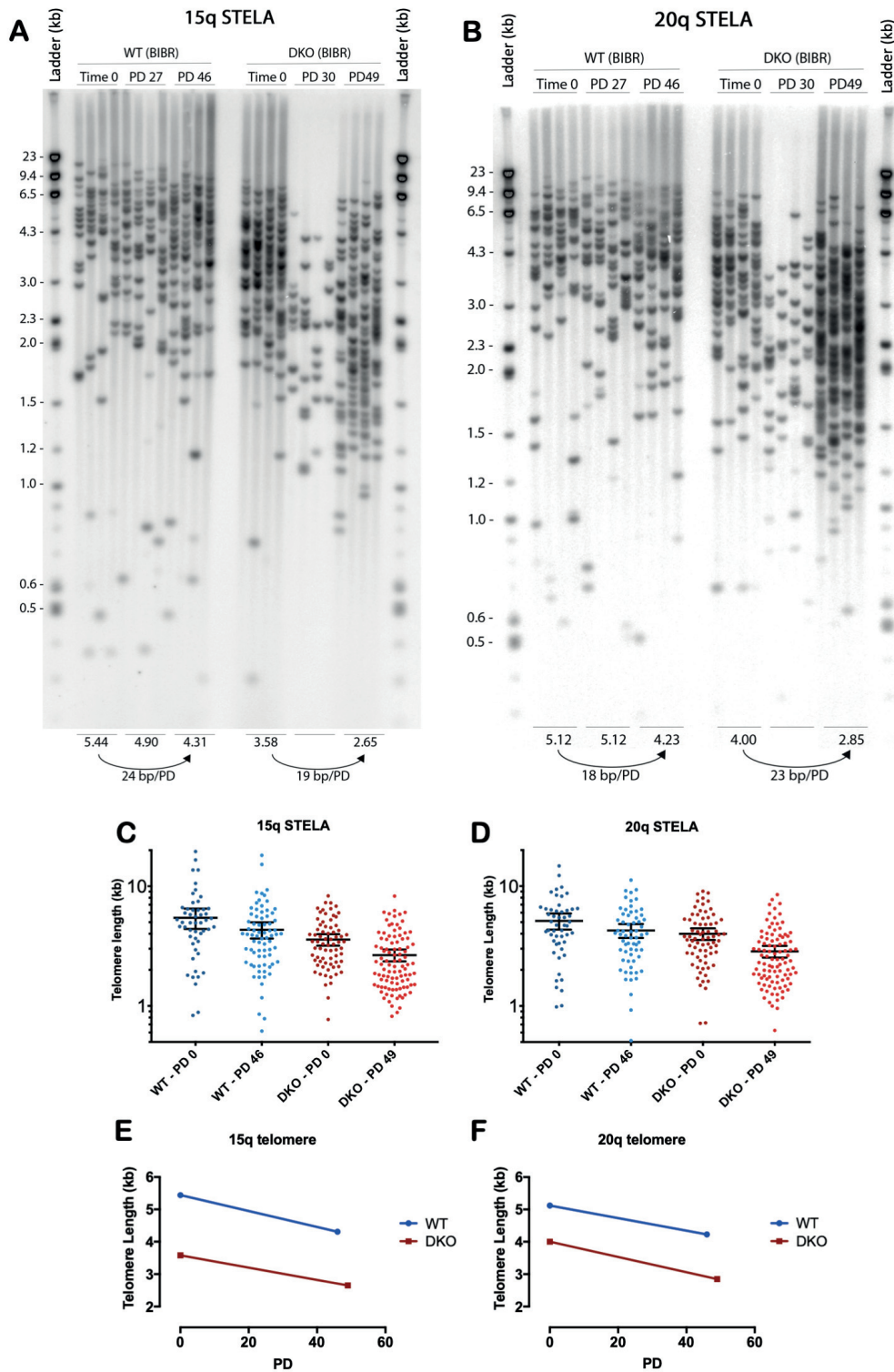
### **Telomere length dynamics are changed in DKO cells**

The telomere length is defined by the equilibrium between telomere elongation and telomere shortening. As HCT116 WT and DKO cells have different telomere lengths, they necessarily have distinctive telomere elongation or shortening rates.

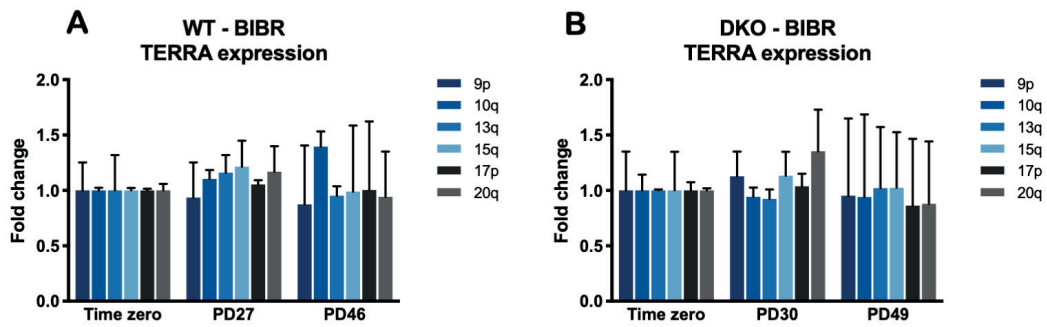
To test whether the telomere shortening is increased in DKO cells, we treated WT and DKO cells with 20  $\mu$ M of the non-competitive telomerase inhibitor BIBR1532 (Pascolo et al., 2002). Since the 15q TERRA is upregulated in DKO cells and 20q is not, we measured the length of these telomeres by STELA. After 46 or 49 population doublings in the presence of BIBR1532, WT cells showed a shortening rate of 24 and 18 bp per population doubling (bp/PD) for the 15q and 20q telomeres, respectively, while DKO cells shortened the same telomeres at rates of 19 and 23 bp/PD (Figure 13). TERRA expression was not changed by BIBR1532 treatment (Figure 14), as previously described (Farnung et al., 2012).

Due to the lack of difference in the shortening rates of the two cell lines, we tested if the telomere elongation rate was decreased in DKO cells. For that, we transduced both cell lines with retroviruses expressing a POT1 mutant lacking its OB fold (POT1 $\Delta$ OB) (Figure 15A). This mutant is no longer able to sequester the telomeric ssDNA either directly or through T-loop formation (Loayza and de Lange, 2003). The ssDNA then becomes more accessible to telomerase and we can measure the elongation rate by telomerase.





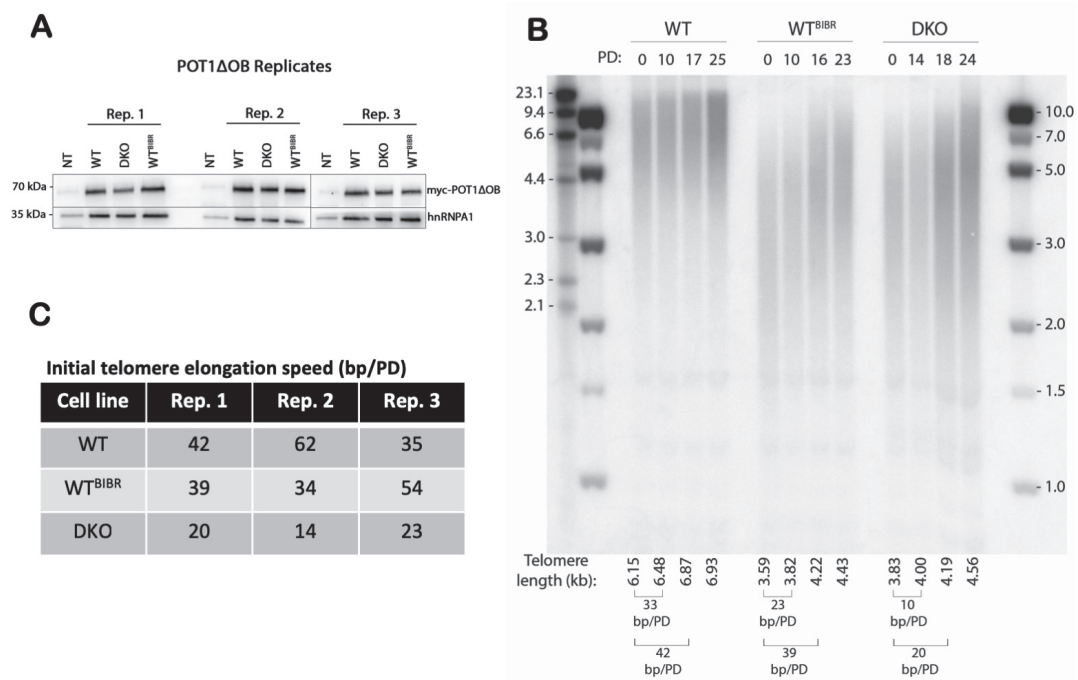
**Figure 13: WT and DKO cells show similar rates of telomere shortening.** 15q (A) and 20q (B) STELAs were performed with genomic DNA of WT or DKO cells grown for up to 49 population doublings in the presence of BIBR1532. Average telomere length from 4 independent PCR reactions are indicated. The average length of DKO PD30 is not calculated due to insufficient number of bands. The same data is plotted in (C) and (D) (geometric mean  $\pm$  95% confidence interval). 15q (E) and 20q (F) telomere shortening rates expressed as average telomere length after 46 and 49 PDs in the presence of BIBR1532.



**Figure 14: TERRA levels remain stable during BIBR1532 treatment.** HCT116 WT (A) and DKO cells (B) were treated for up to 49 population doublings with BIBR1532 or DMSO. Graphs show only BIBR1532-treated samples. The fold change was calculated relative to the TERRA expression in DMSO-treated samples. TERRA levels before the treatment (“Time zero”) were adjusted to 1. Two technical replicates are shown.

Since the difference in the initial telomere length between WT and DKO cells is substantial, we included WT cells that were treated with BIBR1532 for about 70 population doublings until reaching a telomere length similar to that of DKO cells (WT<sup>BIBR</sup>). In three independent experiments, DKO cells presented consistently lower telomere elongation rates (Figure 15).

Taken together, our results suggest that a decreased telomere elongation rate, rather than an increased shortening, is the reason behind the short telomeres phenotype in DKO cells.



**Figure 15: Telomere elongation rate is decreased in DKO cells.** HCT116 WT, DKO or WT treated for about 70 PD with BIBR1532 (WT<sup>BIBR</sup>) were transduced with retroviruses expressing POT1ΔOB. **(A)** Levels of myc-POT1ΔOB expression in the different cell lines in 3 biological replicates. **(B)** Representative TRF (replicate 1). The average telomere length in each lane was calculated with Prism 8. **(C)** Telomere elongation speed in bp/PD from PD 0 to PD 17, 16 or 18 for WT, WT<sup>BIBR</sup> and DKO, respectively, in three replicates.

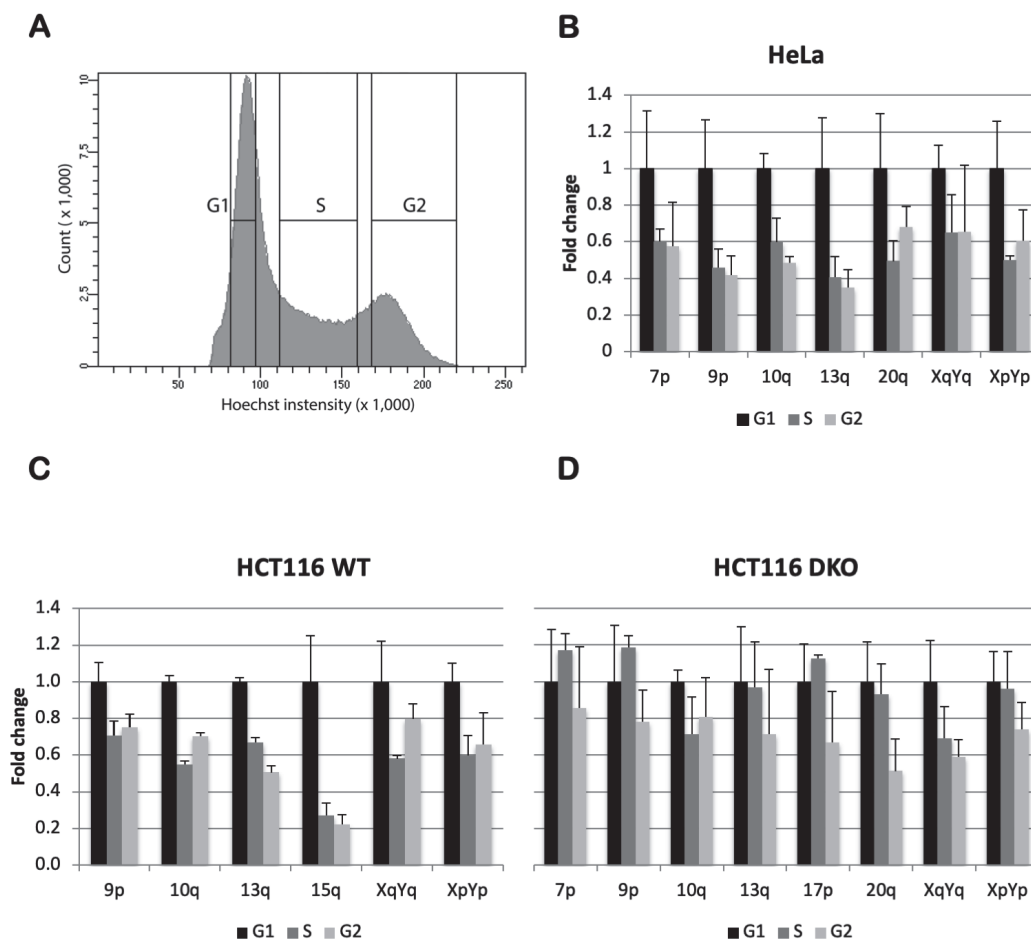
## DKO cells fail to downregulate TERRA in S-phase

TERRA transcripts are detected throughout the cell cycle, but reach minimum levels in the transition between S-phase and G2. Since TERRA has been shown to inhibit telomerase activity *in vitro*, it is hypothesized that this downregulation is important to release telomerase, allowing the extension of telomeres (Porro et al., 2010; Redon et al., 2010). We therefore tested if DKO cells are able to reduce TERRA levels in S-phase.

WT and DKO cells were treated with Hoechst 33342 for 30 minutes and sorted by FACS according to their DNA content into G1, S and G2 phase (Figure 16A). Since the DNA content of late S-phase and G2 cells is very similar, part of cells present in our G2 fraction might in fact be in the end of S-phase.

Immediately after the sorting, total RNA was isolated and TERRA levels were measured by RT-qPCR. HeLa cells were included as a control, since the cell cycle

regulation of TERRA in these cells was already described (Flynn et al., 2015; Porro et al., 2010). As expected, TERRA levels were decreased during S-phase in HeLa and HCT116 WT cells (Figure 16B-C). We did not detect an increase in TERRA in the G2 fraction, possibly due to contamination with late S-phase cells. DKO cells showed no decrease in the expression of the transcripts tested, perhaps with exception of the XqYq (Figure 16D). The downregulation of TERRA, or the lack thereof, was detected independently of the presence of a CpG island promoter, indicating that methylation is unlikely to play a role in this process. Further experiments are needed to assign this downregulation failure to increased transcription or to decreased degradation of TERRA.



**Figure 16: DKO cells fail to downregulate TERRA in S-phase.** (A) Example of cell sorting profile of cells stained with Hoechst 33342. G1, S and G2 phases were gated as shown. TERRA expression in the different cell cycle phases in HeLa (B), HCT116 WT (C) and DKO cells (D) was calculated by RT-qPCR. Two technical replicates are shown.

## TERRA levels are not increased in DNMT3B KO cells

The correlation between telomere length and TERRA expression is not only seen in DKO cells, but also in those of patients suffering from ICF syndrome type I. Although most patients present mutations in the *DNMT3b* gene, Nergadze and colleagues were unable to see any increase in TERRA levels in a DNMT3B KO clone using Northern blots (Nergadze et al., 2009). Our laboratory could also not detect an increase in TERRA by RT-qPCR in this particular clone (Grolimund, 2013). Additionally, this clone has much longer telomeres than the parental cell line (Grolimund, 2013), which could increase heterochromatin formation at the subtelomeres and reduce TERRA expression (Arnoult et al., 2012). Thus, we sought to verify the effect of DNMT3B KO on TERRA expression by creating new KO clones using CRISPR/Cas9.

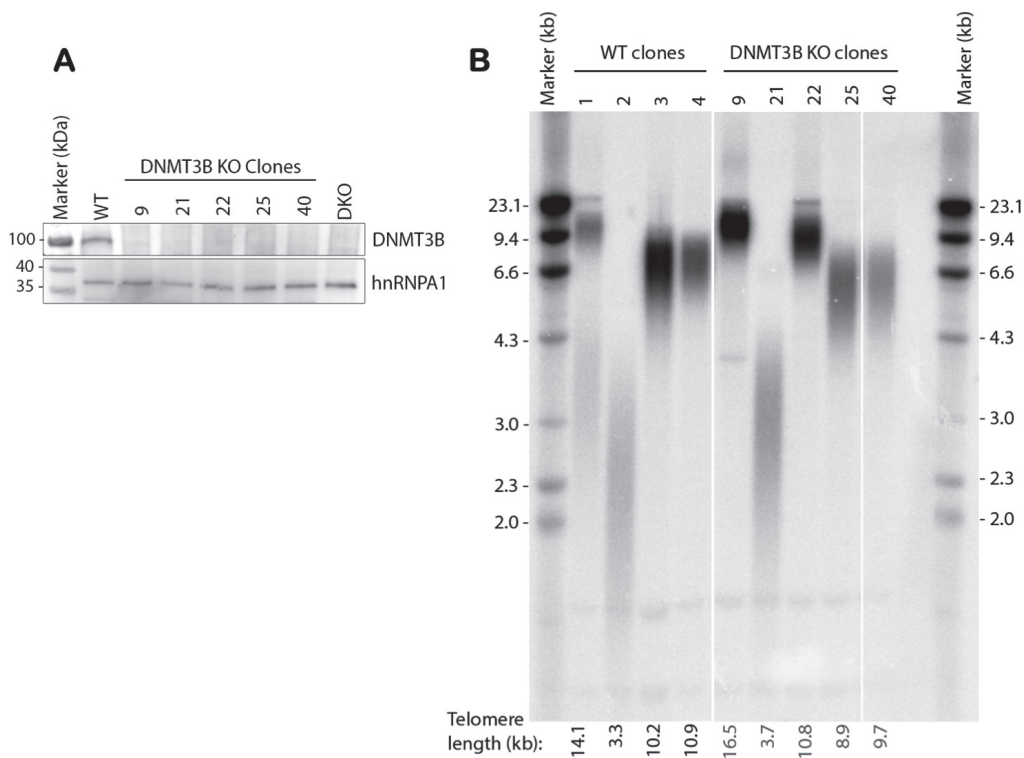
Our CRISPR approach was to target two pairs of gRNAs to exon 3 of the *DNMT3B* gene, downstream of the start codon of the main isoforms for the enzyme. The knockouts were screened by PCR and Western blotting, followed by sequencing. A single nucleotide mutation in the position 18,970 (from the TSS) allowed us to distinguish between the two alleles. We were able to generate five DNMT3B KO clones with different telomere lengths (Figure 17). In all clones, the repair of the damage caused by the Cas9 generated frameshifts and premature stop codons (Table 4).

Table 4: Mutations in *DNMT3B* gene of KO clones. The position of the mutation counting from the transcription start site is indicated within parenthesis.

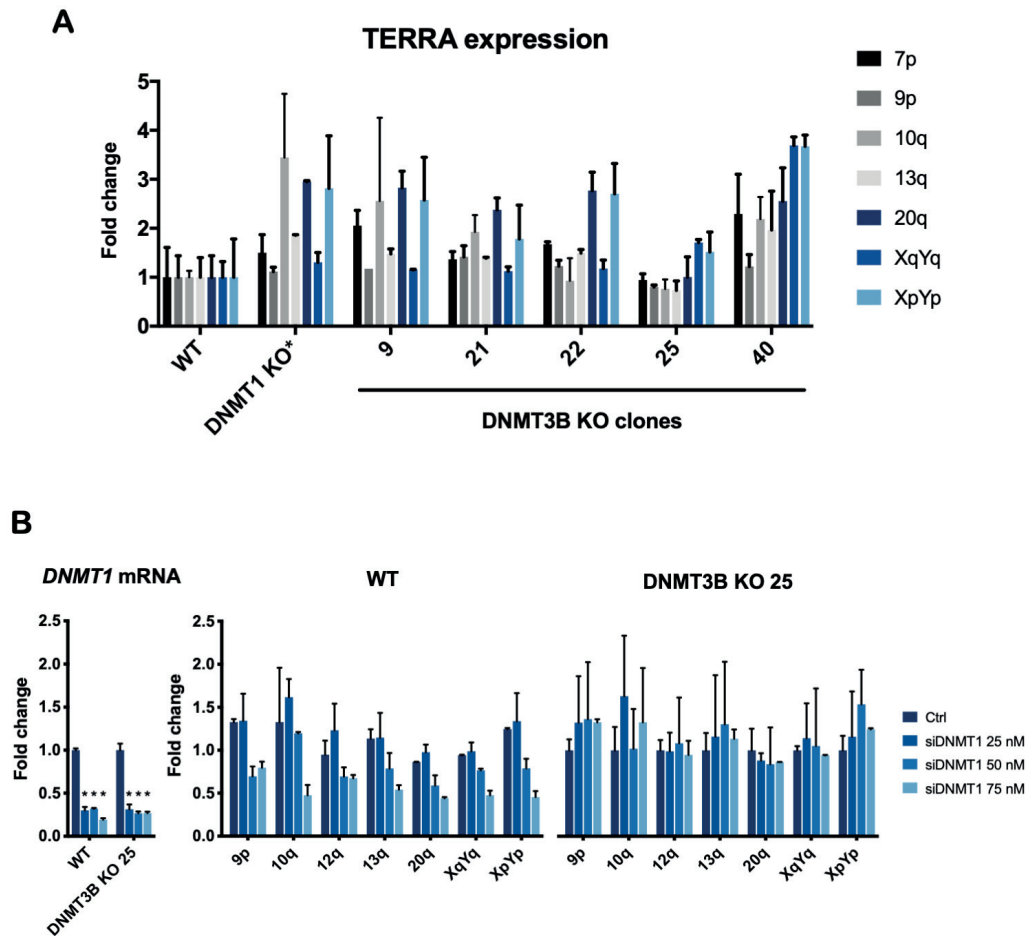
Clone	Allele 1	Allele 2
9	41 nt deletion (18,650)	41 nt deletion (18,650)
21	1 nt deletion (18,650)	1 SNP, 23 nt deletion (18,617; 18,662)
22	41 nt deletion (18,610)	41 nt deletion (18,610)
25	34 nt deletion (18,654)	82 nt deletion, 4 nt insertion, 8 nt deletion (18,664; 18,622: 18,541)
40	34 nt deletion (18,623)	34 nt deletion (18,623)

We then measured TERRA expression in the new DNMT3B KO clones by RT-qPCR (Figure 18). As seen in the original DNMT3B KO clone, TERRA levels were only slightly changed and did not reflect the increase seen in ICF1-derived cells (Toubiana et al., 2018). As these clones present very varied telomere length, we do not believe that telomere length affects this observation.

We also used CRISPR/Cas9 to try to create new DKO clones. For that, DNMT3B KO clones were transfected with different gRNAs against DNMT1. After seeding more than 4,000 clones and screening hundreds of survivals, we were unable to find viable DKO clones, confirming that DNMT1 deletion is lethal in human cells (Liao et al., 2015). However, we tried to simulate the DKO situation by knocking down *DNMT1* using siRNAs. *DNMT1* KD in both WT and DNMT3B KO clone 25 failed to increase TERRA expression (Figure 18B). Therefore, the reason behind the high TERRA expression in the HCT116 DKO cells remains to be determined.



**Figure 17: Generation of HCT116 DNMT3B KO cells.** HCT116 cells were transfected with pairs of gRNAs against DNMT3B. Single cell clones were expanded and screened by Western blot (A). KO clones were confirmed by sequencing. (B) Telomere length of four random WT clones and five DNMT3B KO clones. Average telomere length is indicated below each lane.

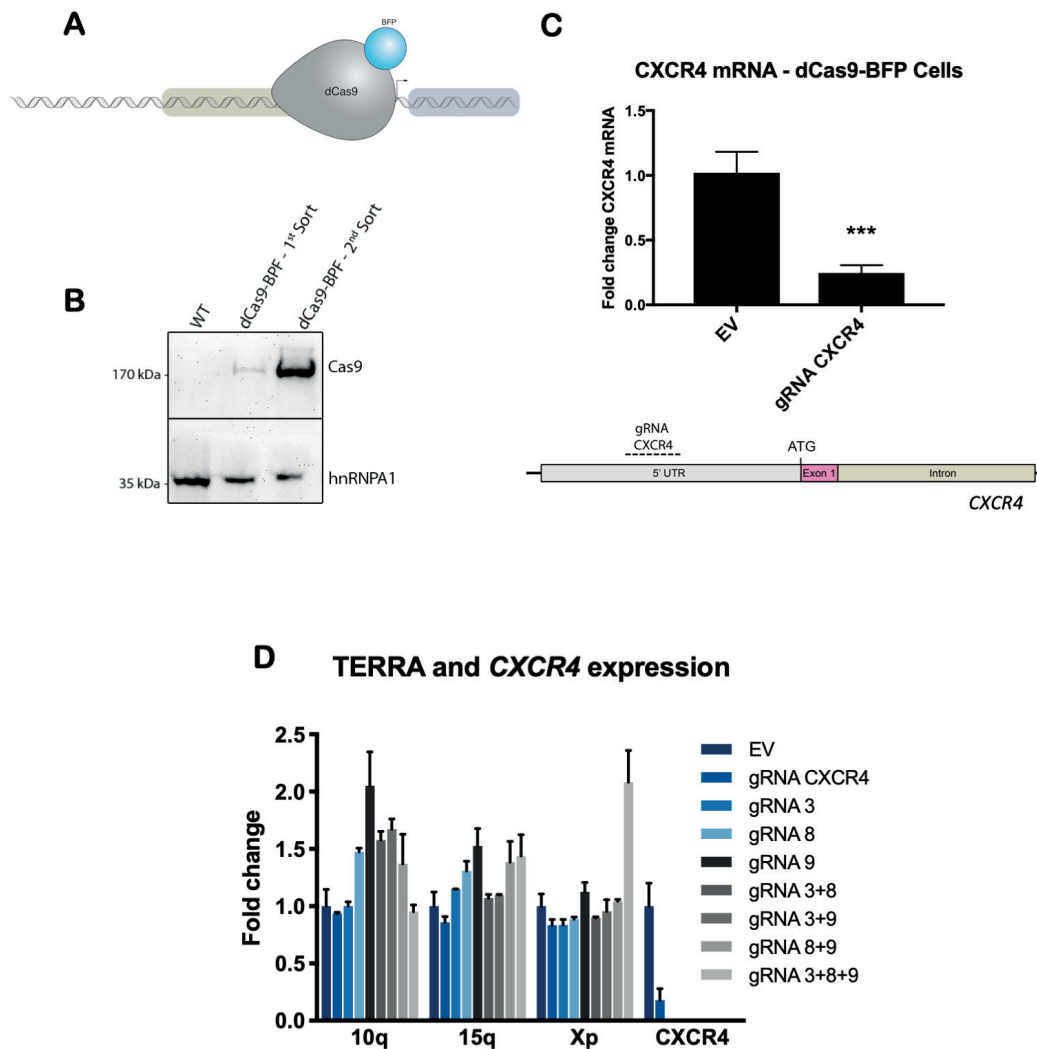


**Figure 18: DNMT3B KO does not increase TERRA expression.** (A) Expression of different TERRA species in DNMT3B KO clones and DNMT1 KO\* (DNMT1  $\Delta$ exons3-5/ $\Delta$ exons3-5). Expression in HCT116 WT population was adjusted to 1. Three technical replicates are shown. (B) TERRA expression in DNMT3B KO 25 upon depletion of DNMT1 for 3 days using varying amounts of siRNA. \* $p < 0.01$ ,  $n = 3$ .

## Tentative downregulation of TERRA using CRISPR interference

TERRA has been shown to inhibit telomerase expression *in vitro* (Redon et al., 2010; Schoeftner and Blasco, 2008) and to cause telomere shortening in *Saccharomyces cerevisiae* (Luke et al., 2008; Pfeiffer and Lingner, 2012). We therefore hypothesized that the high TERRA levels in HCT116 DKO and ICF1 cells could cause the short telomeres seen in these cells, either by inhibiting telomerase or by stimulating end resection.

To determine if TERRA contributes to telomere length regulation, we tried to decrease TERRA expression using CRISPR interference and measure the evolution of the telomere length (Figure 19A). We prepared lentiviruses expressing dCas9-BFP and transduced them into HeLa cells.

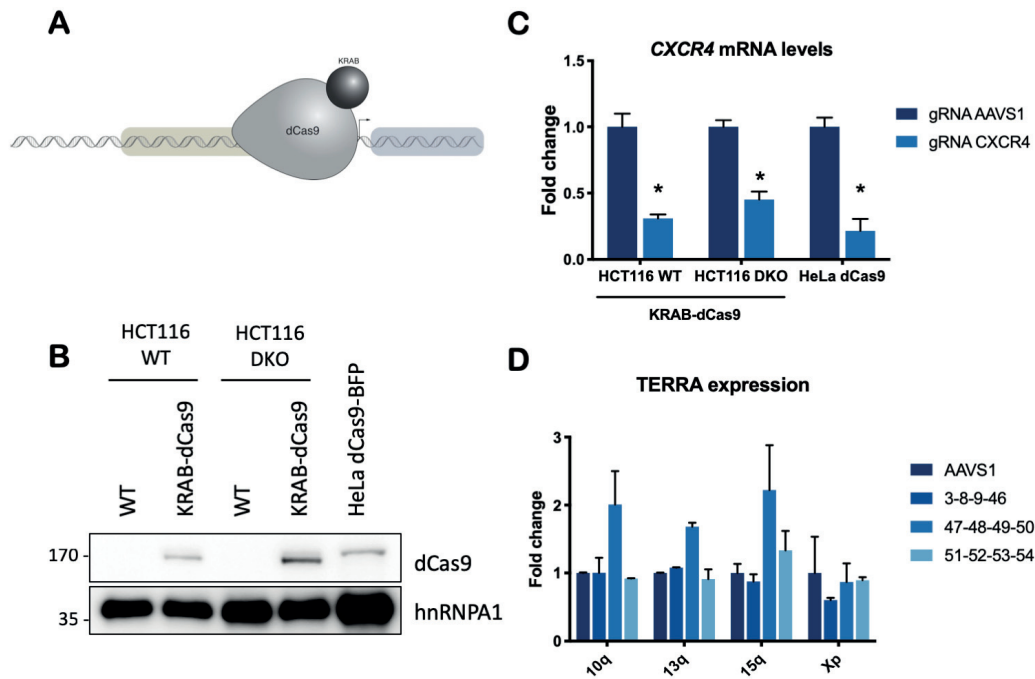


**Figure 19: CRISPR interference with dCas9 does not decrease TERRA levels.** (A) Schematic representation of CRISPR interference (CRISPRi) system, in which the dCas9 alone is used to decrease gene expression. gRNAs targeted to the subtelomeres were used to try to repress TERRA transcription. (B) The BFP fused to dCas9 was used to sort the 10% brightest cells (1<sup>st</sup> sort). Sorted cells were expanded and sorted again for the 10% brightest (2<sup>nd</sup> sort). (C) A gRNA against the CXCR4 gene was used to test the efficiency of the CRISPRi system in the double sorted cells (Gilbert et al., 2013). The position of the gRNA is represented below. CXCR4 mRNA levels were measured by RT-qPCR.  $p < 0.001$ . (D) A representative experiment of many trials performed to downregulate TERRA expression, measured by RT-qPCR. Two technical replicates are shown.



The cells were sorted twice for the 10% brightest BFP-expressing cells (Figure 19B). To characterize the cell line, we transfected the cells with a gRNA previously shown to downregulate the *CXCR4* gene (Gilbert et al., 2013). This gRNA caused 80% downregulation of the *CXCR4* mRNA (Figure 19C). We tried to downregulate TERRA using multiple gRNAs alone or in combination. However, we could never detect a significant decrease in TERRA expression as measured by RT-qPCR. An example is shown in figure 19D.

Since TERRA expression is already very low and sometimes poses a challenge to detection by RT-qPCR, we decided to repeat the experiments in HCT116 WT and DKO cells. This time, we used dCas9 fused to the transcription repressor KRAB (Krüppel associated box) domain (Gilbert et al., 2014) (Figure 20A). The expression levels of KRAB-dCas9 were higher than that of dCas9-BFP in HeLa cells (Figure 20B), but they failed to translate into a more efficient downregulation of our control gene (Figure 20C). In DKO cells, we were not able to achieve more than 60% reduction of *CXCR4*. Nevertheless, we still tried to downregulate TERRA by using gRNAs targeting the subtelomeric CpG islands (gRNAs number 3, 8, 9, 46, 47, 50-52), to the RNA polymerase II binding site (49) or to CTCF binding sites (53-54). Even using a combination of multiple gRNAs, no decrease in TERRA levels was observed (Figure 20D). It is known that the KRAB domain represses gene transcription by inducing DNA methylation via DNMTs and other methyltransferases (Oleksiewicz et al., 2017a; Wiznerowicz et al., 2007). Therefore, the lack of DNMT3B in DKO cells could partially explain why the CRISPR interference was unsuccessful in these cells.



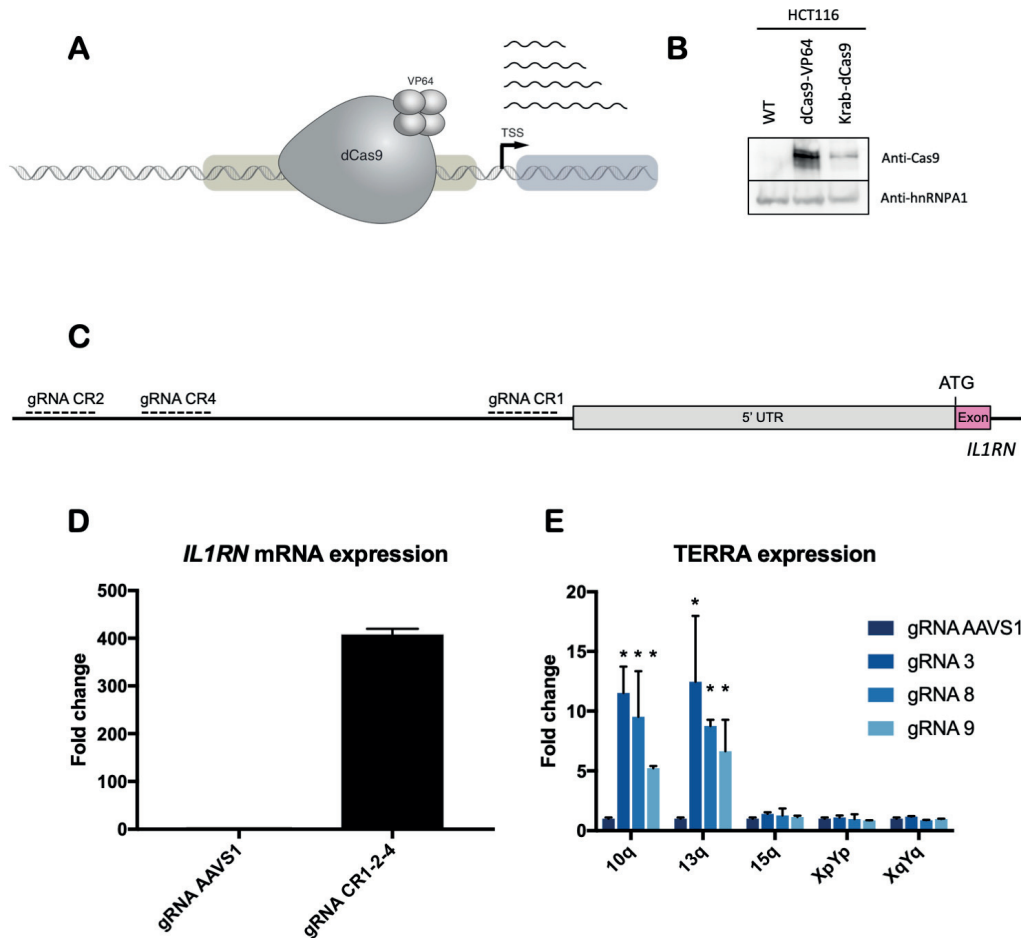
**Figure 20: KRAB-dCas9 does not decrease TERRA levels in HCT116 DKO cells.** (A) Schematic representation of the KRAB-dCas9 CRISPRi system, where gRNAs were used to bring the KRAB-dCas9 to the subtelomeres in order to repress TERRA transcription. (B) Western blot showing the expression level of KRAB-dCas9 in comparison with dCas9-BFP. (C) RT-qPCR for the CXCR4 control gene. Expression in control samples (gRNA AAVS1) was adjusted to 1.  $N = 2$ , \*  $p < 0.05$ . (D) Representative experiment in which combinations of gRNAs were transfected in DKO cells expressing KRAB-dCas9. TERRA levels were measured by RT-qPCR. Two technical replicates are shown.

## TERRA upregulation using CRISPR-based gene activation

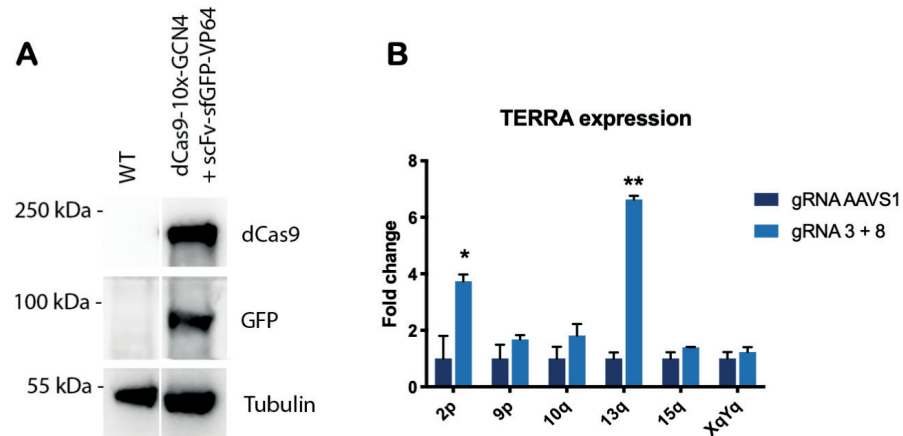
Following the unsuccessful trials of TERRA downregulation, we set out to perform the opposite experiment: overexpress TERRA and monitor possible telomere shortening. To achieve TERRA upregulation, we produced lentiviruses expressing dCas9 fused to four copies of the transcription activator VP16 (dCas9-VP64) as previously described (Perez-Pinera et al., 2013) (Figure 21A). After blasticidin selection, the expression of the construct was stronger than that of KRAB-dCas9 in HCT116 cells (Figure 21B).

To test the functionality of the system, we transfected the cells with three gRNAs against the promoter of the *IL1RN* gene (Perez-Pinera et al., 2013) (Figure 21C). Concomitant expression of the gRNAs led to a 400-fold increase in the mRNA levels of the target gene (Figure 21D). We transiently transfected different

gRNAs against subtelomeric sequences and measured TERRA expression by RT-qPCR. Three gRNAs targeted to the subtelomeric CpG islands triggered 5 to 10-fold upregulation of 10q and 13q TERRAs (Figure 21E). We tried to achieve higher levels of TERRA activation using the dCas9-Suntag, but it did not perform better than the dCas9-VP64 in our hands (Figure 22).



**Figure 21: RNA-guided CRISPR activation leads to 10-fold increase in 10q and 13q TERRAs. (A)** Schematics of the CRISPR activation system. gRNAs target the dCas9-VP64 protein to the locus of interest, such as the subtelomeric CpG islands, where it induces the expression of the downstream gene if correctly positioned. **(B)** Expression of dCas9-VP64 in HCT116 cells in comparison with KRAB-dCas9 expression. **(C)** Illustration of the position of three gRNAs targeted against the IL1RN gene promoter (Perez-Pinera et al., 2013), which were used in combination in **(D)** to induce IL1RN mRNA expression. **(E)** Transient transfection of gRNAs 3, 8 or 9 in HCT116 dCas9-VP64 cells, leading to increased expression of 10q and 13q TERRAs.  $N = 3$ ,  $*p < 0.05$ .



**Figure 22: dCas9-Suntag activation system does not lead to higher TERRA induction than dCas9-VP64.** (A) Lentiviruses expressing dCas9-10x-GCN4 and scFv-sfGFP-VP64 were transduced in HCT116 cells. Expression of both constructs was detected by Western blot with anti-(d)Cas9 and anti-GFP antibodies. (B) RT-qPCR measurement of TERRA expression after transient transfection of a control gRNA (AAVS1) or gRNAs against TERRA promoters (3 and 8). Control TERRA levels were normalized to 1.  $N = 3$ , \*  $p < 0.05$ , \*\*  $p < 0.01$ .

The correct position of the VP64 in the regulatory regions of the gene of interest is crucial for its gene activation performance (Maeder et al., 2013; Perez-Pinera et al., 2013). Therefore, although these gRNAs bind to a number of subtelomeres (Table 5), we were only able to detect changes in the 10q and 13q TERRAs. In addition, the chromatin needs to be sufficiently accessible to the gRNAs and to the dCas9 (Chen et al., 2017).

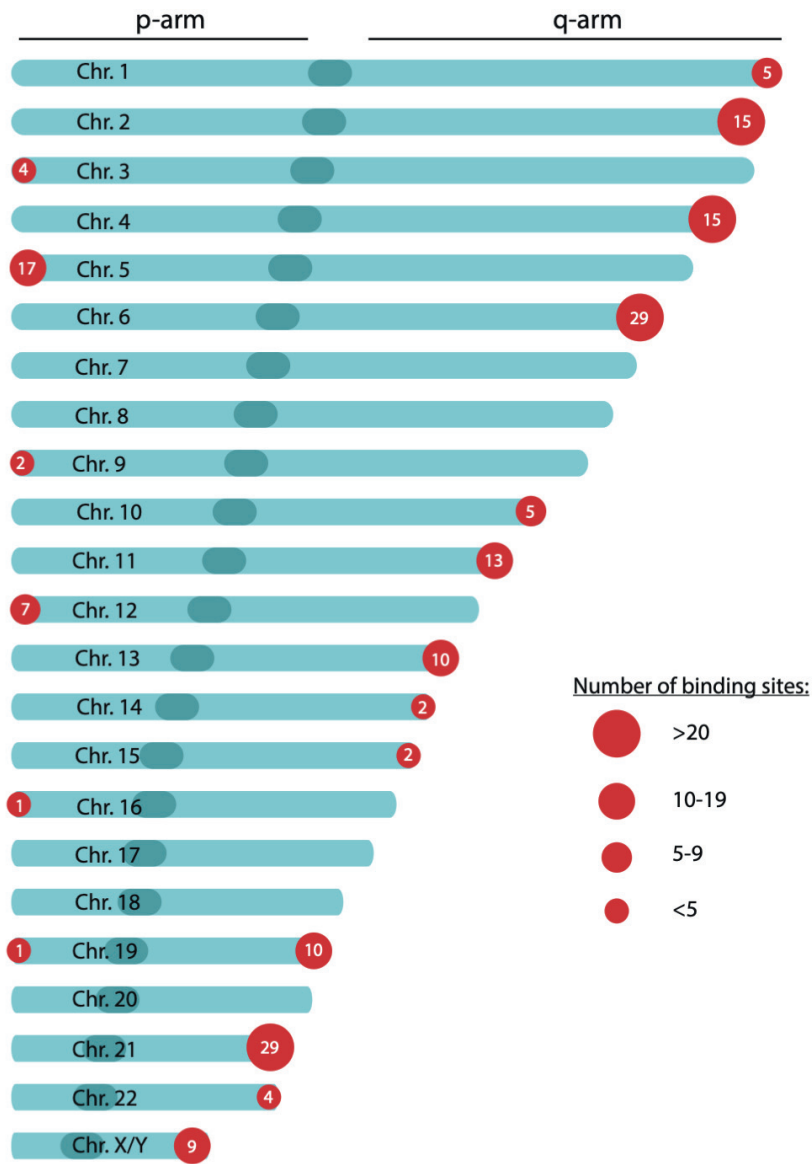
Table 5: Number of binding sites of gRNAs 3, 8 and 9

Chromosome end	gRNA 3	gRNA 8	gRNA 9
1q	2	3	3
2p			10
2q	7	8	8
3p		4	4
4q	7	8	8
5p		17	17
6q	14	15	15
9p		2	2
10q	2	3	3
11q		13	13
12p		7	7
13q	5	5	5
14q	2		
15q		2	2
16p		1	1
19p		1	1
19q		10	1
21q	15	14	14
22q	2	2	2
XqYq		9	7
<b>Total</b>	<b>9</b>	<b>18</b>	<b>19</b>

### **Long-term TERRA upregulation does not cause telomere shortening *in cis***

After determining the best gRNAs to upregulate 10q/13q TERRA, we created a stable cell line overexpressing these specific TERRA molecules to evaluate the effect on telomere length. We produced lentiviruses expressing the gRNAs of interest and transduced gRNAs 3 and 8 (targeting TERRA promoters) or a gRNA targeting the AAVS1 locus in HCT116 cells already expressing dCas9-VP64. The possible binding sites for gRNA 3 and 8 are shown in figure 23. gRNA 9 was not used because it is complementary to gRNA 8 and this could interfere with their performances.

### Localization of Binding Sites for gRNA 3 and 8:

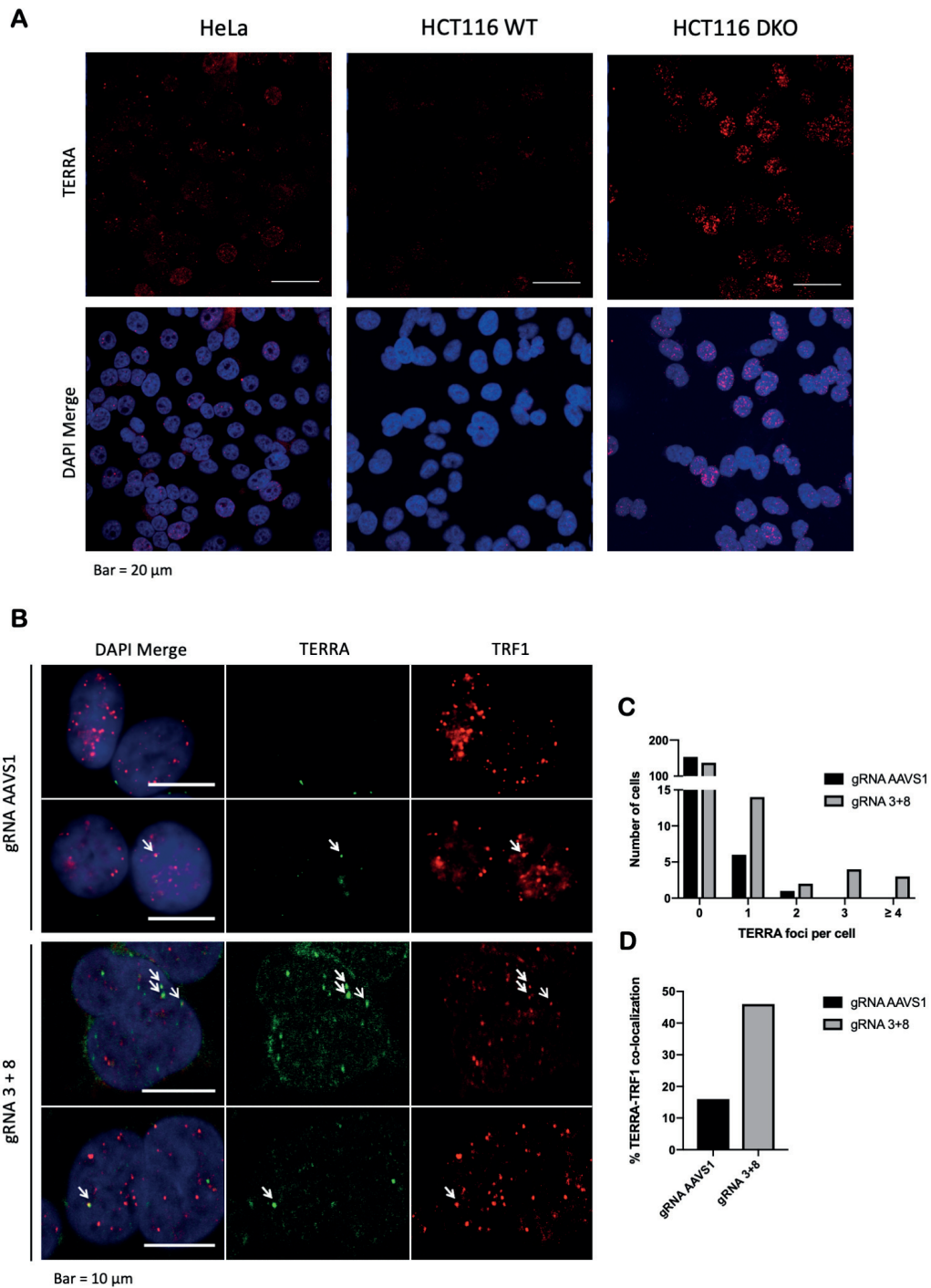


**Figure 23: Illustration of the possible binding sites for gRNAs 3 and 8 in the human genome.** The actual binding does not depend only on the presence of the target sequence, but also of other factors, such as its accessibility, presence of other proteins, chromatin state, and nearby binding of other dCas9-gRNA complexes. These numbers are likely an overestimation of the real binding events.

We next performed TERRA-FISH and TRF1-IF to confirm that the overexpressed TERRA is also localizing to telomeres. As shown in Figure 24A, HCT116 WT cells present fewer foci than HeLa cells, with the vast majority of the cells devoid of any distinguishable TERRA focus. In dCas9-VP64 HCT116 cells expressing the control gRNA, only 8 TERRA foci were identified, while 43 foci were counted in 160 cells expressing gRNAs to overexpress TERRA (Figure 24B-C). Besides having more TERRA foci, these also localized more frequently to telomeres in cells expressing gRNAs 3 and 8 (Figure 24D).

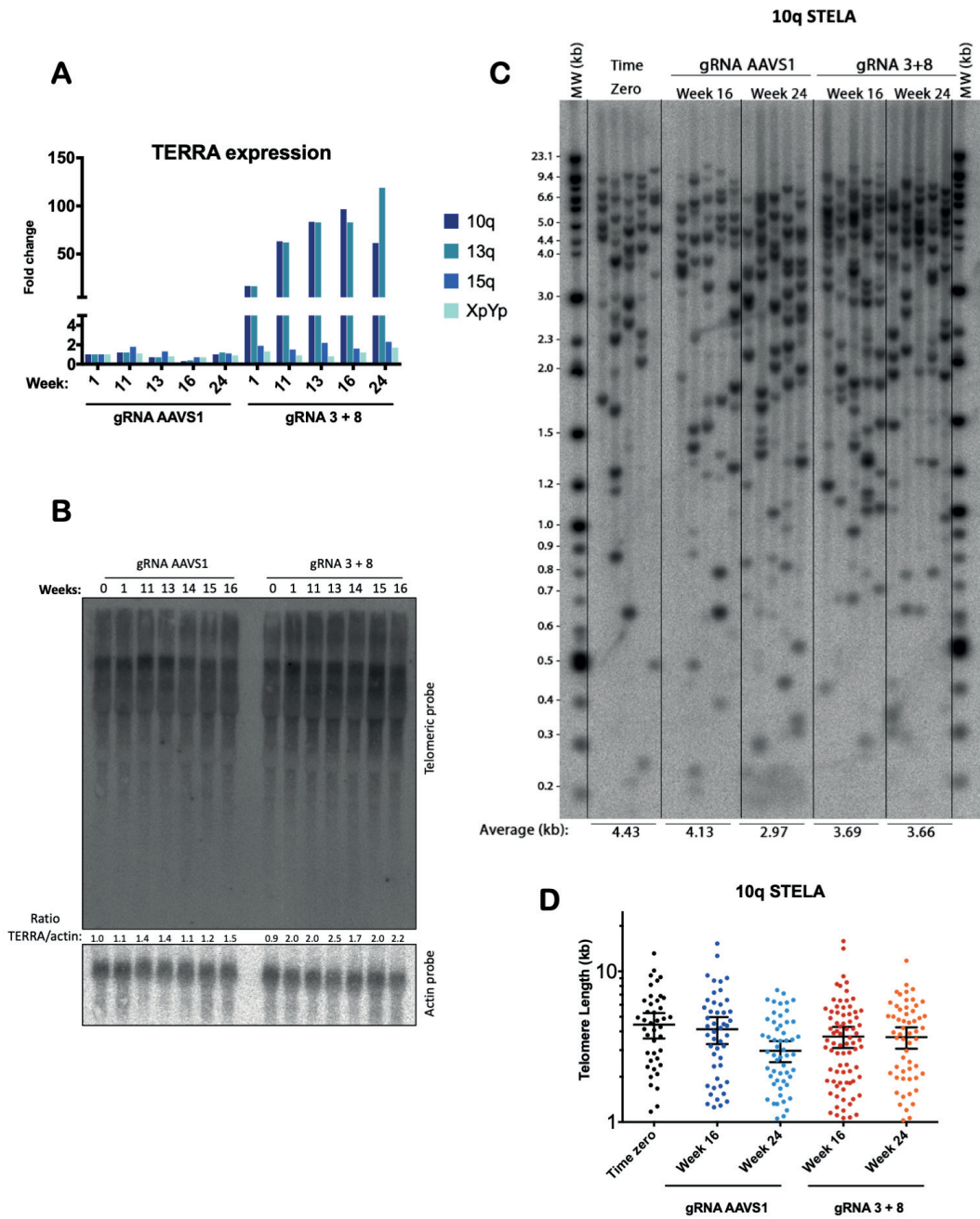
In the weeks immediately following the transduction of gRNA lentiviruses, TERRA increase was similar to those obtained by transient transfection of gRNAs. However, after several weeks, the overexpression was enhanced, reaching almost 100-fold of the original values in week 13 (Figure 25A). Northern blot analysis revealed a 2-fold increase in total TERRA, which is expected since the expression of only 10q and 13q TERRAs was affected (Figure 25B).

To verify if 100-fold 10q/13q TERRA expression could cause telomere shortening *in cis*, we developed 10q STELA. After 24 weeks, we detected small changes in telomere length in both cell lines. The average length of 10q telomeres fluctuated between 4.5 and 3 kb during the course of the experiment. The reduced telomere length seen in the control at 24 weeks, was also likely recovered in the following week, but we have not followed these cell lines further. Similarly, we did not detect significant changes in the 10q telomeres when 10q/13q TERRAs were overexpressed (Figure 25C-D). Thus, this level of TERRA overexpression does not cause telomere shortening *in cis*.



**Figure 24: Overexpressed TERRA co-localizes with TRF1.** (A) TERRA-FISH of HeLa, HCT116 WT and DKO cells. (B) TERRA-FISH and TRF1-IF on HCT116 dCas9-VP64 cells expressing gRNA AAVS1 (control) or 3 and 8 (TERRA overexpression). Arrows indicate co-localization events. (C) Distribution of number of TERRA foci per cell. 160 cells were counted per condition. (D) Quantification of co-localization events between TERRA and TRF1.



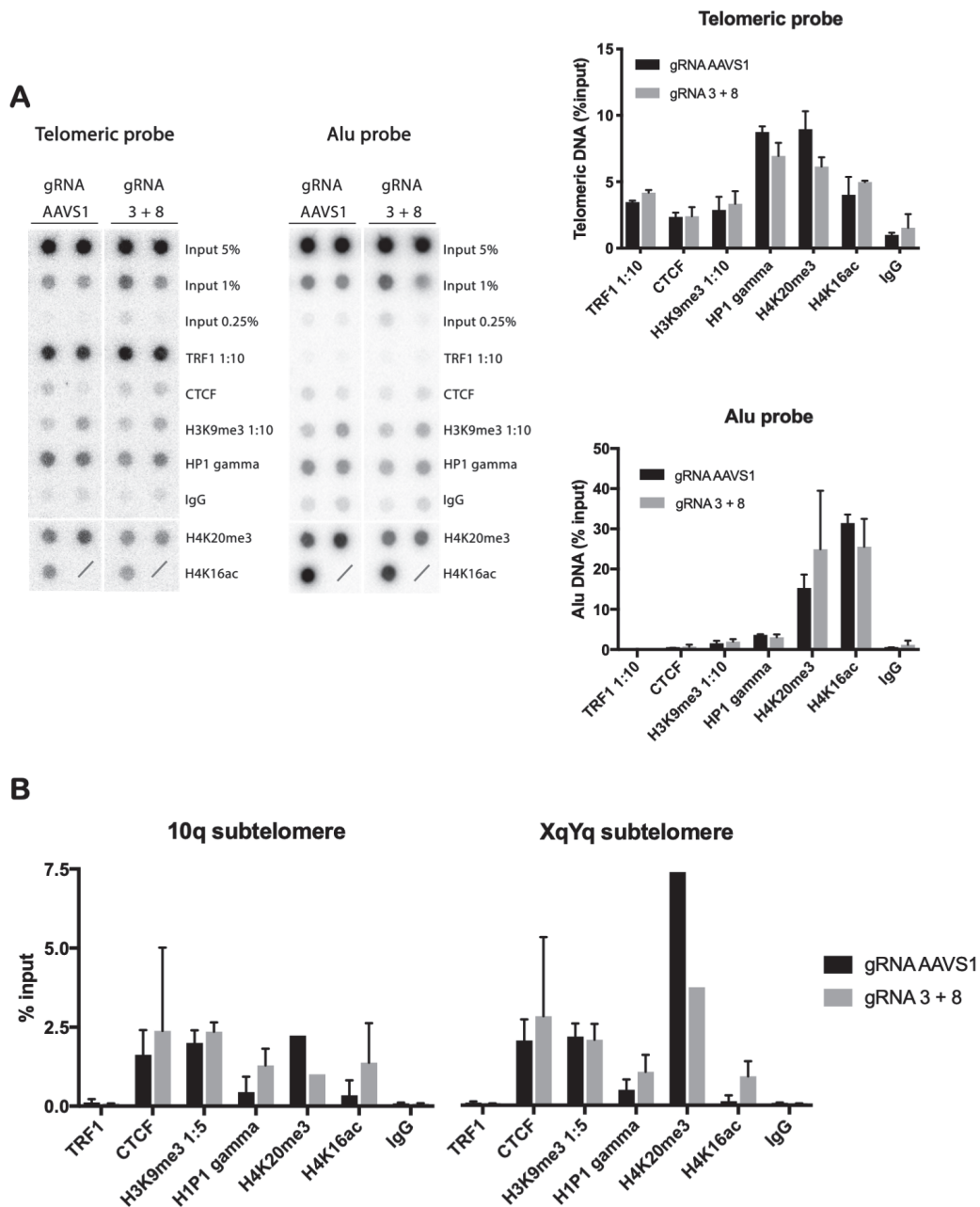


## **Changes in telomeric and subtelomeric chromatin caused by TERRA**

TERRA was shown to contribute to the establishment of telomeric heterochromatin. A direct interaction between TERRA and SUV39H1 has been shown and proposed to be responsible for H3K9me3 deposition at uncapped telomeres (Porro et al., 2014). Likewise, clones showing reduced TERRA levels tend to have less H3K9me3, H4K20me3, H3K4me3 and H3K27me3 at their telomeres (Montero et al., 2018). Since we induce 100-fold overexpression of 10q/13q TERRAs, we questioned if any changes in telomeric or subtelomeric chromatin could be detected.

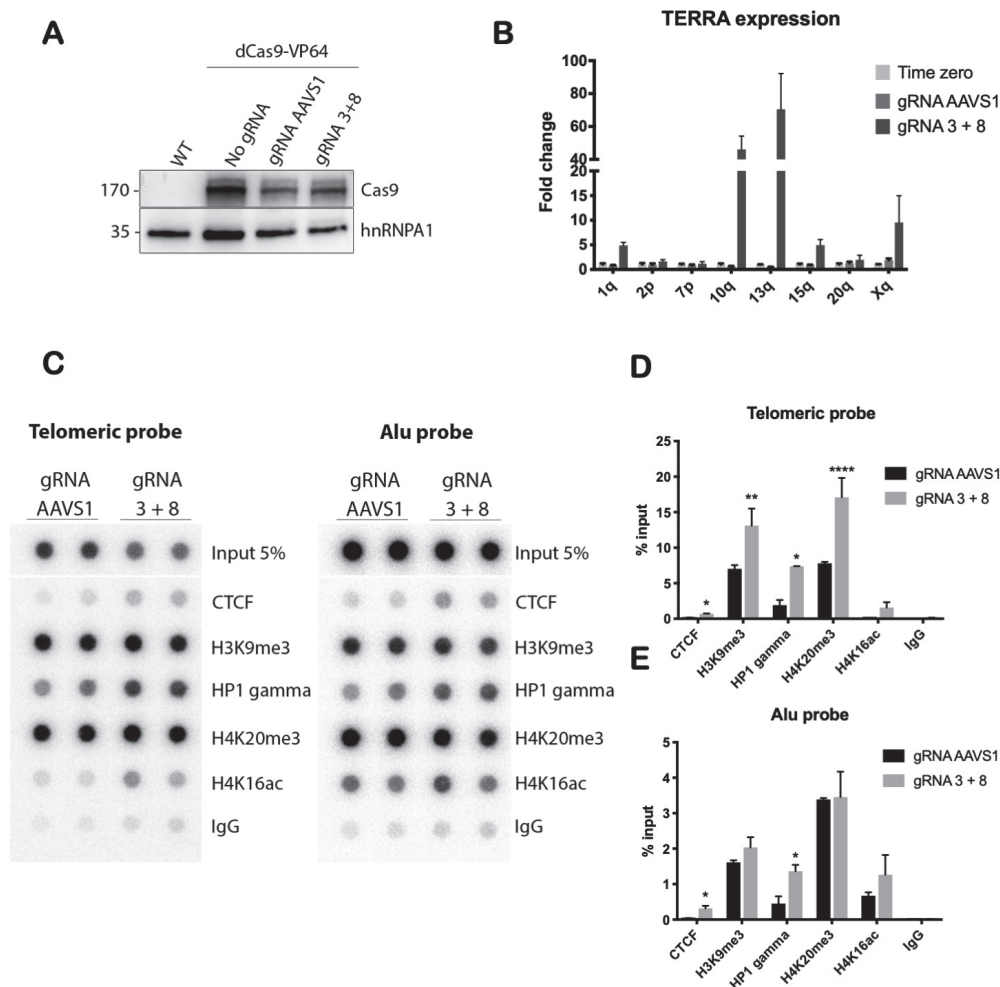
To test that, we measured the levels of H3K9me3, HP1 gamma, H4K20me3 and H4K16ac at the telomeres of dCas9-VP64 HCT116 cells. TERRA overexpression did not modify telomeric and subtelomeric chromatin as measured by ChIP followed by dot blot and qPCR (Figure 26). ChIP-qPCR in HCT116 cells can be challenging, as seen by the huge variation in our results. Therefore, we decided to establish the TERRA activation system in HeLa cells.

dCas9-VP64 and gRNA stable expression in HeLa cells was achieved using the same lentiviruses employed to create the HCT116 cell lines (Figure 27A). Three weeks after gRNA insertion, we could already detect a significant increase in 10q and 13q TERRAs expression (Figure 27B). Surprisingly, moderate to low increases in 1q, 15q and Xq were also detected. gRNA 3 and 8 expression for 4 to 5 weeks led to an increase in the heterochromatin H3K9me3 and H4K20me3 marks at telomeres (Figure 27C-D). CTCF and HP1 gamma signals also increased, but similar changes were seen in the Alu DNA, suggesting that this is not a telomere specific phenotype (Figure 27E).

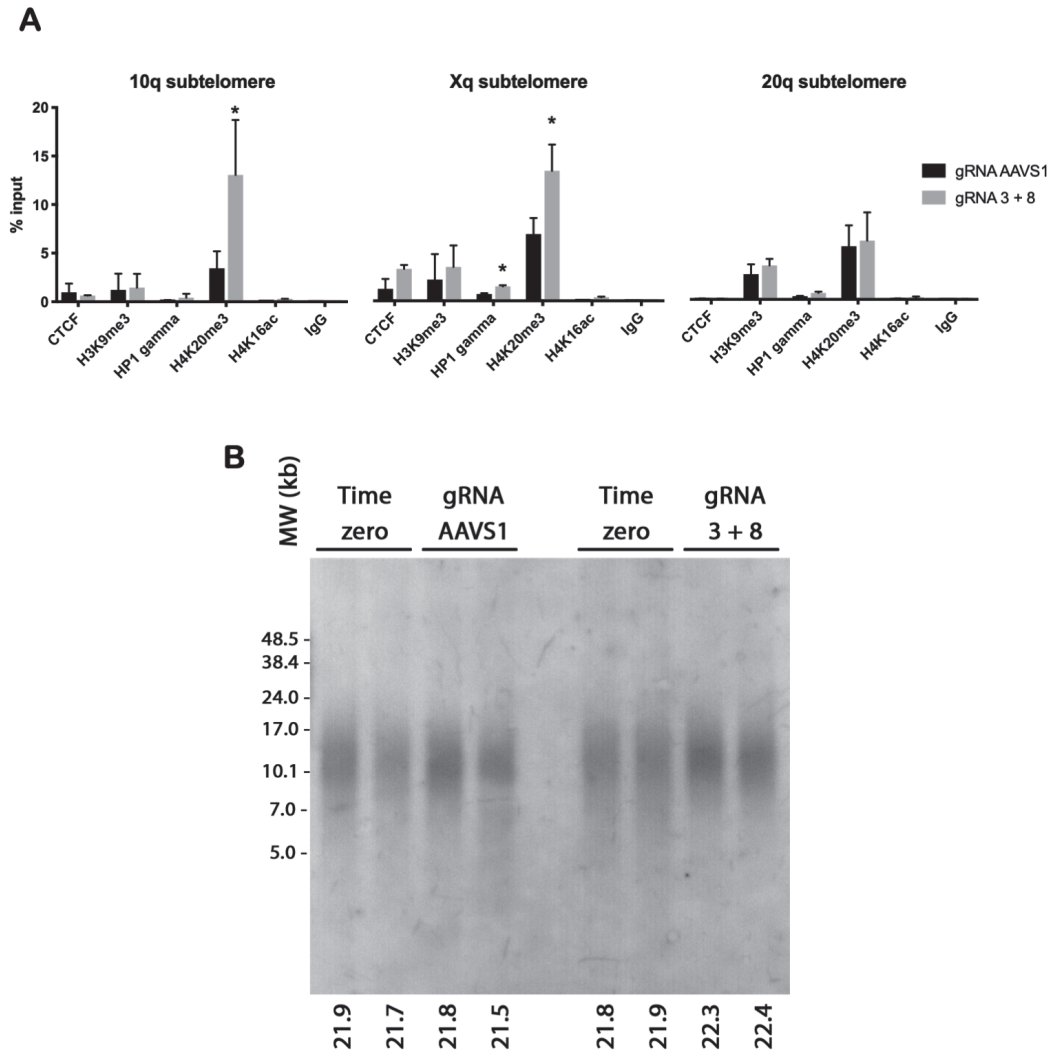


**Figure 26: Overexpression of 10q/13q TERRA does not change telomeric and subtelomeric chromatin structure in HCT116 cells. (A) ChIP-dot blot of HCT116 dCas9-VP64 cells expressing gRNA AAVS1 (control) or 3 and 8 (TERRA overexpression) for 20 to 24 weeks. Telomeric and Alu DNA signals are plotted on the right hand-side.  $N = 3$ , except for H4K16ac ChIP, where  $n = 2$ . (B) ChIP-qPCR of the same cell lines using primers for 10q and XqYq subtelomeres as described in Materials and Methods.  $N = 3$ , except for H4K20me3 ChIP, where  $n = 1$ .**

ChIP-qPCR also showed higher levels of heterochromatin, namely H4K20me3 at subtelomeres (Figure 28A). The H4K20me3 deposition seems to happen exclusively *in cis*, since we detected it in 10q and Xq subtelomeres, but not in the 20q. This effect is unlikely to be caused by the binding of the dCas9-VP64 to the subtelomeres, since one would rather expect that the transcription activator would increase euchromatin marks, such as H4K16ac deposition. Similarly, telomere length also remained stable (Figure 28B), therefore changes in telomere length could not account for subtelomeric chromatin changes.



**Figure 27: TERRA overexpression causes heterochromatin formation in HeLa cells. (A)** Lentiviruses for expression of dCas9-VP64, and gRNAs AAVS1, 3 and 8 were transduced in HeLa cells. dCas9-VP64 expression levels were verified by Western blot. **(B)** TERRA expression on HeLa dCas9-VP64 cells after 3 weeks of gRNA expression. TERRA levels of cells collected immediately before transduction of gRNAs ("Time zero") were adjusted to 1. Two technical replicates are shown. **(C)** ChIP-dot blot of HeLa dCas9-VP64 cells expressing gRNA AAVS1 (control) or 3 and 8 (TERRA overexpression) for 4 to 5 weeks. Telomeric **(D)** and Alu **(E)** DNA signals are plotted on the right hand-side. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ ,  $n = 2$ .



**Figure 28: TERRA changes subtelomeric chromatin in cis in HeLa cells. (A)** ChIP-qPCR of HeLa dCas9-VP64 cells expressing gRNA AAVS1 (control) or 3 and 8 (TERRA overexpression) for 4 to 5 weeks. \*  $p < 0.05$ ,  $n = 3$ . **(B)** TRF of HeLa dCas9-VP64 expressing the indicated gRNAs for 5 weeks. Average telomere lengths are indicated below the gel.

### 3.3 Discussion and perspectives

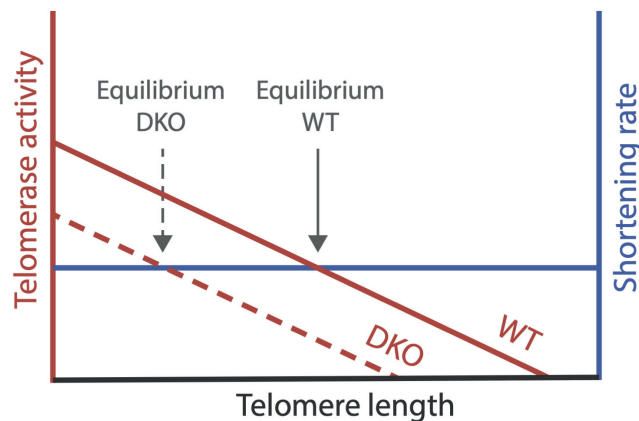
Several reports support that TERRA expression and telomere length are connected. However, the nature of this relationship is complex, depends on the organism, the method used to manipulate TERRA expression and whether telomeres are maintained by telomerase or ALT.

Here we show that telomeres from HCT116 DKO cells are shorter independently of TERRA expression at individual telomeres. To demonstrate that, we developed STELA for the 9p, 15q, 17p and 20q telomeres, and compared their telomere length to their TERRA expression, as measured by RT-qPCR. Even without TERRA overexpression, the 17p and 20q telomeres present the same reduction in telomere length (about 30 to 40%) as their highly expressed counterparts (Figure 12). This suggests that, if TERRA regulates telomere length, it does not do so exclusively *in cis*. As 5'-UUAGGG-3' repeats inhibit telomerase *in vitro* (Redon et al., 2010), it is plausible that the overall levels of TERRA influence telomere length *in trans*.

We also showed that HCT116 DKO cells have difficulties to elongate their telomeres. Upon overexpression of POT1 $\Delta$ OB, one would expect a cell line with such short telomeres to rapidly re-elongate their telomeres. However, the DKO showed a lower elongation rate when compared to WT cells. In contrast to our findings, a previous study suggested that telomerase can efficiently elongate highly transcribed telomeres (Farnung et al., 2012). In this publication, overexpressed hTERT elongated both WT and DKO telomeres at about 200 bp/PD. Furthermore, a TRAP assay revealed no differences in telomerase performance when extracts from DKO cells were used. These results need to be interpreted carefully, since measuring telomere elongation *in vitro* or upon hTERT overexpression may not reflect the *in vivo* situation. In fact, upon POT1 $\Delta$ OB overexpression, which does not interfere with TERRA or telomerase levels, DKO cells do present lower elongation rates (Figure 15).

Additionally, DKO telomeres do not erode faster than WT telomeres. Both cell lines showed a telomere shortening rate of about 20 bp/PD in the presence of 20

$\mu\text{M}$  of BIBR1532. This rate is consistent with published shortening rates for HCT116 cells treated similarly (Farnung et al., 2012). A faster shortening could probably be achieved with higher telomerase inhibition, but higher concentrations of BIBR1532 caused cell cycle arrest and cell death in our hands. Combining the experiments performed to determine telomere elongation and shortening, we conclude that the reduced elongation rate is likely the cause of the short telomere length in DKO cells (Figure 29).



**Figure 29: Proposed telomere length dynamics in HCT116 WT and DKO cells.** While the BIBR1532 experiments showed that both cell lines have the same telomere shortening rate, the elongation rate, reflected by telomerase activity, is consistently lower in DKO cells. Due to this shift in equilibrium, telomeres from DKO cells achieve stability at a lower length.

One of the possible causes for the decreased telomere elongation rate in DKO cells could be sustained TERRA levels in S-phase. Since TERRA inhibits telomerase *in vitro* (Redon et al., 2010; Schoeftner and Blasco, 2008), we hypothesize that failure to downregulate TERRA in S-phase could maintain telomerase inhibition and hamper the extension of short telomeres. Sagie and colleagues showed earlier that ICF1 lymphoblastoid cells fail to downregulate TERRA in S-phase, and maintain telomeric R-loops throughout the cell cycle (Sagie et al., 2017). However, they could also not detect cell cycle regulation of TERRA in WT lymphoblastoid cells, raising the possibility that this is a cell-specific phenotype, rather than an observation relevant to the disease. Since they

show that ICF1 cells have 2 to 3-fold more TERRA:DNA hybrids than WT cells at some telomeres, the authors speculate that the excessive hybrids could contribute to the short telomeres in the syndrome. We were able to detect clear differences in the cell cycle regulation of TERRA between HCT116 WT and DKO (Figure 16), but we currently do not know if that triggers the different telomere elongation rates and if R-loops are concomitantly increased.

Since DKO cells do not reduce TERRA levels in S-phase, we suspected that this reduction is dependent on DNA methyltransferases. Indeed, DNA methylation of subtelomeric CpG islands was responsible for the cell cycle regulation of TERRA, then we should see sustained TERRA levels in DKO cells in S-phase. However, this hypothesis would entail that transcription of CpG-negative chromosome ends, such as 7p, 20q and XpYp, would remain stable in WT cells during S-phase, which is not the case (Figure 16). Since reduction in DNA methylation can have genome-wide effects, it is also possible that the pathways responsible for TERRA transcription or degradation are disturbed by *DNMT3B* deletion. The pathway responsible for TERRA degradation is still poorly understood.

Even though ICF1 and DKO cells have in common short telomeres and high TERRA expression, it is difficult to directly link these phenotypes. The DKO cells could have been extremely valuable to help understanding the disease, but this cell line does not correctly reflect ICF1 cells because (1) it required an additional mutation in *DNMT1* to show increased TERRA expression, and (2) it is a single clone, which could have had very short telomeres even before TERRA expression increased. In an attempt to clarify this relationship, we created five new *DNMT3B* KO clones using CRISPR/Cas9. The clones, however, displayed the same variation of telomere length and TERRA expression which would be expected in WT cells. This, combined with failure of ectopically expressed *DNMT3B* to rescue subtelomeric methylation in ICF cells (Yehezkel et al., 2013), suggests that the subtelomeric methylation pattern created by *DNMT3B* might be established exclusively during embryonic development, and can no longer be modified in somatic cells. Another major difference between ICF and DKO cells is the context in which subtelomeric hypomethylation occurs. ICF cells fail to undergo methylation during embryonic development and therefore mutations in the



DNMT3B *de novo* methyltransferase is sufficient to cause the phenotype. In contrast, HCT116 DKO are derived from a somatic cell line, HCT116 WT, which was already highly methylated in TERRA promoters. Since DNMT3B is not involved in methylation maintenance, it is expected that its KO does not cause major changes in methylation patterns, which would require additional disruption of DNMT1 (Liao et al., 2015).

Unfortunately, creation of additional DKO clones is impossible due to the lethality of *DNMT1* knockout. Even though CRISPR/Cas9 permits us to generate the same *DNMT1*  $\Delta$ exons3-5/ $\Delta$ exons3-5 mutation present in the DKO clone used here, this does not lead us to any definitive conclusions. If the new clones do not behave like the original DKO clone, then it is clear that this observation was just a clonal variation. If they do, then the telomere shortening might still be an indirect effect of reduced methylation levels. In both cases, we do not learn more about TERRA's role in telomere length.

To try to simulate the DKO situation, we knocked down DNMT1 in a DNMT3B KO clone using siRNAs. Even though the *DNMT1* mRNA was significantly reduced, no changes in TERRA levels were observed (Figure 18B). It is possible that the residual expression and activity of DNMT1 was enough to maintain methylation of subtelomeric promoters and hamper TERRA overexpression.

A more direct experiment to determine if TERRA causes the short telomere phenotype in DKO cells is to downregulate TERRA and observe if telomeres elongate. Long term downregulation of TERRA by conventional methods, such as siRNAs, shRNAs or LNA gapmers, is unfeasible due to sequence similarities with the telomeric tract. Therefore, we tried to use CRISPRi, with constant expression of dCas9 or KRAB-dCas9 and gRNAs. More than a hundred combinations of gRNAs were tested, but none led to a significant reduction in TERRA levels. In HeLa and HCT116 WT cells, this could have happened due to their already low TERRA levels or simply because the gRNA was not perfectly positioned, making the gRNA-dCas9 complex innocuous or allowing it to be removed during transcription. DKO cells, although transcribing TERRA ferociously, failed to properly downregulate the control gene. Therefore, we could also not induce

TERRA downregulation in DKO cells using this system (Figure 20). Perhaps the absence of DNMT3B, which deposits the *de novo* methylation and assists KRAB-mediated gene repression (Oleksiewicz et al., 2017b), is one of the reasons for this negative result. The elevated transcriptional activity at subtelomeres can also influence positively or negatively the performance of the dCas9. On one hand, the open chromatin configuration may facilitate the gRNA and dCas9 binding. On the other hand, if these loci are permanently swept by the transcription machinery and transcription factors, these may outcompete or remove the CRISPRi complexes.

When TERRA downregulation failed, we exploited the opposite approach: increasing TERRA transcription in HCT116 WT cells to verify if telomeres become shorter. To do so, we tested both the classical dCas9-VP64 and the dCas9-Suntag system. After testing once again several combinations of gRNA, we found gRNAs 3, 8 and 9 to induce the highest levels of TERRA overexpression. Surprisingly, the dCas9-VP64 slightly outperformed the Suntag system, reaching up to 10-fold upregulation of 10q and 13q TERRAs when the gRNA was transiently delivered (Figure 21 and 22).

To better evaluate the long-term effects of TERRA upregulation, both dCas9-VP64 and gRNAs 3 and 8 were integrated in the genome of HCT116 cells using lentiviruses. The initial TERRA upregulation was very similar to that seen in the transient transfections. However, the upregulation tended to increase with time, stabilizing between 70 and 100-fold after 13 weeks. One possible explanation for this observation could be a local chromatin remodelling caused by the constant presence of the dCas9-VP64 in the subtelomere loci.

Even with such levels of TERRA overexpression, no changes in telomere length were detected after 24 weeks (Figure 25). Previously, a 10 to 20-fold increase in expression of an artificial inducible telomere (tiTEL) also did not lead to telomere shortening *in cis* (Farnung et al., 2012). If TERRA would cause telomere shortening by inhibiting telomerase, perhaps the levels of TERRA attained with these experiments were not sufficient. Northern blot analysis of HCT116 DKO and ICF1 cells reveal that they have at least 50 to 100-fold more TERRA than

control cells (Nergadze et al., 2009; Yehezkel et al., 2008), while we were only able to reach a 2-fold increase (Figure 25). To induce higher expression of TERRA, a single copy of a construct containing a strong promoter harbouring an artificial TERRA could be integrated in the genome by CRISPR/Cas9. The disadvantage of this approach is that it does not express TERRA from an endogenous locus, which could affect the results.

Since TERRA was previously shown to participate in heterochromatin formation at telomeres (Porro et al., 2014), we evaluated if the state of telomeric chromatin of HCT116 cells overexpressing 10q and 13q TERRAs has changed. In the HCT116 cell lines, we did not see any changes in H3K9me3, HP1 gamma, H4K20me3 or H4K16ac after 20 to 24 weeks of TERRA overexpression. Similarly, no changes in subtelomeric chromatin were detected (Figure 26). However, when repeating these experiments with HeLa cells, we found increased levels of H3K9me3 and H4K20me3 at telomeres (Figure 27). The heterochromatin deposition was also reflected in the 10q and Xq subtelomeres, but not in the 20q (Figure 28). The different behaviours of HCT116 and HeLa cells could perhaps be due to their different levels of TERRA overexpression. While HCT116 overexpress only 10q and 13q TERRA with the CRISPR activation system, HeLa cells overexpress 5 out of 8 TERRAs measured, which could lead to a more significant increase in general TERRA levels.

A previous report shows that TERRA binds directly the chromodomain of H3K9 histone methyltransferase SUV39H1, which is responsible for H3K9 trimethylation (Porro et al., 2014). TERRA could interact with SUV39H1 and recruit it to telomeres to induce local heterochromatin formation. This is consistent with our observations that TERRA overexpression leads to H3K9me3 deposition at telomeres. In parallel, we also detected increased H4K20me3 at telomeres. SUV420H2 is the main responsible for catalysing H4K20me3 deposition (Fraga et al., 2005; Schotta et al., 2008) and deficiency of this enzyme causes a 10% increase in telomere length of mouse embryonic fibroblasts (MEFs) (Benetti et al., 2007). Both SUV39H1 and SUV420H2 have in common the presence of a SET domain, very commonly found in histone methyltransferases. However, SUV420H2 does not possess a chromodomain, domain by which

TERRA was shown to bind to SUV39H1 (Porro et al., 2014). The presence of a chromodomain in SUV420H2 would directly suggest that it could bind TERRA, but we might still find that TERRA and SUV420H2 interact through another domain. We currently do not know if this enzyme is directly responsible for the H4K20me3 deposition seen in our experiments. To test that, we will knockdown SUV420H2 and verify if the deposition of this specific heterochromatin is diminished upon TERRA overexpression. If that is the case, we will test if TERRA directly interacts with SUV420H2 by RNA EMSA. Similarly to SUV39H1, TERRA could recruit SUV420H2 to induce heterochromatin formation at telomeres. The final result of this recruitment is unclear, but could serve as a measure to protect telomeres from DDR or to create a feedback loop so that TERRA can regulate its own expression. Supporting the former is the observation that H4K20me3 deposition happens only on those subtelomeres which overexpress TERRA (Figure 28).

# Final considerations

In this thesis, we tackled several fundamental questions in the field of telomere biology. We tried to determine if there is a direct correlation between TERRA expression and telomere length, the factors influencing TERRA transcription, the consequences of TERRA overexpression, and the cause the short telomere phenotype in DKO cells.

In chapter 2, we performed a deep analysis of TERRA expression in several cell lines and concluded that many TERRA species are expressed. Of utmost importance, we demonstrated with two different methods that the 20q TERRA, previously suggested to be the major TERRA, is not the most transcribed one. We also uncovered previously unknown transcription factors influencing TERRA transcription, namely ZNF148, PLAG1, ZFX and EGR1.

In chapter 3, we evaluated the telomere length dynamics of HCT116 DKO cells and discovered that it is their decreased telomere elongation the probable faulty mechanism behind their short telomeres. Additionally, we showed that DKO cells fail to timely downregulate TERRA in S-phase, which could be a major contributor to the reduced telomere elongation rates in this cell line.

We also combined STELA and RT-qPCR to show that TERRA expression and telomere length do not correlate in DKO cells. Importantly, these data suggest that, if TERRA regulates telomere length, it is able to do so *in trans*.

Here we described a method to induce overexpression of endogenous TERRA using CRISPR activation. Although the increased levels of TERRA did not lead to telomere length changes, we show that TERRA plays a role in the deposition of H3K9me3 and H4K20me3 at telomeres. To our knowledge, this was the first time that TERRA has been implicated in H4K20me3 accumulation at telomeres or subtelomeres. Taken together, our results contribute to the understanding of the role of the telomeric long non-coding RNA TERRA in telomere biology.



# Materials and Methods

## Cell culture and transfection

HeLa, U2OS, HLF (passages 10 to 16), and HCT116 WT and DKO human colon carcinoma cells (kind gifts from B. Vogelstein) were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum and 100 U/mL of penicillin/streptomycin. Cells were incubated at 37°C with 5% CO<sub>2</sub>.

## Plasmids and gRNAs

To induce TERRA upregulation, the lentiviral vector Lenti-dCas-VP64\_Blast (gift from Feng Zhang, Addgene #61425) and gRNA-containing plasmids (vector from George Church, Addgene #41824) were integrated in HCT116 and HeLa cells. A puromycin resistance gene (puromycin N-acetyltransferase) was cloned into the gRNA-expressing plasmid. gRNAs were cloned into the modified vector using Gibson assembly (New England Biolabs) as previously published (Mali et al., 2013). For knockouts, oligos containing the gRNA target sequences were cloned using restriction enzymes into the plasmid pSpCas9(BB)-2A-Puro (Addgene #48139) as previously described (Ran et al., 2013). The gRNA sequences are available in Table 10. Suntag plasmids were a gift from Ron Vale (Addgene plasmid #60903 and #60904). pLPC-myc-hPOT1 delta OB was a gift from Titia de Lange (Addgene # 13241).

## Lentiviral vector production and titration

Lentiviral vectors were prepared as previously described (Wiznerowicz et al., 2007). 24h prior to transfection, 11 million 293T cells per dish were plated in 11x 15 cm dishes. 55 µg of pMD2-G and 102 µg of pCMVR8.74 plasmids were

mixed with 157  $\mu\text{g}$  of the transfer vector of interest and transfected using calcium phosphate transfection protocol (CalPhos Mammalian Transfection Kit, Clontech). Cells were incubated overnight with the transfection mix. The supernatant was then discarded and fresh medium was added to the cells. The lentivirus-containing medium was collected and replaced twice in total during 24h. Harvested supernatant was cleared through a 0.22  $\mu\text{m}$  filter unit (Stericup, Millipore) and concentrated by ultracentrifugation for 2h at 50,000  $\times$  g, 16°C (Beckman Coulter Ultra-Clear tubes No. 344058, Rotor SW32Ti). The pellet was resuspended in 500  $\mu\text{L}$  of PBS and frozen at -80°C. Lentivirus titration was performed in HCT116 cells using qPCR as previously described (Barde et al., 2010).

## **Retroviral vector production and transduction**

Seven million 293T cells were seeded in 10 cm dishes 8h before transfection. Cells were transfected with 2  $\mu\text{g}$  pLPC-Nmyc POT1 $\Delta$ OB and 2  $\mu\text{g}$  pcl-Ampho retrovirus packaging vector and 10  $\mu\text{L}$  of lipofectamine 2000 (Invitrogen). 16h after transfection the supernatant was discarded and 10 mL of fresh medium was added. Retrovirus-containing medium was collected 48h post-transfection. Supernatant was cleared through a 0.22  $\mu\text{m}$  syringe filter (Millex-GP, Millipore). This medium was immediately used for transduction.

Six million cells were collected and resuspended in 6 mL of retrovirus-containing medium in the presence of 6  $\mu\text{g}/\text{mL}$  of polybrene. Cells were then seeded in a 10 cm dish. 24h later, 5 mL of fresh medium was added to the dish. 48h post-transduction, cells were split into a 15 cm dish and selected for 12 days with 1  $\mu\text{g}/\text{mL}$  of puromycin.



## Western blot

100,000 cells were incubated at 95°C for 5 min in 2x Laemmli buffer with 100 nM DTT and immediately placed on ice. Proteins were separated on a 4-15% or 7.5% (for DNMT3B) Mini Protean TGX (BioRad) followed by wet transfer onto a 0.2 µm nitrocellulose membrane (Amersham Protran, GE Healthcare). Standard immunoblot protocols were used with the following antibodies: anti-Cas9/anti-dCas9 (1:1,000, G410 Diagenode), anti-hnRNP A1 (1:5,000, 4B10, sc-32301, Santa Cruz), anti-DNMT3B (sc-376043, Santa Cruz), anti-myc (1:1,000, Sigma 9E10), anti-CTCF (1:1,000, Millipore 07-729) and anti-ZNF148 (1:1:1,000), sc-137171, Santa Cruz). HRP-conjugated secondary antibodies anti-mouse (W4021, Promega), anti-goat (sc-2953, Santa Cruz Biotechnology) and anti-rabbit (W4011, Promega) were used at 1:5,000 to reveal the signal of the primary antibody. ECL Spray (Advansta) was used to develop the signal, which was detected using a FluorChem 8900 (Alpha Innotech) or Fusion FX (Vilber).

## Telomere restriction fragment analysis (TRF)

Genomic DNA from 3 to 5 million cells was isolated using Wizard genomic DNA kit (Promega). TRF analysis of HeLa cells was performed by digesting genomic DNA with 30 U of RsaI and 50U of HindI as previously described (Cristofari and Lingner, 2006). For HCT116 cells, minor modifications were introduced: 3 µg of genomic DNA was digested overnight at 37°C with 2.5 U of RsaI, HinfI, AluI, MnlI and HphI (New England Biolabs). Samples were loaded on a 15 cm long 0.8% agarose gel and run at 30V for 16h. Gels were denatured in 0.5 M NaOH and 1.5M NaCl, and neutralized in 1.5 M NaCl and 0.5 M Tris-HCl pH 7.5. In-gel hybridization was performed overnight with P32 radiolabelled telomeric probe (Grolimund et al., 2013) in Church buffer (1% bovine serum albumin, 1 mM EDTA, 0.5 M phosphate buffer, 7% SDS). Radioactive signal was detected with Amershan Typhoon and quantified in AIDA software version 4.06.034. Averages and *P*-values (t-tests corrected for multiple comparisons using the Holm-Sidak method) were calculated using the Prism 8 software.

## Single telomere length analysis (STELA)

STELA was performed as previously described (Baird et al., 2003). Primers used are listed in Table 9. Briefly, 3 to 5 µg of genomic DNA was digested with 100 U of EcoRI-HF (New England Biolabs) overnight at 37°C. Digested DNA was purified with GeneClean Turbo kit (MP Biomedicals). DNA concentration was measured with the Qubit dsDNA High Sensitivity Assay Kit (Thermo Fischer Scientific). 100 ng of DNA was mixed with 1 µM telorette 3 and incubated at 60°C for 10 min. The reaction was cooled down to room temperature before addition of 40 U of T4 ligase (New England Biolabs) and 1 mM ATP in a 10 µL reaction. Ligations were incubated for 12h at 35°C and then heat inactivated for 15 min at 70°C. 200 to 600 pg of ligated DNA was used per PCR amplification. Reactions were done in 15 µL containing 75 mM Tris-HCl (pH 8.8), 20 nM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.01% Tween 20, 1.5 mM MgCl<sub>2</sub>, 0.5 µM TelTail primer, 0.5 µM forward primer, 0.3 mM dNTPs, 1.5 U of Taq Thermoprime Polymerase (ABgene) and 0.15 U of Pwo polymerase (Roche). PCR conditions were the following: 94°C for 2 minutes (initial denaturation), 25 cycles of 94°C for 15 seconds, 65°C for 30 seconds, 68°C for 10 minutes, followed by a 20-minutes final extension at 94°C. PCR reactions were mixed with 3 µL of 6x DNA loading buffer (New England Biolabs) and electrophoresed in a 0.8% agarose gel at 55V for 15 hours. Gels were denatured in 0.5 M NaOH and 1.5M NaCl, neutralized in 1.5 M NaCl and 0.5 M Tris-HCl pH 7.5. DNA was transferred by capillarity in 10x SSC buffer onto a nylon membrane (Hybond N+, GE Healthcare). DNA and membrane were UV-crosslinked and blocked in Church buffer (1% BSA, 1 mM EDTA, 0.5 M phosphate buffer, 7% SDS). Membranes were incubated with P32 radiolabelled telomeric probe overnight at 55°C. Washes were done at 60°C with 2x SSC 0.2% SDS (2 x 30 minutes) and 0.2x SSC 0.2% SDS (2 x 30 minutes). The molecular weight of each band was calculated using AIDA software (version 4.06.034). Averages and *P*-values (t-tests corrected for multiple comparisons using the Holm-Sidak method) were calculated using the Prism 8 software.

## Cell cycle sorting

HeLa, HCT116 WT and DKO cells were seeded in 15 cm dishes at 25-30% confluency the day before the experiment. 5 µg/mL Hoechst 33342 was added for 30 min to each dish. Cells were harvested and sorted at 4°C by a MoFlo Astrios EQ flow cytometer according to their DNA content. Immediately after sorting, cells were centrifuged and lysed in RNA lysis buffer (RA1, Macherey-Nagel) and used for RT-qPCR as described below.

## RT-qPCR

TERRA RT-qPCR was performed as previously described (Feretzaki and Lingner, 2017). Briefly, RNA from 3 to 5 million cells was isolated using NucleoSpin RNA (Macherey-Nagel) and subjected to two on-column and one in-solution DNase digestions. 3 µg of RNA was reverse transcribed using 200 U SuperScript III Reverse transcriptase (Thermo Fischer Scientific), GAPDH and TERRA reverse primers. Reverse transcription was performed at 55°C for 1h, followed by heat inactivation at 70°C for 15 minutes. 10% of the reaction was mixed with 2X Power SYBR Green PCR Master mix (Applied Biosystems) and 0.5 µM forward and reverse qPCR primers (see Table 6 for sequences). qPCR reaction consisted of 10 minutes at 95°C for activation of the DNA Polymerase and 40 cycles at 95°C for 15s and 60°C for 1 min in an Applied Biosystems 7900HT Fast Real-Time System. For comparison between two samples, averages and *P*-values were calculated using t-tests corrected for multiple comparisons using the Holm-Sidak method. For comparing multiple treatments, statistical significance was calculated using two-way ANOVA followed by Tukey's multiple comparison in Prism 8 software.

## **TERRA absolute quantification**

TERRA absolute quantification was done in collaboration with Marianna Feretzaki. Primers used for TERRA qPCR were used to amplify genomic DNA from HeLa or HCT116 cells. PCR products were cloned into pCR-4Blunt-TOPO plasmid (Invitrogen). Specificity of the products was verified by Sanger sequencing. Plasmids were quantified by Qubit dsDNA High Sensitivity Assay Kit (Thermo Fischer Scientific) and used to create a standard curve ranging from 1 to  $10^8$  copies per qPCR reaction. qPCR was done as described above using the standard curve and sample cDNA. The number of copies of a specific TERRA species was extrapolated comparing the CT value of the sample to that of the standard curve.

To be more precise with the quantification of TERRA copies per cell, we cloned 400 to 500 bp of subtelomeric sequences of the 10q, 15q and 20q subtelomeres, followed by 540 bp of telomeric repeats (90 x TTAGGG). A T7 promoter was used for *in vitro* transcription of TERRA. The following RNA product with the correct length was gel eluted and purified. RNA was diluted to create a standard curve and the qPCR was performed as described above. To calculate the number of TERRA copies per cell, we assumed the RNA isolation protocol has 100% efficiency. Primers sequences are listed in Table 7.

## **Northern blot**

10  $\mu$ g of RNA were loaded on a 1.2% formaldehyde agarose gel and separated by electrophoresis as previously described (Azzalin et al., 2007). RNA was transferred overnight to a nylon membrane (Hybond N+, GE Healthcare) in 0.5X TBE. The membrane was UV-crosslinked and pre-hybridized for 1h with Church buffer. The membrane was incubated overnight at 50°C with P32-radiolabelled telomeric probe (Grolimund et al., 2013). After reading the telomeric signal, the membrane was stripped and probed with a 5' P32-labeled  $\beta$ -actin probe (5'-AGTCCGCCTAGAAGCATTTG -3') as loading control.

## **IF and TERRA FISH**

For combined TRF1 IF and TERRA FISH, cells were grown in 18x18 mm coverslips overnight, washed with PBS and pre-extracted for 7 min at 4°C with CSK buffer (100 mM NaCl, 300 mM sucrose, 3 mM MgCl<sub>2</sub>, 10 mM PIPES pH 7, 0.5% triton X-100) supplemented with 10 mM ribonucleoside-vanadyl complex (RVC). Cells were washed with PBS and fixed for 10 min at room temperature with 4% paraformaldehyde pH 7.4. After 3 x 5 minutes washes with PBS, cells were blocked for 30 min in humid chamber with IF blocking solution (10% normal goat serum, 2% BSA in PBS) supplemented with RVC. After removal of excessive blocking solution, cells were incubated 1h with anti-TRF1 antibody (homemade serum 448, 1:1,000) diluted in IF blocking solution. Primary antibodies were washed 3x with 0.05% Tween-20 in PBS. Secondary antibody Alexa-633 anti-rabbit (Thermo Fischer A-21070, 1:500) was incubated for 45 min. After 3 additional washes with 0.05% Tween-20 in PBS, cells were fixed for 10 min at room temperature with 4% paraformaldehyde pH 7.4. RNase digestion with 0.1 mg/mL of RNaseA (Sigma) was performed in negative control samples for 1h at 37°C.

Cells were dehydrated in a 70%, 85%, 100% ethanol series for 5 min with each ethanol concentration. Coverslips were air-dried and incubated upside down in 10 µL of TERRA FISH probe. The probe was prepared by random labelling of 1 µg of a 300 bp telomeric sequence in the presence of 0.5 mM dATP, 0.5 mM dTTP, 1 mM Cy3-dCTP (Perkin-Elmer) and 80 U of Klenow (Invitrogen) for 30 min at 37°C. The probe was purified using DNA Clean & Concentrator kit (Zymo Research) and eluted in 15 µL. The probe was added to 85 µL of hybridization mix (2x SSC, 2 mg/mL BSA, 50% ultrapure formamide, 10% dextran sulfate, 10 mM RVC) and heated to 80°C for 10 min before being chilled on ice. The coverslips and probe were incubated overnight at 37°C in a humid chamber in the dark. Washes were done as follows: 3 washes with 2X SSC, 50% formamide at 39°C for 5 min, 3 washes with 2X SSC at 39°C for 5 min (DAPI at 100 ng/mL was

added to the second wash), one wash with 2X SSC at room temperature for 5 min.

Coverslips were mounted on slides with Vectashield (Vector Laboratories) and imaged with a Zeiss LSM700 confocal microscope.

## **DNMT3B CRISPR/Cas9 deletion**

gRNAs against DNMT3B were cloned into plasmid pSpCas9(BB)-2A-Puro (PX459) as previously described (Ran et al., 2013). gRNA sequences are listed in Table 10.

1.5 million HCT116 cells were transfected in 6-well plates with 1 µg of gRNA DNMT3-1 and 1 µg of gRNA DNMT3-2, or with 1 µg of gRNA DNMT3-3 and 1 µg of gRNA DNMT3-4, together with 10 µL of lipofectamine. Transfections were repeated 3x per gRNA pair. Transfections were pooled 8h after transfection in a 10 cm dish. A 3-day selection with 1 µg/mL of puromycin started 36h after transfection. One week after transfection, single cells were plated in 96-well plates using limited dilution. 3 weeks after plating, clones were screened for DNMT3B deletion by Western blot. Knockouts were confirmed by genotyping (primers in Table 11) and sequencing.

## **siRNA depletion**

siRNAs are listed in Table 8. siRNAs were transfected using calcium phosphate transfection. Cells were harvested 72h later and immediately lysed in RNA lysis buffer (RA1, Macherey-Nagel) for RT-qPCR.

## ChIP-dot blot and ChIP-qPCR

10 million cells were crosslinked for 10 min at room temperature in 2 mL of 1% methanol-free formaldehyde in PBS supplemented with 5% FBS. Formaldehyde was quenched for 10 min with 8 mL of 250 mM Tris pH 8.0 in PBS. Cells were washed twice with PBS and kept hereafter at 4°C until the crosslink reversal. Cells were lysed for 10 min with LB1 (50 mM HEPES-KOH pH 7.4, 140 mM NaCl, 1 mM EDTA, 1% glycerol, 0.5% NP40, 0.25% Triton X-100) on wheel. After spinning for 5 min at 1,700 x g, the pellets were resuspended in LB2 (10 mM Tris pH 8, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA) for 10 min on wheel. Pellets were washed twice in 0.5 mL SDS shearing buffer (10 mM Tris pH 8, 1 mM EDTA, 0.15% SDS) before final resuspension in 1 mL of the same buffer. LB1, LB2 and SDS shearing buffers were supplemented with EDTA-free, protease inhibitor complex (Roche) at all times. Sonication was performed in 1 mL sonication vials with AFA fiber at a concentration of 10 million cells per mL. Sonication was performed in a Focused-Ultrasonicator (E220, Covaris) set to 5% duty cycle, 140 W power and 200 cycles per burst. HCT116 cells were sonicated for 12 min and HeLa cells for 16 min to achieve optimal fragment length (100 to 500 bp).

Sonicated samples were centrifuged at 20,000 x g for 10 min. NaCl and Triton X-100 were added to the supernatants to adjust IP conditions to 10 mM Tris pH 8, 1 mM EDTA, 0.15% SDS, 1% Triton X-100, 150 mM NaCl. Each IP was done overnight on rotating wheel using 1.8 mio cells (200 µL) and 4 µg of each antibody. Antibodies used were: CTCF (Millipore 07-729), TRF1 (homemade, affinity-purified from serum), H3K9me3 (ab 8898), H4K20me3 (ab9053), HP1 gamma (ab10480), H4K16ac (Millipore 07-329), normal rabbit IgG (Santa Cruz sc-2027). The next day, protein A/G magnetic beads (Dynabeads, Thermo Scientific) were washed in 5 mg/mL BSA in PBS and resuspended to their original volume in PBS. 20 µL of beads were added to each IP and incubated for 3h on wheel. Beads were washed twice with low salt wash buffer (10 mM Tris HCl pH 8, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.15% SDS), once with high salt wash buffer (10 mM Tris HCl pH 8, 500 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.15% SDS), LiCl wash buffer (10 mM Tris HCl pH 8, 1 mM EDTA,

250 mM LiCl, 1% NP-40, 1% NaDOC) and TE buffer (10mM Tris-HCl, 1mM EDTA). Beads were incubated for 5 min on wheel between each wash. Washed beads were resuspended in crosslink reversal buffer (20 mM Tris HCl pH 8, 1% SDS, 100 mM NaHCO<sub>3</sub>, 0.5 mM EDTA, 200 µg/mL RNase A) and incubated at 65°C overnight with rotation.

DNA was purified using the NucleoSpin Gel and PCR Clean-up (Macherey-Nagel) with NTB buffer and eluted in 50 µL or 100 µL of MilliQ water for qPCR or dot blot, respectively. qPCR was done in duplicates with 2 µL of eluate, 2 µL ddH<sub>2</sub>O, 5 µL of DNA 2X Power SYBR Green PCR Master mix (Applied Biosystems) and 0.5 µM forward and reverse qPCR primers (0.5 µL each). qPCR reaction for done with 10 minutes at 95°C for activation of the DNA Polymerase and 40 cycles at 95°C for 15s and 60°C for 1 min in an Applied Biosystems 7900HT Fast Real-Time System.

For dot blot, DNA incubated at 95°C for 5 min and chilled immediately after on ice. Samples were spotted onto a Hybond N+ nylon membrane (GE Healthcare) using a Bio-Rad dot blot apparatus. DNA and membrane were UV-crosslinked and blocked in Church buffer (1% BSA, 1 mM EDTA, 0.5 M phosphate buffer, 7% SDS) at 65°C. Incubation with P<sup>32</sup> radiolabeled telomeric probe was done overnight at 65°C. Washes were done with 1x SSC 0.5% SDS (2 x 1h). Incubation with the 5' P<sup>32</sup>-labeled Alu probe (5'-GTGATCCGCCCGCCTCGGCCTCCCAAAGTG-3') was done at 50°C.

The intensity of each dot was calculated using AIDA software (version 4.06.034). Averages and *P*-values (t-tests corrected for multiple comparisons using the Holm-Sidak method) were calculated using the Prism 8 software.

Table 6: qPCR primers

Primer	Sequence (5' - 3')
TERRA RT Primer	CCCTAACCTAACCTAACCTAACCTAA
1q Fw	GCATTCCTAATGCACACATGAC
1q Rev	ACCCTAACCCGAACCCTA
2p Fw	GTAAAGGCGAAGCAGCATTCTCC
2p Rev	TAAGCCGAAGCCTAACTCGTGTC



7p Fw	GGAGGCTGAGGCAGGAGAA
7p Rev	CAATCTCGGCTCACCACAATC
9p Fw	GAGATTCTCCAAGGCAAGG
9p Rev	ACATGAGGAATGTGGGTGTTAT
10q Fw	ATGCACACATGACACCCTAAA
10q Rev	TACCCGAACCTGAACCCTAA
12q Fw	ATTTCCCGTTTTCCACACTGA
12q Rev	CTGTTTGCAGCGCTGAATATTC
13q Fw	CTGCCTGCCTTTGGGATAA
13q Rev	AAACCGTTCTAACTGGTCTCTG
15q Fw	TGCAACCGGGAAAGATTTTATT
15q Rev	GCGTGGCTTTGGGACAAC
17p Fw	GGGACAGAAGTGGATAAGCTGATC
17p Rev	GATCCCACTGTTTTTATTACTGTTCT
18p Fw	TACCTCGCTTTGGGACAAC
18p Rev	CCTAACCCCTCACCTTCTAAC
20q Fw	GCAGCTTCTCAGCACAC
20q Rev	TTTGTTCACGTGTCGATGCG
Xq Fw	AGCAAGCGGGTCCTGTAGTG
Xq Rev	GGTGGAACTTCAGTAATCCGAAA
CXCR4 Fw	GAAGCTGTTGGCTGAAAAGG (Gilbert et al., 2013)
CXCR4 Rev	CTCACTGACGTTGGCAAAGA (Gilbert et al., 2013)
IL1RN Fw	GGAATCCATGGAGGGAAGAT (Perez-Pinera et al., 2013)
IL1RN Rev	TGTTCTCGCTCAGGTCAGTG (Perez-Pinera et al., 2013)
GAPDH Fw	GCGTGGCTTTGGGACAAC
GAPDH Rev	AGCCACATCGCTCAGACAC
CTCF Fw	GACCCACCCCTTCTTCAGATG (Deng et al., 2012)
CTCF Rev	CCACAGCAGCCTCTGCTTCT (Deng et al., 2012)
ATRX Fw	GGTGAAGGAAATGTGGATG
ATRX Rev	CTGGGCTGCTTGGATTAGAA
PLAG1 Fw	ATTGTGATCGCCGGTTCTAC
PLAG Rev	GATCCTTTCGCCCAAATCTCT
EGR1 Fw	CGCAAGAGGCATACCAAGAT
EGR1 Rev	GTAGGAAGAGAGAGAGGAGGTG

ZFX Fw	GTGCCCTCTTGCACATAGAT
ZFX Rev	GATGTCCATCAGGGCCAATAA
ZNF148 Fw	GCAGGCTTTGGACAGAACTA
ZNF148 Rev	GATTTGGGAGGGTCTGGTTATC
WASH F1	GTGTGGAGACCAGCTTCAA
WASH R1	CCATGAAGTACAGGCAGACAA
WASH F2	TGACATGGACACAGCCA
WASH R2	CCTCAAGCCAGCCTTCC
WASH F3	ACTCTGCTAGAGTCCATCCG
WASH R3	TGCTGCTTCTCCAGCTTTC

Table 7: Primers used to create plasmids for TERRA absolute quantification

Primer	Sequence (5' - 3')
2p – Plasmid curve F	GTAAAGGCGAAGCAGCATTCTCC
2p – Plasmid curve R	TAAGCCGAAGCCTAACTCGTGTC
9p – Plasmid curve F	GAGATTCTCCCAAGGCAAGG
9p – Plasmid curve R	ACATGAGGAATGTGGGTGTTAT
17p – Plasmid curve F	GGGACAGAAGTGGATAAGCTGATC
17p – Plasmid curve R	GATCCCCTGTTTTTATTACTGTTCCCT
20q – Plasmid curve F	GCAGCTTTCTCAGCACAC
20q – Plasmid curve R	TTTGTTCACTGTCGATGCG
10q – RNA curve F	GCGGCCGCCGGTGAATAAAATCTTTCCC
10q – RNA curve R	GGATCCCCCTCACCTAACCCCTCACC
15q – RNA curve F	GAACGCGTGTGGATCATTCTCCTCAGGTCAGACCCG
15q – RNA curve R	CTAACCTAAGGATCCTAACCGTGACCCTGACCCCG
20q – RNA curve F	ATGATGGCGGCCGCTGCGACGGGCGTCTCCG
20q – RNA curve R	TATATAGGATCCCCCTGATTATTCAGGGCTGCA

Table 8: siRNA sequences

siRNA	siRNA Target Sequence
DNMT1	TCTGTCCGTTACATGTGT
CTCF_5	TGCGATTACGCCAGTGTAGAA
CTCF_8	CACGTCCGAATACCATGGCAA
CTCF_7	AGGGTGATTATGAGTGGTTCA
ATRX_6	AGCAGCTACAGTGACGACTAA
ATRX_5	ACCGCTGAGCCCATGAGTGAA

ATRX_12	AGGGTTGTGGATTATCTTCCA
PLAG1_8	TCGCCGGTTCTACACCCGAAA
PLAG1_12	CAGAGTCGATATATAGGTAGA
PLAG1_13	TACCTTTATAGTTATCGATCA
EGR1_4	CCCAATTACTATTCCCTTTGA
EGR1_5	CAAACCAATGGTGATCCTCTA
EGR1_7	CCCGTCGGTGGCCACCACGTA
ZFX_7	CAGAGAACGGGTGCTAATTAT
ZFX_9	TAGGTCTGAATCAGTGTCATA
ZFX_10	TAGAGGACTAGTAGATGCTAA
ZNF148_5	AAACTCGGGATCAGATACAAA
ZNF148_5	TAAGCGTAAAGCAGGAAATTA
ZNF148_6	TCGGGCAGTTATACTTAAGCA

Table 9: STELA primers

Primer name	Sequence (5' - 3')
Telorete 3	TGCTCCGTGCATCTGGCATCCCTAACC (Baird et al., 2003)
TelTail	TGCTCCGTGCATCTGGCATC (Baird et al., 2003)
9p Fw Primer	GAGATTCTCCAAGGCAAGG
10q Fw Primer	ATGCACACATGACACCCTAAA
15q Fw Primer	CAGCGAGATTCTCCCAAGCTAAG (Farnung et al., 2012)
17p Fw Primer	GGGACAGAAGTGGATAAGCTGATC
20q Fw Primer	GCAGCTTTCTCAGCACAC

Table 10: gRNA target sequences

gRNA	gRNA sequence (PAM)
AAVS1	GGGGCCACTAGGGACAGGAT(TGG)
TERRA 3 - CpG islands	GCAGGCGCAGAGAGGCG(NGG)
TERRA 8 - CpG islands	CCGGCGCAGGCGCAGAG (NGG)
TERRA 9 - CpG islands	CCTCTCTGCGCCTGCGC (NGG)
TERRA 46 - CpG islands	GCGCCGGCGCAGGCGCAGAG (NGG)
TERRA 47 - CpG islands	CGCAGAGAGGCGCGCCGCGC (NGG)
TERRA 48 - CTCF sites	CACGCCGCTGCTGGCAGCT (NGG)
TERRA 49 - RNAPOL2 sites	CGCGCAGAGACACAGTCCC (NGG)
TERRA 50 - between CTCF and RNAPOL2 sites	AGACACATGCTAGCGGTCC(NGG)
TERRA 51 - CpG islands	GCGCCTCTCTGCGCCTGCGC (NGG)
TERRA 52 - CpG islands	TCTCTGCGCCTGCGCCGGCG (NGG)

TERRA 53 – CTCF sites	TGTCCCTAGCTGCCAGCAGG (NGG)
TERRA 54 – CTCF sites	TTGCCCTAGTCGCCAGCAGG (NGG)
CXCR4	GCAGGTAGCAAAGTGACGCCGA (NGG) (Gilbert et al., 2013)
IL1RN CR1	TGTACTCTCTGAGGTGCTC( NGG) (Perez-Pinera et al., 2013)
IL1RN CR2	ACGCAGATAAGAACCAGTT (NGG) (Perez-Pinera et al., 2013)
IL1RN CR4	GAGTCACCCTCCTGGAAAC (NGG) (Perez-Pinera et al., 2013)
DNMT3B-1	AGACTCGATCCTCGTCAACG (NGG)
DNMT3B-2	CGACTCGCCCCCAATCCTGG (NGG)
DNMT3B-3	GTCGGAGGACTGGTCGCTGC (NGG)
DNMT3B-4	CCGGGGTGCGGATAGCCTCC (NGG)
ZNF148-1	G AGATCGAAGTATGCCTCACC (NGG)
ZNF148-2	G TCTGGCCAGTCTACTGTGTC (NGG)
ZNF148-3	G AGTGCATACTGTAGTCCTTG (NGG)
ZNF148-4	G CCACATTCATCACAGCGAAA (NGG)
ZNF148-5	G AAACAGTATTCACACTGGTA (NGG)

Table 11: Genotyping primers

<b>Primer name</b>	<b>Sequence (5' – 3')</b>
DNMT3B Fw	GGGTACTTGGGCAAGAGCAT
DNMT3B Rev	CCCAAAGCCTCCAGTTGTCT

# References

- Aguilera, A., and García-Muse, T. (2012). R Loops: From Transcription Byproducts to Threats to Genome Stability. *Mol. Cell* *46*, 115–124.
- Aguilera, A., and Gómez-González, B. (2008). Genome instability: a mechanistic view of its causes and consequences. *Nat. Rev. Genet.* *9*, 204–217.
- Ahmed, W., and Lingner, J. (2018). PRDX1 and MTH1 cooperate to prevent ROS-mediated inhibition of telomerase. *Genes Dev.* *32*, 658–669.
- Akutagawa, O., Nishi, H., Kyo, S., Terauchi, F., Yamazawa, K., Higuma, C., Inoue, M., and Isaka, K. (2008). Early growth response-1 mediates downregulation of telomerase in cervical cancer. *Cancer Sci.* *99*, 1401–1406.
- Arnoult, N., Van Beneden, A., and Decottignies, A. (2012). Telomere length regulates TERRA levels through increased trimethylation of telomeric H3K9 and HP1 $\alpha$ . *Nat. Struct. Mol. Biol.* *19*, 948–956.
- Arora, R., Lee, Y., Wischnewski, H., Brun, C.M., Schwarz, T., and Azzalin, C.M. (2014). RNaseH1 regulates TERRA-telomeric DNA hybrids and telomere maintenance in ALT tumour cells. *Nat. Commun.* *5*, 5220.
- Aström, A.K., Voz, M.L., Kas, K., Röijer, E., Wedell, B., Mandahl, N., Van de Ven, W., Mark, J., and Stenman, G. (1999). Conserved mechanism of PLAG1 activation in salivary gland tumors with and without chromosome 8q12 abnormalities: identification of SII as a new fusion partner gene. *Cancer Res.* *59*, 918–923.
- Azzalin, C.M., Reichenbach, P., Khoriauli, L., Giulotto, E., and Lingner, J. (2007). Telomeric Repeat-Containing RNA and RNA Surveillance Factors at Mammalian Chromosome Ends. *Science* *318*, 798–801.
- Bah, A., Wischnewski, H., Shchepachev, V., and Azzalin, C.M. (2012). The telomeric transcriptome of *Schizosaccharomyces pombe*. *Nucleic Acids Res.* *40*, 2995–3005.
- Baird, D.M., Rowson, J., Wynford-Thomas, D., and Kipling, D. (2003). Extensive allelic variation and ultrashort telomeres in senescent human cells. *Nat. Genet.* *33*, 203–207.
- Barde, I., Salmon, P., and Trono, D. (2010). Production and Titration of Lentiviral Vectors. In *Current Protocols in Neuroscience*, J.N. Crawley, C.R. Gerfen, M.A. Rogawski, D.R. Sibley, P. Skolnick, and S. Wray, eds. (Hoboken, NJ, USA: John Wiley & Sons, Inc.), p.
- Benetti, R., Gonzalo, S., Jaco, I., Schotta, G., Klatt, P., Jenuwein, T., and Blasco, M.A. (2007). Suv4-20h deficiency results in telomere elongation and derepression of telomere recombination. *J. Cell Biol.* *178*, 925–936.
- Blasco, M.A. (2007). The epigenetic regulation of mammalian telomeres. *Nat. Rev. Genet.* *8*, 299–309.
- Bogdanović, O., and Veenstra, G.J.C. (2009). DNA methylation and methyl-CpG binding proteins: developmental requirements and function. *Chromosoma* *118*, 549–565.
- Boule, J.-B., and Zakian, V.A. (2007). The yeast Pif1p DNA helicase preferentially unwinds RNA DNA substrates. *Nucleic Acids Res.* *35*, 5809–5818.

- Brown, W.R., MacKinnon, P.J., Villasanté, A., Spurr, N., Buckle, V.J., and Dobson, M.J. (1990). Structure and polymorphism of human telomere-associated DNA. *Cell* *63*, 119–132.
- Campisi, J. (2005). Senescent Cells, Tumor Suppression, and Organismal Aging: Good Citizens, Bad Neighbors. *Cell* *120*, 513–522.
- Chavez, A., Tuttle, M., Pruitt, B.W., Ewen-Campen, B., Chari, R., Ter-Ovanesyan, D., Haque, S.J., Cecchi, R.J., Kowal, E.J.K., Buchthal, J., et al. (2016). Comparison of Cas9 activators in multiple species. *Nat. Methods* *13*, 563–567.
- Chen, X., Liu, J., Janssen, J.M., and Gonçalves, M.A.F.V. (2017). The Chromatin Structure Differentially Impacts High-Specificity CRISPR-Cas9 Nuclease Strategies. *Mol. Ther. Nucleic Acids* *8*, 558–563.
- Chouery, E., Abou-Ghoch, J., Corbani, S., El Ali, N., Korban, R., Salem, N., Castro, C., Klayme, S., Azoury-Abou Rjeily, M., Khoury-Matar, R., et al. (2012). A novel deletion in *ZBTB24* in a Lebanese family with immunodeficiency, centromeric instability, and facial anomalies syndrome type 2. *Clin. Genet.* *82*, 489–493.
- Chu, H.-P., Cifuentes-Rojas, C., Kesner, B., Aeby, E., Lee, H., Wei, C., Oh, H.J., Boukhali, M., Haas, W., and Lee, J.T. (2017). TERRA RNA Antagonizes ATRX and Protects Telomeres. *Cell* *170*, 86–101.e16.
- Colgin, L.M., Baran, K., Baumann, P., Cech, T.R., and Reddel, R.R. (2003). Human POT1 Facilitates Telomere Elongation by Telomerase. *Curr. Biol.* *13*, 942–946.
- Costa, V., Casamassimi, A., Roberto, R., Gianfrancesco, F., Matarazzo, M.R., D’Urso, M., D’Esposito, M., Rocchi, M., and Ciccodicola, A. (2009). DDX11L: a novel transcript family emerging from human subtelomeric regions. *BMC Genomics* *10*, 250.
- Court, R., Chapman, L., Fairall, L., and Rhodes, D. (2005). How the human telomeric proteins TRF1 and TRF2 recognize telomeric DNA: a view from high-resolution crystal structures. *EMBO Rep.* *6*, 39–45.
- Cristofari, G., and Lingner, J. (2006). Telomere length homeostasis requires that telomerase levels are limiting. *EMBO J.* *25*, 565–574.
- Cubiles, M.D., Barroso, S., Vaquero-Sedas, M.I., Enguix, A., Aguilera, A., and Vega-Palas, M.A. (2018). Epigenetic features of human telomeres. *Nucleic Acids Res.* *46*, 2347–2355.
- Cusanelli, E., Romero, C.A.P., and Chartrand, P. (2013). Telomeric Noncoding RNA TERRA Is Induced by Telomere Shortening to Nucleate Telomerase Molecules at Short Telomeres. *Mol. Cell* *51*, 780–791.
- Dai, X., Huang, C., Bhusari, A., Sampathi, S., Schubert, K., and Chai, W. (2010). Molecular steps of G-overhang generation at human telomeres and its function in chromosome end protection. *EMBO J.* *29*, 2788–2801.
- Deaton, A.M., and Bird, A. (2011). CpG islands and the regulation of transcription. *Genes Dev.* *25*, 1010–1022.
- de Greef, J.C., Wang, J., Balog, J., den Dunnen, J.T., Frants, R.R., Straasheijm, K.R., Aytekin, C., van der Burg, M., Duprez, L., Ferster, A., et al. (2011). Mutations in *ZBTB24* Are Associated with Immunodeficiency, Centromeric Instability, and Facial Anomalies Syndrome Type 2. *Am. J. Hum. Genet.* *88*, 796–804.

- Déjardin, J., and Kingston, R.E. (2009). Purification of Proteins Associated with Specific Genomic Loci. *Cell* 136, 175–186.
- Denchi, E.L., and de Lange, T. (2007). Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1. *Nature* 448, 1068–1071.
- Deng, Y., Chan, S.S., and Chang, S. (2008). Telomere dysfunction and tumour suppression: the senescence connection. *Nat. Rev. Cancer* 8, 450–458.
- Deng, Z., Norseen, J., Wiedmer, A., Riethman, H., and Lieberman, P.M. (2009). TERRA RNA Binding to TRF2 Facilitates Heterochromatin Formation and ORC Recruitment at Telomeres. *Mol. Cell* 35, 403–413.
- Deng, Z., Wang, Z., Stong, N., Plasschaert, R., Moczan, A., Chen, H.-S., Hu, S., Wikramasinghe, P., Davuluri, R.V., Bartolomei, M.S., et al. (2012). A role for CTCF and cohesin in subtelomere chromatin organization, TERRA transcription, and telomere end protection: Chromatin organization of human subtelomeres. *EMBO J.* 31, 4165–4178.
- Diman, A., Boros, J., Poulain, F., Rodriguez, J., Purnelle, M., Episkopou, H., Bertrand, L., Francaux, M., Deldicque, L., and Decottignies, A. (2016). Nuclear respiratory factor 1 and endurance exercise promote human telomere transcription. *Sci. Adv.* 2, e1600031.
- Doksani, Y., Wu, J.Y., de Lange, T., and Zhuang, X. (2013). Super-Resolution Fluorescence Imaging of Telomeres Reveals TRF2-Dependent T-loop Formation. *Cell* 155, 345–356.
- Drosopoulos, W.C., Kosiyatrakul, S.T., and Schildkraut, C.L. (2015). BLM helicase facilitates telomere replication during leading strand synthesis of telomeres. *J. Cell Biol.* 210, 191–208.
- Egger, G., Jeong, S., Escobar, S.G., Cortez, C.C., Li, T.W.H., Saito, Y., Yoo, C.B., Jones, P.A., and Liang, G. (2006). Identification of DNMT1 (DNA methyltransferase 1) hypomorphs in somatic knockouts suggests an essential role for DNMT1 in cell survival. *Proc. Natl. Acad. Sci.* 103, 14080–14085.
- Ehrlich, M., Jackson, K., and Weemaes, C. (2006). Immunodeficiency, centromeric region instability, facial anomalies syndrome (ICF). *Orphanet J. Rare Dis.* 1, 2.
- Fang, J., Jia, J., Makowski, M., Xu, M., Wang, Z., Zhang, T., Hoskins, J.W., Choi, J., Han, Y., Zhang, M., et al. (2017). Functional characterization of a multi-cancer risk locus on chr5p15.33 reveals regulation of TERT by ZNF148. *Nat. Commun.* 8, 15034.
- Farnung, B.O., Giulotto, E., and Azzalin, C.M. (2010). Promoting transcription of chromosome ends. *Transcription* 1, 140–143.
- Farnung, B.O., Brun, C.M., Arora, R., Lorenzi, L.E., and Azzalin, C.M. (2012). Telomerase Efficiently Elongates Highly Transcribing Telomeres in Human Cancer Cells. *PLOS ONE* 7, e35714.
- Feretzi, M., and Lingner, J. (2017). A practical qPCR approach to detect TERRA, the elusive telomeric repeat-containing RNA. *Methods* 114, 39–45.
- Flynn, R.L., Centore, R.C., O’Sullivan, R.J., Rai, R., Tse, A., Songyang, Z., Chang, S., Karlseder, J., and Zou, L. (2011). TERRA and hnRNPA1 orchestrate an RPA-to-POT1 switch on telomeric single-stranded DNA. *Nature* 471, 532–536.
- Flynn, R.L., Cox, K.E., Jeitany, M., Wakimoto, H., Bryll, A.R., Ganem, N.J., Bersani, F., Pineda, J.R., Suvà, M.L., Benes, C.H., et al. (2015). Alternative lengthening of telomeres renders cancer cells hypersensitive to ATR inhibitors. *Science* 347, 273–277.

- Fraga, M.F., Ballestar, E., Villar-Garea, A., Boix-Chornet, M., Espada, J., Schotta, G., Bonaldi, T., Haydon, C., Ropero, S., Petrie, K., et al. (2005). Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nat. Genet.* *37*, 391–400.
- Galan-Caridad, J.M., Harel, S., Arenzana, T.L., Hou, Z.E., Doetsch, F.K., Mirny, L.A., and Reizis, B. (2007). Zfx Controls the Self-Renewal of Embryonic and Hematopoietic Stem Cells. *Cell* *129*, 345–357.
- Gao, X.-H., Liu, Q.-Z., Chang, W., Xu, X.-D., Du, Y., Han, Y., Liu, Y., Yu, Z.-Q., Zuo, Z.-G., Xing, J.-J., et al. (2013). Expression of ZNF148 in different developing stages of colorectal cancer and its prognostic value: A large Chinese study based on tissue microarray. *Cancer* *119*, 2212–2222.
- Garcia-Exposito, L., Bournique, E., Bergoglio, V., Bose, A., Barroso-Gonzalez, J., Zhang, S., Roncaioli, J.L., Lee, M., Wallace, C.T., Watkins, S.C., et al. (2016). Proteomic Profiling Reveals a Specific Role for Translesion DNA Polymerase  $\eta$  in the Alternative Lengthening of Telomeres. *Cell Rep.* *17*, 1858–1871.
- Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., et al. (2013). CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell* *154*, 442–451.
- Gilbert, L.A., Horlbeck, M.A., Adamson, B., Villalta, J.E., Chen, Y., Whitehead, E.H., Guimaraes, C., Panning, B., Ploegh, H.L., Bassik, M.C., et al. (2014). Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* *159*, 647–661.
- Gössling, K.L., Schipp, C., Fischer, U., Babor, F., Koch, G., Schuster, F.R., Dietzel-Dahmen, J., Wieczorek, D., Borkhardt, A., Meisel, R., et al. (2017). Hematopoietic Stem Cell Transplantation in an Infant with Immunodeficiency, Centromeric Instability, and Facial Anomaly Syndrome. *Front. Immunol.* *8*.
- Graf, M., Bonetti, D., Lockhart, A., Serhal, K., Kellner, V., Maicher, A., Jolivet, P., Teixeira, M.T., and Luke, B. (2017). Telomere Length Determines TERRA and R-Loop Regulation through the Cell Cycle. *Cell* *170*, 72-85.e14.
- Griffith, J.D., Comeau, L., Rosenfield, S., Stansel, R.M., Bianchi, A., Moss, H., and de Lange, T. (1999). Mammalian Telomeres End in a Large Duplex Loop. *Cell* *97*, 503–514.
- Grill, S., Tesmer, V.M., and Nandakumar, J. (2018). The N Terminus of the OB Domain of Telomere Protein TPP1 Is Critical for Telomerase Action. *Cell Rep.* *22*, 1132–1140.
- Grolimund, L. (2013). A Quantitative Telomeric Chromatin Isolation Protocol (Q-TIP) to Characterize Telomere Composition Changes. EPFL Thesis 6022.
- Grolimund, L., Aeby, E., Hamelin, R., Armand, F., Chiappe, D., Moniatte, M., and Lingner, J. (2013). A quantitative telomeric chromatin isolation protocol identifies different telomeric states. *Nat. Commun.* *4*.
- Hagleitner, M.M., Lankester, A., Maraschio, P., Hulten, M., Fryns, J.P., Schuetz, C., Gimelli, G., Davies, E.G., Gennery, A., Belohradsky, B.H., et al. (2007). Clinical spectrum of immunodeficiency, centromeric instability and facial dysmorphism (ICF syndrome). *J. Med. Genet.* *45*, 93–99.
- Hamperl, S., and Cimprich, K.A. (2016). Conflict Resolution in the Genome: How Transcription and Replication Make It Work. *Cell* *167*, 1455–1467.



- Harel, S., Tu, E.Y., Weisberg, S., Esquilin, M., Chambers, S.M., Liu, B., Carson, C.T., Studer, L., Reizis, B., and Tomishima, M.J. (2012). ZFX Controls the Self-Renewal of Human Embryonic Stem Cells. *PLoS ONE* 7, e42302.
- Heaphy, C.M., Subhawong, A.P., Hong, S.-M., Goggins, M.G., Montgomery, E.A., Gabrielson, E., Netto, G.J., Epstein, J.I., Lotan, T.L., Westra, W.H., et al. (2011a). Prevalence of the Alternative Lengthening of Telomeres Telomere Maintenance Mechanism in Human Cancer Subtypes. *Am. J. Pathol.* 179, 1608–1615.
- Heaphy, C.M., de Wilde, R.F., Jiao, Y., Klein, A.P., Edil, B.H., Shi, C., Bettegowda, C., Rodriguez, F.J., Eberhart, C.G., Hebbar, S., et al. (2011b). Altered Telomeres in Tumors with ATRX and DAXX Mutations. *Science* 333, 425–425.
- Henson, J.D., Cao, Y., Huschtscha, L.I., Chang, A.C., Au, A.Y.M., Pickett, H.A., and Reddel, R.R. (2009). DNA C-circles are specific and quantifiable markers of alternative-lengthening-of-telomeres activity. *Nat. Biotechnol.* 27, 1181–1185.
- Horn, S., Figl, A., Rachakonda, P.S., Fischer, C., Sucker, A., Gast, A., Kadel, S., Moll, I., Nagore, E., Hemminki, K., et al. (2013). TERT Promoter Mutations in Familial and Sporadic Melanoma. *Science* 339, 959–961.
- Huang, F.W., Hodis, E., Xu, M.J., Kryukov, G.V., Chin, L., and Garraway, L.A. (2013). Highly Recurrent TERT Promoter Mutations in Human Melanoma. *Science* 339, 957–959.
- Huffman, K.E., Levene, S.D., Tesmer, V.M., Shay, J.W., and Wright, W.E. (2000). Telomere Shortening Is Proportional to the Size of the G-rich Telomeric 3'-Overhang. *J. Biol. Chem.* 275, 19719–19722.
- Jacob, S., Nayak, S., Kakar, R., Chaudhari, U.K., Joshi, D., Vundinti, B.R., Fernandes, G., Barai, R.S., Kholkute, S.D., and Sachdeva, G. (2016). A triad of telomerase, androgen receptor and early growth response 1 in prostate cancer cells. *Cancer Biol. Ther.* 17, 439–448.
- Jeanpierre, M., Turleau, C., Aurias, A., Prieur, M., Ledest, F., Fischer, A., and Viegas-Pequignot, E. (1993). An embryonic-like methylation pattern of classical satellite DNA is observed in ICF syndrome. *Hum. Mol. Genet.* 2, 731–735.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* 337, 816–821.
- Karlseder, J., Hoke, K., Mirzoeva, O.K., Bakkenist, C., Kastan, M.B., Petrini, J.H.J., and Lange, T. de (2004). The Telomeric Protein TRF2 Binds the ATM Kinase and Can Inhibit the ATM-Dependent DNA Damage Response. *PLoS Biol.* 2, e240.
- Kearns, N.A., Pham, H., Tabak, B., Genga, R.M., Silverstein, N.J., Garber, M., and Maehr, R. (2015). Functional annotation of native enhancers with a Cas9-histone demethylase fusion. *Nat. Methods* 12, 401–403.
- Kelleher, C., Kurth, I., and Lingner, J. (2005). Human Protection of Telomeres 1 (POT1) Is a Negative Regulator of Telomerase Activity In Vitro. *Mol. Cell. Biol.* 25, 808–818.
- Konermann, S., Brigham, M.D., Trevino, A.E., Joung, J., Abudayyeh, O.O., Barcena, C., Hsu, P.D., Habib, N., Gootenberg, J.S., Nishimasu, H., et al. (2014). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* 517, 583–588.

- Koskas, S., Decottignies, A., Dufour, S., Pezet, M., Verdel, A., Vourc'h, C., and Faure, V. (2017). Heat shock factor 1 promotes TERRA transcription and telomere protection upon heat stress. *Nucleic Acids Res.*
- Krones-Herzig, A., Mittal, S., Yule, K., Liang, H., English, C., Urcis, R., Soni, T., Adamson, E.D., and Mercola, D. (2005). Early Growth Response 1 Acts as a Tumor Suppressor *In vivo* and *In vitro* via Regulation of p53. *Cancer Res.* *65*, 5133–5143.
- Lam, Y.C., Akhter, S., Gu, P., Ye, J., Poulet, A., Giraud-Panis, M.-J., Bailey, S.M., Gilson, E., Legerski, R.J., and Chang, S. (2010). SNMIB/Apollo protects leading-strand telomeres against NHEJ-mediated repair. *EMBO J.* *29*, 2230–2241.
- Lee, M., Teber, E.T., Holmes, O., Nones, K., Patch, A.-M., Dagg, R.A., Lau, L.M.S., Lee, J.H., Napier, C.E., Arthur, J.W., et al. (2018). Telomere sequence content can be used to determine ALT activity in tumours. *Nucleic Acids Res.* *46*, 4903–4918.
- Li, B., Oestreich, S., and de Lange, T. (2000). Identification of Human Rap1. *Cell* *101*, 471–483.
- Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* *69*, 915–926.
- Liao, J., Karnik, R., Gu, H., Ziller, M.J., Clement, K., Tsankov, A.M., Akopian, V., Gifford, C.A., Donaghey, J., Galonska, C., et al. (2015). Targeted disruption of DNMT1, DNMT3A and DNMT3B in human embryonic stem cells. *Nat. Genet.* *47*, 469–478.
- Linaropoulou, E.V., Parghi, S.S., Friedman, C., Osborn, G.E., Parkhurst, S.M., and Trask, B.J. (2007). Human Subtelomeric WASH Genes Encode a New Subclass of the WASP Family. *PLoS Genet.* *3*, e237.
- Lingner, J., Cooper, J.P., and Cech, T.R. (1995). Telomerase and DNA end replication: no longer a lagging strand problem? *Science* *269*, 1533–1534.
- Loayza, D., and de Lange, T. (2003). POT1 as a terminal transducer of TRF1 telomere length control. *Nature* *423*, 1013–1018.
- Loayza, D., Parsons, H., Donigian, J., Hoke, K., and de Lange, T. (2004). DNA binding features of human POT1: a nonamer 5'-TAGGGTTAG-3' minimal binding site, sequence specificity, and internal binding to multimeric sites. *J. Biol. Chem.* *279*, 13241–13248.
- Lovejoy, C.A., Li, W., Reisenweber, S., Thongthip, S., Bruno, J., de Lange, T., De, S., Petrini, J.H.J., Sung, P.A., Jasin, M., et al. (2012). Loss of ATRX, Genome Instability, and an Altered DNA Damage Response Are Hallmarks of the Alternative Lengthening of Telomeres Pathway. *PLoS Genet.* *8*, e1002772.
- Luke, B., Panza, A., Redon, S., Iglesias, N., Li, Z., and Lingner, J. (2008). The Rat1p 5' to 3' Exonuclease Degrades Telomeric Repeat-Containing RNA and Promotes Telomere Elongation in *Saccharomyces cerevisiae*. *Mol. Cell* *32*, 465–477.
- Maeder, M.L., Linder, S.J., Cascio, V.M., Fu, Y., Ho, Q.H., and Joung, J.K. (2013). CRISPR RNA-guided activation of endogenous human genes. *Nat. Methods* *10*, 977–979.
- Majerská, J., Redon, S., and Lingner, J. (2017). Quantitative telomeric chromatin isolation protocol for human cells. *Methods* *114*, 28–38.
- Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-Guided Human Genome Engineering via Cas9. *Science* *339*, 823–826.

- Marcand, S., Brevet, V., and Gilson, E. (1999). Progressive cis-inhibition of telomerase upon telomere elongation. *EMBO J.* 18, 3509–3519.
- Mazzolini, R., González, N., Garcia-Garijo, A., Millanes-Romero, A., Peiró, S., Smith, S., García de Herreros, A., and Canudas, S. (2018). Snail1 transcription factor controls telomere transcription and integrity. *Nucleic Acids Res.* 46, 146–158.
- Mohamad, T., Kazim, N., Adhikari, A., and Davie, J.K. (2018). EGR1 interacts with TBX2 and functions as a tumor suppressor in rhabdomyosarcoma. *Oncotarget* 9.
- Montero, J.J., Lopez de Silanes, I., Grana, O., and Blasco, M.A. (2016). Telomeric RNAs are essential to maintain telomeres. *Nat. Commun.* 7, 12534.
- Montero, J.J., López-Silanes, I., Megías, D., F. Fraga, M., Castells-García, Á., and Blasco, M.A. (2018). TERRA recruitment of polycomb to telomeres is essential for histone trimethylation marks at telomeric heterochromatin. *Nat. Commun.* 9.
- Moravec, M., Wischnewski, H., Bah, A., Hu, Y., Liu, N., Lafranchi, L., King, M.C., and Azzalin, C.M. (2016). TERRA promotes telomerase-mediated telomere elongation in *Schizosaccharomyces pombe*. *EMBO Rep.* 17, 999–1012.
- Nergadze, S.G., Farnung, B.O., Wischnewski, H., Khoraiuli, L., Vitelli, V., Chawla, R., Giulotto, E., and Azzalin, C.M. (2009). CpG-island promoters drive transcription of human telomeres. *RNA* 15, 2186–2194.
- Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development. *Cell* 99, 247–257.
- Oleksiewicz, U., Gładych, M., Raman, A.T., Heyn, H., Mereu, E., Chlebanowska, P., Andrzejewska, A., Sozańska, B., Samant, N., Fąk, K., et al. (2017a). TRIM28 and Interacting KRAB-ZNFs Control Self-Renewal of Human Pluripotent Stem Cells through Epigenetic Repression of Pro-differentiation Genes. *Stem Cell Rep.* 9, 2065–2080.
- Oleksiewicz, U., Gładych, M., Raman, A.T., Heyn, H., Mereu, E., Chlebanowska, P., Andrzejewska, A., Sozańska, B., Samant, N., Fąk, K., et al. (2017b). TRIM28 and Interacting KRAB-ZNFs Control Self-Renewal of Human Pluripotent Stem Cells through Epigenetic Repression of Pro-differentiation Genes. *Stem Cell Rep.* 9, 2065–2080.
- Paeschke, K., Bochman, M.L., Garcia, P.D., Cejka, P., Friedman, K.L., Kowalczykowski, S.C., and Zakian, V.A. (2013). Pif1 family helicases suppress genome instability at G-quadruplex motifs. *Nature* 497, 458–462.
- Palm, W., and de Lange, T. (2008). How Shelterin Protects Mammalian Telomeres. *Annu. Rev. Genet.* 42, 301–334.
- Park, Y.-J., Kim, E.K., Bae, J.Y., Moon, S., and Kim, J. (2016). Human telomerase reverse transcriptase (hTERT) promotes cancer invasion by modulating cathepsin D via early growth response (EGR)-1. *Cancer Lett.* 370, 222–231.
- Pascolo, E., Wenz, C., Lingner, J., Huel, N., Priepke, H., Kauffmann, I., Garin-Chesa, P., Rettig, W.J., Damm, K., and Schnapp, A. (2002). Mechanism of human telomerase inhibition by BIBR1532, a synthetic, non-nucleosidic drug candidate. *J. Biol. Chem.* 277, 15566–15572.
- Passantino, R., Antona, V., Barbieri, G., Rubino, P., Melchionna, R., Cossu, G., Feo, S., and Giallongo, A. (1998). Negative regulation of beta enolase gene transcription in embryonic

muscle is dependent upon a zinc finger factor that binds to the G-rich box within the muscle-specific enhancer. *J. Biol. Chem.* *273*, 484–494.

Perez-Pinera, P., Kocak, D.D., Vockley, C.M., Adler, A.F., Kabadi, A.M., Polstein, L.R., Thakore, P.I., Glass, K.A., Ousterout, D.G., Leong, K.W., et al. (2013). RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat. Methods* *10*, 973–976.

Pfeiffer, V., and Lingner, J. (2012). TERRA Promotes Telomere Shortening through Exonuclease 1-Mediated Resection of Chromosome Ends. *PLoS Genet* *8*, e1002747.

Porro, A., Feuerhahn, S., Reichenbach, P., and Lingner, J. (2010). Molecular Dissection of Telomeric Repeat-Containing RNA Biogenesis Unveils the Presence of Distinct and Multiple Regulatory Pathways. *Mol. Cell. Biol.* *30*, 4808–4817.

Porro, A., Feuerhahn, S., Delafontaine, J., Riethman, H., Rougemont, J., and Lingner, J. (2014). Functional characterization of the TERRA transcriptome at damaged telomeres. *Nat. Commun.* *5*, 5379.

Prescott, J., and Blackburn, E.H. (1997). Telomerase RNA mutations in *Saccharomyces cerevisiae* alter telomerase action and reveal nonprocessivity in vivo and in vitro. *Genes Dev.* *11*, 528–540.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. (2013). Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* *152*, 1173–1183.

Rai, R., Chen, Y., Lei, M., and Chang, S. (2016). TRF2-RAP1 is required to protect telomeres from engaging in homologous recombination-mediated deletions and fusions. *Nat. Commun.* *7*.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* *8*, 2281–2308.

Redon, S., Reichenbach, P., and Lingner, J. (2010). The non-coding RNA TERRA is a natural ligand and direct inhibitor of human telomerase. *Nucleic Acids Res.* *38*, 5797–5806.

Rhee, I., Jair, K.W., Yen, R.W., Lengauer, C., Herman, J.G., Kinzler, K.W., Vogelstein, B., Baylin, S.B., and Schuebel, K.E. (2000). CpG methylation is maintained in human cancer cells lacking DNMT1. *Nature* *404*, 1003–1007.

Rhee, I., Bachman, K.E., Park, B.H., Jair, K.-W., Yen, R.-W.C., Schuebel, K.E., Cui, H., Feinberg, A.P., Lengauer, C., Kinzler, K.W., et al. (2002). DNMT1 and DNMT3b cooperate to silence genes in human cancer cells. *Nature* *416*, 552–556.

Riethman, H. (2008). Human subtelomeric copy number variations. *Cytogenet. Genome Res.* *123*, 244–252.

Riethman, H., Ambrosini, A., and Paul, S. (2005). Human subtelomere structure and variation. *Chromosome Res.* *13*, 505–515.

Rudenko, G., and Van der Ploeg, L.H. (1989). Transcription of telomere repeats in protozoa. *EMBO J.* *8*, 2633–2638.

Sagie, S., Ellran, E., Katzir, H., Shaked, R., Yehezkel, S., Laevsky, I., Ghanayim, A., Geiger, D., Tzukerman, M., and Selig, S. (2014). Induced pluripotent stem cells as a model for telomeric abnormalities in ICF type I syndrome. *Hum. Mol. Genet.*

- Sagie, S., Toubiana, S., Hartono, S.R., Katzir, H., Tzur-Gilat, A., Havazelet, S., Francastel, C., Velasco, G., Chédin, F., and Selig, S. (2017). Telomeres in ICF syndrome cells are vulnerable to DNA damage due to elevated DNA:RNA hybrids. *Nat. Commun.* *8*, 14015.
- Schmidt, J.C., and Cech, T.R. (2015). Human telomerase: biogenesis, trafficking, recruitment, and activation. *Genes Dev.* *29*, 1095–1105.
- Schmidtman, E., Anton, T., Rombaut, P., Herzog, F., and Leonhardt, H. (2016). Determination of local chromatin composition by CasID. *Nucleus* *7*, 476–484.
- Schmutz, I., Timashev, L., Xie, W., Patel, D.J., and de Lange, T. (2017). TRF2 binds branched DNA to safeguard telomere integrity. *Nat. Struct. Mol. Biol.* *24*, 734–742.
- Schoeftner, S., and Blasco, M.A. (2008). Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. *Nat. Cell Biol.* *10*, 228–236.
- Schotta, G., Sengupta, R., Kubicek, S., Malin, S., Kauer, M., Callen, E., Celeste, A., Pagani, M., Opravil, S., De La Rosa-Velazquez, I.A., et al. (2008). A chromatin-wide transition to H4K20 monomethylation impairs genome integrity and programmed DNA rearrangements in the mouse. *Genes Dev.* *22*, 2048–2061.
- Sfeir, A., Kosiyatrakul, S.T., Hockemeyer, D., MacRae, S.L., Karlseder, J., Schildkraut, C.L., and de Lange, T. (2009). Mammalian Telomeres Resemble Fragile Sites and Require TRF1 for Efficient Replication. *Cell* *138*, 90–103.
- Sfeir, A., Kabir, S., van Overbeek, M., Celli, G.B., and de Lange, T. (2010). Loss of Rap1 Induces Telomere Recombination in the Absence of NHEJ or a DNA Damage Signal. *Science* *327*, 1657–1661.
- Shechner, D.M., Hacısuleyman, E., Younger, S.T., and Rinn, J.L. (2015). Multiplexable, locus-specific targeting of long RNAs with CRISPR-Display. *Nat. Methods* *12*, 664–670.
- Smirnova, A., Gamba, R., Khorauli, L., Vitelli, V., Nergadze, S.G., and Giulotto, E. (2013). TERRA expression levels do not correlate with telomere length and radiation sensitivity in human cancer cell lines. *Cancer Mol. Targets Ther.* *3*, 115.
- Smogorzewska, A., Karlseder, J., Holtgreve-Grez, H., Jauch, A., and de Lange, T. (2002). DNA Ligase IV-Dependent NHEJ of Deprotected Mammalian Telomeres in G1 and G2. *Curr. Biol.* *12*, 1635–1644.
- Solovei, I., Gaginskaya, E.R., and Macgregor, H.C. (1994). The arrangement and transcription of telomere DNA sequences at the ends of lampbrush chromosomes of birds. *Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol.* *2*, 460–470.
- Spada, F., Haemmer, A., Kuch, D., Rothbauer, U., Schermelleh, L., Kremmer, E., Carell, T., Längst, G., and Leonhardt, H. (2007). DNMT1 but not its interaction with the replication machinery is required for maintenance of DNA methylation in human cells. *J. Cell Biol.* *176*, 565–571.
- Takai, K.K., Kibe, T., Donigian, J.R., Frescas, D., and de Lange, T. (2011). Telomere Protection by TPP1/POT1 Requires Tethering to TIN2. *Mol. Cell* *44*, 647–659.
- Tanenbaum, M.E., Gilbert, L.A., Qi, L.S., Weissman, J.S., and Vale, R.D. (2014). A protein tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* *159*, 635–646.

- Toubiana, S., Velasco, G., Chityat, A., Kaindl, A.M., Hershtig, N., Tzur-Gilat, A., Francastel, C., and Selig, S. (2018). Subtelomeric methylation distinguishes between subtypes of Immunodeficiency, Centromeric instability and Facial anomalies syndrome. *Hum. Mol. Genet.* *27*, 3568–3581.
- Tsui, C., Inouye, C., Levy, M., Lu, A., Florens, L., Washburn, M.P., and Tjian, R. (2018). dCas9-targeted locus-specific protein isolation method identifies histone gene regulators. *Proc. Natl. Acad. Sci.* *115*, E2734–E2741.
- Van Beneden, A., Arnoult, N., and Decottignies, A. (2013). Telomeric RNA expression: length matters. *Cancer Mol. Targets Ther.* 178.
- Van Ly, D., Low, R.R.J., Frölich, S., Bartolec, T.K., Kafer, G.R., Pickett, H.A., Gaus, K., and Cesare, A.J. (2018). Telomere Loop Dynamics in Chromosome End Protection. *Mol. Cell* *71*, 510–525.e6.
- Vojta, A., Dobrinić, P., Tadić, V., Bočkor, L., Korać, P., Julg, B., Klasić, M., and Zoldoš, V. (2016). Repurposing the CRISPR-Cas9 system for targeted DNA methylation. *Nucleic Acids Res.* gkw159.
- Vrbsky, J., Akimcheva, S., Watson, J.M., Turner, T.L., Daxinger, L., Vyskot, B., Aufsatz, W., and Riha, K. (2010). siRNA-Mediated Methylation of Arabidopsis Telomeres. *PLoS Genet.* *6*, e1000986.
- Wang, F., Podell, E.R., Zaug, A.J., Yang, Y., Baciú, P., Cech, T.R., and Lei, M. (2007). The POT1–TPP1 telomere complex is a telomerase processivity factor. *Nature* *445*, 506–510.
- Watson, J.D. (1972). Origin of Concatemeric T7DNA. *Nature. New Biol.* *239*, 197–201.
- Wellinger, R.J., and Zakian, V.A. (2012). Everything You Ever Wanted to Know About *Saccharomyces cerevisiae* Telomeres: Beginning to End. *Genetics* *191*, 1073–1105.
- Wiznerowicz, M., Jakobsson, J., Szulc, J., Liao, S., Quazzola, A., Beermann, F., Aebischer, P., and Trono, D. (2007). The Krüppel-associated Box Repressor Domain Can Trigger de Novo Promoter Methylation during Mouse Early Embryogenesis. *J. Biol. Chem.* *282*, 34535–34541.
- Wu, Y., and Zakian, V.A. (2011). The telomeric Cdc13 protein interacts directly with the telomerase subunit Est1 to bring it to telomeric DNA ends in vitro. *Proc. Natl. Acad. Sci.* *108*, 20362–20369.
- Wu, P., van Overbeek, M., Rooney, S., and de Lange, T. (2010). Apollo Contributes to G Overhang Maintenance and Protects Leading-End Telomeres. *Mol. Cell* *39*, 606–617.
- Wu, P., Takai, H., and de Lange, T. (2012). Telomeric 3' Overhangs Derive from Resection by Exo1 and Apollo and Fill-In by POT1b-Associated CST. *Cell* *150*, 39–52.
- Xi, L., and Cech, T.R. (2014). Inventory of telomerase components in human cells reveals multiple subpopulations of hTR and hTERT. *Nucleic Acids Res.* *42*, 8565–8577.
- Yasuhara, T., Kato, R., Hagiwara, Y., Shiotani, B., Yamauchi, M., Nakada, S., Shibata, A., and Miyagawa, K. (2018). Human Rad52 Promotes XPG-Mediated R-loop Processing to Initiate Transcription-Associated Homologous Recombination Repair. *Cell* *175*, 558–570.e11.
- Yehezkel, S., Segev, Y., Viegas-Péquignot, E., Skorecki, K., and Selig, S. (2008). Hypomethylation of subtelomeric regions in ICF syndrome is associated with abnormally short telomeres and enhanced transcription from telomeric regions. *Hum. Mol. Genet.* *17*, 2776–2789.

Yehezkel, S., Shaked, R., Sagie, S., Berkovitz, R., Shachar-Bener, H., Segev, Y., and Selig, S. (2013). Characterization and rescue of telomeric abnormalities in ICF syndrome type I fibroblasts. *Front. Oncol.* 3.

Zatkova, A., Rouillard, J.-M., Hartmann, W., Lamb, B.J., Kuick, R., Eckart, M., von Schweinitz, D., Koch, A., Fonatsch, C., Pietsch, T., et al. (2004). Amplification and overexpression of the IGF2 regulator PLAG1 in hepatoblastoma. *Genes. Chromosomes Cancer* 39, 126–137.

Zhang, C.Z.Y., Chen, G.G., and Lai, P.B.S. (2010). Transcription factor ZBP-89 in cancer growth and apoptosis. *Biochim. Biophys. Acta BBA - Rev. Cancer* 1806, 36–41.





# Acknowledgements

First, I would like to thank Prof. Joachim Lingner for opening a position for me in his laboratory when I was still a master's student, and for giving me the opportunity to continue my scientific career as a PhD student. I am very grateful to him for believing in me and for all the support during these past 6 years!

I would like to thank all past and current lab members for the great discussions, help and support. I would not have been able to perform this work without the precious help from our post-docs, Galina, Wareed and Gabriela, who would always be available to answer my questions, even when they had little or no time. Special thanks go to Marianna, for the collaborations and for sharing many TERRA-related difficulties with me. I thank Anna Briod for sharing the challenges, apprehensions and excitement during this long PhD path, for putting me back on track when necessary and for encouraging me. I thank Rita, Anna Näger and Trang for sharing countless lunches, for the comradeship and for adding a lot of fun to this adventure. To Thomas, thanks for everything! I am eternally grateful for the support inside and outside the lab and for making me smile even in the hardest moments.

I also thank my mentor, Prof. Oliver Hantschel, for the support and advice during all these years. I thank Nicole and Tatiana for taking care of essential administrative aspects of my PhD and making this journey as smooth as possible.

I am very grateful to the Flow Cytometry, Microscopy and Mass Spectrometry core facilities at EPFL for making several experiments possible.

I specially thank the Boehringer Ingelheim Fonds for the financial support during my PhD and for broadening my horizons.

I am very grateful and honoured that Prof. Sara Selig, Prof. Ana Claudia Marques, Prof. David Suter and Prof. Kei Sakamoto accepted to be part of my thesis committee and gave me suggestions to improve this thesis.

I will forever be indebted to my parents, Paulo and Vera, for giving me all the opportunities they did not have, for sacrificing all they could to give me a good education and for enduring the large distance between us.

I am most grateful to my fiancé, Andreas, for the unconditional support, for the welcoming arms and for being my stronghold in difficult times. Thank you for sharing so many great moments and for discovering the world with me! Finally, I would like to thank our cats, Lee, Lola and Cowboy, who always cheered me up, comforted me and understood when I had no time to rub their bellies properly.

# Patricia RENCK NUNES

## Education

---

**Ecole Polytechnique  
Fédérale, Lausanne,  
Switzerland**

Ph.D. in Molecular Life Sciences

April 2014 – Present

**University of Lausanne,  
Lausanne, Switzerland**

M.Sc. in Medical Biology

September 2012 – January 2014

**Federal University of  
Rio Grande do Sul, Por-  
to Alegre, Brazil**

B.Sc. in Biomedicine

March 2008 – December 2011

## Research Experience

---

### **Ecole Polytechnique Fédérale de Lausanne, Switzerland**

January 2013 - Present

Developed tools to analyze the transcription of telomeres in disease-relevant models. Performed genetic manipulation experiments, including gene knockouts and knockins (CRISPR-Cas9), as well as lentiviral and retroviral transgene expression. Set up a proteomics approach (BioID) to identify proteins directly involved in telomere transcription.

Routinely performed **flow cytometry, RT-qPCR, western, southern and northern blotting.**

PhD partially supported by a fellowship from Boehringer Ingelheim Fonds.

---

### **University of Lausanne, Switzerland**

September 2012 - January 2013

Master project entitled “Establishment of Conditional Knockout Cell Lines by CRISPR-Cas9-Mediated Genome Engineering”, winner of the **Prix de Faculté in 2014.**

Developed tools to optimize gene editing using the CRISPR-Cas9 approach. Performed molecular cloning, transgene expression in different human cell lines and routinely used flow cytometry, western blotting and RT-qPCR to analyze the results.

---

### **Agency for Science, Technology and Research, Singapore**

January 2012 – July 2012

As a research intern, I worked with prof. Philipp Kaldis and Dr. Kasim Diril on the effects on constitutive Cdk1 acti-

---

vation on cell cycle progression and DNA damage. Performed mouse tail **genotyping**, isolation of **mouse embryonic fibroblasts**, **comet assays** and flow cytometry.

---

## Federal University of Rio Grande do Sul, Brazil

June 2009 – December 2011

Part-time research activities in the laboratory of cellular physiology of prof. Paulo Ivo. Participated in the establishment of a protocol to evaluate the effects of different levels of physical activity on the immunity of Wistar rats. Set up culture of primary lymphocytes and performed western blot analysis of different tissues. Developed a protocol for isolation and **culture of primary rat and mouse hepatocytes**.

---

## Publications

---

“Premature activation of Cdk1 leads to mitotic events in S phase and embryonic lethality.”

*Szmyd R, Niska-Blakie J, Diril MK, **Renck Nunes P**, Tzelepis K, Lacroix A, van Hul N, Deng LW, Matos J, Dreesen O, Bisteau X, Kaldis P. Oncogene. 2018*

“Acute exercise boosts cell proliferation and the heat shock response in lymphocytes: correlation with cytokine production and extracellular-to-intracellular HSP70 ratio.”

*Heck TG, Scomazzon SP, **Nunes PR**, Schöler CM, da Silva GS, Bittencourt A, Faccioni-Heuser MC, Krause M, Bazotte RB, Curi R, Homem de Bittencourt PI Jr. Cell Stress Chaperones. 2017*

“Short-term but not long-term hypoglycemia enhances plasma levels and hepatic expression of HSP72 in insulin-treated rats: an effect associated with increased IL-6 levels but not with IL-10 or TNF- $\alpha$ .”

*Ludwig MS, Minguetti-Câmara VC, Heck TG, Scomazzon SP, **Nunes PR**, Bazotte RB, Homem de Bittencourt PI Jr. Mol Cell Biochem. 2014*

---

## Languages

English	C2
French	B2
German	B1
Spanish	B1
Portuguese	Native

---

## Personal Data

---

Nationality: Brazilian  
Work Permit: B  
Date of Birth: 05.07.1990  
Civil Status: Single  
Driver's License: B



