

TERRA Transcription Profiles in Human Cancer Cells

Miguel Filipe Martins Louro Alves
miguel.alves@tecnico.ulisboa.pt

Thesis to obtain the Master of Science Degree in Biomedical Engineering
Supervisors: Dr. Claus Maria Azzalin and Prof. Cláudia Lobato da Silva

Instituto Superior Técnico, Lisboa, Portugal
January 2021

Abstract

TERRA is a long noncoding RNA transcribed from eukaryotic telomeres. It is a major player in the telomerase-independent Alternative Lengthening of Telomeres (ALT) pathway, utilized by approximately 15% of human cancer cells to counteract the telomere shortening that occurs with multiple rounds of cell division, thereby maintaining their immortality. Several ALT cells lines contain elevated levels of TERRA. However, the TERRA transcription profile is still poorly understood, regarding both the levels of TERRA transcription occurring in different cell lines, and the chromosome ends from which TERRA predominantly originates. By long-read sequencing of TERRA, the Azzalin lab found that TERRA levels from chromosome ends 7q and 12q seem to be especially elevated in ALT cells, and hypothesized that quantification of TERRA from these ends could serve as an ALT cancer diagnosis tool. In this project, the TERRA long-read sequencing data was validated by RT-qPCR in a panel of ALT and telomerase-positive cells. TERRA was found to be elevated in U2OS ALT cells at several chromosome ends, particularly 7q, 12q, 16p and 5p. TERRA from these ends was quantified at between 10 and 20 molecules per cell, an average 5-to-10-fold increase compared to telomerase-positive cells. It is proposed that this increase is not sufficient for establishment of a reliable 7q and 12q TERRA diagnosis threshold to discriminate ALT cells. The 7q and 12q ends do not contain the previously characterized CpG dinucleotide-rich TERRA promoter sequences. As such, the second part of this project consisted in the identification, isolation and characterization of a putative new TERRA promoter sequence at these chromosome ends.

Keywords: TERRA, telomere transcription, Alternative Lengthening of Telomeres, RT-qPCR

Acknowledgement: This document was written and made publicly available as an institutional academic requirement and as a part of the evaluation of the MSc thesis in Biomedical Engineering of the author at Instituto Superior Técnico. The work described herein was performed at Instituto de Medicina Molecular João Lobo Antunes (Lisbon, Portugal), during the period of January-December 2019, under the supervision of Dr. Claus Maria Azzalin and Dr. Joana Rodrigues. The thesis was co-supervised at Instituto Superior Técnico by Prof. Cláudia Silva Lobato.

1. Introduction

1.1. Theoretical Background

Telomeres are heterochromatic nucleoprotein structures that cap the ends of linear eukaryotic chromosomes, and protect them from degradation and inappropriate recombination events due to erroneous activation of DNA damage repair (DDR) processes [1]–[3]. The semiconservative DNA replication machinery is unable to fully replicate the extremities of linear DNA, due to the failure of lagging strand synthesis to fully replicate the parental strand. In somatic cells that lack telomere length maintenance mechanisms, this results in progressive loss of telomeric sequence, with each round of replication, a phenomenon known as the end replication problem [1],[4],[5]. Progressive erosion of the telomeric tract sets a finite lifespan in human somatic tissues. After multiple rounds of cell division, telomeres become critically short and lose their protective

function. They become dysfunctional and are recognized as double stranded breaks (DSBs), leading to the induction of DDR pathways. Sustained telomeric DDR triggers replicative senescence, a state of permanent cell cycle arrest [2], [6], [7]. The suppression of the proliferative capacity of somatic cells is a crucial barrier against cancer development and progression [2],[3]. Since cancer cells rely on their immortal potential to maintain aberrant growth, developing mechanisms to counteract telomere shortening and achieve immortality is one of the hallmarks of cancer [8]. Most human cancer cells compensate for the end replication problem by reactivating the reverse transcriptase telomerase[9],[10]. A distinct subset of tumors, approximately 10-15%, are telomerase-negative and employ a different mechanism to elongate telomeres, a specialized homology-directed repair (HDR) pathways known as Alternative Lengthening of Telomeres (ALT) [11],[12]. ALT is commonly found in cancers of mesenchymal or epithelial origin,

including some osteosarcomas, liposarcomas, glioblastomas and astrocytomas [13]. Many of these cancers have particularly poor clinical prognosis. The ALT pathway relies on break-induced replication (BIR) [14]. BIR is a conservative DNA synthesis-based repair process that is engaged at one-ended DNA DSBs and arrested replication forks, and uses one telomere as a copy template to extend another [15]. ALT is dependent on HDR of damaged telomeric DNA, which means that at least a subset of telomeres in ALT cells is maintained physiologically damaged at all times. This sustained damage derives from an ALT-specific Telomeric Replication Stress (ATRS) [16]. This study focuses on a major activator of ALTRs and thus the ALT mechanisms, the long noncoding TERRA (Telomeric Repeat-containing RNA) [17]. TERRA is transcribed at telomeres, from subtelomeric transcription start sites (TSSs) towards the chromosome ends, by RNA Polymerase II (RNAPII), using the C-rich telomeric strand as template. TERRA molecules comprise chromosome-specific subtelomeric sequences at their 5' end and terminate with tandem arrays of G-rich telomeric repeats (5'-UUAGGG-3') [18]. TERRA transcripts interact directly with telomeric DNA forming DNA:RNA hybrids known as R-loops. An R-loop is formed by separating the two strands of dsDNA, with one binding to the complementary RNA and the other being displaced. Telomeric R-loops (telR-loops) are major triggers of ATRS, and, consequently, of the ALT mechanism, as they can physically interfere with replication fork progression, and prime the initiation of BIR [7],[19]. Hence, TERRA plays a pivotal role in sustaining telomere elongation in ALT cells through telR-loop formation. Accordingly, ALT cells have been shown to have elevated levels of TERRA. In humans, most TERRA transcription start sites are within CpG islands, high CG dinucleotide content regions proximal to the telomeric repeat tract [20]. Roughly 75% of human chromosome ends have been shown to contain CpG islands [21]. A human subtelomeric TERRA promoter has been characterized within CpG islands, being a repetitive region conserved across telomeres, known as the "61-29-37" bp repeats [20]. TERRA has been shown to be transcribed from several chromosome ends, with some ends potentially expressing considerably more TERRA transcripts than others. Nonetheless, conflicting reports exist on TERRA transcription across chromosome ends [20]-[23]. There is currently no consensus on which chromosome ends produce highest amount of TERRA transcripts, or if a subset of chromosome ends could be responsible for the majority of total TERRA transcripts. These chromosome ends could represent potential targets for TERRA-focused clinical approaches for ALT cancers, on both the therapeutical and diagnostic side.

1.2. ONT-seq and hypothesis

To obtain TERRA expression profiles for every human chromosome end, the Azzalin lab at IMM performed Oxford Nanopore long-read sequencing (ONT-seq) [24] of a cDNA library produced from purified TERRA RNA from two ALT (U2OS, SAOS-2) and two telomerase-positive (HEK293T, HeLa) cell lines. Two chromosome ends, the long arms of chromosomes 7 and 12 (7q and 12q chromosome ends) emerged as producing the most abundant TERRA transcripts in ALT cells. The 7q and 12q chromosome ends have high homology, and contain a subtelomeric CpG island sequence proximal to the telomeric repeats that is highly conserved

between the two ends. (7q_12q subtelomeric sequence). It is interesting to note that the 7q and 12q chromosome ends do not contain the previously characterized 61-29-37 repeats TERRA promoter [20]. This indicates that the ONT-seq results could be unveiling new TERRA promoters within subtelomeric CpG islands. The high levels of TERRA transcribed from 7q and 12q compared to other ends in ALT cells and to the same end in telomerase-positive cells, makes these telomeres compelling targets for clinical purposes. The Azzalin lab hypothesized that quantification of 7q- and 12q-TERRA might represent a valuable and sensitive tool for ALT diagnosis. The only currently available diagnosis assay is the C-circle assay, which relies on detection of extra-chromosomal circular DNA fragments (c-circles) that are exclusive to ALT cells [25]. However, reverse transcription RT-qPCR quantification of chromosome end-specific TERRA could be an excellent complement to the c-circle assay.

The TERRA ONT-seq results served as the launching pad of this project. With 7q and 12q chromosome ends suggested as significant players in TERRA expression, the first goal of this project was to validate TERRA quantification as a new diagnosis tool by measuring and comparing the levels of 7q- and 12q-TERRA in one ALT and one telomerase-positive cancer cell line. The working hypothesis was that, for this diagnosis tool to be viable, 7q_12q TERRA levels in ALT cells should be at least 50 to 100-fold higher than in telomerase-positive cell lines. This would allow for the establishment of a diagnosis threshold that ensures reliability and sensitivity, with little risk of false positives or negatives [26]. The second goal of this project moved beyond the focus of quantifying TERRA expression with a potential diagnostic goal, into a more extensive validation of the ONT-seq results through RT-qPCR. A panel of 4 cell lines was studied, the same cell lines used in the ONT-seq experiment, ALT U2OS and SAOS2 osteosarcoma and telomerase-positive HeLa cervical cancer and HEK embryonic kidney cells. Finally, the third goal was the identification and characterization of a new TERRA promoter sequence present within the CpG islands of 7q and 12q telomere ends.

2. Results

2.1. Validation of 7q_12q Quantification as Diagnostic Tool

TERRA from the long arms of chromosomes 7 and 12 (7q_12q TERRA) was quantified by RT-qPCR and compared between a human ALT osteosarcoma cell line (U2OS) and a human telomerase-positive osteosarcoma cell line (HOS). Absolute and relative quantification of TERRA was performed in U2OS and HOS cells in triplicate. Two distinct qPCR primer sets (primer pair 1 and 2) were used in qPCR to target two distinct amplicons downstream of the TERRA TSS.

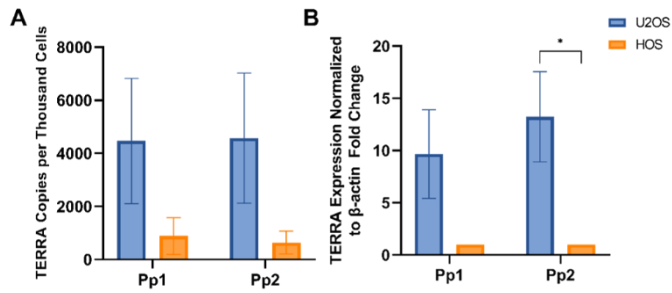


Figure 1: qPCR quantification of TERRA from chromosomes 7q and 12q in U2OS and HOS cells. Reverse transcription performed with primer o1 (C-rich RT primer from Table 2). (A) Absolute quantification. TERRA Copies per Thousand Cells displayed on y axis. (B) Relative quantification. All results normalized to β -actin. Fold Change in TERRA expression normalized to β -actin displayed on y axis. Bars and error bars represent means and SDs from three independent experiments. *P* values were calculated with a two-tailed student's *t*-test. * $p \leq 0.05$

Considering the absolute quantification (Figure 1A), U2OS cells expressed an average of about 4500 7q_12q TERRA copies per thousand cells and telomerase positive HOS cells expressed about 900. The results were similar for both sets of qPCR primers. While 7q_12q TERRA expression is clearly higher in U2OS than HOS cells, the fold increase in U2OS relative to HOS is less than ten-fold. The variability between biological replicates is also high in both U2OS and HOS. As for the relative quantification (Figure 1B) - which was performed by normalizing TERRA to the β -actin housekeeping gene - a 10-fold increase in 7q_12q TERRA is observed in U2OS relative to HOS, using primer pair 1, and about 13-fold increase using primer pair 2. The variability between biological replicates in both U2OS and HOS remains quite elevated.

A possible source of bias in these results comes from the choice of primer for cDNA synthesis during the reverse transcription step. The reverse primer was designed to be complementary to the TERRA (5'-UUAGGG-3') sequence derived from the telomeric 5'-(CCCTAA)-3' repeats (primer o1 from Table 2), which were assumed to be identical in both U2OS and HOS cell lines. However, it has been reported that ALT cells often have degenerated repeats at their telomeres [27],[28]. Thus, reverse transcription from U2OS TERRA could be less efficient when compared to TERRA from HOS. Such difference in reverse transcription efficiency could bias the final qPCR results, leading to an underestimation of U2OS TERRA expression. To test this, the RT-qPCR was repeated in identical conditions but using a reverse transcription primer that annealed to the subtelomeric part of the 7q_12q TERRA transcripts (primer o2 from Table 2), right upstream of to the telomeric repeats, assuring that cDNA still covers the entire qPCR target region for amplification. The results of the second set of RT-qPCRs are shown in Figure 2.

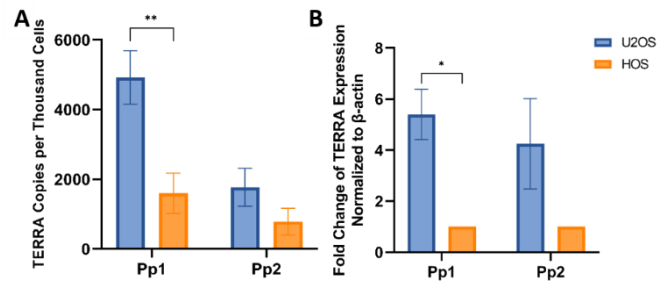


Figure 2: qPCR quantification of TERRA from chromosomes 7q and 12q in U2OS and HOS cells. Reverse transcription performed with primer o2 (7q_12q subtelomeric RT primer from Table 2). (A) Absolute quantification. TERRA Copies per Thousand Cells displayed on y axis. (B) Relative quantification. All results normalized to β -actin. Fold Change in TERRA expression normalized to β -actin displayed on y axis. Bars and error bars represent means and SDs from three independent experiments. *P* values were calculated with a two-tailed student's *t*-test. * $p \leq 0.05$, ** $p \leq 0.01$

Considering the absolute quantification (Figure 2A), HOS contain about 1600 7q_12q TERRA copies per thousand cells with primer pair 1, and 800 copies with primer pair 2. The results are similar to those using the C-rich primer o1 in Figure 1A (600 copies per thousand cells), indicating that the two primers prime reverse transcription with similar efficiencies for HOS 7q_12q TERRA. 7q_12q U2OS TERRA expression was also similar for primer pair 1 (about 5000 copies per thousand cells), when compared to the results using the C-rich RT primer (4500 copies per thousand cells). For primer pair 2, a small decrease occurred, with TERRA levels being around 1800 copies per thousand cells. The relative quantification results (Figure 2B), also displayed a lower 7q_12q TERRA expression fold using the 7q_12q subtelomeric RT primer, when compared to the C-rich RT primer. A 5-fold 7q_12q TERRA increase was observed for primer pair 1 and 4-fold increase for primer pair 2. Overall, very similar 7q_12q TERRA levels were observed using cDNA reverse transcribed using the C-rich telomeric RT primer compared to the 7q_12q subtelomeric derived reverse primer. This suggests that reverse transcription efficiency is not affected by primer choice. Therefore, either there is no significant degeneration of U2OS telomeric repeats, or, if there is, the RT can still efficiently reverse transcribe TERRA from these telomeres. The highest 7q_12q TERRA fold-increase in U2OS when compared to HOS was of 13-fold. This is much lower than the 50-100-fold minimum that has hypothesized to be necessary for establishment of a diagnosis threshold. This, combined with the high variability between biological replicates in both cell lines, indicates that it might not be possible to establish a threshold that could confidently segregate ALT and non-ALT cancers based on 7q_12q TERRA expression. Therefore, 7q_12q TERRA quantification may not be viable as an ALT cancer diagnosis tool.

The high variability in detected 7q_12q TERRA transcripts between samples can be explained by two factors. The first is the possibility of naturally high fluctuations of TERRA in ALT cells. TERRA has been shown to increase in response to DNA damage induced by TRF2 depletion, likely in order to help sustain the telomeric DNA damage response [29]. DNA damage levels fluctuates under action of a complex network of activators and repressor, and TERRA may fluctuate in response to these alterations. It is plausible that moderate

TERRA fluctuations (both total and 7q_12q specific) can occur in ALT populations in response to even minor stresses that the population is exposed to. Another reason for variability between samples are technical factors in the RT-qPCR experiment. The reverse transcription process of each of the RNA is assumed to have identical efficiency for quantification purposes, but minor changes in RT efficiency from sample to sample might occur, which can add to a significant error when dealing with an RNA expressed at low levels, such as TERRA. Therefore, this RT-qPCR assay is difficult to fully standardize and is expected to have high variability between biological replicates.

It is worth noting that quantification of total TERRA, rather than chromosome end specific TERRA, could potentially be a viable diagnosis tool. This could be achieved by RT-qPCR combined with rolling circle amplification (RCA) using padlock probes [30]. A padlock probe is an oligonucleotide that, via a ligation reaction, forms circular DNA when hybridizing to specific target RNA. The region of the padlock probe that does not participate in target RNA hybridization contains generic primer sequences for amplification. Padlock probes can be designed to target telomeric repeats in TERRA. Circularized padlock probes can be amplified by RCA, and then, real time qPCR is performed, allowing for quantitative assessment of total telomeric repeats in the RNA sample. This could result in accurate estimation of TERRA transcripts. It is possible that, unlike in the case of chromosome end-specific TERRA, the fold-difference in total TERRA repeats could be as high as 50-to-100-fold, and allow for establishment of a diagnosis threshold.

2.2. Validation of ONT-seq by RT-qPCR

The second goal of this project was a broader validation of the ONT-seq results in a panel of 4 cell lines: U2OS and SAOS2 ALT cells, and HeLa and HEK293T telomerase-positive cells. Five chromosome ends were chosen for validation, 5p, 7q_12q, 15q, 16p and 20q. These 5 ends represent a range of TERRA expression, from high (7q and 12q) to moderate (5p and 16p) to low (15q and 20q). The ONT-seq results for these chromosome ends are shown in Figure 3.

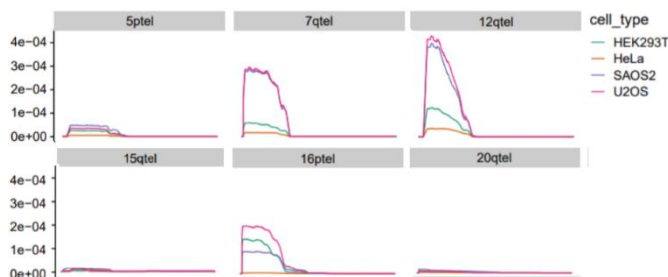


Figure 3: Oxford ONT-seq of TERRA expression for a selection of chromosome ends. x axis displays the 2000 bp after the telomeric repeats in each subtelomere, with numbering beginning from 0 immediately after the telomeric repeats. y axis is TERRA long-read coverage, on a log scale.

TERRA levels were measured by RT-qPCR by absolute quantification. The RT-qPCR validation results are shown in Figure 4.

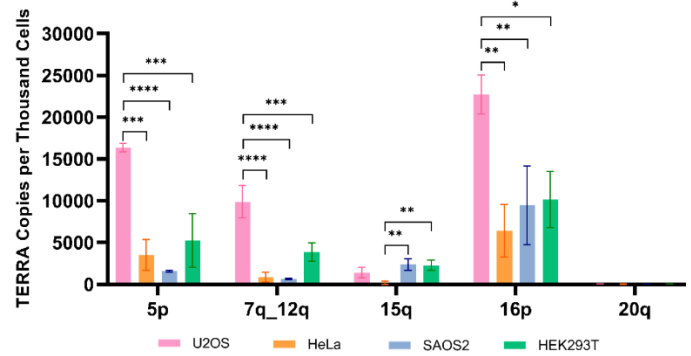


Figure 4: qPCR quantification of TERRA from chromosomes 5p, 7q, 12q, 15q, 16p and 20q in U2OS, HeLa, SAOS-2 and HEK293T cells. Absolute quantification. TERRA Copies per Thousand Cells displayed on y axis. Bars and error bars represent means and SDs from three independent experiments. P values were calculated with a one-way ANOVA test. * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$, **** $P \leq 0.0001$

Considering the RT-qPCR results (Figure 4), chromosome ends 5p and 16p showed higher TERRA expression across all cell lines, when compared to the ONT-seq results, even surpassing 7q_12q TERRA levels. However, the standard curves obtained from these two chromosome ends were less linear, with a low R^2 value (0.93 for 5p and 0.9 for 16p). These curves were also not representative of the 1:10 dilutions made. A 10-fold difference in concentration should result in a 3.3 CT value difference between curves. However, the points in the 5p and 16p curves were, at times, separated by 5 CTs. Therefore, the results for these two chromosome ends are less reliable and likely biased these results.

As for the remaining chromosome ends, 7q_12q, 15q and 20q, the results are similar to the ONT-seq with regard to U2OS, HEK2937 and HeLa. However, the outlier was SAOS2 osteosarcoma cells, which, despite being an ALT cell line, had much lower TERRA expression than predicted by ONT-seq, particularly in 7q_12q ends. SAOS2 TERRA is also underexpressed in 5p and 16p compared to U2OS. Northern blot analyses of total TERRA performed in Azzalin lab have also showed that SAOS2 have considerably higher overall levels of TERRA than HeLa and HEK293T, although lower than U2OS. Several possibilities for this discrepancy were tested and ruled out. Primer annealing efficiency was proved to be similar in all cell lines via a control qPCR with known amounts of genomic DNA. Telomeric repeat degeneration in SAOS2 was ruled out because the results of the RT-qPCR were similar when using cDNA from a reverse transcription made with a subtelomeric primer. And secondary structures in the SAOS2 sequence biasing reverse transcription efficiency between the 2 assays were ruled out by using matching temperatures for the reverse transcription reaction in RT-qPCR and ONT-seq.

It is possible that this SAOS2 cell population simply suffered changes in DNA damage levels that resulted in a downregulating of TERRA expression. It is also possible that there are limitation in the ONT-seq results. Before sequencing, an RNA purification process was performed to produce TERRA-enriched RNA. A TERRA-enriched sample is less complex than a sample of total RNA and less complex samples will likely lead to a more efficient cDNA synthesis, because there is less interference from other molecules. In the ONT-seq

experiment, TERRA was also purified from nuclear RNA, while in the RT-qPCR, it was purified from total RNA. These differences in the ONT-seq process could have resulted in a bias facilitating detection of 7q_12q TERRA in SAOS2 cells. Alternatively, the normalization and smoothing process during bioinformatics analysis of the sequencing data could have caused a bias. In depth troubleshooting of ONT-seq data is beyond the scope of this thesis. However, further work is necessary to explain the discrepancy between the two approaches.

There is, however, a series of conclusions that can definitively be made from this set of RT-qPCR assays. First, TERRA was confirmed to be expressed from multiple chromosome ends in all analyzed human cell lines. This falls in line with the popular TERRA transcription model that postulates the existence of TERRA loci at multiple chromosome ends. Moreover, the 7q_12q conserved sequence was confirmed to be a main source of TERRA transcripts. These chromosome ends do not contain the previously characterized 61-29-37 bp promoters, that guide active TERRA transcription from other chromosome ends. This led to the establishment of the hypothesis that a previously uncharacterized TERRA promoter in this conserved sequence might be actively transcribing TERRA *in vivo*. Thus, the next part of this project consisted in the characterization of the 7q and 12q sequence promoter. Furthermore, the effects of DNA methylation on 7q_12q TERRA transcription was also studied.

2.3. 7q_12q TERRA Transcription Regulation by DNA Methylation

The first step taken to study active TERRA transcription by 7q and 12q subtelomeric CpG island putative promoters was to study the impact of DNA methylation on transcription activity. ALT cell subtelomeric CpG islands have been shown to be hypomethylated [20], and increased transcription of TERRA in ALT cells is believed to be a result of this hypomethylation [16]. Nergadze et al. analyzed the methylation state of 61-29-37 repeats in human HCT116 cells knocked-out for DNMT1 and DNMT3b (double KO [DKO]) DNA methyltransferases [20]. Concomitant disruption of DNMT1 and DNMT3b genes was shown to completely abolish methylation at 61-29-37 repeat CpG dinucleotides, implying that DNMT1 and DNMT3b enzymes cooperatively maintain DNA methylation at 61-29-37 repeats in HCT116 cells. This decreased methylation state resulted in TERRA upregulation in 61-29-37 bp containing subtelomeres in HCT116 DKO cells [20]. In this way, DNA methylation has been shown to be a crucial suppressor of TERRA expression from the previously characterized 61-29-37 repeats TERRA promoters.

For this study, the hypothesis was that a decreased methylation state could have the same impact on the 7q_12q putative promoter sequence. As such, HCT116 cells and HCT116 DKO cells were chosen to study this effect at the 7q and 12q chromosome ends. RT-qPCR was performed with 7q_12q primers on the parental telomerase-positive cancer cell line HCT116 and HCT116 DKO cells. Absolute (Figure 5A) and relative (Figure 5B) quantifications were performed. Both analyses showed that DKO hypomethylated cells express more TERRA from 7q and 12q telomeres relative to the parental cells. However, only a very minor increase, about 2-fold in absolute

quantification and 1.5-fold in relative quantification, was observed, which is lower than what observed in previous studies with these cell lines [20], [21]. For instance, Lingner et al. reported a 20-fold increase in DKO TERRA expression relative to HCT116 at 12q chromosome end normalized to GAPDH.

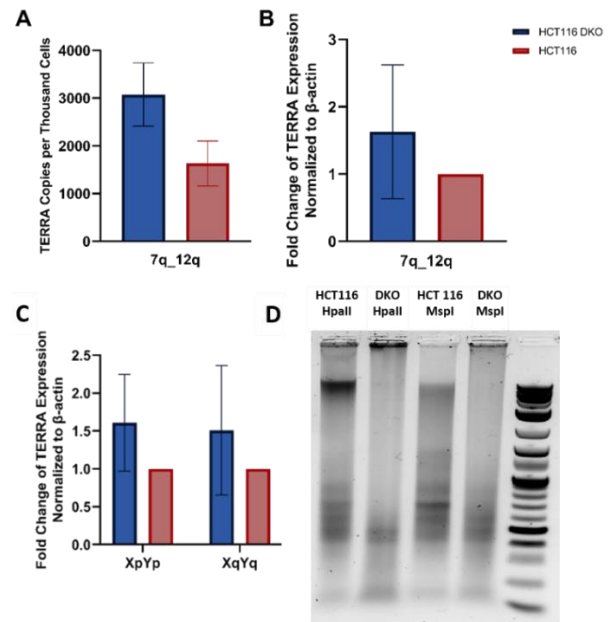


Figure 5: qPCR quantification of TERRA from chromosomes 7q and 12q in HCT116 and HCT116 DKO cells. Reverse transcription performed with primer o1 (C-rich RT primer from Table 2). (A) Absolute quantification. TERRA Copies per Thousand Cells displayed on y axis. (B) Relative quantification. All results normalized to β -actin. Fold Change in TERRA expression normalized to β -actin displayed on y axis. (C) qPCR quantification of TERRA from chromosomes XpYp and XqYq in HCT116 and HCT116 DKO cells. Reverse transcription performed with primer o1 (C-rich RT primer from Table 2). Relative quantification. All results normalized to β -actin. Fold Change in TERRA expression normalized to β -actin displayed on y axis. Bars and error bars represent means and SDs from three independent experiments. P values were calculated with a two-tailed student's t-test. (D) 0.8% agarose gel run of HCT116 and HCT116 DKO gDNA digested with methylation sensitive HpaII enzyme or methylation insensitive isoschizomer MspI

To validate the cell lines, an experiment was performed to verify the high methylation state of HCT116 cells and demethylated state of HCT116 DKO cells (Figure 5D). Genomic DNA from both cell lines was digested with the methylation-sensitive HpaII restriction enzyme or with its methylation-insensitive isoschizomer MspI, and the digested DNA was run in an agarose gel. HpaII digested HCT116 DNA displays a prominent genomic DNA band at high molecular weights, while HpaII digested DKO DNA displays only a smear. Both MspI digested controls display a smear since MspI digests both methylated and non-methylated DNA. This confirms that HCT116 cell DNA is methylated *in vivo*, but not DKO cell DNA.

The 7q_12q result was controlled by performing RT-qPCR for TERRA expression on a chromosome end known to have a CpG island with TERRA 61-29-37 bp repeats promoter, positive control XqYq, and one known to not contain a CpG island, negative control XpYp (Figure 5C). However, also for these

chromosome ends, TERRA fold-increase between HCT116 and DKO cells was very minor.

The XpYp and XqYq controls suggest that there is an unforeseen irregularity in these HCT116 and DKO populations and definitive conclusions cannot be taken from the RT-qPCR result. This assay should be repeated with new populations of a wild-type cell line and cell lines knocked out for DNA methyltransferases. In any case, DNA methylation regulation of TERRA expression in 7q and 12q chromosome ends still requires further investigating.

2.4. 7q and 12q Putative Promoter Isolation

The next objective of the 7q and 12q putative promoter study was to isolate the 7q and 12q subtelomeric sequences for functional assays to be performed. Bioinformatics analysis was performed using the CpG search function of the MethPrimer 2.0 tool (Figure 6A and Figure 7A) [31] and the CpGPlot/CpGReport at the European molecular biology open software suite program (Figure 6B and Figure 7B) [32]. The analysis predicted the existence of a 332 bp CpG island in the 7q subtelomere 340 bp away from the first telomeric repeats, and a 249 bp CpG island in 12q located 362 bp away from the first telomeric repeat (Figure 8).

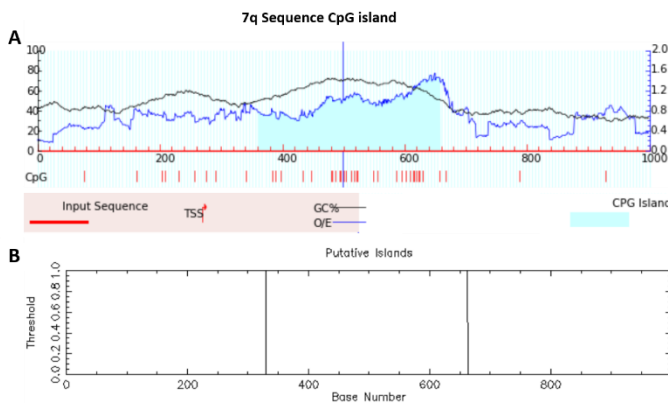


Figure 6: First 1000 bp of the 7q subtelomere immediately following the telomeric repeats analyzed using two bioinformatics tools (A) CpG function of the MethPrimer 2.0 tool. Putative CpG island marked in blue. Predicted TSSs indicated as red vertical bars. CpG island defined as sequence over 100bp long with >50% GC content and >0.6 Observed/Expected Ratio (B) CpGPlot/CpGReport at the European molecular biology open software suite program. Putative CpG island indicated between two vertical black bars. CpG island defined as sequence over 200bp long with >50% GC content and >0.6 Observed/Expected Ratio

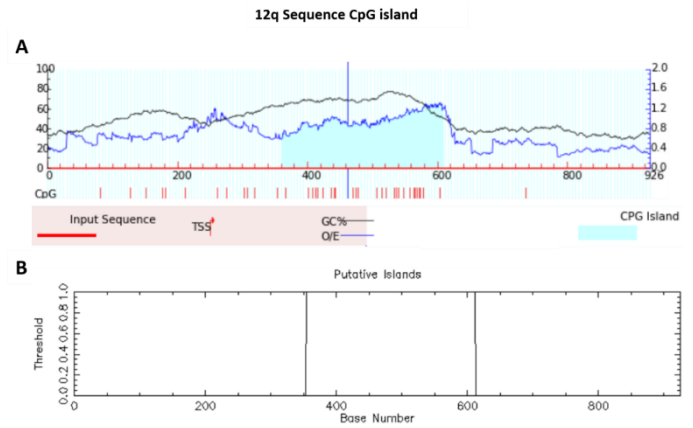


Figure 7: First 1000 bp of the 12q subtelomere immediately following the telomeric repeats analyzed using two bioinformatics tools (A) CpG function of the MethPrimer 2.0 tool. Putative CpG island marked in blue. Predicted TSSs indicated as red vertical bars. CpG island defined as sequence over 100bp long with >50% GC content and >0.6 Observed/Expected Ratio (B) CpGPlot/CpGReport at the European molecular biology open software suite program. Putative CpG island indicated between two vertical black bars. CpG island defined as sequence over 200bp long with >50% GC content and >0.6 Observed/Expected Ratio

To isolate each sequence, a 697 bp 7q amplicon containing the 332 bp CpG island and a 912 bp 12q amplicon containing the 249bp CpG island were amplified by PCR using genomic DNA from U2OS and HeLa cells. An array of combinations of PCR buffers and primer annealing temperatures were tested in order to find the ideal parameters for amplification. Phusion GC Buffer, which is optimized to improve performance of Phusion DNA Polymerase on GC-rich templates, or templates with complex secondary structures, proved to be the most efficient buffer. Furthermore, 70 °C was the only annealing temperature that resulted in amplification, likely because it is high enough to resolve secondary structures yet below the primer annealing temperature.

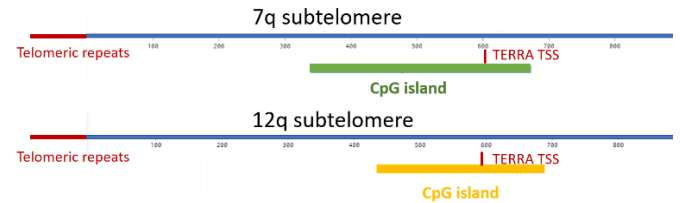


Figure 8: Representation of a portion of the 7q and 12q subtelomeric region. Subtelomeric region indicated in blue and telomeric region indicated in red. 7q CpG island indicated in green, and 12q CpG island indicated in yellow. TERRA TSSs predicted by ONT-seq indicated as red vertical bars.

7q and 12q CpG island-containing sequences were amplified from each cell line. A 7q amplicon from HeLa and a 12q amplicon from U2OS were chosen for cloning into the pjet1.2 plasmid. Sequences were confirmed by Sanger sequencing (Eurofin Genomics). There is a currently ongoing assay being performed in the Azzalin lab to clone both sequences into a reporter plasmid upstream of an eGFP cDNA. This will allow to test for sequence promoter activity. The new data gained on TERRA CpG islands at 7q and 12q subtelomeres can also serve as a foundation for a variety of future assays, such as a knockout of the CpG sequence in 7q and/or 12q and

subsequent analysis of telomere length and DNA damage response.

2.5. 7q_12q Depletion by CRISPRi Assay

The last attempt to validate the 7q and 12q subtelomeric sequences as bona fide TERRA promoters consisted in the silencing of these regions through a sgRNA-guided CRISPR-interference (CRISPRi) system. The CRISPRi system can modulate gene expression without altering the target DNA sequences. It consists of a catalytically inactive Cas9 (dCas9) fused with a transcriptional repressor domain known as the Kruppel-associated box (KRAB), which silences the target loci [33]. dCas9 strongly binds to the DNA target sequence and this tight binding interferes with the activity of transcription factors and RNAPII. KRAB works by recruiting additional co-repressor proteins such as KRAB-box-associated protein-1 (KAP-1) and epigenetic readers such as heterochromatin protein 1 (HP1), thus leading to epigenetic reprogramming of histone modifications and repressing RNAPII [34]. Fusing the KRAB repressor to dCas9 results in a stronger and more specific gene repressor than dCas9 alone [34].

U2OS cells were transfected with plasmids containing dCas9-KRAB plasmid and a sgRNA complementary to 7q (CRISPRi_7q U2OS) or 12q (CRISPRi_12q U2OS) sequences upstream of the TERRA TSS, or with an empty dCas9-KRAB vector (CRISPRi-EV U2OS). The results of the RT-qPCR with RNA extracted from the three U2OS populations can be seen in Figure 9. If the targeted 7q and 12q subtelomeric sequences were indeed TERRA promoters one should observe a decrease in TERRA transcribed from these telomeres upon CRISPRi-derived silencing.

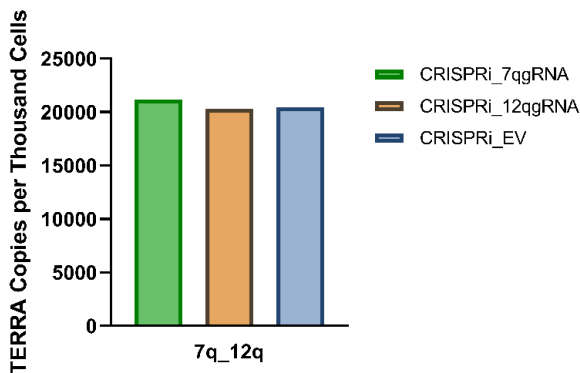


Figure 9: qPCR quantification of TERRA from chromosomes 7q and 12q in CRISPRi_7q, CRISPRi_12q and CRISPRi_EV U2OS cells. Reverse transcription performed with primer o1 (C-rich RT primer from Table 2). Absolute quantification. TERRA Copies per Thousand Cells displayed on y axis.

Approximately the same TERRA levels were measured in CRISPRi_7q and CRISPRi_12q cells as in the negative control CRISPRi-EV U2OS. It is possible that this is evidence for low activity of this CpG sequence as a promoter in vivo. However, that does not match the ONT-seq results showing TERRA transcript coverage in this region, nor the extensive RT-qPCR validation. It is more likely that there are limitations in the CRISPRi silencing protocol. This CRISPRi system and protocol have been used successfully for gene inactivation in the Azzalin lab before. However, in this experiment the designed sgRNAs

annealed to DNA sequences approximately 50 bp (for 7q) and 30 bp (for 12q) upstream of the TERRA TSS suggested by the ONT-seq analysis. sgRNA that annealed on top of or immediately downstream of the TSS are the preferred ideal in these assays. However, due to the requirements of gRNA design it was not possible to design sgRNAs that annealed in the ideal locations and had high efficacy and specificity scores. It is possible that the dCas9 protein is interacting with the DNA too far upstream from where TERRA transcription begins, and thus not efficiently interfering with TERRA transcription.

3. Conclusion

This project was designed to better our understanding of TERRA transcription profiles in human cancer cells. Through extensive and accurate RT-qPCR measurement of TERRA, validating previous long-read RNA-seq results, TERRA transcription was studied in a panel of ALT and telomerase-positive cells at multiple chromosome ends. TERRA was confirmed to be expressed from multiple chromosome ends in all analyzed cell lines, supporting the notion that TERRA originates from several telomeres. Chromosome-specific TERRA quantification was postulated to become a potential ALT diagnosis procedure. However, expression was shown to be prone to fluctuations and the fold-increase in ALT cells is likely not sufficient for a robust diagnosis method to be established. Elevated expression of TERRA was found in U2OS ALT cells at 7q and 12q chromosome ends, leading to the establishment of the hypothesis that these ends contain a previously uncharacterized promoter actively transcribing TERRA. The subtelomeric CpG island sequence containing a TERRA transcription start site, which is conserved between the two ends, was isolated. The first steps were taken to characterize its promoter activity and further assays will be performed in the Azzalin lab with this goal. This project helped gain new insights into the profile of TERRA transcription in ALT cells and helped guide an ongoing search for further sites of active TERRA transcription. The identified TERRA loci could represent potential targets for TERRA-focused approaches. Research into this area is of high importance, since ALT cancers reliant on TERRA transcription, currently affect millions of people, have generally poor prognosis and are resistant to available treatment protocols.

4. Materials and methods

4.1. Cell Lines and Culture Conditions

ALT cell lines used in this project were U2OS and SAOS2 osteosarcoma cells. Telomerase-positive cell lines used were HOS osteosarcoma, HeLa cervical cancer, HEK293T embryonic kidney and HCT116 and HCT 116 DKO colorectal carcinoma cells. U2OS, HeLa and HEK cells were cultured in high glucose DMEM, GlutaMAX (Thermo Fisher Scientific) supplemented with 5% tetracycline-free fetal bovine serum (Pan BioTech) and 100 U/ml penicillin-streptomycin (Thermo Fisher Scientific). HCT116 and HCT116 DKO cells were cultured in the same conditions except supplemented with 10% fetal bovine serum. SAOS2 and HOS cells were cultured in high glucose DMEM/F12, GlutaMAX (Thermo Fisher Scientific), supplemented with 5% tetracycline-free fetal bovine serum (Pan BioTech), 100 U/ml penicillin-streptomycin (Thermo Fisher Scientific) and nonessential amino acids

(Thermo Fisher Scientific). The cells were maintained in 5% CO₂ incubator at 37 °C.

4.2. RNA Preparation

RNA from 4 to 5 million cultured cells was isolated using NucleoSpin RNA (Macherey-Nagel). Two in-solution DNase digestions were performed to remove any trace of genomic DNA contaminants. RNAs were re-hydrated in 82 µL RNase-free H₂O. RNA concentrations were measured on a NanoDrop spectrophotometer. RNA was stored at -80 °C.

4.3. Genomic DNA Preparation

gDNA was isolated using Wizard® Genomic DNA Purification Kit (Promega). DNAs were rehydrated in H₂O by incubation at 65°C for 1 hour, or overnight at 4°C. DNA concentrations were measured on a NanoDrop spectrophotometer DNA was stored at 4°C.

4.4. RT-qPCR

TERRA RT-qPCR was performed as previously described [35]. Briefly, 3 µg of RNA was reverse transcribed using 200 U SuperScript IV Reverse transcriptase (Thermo Fischer Scientific), β-actin and TERRA reverse primers. Reverse transcription was performed at 50°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 S1 for sequences). qPCR consisted of 10 minutes at 95°C followed by 45 cycles at 95°C for 15s and 60°C for 30s in a Rotor Gene 6000 System (Corbett). Primers were designed to anneal to the subtelomeric DNA sequence of a chromosome end proximal to the telomeric repeats and amplify a 90-150 bp target downstream of the TERRA TSS. Primers were designed with the online software primer3. Primer melting temperatures (T_m) were in the 64-69 °C range. Primers were confirmed not to form cross primer dimers, using Multiple Primer Analyzer (ThermoFisher Scientific). Primer efficiencies were tested by performed qPCR on a 5-point dilution series of gDNA (from 10 ng/µl to 0,001 ng/µl). Average CTs values for each dilution were calculated. The log value of the starting quantity was calculated. The slope of the regression curve between the log values and the average CT values was calculated in Microsoft Excel. Primer Efficiency was calculated using the following formula.

$$E = \left(10^{\frac{-1}{\text{slope}}} - 1 \right) \times 100$$

4.5. TERRA Absolute Quantification

To measure the number of TERRA-derived cDNA molecules in the cDNA sample, absolute quantification was performed. For this, a standard curve was made with a standard in known concentrations. Standards were designed to be identical to 5 distinct 300-500 bp subtelomeric regions (5p, 7q_12q, 15q, 16p, 20q), as close as possible to the 5'-(CCCTAA)-3' repeats and downstream of the TERRA TSS predicted by ONT-seq. Thus, each 300-500 bp standard contained the 90-150 bp fragment targeted for amplification by qPCR and is identical to the cDNA obtained from TERRA transcribed in that chromosome end. Standards were ordered as gBlocks Gene Fragments from Integrated DNA Technologies. The gBlock

sequences were cloned into a pJET1.2/Blunt plasmid (Thermo Scientific pJET Cloning Kit) and verified by Sanger Sequencing (Lighrun Tube sent to Eurofin Genomics). The plasmid concentrations were measured by Nanodrop. The plasmids containing the gBlocks were used to create a 5-point standard curve. The number of standard copies in each dilution was calculated using the following formula:

$$\text{Standard copy number} = \frac{6.02 \times 10^{23} (\text{copies mol}^{-1}) \times \text{DNA amount (g)}}{\text{DNA length (bp)} \times 660 (\text{g mol}^{-1} \text{bp}^{-1})}$$

The standard dilution series was amplified by qPCR at the same time as TERRA cDNA. The average CT values of the standard dilutions were determined. A regression curve was made between the log values of the copy numbers in the standard dilutions and the average CT values standard dilutions was calculated in Microsoft Excel. Finally, the CTs of the target samples were plotted on to this standard curve to obtain TERRA copy numbers To calculate the copy number per cell, cells were counted before RNA extraction and the total RNA yielded from the extraction was quantified with Nanodrop. Thus, the ratio of RNA yielded per thousand cells, for each cell line, was calculated. Known RNA mass (4 to 5 ug) was used in the RT reaction and RT reaction was assumed to be 100% efficient, allowing for estimation of the number of cells corresponding to the nanograms of cDNA present in one qPCR tube. Finally, copy number per thousand cells was calculated with the following formula:

$$\text{Copies per thousand cells} = \frac{\text{TERRA Molecules per qPCR reaction}}{\text{Thousands of cells corresponding to ngs of cDNA in qPCR}}$$

4.6. DNA Sequence Analysis

To analyze CpG dinucleotide contents and predict CpG islands, the CpGPlot/CpGReport at the European molecular biology open software suite program (EMBOSS) (https://www.ebi.ac.uk/Tools/seqstats/emboss_cpplot/) was utilized. The results were validated using the CpG function of the MethPrimer 2.0 tool at Li Lab (Peking Union Medical College Hospital) (<http://www.urogene.org/cgi-bin/methprimer2/MethPrimer.cgi>)

4.7. DNA Methylation Analysis

Genomic DNA was extracted from HCT116 and HCT116 DKO cells. 5 ug of DNA from both cell lines was digested in 40 µl total volume, with 1 µl (or 20 U) of HpaII (CpG methylation-sensitive enzyme), in 1X Cutsmart Buffer, overnight. Another 5 ug of DNA from both cell lines was digested in identical conditions with MspI (CpG methylation-insensitive enzyme). The digested DNA was run on a 0,8% agarose gel and an image was taken with ChemiDoc XRS+ System (Bio-Rad).

4.8. 7q and 12q Subtelomeric Sequence Isolation by PCR

Genomic DNA was extracted from U2OS, HeLa, HEK and SAOS2. Using primer3 a single forward primer and two distinct reverse primers were designed to amplify the 7q subtelomeric CpG island sequence and the 12q subtelomeric sequence. Using these two primer pairs sharing the same forward primer, two partially overlapping subtelomeric sequences were targeted for amplification, one being approximately 200 bp longer than the other. The two sets of primer pairs did not allow to distinguish between 7q and 12q sequences, due to the high

homology between the sequences. Instead, 7q and 12q sequences were distinguished from each other after amplification by the presence of a 30 bp sequence in the 12q sequence that is not present in the 7q sequence. The sequences were amplified from 50 ng of template DNA with 0,5 µl of Phusion Polymerase Enzyme (2 U/ml) and 10 µl of GC Buffer (5X). PCR reactions consisted of 98 °C for 30 s (1 cycle), 98 °C for 10/15 s (30-45 cycles), 72 °C for 20 sec (30-45 cycles), 68/70/72 °C for 54 s (30-45 cycles) and 72 °C for 5 min (1 cycle). The amplified sequences were cloned into a pJET1.2/Blunt plasmid (Thermo Scientific pJET Cloning Kit) and verified by Sanger Sequencing (Lightrun Tube sent to Eurofin Genomics). The plasmid concentrations were measured by Nanodrop.

4.9. CRISPRi Assay

20 bp sgRNAs were designed with CRISPR-ERA webtool (<http://crispr-era.stanford.edu/>). For each of the 7q and 12q sequences, a pairs of primers was designed that annealed with each other, forming a duplex, while leaving 4 basepair long single stranded ends on each side. These ends are complementary to BsaI cuts that result from digestion of the CRISPRi empty vector. sgRNAs were annealed to each other by adding, in a 1.5 ml Eppendorf, 1 µl forward sgRNA (0,5 µg/µl), 1 µl reverse gRNA and 48 µl TE (pH=8.0). This reaction mixture was incubated at 95 °C for 5 min and then cooled slowly to room temperature. The backbone plasmid to be used for CRISPRi depletion strategy was gifted to Azzalin lab by Prof. Ulrike Kutay. It already contained the deactivated Cas9 protein (dCas9), KRAB repressor and sgRNA scaffold. 2 µg of

backbone plasmid were digested with BsaI-HF enzyme in a 20 µl total volume. The linearized plasmid was purified using Zymoclean Gel DNA Recovery Kit. Plasmid was eluted in 20 µl dd H₂O. Duplex sgRNAs were cloned into the backbone dCas9 plasmid between BsaI sites, in Stbl3 E. Coli. Plasmids were transfected into U2OS cells using Lipofectamine RNAiMAX (Invitrogen). 6 µg of plasmid and 15 µl of Lipofectamine were separately diluted in Opti-MEM (Thermo Fisher Scientific). The two solutions were added together, briefly mixed by pipetting and incubated at room temperature for 10 min. These mixtures were then added to 10cm plate cell cultures. Medium was changed 5 h after adding the transfection mixture. Puromycin was added to a final concentration of 1 µg/mL, 24 h after selection. 48 h after selection, RNA was isolated from the cells for RT-qPCR.

4.10. Statistical Analysis

Two-tailed Student's t-test was performed in Microsoft Excel for a direct comparison of two groups. One-way analysis of variance (one-way ANOVA) was employed in GraphPad Prism 8 for comparison of more than two groups, followed by Tukey's HSD treatment to adjust the pairwise statistical significance. *P* values are indicated as: **P* ≤ 0.05, ***P* ≤ 0.01, ****P* ≤ 0.001. Error bars represent SD.

Annex

Table 1: oligos used for PCR amplification of 7q and 12q CpG islands

Oligo	Sequence (5' -> 3')
7q_12q_CpG_island_FW1	ATGAGCAATGGGGTGTGTTATATT TCGGTGTCAT
7q_12q_CpG_island_RV1	GGGTGGCATATTTTGGTCTTATACA CTGTGTTCCA
7q_12q_CpG_island_RV2	ATGTGCACAGAGGTAGCTTTTAAAT GGGGATTCC

Table 2: Reverse Transcription Primers used in this study

Number	Primer	Sequence (5' -> 3')
o1	C-rich RT primer	CCCTAACCCCTAACCCCTAACCCCTAACCCCTAA
o2	7q_12q subtelomeric RT primer	GCACGGCAGTGTGCGGATTG

Table 3: qPCR Primers chosen for use in this study following primer efficiency testing

Number	Primer	Sequence (5' -> 3')	T _m (°C)
1	5p Fw	CATTTGTCTTCCGAGGCTGC	67,3
2	5p Rev	CAAAAGCCCCTCTGAATCCTG	67,1
3	7q_12q Fw 1	ATTTGTCTTCCGAGGCTGC	64,7
5	7q_12q Fw 2	GAGCAATGGGGTGTGTCATA	63,9
6	7q_12q Rev	GTTCAGTGTGGAAAACGGGA	65,3
7	15q Fw	CGATGCTGCAACTGGACCC	69,6
8	15q Rev	GCGGTTCACTGCGGAAAAC	68,2
9	16p Fw	CCCGTTTTCCACACTGAACC	66,9
10	16p Rev	CTCCTAAGCACAGACGTTGGG	66,2
11	20q Fw	CAATAATTGAAAAGCCGGGC	68,3
12	20q Rev	GGCCACCACTATAAGCAACAG	63,7

References

- [1] N. Hug and J. Lingner, "Telomere length homeostasis," *Chromosoma*, pp. 413–425, 2006, doi: 10.1007/s00412-006-0067-3.
- [2] A. Prieur and D. S. Peiper, "Cellular senescence in vivo: a barrier to tumorigenesis," *Curr. Opin. Cell Biol.*, vol. 20, no. 2, pp. 150–155, 2008, doi: 10.1016/j.ceb.2008.01.007.
- [3] Y. Deng, S. S. Chan, and S. Chang, "Telomere dysfunction and tumour suppression: The senescence connection," *Nat. Rev. Cancer*, vol. 8, no. 6, pp. 450–458, 2008, doi: 10.1038/nrc2393.
- [4] R. J. O'Sullivan and J. Karlseder, "Telomeres: Protecting chromosomes against genome instability," *Nat. Rev. Mol. Cell Biol.*, vol. 11, no. 3, pp. 171–181, 2010, doi: 10.1038/nrm2848.
- [5] M. Z. Levy, R. C. Allsopp, A. B. Futcher, C. W. Greider, and C. B. Harley, "Telomere end-replication problem and cell aging," *J. Mol. Biol.*, vol. 225, no. 4, pp. 951–960, 1992, doi:

10.1016/0022-2836(92)90096-3.

- [6] A. J. Cesare and R. R. Reddel, "Alternative lengthening of telomeres: Models, mechanisms and implications," *Nat. Rev. Genet.*, vol. 11, no. 5, pp. 319–330, 2010, doi: 10.1038/nrg2763.
- [7] R. Arora, Y. Lee, H. Wischniewski, C. M. Brun, T. Schwarz, and C. M. Azzalin, "RNaseH1 regulates TERRA-telomeric DNA hybrids and telomere maintenance in ALT tumour cells," *Nat. Commun.*, vol. 5, pp. 1–11, 2014, doi: 10.1038/ncomms6220.
- [8] D. Hanahan and R. A. Weinberg, "Hallmarks of cancer: The next generation," *Cell*, vol. 144, no. 5, pp. 646–674, 2011, doi: 10.1016/j.cell.2011.02.013.
- [9] N. W. Kim *et al.*, "Specific Association of Human Telomerase Activity with Immortal Cells and Cancer Published by: American Association for the Advancement of Science" *Science (80-)*, vol. 266, no. 5193, pp. 2011–2015, 1994.
- [10] C. W. Greider, "Telomeres, Telomerase and Senescence," *BioEssays*, no. 1, pp. 66–71, 1990.
- [11] B. P. O. DeRisi Joseph, Penland Lolita, S. Tyagi, F. R. Kramer, N. P. Group, and B. P. O. DeRisi Joseph, Penland Lolita, "Evidence for an Alternative Mechanism for Maintaining Telomere Length in Human Tumors and Tumor-Derived Cell Lines," *Nat. Publishing Gr.*, vol. 4, pp. 303–308, 1996, doi: 10.1038/nm0798-822.
- [12] J. D. Henson, A. A. Neumann, T. R. Yeager, and R. R. Reddel, "Alternative lengthening of telomeres in mammalian cells," *Oncogene*, vol. 21, no. 4 REV. ISS. 1, pp. 598–610, 2002, doi: 10.1038/sj/onc/1205058.
- [13] C. M. Heaphy *et al.*, "Prevalence of the Alternative Lengthening of Telomeres Telomere Maintenance Mechanism in Human Cancer Subtypes," *Short Commun.*, vol. 179, no. 4, pp. 1608–1615, 2011, doi: 10.1016/j.ajpath.2011.06.018.
- [14] J. M. Zhang and L. Zou, "Alternative lengthening of telomeres: From molecular mechanisms to therapeutic outlooks," *Cell Biosci.*, vol. 10, no. 1, pp. 1–9, 2020, doi: 10.1186/s13578-020-00391-6.
- [15] J. Kramara, B. Osia, and A. Malkova, "Break-Induced Replication: The Where, The Why, and The How.," *Trends Genet.*, vol. 34, no. 7, pp. 518–531, Jul. 2018, doi: 10.1016/j.tig.2018.04.002.
- [16] B. Domingues-Silva, B. Silva, and C. M. Azzalin, "Alternative Functions for Human FANCM at Telomeres," *Front. Mol. Biosci.*, vol. 6, no. September, pp. 1–7, 2019, doi: 10.3389/fmolb.2019.00084.
- [17] C. M. Azzalin, P. Reichenbach, L. Khoriant, E. Giulotto, and J. Lingner, "Telomeric Repeat-Containing RNA and RNA Surveillance Factors at," vol. 318, no. November, pp. 798–801, 2007.
- [18] C. M. Azzalin and J. Lingner, "Telomere functions grounding on TERRA firma," *Trends Cell Biol.*, vol. 25, no. 1, pp. 29–36, 2015, doi: 10.1016/j.tcb.2014.08.007.
- [19] A. Aguilera and T. García-Muse, "R Loops: From Transcription Byproducts to Threats to Genome Stability," *Mol. Cell*, vol. 46, no. 2, pp. 115–124, 2012, doi: 10.1016/j.molcel.2012.04.009.
- [20] S. G. Nergadze *et al.*, "CpG-island promoters drive transcription of human telomeres," *Rna*, vol. 15, no. 12, pp. 2186–2194, 2009, doi: 10.1261/rna.1748309.
- [21] M. Feretzaki, P. R. Nunes, and J. Lingner, "Expression and differential regulation of human TERRA at several chromosome ends," *Rna*, vol. 25, no. 11, pp. 1470–1480, 2019, doi: 10.1261/rna.072322.119.
- [22] A. Porro, S. Feuerhahn, J. Delafontaine, H. Riethman, J. Rougemont, and J. Lingner, "Functional characterization of the TERRA transcriptome at damaged telomeres," *Nat. Commun.*, vol. 5, pp. 1–13, 2014, doi: 10.1038/ncomms6379.
- [23] D. Montero, Silanes, O. Gran, and M. A. Blasco, "Telomeric RNAs are essential to maintain telomeres," pp. 1–13, 2016, doi: 10.1038/ncomms12534.
- [24] C. Soneson, Y. Yao, A. Bratus-Neuenschwander, A. Patrignani, M. D. Robinson, and S. Hussain, "A comprehensive examination of Nanopore native RNA sequencing for characterization of complex transcriptomes," *Nat. Commun.*, vol. 10, no. 1, p. 3359, 2019, doi: 10.1038/s41467-019-11272-z.
- [25] J. D. Henson *et al.*, "DNA C-circles are specific and quantifiable markers of alternative-lengthening-of-telomeres activity," *Nat. Biotechnol.*, vol. 27, no. 12, pp. 1181–1185, 2009, doi: 10.1038/nbt.1587.
- [26] A. M. Simundic, "Measures of Diagnostic Accuracy: Basic Definitions," *Ejifcc*, vol. 19, no. 4, pp. 203–211, 2009.
- [27] M. Lee *et al.*, "Telomere extension by telomerase and ALT generates variant repeats by mechanistically distinct processes," *Nucleic Acids Res.*, vol. 42, no. 3, pp. 1733–1746, 2014, doi: 10.1093/nar/gkt1117.
- [28] A. Sommer and N. J. Royle, "ALT: A multi-faceted phenomenon," *Genes (Basel)*, vol. 11, no. 2, pp. 1–13, 2020, doi: 10.3390/genes11020133.
- [29] A. Porro, S. Feuerhahn, and J. Lingner, "TERRA-Reinforced Association of LSD1 with MRE11 Promotes Processing of Uncapped Telomeres," *Cell Rep.*, vol. 6, no. 4, pp. 765–776, 2014, doi: 10.1016/j.celrep.2014.01.022.
- [30] J. Gong, Y. Li, T. Lin, X. Feng, and L. Chu, "Multiplex real-time PCR assay combined with rolling circle amplification (MPRP) using universal primers for non-invasive detection of tumor-related mutations," *RSC Adv.*, vol. 8, no. 48, pp. 27375–27381, 2018, doi: 10.1039/C8RA05259J.
- [31] L. C. Li and R. Dahiya, "MethPrimer: Designing primers for methylation PCRs," *Bioinformatics*, vol. 18, no. 11, pp. 1427–1431, 2002, doi: 10.1093/bioinformatics/18.11.1427.
- [32] P. Rice, L. Longden, and A. Bleasby, "EMBOSS: The European Molecular Biology Open Software Suite," *Trends Genet.*, vol. 16, no. 6, pp. 276–277, 2000, doi: 10.1016/S0168-9525(00)02024-2.
- [33] K. M. Parsi, E. Hennessy, N. Kearns, and R. Maehr, "Using an Inducible CRISPR-dCas9-KRAB Effector System to Dissect Transcriptional Regulation in Human Embryonic Stem Cells," *Methods Mol. Biol.*, vol. 1507, pp. 221–233, 2017, doi: 10.1007/978-1-4939-6518-2_16.
- [34] M. H. Larson, L. A. Gilbert, X. Wang, W. A. Lim, J. S. Weissman, and L. S. Qi, "CRISPR interference (CRISPRi) for sequence-specific control of gene expression," *Nat. Protoc.*, vol. 8, no. 11, pp. 2180–2196, 2013, doi: 10.1038/nprot.2013.132.
- [35] M. Feretzaki and J. Lingner, "A practical qPCR approach to detect TERRA, the elusive telomeric repeat-containing RNA," *Methods*, vol. 114, pp. 39–45, 2017, doi: 10.1016/j.ymeth.2016.08.004.