EPFL

Thèse n° 9544

Regulation and impact of TERRA R-loops at human telomeres

Présentée le 31 mars 2023

Faculté des sciences de la vie
Unité du Prof. Lingner
Programme doctoral en approches moléculaires du vivant

pour l'obtention du grade de Docteur ès Sciences

par

Rita VALADOR FERNANDES

Acceptée sur proposition du jury

Prof. E. Oricchio, présidente du jury
Prof. J. Lingner, directeur de thèse
Prof. S. Smith, rapporteuse
Prof. C. M. Azzalin, rapporteur
Prof. D. Suter, rapporteur

École polytechnique fédérale de Lausanne

2023

Acknowledgments

"The potential people who could have been here in my place but who will in fact never see the light of day outnumber the sand grains of Arabia. Certainly those unborn ghosts include greater poets than Keats, scientists greater than Newton. We know this because the set of possible people allowed by our DNA so massively exceeds the set of actual people. In the teeth of these stupefying odds it is you and I, in our ordinariness, that are here."

Richard Dawkins – evolutionary biologist, science communicator

No matter the serendipity of life, it is the combination of special people that I have encountered through my existence that has led me to where I am now. To all of them I owe my deepest appreciation.

Particularly, I would like to express my gratitude to Prof. Joachim Lingner, for the opportunity to pursue my PhD studies in his lab. Joachim's analytical guidance and scientific input were essential, and his outstanding patience and enthusiastic support made the whole journey so much more enjoyable. I am especially thankful for Joachim's trust and encouragement, so vital to surpass the multiple failures inherent to laboratory work, and fundamental to face the apprehension of public speaking.

I also wish to thank Prof. Elisa Oricchio, my kind mentor and president of my thesis committee, as well as Prof. David Suter for his constructive input during my candidacy exam, as well as department seminars, and for his participation as a member of my thesis committee. I am also thankful to Prof. Susan Smith and Prof. Claus Azzalin, for the generous interest in my research project demonstrated in scientific conferences, and their participation as members of my thesis committee. I would also like to acknowledge the constructive involvement of Prof. Françoise Stutz and Prof. Oliver Mühlemann in my PhD candidacy exam and mid-thesis meeting, respectively. Furthermore, I thank all members of the NCCR network with whom I have shared my research, making academic life exciting.

Daily lab life would not have been possible without all the supportive group mates that have helped me through the years. I owe a special thanks to Marianna, who taught me most lab procedures and scientific reasoning, patiently replied to my extensive and detailed questions, and shared her inspiring experience. In addition, Marianna has allowed me to participate in her project, and her work has laid the foundation for my main PhD project. For that I am thankful. I would also like to thank Wareed, Galina and Efty, three experienced and inspiring post-docs always available to explain and discuss all kinds of scientific questions. They are role models defining the finest standards of lab work, but also realistic and humble examples of lab-life, with a comforting word when experiments would fail. I am also particularly lucky to have shared most of my doctoral studies with Anna Näger and Trang. Anna's friendship, kindness and contagious laughter always lightened the mood, even in the darkest days, and Trang's relaxed and light-hearted approach to life always helped to put things into perspective. I was also fortunate to have shared unforgettable moments with other Lingner lab post-docs and PhD students, in and outside the lab, including Patricia – who was so welcoming from the beginning –, Aleks, Anna Briod, Gabriela, Suna and Samah. I deeply appreciate all their constructive scientific feedback, their encouragement and all memorable conversations over lunch. Moreover, I wish to thank Tom, who kept everyone's spirits high, while taking care of the lab. I am also grateful to Tom for taking over different *in vitro* experiments, enduring all my comments and requests. I would also like to acknowledge the work developed by my internship student Stanislas Bonnin, who cloned shRNAs for downregulation of THOC subunits. I am thankful for Galina's comments to this thesis introduction and the time dedicated by Tom and Samah to translate the abstract.

I would like to leave my appreciation for all the support provided by different EPFL core facilities, including Bioimaging and Optics Platform (BIOP), Gene Expression (GECF) and Flow Cytometry (FCCF) core facilities. I would also like to thank Tatiana Dubi, Nicole and Natalia, for making all

administrative duties a breeze, and to Barbara Grisoni, for all the support that allowed me to fulfill and enjoy my teaching duties.

My enthusiasm for natural sciences in general, and cell and molecular biology in particular was nourished by different teachers, from primary school to university. They have allowed me to see the infinite beauty in something invisible to the naked eye. I would also like to express my sincere gratitude to my Master's thesis supervisor Linda Vidarsdóttir, who was my first PhD student role-model and who always shared with me the wonders of academic life.

All the words in the world would be insufficient to thank my friends and family. Very briefly:

Thanks to Martocas, for sharing the passion for arts and crafts, but also the frustrations of PhD life.

Thanks to Vanessa, Tio Tópê, Ana, Tia Sara, Sérgio, Simião, Vasco, Inês, Mariana, Mafalda, Sónia, Tózé, Isabel, Augusto and Xaninha for all the comforting words that always warm my heart in the coldest days away from home.

Thanks to my second family, my parents-in-law Nina, Rui and Fernanda, and my sisters- and brothers-in-law.

Thanks to Sofia, Ricardo and Matilde, for facilitating the settlement in Lausanne.

Thanks to Wally The Cat for keeping my spirits up.

Thanks to Fenhanha, for all the love and all the silly but touching moments.

Thanks to my grandparents Alda, António, Laura e Ilídio, who have defined a major part of who I am, and who have always expressed their love and support, regardless of where they are and where I go.

Thanks to my brother Guilherme, who always manages to find the most unconventional ways to express his friendship.

Thanks to my best-friend and life partner Pedro, for his patient and encouraging support, making me step out of the comfort-zone, and for his colossal love.

Finally, I would like to deeply thank my parents Margarida and Miguel, for their unconditional love and heartening support.

Thank you! *Obrigada*!

Abstract

Telomeres are the nucleoprotein structures found at the ends of linear chromosomes. They ensure that the termini of chromosomes are not inappropriately recognized as sites of DNA damage, and are therefore crucial for genome stability. In spite of the heterochromatin-characteristic features present at telomeres and subtelomeric regions, telomeres are transcribed into the telomeric repeat containing RNA – TERRA.

The transcription of TERRA stems from promoters residing within subtelomeric sequences in most chromosome arms. It proceeds through the telomeric repetitive tract, using the C-rich telomeric strand as template. TERRA has been implicated in multiple processes, such as the regulation of the heterochromatic structure of telomeres, and the modulation of telomerase – the ribonucleoprotein complex which extends telomeric DNA. Notably, TERRA is capable of hybridizing telomeric DNA, forming three-stranded structures termed R-loops, which comprise a DNA:RNA hybrid and a displaced DNA strand. Remarkably, many functions attributed to TERRA have been shown to depend on TERRA R-loops, including the stimulation of telomere elongation mediated by homologous-recombination, as well as the interference with the progression of the DNA replication machinery for appropriate duplication of telomeric DNA.

Aiming to elucidate the regulation of TERRA recruitment to telomeres, a reporter system was developed for the study of ectopically-expressed TERRA-like RNA molecules. Employing this system revealed that TERRA can associate with chromosome ends post-transcriptionally *in trans*, in different phases of the cell cycle, forming R-loop structures preferentially at short telomeres.

The study of the regulation of TERRA association with telomeres is of utmost relevance, given that the excess or scarcity of TERRA R-loops can impinge on telomeric stability and maintenance, thus potentially affecting cellular life span and tumorigenesis. Importantly, we show that TERRA R-loops are mediated by the DNA recombinase RAD51, which directly binds TERRA and stimulates TERRA invasion of telomeric DNA. On the other hand, we also report that the association of TERRA with telomeres through R-loops is counteracted by the THO complex (THOC) – a protein complex which links transcription with RNA processing and was previously shown to restrain the accumulation of DNA:RNA hybrids throughout the genome. We found that THOC counteracts R-loops formed at telomeres co-transcriptionally and also post-transcriptionally *in trans* – when TERRA is ectopically-expressed. We demonstrate that THOC binds nucleoplasmic TERRA, and that RNaseH1 loss increases THOC telomeric occupancy. Additionally, RAD51-mediated TERRA R-loop formation results in telomere fragility – which is indicative of defects in the semiconservative replication of telomeric DNA. Concurrently, THOC counteracts R-loop-derived telomeric fragility. Particularly, we found that THOC restrains telomeric fragility at telomeres replicated by lagging strand synthesis and mainly by leading strand synthesis. Finally, we observed that THOC suppresses telomeric sister chromatid exchange and C-circle accumulation in ALT cells, which employ a recombination-mediated telomere maintenance mechanism.

Overall, this work contributes to the comprehensive characterization of the regulation of TERRA association with telomeres as R-loops, and explores the impact of the accumulation of these structures in telomere integrity.

Keywords

Telomeres / TERRA / R-loops / DRIP / RAD51 / THO complex / telomeric fragility / recombination / C-circles

Résumé

Les télomères sont des structures nucléoprotéiques présentent aux extrémités des chromosomes linéaires. Ils permettent d'éviter la reconnaissance des extrémités chromosomiques comme des cassures doubles brins, et donc ont un rôle important dans la stabilité du génome. Malgré les marques hétérochromatiques présentent aux télomères et aux régions subtélomériques, les télomères sont transcrits en un ARN appelé TERRA.

TERRA est transcrit à partir des promoteurs situés dans les régions subtelomériques. La plupart des régions subtélomériques peuvent générer ce transcrit qui utilise comme matrice le brin riche en cytosine. TERRA a des nombreuses fonctions tel que la régulation de la structure hétérochromatique des télomères, ou la régulation de l'activité de la télomérase – une ribonucléoprotéine qui permet d'ajouter des séquences répétées d'ADN télomériques. Remarquablement, TERRA peut former un hybride ADN:ARN qui provoque l'ouverture de l'ADN double brin et constitue ce qui est appelé une R-loop. Plusieurs fonctions attribuées à TERRA proviennent de cette capacité à former des R-Loops, parmi lesquelles la stimulation de l'élongation des télomères provenant de la recombinaison homologue, ainsi que l'interférence du processus de réplication responsable de la duplication de l'ADN.

Dans le but d'élucider la régulation du recrutement de TERRA aux télomères, un système rapporteur a été développé au laboratoire pour l'étude des molécules de TERRA ectopiquement exprimées. En utilisant ce système, nous avons mis en évidence que TERRA pouvait s'associer de façon post-transcriptionnel en *trans*, aux différentes phases du cycle cellulaire. De plus, les R-loops sont préférentiellement formés aux télomères courts.

L'étude de la régulation de l'association de TERRA aux télomères est de toute importance étant donné que l'excès ou la rareté des R-loops peut avoir un effet délétère sur la stabilité et la maintenance des télomères, affectant donc potentiellement la tumorogénèse et la durée de vie cellulaire de façon générale. Nous avons démontré que les R-loops sont contrôlées par l'ADN recombinase RAD51, qui se lie directement à TERRA et stimule l'invasion de l'ADN télomérique par ce dernier. Nous avons démontré également que l'association de TERRA aux télomères est contrecarrée par le THO complexe (THOC) – un complexe protéique à la fois important pour la transcription et la machinerie de régulation de l'ARN, et dont il a été démontré précédemment qu'il limite l'accumulation d'hybrides ADN:ARN dans tout le génome. Nous avons observé que THOC s'oppose aux R-loops formés au niveau des télomères de manière co-transcriptionnelle et également post-transcriptionnelle en *trans* - lorsque TERRA est exprimé de manière ectopique. Nous avons démontré que THOC se lie à TERRA dans le nucléoplasme et que la déplétion de la RNaseH1 augmente la présence de THOC1 aux télomères. De plus, la formation des R-loops dépendante de RAD51 induit le phénomène de fragilité des télomères, indicatif d'un problème dans le processus semi-conservatif de la réplication de l'ADN télomérique. En même temps, THOC contrecarre la fragilité provenant de la présence des R-loops. Particulièrement, nous avons observés que THOC limite la fragilité télomérique au niveau des télomères répliqués par synthèse de brin retardé et principalement par synthèse de brin avancé. Enfin, nous avons observé que THOC joue un rôle important dans les cellules ALT où il limite le processus de recombinaison, et particulièrement l'échange entre chromatides sœurs et l'accumulation de C-circles.

De manière générale, ce travail contribue à la caractérisation de la régulation de l'association de TERRA aux télomères sous la forme de R-loops et explore l'impact de l'accumulation de ces structures pour l'intégrité des télomères.

Mots-clés

Télomères / TERRA / R-Loops / DRIP / RAD51 / THO complexe / fragilité des télomères / recombinaison / C-circles

Contents

Chapter 1 Introduction

1.1. Telomere structure, function and maintenance

Interest in the ends of linear chromosomes emerged in the 1930s from the fundamental observation that physiological chromosome ends must be distinct from chromosome breaks induced by X-rays, which often result in detrimental rearrangements (Muller, 1938; McClintock, 1939, 1941). The term telomere was then coined by H. Muller – from the Greek "end part" – referring to the termini of linear chromosomes, and it was postulated that a protective role was associated with these structures.

Following the discovery of the DNA double helix structure, a theory of the mechanism for DNA duplication was established, ensuring appropriate segregation of the genetic material into two identical daughter cells originating from a single cell (Watson & Crick, 1953; Meselson & Stahl, 1958). In the case of linear chromosomes, this mechanism implied that with each replication cycle DNA molecules would shorten, resulting in loss of genetic information (Watson, 1972). This is recognized as the end-replication problem. Indeed, it was discovered that there is a finite number of duplications that can be undertaken by somatic cells in culture, before they cease to divide – known as the Hayflick limit (Hayflick, 1965). Consistent with the idea that telomeres play a protective function, it was suggested that the terminally-localized genes – designated "telogenes" – would be sacrificed during successive cell divisions, functioning as "buffers" that would be gradually exhausted until a fixed number of somatic cell duplications is reached (Olovnikov, 1973). Later on, it was understood that this loss of genetic material was justified not only by a replication-inherent problem, but due to post-replicative processing of chromosome ends (see chapter 1.1.2).

Remarkably, the function of telomeres is tightly linked to its structure (Szostak & Blackburn, 1982; Lundblad & Szostak, 1989). In a ciliated protozoan, it was originally found that the termini of linear ribosomal RNA genes are composed of a repeated hexanucleotide sequence of variable length (Blackburn & Gall, 1978). This repetitive feature has been found to be a conserved attribute of telomeres across eukaryotes, displaying a tandem array of guanine-rich repeats – underscoring the tight link between telomeric structure and function.

In human cells, telomeres consist of $(TTAGGG)_n$ repeats oriented 5' to 3' towards the end of chromosomes (Figure 1.1), stretching from ca. 2 to 15 kb long sequences (Moyzis *et al*, 1988; Hastie *et al*, 1990; de Lange *et al*, 1990). In various species, including human, all telomeres terminate into a 3' overhang, with a variable length of 50 to 300 nt, where the G-rich strand is longer than the C-rich strand (Klobutcher *et al*, 1981; Henderson & Blackburn, 1989; Makarov *et al*, 1997).

Subtelomeres designate the sequences located immediately upstream of the repetitive telomeric tract. They bear chromosome-arm-specific sequences and comprise telomeric-like repeats and segmentally duplicated tracts (Riethman *et al*, 2005). A subset of subtelomeres contains a conserved tandem repeat-containing region termed "61-29-37 repeats" (Nergadze *et al*, 2009). Several human subtelomeres contain promoters that drive transcription of the telomeric long noncoding RNA TERRA – Telomeric Repeat containing RNA (Azzalin *et al*, 2007) (Figure 1.1) (see chapter 1.2).

1.1.1. The telomere-specific shelterin complex and chromosome end protection

Several proteins have been found at chromosome ends. The most abundant protein complex specifically accumulated at human telomeres – named shelterin complex – comprises six distinct subunits (de Lange, 2005) (Figure 1.1). Two shelterin proteins – TRF1 and TRF2 (telomeric repeat binding factor 1 and 2) specifically bind double-stranded (ds) TTAGGG repeats as homodimers, through Myb domains (Zhong *et al*, 1992; Chong *et al*, 1995; Bilaud *et al*, 1997; Broccoli *et al*, 1997; König *et al*, 1998). The shelterin Rap1 (repressor-activator protein 1) is a constitutive binding partner of TRF2 (Li *et al*, 2000). TRF1 and TRF2 are linked by TIN2 (TRF1-interacting nuclear protein 2) (Kim *et al*, 1999; Houghtaling *et al*, 2004; Liu *et al*, 2004a). In turn, TIN2 associates with TPP1 (combination of TINT1,

PTOP and PIP1) (Houghtaling *et al*, 2004; Liu *et al*, 2004b; Ye *et al*, 2004), which recruits POT1 (protection of telomeres 1) (Baumann & Cech, 2001; Liu *et al*, 2004b; Ye *et al*, 2004). The shelterin POT1 has the particularity of binding the single stranded (ss) G-rich telomeric strand, through OB (oligonucleotide-binding) folds. A crucial purpose of the telomere-specific shelterin complex is ensuring that chromosome ends are not inappropriately recognized as dsDNA breaks, thus avoiding a promiscuous activation of the DNA damage signalling cascade (see below) (Figure 1.1).

The single-stranded 3' overhang found at the termini of telomeres has been shown to invade the double-stranded telomeric tract – at a variable position –, hybridizing with the complementary C-rich strand, leading to displacement of the G-rich strand. This characteristic lasso-like structure – termed "T-loop" (Figure 1.1) – is thought to provide a mechanism for "sequestering" the natural ends of chromosomes from DNA damage signalling (Griffith *et al*, 1999). T-loop dynamics are coordinated by TRF2 (Griffith *et al*, 1999; Doksani *et al*, 2013). This shelterin regulates T-loop unwinding by RTEL1 helicase for appropriate telomere replication in S-phase, and blocks unscheduled T-loop dismantlement throughout the rest of the cell cycle (Vannier *et al*, 2012; Sarek *et al*, 2015, 2019). When tucked into the T-loop conformation, chromosome ends are "concealed" from the MRN (Mre11-NBS1-RAD50) complex – one of the first sensors and responders to dsDNA breaks –, averting the activation of the DNA damage response (DDR) via ATM kinase (Ataxia-telangiectasia mutated) pathway and c-NHEJ (classical non-homologous end joining) signalling by Ku70/80 (Doksani *et al*, 2013). Additionally, TRF2 directly prevents the oligomerization of Ku70/80 (Ribes-Zamora *et al*, 2013) and inhibits the accumulation of 53BP1 (p53-binding protein 1) at dsDNA breaks (Okamoto *et al*, 2013), therefore counteracting c-NHEJ-mediated repair. TRF2 deficiency results in phosphorylation of ATM downstream effectors, leading to the formation of so-called telomere dysfunction-induced foci (TIFs) – visualized by immunofluorescence detection of 53BP1 or γH2AX (Takai *et al*, 2003). In addition, a notorious phenotype observed in TRF2-depleted cells is the formation of chromosome end-to-end fusions, as telomeres become substrates for c-NHEJ-mediated repair (van Steensel *et al*, 1998; Celli & de Lange, 2005). This leads to considerable chromosomal instability through deleterious breakage-fusion-bridge cycles (McClintock, 1941). Co-depletion of TRF2 and TPP1 in mouse embryonic fibroblasts (MEFs) was suggested to induce the alternative-NHEJ pathway, also resulting in increased chromosome fusions (Rai *et al*, 2010).

The shelterin subunit POT1 is recruited to telomeres via TPP1 and specifically binds the single stranded 3' overhang (Baumann & Cech, 2001; Liu *et al*, 2004b; Ye *et al*, 2004). The elevated concentration of POT1 at chromosomes ends excludes binding of ss-telomeric DNA by RPA (replication protein A). In the case of Pot1a loss (one of the two Pot1 proteins identified in mouse cells), RPA binds ssDNA with high affinity, which elicits ATR (ataxia telangiectasia and Rad3 related) kinase activation and phosphorylation of downstream targets, eventually culminating in cell cycle arrest (Denchi & de Lange, 2007). In human and mouse cells, conditional deletion of POT1 (or Pot1a) further revealed a role for this shelterin in suppressing detrimental homology-directed repair (HDR) at telomeres (Wu *et al*, 2006; Glousker *et al*, 2020). POT1 – and subsequent RPA – replacement by ssDNA-binding factors involved in classical recombination – such as the DNA recombinase RAD51 – may be envisioned as a possible mechanism responsible for eliciting HDR at telomeres devoid of POT1 (Glousker *et al*, 2020).

ATR activation at telomeres may not only stem from unprotected ssDNA at telomeric overhangs, but also – for example – from exposed ssDNA stretches at damaged or stalled replication forks or other structures formed at telomeres, such as DNA:RNA hybrids (see chapter 1.2.2.1). The shelterin TRF1 contributes to the suppression of ATR activation at telomeres during S-phase, given its key role in suppressing replication defects at telomeres (Sfeir *et al*, 2009) (see chapter 1.1.2). Furthermore, loss of TRF1 in MEFs results in ATR-mediated TIF accumulation (Sfeir *et al*, 2009).

Overall, the shelterin complex constitutes the central solution for the end-protection problem, ensuring that the ends of linear chromosomes are not threatened by inappropriate DDR activation throughout the cell cycle.

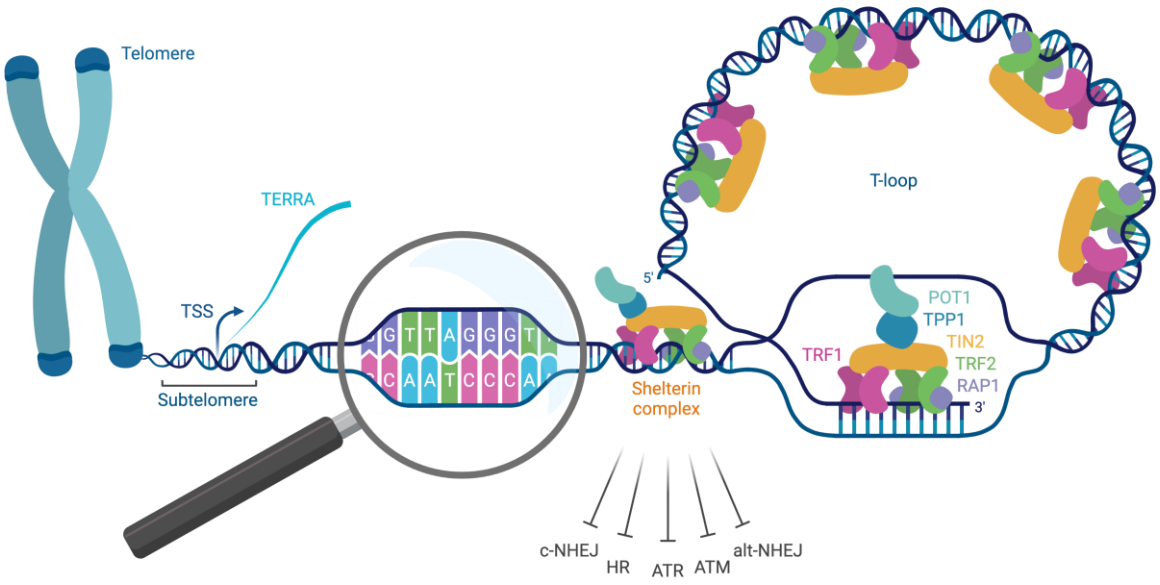


Figure 1.1. The nucleoprotein structure of chromosome ends.

Telomeres are the nucleoprotein structures located at chromosome ends. They consist of an array of 5'-TTAGGG-3' repeats of variable length (highlighted by the magnifying glass). The telomeric single-stranded 3' overhang invades the double-stranded telomeric tract, hybridizing with the C-rich strand, leading to displacement of the G-rich strand – forming the T-loop. This structure is essential for maintenance of telomeric integrity, and is transiently dismantled for appropriate replication of the telomeric DNA. Telomeres are coated by the specialized shelterin complex, comprising double-stranded DNA-binding proteins TRF1 and TRF2, POT1 which binds the single stranded telomeric overhang, TIN2 and TPP1 which connect POT1 to TRF1 and TRF2, and TRF2-bound Rap1. The shelterin complex ensures that chromosome ends are not inappropriately recognized as sites of DNA damage, preventing activation of several repair pathways, including homology-directed repair (HR), ATR and ATM signalling, and canonical and alternative non-homologous end joining (c- and alt-NHEJ). The sequences located immediately upstream of the repetitive telomeric tract are designated subtelomeres, and comprise unique sequences for each chromosome arm. Several subtelomeres have been found to contain promoters driving the transcription of the telomeric long noncoding RNA TERRA (TSS indicates transcriptional start site). *This illustration was created with BioRender.com.*

1.1.2. Telomere shortening and telomere fragility

As predicted upon the conceptualization of the mechanism for DNA duplication, chromosome ends shorten with every cell division. This is partially due to the fact that DNA polymerases require a free 3' OH group as a site for nucleotide addition. Because of this inability for *de novo* synthesis, and since the replication machinery at telomeres moves predominantly towards chromosome ends (Drosopoulos *et al*, 2012), replication must occur discontinuously to duplicate the telomeric G-rich strand oriented 5' to 3'. It involves the generation of short RNA primers by primase, which are extended by DNA polymerases α and δ, synthesizing ca. 200 bp stretches at a time. The RNA primers are then removed, leaving gaps which can be filled-in and ligated (MacNeill, 2012). Due to the removal of the distal-most RNA primer, a short 3' overhang is formed. As for the telomeric C-rich strand oriented 3' to 5', replication occurs in a continuous manner by leading strand synthesis by DNA polymerase ε, leaving a blunt end (MacNeill, 2012). However, all chromosome ends within each cell have 3' single-stranded overhangs, rather than blunt ends (Klobutcher *et al*, 1981; Henderson & Blackburn, 1989; Wellinger *et al*, 1993; Makarov *et al*, 1997; McElligott & Wellinger, 1997). Furthermore, in the case of mouse and human cells, the telomere shortening observed for each population doubling exceeds that attributable

to the removal of the RNA primer alone (Huffman *et al*, 2000). Therefore, chromosome ends were found to be subjected to a post-replicative processing that generates 3' overhangs.

Formation of 3' overhangs at telomeres is initiated by TRF2 and regulated by POT1 (Wu *et al*, 2012). It involves 5' resection by Apollo nuclease – recruited by TRF2 – at the parental strand duplicated by leading strand replication, followed by extensive 5' resection by Exonuclease 1 at both telomeres. POT1 then engages the CST (CTC1-STN1-TEN1) complex and DNA polymerase α-primase for C-strand fill-in synthesis (Dai *et al*, 2010; Wu *et al*, 2012), leaving both ends with a 3' overhang. This process results in one chromatid with equal telomere length – following lagging strand synthesis – as the parental molecule, while leading strand synthesis and nuclease processing generates a telomere shorter than the one contained in the parental molecule. In different human cell lines, a correlation can be established between 3' overhang length and the shortening rate (Huffman *et al*, 2000).

Telomeres are difficult to replicate regions. The hexameric repeats that compose telomeres and their GC-rich nature represent a risk for replication fork slippage (de Lange *et al*, 1990). In addition, it has been suggested that the progression of the replication bubble at telomeres is mostly unidirectional – from subtelomeres towards telomere ends (Drosopoulos *et al*, 2012). This implies that the G-rich telomeric strand is replicated by lagging strand synthesis, increasing the chances for bare G-rich ssDNA to form stable G-quadruplex (G4) structures, which can impede the progression of the replication fork (Granotier *et al*, 2005; Lerner & Sale, 2019). Moreover, G4s can also form at the displaced DNA strand of R-loops (Duquette *et al*, 2004) – which are structures formed at telomeres when the telomeric RNA TERRA hybridizes with the C-rich DNA strand. Telomeric transcription as well as the accumulation of telomeric R-loops constitute major hurdles to proper telomere replication (Balk *et al*, 2013; Arora *et al*, 2014; Sagie *et al*, 2017) (see chapter 1.2.2.1). T-loops – formed at telomere ends to protect telomeric 3' overhangs – are normally unwound in S-phase, allowing access of the replication machinery to chromosome termini. T-loop persistence constitutes a major threat to telomere replication, resulting in excision by the SLX4 nuclease and culminating in telomere loss (Vannier *et al*, 2012; Sarek *et al*, 2015). Positive supercoiling formed ahead of the replication fork due to unwinding by the replicative helicase also constitutes another source of topological stress at the telomere, which must be resolved for faithful DNA duplication (Bermejo *et al*, 2007; Ye *et al*, 2010). Furthermore, *in vitro* experiments have suggested that the presence of the shelterin complex at telomeres must be finely regulated so that its role in promoting telomeric replication is fulfilled, while tight binding to telomeric DNA does not impair replication fork progression (Ohki & Ishikawa, 2004; Sfeir *et al*, 2009). Besides the multiple features identified at telomeres that can potentially constitute obstacles for complete and faithful telomere duplication, the unidirectionality of the replication fork at telomeres means that, in the eventuality of replication stalling and fork collapse, a convergent fork is unlikely to sustain its rescue.

Challenges encountered by the replication machinery may give rise to defects that can be observed in metaphase chromosomes by fluorescence *in situ* hybridization (FISH). This phenotype can be induced by the replication inhibitor aphidicolin at so-called fragile sites, by delaying replication at these genomic loci and thus affecting chromatin condensation (Glover *et al*, 2017). Similarly, replication stress at chromosome ends promotes the formation of fragile telomeres, manifested as smeary or multiple telomeric FISH signals per chromosome end (Sfeir *et al*, 2009). While the molecular source of this feature is not fully elucidated, it is thought to originate from incomplete DNA replication and the subsequent repair processes of damaged replication forks – leaving ssDNA gaps –, or partial chromatin condensation (reviewed in (Glousker & Lingner, 2021)).

Several factors that enable replication fork progression have been shown to counteract telomere fragility, for example by regulating the accumulation of topological obstructions such as G4s, R-loops or the T-loop, or elicit/participate in DNA repair at telomeres. The shelterin component TRF1 constitutes a paradigmatic example. TRF1 has been shown to be a major regulator of telomeric fragility by recruiting the helicase BLM for unwinding of G4s (Sfeir *et al*, 2009; Zimmermann *et al*, 2014), as well as counteracting the TRF2-mediated formation of TERRA R-loops (Lee *et al*, 2018). In addition, it was proposed that TRF1 interacts and functions together with TIMELESS – which is part of the fork protection complex, stabilizing the replisome (Leman *et al*, 2012). It was also suggested that TRF1

counteracts telomere fragility by mediating the binding of DNA topoisomerase II α to telomeres, which removes topological stress during replication (d'Alcontres *et al*, 2014). Moreover, a recent study in MEFs showed that TRF1 interacts with and recruits TFIIH to telomeres. There, TFIIH acts as a key factor facilitating DNA replication, thus preventing telomeric fragility. While TFIIH is a key component involved in transcriptional initiation and nucleotide excision repair (NER), its TRF1-mediated role in telomere protection appears to involve a noncanonical function (Yang *et al*, 2022).

1.1.3. Telomere maintenance

Inadequate telomere maintenance results in insufficient shelterin binding and the inability to form a T-loop, leading to activation of ATM- and ATR-dependent checkpoint signaling and cellular senescence or apoptosis (Yu *et al*, 1990; Maciejowski & de Lange, 2017). Therefore, telomere elongation mechanisms have evolved as a form of DNA repair localized to chromosome ends, to compensate for telomere shortening in rapidly proliferating cells – such as stem cells, germ line cells or tumour cells.

Most eukaryotes resort to the specialized enzyme telomerase to counteract telomere attrition. Telomerase is a ribonucleoprotein complex which extends telomeres, adding telomeric repeats *de novo* – one nucleotide at a time. It is comprised of a reverse transcriptase protein subunit (TERT) and an RNA component (TERC) which serves as a primer and a template to specify the sequence added to chromosome ends (Greider & Blackburn, 1985, 1987; Feng *et al*, 1995; Lingner *et al*, 1997; Nakamura *et al*, 1997).

During human embryonic development, telomerase activity is stringently suppressed in the large majority of somatic cells, either by epigenetic-mediated transcriptional repression, or processing of hTERT RNA or hTR (human TERC) (Cong *et al*, 2002). Upon critical shortening, the protective function of telomeres and the shelterin complex succumbs, DNA damage signaling cascades are triggered, and cells eventually subside into cell cycle arrest and replicative senescence (Fagagna *et al*, 2003). Thus, cessation of telomerase activity provides a molecular expiration date dictating the proliferative lifespan of human somatic cells, acting as a tumor suppressor. In the majority of tumors – 85 to 90% of cancers –, transformed cells restore telomerase expression and activity – often by acquiring mutations in hTERT promoter (Sharma & Chowdhury, 2022) – to sustain high cell division rates and avert telomere-derived cellular senescence (Kim *et al*, 1994; Shay & Bacchetti, 1997).

The alternative lengthening of telomeres (ALT) is a telomerase-independent pathway, which resorts to homology-directed recombination between telomeres (Bryan *et al*, 1995; Dunham *et al*, 2000). It is used for telomere maintenance in 10-15% of cancer types – with a higher prevalence in aggressive tumours with mesenchymal cell origin (Henson *et al*, 2005; Heaphy *et al*, 2011a). In ALT cells, break-induced replication (BIR)-mediated telomere synthesis requires the RFC-PCNA-Polδ replisome and is instigated by exacerbated and unresolved replication stress (Dilley *et al*, 2016; Roumelioti *et al*, 2016; Lu & Pickett, 2022). Telomere processing by ALT occurs upon telomere clustering in ALT-associated PML (promyelocytic leukemia) bodies (APBs), and is thought to follow two mechanisms with converse RAD52-dependency (Yeager *et al*, 1999; Zhang *et al*, 2019). Recombination intermediates can be dissolved by the BLM-TOP3A-RMI (BTR) complex – which is essential for ALT activity –, promoting telomere extension with no exchange of telomeric DNA, or resolved by SLX4–SLX1–ERCC4, where telomere extension is aborted and crossover events may occur (Wu & Hickson, 2003; Svendsen *et al*, 2009; Sobinoff *et al*, 2017). This results in telomeric sister chromatid exchanges (T-SCEs) and a distinctive telomere length heterogeneity (Bryan *et al*, 1995; Londoño-Vallejo *et al*, 2004) – two hallmarks of ALT cells. Extra-chromosomal telomeric repeats (ECTR) in the form of C-rich circles – likely formed as a by-product of recombination – are also commonly detected in ALT-immortalized cells (Ogino *et al*, 1998; Tokutake *et al*, 1998).

While characteristic features of ALT activity have been identified, the precise molecular trigger of this telomere maintenance mechanism is still ill-defined. Many ALT lines have been shown to have a G2/M checkpoint deficiency, but this could not be correlated with defects in the ATM or ATR signalling pathways (Lovejoy *et al*, 2012). In addition, mutations in the genes encoding the histone variant H3.3

or the chromatin remodelers ATRX (α-thalassemia/mental retardation X-linked) and DAXX (death-domain-associated protein) are commonly found in ALT tumour cells (Heaphy *et al*, 2011b; Schwartzentruber *et al*, 2012). Loss of ATRX or DAXX affects H3.3 deposition at telomeres (Lewis *et al*, 2010), which is thought to modify the telomeric heterochromatic status, contributing to telomere recombination. Other mechanisms implicating ATRX in replication progression or telomere cohesion have also been suggested as possible explanations for the correlation of ALT with ATRX deficiency (Pickett & Reddel, 2015). Nonetheless, loss of any of these factors appears to be insufficient to activate ALT (Lovejoy *et al*, 2012). *In vitro* activation of ALT has been achieved through downregulation of the ASF1a and ASF1b histone chaperones (O'Sullivan *et al*, 2014). It results in *de novo* appearance of ALT features such as APBs, ECTRs, T-SCEs accompanied by increased heterogeneous telomere length, and even in repression of hTERT expression and activity (O'Sullivan *et al*, 2014). Even though ASF1 is expressed in ALT cells, these results emphasize the notion that an altered chromatin organization underlies ALT activation (O'Sullivan *et al*, 2014). Furthermore, another key factor with a demonstrated contribution to triggering/sustaining ALT activity is the telomere-derived transcript TERRA (see below).

1.2. The telomeric repeat-containing lncRNA: TERRA

1.2.1. Telomeric transcription

The telomere position effect (TPE) has been described using yeast, over the observation that the expression of reporter genes is reversibly repressed when they are placed adjacently to telomeric repeats (Gottschling *et al*, 1990). TPE has also been described in human cells (Baur *et al*, 2001). This phenomenon is consistent with the fact that telomeres and subtelomeric regions are enriched in features characteristic of heterochromatic domains. At mammalian chromosome ends these features include DNA methylation, trimethylation of lysines 9 and 27 at histone H3, trimethylation of lysine 20 at histone H4 and hypoacetylation of H3 and H4, and binding by HP1α (heterochromatin protein 1) (Blasco, 2007; Tardat & Déjardin, 2018). These observations supported the assumption that telomeres are devoid of transcriptional activity. However, telomeres were found to be transcribed into the telomeric repeat containing RNA (TERRA) (Figure 1.2) (Azzalin *et al*, 2007).

TERRA was detected in human (Azzalin *et al*, 2007), mouse and Chinese hamster cells (Azzalin *et al*, 2007; Schoeftner & Blasco, 2008), zebrafish (Schoeftner & Blasco, 2008) and other eukaryotes including plants (Vrbsky *et al*, 2010), yeast (Luke *et al*, 2008; Bah *et al*, 2012; Greenwood & Cooper, 2012) and protozoa (Nanavaty *et al*, 2017), demonstrating its conservation across species.

The transcription of TERRA is driven by RNA Polymerase II. It stems from promoters residing within subtelomeric sequences at most chromosome arms (Figure 1.1), and proceeds through the telomeric tract, using the C-rich telomeric strand as template (Azzalin *et al*, 2007; Schoeftner & Blasco, 2008; Nergadze *et al*, 2009; Porro *et al*, 2014; Feretzaki *et al*, 2019). In human cells, a subset of TERRA promoters comprises CpG islands which are methylated *de novo* during development and remain methylated in most somatic cells (Nergadze *et al*, 2009; Porro *et al*, 2014; Feretzaki *et al*, 2019). The DNA methyltransferases DNMT1 and -3B are responsible for methylation of subtelomeric CpG islands, inhibiting the expression of a fraction of TERRA molecules (Yehezkel *et al*, 2008). Also the shelterin TRF2 has been shown to repress TERRA transcription (Porro *et al*, 2014). Many transcription factors have been identified as transcriptional regulators of TERRA (Feretzaki *et al*, 2019), including CTCF (CCCTC-binding factor) and the cohesion subunit RAD21, which have been show to bind to subtelomeres and sustain telomeric transcription (Deng *et al*, 2012).

1.2.2. Molecular features, regulation and functions of TERRA

Given that TERRA transcription is initiated at subtelomeric sequences and extends through telomeric repeats at several chromosome arms, TERRA refers to a heterogeneous class of long non-coding RNAs (lncRNAs), containing chromosome arm-specific subtelomeric-derived sequences at the 5' end, followed by an array of UUAGGG repeats (Porro *et al*, 2010; Feretzaki *et al*, 2019). In human cells,

TERRA molecules range in size from about 100 bases up to 9 kb and are not translated into known functional proteins – justifying the long non-coding RNA designation (Azzalin *et al*, 2007; Statello *et al*, 2021). The length heterogeneity is thought to arise from the lack of a specified transcription termination site, resulting in variable processing through telomeric repeats by the RNA polymerase. TERRA is mainly localized in the nucleus (Azzalin *et al*, 2007), and similarly to mRNAs generated by RNA Polymerase II, TERRA contains a 7-methylguanosine cap at the 5' end, which is thought to have a protecting role from 5'-to-3' exonucleases (Shatkin, 1976; Porro *et al*, 2010). Only a minor fraction of TERRA molecules is polyadenylated (< 10%), which is more stable – with a half-life exceeding 8 hours – than the poly(A)-negative TERRA majority – with a half-life of ~3 hours (Porro *et al*, 2010). While a considerable fraction of poly(A)-negative TERRA is found in association with chromatin – which is thought to have major implications for TERRA function (see below) –, polyadenylated TERRA is mainly found in the nucleoplasm (Porro *et al*, 2010) (Figure 1.2). Analysis of the 5' subtelomeric-derived region of TERRA did not reveal any splicing events undertaken during TERRA biogenesis (Porro *et al*, 2014). Recently, the subtelomeric-derived region of human TERRA has been shown to contain the N^6-Methyladenosine (m^6A) modification, which is catalysed by the m^6A methyltransferase METTL3 and protected by the YTHDC1 reader (Chen *et al*, 2022). The m^6A post-transcriptional modification was shown to contribute to the stability of TERRA, without affecting its transcription (Chen *et al*, 2022).

The telomeric structure and composition is modulated throughout the cell cycle, namely to enable progression of the replication machinery, and access and extension by telomerase. Concurrently, cellular levels of TERRA are also subjected to cell cycle regulation. RT-qPCR analysis of the abundance of different TERRA sub-species in synchronized human HeLa cells demonstrated that the highest TERRA levels are measured in early G1, gradually decreasing in late G1 and S phases, and reaching the lowest levels in late S/G2. As cells progress through G2 and mitosis into G1, TERRA levels progressively increase (Porro *et al*, 2010). Whether this cyclic process depends on transcriptional or post-transcriptional regulation in human cells is still unexplored. In budding yeast, maximum TERRA levels are reached in S-phase, decreasing throughout S and G2 phases (Graf *et al*, 2017).

Cell cycle regulation of TERRA in human cells has been hypothesized to contribute to telomere length maintenance by telomerase (Figure 1.2) – which extends chromosome ends in late S-phase, when TERRA levels reach the lowest levels (Tomlinson *et al*, 2006; Porro *et al*, 2010). *In vitro* experiments have shown that TERRA is a strong inhibitor of telomerase activity, base pairing with hTR and interacting with the catalytic component hTERT (Redon *et al*, 2010). In addition, a fine balance between TERRA and heterogeneous nuclear ribonucleoprotein A1 (hnRNPA1) is thought to regulate telomerase access to telomeres. hnRNPA1 is a major TERRA-binding protein, which also inhibits telomerase *in vitro* and is detected at telomeres (Redon *et al*, 2013). When in excess over hnRNPA1, TERRA may bind and directly inhibit telomerase activity, while when TERRA is at equilibrium with hnRNPA1, telomerase may be released from binding by TERRA and hnRNPA1 – which interact with each other –, thus accessing and elongating chromosome ends (Redon *et al*, 2013). Furthermore, the dynamic interplay between TERRA and hnRNPA1 has been proposed to promote POT1 binding to the telomeric 3' overhang. It occurs through displacement of the strong ssDNA-binding complex RPA by hnRNPA1 – which in turn is not bound by downregulated TERRA in late S-phase (Flynn *et al*, 2011). Re-association of hnRNPA1 with TERRA after S-phase has been hypothesized to then liberate the G-rich overhang for POT1 access. POT1 binding to telomeres has been suggested to counteract telomerase activity *in vivo* and *in vitro* (Loayza & De Lange, 2003; Kelleher *et al*, 2005), therefore TERRA may exert a direct and POT1-mediated indirect negative regulation over telomerase. Of note, while TERRA together with hnRNPA1 may contribute to POT1 association with telomeres by mediating RPA displacement, an alternative model suggests that tethering of POT1 to telomeres by TPP1 and TIN2 is sufficient for an RPA-to-POT1 switch (Takai *et al*, 2011).

Conversely, in yeast, *in vivo* experiments point towards a positive control of telomerase activity by TERRA. In *S. cerevisiae*, TERRA levels have been shown to be elevated in cells with short telomeres, where TERRA clusters with *TLC1* (yeast telomerase RNA) in early S-phase. TERRA then specifically coordinates telomerase recruitment to short telomeres *in cis* – from which TERRA was transcribed –,

promoting their preferential elongation (Cusanelli *et al*, 2013). In *S. pombe*, telomere shortening promotes accumulation of polyadenylated TERRA. Upon transcriptional induction of telomeres, polyadenylated TERRA was found to stimulate the association of telomerase with short telomeres, promoting their elongation (Moravec *et al*, 2016). However, this was solely achieved when TERRA induction was coupled with a histone deacetylase inhibitor treatment, which decreases chromatin compaction (Moravec *et al*, 2016).

TERRA levels are not only regulated in a cell cycle-dependent manner, but are also inversely correlated with telomere length. In human cells, TERRA levels have been shown to decrease upon telomere elongation, through transcriptional repression mediated by SUV39H1 H3K9 histone methyltransferase and HP1 proteins (Arnoult *et al*, 2012). Despite a lower abundance of TERRA in cells with longer telomeres, the UUAGGG content of these TERRA molecules was shown to be higher. In turn, a rise in the size of TERRA was hypothesized to serve as a recruitment platform for histone-modifying enzymes (Figure 1.2), allowing re-establishment of H3K9me3 at telomeres, in a cell cycle-regulated and telomere length-dependent manner (Arnoult *et al*, 2012). This supports a model in which TERRA contributes to the regulation of telomeric heterochromatin through a feedback mechanism. Also upon telomere uncapping following TRF2 depletion in human cells, transcriptional repression of TERRA is lost (Porro *et al*, 2014). Elevated TERRA in TRF2-depleted cells directly interacts with SUV39H1, which maintains telomere chromatin compaction by methylating H3K9. In addition, recruitment of SUV39H1 by TERRA was proposed to sustain end-to-end telomeric fusions prompted by TRF2 deficiency (Porro *et al*, 2014).

In budding yeast, TERRA levels were also found to be increased in cells with shorter telomeres. This process is thought to be regulated by the nuclear 5'-3' RNA exonuclease Rat1p, which is preferentially enriched at longer telomeres where it degrades TERRA (Luke *et al*, 2008; Graf *et al*, 2017). In turn, TERRA accumulation at critically short telomeres elicits DNA damage signalling, promoting recruitment of factors involved in recombination-mediated repair such as RAD51, thus stimulating HDR-mediated telomere elongation (Graf *et al*, 2017) (Figure 1.2). Importantly, stimulation of HDR at telomeres by TERRA has been shown to require direct hybridization of TERRA with the telomeric tract, forming structures termed R-loops (Balk *et al*, 2013; Graf *et al*, 2017) (see chapter 1.2.2.1).

1.2.2.1. Regulation and impact of TERRA R-loops

R-loops are intermolecular three-stranded structures that form when an RNA strand invades the dsDNA helix and hybridizes with the complementary strand forming a DNA:RNA hybrid, leaving a displaced DNA strand. R-loops can range from 100 bp to over 2 kb long (García-Muse & Aguilera, 2019). These structures have been detected throughout the genome and are thought to mainly occur during transcription *in cis*, due to sequence complementarity between the nascent transcript and the template DNA strand (Reaban *et al*, 1994; Yu *et al*, 2003; Crossley *et al*, 2019; Lafuente-Barquero *et al*, 2020). Studies in yeast (Wahba *et al*, 2013) and plant (Ariel *et al*, 2020) revealed that R-loops can also be formed post-transcriptionally, *in trans* – at genomic regions distinct from the loci where the RNA was transcribed from. R-loops tend to form in genomic regions with high GC content and GC skew, and the displaced ssDNA strand is prone to form G4s, which in turn are thought to stabilize the DNA:RNA hybrid (Duquette *et al*, 2004; Miglietta *et al*, 2020).

R-loops have been demonstrated to be crucial intermediates in several processes (reviewed in (Aguilera & García-Muse, 2012)), namely in immunoglobulin (Ig) class-switch recombination (Yu *et al*, 2003), sequence-specific targeting by the bacterial CRISPR-Cas9 DNA endonuclease (Jinek *et al*, 2012), or priming of mitochondrial DNA replication (Xu & Clayton, 1996). In addition, R-loops formed across the genome have been implicated in the regulation of gene expression, being involved in transcriptional activation or silencing – by stimulating binding of transcription factors to promoters where R-loops can accumulate, or by "masking" promoters from transcription factors –, or in promoting transcriptional termination by RNA Polymerase (reviewed in (García-Muse & Aguilera, 2019)). On the other hand, R-loops can pose an obstacle to the elongating transcription machinery, as well as an impairment to DNA

replication, potentially resulting in accumulation of DNA damage and genome instability in the form of transcription-associated recombination (TAR) (Huertas & Aguilera, 2003; Wellinger *et al*, 2006; Gan *et al*, 2011a; Domínguez-Sánchez *et al*, 2011). While R-loops can directly slow down or obstruct the replication fork leading to fork stall and collapse, the chromatin changes promoted by R-loop accumulation may also constitute a barrier to DNA replication, and contribute to genome instability (García-Pichardo *et al*, 2017). Exposure and consequent vulnerability of the ssDNA strand within R-loops, and nucleolytic processing of R-loops have also been described as R-loop-derived sources of DNA damage (reviewed in (Brickner *et al*, 2022)).

At telomeres, R-loops form when TERRA RNA invades the telomeric dsDNA, base pairing with the C-rich strand, leaving the G-rich DNA strand displaced (Figure 1.2). R-loops were first detected at yeast telomeres (Balk *et al*, 2013; Graf *et al*, 2017), resorting to a chromatin immunoprecipitation approach, with the DNA:RNA hybrid-recognizing S9.6 antibody (Boguslawski *et al*, 1986). Also at human telomeres the formation and regulation of R-loops has been the focus of growing interest (Arora *et al*, 2014; Sagie *et al*, 2017; Silva *et al*, 2019; Yadav *et al*, 2022, 1; Kaminski *et al*, 2022).

Notably, several functions associated with the role of TERRA in telomere maintenance have been shown to depend on R-loops. As aforementioned, in telomerase-negative budding yeast, induction of HDR following increased expression and accumulation of TERRA at short telomeres was shown to involve R-loops, as it is sensitive to RNaseH1 overexpression (Graf *et al*, 2017). RNaseH1 is part of the RNaseH ribonuclease family which specifically hydrolyses the RNA moiety in R-loops (Cerritelli & Crouch, 2009). Additionally, it was reported that another member of this family – RNaseH2 – is preferentially recruited to long telomeres where it degrades TERRA R-loops in late S-phase, but is lost at short telomeres where R-loop build-up is required for recombination-mediated telomere re-elongation, avoiding replicative senescence (Graf *et al*, 2017; Misino *et al*, 2022) (Figure 1.2). At short telomeres, Npl3 – previously implicated in transcription and R-loop regulation – was shown to interact with TERRA-mediated R-loops, stabilizing these structures (Pérez-Martínez *et al*, 2020).

Numerous studies have demonstrated the fundamental role of TERRA R-loops specifically in human ALT cancer cells – which depend on recombination to avert telomere shortening and subsequent replicative senescence. ALT cells are characterized by a reduced chromatin compaction and elevated telomeric transcription and telomeric R-loops, compared to telomerase-positive cell lines (Episkopou *et al*, 2014; Arora *et al*, 2014). In addition, loss of the chromatin remodeling protein ATRX in ALT cells compromises cell cycle regulation of TERRA (Flynn *et al*, 2015). Thus, TERRA lingers throughout S-phase and G2, sustaining elevated telomeric R-loop levels and contributing to the elevated replication stress characteristic of ALT telomeres (Arora *et al*, 2014; Flynn *et al*, 2015; Silva *et al*, 2019; Lu & Pickett, 2022).

The role of TERRA in recombination-mediated telomere maintenance is inherently dependent on an extremely tight control of telomeric R-loops (Figure 1.2), as supported by the growing number of reports demonstrating that dysregulation of factors which control telomeric R-loops leads to disproportionate and possibly detrimental ALT activity levels. The ribonuclease RNaseH1 abounds at ALT telomeres, where it directly regulates R-loops by degrading TERRA molecules engaged in DNA:RNA hybrids. Loss of RNaseH1 leads to an excess of R-loops, which triggers an aberrant accumulation of phosphorylated RPA (pSer33RPA) at telomeres – indicative of high replication stress (Olson *et al*, 2006) –, increased telomeric leading strand fragility, C-circles and rapid telomere loss (Arora *et al*, 2014). Strikingly, overexpression of RNaseH1 is also disadvantageous, as it prevents R-loop-derived replication stress from sustaining homology-directed recombination for telomere elongation, thus also resulting in telomere loss derived from telomere shortening (Arora *et al*, 2014). *In vivo* and *in vitro* experiments have shown that the Fanconi anemia-associated ATPase/translocase FANCM also restrains exacerbated TERRA levels and R-loops at telomeres in ALT cells, likely through its translocase activity (Silva *et al*, 2019; Hodson *et al*, 2022). Depletion of FANCM leads to an escalation of replication defects at telomeres as a consequence of an accumulation of R-loops, increased DNA damage signaling and elevated ALT activity. Notably, FANCM deficiency selectively compromises the viability of ALT cells (Silva *et al*, 2019). Similarly, NONO and SFPQ proteins, which have been implicated in different steps of RNA

biogenesis, have also been proposed to suppress TERRA R-loops in ALT cells (Petti *et al*, 2019). NONO/SFPQ heterodimers were identified as TERRA binding proteins. Their depletion in ALT cells promotes the accumulation of DNA:RNA hybrid signals at telomeres. In addition, it increases the frequency of some ALT features, including an elevated localization of phosphorylated RPA at telomeres – which was partially rescued by RNaseH1 overexpression –, increased APBs and telomeric sister chromatid exchanges. Loss of NONO increases telomeric fragility specifically at telomeres replicated by leading strand synthesis (Petti *et al*, 2019). Interestingly, the DNA recombination factor BRCA1 was shown to bind TERRA R-loops and mediate their repair. Depletion of BRCA1 in different human cell lines (including ALT cells) resulted in R-loop-dependent accumulation of DNA damage, accompanied by elevated frequency of telomeric fragility and telomere loss (Vohhodina *et al*, 2021). In ALT cells, XPF – an endonuclease implicated in nucleotide excision repair – is recruited to telomeres via TERRA R-loops. There, it has been proposed to counteract R-loop accumulation by cleaving R-loop-containing DNA, forming DSBs. Processing of R-loops by XPF presumably activates BIR-mediated telomere synthesis by mediating BRCA1 and RAD51 recruitment to telomeres, contributing to proliferation of ALT cells (Sollier *et al*, 2014; Guh *et al*, 2022). Recently, RAD51AP1 (RAD51-associated protein 1), which has previously been demonstrated to be a crucial mediator of ALT (Gonzalez *et al*, 2019), was shown to bind TERRA and contribute to the formation of TERRA R-loops (Yadav *et al*, 2022; Kaminski *et al*, 2022).

Experiments involving direct manipulation of TERRA transcription or stability in ALT cells have also contributed to the perception that a finely-tuned balance of R-loop-induced replication stress is required for triggering HR by break-induced replication, allowing telomere elongation without entirely threatening telomere integrity. Particularly, in ALT cells with engineered Transcription Activator-Like Effectors (TALEs) fused to transcription repressor domains targeting some subtelomeres, TERRA transcription is blocked, which presumably constrains R-loop formation. This is correlated with a reduced DNA damage checkpoint signalling at telomeres (as accessed through colocalization of pSer33RPA and γH2AX with telomeres), impaired ALT activity (decreased APBs, telomere synthesis and POLD3 localization to telomeres) and a rise of telomere loss (Silva *et al*, 2021). Downregulation of TERRA RNA with an RNA-targeting Cas9 system in ALT cells has led to similar results (Guh *et al*, 2022), indicating that it is not only a high telomeric transcription rate *per se* that contributes to the ALT pathway, but TERRA RNA itself. Conversely, the use of TALEs fused to a transcriptional activator enhanced TERRA transcription in ALT cells, and resulted in amplified damage checkpoint signalling at telomeres (further recruitment of pSer33RPA and γH2AX to telomeres) and elevated ALT activity (increased APBs, telomere synthesis and POLD3 localization to telomeres). Interestingly, despite the increase in telomere synthesis – as detected through EdU incorporation at telomeres –, overexpression of TERRA in ALT cells results in accumulation of telomere-free ends (Silva *et al*, 2022). This observation, which had also been reported in experiments with inhibition of TERRA transcription, illustrates the notion that TERRA levels must be kept at precise levels to efficiently sustain BIR-mediated telomere maintenance.

Furthermore, interfering with the methylation status and consequent stability of TERRA in ALT cells was linked to lower frequency of telomeric recombination events, which was rescued when R-loop levels were restored by RNaseH1 depletion. Inadequate methylation of TERRA culminated in telomere loss and increased frequency of telomere fusions (Chen *et al,* 2022).

Altogether, dysregulation of factors which regulate the abundance of TERRA R-loops, as well as direct modulation of TERRA transcriptional levels or TERRA stability in ALT cells have provided cumulative evidence on the significance of TERRA and TERRA R-loops in cells which resort to the ALT pathway for telomere regulation.

Several other factors have been shown to regulate TERRA R-loops in human telomerase-positive cells, impacting telomere stability. The FEN1 endonuclease – which has a canonical function in lagging strand DNA replication –, has been shown to prevent leading strand fragility specifically derived from telomeric DNA:RNA hybrids (Teasley *et al*, 2015). *In vitro* experiments demonstrated that the shelterin component TRF2 can bind TERRA and promote R-loop formation, and TRF2 overexpression in human cells leads to increased R-loop detection at telomeres. Conversely, loss of the shelterin subunit TRF1 was shown to result in elevated TERRA R-loop levels in a TRF2-dependent manner and R-loop-derived

telomere loss (Lee *et al*, 2018). In human cells, conditional deletion of POT1 leads to an accumulation of TERRA R-loops and correlates with an induction of ALT-associated features, including formation of APBs, C-circles and strong activation of telomeric recombination (Glousker *et al*, 2020). UPF1 is an ATPase and 5'-3' helicase involved in the regulation of mRNAs containing premature stop codons through the nonsense mediated RNA decay (NMD) pathway. UPF1 association with telomeres is stimulated by ATR and by telomere elongation. Interestingly, loss of UPF1 in HeLa cells increases TERRA association with telomeres, and causes DNA damage signaling activation, as well as frequent loss of telomeres replicated by leading strand synthesis (Azzalin *et al*, 2007; Chawla *et al*, 2011). The helicase RTEL1 – with a crucial role in dismantling T-loops in S-phase – has been shown to bind G4-containing TERRA RNA, promoting its association with chromosome ends, and thus hindering transcription from telomeres harboring R-loops. Additionally, deletion of RTEL1 promotes telomere loss, limits growth capacity and leads to premature senescence. Strikingly, complementation of RTEL1 deletion with RTEL1 mutants deficient for either TERRA binding or helicase activity failed to rescue the senescence phenotype (Ghisays *et al*, 2021). Whether the defects observed at telomeres in cells with deficiency of POT1, UPF1 or RTEL1 depend on TERRA is not yet fully elucidated.

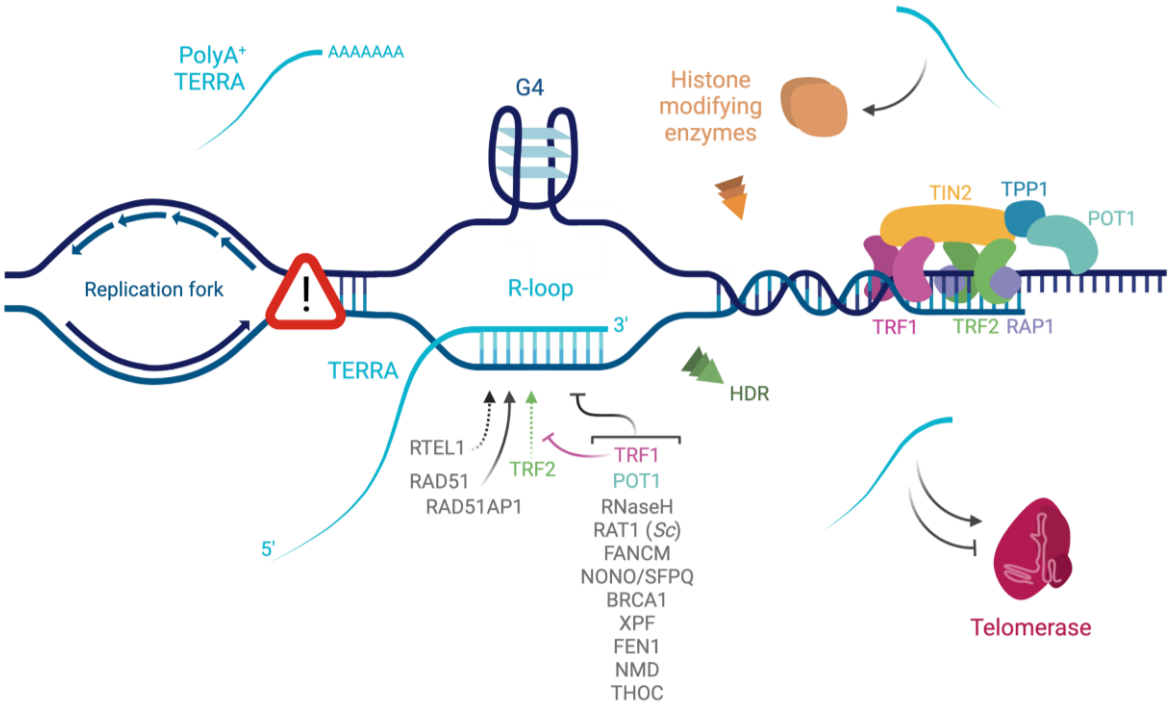


Figure 1.2. The telomeric lncRNA TERRA and its proposed functions at chromosome ends; Regulators of TERRA R-loops.

Telomeres are transcribed into the long non-coding RNA TERRA - Telomeric Repeat containing RNA. The transcription of TERRA stems from subtelomeric promoters and is extended through the telomeric tract. TERRA molecules have a variable size and are mainly localized in the nucleus. TERRA contains a 5'-m^7G cap, and was found to carry the m^6A modification (not depicted). Only a minor fraction of TERRA is polyadenylated (< 10%) – PolyA$^+$ TERRA. Polyadenylated TERRA is commonly detected in the nucleoplasm, while non-polyadenylated TERRA RNAs are often found in association with the telomeric chromatin, forming R-loops. These structures are formed when TERRA invades the telomeric dsDNA, base pairing with the C-rich strand – forming a DNA:RNA hybrid –, leaving a displaced DNA strand. The displaced G-rich ssDNA strand has the propensity to form G-quadruplex (G4) structures. TERRA has been proposed to regulate the recruitment or activity of telomerase – while some studies indicate that TERRA may inhibit telomerase, others suggest that TERRA guides telomerase to short telomeres for their preferential elongation. TERRA has also been implicated in the regulation of the chromatin composition of telomeres, serving as a recruitment platform for histone-modifying

enzymes, particularly when telomeres become short or uncapped. TERRA R-loops have been shown to interfere with the progression of the replication fork at telomeres, affecting their integrity. Additionally, TERRA R-loops can stimulate homology-directed recombination (HDR)-mediated telomere elongation. Given the relevance of TERRA R-loops in different biological contexts, numerous factors have been implicated in their regulation. Proteins promoting or counteracting the accumulation of TERRA R-loops are listed. *This illustration was created with BioRender.com.*

Elevated transcriptional levels of TERRA have been correlated with short telomere length in cells from patients affected by the rare autosomal recessive ICF (Immunodeficiency, Centromeric instability and Facial anomalies) syndrome type I. Increased TERRA has been explained by severe hypomethylation of subtelomeres as a consequence of mutations in the gene encoding the major DNA methyltransferase DNMT3B (Yehezkel *et al*, 2008), although it has been shown that loss of DNMT3B is not sufficient to upregulate TERRA, requiring co-depletion of DNMT1 (Nergadze *et al*, 2009). In ICF samples, TERRA cell cycle regulation is lost, resulting in high R-loop detection throughout the cell cycle, particularly involving the telomeric-hexameric repeats (Sagie *et al*, 2017). Remarkably, ICF cells display high DNA damage marks at chromosome ends, which are reduced by RNaseH1 overexpression. Therefore, the prominent DNA damage signaling detected at telomeres of ICF cells appears to arise as a consequence of accumulated TERRA R-loops across the cell cycle, presumably interfering with the progression of the replication machinery through telomeres, and resulting in the short telomere phenotype and premature replicative senescence characteristic of ICF.

As described in budding yeast and human ALT cells, TERRA R-loops act as an essential trigger for telomere elongation. On the other hand, in ICF it may pose as a harmful obstacle, inducing telomere shortening rather than contributing to recombination-mediated telomere extension. Given the contrasting outcomes of TERRA R-loop-derived replication stress, it raises the question on how sufficient TERRA R-loops are to modulate telomere length. Altogether, it suggests that the consequences of R-loop formation at telomeres are highly context-dependent – including telomere chromatin composition.

1.3. The THO and TREX complexes

1.3.1. THOC/TREX in yeast

Over 30 years ago, Andrés Aguilera and Hannah Klein identified an *hpr1 S. cerevisiae* mutant, which confers an increased intrachromosomal recombination phenotype, and a mild reduction in growth rate (Aguilera & Klein, 1990). Further characterization of *hpr1*Δ cells showed a transcriptional elongation defect at the engineered bacterial *lacZ* coding region – irrespectively of the type of promoter –, while no hurdles were detected in the transcription of an examined set of endogenous yeast genes. Depletion of *HPR1* did, however, also confer sensitivity to a transcription elongation inhibitor in endogenous yeast DNA sequences (Chávez & Aguilera, 1997).

In a quest to find yeast factors which – upon overexpression – would suppress the transcriptional defects observed in *hpr1*Δ, Tho2 was identified (as in Suppressor of Transcription Defects of *hpr1*Δ Mutants by Overexpression). Analysis of *tho2*Δ cells revealed that such mutants grow slower than wildtype and show a striking hyperrecombination phenotype (Piruat & Aguilera, 1998). Such effect in recombination seemed to arise from a general role for Tho2 in transcriptional elongation of yeast genes transcribed by RNA Polymerase II, independently from whether genes are located in plasmids or chromosomes, and whether they are constitutive or regulated (Piruat & Aguilera, 1998). Interestingly, the manifestations of *tho2*Δ mutation resembled those of *hpr1*Δ, suggesting Hpr1 and Tho2 act in the same biological process.

Indeed, the multisubunit THO complex (THOC) was readily identified as a functional unit in yeast (Chávez *et al*, 2000). This protein complex comprises Hpr1, Tho2, as well as Mft1 (mitochondrial fusion targeting 1), Thp2 (for *Tho2*/*Hpr1* phenotype) and Tex1 (TREX component 1) (Chávez *et al*, 2000; Pühringer *et al*, 2020) (Figure 1.3A). Expression of GFP-fusion proteins suggested that the THO complex

is mainly nuclear, and similarly to *hpr1*Δ and *tho2*Δ, *mft1*Δ and *thp2*Δ lead to impaired transcription through *lacZ*, and hyperrecombination dependent on transcription (Chávez *et al*, 2000). Concordantly, THOC components were found to interact with RNA Polymerase II in yeast (Chang *et al*, 1999) and mutations in the RNA Polymerase II transcriptional machinery were found to rescue the hyperrecombination phenotype characteristic of *hpr1*Δ cells (Fan *et al*, 1996).

Transcriptional defects of yeast THOC mutants were initially detected in G+C-rich genes – such as the bacterial *lacZ* reporter or the yeast *YAT1* –, or in long ORFs – such as the yeast *LYS2* fragment –, while only a slight defect was detected in the transcription of *LAC4* – an ortholog of *lacZ* with a lower G+C content (Chávez *et al*, 2001). Therefore, length and G+C content were proposed to be preferential elements influencing the requirement of THOC (namely Hpr1) for efficacious transcription in *S. cerevisiae* (Chávez *et al*, 2001). On the other hand, it was proposed that long internal tandem repeats were instead the key feature found in yeast genes that require THOC for transcription, as for example appropriate transcription of the wild-type *FLO11* yeast gene containing internal tandem repeats depends on a functional THO complex, while transcription of the same gene without the repeats is insensitive to THOC expression levels (Voynov *et al*, 2006). This is consistent with the fact that an *hpr1* mutant was primarily identified on the basis of increased recombination frequency between engineered tandem repeats (Aguilera & Klein, 1988). Additionally, yeast genes whose expression is most affected by depletion of THOC components do not necessarily exhibit a high G+C content nor a remarkable length (Voynov *et al*, 2006).

Along with the role of THOC in transcription, this oligomeric complex was found to play a part in RNA nuclear export as an element of the larger TREX complex (standing for transcription and export), together with Yra1 and Sub2 (Sträßer *et al*, 2002; Jimeno *et al*, 2002) (Figure 1.3A). All components of the THO complex were found to genetically interact with *SUB2* in a synthetic lethal screen with mutant *sub2* (Sträßer *et al*, 2002), which had in turn been found to be related to *YRA1* – two yeast genes encoding proteins which interact with each other and are involved in messenger RNA export (Sträßer & Hurt, 2000, 2001). In parallel, SUB2 was isolated as a multicopy suppressor of the transcriptional defect observed in *hpr1*Δ, and together with *yra1*Δ, *sub2*Δ cells displayed similar transcriptional problems and hyperrecombination phenotypes as THOC mutants (Jimeno *et al*, 2002). Additionally, Yra1, Sub2 and THOC components were found to physically interact independently of RNA, suggesting that – as previously described (Lei *et al*, 2001) – export factors are recruited to RNAs in a transcription-coupled manner (Sträßer *et al*, 2002).

While Sub2 had previously been implicated in the spliceosome assembly process (Libri *et al*, 2001; Zhang & Green, 2001), it was proposed to be involved in recruiting Yra1 to both intron-containing and intron-lacking mRNAs for export (Sträßer & Hurt, 2001). Likewise, THOC components were found to be required for efficient export of mRNAs derived from intronless genes (Sträßer *et al*, 2002; Rondón *et al*, 2003), being crucial for the efficient recruitment of Sub2 and Yra1 to elongating RNA Polymerase II (Zenklusen *et al*, 2002).

Yeast cells harbouring a mutated TREX complex show an irregular nuclear export of polyadenylated mRNAs, thought to derive from abnormal 3'end mRNA processing. Interestingly, in the case of a pool of heat shock *HSP104* transcripts, the observed 3'end-truncation in TREX mutant strains was suggested to stem from nuclear RNA degradation, rather than from a transcription elongation defect – as previously described for other transcripts regulated by THOC/TREX. In this case, incomplete RNAs result from degradation by an exosome-associated exonuclease, while RNAs that are not degraded remain trapped at the transcription site (Libri *et al*, 2002; Zenklusen *et al*, 2002; Saguez *et al*, 2008).

Importantly, strains lacking a functional THO complex, exhibit a striking accumulation of genome-wide DNA:RNA hybrids, as determined by the sensitivity of nucleic acids extracted from THOC mutant cells to different nucleases (Huertas & Aguilera, 2003). Such structures were found to accumulate toward the 3' end of genes (Huertas & Aguilera, 2003), and when formed co-transcriptionally pose an obstacle to elongating RNA Polymerase II, impairing transcription elongation (Bentin *et al*, 2005; Tous & Aguilera, 2007; Gómez-González *et al*, 2011). Additionally, if a hammerhead ribozyme self-cleaves nascent RNAs during transcription, the hyperrecombination phenotype observed in THOC mutants is

suppressed. Therefore, DNA:RNA hybrids are thought to be one of the main sources of so-called transcription-associated recombination (Huertas & Aguilera, 2003). Such hyperrecombination has been proposed to stem from stalling of the replication fork (Gan *et al*, 2011b; Wellinger *et al*, 2006) or from RNA polymerase obstruction by DNA:RNA hybrids. In addition, the vulnerability of ssDNA stretches formed in the displaced DNA strand of R-loops to damage agents, inducing the formation of DNA breaks, is also thought to contribute to recombinogenic events (Huertas & Aguilera, 2003). Moreover, detailed analysis of DNA:RNA hybrids in THOC mutants throughout the cell cycle suggested that the THO complex counteracts R-loop accumulation in G1 and S-phases of the cell cycle (San Martin-Alonso *et al*, 2021). Overall, TAR is strongly associated to genetic instability, placing the THOC/TREX complex as a key player in the maintenance of genome integrity.

Unexpectedly, overexpression of the TREX adaptor protein Yra1 was shown to increase DNA:RNA hybrids – presumably by binding and stabilizing such structures –, thus contributing to transcription-replication collisions (Gavaldá *et al*, 2016; García-Rubio *et al*, 2018). Therefore, overriding Yra1 tight stoichiometric regulation (Rodríguez-Navarro *et al*, 2002; Preker & Guthrie, 2006) leads to genomic instability and has a negative impact on cell viability in yeast (Gavaldá *et al*, 2016; García-Rubio *et al*, 2018).

Altogether, the yeast TREX complex seems to be recruited to transcribing genes, independently from splicing, traveling with elongating RNA polymerase II. Additionally, it couples transcription with appropriate RNA processing, counteracting the accumulation of DNA:RNA hybrids, and contributing to appropriate RNA export from the nucleus to the cytoplasm (role in nuclear export reviewed in (Iglesias & Stutz, 2008)).

1.3.2. THOC/TREX in human

Upon the discovery of the THOC/TREX complex in yeast, a search for structural and functional homologues in human cells was launched. The human THO complex was characterized as comprising six essential subunits (Figure 1.3A). It includes THOC1, which had previously been discovered as a protein associated with the retinoblastoma tumour suppressor protein (Rb1) (Durfee *et al*, 1994), and which was identified as a member of human THOC on the basis of resemblance with the yeast homologue Hpr1 (Li *et al*, 2005). Resorting to human protein data bases, THOC2 was found to be a homologue of yeast Tho2 (West *et al*, 2000) – where the C-terminal domain of yeast Tho2 is thought to be essential for nucleic acid binding (Peña *et al*, 2012). In addition, THOC3 was identified as a homologue of yeast Tex1 (Söding *et al*, 2005). More recently, the homology between yeast Thp2 and Mft1 with human THOC5 and THOC7, respectively, was determined by structural analysis of yeast and human complexes (Pühringer *et al*, 2020). THOC6 – originally identified as part of THOC in *Drosophila* cells (Rehwinkel *et al*, 2004) – is also a component of the human THO complex, with no identified yeast homologue. Similarly to yeast, human THOC was found to be included in the larger TREX complex, together with the splicing factor UAP56/DDX39B (a homologue of yeast Sub2), and the RNA nuclear export protein Aly/ALYREF/THOC4 (homologue of Yra1) (Fleckner *et al*, 1997; Zhang & Green, 2001; Zhou *et al*, 2000; Luo *et al*, 2001) (Figure 1.3A). Taken together, the human TREX complex assembles as a tetramer, with four units of each composing subunit (Figure 1.3B) (while it seems to form a dimer in yeast) (Pühringer *et al*, 2020). Excitingly, there seems to be a high degree of structural evolutionary conservation of TREX proteins, consistent with a key function in genome maintenance across species.

The role of THOC/TREX in human cells was described to resemble, to some extent, that observed in *S. cerevisiae*. Namely, a correlation was established between THOC1 or THOC2 depletion and a decreased growth rate, in different analysed human cell lines – namely HeLa, Hek293 and U2OS cells (Li *et al*, 2005). Also in human cells, THOC was found to interact with elongating RNA Polymerase II (Li *et al*, 2005). siRNA-mediated depletion of subunits THOC1 or THOC2 was shown to impact the expression of some endogenous human genes, as well as the reporter gene *lacZ*. *lacZ* expression was reduced upon THOC1 depletion, irrespectively of the promoter driving *lacZ* expression in human cells, indicating that the observed transcriptional defect is presumably not due to a faulty promoter-

dependent transcriptional initiation. Additionally, upon THOC1 depletion, the density of RNA Polymerase II along the *lacZ* gene was reduced (Li *et al*, 2005). These data, together with the use of a tandem transcriptional elongation reporter system (Domínguez-Sánchez *et al*, 2011), are consistent with a role of human THOC in transcriptional elongation.

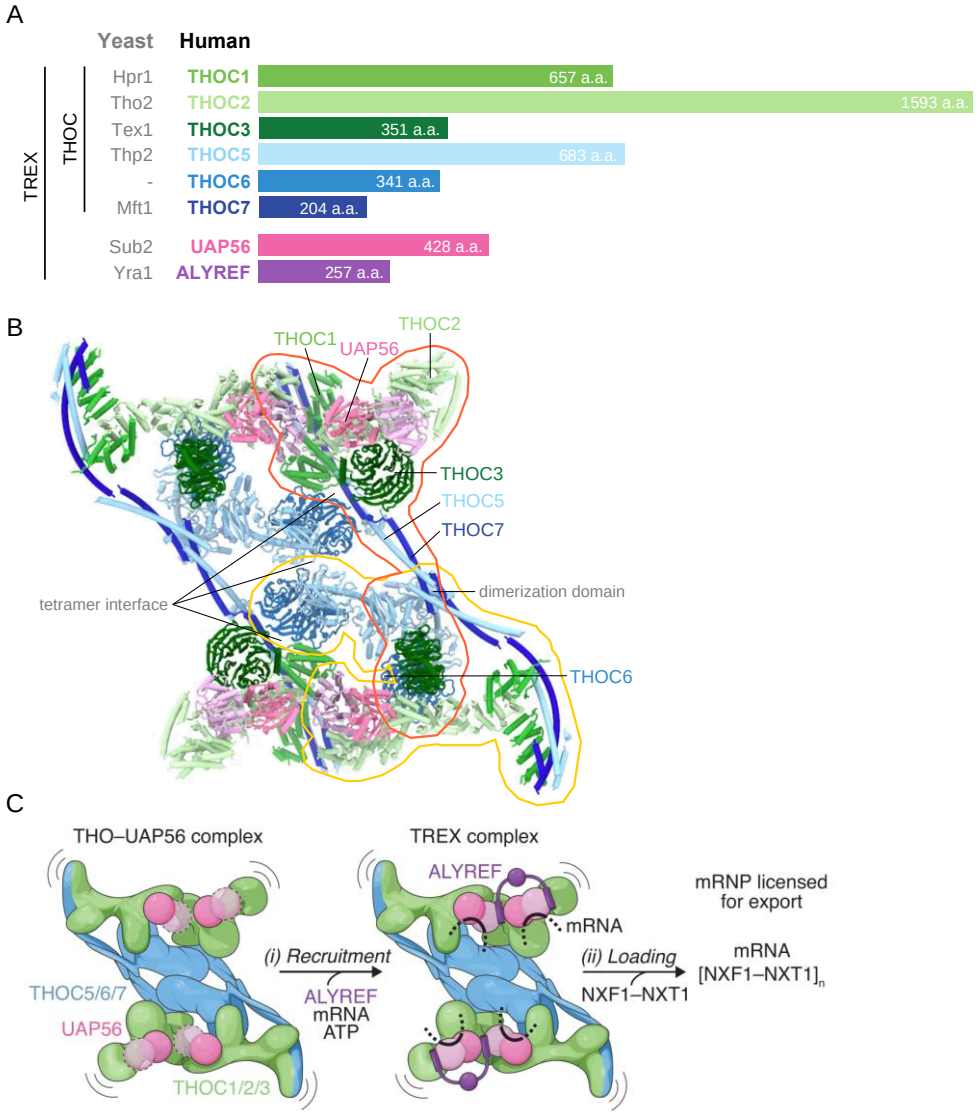


Figure 1.3. Composition and structure of the THO and TREX complexes.

A Subunits of THO and TREX complexes. Human proteins are colour-coded matching the structures depicted in B and C. Respective homologue yeast proteins are indicated in grey. The number of amino acids (a.a.) of each human protein is indicated, and each bar's length attempts to reflect the relative size between components.

B Front view of the human 14-subunit THOC-UAP56 asymmetric dimers, forming a tetramer with 28 subunits. Red and yellow outlines highlight a monomer each. Seven main subunits of TREX are labelled in one of the four monomers (monomer outlined in red). A dimerization domain and the tetramer interface are indicated. The obtained structure was modelled from cryo-electron microscopy and combined with a previously obtained low-resolution yeast THOC-Sub2 crystal structure (for UAP56 RecA1 lobe depicted in light pink). Depicted THOC2 subunits only comprise residues 1-1203, since the disordered C-terminal domain was truncated for recombinant protein expression. The TREX component ALYREF was not modelled in this structure.

C Depiction of TREX structure and its proposed function in nuclear export. The THO complex binds UAP56, forming a tetramer. It is recruited to RNA molecules (depicted as black lines) and engages the

nuclear export adaptor ALYREF (i), and presumably other adaptor proteins. In turn, it may promote binding of other factors, licensing the RNA for nuclear export (ii). The TREX complex depiction including ALYREF (here represented in purple) was obtained by preparing a homology model of UAP56 bound to RNA and ALYREF, based on a previously obtained yeast crystal structure, and superimposing this model onto the human THO–UAP56 tetramer structure.

Figure modified from Pühringer *et al*, 2020.

Remarkably, the role of THOC within the TREX complex in coupling transcription with RNA export seems to be conserved from yeast to human, however, in higher eukaryotes it is thought to depend on RNA splicing (which occurs co-transcriptionally). In contrast to yeast, human THOC was proposed to be recruited to elongating transcription in an indirect manner – via the splicing machinery. I*n vitro* experiments have suggested that human THOC exclusively associates with spliced RNA, but not unspliced pre-mRNA (Masuda *et al*, 2005). Additionally, immunofluorescence staining experiments revealed that THOC shows a diffuse nuclear pattern and localizes with splicing factors in nuclear speckle domains (where splicing takes place) (Masuda *et al*, 2005). In order to assemble the larger TREX complex, THOC associates with the splicing factor UAP56 (Li *et al*, 2005), which interacts with Aly in an ATP-dependent manner (Taniguchi & Ohno, 2008), which in turn are thought to associate with the exon junction complex (EJC) (Merz *et al*, 2007). In contrast, THOC1, UAP56 and Aly were also found to interact with unspliced, 5' capped AdML RNA (Merz *et al*, 2007). In line with this observation, TREX is thought to be recruited to RNA not only in a splicing-dependent manner, but also via the RNA 5' cap site, through THOC and Aly association with the cap-binding complex (CBC) component CBP80 (Cheng *et al*, 2006; Chi *et al*, 2013). In addition, TREX components were recently found to interact with the m^6A methyltransferase complex. This places the m^6A modification of RNA as yet another feature presumably contributing to the recruitment of TREX proteins to RNA – a process which was shown to be crucial for efficient RNA export (Lesbirel *et al*, 2018).

Following recruitment to RNA, TREX becomes essential for appropriate RNP (ribonucleoprotein complex) formation, engaging the export factor NXF1-NXT1 and sustaining RNA nuclear export (Figure 1.3C) (reviewed in (Köhler & Hurt, 2007; Carmody & Wente, 2009)). Accordingly, depletion of THOC or TREX components results in a noticeable nuclear accumulation of polyadenylated RNAs (Domínguez-Sánchez *et al*, 2011; Chi *et al*, 2013).

The first evidence of a role of human THOC in the regulation of DNA:RNA hybrids in human cells arose from the quantification of the reporter *lacZ* RNA levels in nuclease-treated nucleic acid extracts from human cells. Analysis of nucleic acids from cells depleted of THOC1 showed an increased sensitivity to RNaseH digestion (which specifically degrades the RNA moiety of DNA:RNA hybrids) prior to reverse transcription, compared to extracts from control cells (Li *et al*, 2005). Furthermore, downregulation of THOC subunits led to an accumulation of DNA damage, as examined by an increase in tail moment in comet assay (single-cell electrophoresis), and a higher percentage of cells displaying γH2AX and 53BP1 foci. Importantly, staining of such DNA damage markers in cells with reduced THOC levels was sensitive to RNaseH1 and 2, indicating that DNA breaks are prompted by R-loops linked to loss of THOC (Domínguez-Sánchez *et al*, 2011). Analogously to the hyperrecombination phenotype documented in yeast THOC mutants – although to a lower extent –, depletion of THOC subunits in human cells induces an elevated recombination frequency, as measured with a direct-repeat recombination reporter construct (Domínguez-Sánchez *et al*, 2011). Intriguingly, DNA combing experiments suggest that loss of THOC1 results in apparently longer replicons. This was interpreted as a possible failure in triggering DNA damage response checkpoint, or as a direct consequence of delayed replication termination, or even as an indirect outcome of a lower density of RNA Polymerase II, due to unsuccessful transcriptional elongation (Domínguez-Sánchez *et al*, 2011). Overall, this underscores the role of human THOC/TREX in the maintenance of genome stability, by counteracting the accumulation of DNA:RNA hybrids.

Nevertheless, the exact mechanism through which THOC/TREX prevents the accumulation of DNA:RNA hybrids is still not well elucidated. The observation that downregulation of several RNA-

binding and RNA-processing proteins contributes to the accumulation of R-loops is consistent with the notion that inadequate formation of RNP particles prompts RNA hybridization with DNA (reviewed in (Aguilera, 2005) – which is most prone to occur co-transcriptionally, due to sequence complementarity and an open chromatin state during transcriptional elongation.

In addition, a relationship between THOC and chromatin modifiers unravelled a new putative explanation for the mechanism undertaken by THOC in restraining R-loops (Salas-Armenteros *et al*, 2017). In this study, THOC1 was shown to physically interact with components of the Sin3 deacetylase complex. Depletion of THOC1 led to a decrease in histone deacetylase activity and a tenuous increase in cellular histone acetylation levels. In turn, depletion of Sin3A scaffold histone deacetylase or chemical inhibition of histone deacetylation led to elevated DNA:RNA hybrid levels, as well as DNA damage levels – which are dependent on active transcription and sensitive to RNaseH1 overexpression, thus likely being a consequence of R-loop build-up. In parallel, a histone acetyltransferase inhibitor reduced DNA:RNA hybrid levels, as well as cellular DNA damage detection by γH2AX staining and comet assay. Altogether it was proposed that the THO complex directs histone hypo-acetylation via Sin3A, transiently closing chromatin during transcriptional elongation, and thus preventing co-transcriptional accumulation of R-loops (Salas-Armenteros *et al*, 2017). Nonetheless, while deacetylase activity was decreased upon THOC1 and/or Sin3A siRNA-mediated downregulation, analysis of AcH3 levels at specific genomic loci did not fully correlate with R-loop accumulation under such conditions (Salas-Armenteros *et al*, 2017). Therefore, further studies will be required to conclusively determine whether chromatin modifications form the direct basis for the role of THOC in preventing R-loops.

Along with the notion that THOC/TREX can offset the accumulation of R-loops by preventing the hybridization of RNA with template DNA in the first place – either by directly binding the RNA, and/or by directing chromatin accessibility –, it has been shown that the core TREX component UAP56 has the capacity to resolve unscheduled R-loops (Pérez-Calero *et al*, 2020). UAP56 is a DEAD-box RNA-dependent ATPase (Shen *et al*, 2007), which was shown to be recruited to most transcribed DNA regions – possibly dependently on splicing (Luo *et al*, 2001). *In vitro*, UAP56 displays a strong DNA:RNA helicase activity, unwinding blunt-ended hybrids and hybrids displaying 5' or 3' overhangs, as well as an RNA-DNA flap structure which mimics an R-loop. Accordingly, *in vivo* experiments showed that depletion of UAP56 increases R-loops genome-wide, which is accompanied by elevated DNA damage levels, specifically dependent on active transcription and R-loops (as γH2AX staining and comet assay revealed a sensitivity to the transcription elongation inhibitor cordycepin and to RNaseH1 overexpression in UAP56 depleted condition). Furthermore, loss of UAP56 led to an increase in γH2AX most prominently in S and G2 phases of the cell cycle, as well as a lower replication fork velocity, and a higher fork asymmetry rescued by RNaseH1 overexpression, indicative of replication defects as a consequence of R-loop persistence (Pérez-Calero *et al*, 2020). Interestingly – as shown for THO subunit 1 (Salas-Armenteros *et al*, 2017) –, UAP56 interacts with the Sin3A scaffold histone deacetylase, and co-depletion of UAP56 and Sin3A results in elevated R-loop levels, compared to control cells, although to a lower extent than individual depletions (Pérez-Calero *et al*, 2020).

1.3.3. Involvement of THOC/TREX in development and pathological contexts: a brief overview

Given that the highly conserved THOC/TREX participates in very fundamental molecular processes – including RNA biogenesis and export –, it is anticipated that these complexes are involved, to some degree, in developmental and pathological contexts, across different species.

While the TREX component Yra1 is essential for viability in yeast cells (Portman *et al*, 1997), strains with loss of THOC components are viable, despite a slight decrease in growth rate (Aguilera & Klein, 1990; Piruat & Aguilera, 1998). On the other hand, a different scenario is observed in mammals. In mice, homozygous THOC1 null embryos are not recovered, indicating that THOC1 is required for early embryonic development (Wang *et al*, 2006). Particularly, development is blocked upon implantation, and *in vitro* culture of embryos results in fundamental defects (Wang *et al*, 2006). In addition, in mouse

embryonic stem-cells, THOC subunit 5 was shown to preferentially interact with pluripotency-related RNAs such as Nanog and Sox2 (Wang *et al*, 2013). Through its role in RNA export, THOC5 is thought to contribute to embryonic stem cell self-renewal, while at the onset of differentiation, THOC5 is downregulated, thus leading to a depletion of pluripotency-related factors, facilitating exit from self-renewal (Wang *et al*, 2013). Interestingly, inducible depletion of THOC1 in adult mice appears to have a tissue/cell type-dependent effect. For example, at the gastro-intestinal tract, it is small intestine cells which seem to be the most disrupted, with compromised proliferation and viability upon inducible downregulation of THOC1 (Pitzonka *et al*, 2013). Also, at the hematopoietic system, specifically myeloid lineages appear to have the most pronounced sensitivity to loss of THOC1, presumably owing to their high rate of proliferation (Pitzonka *et al*, 2014). In the male mouse, THOC1 deficiency compromises gametogenesis – likely as a consequence of defects in the regulation of transcripts involved in postnatal differentiation of testes –, rendering the animals infertile (Wang *et al*, 2009). In human, THOC5 was implicated in osteoclast differentiation (Mun *et al*, 2022) – a process required for constant maintenance and repair of bones. Therefore, the impact of THOC in adult tissue appears to be highly context specific, possibly because different subsets of RNAs are regulated by THOC in different tissues/cell types.

Intriguingly, it has been hypothesized that some cancer cells require higher levels of THOC components for their viability than cells which have not undergone neoplastic transformation (Li *et al*, 2007). For instance, THOC1 is overexpressed in breast cancer cells, compared to normal epithelial cells, and its levels correlate with tumour size and metastases (Guo *et al*, 2005). Loss of THOC1 impaired the growth and viability of oncogene-transformed human fibroblasts – presumably by promoting apoptosis –, while control fibroblasts were not affected. Also in human epithelial cancer cell lines – including HeLa cells – viability is slightly affected and γH2AX cellular levels are increased upon THOC1 downregulation (Li *et al*, 2007). Additionally, in hepatocellular carcinoma, THOC1 is outstandingly upregulated, and is thought to promote proliferation, contributing to tumorigenesis (Cai *et al*, 2020). Importantly, THOC1 loss was shown to sensitize hepatocellular carcinoma cells to the chemotherapeutic cisplatin (Cai *et al*, 2020), raising the possibility of exploring THOC subunits as potential therapeutic targets. In contrast, overexpression of THOC1 in lung cancer cell lines, was demonstrated to inhibit cell proliferation and promote G2/M cell cycle arrest, culminating in apoptosis, likely by affecting the cell cycle regulators cyclin A1 and B1, and pro-apoptotic factors Bax and caspase-3 (Wan *et al*, 2014).

Overall, considering the extent of biological contexts THOC/TREX has been implicated in, the elucidation of its molecular function is of great relevance.

1.3.4. THOC/TREX at yeast and human telomeres

A few studies have been undertaken to decipher a putative (direct or indirect) role of THOC components in telomere biology. For example, *S. cerevisiae hpr1* or *tho2* mutants were shown to have longer telomeres. This elongation was shown to occur in a telomerase-dependent manner, and irrespectively of recombination, since cells lacking Rad52 (crucial for homologous recombination) still displayed such elongation (Yu *et al*, 2012). Interestingly, this effect in telomere length caused by loss of Hpr1 and Tho2 was proposed to result from a decrease in Rif1 mRNA and protein levels (Yu *et al*, 2012) – a factor which interacts with the shelterin Rap1 and is a negative regulator of telomere length (Hardy *et al*, 1992). This study did not evaluate the presence of THOC components at yeast telomeres, and reports what seems to be an indirect consequence of loss of THOC, given its role in global RNA biogenesis.

Also the impact of the TREX component Sub2 was briefly evaluated at yeast telomeres. Overexpression of Sub2 was proposed to rescue telomeric silencing, as measured by a colony survival assay reflecting the expression level of the telomeric positioning effect reporter URA3 (Lahue *et al*, 2005). While this can result from a direct or indirect pathway (for example by impacting other telomere-related factors, as described above for some THOC subunits), Sub2 proteins were in fact shown to associate with yeast telomeres (Lahue *et al*, 2005), recognizing also the possibility of a direct role of THOC/TREX at telomeres.

In line with this notion, ChIP experiments in *S. cerevisiae* cells revealed that THOC subunits Hpr1, Tho2, Mft1 and Thp2 are found at most chromosome ends (Pfeiffer *et al*, 2013). Interestingly, deletion of *HPR1* or *THP2* increased the levels of TERRA-R-loops detected at yeast telomeres, without significantly changing TERRA levels (Balk *et al*, 2013; Pfeiffer *et al*, 2013). Furthermore – and in contrast to the study of Yu and colleagues (Yu *et al*, 2012) regarding loss of *HPR1* and *THO2* –, *THP2* deletion resulted in shortening of telomeres. However, this effect could not be rescued by RNAseH1 overexpression combined with *THP2* deletion. Instead, telomere shortening caused by loss of *THP2* was sensitive to deletion of the DNA specific Exonuclease 1 (*EXO1*), suggesting that Thp2 somehow protects telomeres from Exo1 (Pfeiffer *et al*, 2013). In addition, Thp2 is presumably required for efficient replication of telomeres when replication stress is induced (Pfeiffer *et al*, 2013). Overall, different pathways necessary for telomeric homeostasis may require the THO component Thp2, in a direct and/or indirect manner.

In parallel with the global role of THOC/TREX components in counteracting accumulation of R-loops, overexpression of the TREX adaptor protein Yra1 in *S. cerevisiae* cells actually leads to increased DNA:RNA hybrids across the genome, presumably by binding and stabilizing such structures (Gavaldá *et al*, 2016; García-Rubio *et al*, 2018). This effect was also observed at the telomeric tract, both in telomerase-positive and -negative strains. Despite the increased R-loop-mediated recombination frequency in Yra1 overexpressing cells, rapid shortening of telomeres was observed under these circumstances (independently from telomerase), as well as an accelerated senescence phenotype (Gavaldá *et al*, 2016; García-Rubio *et al*, 2018). Altogether, THOC/TREX components appear to play a role in proper maintenance of yeast telomeres, and disruption of different subunits appears to lead to highly diverse phenotypes.

Resorting to a mass spectrometry-based quantitative telomeric chromatin isolation protocol (QTIP) in human cells (HeLa), THO subunits 1-7, as well as the TREX component Aly were found to be associated with human telomeres (Grolimund *et al*, 2013), meaning THOC/TREX occupancy at telomeres seems to be evolutionarily conserved. Comparison between cell clones with long and short telomere length did not detect any differences in the telomeric occupancy of THOC/TREX components (Grolimund *et al*, 2013). The TREX component DDX39B was also identified at human telomeres (Hek293E cells) in a subsequent QTIP experiment (Lin *et al*, 2021).

In human cells, the TREX complex is generally recruited to elongating nascent RNA via the splicing apparatus (Masuda *et al*., 2005). However, while THOC/TREX components have been detected at human telomeres along with some splicing factors, there is thus far no evidence supporting that the telomeric RNA TERRA is spliced. It must be noted that this evaluation is delicate due to the repetitive nature of TERRA. Analysis of the subtelomeric-derived regions of TERRA (comprised between TERRA's 5' end and the start of the UUAGGG repetitive tract) did not reveal any splicing events in TERRA (Porro *et al*, 2014). Therefore, the criteria determining THOC/TREX recruitment to human telomeres and the function of THOC/TREX at chromosome ends requires further elucidation.

1.4. Thesis outline

The telomeric lncRNA TERRA has previously been implicated in several aspects concerning telomere maintenance and stability. Notably, important roles carried out by TERRA – including interference with replication of telomeric DNA, as well as stimulation of homologous recombination at telomeres – rely on direct DNA:RNA interactions with telomeric DNA. Therefore, this study aims at extending current knowledge on the regulation and impact of R-loops formed at chromosome ends.

Chapter 2 presents the work spearheaded by Marianna Feretzaki (Feretzaki *et al*, 2020), to which I contributed as a third author. This project entailed the development of a reporter system for the visualization of ectopically-expressed TERRA-like molecules, which enabled the scrutiny of the association of TERRA with telomeres. Namely, it allowed the identification of the domain within TERRA RNA that is necessary for the localization of TERRA to chromosome ends. Additionally, it was demonstrated that TERRA preferential association with telomeres inversely correlates with telomere

length, and that telomeric R-loops can form post-transcriptionally *in trans*, in a RAD51-mediated manner. The reporter system developed in this project and subsequent findings laid the foundation for the following chapters presented in this thesis.

Chapter 3 consists of a published article (Valador Fernandes *et al*, 2021) intended at placing the results described in Feretzaki *et al*, 2020 into context. In this article, we listed several regulators of telomeric R-loops, and discussed putative consequences of RAD51-mediated R-loop formation and accumulation at the telomeric tract. In addition, control and complementing experiments to the work presented in Chapter 2 were included, namely the analysis of the cell cycle distribution of RAD51-depleted cells, and evaluation of expression and telomeric localization of ectopically-expressed TERRA molecules during S-phase.

Chapter 4 comprises a published detailed description and illustration of the DNA:RNA immunoprecipitation protocol optimized for detection of R-loops at chromosome ends, either as a bulk – through dot blot analysis of immunoprecipitated telomeric DNA –, or at individual subtelomeres – by qPCR (Glousker, Valador Fernandes, Feretzaki *et al*, 2022).

The THO complex (THOC) has previously been identified at human telomeres. It couples transcription with RNA processing, restraining the accumulation of co-transcriptional DNA:RNA hybrids throughout the genome. In Chapter 5 we investigated how THOC impacts the association of TERRA with telomeres in human cells. Resorting to the TERRA reporter system described in Chapter 2, we explored the role of THOC in the post-transcriptional regulation of telomeric R-loops. We propose a model in which THOC plays a dual role at telomeres, where it directly prevents the association of TERRA with telomeric DNA – co- and post-transcriptionally – by binding it, and also contributes to the resolution of DNA:RNA hybrid structures formed at telomeres. Furthermore, we examined the impact of THOC in restricting telomere fragility, as well as telomeric sister chromatid exchange and C-circle accumulation, implicating the THO complex in telomere homeostasis.

Chapter 2 RAD51-dependent recruitment of TERRA lncRNA to telomeres through R-loops

Marianna Feretzaki, Michaela Pospisilova, Rita Valador Fernandes, Thomas Lunardi, Lumir Krejci and Joachim Lingner
Nature 587, pages 303–308 (2020)
https://doi.org/10.1038/s41586-020-2815-6
Reprinted with permission from all authors

2.1 Abstract

"Telomeres — repeated, noncoding nucleotide motifs and associated proteins that are found at the ends of eukaryotic chromosomes—mediate genome stability and determine cellular lifespan. Telomeric-repeat-containing RNA (TERRA) is a class of long noncoding RNAs (lncRNAs) that are transcribed from chromosome ends; these RNAs in turn regulate telomeric chromatin structure and telomere maintenance through the telomere-extending enzyme telomerase and homology-directed DNA repair. The mechanisms by which TERRA is recruited to chromosome ends remain poorly defined. Here we develop a reporter system with which to dissect the underlying mechanisms, and show that the UUAGGG repeats of TERRA are both necessary and sufficient to target TERRA to chromosome ends. TERRA preferentially associates with short telomeres through the formation of telomeric DNA–RNA hybrid (R-loop) structures that can form *in trans*. Telomere association and R-loop formation trigger telomere fragility and are promoted by the recombinase RAD51 and its interacting partner BRCA2, but counteracted by the RNA-surveillance factors RNaseH1 and TRF1. RAD51 physically interacts with TERRA and catalyses R-loop formation with TERRA *in vitro*, suggesting a direct involvement of this DNA recombinase in the recruitment of TERRA by strand invasion. Together, our findings reveal a RAD51-dependent pathway that governs TERRA-mediated R-loop formation after transcription, providing a mechanism for the recruitment of lncRNAs to new loci *in trans*."

2.2 Highlights

- Development of a reporter system which allows visualization of ectopically-expressed GFP-tagged TERRA-like molecules.
- The UUAGGG repetitive tract of TERRA is necessary and sufficient for the localization of TERRA to chromosome ends.
- Association of TERRA with telomeres occurs preferentially at telomeres with shorter telomere length.
- Ectopically-expressed TERRA molecules can hybridize with telomeres, forming DNA-RNA hybrids (R-loops) *in trans*.
- The recruitment of TERRA to telomeres and formation of telomeric R-loops depend on the DNA recombinase RAD51 and its interacting partner BRCA2.
- RNA-surveillance factors, DNA:RNA hybrid-specific endonuclease RNaseH1 and shelterin component TRF1 counteract TERRA association with chromosome ends.
- RAD51 binds TERRA in nuclear extracts and *in vitro*.
- RAD51 catalyses strand invasion of telomeric DNA *in vitro*.

2.3 Author contributions

"M.F. and J.L. conceived the study. M.F. and R.V.F. executed all cell and molecular biology experiments. M.P. performed all biochemistry experiments and T.L. some EMSA experiments. L.K.

conceived the RAD51-based biochemistry experiments and advised on the text. J.L. and M.F. wrote the paper."

Detailed contributions of R.V.F:

- Performed amplification and purification of PP7-TERRA constructs and respective controls.
- Performed amplification and purification of wildtype and II3A mutant RAD51-expressing constructs.
- Contributed to cell and molecular biology experiments, including cell culture, steps in immunofluorescence experiments, western blotting, and optimization of DRIP-qPCR.
- Performed experiments requested by reviewers that did not arrive to conclusive results and were consequently not included in the article. These experiments included evaluation of the effect of BRCA2 depletion, as well as RAD51 strand-invasion mutant expression in telomeric R-loops by DRIP dot-blot in cells with longer and shorter telomere length; troubleshooting of siRNA-mediated depletion of HOP2 and MND1; design of guide RNAs for CRISPR-Cas9-mediated downregulation of HOP2 and MND1.
- Re-organized main and extended figures, adjusted figure legends and prepared source data documents for the manuscript's final submission.

Article

RAD51-dependent recruitment of TERRA lncRNA to telomeres through R-loops

https://doi.org/10.1038/s41586-020-2815-6

Received: 13 October 2019

Accepted: 16 July 2020

Published online: 14 October 2020

Check for updates

Marianna Feretzaki[1], Michaela Pospisilova[2], Rita Valador Fernandes[1], Thomas Lunardi[1], Lumir Krejci[2,3]✉ & Joachim Lingner[1]✉

Telomeres—repeated, noncoding nucleotide motifs and associated proteins that are found at the ends of eukaryotic chromosomes—mediate genome stability and determine cellular lifespan[1]. Telomeric-repeat-containing RNA (TERRA) is a class of long noncoding RNAs (lncRNAs) that are transcribed from chromosome ends[2,3]; these RNAs in turn regulate telomeric chromatin structure and telomere maintenance through the telomere-extending enzyme telomerase[4–6] and homology-directed DNA repair[7,8]. The mechanisms by which TERRA is recruited to chromosome ends remain poorly defined. Here we develop a reporter system with which to dissect the underlying mechanisms, and show that the UUAGGG repeats of TERRA are both necessary and sufficient to target TERRA to chromosome ends. TERRA preferentially associates with short telomeres through the formation of telomeric DNA–RNA hybrid (R-loop) structures that can form in *trans*. Telomere association and R-loop formation trigger telomere fragility and are promoted by the recombinase RAD51 and its interacting partner BRCA2, but counteracted by the RNA-surveillance factors RNaseH1 and TRF1. RAD51 physically interacts with TERRA and catalyses R-loop formation with TERRA in vitro, suggesting a direct involvement of this DNA recombinase in the recruitment of TERRA by strand invasion. Together, our findings reveal a RAD51-dependent pathway that governs TERRA-mediated R-loop formation after transcription, providing a mechanism for the recruitment of lncRNAs to new loci in *trans*.

TERRA is transcribed from numerous chromosome ends, and comprises both subtelomeric sequences and telomeric repeats. More than 50% of TERRA is associated with chromatin[9]. To investigate how TERRA is recruited to or retained at telomeres, we generated a plasmid encoding 24 copies of the stem-loop of phage PP7 (ref. [10]) under the control of the tetracycline-inducible (TET) promoter, followed by 90 TTAGGG repeats (Fig. 1a). To generate full-length TERRA transcripts, we also cloned the human chromosome Xq and 15q subtelomeric regions containing the TERRA start sites between the PP7 stem-loops and the TTAGGG repeats. The constructs were then transiently transfected into HeLa clones that were constitutively expressing the PP7 coat protein fused to GFP (PCP–GFP) and a nuclear-localization signal. PCP–GFP exhibited a diffuse signal in the nucleus but formed nuclear foci upon expression of the PP7 stem-loops, which are bound by PCP and can gather up to 48 PCP–GFP molecules per RNA. These foci did not co-localize with telomeres (Fig. 1b). The fusion of the subtelomeric region of 15q or Xq TERRA to the stem-loops did not promote substantial trafficking of the PP7 foci to telomeres. However, when the telomeric TTAGGG repeats were fused downstream of PP7, co-localization with telomeres occurred, as analysed by conventional and confocal imaging (Fig. 1b), indicating that the 5′-UUAGGG-3′ repeats of TERRA drive telomere association. The full-length PP7-tagged 15q and Xq chimaeric TERRA also showed marked co-localization with telomeres (Fig. 1b and Extended Data Fig. 1a). Therefore, chimaeric TERRAs that originated from a plasmid were directed to telomeres in *trans*.

To eliminate possible confounding effects due to the high plasmid copy number or increased levels of transgenic TERRA, we used CRISPR–Cas9 technology to integrate the chimaeric TERRA constructs into the genome at the adeno-associated-virus integration site 1 (AAVS1) on chromosome 19, which represents a safe harbour for transgene expression[11] (Extended Data Fig. 1b). Following isolation of clones, we confirmed monoallelic site-specific integration of the full constructs by polymerase chain reaction (PCR) and sequencing. These TERRA expression levels were lower than the levels of expression from plasmids, giving one to three foci—indicative of displacement from the transcription site. But, similar to the results obtained upon transient transfection, the PP7 loops formed nuclear foci, and only when fused to 5′-UUAGGG-3′ repeats did the chimaeric RNAs co-localize with telomeres (Extended Data Fig. 1c, d).

Shorter telomeres recruit more TERRA

In *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, short telomeres recruit more TERRA, possibly to facilitate telomere maintenance through recombination or telomerase recruitment[5,6,8]. To explore the putative roles of telomere length in TERRA recruitment

[1]Swiss Institute for Experimental Cancer Research (ISREC), School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. [2]Department of Biology and National Centre for Biomolecular Research, Masaryk University, Brno, Czech Republic. [3]International Clinical Research Center, St Anne's University Hospital, Brno, Czech Republic.
✉e-mail: lkrejci@chemi.muni.cz; joachim.lingner@epfl.ch

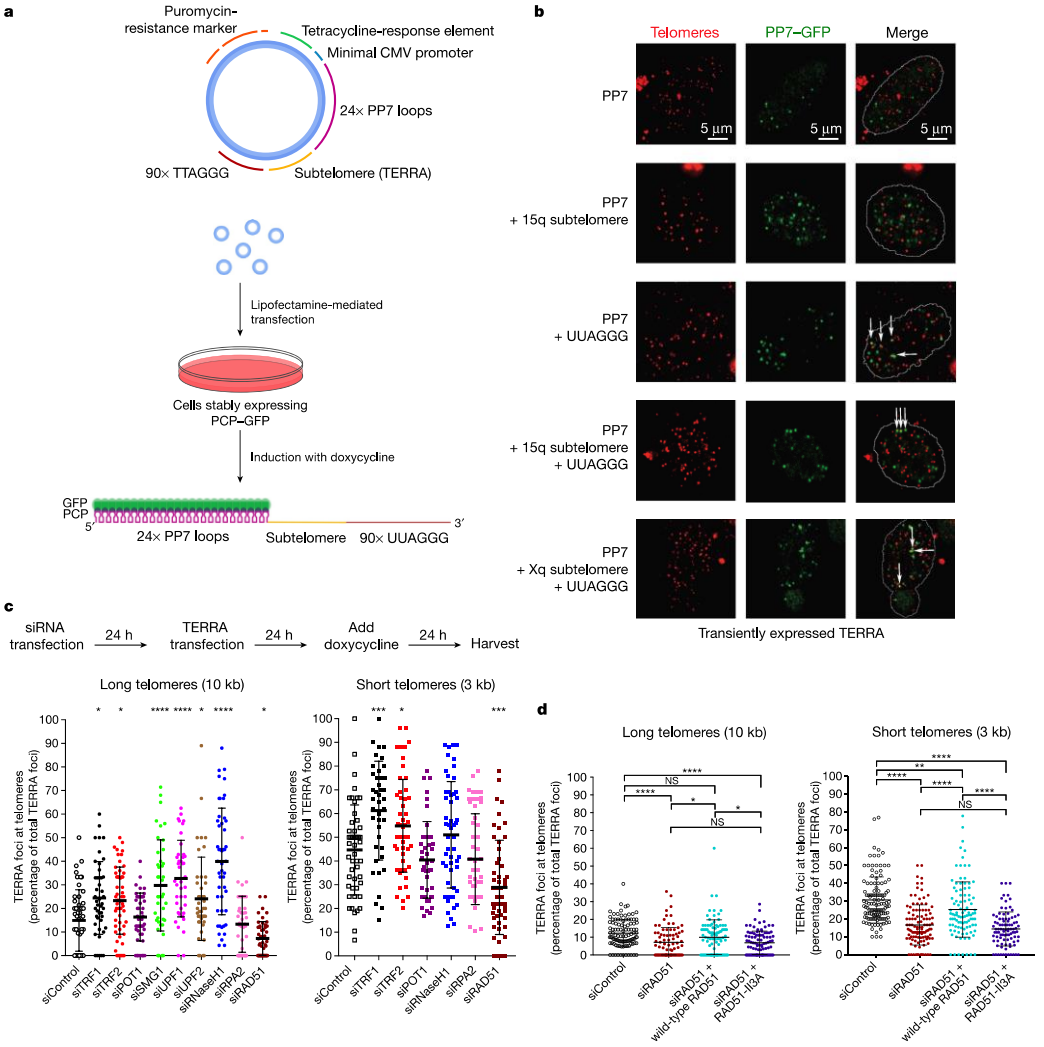


Fig. 1 | Transgenic TERRA associates with telomeres in a manner that depends on RAD51. a, Chimaeric TERRA construct (top) and assay for TERRA localization (bottom). Using lipofectamine, plasmids were transfected into HeLa cells constitutively expressing PCP–GFP protein. After 24 h, transcription of TERRA was induced with doxycycline. Chimaeric TERRA, which is recognized and bound by PCP–GFP, was analysed 24 h after induction. CMV, cytomegalovirus. **b**, Fluorescence in situ hybridization (FISH) analysis of telomeric DNA (red) and immunofluorescence of GFP (green) were used to assess the co-localization of transiently expressed PP7 constructs with telomeres. Confocal images are shown. White arrows indicate co-localization of PP7 foci with telomeric signals. **c**, To identify proteins involved in the localization of TERRA at telomeres, we used siRNAs to target factors implicated in TERRA and telomere biology. HeLa clones with long (10-kilobase average) and short (3 kb) telomeres were transfected first with siRNA pools and then with chimaeric TERRA. The percentage of co-localizing TERRA foci was assessed by telomeric FISH combined with GFP immunofluorescence. $n = 2$ biologically independent experiments; more than 40 nuclei were analysed per condition; data are means ± s.d. One-way analysis of variance (ANOVA) with Dunnett's multiple comparisons test was used, comparing all conditions with control siRNA (siControl): $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; $^{****}P < 0.0001$. **d**, The enzymatic activity of RAD51 is required to recruit TERRA. Endogenous RAD51 was depleted with siRNA, and wild-type RAD51 or the RAD51 II3A mutant was expressed from plasmid containing complementary DNA. The co-localization of TERRA with telomeres was assessed as in **c**. $n = 3$ biologically independent experiments; more than 80 nuclei were analysed per condition; data are means ± s.d. A two-tailed unpaired *t*-test was used to calculate *P*-values: $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; $^{****}P < 0.0001$.

in human cells, we isolated individual HeLa clones that constitutively expressed the PCP–GFP and measured telomere lengths by telomere restriction fragment length (TRF) analysis (Extended Data Fig. 2). Cells carrying short telomeres recruited TERRA much more efficiently than cells with long telomeres, as seen upon transient or stable expression of TERRA (Extended Data Fig. 2). Therefore, short telomeres must be

more accessible to recruitment or retention of TERRA; alternatively, long telomeres might contain active systems that expel TERRA. In both experiments the overall expression levels of chimaeric TERRA varied in individual clones, but this did not correlate with telomere length and telomere recruitment of TERRA (Extended Data Fig. 2b, d).

Recombination factors enable TERRA recruitment

To identify proteins involved in the localization of TERRA at telomeres, we performed screens using short interfering RNAs (siRNAs) to target selected factors implicated in TERRA and telomere biology. Using cell lines with long or short telomeres, we transfected first siRNA pools and then the 15q TERRA construct. After inducing TERRA expression with doxycycline, we analysed the cells (Fig. 1c) and evaluated the level of depletion of each factor by using reverse transcription with quantitative PCR (RT–qPCR) or western blotting (Extended Data Fig. 3a, b). We found that the number of transgenic TERRA foci was not affected by individual depletions, and we observed no striking effects on levels of selected endogenous TERRA molecules (Extended Data Fig. 3c, d). Among tested factors, depletion of telomeric repeat factor 1 (TRF1) significantly increased TERRA co-localization at short and long telomeres, while removal of TRF2 led to a milder increase in recruitment (Fig. 1c). Depletion of the nonsense-mediated decay (NMD) factors also stimulated co-localization of TERRA at long telomeres, supporting their crucial role in displacing TERRA from chromosome ends[2]. Similarly, removal of RNaseH1 resulted in a substantial accumulation of TERRA at chromosome ends in cells with long telomeres (Fig. 1c). This result indicated that TERRA recruitment to, or retention at, long telomeres involves the formation of DNA–RNA hybrids. In cells with short telomeres, depletion of RNaseH1 only marginally increased the co-localization of TERRA with telomeres. The roles of NMD factors in cells with short telomeres could not be analysed, as their depletion caused cell death. Notably, depletion of RAD51—which facilitates strand invasion of DNA molecules during homology-directed repair (HDR)—led to a substantial decrease in TERRA recruitment to both long and short telomeres (Fig. 1c, d).

The involvement of RAD51 in TERRA recruitment prompted us to interrogate the role of the BRCA2 protein in TERRA trafficking. BRCA2 loads RAD51 and promotes the displacement of replication protein A (RPA), allowing the formation of stable filaments of RAD51 on single-stranded DNA (ssDNA) that are capable of homology search to facilitate HDR during double-strand-break repair; BRCA2 also protects stalled replication forks[12,13]. We found that depletion of BRCA2 led to a marginal decrease in TERRA recruitment at long, but more prominent reduction at short, telomeres (Extended Data Fig. 4a). However, removal of BRCA2 also diminished total RAD51 protein levels in both cell lines (Extended Data Fig. 4a). Finally, we tested whether RAD51 enzymatic activity is required for TERRA recruitment, taking advantage of a mutation (RAD51 II3A) that allows the protein to retain its DNA-binding but not its strand-invasion activity[14]. In siRAD51-treated cells, the expression of wild-type RAD51 from complementary DNA largely rescued TERRA co-localization with telomeres, but expression of RAD51 II3A did not (Fig. 1d and Extended Data Fig. 4b). Therefore, the enzymatic activity of RAD51 is required for TERRA to associate with telomeres. Overall, these data suggest that the HDR machinery promotes the recruitment of TERRA to telomeres.

TERRA forms R-loops causing telomere fragility

As the association of TERRA with long telomeres was increased upon depletion of RNaseH1, we hypothesized that the transgenic TERRA may form R-loops with telomeres in *trans*. To explore this possibility, we applied the DNA–RNA immunoprecipitation protocol (DRIP), in which the specificity of the S9.6 monoclonal antibody for eight to nine base pairs of DNA–RNA hybrids is exploited[15]. Precipitated nucleic acids were probed for telomeric repeats, and, as a control for specificity, isolated nucleic acids were treated in vitro with RNaseH1 before immunoprecipitation. Abolishment of the signal upon pretreatment with RNaseH1 confirmed the specificity of the assay for telomeric R-loops (Fig. 2a). As expected, we detected R-loops at telomeres in wild-type cells (Fig. 2a). Depletion of RNaseH1 led to an increase in the number of R-loops, while its overexpression to a decrease in R-loops at both long and short telomeres (Fig. 2a and Extended Data Fig. 5a). Overexpression of chimaeric TERRA further increased the frequency of R-loops in cells with both long and short telomeres (Fig. 2a); this frequency again increased upon depletion of RNaseH1 and decreased upon its overexpression (Fig. 2a and Extended Data Fig. 5a), indicating that the transgenic TERRA formed DNA–RNA hybrids.

The DRIP assay (Fig. 2a) could not distinguish to what extent the DRIP signals for transgenic TERRA were derived from R-loops forming within the transgenic plasmid during transcription, or R-loops forming after transcription in *trans* with telomeres. To measure R-loops specifically at telomeres, we used the DRIP samples derived from HeLa cells with long telomeres (Fig. 2a), and determined the presence of four specific chromosome ends by qPCR using subtelomere-specific primers residing in immediate proximity to the terminal 5′-TTAGGG-3′ repeats (Fig. 2b). Telomeric R-loops became detectable at the ends of all four chromosomes upon depletion of RNaseH1 (Fig. 2b). Specifically, 1q, 10q and 13q subtelomeric DNA increased strongly in abundance upon expression of transgenic PP7–15qTERRA, indicating that R-loops had formed at these telomeres with PP7–15qTERRA. The 15q subtelomeric signal was enhanced even more extensively, presumably because of R-loops forming with plasmid DNA containing the 15q sequence. As with wild-type TERRA, overexpression of RNaseH1 almost completely abolished the signals, whereas RNaseH1 depletion increased R-loop abundance. We sequenced the qPCR products (Extended Data Fig. 5b), verifying the identity of products for each chromosome end. Together, these data confirmed that transgenic PP7–15qTERRA associated with telomeres in *trans* through the formation of R-loop structures.

TERRA R-loops have been implicated in interfering with telomere replication[7,16–18]; this is manifested in telomere fragility[19], characterized by the accumulation of telomeric signals in metaphase chromosomes with a smeary or discontinuous appearance (Extended Data Fig. 6a). We found that transgenic TERRA increased telomere fragility (Fig. 2c and Extended Data Fig. 6b), which was suppressed by depletion of RAD51 or overexpression of RNaseH1 but increased by depletion of RNaseH1 (Fig. 2d and Extended Data Fig. 6c, d). These results confirm that, after transcription from plasmids, TERRA forms R-loops at telomeres in *trans*.

RAD51 promotes telomeric R-loop formation

We next used the DRIP assay to test the role of RAD51 in the formation of hybrids containing endogenous TERRA at telomeres in wild-type cell lines (Fig. 2e and Extended Data Fig. 7a). While depletion of RNaseH1 led to the expected mild increase in R-loops, depletion of RAD51 caused a substantial decrease in hybrid accumulation. Even stronger reduction of R-loops was observed in RAD51-depleted cells with short telomeres, which are characterized by higher levels of DNA–RNA hybrids. Therefore, RAD51 promotes the association of endogenous TERRA with telomeres through R-loop formation.

We hypothesized either that RAD51 binds TERRA to catalyse its strand invasion into the telomeric repeats (Extended Data Fig. 7b, lower panel), or that TERRA might hybridize to exposed single-stranded telomeric DNA during RAD51-mediated HDR between telomeric DNA molecules, even in the absence of a physical interaction between TERRA with RAD51 (Extended Data Fig. 7b, upper panel). To explore these hypotheses, we performed native RNA immunoprecipitations using anti-RAD51 antibodies in HeLa cells. As a control, we also included antibodies specific for hnRNPA1, as this protein binds TERRA[4]. Immunoprecipitation of

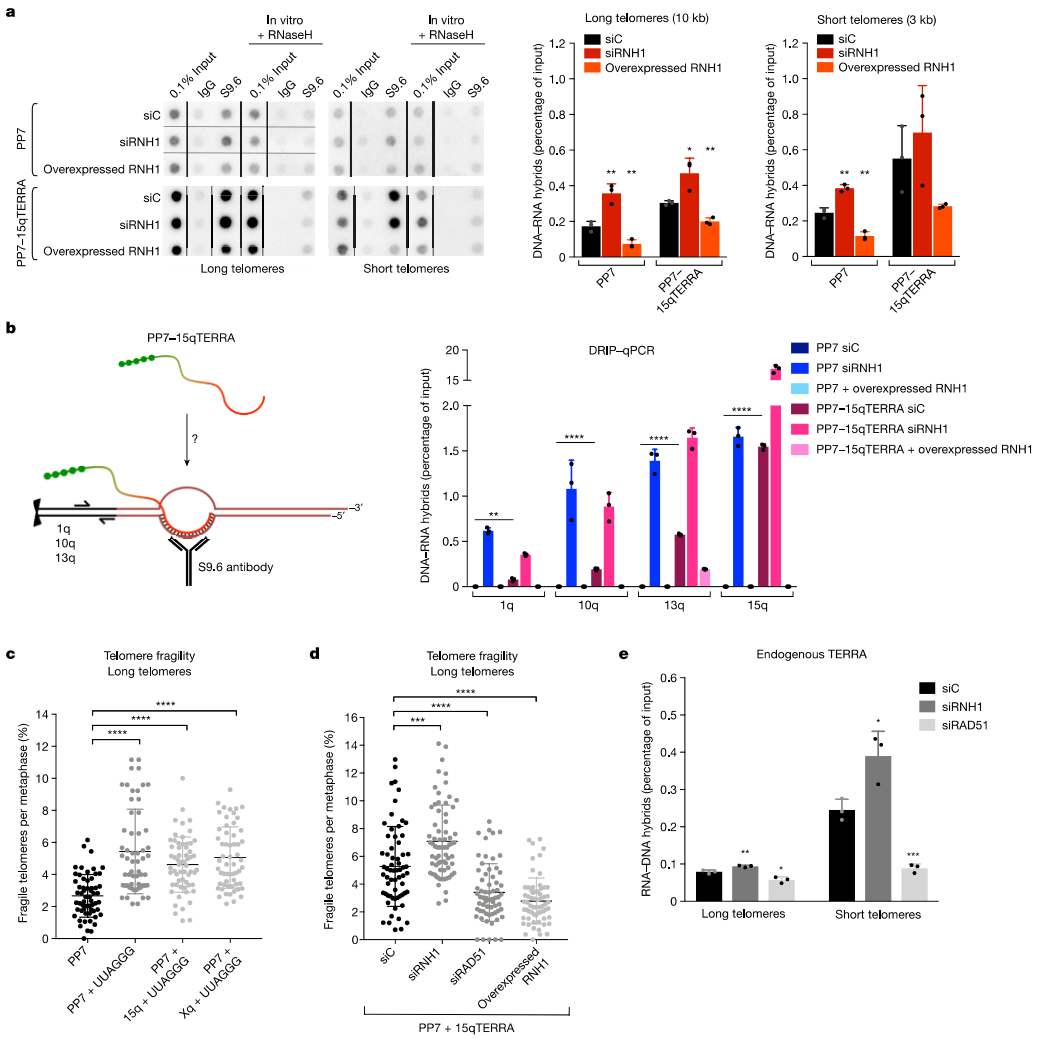


Fig. 2 | TERRA forms R-loops in *trans*, inducing telomere fragility in a RAD51-dependent way. a, Left panels, detection of telomeric hybrids on dot-blots. Telomeric hybrids were isolated from cells with long or short telomeres by DRIP using the S9.6 antibody. Cells expressed either the PP7 stem-loops or the PP7–15qTERRA, and had been transfected with control siRNA (siC), siRNAs against RNaseH1 (RNH1), or an RNH1-overexpressing plasmid. In vitro treatment with RNH1 served as a negative control. Right panels, quantification of signals. **b**, Left, schematic representation showing the expression of PP7–15qTERRA from a plasmid, forming R-loops in *trans* at telomeres 1q, 10q and 13q. Right, telomeric hybrids arising at the indicated chromosome ends were quantified by qPCR using primers that amplify specific subtelomeric DNA found next to the telomeric tracts in cells with long telomeres. **c**, Telomere fragility induced upon expression of TERRA from a plasmid. The fraction of fragile telomeres was quantified on metaphase chromosomes stained with a telomeric FISH probe (Extended Data Fig. 6). **d**, Effects of RNH1 and RAD51 on telomere fragility. In cells expressing PP7-15qTERRA from a plasmid, telomere fragility was quantified upon depletion or overexpression of RNH1, or depletion of RAD51. **e**, Detection of endogenous telomeric R-loops in cells transfected with siC, siRNH1 or siRAD51. For all panels, $n = 3$ biologically independent experiments; data are means ± s.d. Two-tailed unpaired *t*-tests were used to calculate *P*-values: *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$. For **c**, **d**, the numbers of metaphases scored are indicated for each condition in Extended Data Fig. 6b, c.

endogenous RAD51 specifically retrieved TERRA and not the nuclear U1 small nuclear RNA (snRNA; Fig. 3a and Extended Data Fig. 7c), and similar results were observed in the U2OS ALT cell line (Extended Data Fig. 7d). The TERRA signal was sensitive to treatment with RNaseA, showing a specific recovery of RNA, but the RNA signal was insensitive to treatment with DNaseI. Together these results indicate that TERRA associates with RAD51 in vivo.

We wished to determine whether purified RAD51 can bind TERRA directly by carrying out electrophoretic mobility shift assays (EMSAs). We incubated recombinant RAD51 with fluorescently labelled TERRA

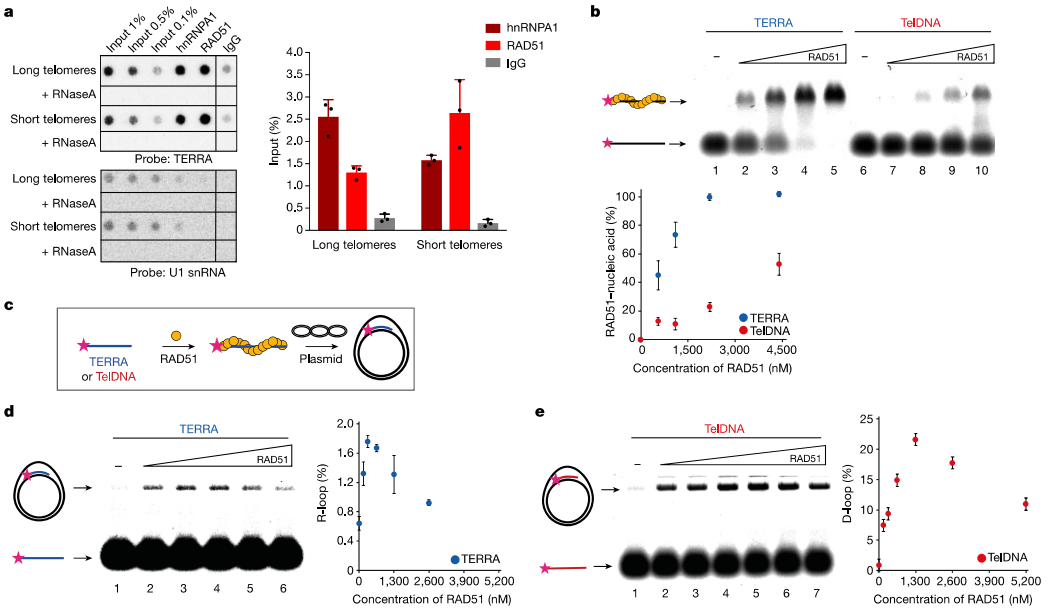


Fig. 3 | RAD51 associates with TERRA and catalyses R-loop formation. a, Native RNA-immunoprecipitation assays using anti-RAD51 and anti-hnRNPA1 antibodies were performed in extracts from HeLa cell clones with long or short telomeres. Western blotting was used to evaluate the efficiency of immunoprecipitation of RAD51 and hnRNPA1 (and the co-immunoprecipitation of BRCA2; Extended Data Fig. 7c, d). Immunoprecipitation-recovered RNA was analysed for TERRA and U1 snRNA. $n = 3$ biologically independent experiments; data are means ± s.d. **b**, The affinity of RAD51 for TERRA and telomeric DNA (TelDNA) oligonucleotides was analysed by electrophoretic mobility shift assay (EMSA). Fluorescently labelled (pink star) RNA or DNA substrates (20 nM) were incubated with increasing concentrations of RAD51 protein (0 nM, 550 nM, 1,100 nM, 2,200 nM and 4,400 nM). Quantification is shown at the bottom. $n = 3$ independent experiments; data are means ± s.d. **c**, Assay for the formation of R-loops and D-loops. **d**, The formation of R-loops depends on the concentration of RAD51: detection on native gel (left) and quantification (right). $n = 3$ independent experiments; data are means ± s.d. **e**, Detection of D-loops on native gel (left) and quantification (right). $n = 3$ independent experiments; data are means ± s.d.

or telomeric DNA oligonucleotides, and resolved the reactions by agarose-gel electrophoresis (Fig. 3b). RAD51 bound the TERRA oligonucleotide with threefold higher affinity than the corresponding telomeric DNA sequence (TelDNA), as with TERRA oligonucleotide a lower protein concentration was required to obtain shifted oligonucleotide–RAD51 complexes (Fig. 3b). The binding of RAD51 to the TERRA oligonucleotide was tighter than to an unrelated RNA (Extended Data Fig. 7e) and also more stable when compared with telomeric DNA, as shown by a stability assay in which pre-formed TERRA–RAD51 or TelDNA–RAD51 complexes were challenged with an excess of unlabelled 49-mer competitor ssDNA (cDNA) (Extended Data Fig. 7f).

To test whether RAD51 promotes R-loop formation in vitro, we carried out a strand-invasion assay using recombinant RAD51 and telomeric RNA and DNA (Fig. 3c–e). Wild-type RAD51, unlike the catalytically dead RAD51 II3A, catalysed both R-loop and D-loop formation in a concentration-dependent manner (Extended Data Fig. 8a). As expected, canonical R-loops were sensitive to treatment with RNaseH1 and were supershifted with the R-loop recognizing S9.6 antibody when combined with anti-mouse IgG (Extended Data Fig. 8b, c). Together, these data support a direct role of the RAD51 protein in localizing TERRA to initiate strand invasion and promote the formation of DNA–RNA hybrids at telomeres.

Discussion

Our data reveal the mechanism by which TERRA is recruited to chromosome ends through RNA strand invasion. The recruitment of TERRA and formation of R-loops depend on the recombinase RAD51 and the 5′-UUAGGG-3′ repeats of TERRA. Although the subtelomeric TERRA sequences were not required for telomere recruitment in our assay, we do not exclude the possibility that they may contribute by targeting endogenous TERRA molecules to their native chromosome ends. The observed base pairing between TERRA and telomeric DNA provides a mechanism for how this long noncoding RNA can encounter its major site of action at chromosome ends. Therefore, the recruitment of TERRA seems to occur through a process that strongly resembles the strand-invasion and homology-search mechanism exploited in all living organisms during DNA repair by HDR, and which is also characteristic of telomere stabilization by the ‘alternative lengthening of telomeres’ (ALT) mechanism in cancer cells. A direct involvement of RAD51 in the strand-invasion reaction is supported by our finding that RAD51 strongly binds TERRA and catalyses strand invasion into telomeric sequences in vitro. However, it is likely that there are important differences between DNA- and RNA-mediated strand invasion with the regard to the mechanism and the requirement for accessory factors, which must be explored in the future.

The increased formation of R-loops and association of RAD51 with short telomeres has previously been shown in yeast[8]. Our findings in human cells are consistent with these observations. Furthermore, we reveal that TERRA can form R-loops post-transcriptionally at telomeres in *trans*. R-loops have generally been assumed to form only during transcription, through the unsuccessful removal of native RNA from its DNA template[20]. However, previous studies in yeast suggested that

R-loops may also form at loci that are distinct from the site of synthesis of the RNA[21]. Our work here suggests that R-loop formation in *trans* plays a major part in telomere homeostasis. Through its mechanism of recruitment, TERRA may represent a scaffold that guides and regulates telomerase, HDR proteins and chromatin modifiers specifically at the chromosome ends that may need their attention. It will be interesting to see which other lncRNAs are recruited to their sites of action through similar mechanisms.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2815-6.

1. Maciejowski, J. & de Lange, T. Telomeres in cancer: tumour suppression and genome instability. *Nat. Rev. Mol. Cell Biol.* **18**, 175–186 (2017).
2. Azzalin, C. M., Reichenbach, P., Khoriauli, L., Giulotto, E. & Lingner, J. Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science* **318**, 798–801 (2007).
3. Schoeftner, S. & Blasco, M. A. Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase **II**. *Nat. Cell Biol.* **10**, 228–236 (2008).
4. Redon, S., Zemp, I. & Lingner, J. A three-state model for the regulation of telomerase by TERRA and hnRNPA1. *Nucleic Acids Res.* **41**, 9117–9128 (2013).
5. Cusanelli, E., Romero, C. A. P. & Chartrand, P. Telomeric noncoding RNA TERRA is induced by telomere shortening to nucleate telomerase molecules at short telomeres. *Mol. Cell* **51**, 780–791 (2013).
6. Moravec, M. et al. TERRA promotes telomerase-mediated telomere elongation in *Schizosaccharomyces pombe*. *EMBO Rep.* **17**, 999–1012 (2016).
7. Arora, R. et al. RNaseH1 regulates TERRA-telomeric DNA hybrids and telomere maintenance in ALT tumour cells. *Nat. Commun.* **5**, 5220 (2014).
8. Graf, M. et al. Telomere length determines TERRA and R-loop regulation through the cell cycle. *Cell* **170**, 72–85 (2017).
9. Porro, A., Feuerhahn, S., Reichenbach, P. & Lingner, J. Molecular dissection of telomeric repeat-containing RNA biogenesis unveils the presence of distinct and multiple regulatory pathways. *Mol. Cell. Biol.* **30**, 4808–4817 (2010).
10. Larson, D. R., Zenklusen, D., Wu, B., Chao, J. A. & Singer, R. H. Real-time observation of transcription initiation and elongation on an endogenous yeast gene. *Science* **332**, 475–478 (2011).
11. Hockemeyer, D. et al. Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat. Biotechnol.* **27**, 851–857 (2009).
12. Jensen, R. B., Carreira, A. & Kowalczykowski, S. C. Purified human BRCA2 stimulates RAD51-mediated recombination. *Nature* **467**, 678–683 (2010).
13. Thorslund, T. et al. The breast cancer tumor suppressor BRCA2 promotes the specific targeting of RAD51 to single-stranded DNA. *Nat. Struct. Mol. Biol.* **17**, 1263–1265 (2010).
14. Mason, J. M., Chan, Y.-L., Weichselbaum, R. W. & Bishop, D. K. Non-enzymatic roles of human RAD51 at stalled replication forks. *Nat. Commun.* **10**, 4410 (2019).
15. Boguslawski, S. J. et al. Characterization of monoclonal antibody to DNA.RNA and its application to immunodetection of hybrids. *J. Immunol. Methods* **89**, 123–130 (1986).
16. Sagie, S. et al. Telomeres in **I**CF syndrome cells are vulnerable to DNA damage due to elevated DNA:RNA hybrids. *Nat. Commun.* **8**, 14015 (2017).
17. Pfeiffer, V., Crittin, J., Grolimund, L. & Lingner, J. The THO complex component Thp2 counteracts telomeric R-loops and telomere shortening. *EMBO J.* **32**, 2861–2871 (2013).
18. Balk, B. et al. Telomeric RNA-DNA hybrids affect telomere-length dynamics and senescence. *Nat. Struct. Mol. Biol.* **20**, 1199–1205 (2013).
19. Sfeir, A. et al. Mammalian telomeres resemble fragile sites and require TRF1 for efficient replication. *Cell* **138**, 90–103 (2009).
20. Crossley, M. P., Bocek, M. & Cimprich, K. A. R-loops as cellular regulators and genomic threats. *Mol. Cell* **73**, 398–411 (2019).
21. Wahba, L., Gore, S. K. & Koshland, D. The homologous recombination machinery modulates the formation of RNA–DNA hybrids and associated chromosome instability. *eLife* **2**, e00505 (2013).

Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized and investigators were not blinded to allocation during experiments and outcome assessment.

Cell culture and transfections

The telomerase-positive cell line HeLa (cervical cancer) was from the American Type Culture Collection (ATCC), and the ALT cell line U2OS (osteosarcoma) from the European Collection of Authenticated Cell Cultures (ECACC). The cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 100 U ml^{-1} of penicillin/streptomycin (Thermo Fisher Scientific). The cell lines were maintained in a 5% CO_2 incubator at 37 °C and were routinely checked to ensure that they remained mycoplasma-free.

HeLa cells were transfected with lipofectamine 2000 according to the manufacturer's instructions (Invitrogen). To induce TERRA transcription, doxycycline (1 μg ml^{-1}) was added to the medium 24 h after transfection, and the cells were harvested 48 h after transfection. To isolate clones containing integrated TERRA transgenes, puromycin (1 μg ml^{-1}) was added to the medium 24 h following transfection using the pSpCas9(BB)-2A-Puro plasmid. Selection was maintained for 5 days. SiRNAs were purchased as pools from Dharmacon (siGENOME SMARTpool; Extended Data Table 1). Cells were transfected with 20 pmol siRNAs using calcium phosphate transfection in 6-well plates in DMEM supplemented with 10% FBS. Cells were harvested 72 h after transfection. Wild-type RAD51 or the RAD51 II3A mutant was expressed from plasmid containing complementary DNA[14].

Lentivirus production and cell transduction

Plasmids pMD2.G (1 μg) and pCMVR8.74 (3 μg) (gifts from D. Trono, EPFL) were mixed with PP7_eGFP (3 μg) (a gift from D. Larson, National Institutes of Health (NIH)) or pLenti CMV rtTA3 Hygro (Addgene catalogue number 26730) for transfection of HEK293T cells in Optimem medium (Thermo Fisher Scientific) using lipofectamine 2000. The transfection mix was incubated overnight and the medium replaced the next day. Supernatants were collected on the next two consecutive days upon centrifugation, and cleared through a 0.22-μm filter unit (Stericup, Millipore), before the viruses were aliquoted and frozen at −80 °C. HeLa cells were transduced with 1 ml of viral particles of pLenti CMV rtTA3 Hygro. Following infection, the cells were selected with hygromycin (200 μg ml^{-1}) for 5 days, before they were infected again with the PP7–GFP viruses. GFP-positive clones with similar GFP intensity were isolated with a FACSAria Fusion sorter by EPFL's flow cytometry core facility.

Cloning of TERRA constructs

TERRA constructs were cloned in the pTRE2puro vector of Clontech's TET-ON system without the corresponding polyA region. The PP7 stem-loops were amplified from Addgene plasmid #61762 (a gift from D. Larson) and introduced into pTRE2puro through in-fusion cloning. To amplify the TTAGGG sequence, we performed PCR using the Phusion Green Hot Start II high-fidelity DNA polymerase (F537S) with no template in a reaction containing: 1× GC Phusion buffer, 0.2 mM dNTPs, 0.4 μM primer 5′-TTAGGGTTAGGGTTAGGGTTAGGGTTAGGG-3′, 0.4 μM primer 5′- CCCTAACCCTAACCCTAACCCTAACCCTAA-3′, and 2 U of polymerase. The PCR consisted of 30 s at 98 °C followed by 10 cycles at 98 °C for 5 s, 60 °C for 10 s, and 72 °C for 15 s. DNA products of variable size were fractionated on a 1.2% agarose gel, and the desired size was excised, extracted with a QIAquick gel extraction kit, and cloned into the pTRE2puro vector containing the 24 copies of PP7 stem-loops. Subtelomeric sequences were amplified using a high-fidelity polymerase from phenol–chloroform-extracted genomic DNA from HeLa cells, and introduced into the corresponding vectors through in-fusion cloning. All constructs were verified by restriction digestion and sequencing. The plasmids were amplified in a homologous-recombination-deficient *E. coli* strain (Stbl3) at 30 °C. All of the primers are listed in Extended Data Table 1.

Generation of integrated TERRA cell lines

The integration of TERRA constructs into the AAVS1 locus was performed as described[22]. The guide RNA for AAVS1 (5′-ACCCCACAG TGGGGCCACTA-3′) and its complementary strand were annealed, and cloned into the pSpCas9(BB)-2A-Puro plasmid acquired from Addgene (catalogue number 48139). The donor template encompassed roughly 800 base pairs of homology arms for AAVS1, which were amplified from HeLa genomic DNA, and cloned into the plasmids containing the control PP7 stem-loop construct and the different TERRA transgenes. HeLa cells were transfected with both the gRNA/Cas9 and donor template vectors. Individual clones were screened by PCR for the presence of the transgene in the AAVS1 locus.

Immunofluorescence and FISH

Cells were grown on glass coverslips by following the culture conditions above. The coverslips were washed twice with 1× phosphate-buffered saline (PBS), fixed with 4% paraformaldehyde for 10 min at room temperature, and washed twice with 1× PBS. The cells were then permeabilized in 1× detergent solution for 5 min (0.1% Triton X-100, 0.02% SDS in 1× PBS), followed by pre-blocking with 2% bovine serum albumin (BSA) in 1× PBS for 10 min. Next the cells were blocked with 10% normal goat serum in 2%BSA/1× PBS for 30 min at room temperature. Coverslips were incubated with primary and then secondary antibodies in blocking solution for 90 min each time at room temperature, and fixed with 4% paraformaldehyde for 5 min at room temperature; the samples were then dehydrated with increasing concentrations of ethanol. For FISH staining, the cells were incubated with hybridization solution containing 10 mM Tris–HCl pH 7.4, 70% formamide, 0.5% blocking reagent and 1/1,000 Cy3 probe, denatured at 80 °C for 3 min, and hybridized for 3 h at room temperature. The cells were washed with wash 1 (10 mM Tris–HCl pH 7.4, 70% formamide) and wash 2 (0.1 M Tris–HCl pH 7.4, 0.15 M NaCl, 0.08% Tween-20) twice, stained with 4′,6-diamidino-2-phenylindole (DAPI), and dehydrated with ethanol. Images were acquired on an Upright Zeiss Axioplan or on a Zeiss LSM700 equipped with an AxioCam MRm B/W. Images were processed and analysed with ImageJ and Adobe Photoshop. All statistical analysis was performed using GraphPad Prism. All figures were created in Adobe Illustrator.

Telomeric FISH on metaphase spreads

Cells were treated with 0.05 μg ml^{-1} demecolcine for 2 h, collected and incubated in hypotonic solution (0.075 M KCl) at 37 °C for 8 min. Swollen cells were collected and fixed in ice-cold methanol:acetic acid (3:1) overnight at 4 °C. The next day, 100 μl of cell suspension was dropped onto slides, incubated at 70 °C for 1 min and air-dried overnight at room temperature. FISH staining was performed as above.

Telomere length analysis

HeLa genomic DNA was isolated with the Wizard genomic DNA purification kit according to the manufacturer's instructions (Promega). Then, 6 μg of genomic DNA was digested with 30 U of *Hinf*I and *Rsa*I overnight at 37 °C. The digested DNA was mixed with 6× MassRuler DNA-loading dye (Thermo Fisher Scientific), loaded on a 0.8% agarose gel in 1× Tris–borate–EDTA (TBE) buffer, and fractionated by gel electrophoresis at 2 V cm^{-1} for 20 h. The gels were dried for two hours at 50 °C in vacuum, treated with denaturation buffer (0.5 M NaOH, 1.5 M NaCl) and neutralization buffer (0.5 M Tris-HCl pH 7.5, 1.5 M NaCl), and then pre-hybridized with Church buffer for 1 h at 50 °C. The gels were hybridized overnight at 50 °C with a randomly labelled TeloC probe as described[23]. The gels were washed for 1 h at 50 °C with 4× saline sodium

citrate (SSC), 4× SSC 0.1% SDS, and 2× SSC 0.1%SDS, exposed to a phosphorimager screen and analysed on a Typhoon phosphorimager (GE).

RT–qPCR

For RT–qPCR of TERRA, 3×10^6 cells were harvested following transfection of the chimaeric TERRA constructs. RNA was isolated with a NucleoSpin RNA kit (Macherey-Nagel). RT–qPCR for 15qTERRA was carried out using our previously described protocol[24]. To assess messenger RNA levels following siRNA transfections, we used ThermoFisher Scientific's SuperScript III reverse transcriptase and Power SYBR Green PCR master mix on an Applied Biosystems 7900HT fast real-time system according to the manufacturer's instructions.

Western blotting

Antibodies are listed in Extended Data Table 2. Protein samples were mixed with 2× Laemmli buffer, boiled for 5 min at 95 °C, run on 4–15% SDS–PAGE precast gels (Mini-PROTEAN TGX Gels, BioRad), transferred to nitrocellulose blotting membranes (Amersham), blocked with blocking solution (3% BSA in 1× PBS, 0.1% Tween-20), and incubated overnight at 4 °C with the corresponding primary antibody. Membranes were washed three times for 5 min each with 1× PBS plus Tween (PBST) buffer, and then incubated for 1 h at room temperature with horseradish-peroxidase-conjugated secondary antibodies (Promega) in blocking solution. Membranes were washed again three times for 5 min with 1× PBST before revealing them with a chemiluminescence detection kit (Western bright electrochemiluminescence, Advansta) and analysing them on a Vilber Fusion FX imaging system.

DNA–RNA immunoprecipitation

Cells at roughly 40% confluence in 6-well plates were transfected first with siRNAs and then, the next day, with plasmids. The cells were harvested on ice 48 h later, counted on a CASY cell counter and washed with 1× PBS; samples were taken for DRIP and western blot analysis. For DRIP, 10^7 cells were dissolved in 175 µl of ice-cold RLN buffer (50 mM Tris-HCl pH 8.0, 140 mM NaCl, 1.5 mM $MgCl_2$, 0.5% NP-40, 1 mM dithiothreitol (DTT), and 100 U ml^{-1} RNasIN PLUS), incubated on ice for 5 min, and centrifuged (300*g*, 2 min, 4 °C). The nuclei were lysed in 500 µl RA1 buffer (NucleoSpin RNA, Macherey-Nagel) containing 5 µl of β-mercaptoethanol, and homogenized by passing them through a 20G $\times 1^{1/2}$ syringe (0.9 mm × 40 mm). The nucleic-acid-containing extracts were mixed with 250 µl H_2O and 750 µl phenol:chloroform:isoamylalcohol (25:24:1) in a Phase Lock Gel heavy (5PRIME) and centrifuged (13,000*g*, 5 min, room temperature). The upper aqueous phase was mixed with 750 µl of ice-cold isopropanol, with addition of NaCl to 50 mM, then incubated on ice for 30 min; precipitated nucleic acids were collected by centrifugation at 10,000*g* for 30 min at 4 °C. The pellets were washed twice with 70% ice-cold ethanol, air-dried, dissolved in 130 µl of H_2O, and sonicated on a Covaris system (10% duty factor, 200 cycles per burst, for 180 s, with an AFA intensifier) to obtain fragments of 100–500 bp. Next, 90 µg of sonicated nucleic acids were mixed with RNaseH1 or H_2O in RNaseH1 buffer (20 mM HEPES-KOH pH 7.5, 50 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT) and incubated at 37 °C for 90 min. The samples were diluted ten times in DIP-1 buffer (10 mM HEPES-KOH pH 7.5, 275 mM NaCl, 0.1% Na-deoxycholate, 0.1% SDS, 1% Triton X-100) and pre-cleared with 80 µl of sepharose protein G beads for 1 h, on a rotating wheel, at 4 °C. One per cent of the nucleic acids were kept as input. Half of the samples (roughly 45 µg of nucleic acids) were incubated with 3 µg of S9.6 antibody or mouse IgG and 40 µl of sepharose protein G beads on a rotating wheel at 4 °C overnight. The next day the samples were washed for 5 min on a rotating wheel at 4 °C with DIP-2 (50 mM HEPES-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA pH 8.0, 1% Triton X-100, 0.1% Na-deoxycholate), DIP-3 (50 mM HEPES-KOH pH 7.5, 500 mM NaCl, 1 mM EDTA pH 8.0, 1% Triton-X100, 0.1% Na-deoxycholate), DIP-4 (10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0, 250 mM LiCl, 1% NP-40, 1% Na-deoxycholate), and TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0). The samples were digested at 65 °C overnight with 10 µg ml^{-1} RNase (DNase-free (Roche)) in a buffer containing 20 mM Tris-HCl pH 8.0, 1% SDS, 0.1 M $NaHCO_3$, 0.5 mM EDTA pH 8.0. The DNA was isolated using the Qiagen PCR clean-up kit. The DNA was pipetted onto a positively charged nylon membrane (Amersham Hybond N+) and telomeric DNA was detected with a telomeric probe as described[25].

RNA immunoprecipitation

Cells were grown to roughly 70% confluence in 15-cm dishes, harvested on ice, counted on a CASY cell counter, washed with 1× PBS, and lysed in RNA-immunoprecipitation RLN buffer (50 mM Tris-HCl pH 8.0, 140 mM NaCl, 1.5 mM $MgCl_2$, 0.5% NP-40, 1 mM DTT, 400 U ml^{-1} RNasIN PLUS, and protease inhibitors (Complete, Roche)). The lysates were pre-cleared with Dynabeads plus protein G for 1 h on a rotating wheel at 4 °C, and incubated overnight with 6 µg of the corresponding antibody on a rotating wheel at 4 °C. Next 35 µl of Dynabeads protein G (ThermoFisher Scientific) that had been preblocked with transfer RNA were added to each mixture and incubated for 2 h on a rotating wheel at 4 °C; this was followed by five washes with RLN buffer supplemented with 6 mM EDTA pH 8.0. The RNA was eluted in 1% SDS, 5 mM EDTA pH 8.0, and 5 mM β-mercaptoethanol at 42 °C for 30 min, followed by 30 min at 65 °C. The RNA was purified using the RNA clean-up protocol of the NucleoSpin RNA kit (Macherey-Nagel). The RNA was denatured at 65 °C for 3 min and pipetted on a positively charged nylon membrane (Amersham Hybond N+). TERRA and U1 snRNA were detected with corresponding probes as described[26].

EMSA

RAD51 protein was purified[27] and diluted in dilution buffer (25 mM Tris-HCl (pH 7.5), 10% (*v*/*v*) glycerol, 0.5 mM EDTA, 50 mM KCl, 1 mM DTT and 0.01% NP40). Increasing concentrations of RAD51 were incubated with 20 nM Cy3-labelled 41mer TERRA (5′-UUAGGGUUAGGGUUAGGGUUAGGGUUAGGGUUAGGGUUAGG-3′), 41mer TelDNA (5′-TTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGG-3′), 41mer non-TelRNA (5′-AGUAUAUAUGAGUAAACUUGGUCUGACAGUUACCAAUGCUU-3′) or 40mer non-TelDNA (5′-AAATTAACAAGTATAATAAGAA ATAGAAACAAGAAATAGA-3′) substrate at 37 °C for 10 min in 10 µl of buffer D (50 mM Tris-HCl (pH 7.5), 50 mM KCl, 1 mM $MgCl_2$, 1 mM ATP). Reaction mixtures were then crosslinked with 0.01% glutaraldehyde for 10 min. Products were resolved using 0.8% TBE agarose gels supplemented with 10 mM KCl at 4 °C for 50 min (6.5 V cm^{-1}). Gels were imaged on a FLA-9000 scanner (Fujifilm) and quantified with Multi Gauge version 3.2 (Fujifilm).

Stability EMSA

Fluorescently labelled 41mer TERRA or 41mer TelDNA substrates (20 nM) were incubated with indicated concentrations of RAD51 at 37 °C for 10 min in 50 mM Tris-HCl (pH 7.5), 50 mM KCl, 1 mM $MgCl_2$ and 1 mM ATP. To challenge assembled RAD51–ssDNA complexes, increasing concentrations of unlabelled 49mer competitor ssDNA (5′-CCTGTTCAAACGCACATATTAAGCATTTC CTGTCATTG GCGGCTAATTC-3′) were added and incubated for another 10 min at 37 °C. Products were crosslinked with 0.005% glutaraldehyde for a further 10 min and resolved using a 0.8% TBE agarose gel supplemented with 10 mM KCl at 4 °C for 50 min (6.5 V cm^{-1}). Gels were imaged on a FLA-9000 scanner (Fujifilm) and quantified with Multi Gauge version 3.2 (Fujifilm).

Assays for R-loop/D-loop formation

Fluorescently labelled 41mer TERRA or 41mer TelDNA (50 nM) were pre-incubated for 10 min at 37 °C with increasing concentrations of RAD51 in 50 mM Tris-HCl (pH 7.5), 1 mM $MgCl_2$ supplemented with 1 mM $CaCl_2$ and 1 mM adenylyl-imidodiphosphate (AMP–PNP). The reaction was started by adding 600 ng pCR4-TOPO vector containing the 15q

subtelomeric sequence followed by 15 copies of TTAGGG repeats. The mixture was incubated for another 10 min at 37 °C and then stopped by adding SDS (0.1%) and proteinase K (0.1 mg ml^{-1}); this was followed by 3-min incubations at 37 °C. Reaction mixtures were separated on 0.8% agarose gel and analysed as described above for EMSA. For digestion of R-loops by RNaseH, the mixtures were incubated for 1 min with phenylmethylsulfonyl fluoride (PMSF; 2 mM) and with EGTA (1.6 mM) at 37 °C to inhibit proteinase K and chelate calcium ions, respectively. The products were digested with RNaseH1 (6.8 mU μl^{-1}, Thermo Scientific) or RNaseH2 (6.8 mU μl^{-1}, New England Biolabs) for 1 min at 37 °C and resolved as above. For antibody-specific supershift of R-loops, the products were formed as described above for RNaseH digestion (but without EGTA) and then incubated with S9.6 antibodies (0.02 $\mu g\ \mu l^{-1}$) for 2 min at 37 °C and/or anti-mouse horseradish-peroxidase-conjugated antibody.

RNaseH1 digestion of telomeric DNA–RNA hybrids

Fluorescently labelled TERRA or non-telomeric RNAs were mixed with their corresponding complementary ssDNA in annealing buffer (10 mM Tris pH 8.5, 50 mM NaCl, 1 mM EDTA). To form DNA–RNA hybrids, the mixture was heated for 5 min at 95 °C, and then gradually cooled to room temperature. The DNA–RNA hybrid (40 nM) was cleaved by RNaseH1 (6.8 mU μl^{-1}) in the presence or absence of 1 mM $CaCl_2$ for 10 min at 37 °C in 50 mM Tris (pH 7.5), 1 mM $MgCl_2$ and 1 mM AMP–PNP. The reaction was stopped by adding SDS/proteinase K, and then incubated for 10 min at 37 °C. Reaction mixtures were loaded on 10% PAGE gel, separated by electrophoresis (90 V for 60 min at 4 °C), scanned using an Image Reader FLA-9000 scanner and quantified by MultiGauge version 3.2 software.

Supershift of DNA–RNA hybrid with S9.6 antibody

Fluorescently labelled TERRA and non-telomeric DNA–RNA hybrids (40 nM) were incubated with 0.015 $\mu g\ \mu l^{-1}$ of S9.6 antibody for 10 min at 37 °C, then crosslinked with 0.01% glutaraldehyde for 10 min at 37 °C and resolved using 0.8% TBE agarose gel at 4 °C for 50 min (6.5 V cm^{-1}). Gels were imaged on a FLA-9000 scanner (Fujifilm) and quantified with Multi Gauge version 3.2 (Fujifilm).

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The data that support the findings of this study are available from the corresponding authors upon reasonable request. Source data are provided with this paper.

22. Ran, F. A. et al. Genome engineering using the CRISPR-Cas9 system. *Nat. Protocols* **8**, 2281–2308 (2013).
23. Grolimund, L. et al. A quantitative telomeric chromatin isolation protocol identifies different telomeric states. *Nat. Commun.* **4**, 2848 (2013).
24. Feretzaki, M. & Lingner, J. A practical qPCR approach to detect TERRA, the elusive telomeric repeat-containing RNA. *Methods* **114**, 39–45 (2017).
25. Porro, A. et al. Functional characterization of the TERRA transcriptome at damaged telomeres. *Nat. Commun.* **5**, 5379 (2014).
26. Porro, A., Feuerhahn, S. & Lingner, J. TERRA-reinforced association of LSD1 with MRE11 promotes processing of uncapped telomeres. *Cell Rep.* **6**, 765–776 (2014).
27. Špírek, M. et al. Human RAD51 rapidly forms intrinsically dynamic nucleoprotein filaments modulated by nucleotide binding state. *Nucleic Acids Res.* **46**, 3967–3980 (2018).

Acknowledgements We thank D. Larson (NIH), D. Bishop (Univ. Chicago), V. Simanis (EPFL), P. Gönczy (EPFL) and D. Trono (EPFL) for providing material. We also thank the members of the BIOP core facility at EPFL, members of the Gönczy laboratory and M. Spirek for technical support and advice and J. Cibulka for recombinant mutant RAD51 protein. M.F. was supported by the European Union's Horizon 2020 research and innovation programme under Marie Skłodowska-Curie grant agreement 702824. J.L.'s laboratory was supported by the Swiss National Science Foundation (SNFS grant 310030_184718), the SNFS-funded National Centres of Competence in Research (NCCR) RNA and disease network (grant 182880), and EPFL. L.K.'s laboratory was supported by the European Structural and Investment Funds, Operational Programme Research, Development and Education 'Preclinical Progression of New Organic Compounds with Targeted Biological Activity' (Preclinprogress) (CZ.02.1.01/0.0/0.0/16_025/0007381); a Wellcome Trust Collaborative Grant 206292/E/17/Z; the Czech Science Foundation (GACR 17-17720S); and the National Program of Sustainability II (MEYS CR, project number LQ1605). Both J.L. and L.K. were also supported by the European Union's Horizon 2020 research and innovation programme under grant agreement 812829.

Author contributions M.F. and J.L. conceived the study. M.F. and R.V.F. executed all cell and molecular biology experiments. M.P. performed all biochemistry experiments and T.L. some EMSA experiments. L.K. conceived the RAD51-based biochemistry experiments and advised on the text. J.L. and M.F. wrote the paper.

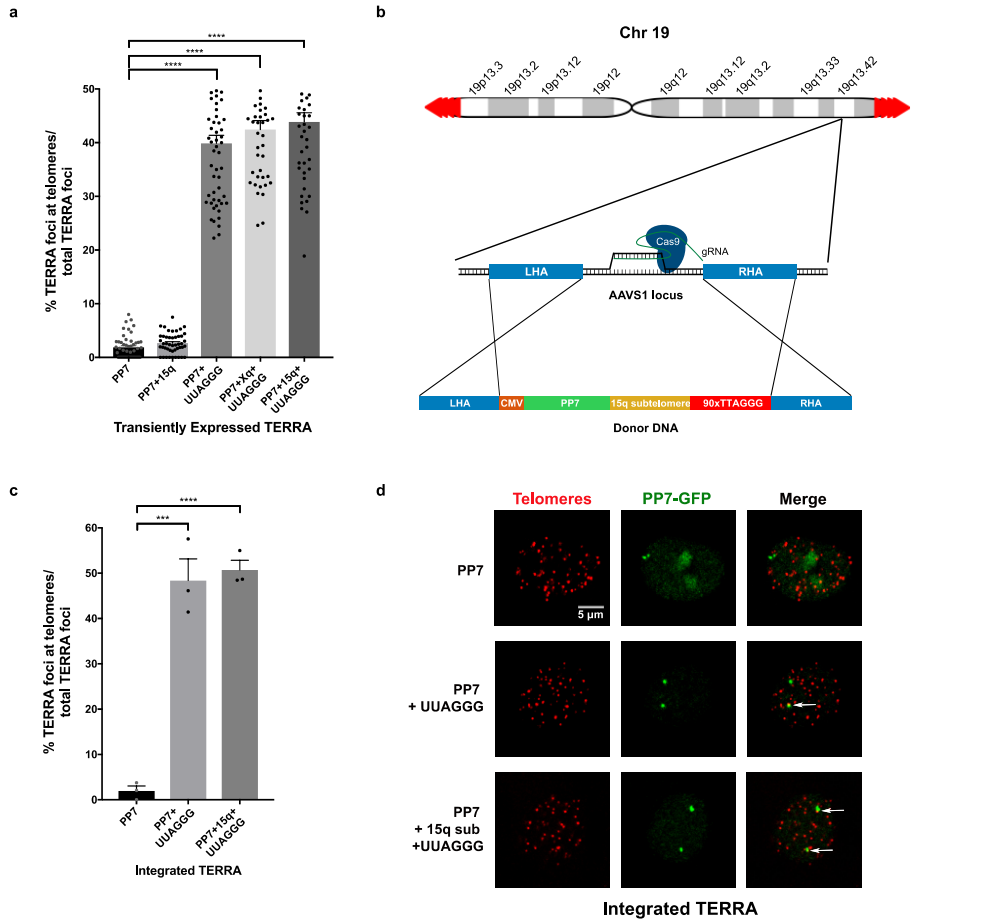
Competing interests The authors declare no competing interests.

Additional information

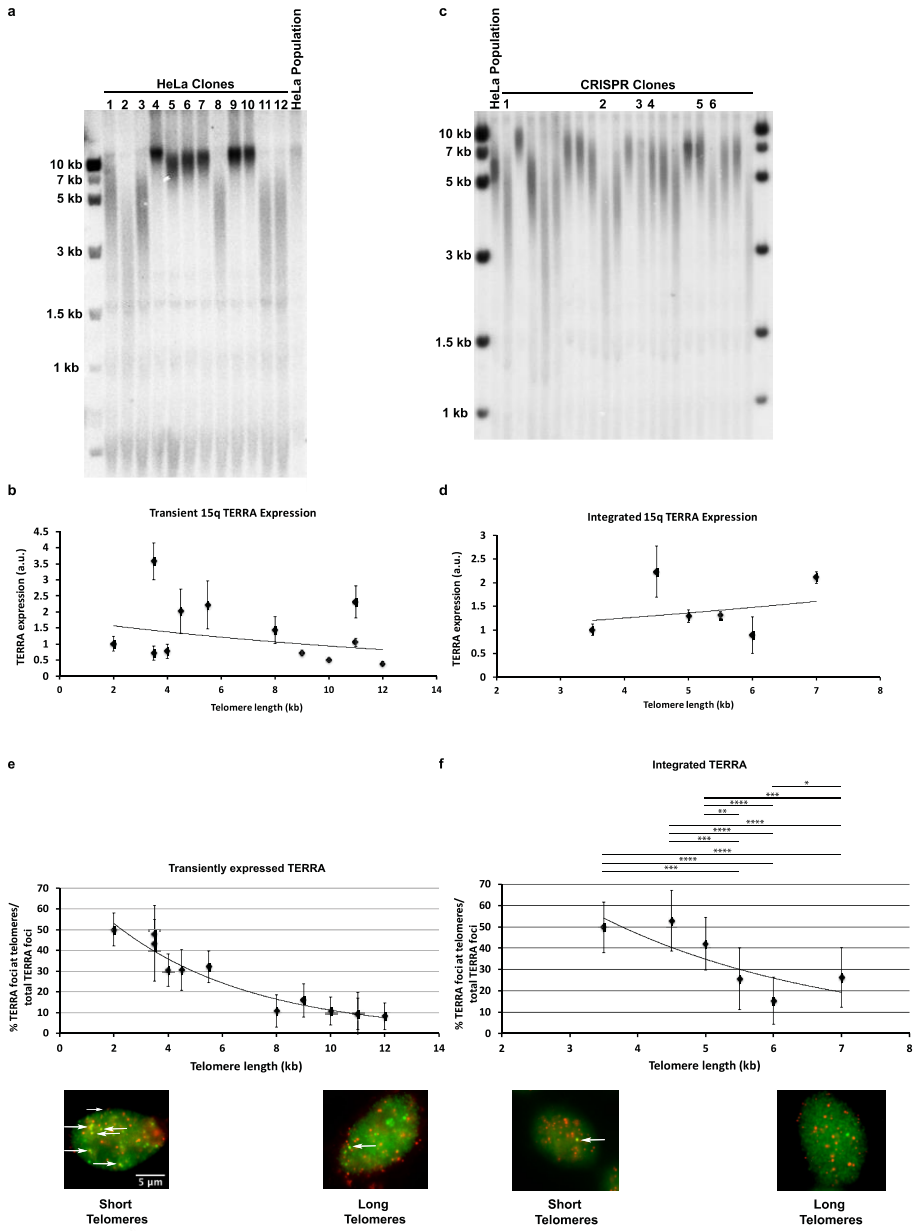
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2815-6.

Correspondence and requests for materials should be addressed to L.K. or J.L.

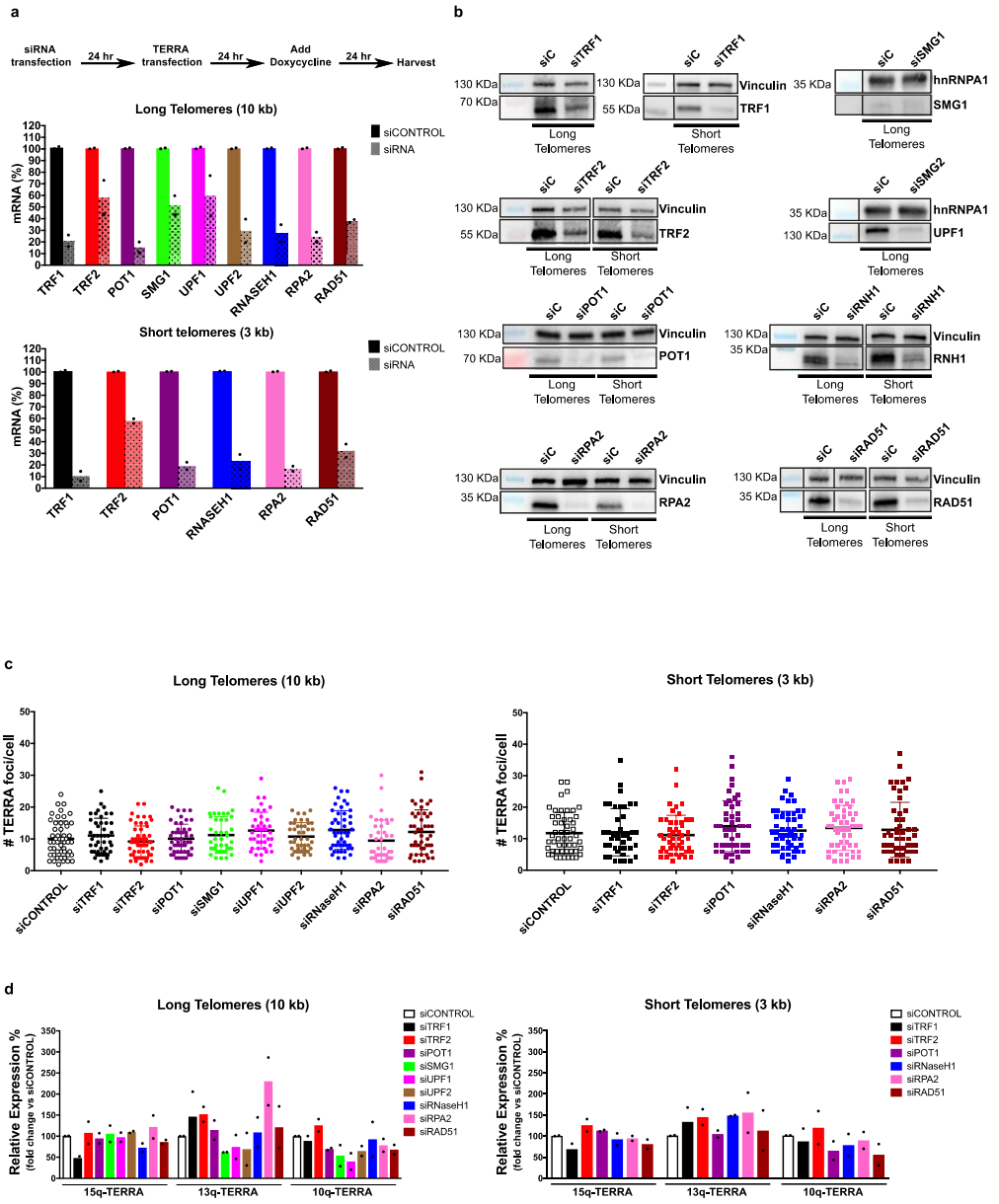
Peer review information *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work.



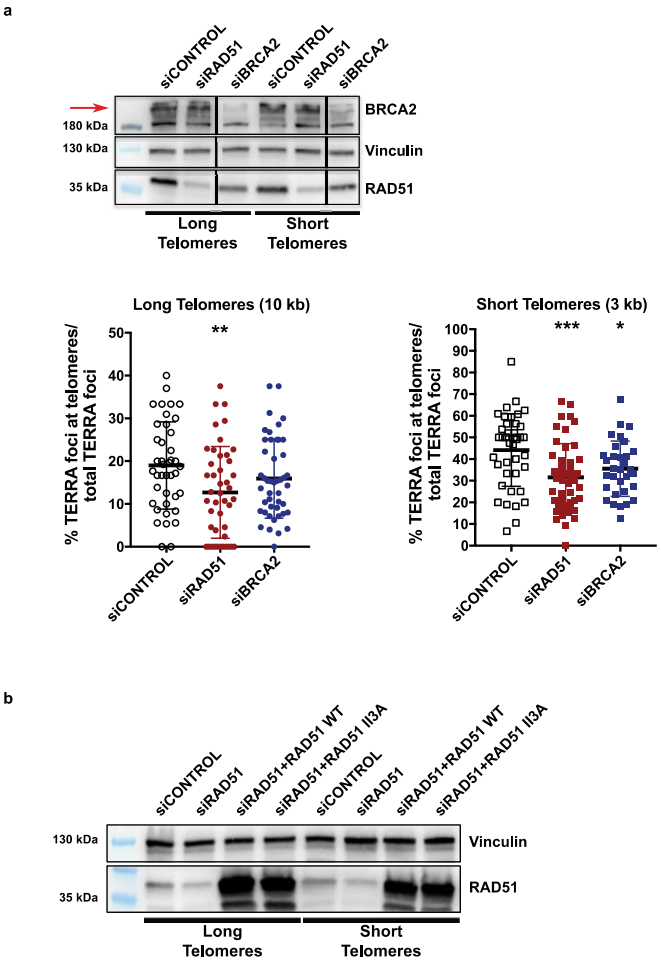
Extended Data Fig. 1 | Co-localization of transgenic TERRA with telomeres. **a**, Quantification of co-localization of transgenic RNA foci expressed from plasmids with telomeres. **b**, Strategy for targeting the AAVS1 locus through CRISPR–CAS9 to integrate chimaeric TERRA constructs into the genome. LHA and RHA, left and right homology arms. **c**, Quantification of co-localization of transgenic RNA foci with telomeres when TERRA was expressed from the AAVS1 locus. **d**, Confocal images obtained when TERRA was expressed from the AAVS1 locus. For **a**, **c**, $n = 3$ biologically independent experiments; data are mean ± s.d.; two-tailed unpaired *t*-test: ***$P < 0.001$; ****$P < 0.0001$.



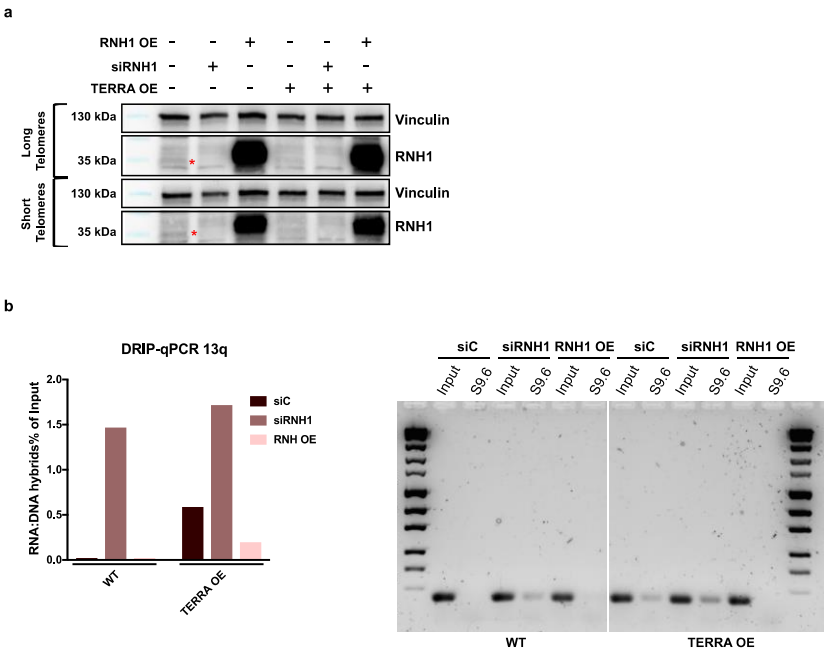
Extended Data Fig. 2 | TERRA associates preferentially with short telomeres. a, TRF analysis of HeLa clones used for transient expression of TERRA. **b**, **d**, TERRA expression levels measured by RT–qPCR relative to the levels of TERRA in the clone with the shortest telomeres. $n = 3$ biologically independent experiments; data are means ± s.d. **c**, TRF analysis of HeLa clones expressing transgenic TERRA from the AAVS1 locus. **e**, HeLa clones of different telomere lengths were transiently transfected with the PP7–15qTERRA construct and co-localization was assessed by immunofluorescence with FISH (bottom). White arrows indicate co-localization of PP7 foci with the telomeric signal. The percentage of TERRA foci that co-localized with telomeres is plotted as a function of telomere length (top). Random colocalization events (roughly 3%) were not subtracted. **f**, Co-localization events as in **e** but for PP7–15qTERRA expressed from the AAVS locus on chromosome 19. $n = 3$ biologically independent experiments; data are means ± s.d. Significant differences are indicated. Two-tailed unpaired *t*-tests were used to calculate *P*-values: $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; $^{****}P < 0.0001$.



Extended Data Fig. 3 | Depletion of factors regulating TERRA trafficking. **a**, Timeline of transfections and cell harvesting. mRNA levels were determined by RT–qPCR for long-telomere and short-telomere cell lines. **b**, For the second replicate of each siRNA screen, the level of depletion was also evaluated on western blots. Vinculin and hnRNPA1 were used as loading controls. **c**, Numbers of TERRA foci per cell are plotted for each depleted factor. $n = 2$ biologically independent experiments; at least 40 nuclei were analysed per condition. **d**, Quantification of endogenous TERRA stemming from the ends of chromosomes 15q, 13q and 10q upon depletion of the indicated factors relative to the negative control (siCONTROL). $n = 2$ biologically independent experiments.

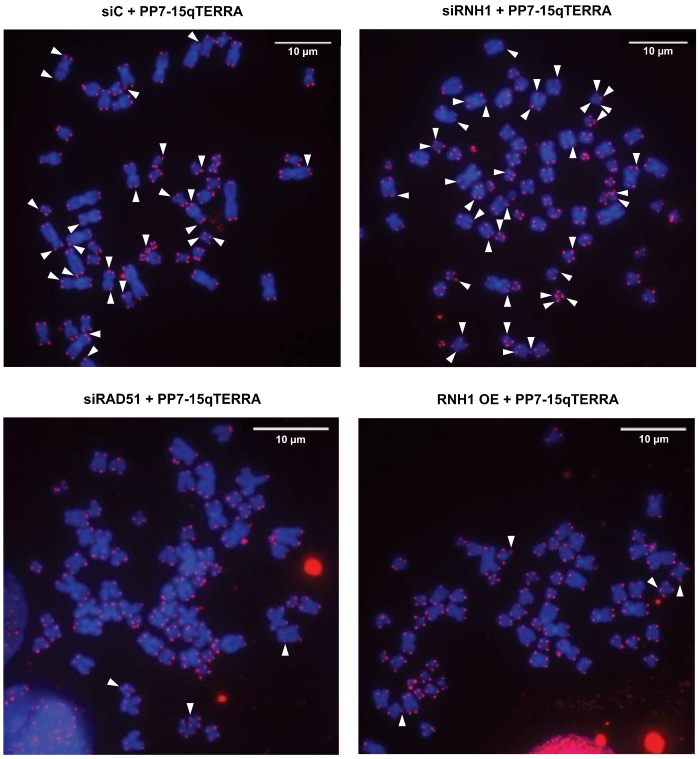


Extended Data Fig. 4 | Depletion of RAD51 and BRCA2, which regulate the association of TERRA with telomeres. a, HeLa clones with long and short telomeres were transfected with siRAD51 and siBRCA2, and then with chimaeric TERRA constructs. Representative immunoblots show RAD51 and BRCA2 depletion, with vinculin as a loading control. Upon depletion, co-localization of TERRA with telomeres was assessed with immunofluorescence–FISH. $n = 2$ biologically independent experiments; at least 54 nuclei were analysed per condition; data are means ± s.d. One-way ANOVA with Dunnett's multiple comparisons test was used, comparing all conditions to siCONTROL: $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$. **b**, Detection of endogenous and transgenic RAD51 on a western blot. Endogenous RAD51 was depleted with siRNA, and wild-type (WT) RAD51 or the RAD51 II3A mutant was expressed from plasmids containing complementary DNA.



Extended Data Fig. 5 | RNaseH1-regulated formation of telomeric R-loops in *trans*. **a**, RNH1 depletion and overexpression (OE) was assessed by western blotting in cell lines with long and short telomeres and upon TERRA overexpression. **b**, Representative example of one DRIP assay followed by qPCR for the end of chromosome 13q. Left, quantification of RNA–DNA hybrids is expressed as a fraction of input. Right, the amplified DNA was run on a gel, then isolated and sequenced.

a



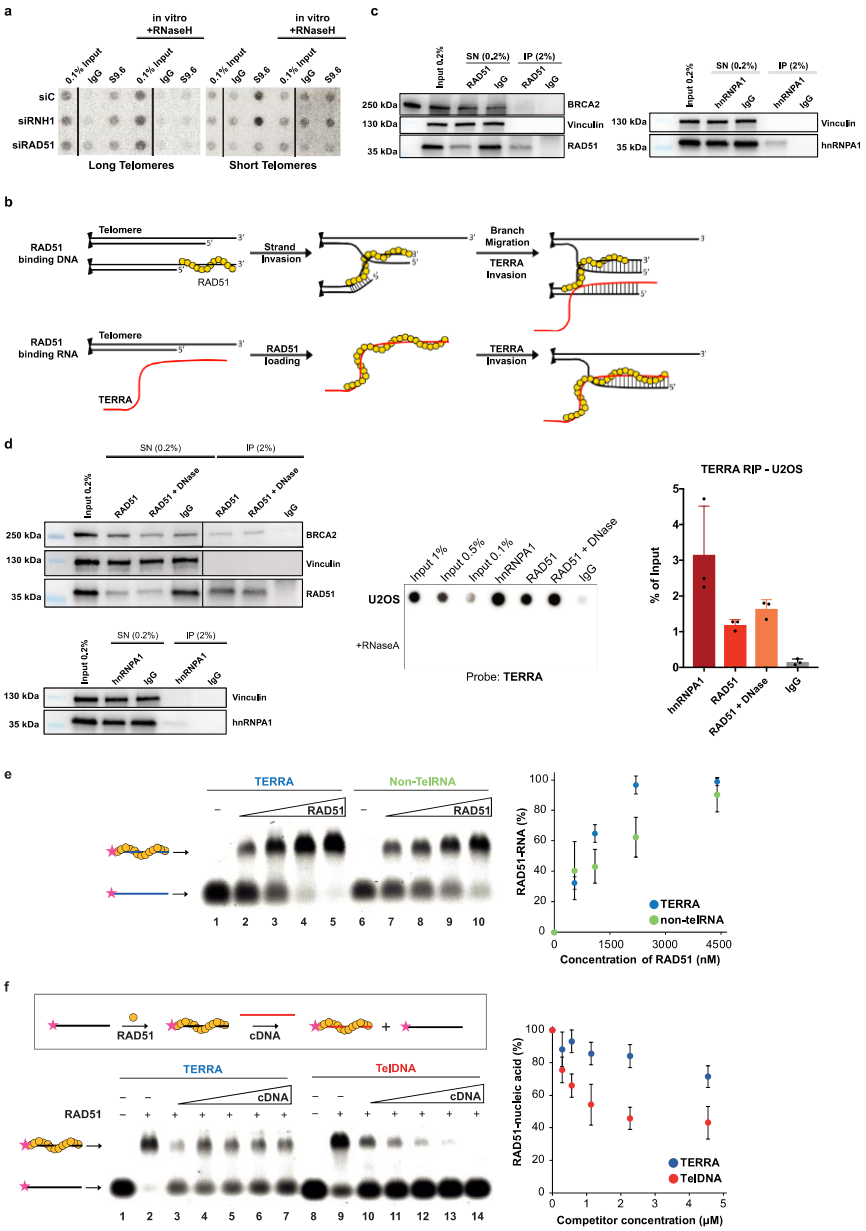
b

Plasmid	Metaphases scored	Telomere Fragility (mean % ± SEM)
PP7	n=59	2.7 ± 0.2
PP7+UUAGGG	n=58	5.4 ± 0.4
PP7+15q+UUAGGG	n=60	4.6 ± 0.2
PP7+Xq+UUAGGG	n=58	5.1 ± 0.3

c

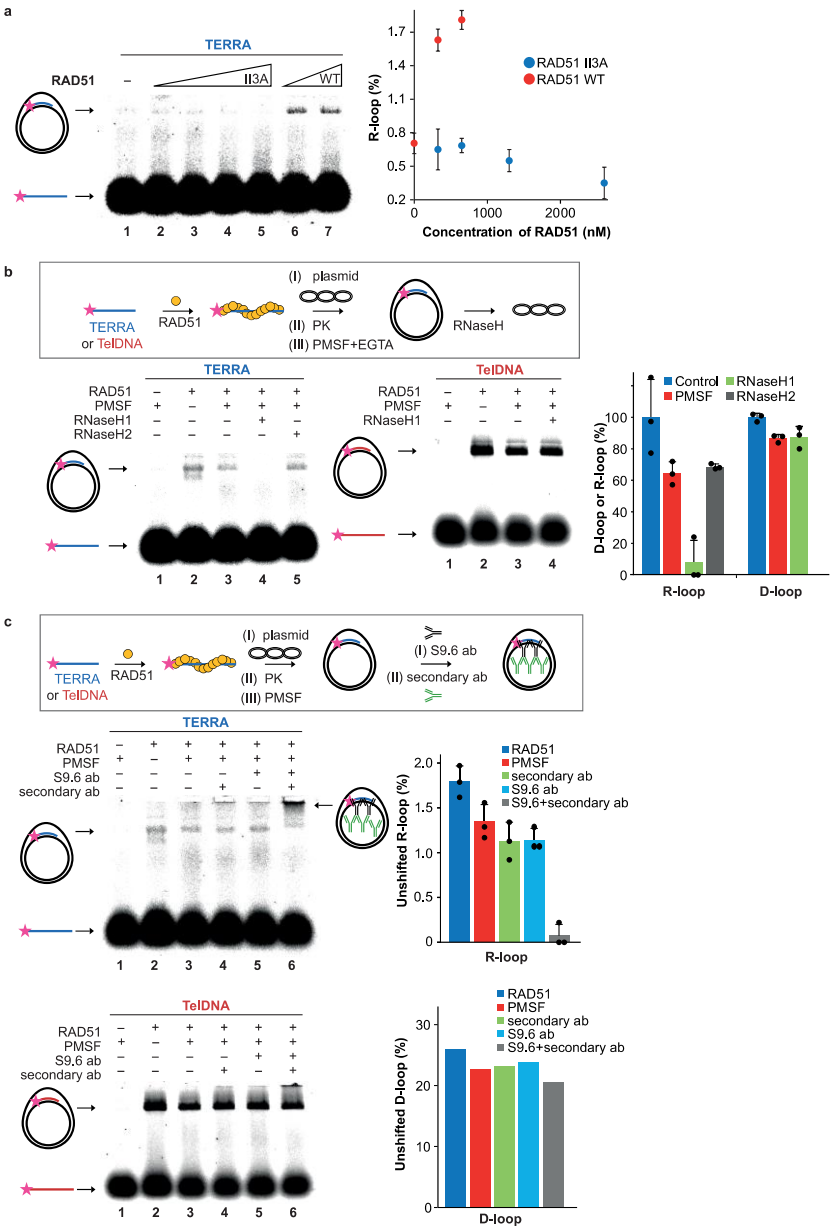
	Condition	Metaphases scored	Telomere Fragility (mean % ± SEM)
+ PP7-15qTERRA	siC	n=64	5.3 ± 0.4
	siRNH1	n=69	7.1 ± 0.3
	siRAD51	n=64	3.4 ± 0.3
	RNH1 OE	n=64	2.8 ± 0.2

Extended Data Fig. 6 | Transgenic TERRA expressed from plasmids induces telomere fragility. a, Representative examples of metaphase chromosomes stained with FISH to visualize telomeres. DNA is stained with DAPI. Fragile telomeres are indicated by white arrowheads. **b**, Quantification of telomere fragility. **c**, Quantification of telomere fragility in cells expressing PP7–15qTERRA in which expression of RNaseH1 (RNH1) and RAD51 was manipulated as indicated. For **b**, **c**, the numbers of metaphases scored over three biologically independent experiments are indicated for each condition as *n*.



Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | RAD51 binds TERRA and promotes R-loop formation. **a**, Detection of endogenous telomeric R-loops on dot-blots as in Fig. 2a. Cells were transfected with siCONTROL (siC), siRNH1 or siRAD51. **b**, Possible models for the roles of RAD51 in mediating TERRA–telomere associations. Upper row, RAD51 binds telomeric single-stranded DNA, inducing strand invasion of another telomere. TERRA hybridizes to the displaced strand upon branch migration. Lower row, RAD51 binds TERRA directly and initiates homology search and strand invasion by TERRA at a telomere. **c**, Western blot analysis of RAD51 and hnRNPA1 upon immunoprecipitation (IP) of native RNA (see Fig. 3a for RNA analysis). SN, supernatant. **d**, Native Immunoprecipitation of RNA was performed in U2OS cell extracts, demonstrating the association of TERRA with RAD51 and hnRNPA1. $n = 3$ biologically independent experiments; data are means ± s.d. **e**, Affinity of RAD51 for TERRA and non-telomeric-RNA (Non-TelRNA) oligonucleotides. Quantification is shown on the right. $n = 3$ independent experiments; data are means ± s.d. **f**, Top, stability EMSA assay. Bottom, 20 nM TERRA (lanes 2-7) or TelDNA (lanes 9-14) oligonucleotides were pre-bound with RAD51 (2.2 μM or 8.8 μM) and challenged with increasing concentrations of unlabelled competitor ssDNA (cDNA; 0.28 μM, 0.56 μM, 1.13 μM, 2.27 μM or 4.54 μM). Quantification is shown on the right. $n = 3$ independent experiments; data are means ± s.d.



Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | RAD51 catalyses the formation of canonical R-loops. **a**, Left, the RAD51 II3A mutant (lines 2–5; 325 nM, 650 nM, 1,300 nM or 2,600 nM) or wild-type RAD51 (lines 6 and 7; 325 nM or 650 nM) was incubated with TERRA oligonucleotide substrate (50 nM), and then plasmid containing the homologous region was added. Right, quantification of R-loops in the presence of wild-type RAD51 or RAD51 II3A. $n = 2$ independent experiments; data are means ± s.d. **b**, Top, the R-loop and D-loop assays. After RAD51-mediated R-loop or D-loop formation, proteins were digested with proteinase K (PK), which was then inactivated with PMSF and EGTA. Bottom left, R-loops or D-loops were detected on native gels as indicated. Treatment with RNaseH1, which degrades the RNA moiety in RNA–DNA hybrid structures, eliminates R-loops but has no effect on D-loops. RNaseH2, which cleaves ribonucleotides in DNA, has no effect, as expected. Right, quantification. $n = 3$ independent experiments; data are means ± s.d. **c**, Top, the R-loop and D-loop assays. Middle and bottom, RAD51-mediated R-loops (middle; $n = 3$ independent experiments; data are means ± s.d.), unlike D-loops (n=1 experiment), are recognized by the S9.6 antibody and supershifted in presence of both S9.6 and anti-mouse IgG. Quantification is shown on the right.

Extended Data Table 1 | Oligonucleotides used herein

Primers used for Chimeric TERRA constructs

Name	Forward	Reverse
Amplification of PP7 stem loops	CCGCGGCCCCGAATTCTATCGATACTCGAGATCCTA	GCTGACTAGAGGATCCACACGCGTTCTCGATAATGAA
Amplification of TTAGGG repeats	TTAGGGTTAGGGTTAGGGTTAGGGTTAGGG	CCCTAACCCTAACCCTAACCCTAACCCTAA
Amplification of 15q subtelomere	GAACGCGTGTGGATCATTCTCCTCAGGTCAGACCCG	CTAACCCTAAGGATCCTAACCGTGACCCTGACCCCG
Amplification of Xq subtelomere	GAACGCGTGTGGATCCCGAGTTGCGTTCTCG	CTAACCCTAAGGATCGCACATGAGGAATGTGGGTG
Amplification of AAVS1 left homology arm	TGACCGGTTGCTTTCTCTGACCAGCATTC	GTGTTAACCACTGTGGGGTGGAGGGGAC
Amplification of AAVS1 right homology arm	CGGATATCACTAGGGACAGGATTGGT	GTGATATCCTGTAGGAAGGGGCAGGAGA

Primers used for RT-qPCR

Name	Forward	Reverse
15q TERRA	CAGCGAGATTCTCCCAAGCTAAG	AACCCTAACCACATGAGCAACG
GAPDH	AGCCACATCGCTCAGACAC	GCCCAATACGACCAAATCC
TRF1	CTTGCCAGTTGAGAACGATATACA	CATCAGGGCTGATTCCAAGG
TRF2	GGGTTATGCAGTGTCTGTCGCG	CAGTGGTGTGAGCTCAGCCT
POT1	CCAAGCTCTGGATCAGTATCATT	CATAGTGGTGTCCTCTC AAATAC
SMG1	CCAAGCACCGTTCCAGGAACTG	CTCTCTTGCACCGCTTTCCCAG
UPF1	CGCAGGGCTACATCTCCATGAG	CTCGTCACCAAGGTAACTGTCCTG
UPF2	GGCTGAGTCTGCAGACACAATGC	GCAGCAAGTTGAGAGGACATGGG
RNaseH1	GGCCAGGCCATCCTTTAAATGTAGG	GCTTGTTCAATGGCTTTGCAGGC
RPA2	GCAGGGAACTTTGGTGGGAATAGC	CCCTTCAGGTCTTGGACAAGCC
RAD51	GAGGAAAGGAAGAGGGGAAACCAG	CCCACTCCATCTGCATTAATGGCG
1q subtelomere	CAGCGTCGCAACTCAAATG	CCCTCACCCTCCATGAGTAATA
10q subtelomere	GCATTCCTAATGCACACATGAC	TACCCGAACCTGAACCCTAA
13q subtelomere	GCACTTGAACCCTGCAATACAG	CCTGCGCACCGAGATTCT
15q subtelomere	AACCCTAACCACATGAGCAACG	GCTGCATTAAAGGGTCCAGT

siRNAs

Gene name	Dharmacon catalog number	siRNA sequences
siControl	D-001206-13-20	UAGCGACUAAACACAUCAA, UAAGGCUAUGAAGAGAUAC,
siTRF1	M-010542-02	AUGUAUUGGCCUGUAUUAG, AUGAACGUGAAUUGCUCAA
siTRF2	M-003546-00	CAAGAUAAACCUAGUGGUA, GGUGAUCCAAAUUCUCAUA,
siPOT1	M-004205-01	GGAAACUGGUCUAAAAUAC, GCCAGUUGAGAACGAUAUA
siSMG1	M-005033-01	GAAGUGGACUGUAGAAGAA, GGAAGCUGCUGUCAUUAUU,
siUPF1	M-011763-01	GGAAUCAGCUAUCAAUGUG, GAAGACAGUACAACCAAUA
siUPF2	M-012993-01	AGAAAGAUGUCAACAGCUA, GAGGCAAAGAAUCGAAAUA,
siRNaseH1	M-012595-00	CAGGAGUACUAGAAGCCUA, UGCAAGAUCUCCACGUUAA
siRPA2	M-017058-00	GUGAAGAUGUUCCCUAUGA, GAGGUUAGCUGCGGAAAGA,
siRAD51	M-003530-04	GGUCAGACAUCCACCAGAA, UAACUUGGCUCAGCUGUAU
siBRCA2	M-003462-01	GCUCCUACCUGGUGCAGUA, UCAAGGUCCCUGAUAAUUA,
		GGAAGUCGACCUCCUUUGA, CAAGAUAACAUCACUGUCA
		GGAACGAGAAUUCUUAAUA, GCAUGUACCUUGUGUAGAA,
		GAAGAUAUUCGAUUAGGAA, GGUCUAGAGAGUUGCGAAU
		GCGCAGAGCCGUAUGCAAA, GAGCUAAACAAUCGGAAGA,
		GCCAGGCCAUCCUUUAAAU, GACAUUCAGUGGAUGCAUG
		GAUCAAUGCACACAUGGUA, CAAAAUAGAUGACAUGACA,
		GAGUGAAGCAGGGAACUUU, GUGGAACAGUGGAUUCGAA
		GAAGCUAUGUUCGCCAUUA, GCAGUGAUGUCCUGGAUAA,
		CCAACGAUGUGAAGAAAUU, AAGCUAUGUUCGCCAUUAA
		GAAACGGACUUGCUAUUUA, GUAAAGAAAUGCAGAAUUC,
		GGUAUCAGAUGCUUCAUUA, GAAGAAUGCAGGUUUAAUA

Extended Data Table 2 | Antibodies used herein

Antibody	Company	Catalogue No	Dilution (Technique)
Anti-GFP	Home-Made		1:1,000 (IF)
Anti-GFP	Merck Millipore	MAB3580	1:1,000 (IF, WB)
Anti-RAD51	Santa-Cruz	sc-8349	1:1,000 (WB)
Anti-RAD51	ABCAM	ab133534	6 µg (IP)
Anti-BRCA2	Merck Millipore	MAB3580	1:1,000 (WB)
Anti-Vinculin	ABCAM	ab129002	1:10,000 (WB)
Anti-hnRNPA1	Santa-Cruz	sc-32301	1:1,000 (WB), 6 µg (IP)
Anti-RNaseH1	GeneTex	GTX-117624	1:1,000 (WB)
Anti-TRF1	Santa-Cruz	sc-6165R	1:1,000 (WB)
Anti-TRF2	Merck Millipore	05-521	1:1,000 (WB)
Anti-POT1	ABCAM	ab124784	1:1,000 (WB)
Anti-SMG1	Abgent	AP8055a	1:500 (WB)
Anti-UPF1	Bethyl Laboratories	A300-037A	1:1,000 (WB)
Anti-RPA2	ABCAM	ab2175	1:1,000 (WB)
Anti-DNA-RNA Hybrid [S9.6]	Kerafast	ENH001	1 µg / 10 µg nucleic acids (DRIP)
Anti-Rabbit IgG (H+L), HRP Conjugate	Promega	W4011	1:10,000 (WB)
Anti-Mouse IgG (H+L), HRP Conjugate	Promega	W4021	1:10,000 (WB)
Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 633	Thermo Fisher	A-21070	1:1,000 (IF)

IF, immunofluorescence; IP, immunoprecipitation; WB, western blot.

Chapter 3 The makings of TERRA R-loops at chromosome ends

Rita Valador Fernandes, Marianna Feretzaki and Joachim Lingner
Cell Cycle 20(18): 1745–1759 (2021)
https://doi.org/10.1080/15384101.2021.1962638

3.1 Abstract

"Telomeres protect chromosome ends from nucleolytic degradation, uncontrolled recombination by DNA repair enzymes and checkpoint signaling, and they provide mechanisms for their maintenance by semiconservative DNA replication, telomerase and homologous recombination. The telomeric long noncoding RNA TERRA is transcribed from a large number of chromosome ends. TERRA has been implicated in modulating telomeric chromatin structure and checkpoint signaling, and in telomere maintenance by homology directed repair, and telomerase – when telomeres are damaged or very short. Recent work indicates that TERRA association with telomeres involves the formation of DNA:RNA hybrid structures that can be formed post transcription by the RAD51 DNA recombinase, which in turn may trigger homologous recombination between telomeric repeats and telomere elongation. In this review, we describe the mechanisms of TERRA recruitment to telomeres, R-loop formation and its regulation by shelterin proteins. We discuss the consequences of R-loop formation, with regard to telomere maintenance by DNA recombination and how this may impinge on telomere replication while counteracting telomere shortening in normal cells and in ALT cancer cells, which maintain telomeres in the absence of telomerase."

3.2 Context/highlights

In this review article, we describe several factors contributing to the regulation of TERRA association with telomeres and telomeric R-loop formation. We discuss putative consequences of R-loop formation and accumulation, including the interference with semi-conservative replication of telomeric DNA and possible stimulation of homology-directed repair.

Furthermore, we contribute to the characterization of the RAD51-dependent pathway directing TERRA-mediated telomeric R-loop formation described in Feretzaki *et al*, 2020 (see Chapter 2). Namely, we show that the cell cycle control that governs endogenous TERRA levels is disrupted when transgenic TERRA is expressed from plasmids with a heterologous promoter. Additionally, we demonstrate that ectopically expressed TERRA is recruited to telomeres in and outside of S phase – as determined by IF-FISH combined with nuclear staining upon pulse labeling with EdU. Moreover, we analyze the cell cycle profile of RAD51-depleted cells, in the absence or presence of ectopically-expressed TERRA. We show that RAD51 depletion induces a slight increase in the percentage of cells in G2/M, and no substantial changes in the percentage of cells in S phase – when endogenous TERRA levels are the lowest in HeLa cells.

3.3 Erratum

In figures 2 and 3 of "The makings of TERRA R-loops at chromosome ends" by Valador Fernandes et al, 2021, the molecular weight marker in α-RAD51 blots shows bands corresponding to 35 and 40 kDa, instead of 25 and 35 kDa as indicated.

3.4 Author contributions

- Optimized and performed all experiments, namely:
  - Cell culture, cell transfections and EdU incorporation.
  - Coupling of EdU click-it with immunofluorescence and FISH.
  - Acquisition and analysis of microscopy data.
  - Ethanol fixation of cells and DAPI staining for flow cytometry.
  - Acquisition and analysis of cell cycle distribution data, assisted by Flow Cytometry Research Core Facility of EPFL.
- Prepared all figures, schemes, and respective legends.
- Contributed to the text written by J.L..

CELL CYCLE
2021, VOL. 20, NO. 18, 1745–1759
https://doi.org/10.1080/15384101.2021.1962638

Taylor & Francis
Taylor & Francis Group

REVIEW

The makings of TERRA R-loops at chromosome ends

Rita Valador Fernandes, Marianna Feretzaki, and Joachim Lingner

School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Swiss Institute for Experimental Cancer Research (ISREC), Lausanne, Switzerland

ABSTRACT

Telomeres protect chromosome ends from nucleolytic degradation, uncontrolled recombination by DNA repair enzymes and checkpoint signaling, and they provide mechanisms for their maintenance by semiconservative DNA replication, telomerase and homologous recombination. The telomeric long noncoding RNA TERRA is transcribed from a large number of chromosome ends. TERRA has been implicated in modulating telomeric chromatin structure and checkpoint signaling, and in telomere maintenance by homology directed repair, and telomerase – when telomeres are damaged or very short. Recent work indicates that TERRA association with telomeres involves the formation of DNA:RNA hybrid structures that can be formed post transcription by the RAD51 DNA recombinase, which in turn may trigger homologous recombination between telomeric repeats and telomere elongation. In this review, we describe the mechanisms of TERRA recruitment to telomeres, R-loop formation and its regulation by shelterin proteins. We discuss the consequences of R-loop formation, with regard to telomere maintenance by DNA recombination and how this may impinge on telomere replication while counteracting telomere shortening in normal cells and in ALT cancer cells, which maintain telomeres in the absence of telomerase.

ARTICLE HISTORY
Received 8 June 2021
Revised 23 July 2021
Accepted 28 July 2021

KEYWORDS
Telomeres; TERRA; R-loops; RAD51; shelterin proteins; homologous recombination

Introduction

Telomeres correspond to the physical ends of eukaryotic chromosomes. They consist of short tandem DNA repeats that are generally rich in guanine and thymine bases. In vertebrates, telomeres consist of 5′-TTAGGG-3′ repeats in the DNA strand containing the 3′ end. The overall length of human telomeres varies between roughly 3,000–15,000 bp. Telomeres have a 3′ overhang of roughly 50–200 nucleotides [1,2] and are associated with several hundred proteins [3,4,5]. Most abundant and best characterized is the shelterin protein complex comprising up to six different polypeptides [6]. TRF1 (telomeric repeat-binding factor 1) and TRF2 bind as homodimers the double stranded part of telomeres. POT1 (protection of telomeres 1) binds specifically to the single-stranded 5′-TTAGGG-3′ repeats which are present at the 3′ end of telomeres [7]. Alternatively, POT1 is thought to bind to single-stranded 5′-TTAGGG -3′ repeats present as displaced strand internally, when telomeres adopt the T-loop configuration (Figure 1). In T-loops, the telomeric 3′ overhang is tucked into the double-stranded part of the telomere, base pairing with the complementary 5′-CCCTAA-3′ repeats [8]. POT1 can be physically linked to TRF1 and TRF2 through protein interactions involving the shelterin components TPP1 and TIN2 (TRF1-interacting nuclear factor 2). TPP1 also enhances the affinity of POT1 for telomeric DNA, and the interaction of POT1 with the other shelterins is required for telomere association of POT1, possibly by increasing its local concentration [9]. Rap1 (repressor activator protein 1) is recruited to vertebrate telomeres via TRF2.

A large number of additional telomeric proteins that are less abundant at chromosome ends have been identified through studies of telomere maintenance by telomerase, genetic and protein interaction screens and upon purification of crosslinked telomeric chromatin, followed by mass spectrometry analysis [3,4,5]. Most of these proteins are not telomere specific and many of them bind to telomeres only under certain conditions. Notably, a significant number of proteins identified at telomeres are linked to RNA

CONTACT Joachim Lingner joachim.lingner@epfl.ch School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Swiss Institute for Experimental Cancer Research (ISREC), Lausanne 1015, Switzerland

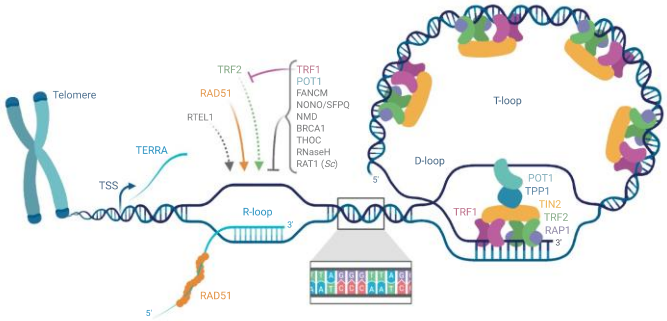


Figure 1. Cartoon of the nucleoprotein structure of chromosome ends; telomeric R-loops and its regulators. Among the numerous proteins associating with telomeres, the shelterin protein complex components are the most abundant, comprising the dsDNA-binding proteins TRF1 and TRF2, the ssDNA-binding protein POT1, TIN2, TPP1 and Rap1. Telomeres have a 3′ overhang, which can invade the dsDNA region, forming a D-loop which is termed T-loop. Telomeres are transcribed into TERRA. TERRA transcription starts at subtelomeric regions, extending toward chromosome ends. The vast majority of TERRA molecules is not polyadenylated and largely colocalizes with telomeres. TERRA binding to telomeres can occur through direct base-pairing with telomeric DNA, forming R-loop structures, leaving a displaced G-rich DNA strand. Factors which are suspected to regulate TERRA association with telomeres are indicated. TERRA association and R-loop formation at telomeres depend on the RAD51 DNA recombinase, which binds TERRA and catalyzes TERRA R-loop formation *in vitro*. Also, the helicase RTEL1 was proposed to stimulate TERRA association with chromosome ends via R-loops. The shelterin TRF2 can also bind TERRA and stimulate R-loop formation *in vitro* – a process counteracted *in vivo* by TRF1. Loss of POT1 was found to result in a striking accumulation of telomeric R-loops, suggesting a role in preventing the accumulation of such structures. Proteins involved in the NMD RNA surveillance pathway, including UPF1, UPF2 and SMG1, were also shown to prevent TERRA association with telomeres. The DNA recombination factor BRCA1 was as well shown to modulate TERRA binding to telomeres, preventing R-loop-associated telomeric DNA damage. The ATPase/translocase FANCM resolves RNA:DNA hybrids *in vitro* and counteracts telomeric R-loop accumulation in ALT cells. The THO multi-subunit complex is present at human and budding yeast telomeres, and was found to prevent R-loop accumulation and telomere shortening in budding yeast. RNase H enzymes – which remove RNA-DNA hybrids through the endonucleolytic cleavage of the engaged RNA moiety – were found to regulate TERRA R-loops, with a prominent role in ALT cells. In *S. cerevisiae*, Rat1 nuclease, as well as RNase H2, are preferentially recruited to long telomeres in S phase, preventing TERRA accumulation and formation of R-loops prone to pose an obstacle to telomeric replication. The Illustration was created with BioRender.com.

metabolism, which can be rationalized by the fact that telomeres are transcribed into the telomeric repeat containing RNA (TERRA) [10,11].

TERRA has been detected in a large number of eukaryotes including humans and other vertebrates [10,11], yeast [12,13,14], plants [15] and protozoa [16]. Thus, telomere transcription occurs despite the fact that telomeric chromatin has heterochromatic characteristics [17,18]. At human telomeres, histone H3 is frequently trimethylated at Lys9 and Lys27, and nucleosomes are narrowly spaced [18] due to the lack of the linker histone H1[3]. The discovery of telomere transcription was also counterintuitive at the time, as the telomere position effect – which reflects the variegated repression of genes experimentally placed next to telomeres – suggested a general absence of transcription at telomeres [19,20]. TERRA transcription proceeds from promoters residing in the subtelomeric regions into the telomeric tract [10,21–23]. Thus, human TERRA starts with sequences stemming from different subtelomeres and it ends with numerous 5′-UUAGGG-3′ repeats [24]. In *Trypanosoma brucei*, telomere transcription occurs mainly at the telomere which contains the active VSG gene [25]. In other eukaryotes where this has been characterized, telomere transcription is common to several or all chromosome ends. The assignment of TERRA molecules to individual chromosome ends is ambiguous in many species including humans, as subtelomeres are rich in repetitive sequences which have not been well annotated. Interestingly, the regulation of TERRA expression

at individual chromosome ends may differ from one to another. In *Saccharomyces cerevisiae* for example, the major double strand telomere-binding protein Rap1p regulates TERRA transcription and degradation [26]. At subtelomeres containing the so-called Y' repeat elements, Rap1p recruits Rif1 and Rif2 to downregulate TERRA transcription. At subtelomeres containing X repeat elements, Rap1-mediated TERRA repression involves the Sir2/3/4 histone deacetylase and Rif1/2 complexes. Also in humans, TERRA transcription has been detected at a large number of chromosome ends. At some telomeres, TERRA promoters harbor CpG islands which are negatively regulated by DNA methyltransferases that can modify these sequences [22,23,27]. A second class of TERRA promoters lacks CpG islands and is insensitive to DNA methylation [23]. In mouse cells, TERRA expression has also been detected at several chromosome ends [28]. In addition, a more abundant noncoding 5′-UUAGGG-3′ repeat containing RNA termed PAR-TERRA which is involved in the pairing of homologous sex chromosomes has been described. The PAR-TERRA primary structure has not been elucidated so far and its transcription has not been demonstrated to proceed into the terminal telomeric repeats, as seen for canonical TERRA transcripts [29]. CHIRT-seq experiments indicate binding of PAR-TERRA throughout the genome as well as at chromosome ends, suggesting functions that are particular to this RNA. We do not cover PAR-TERRA in this review.

Telomeres are dynamic structures. They change their composition during the cell cycle and upon shortening and damage, to mediate DNA damage checkpoint signaling and repair. Telomere shortening is a common phenomenon in human somatic cells, as differentiated cells do not express the catalytic subunit of telomerase hTERT [30]. Short telomeres elicit ATM- and ATR-dependent checkpoint signaling to induce cellular senescence which represses the growth of precancerous lesions that have lost normal growth control [31]. At short telomeres, T-loops are thought to unfold due to low levels of TRF2, which enables ATM activation. Low levels of POT1 allow binding of RPA to single-stranded telomeric DNA, leading to ATR-ATRIP recruitment and checkpoint activation. TERRA is another major player which triggers local remodeling events at individual chromosome ends that may have suffered from damage or severe shortening. TERRA expression is enhanced at telomeres from which TRF2 has been depleted [21] and its expression increases from telomeres when they get shorter in budding yeast [32] and in human cells [33]. Thus, through its accumulation at short or damaged telomeres through increased expression and recruitment, TERRA can act as a recruitment platform for DNA repair enzymes at telomeres that need their attention.

TERRA R-loop formation and regulation at chromosome ends

TERRA is transcribed at human and yeast telomeres by RNA polymerase II. The 5′-UUAGGG-3′ tract of human TERRA is heterogeneous in length. For most TERRA molecules the length varies between 100–400 nucleotides [24], though longer TERRA transcripts approaching the ends of chromosomes can also be detected [27,33]. More than 90% of human TERRA is not polyadenylated, frequently terminating with the sequence 5′-UUAGG -3′[24]. The mechanisms of 3′end formation are not well understood. The 3′ends of the non-polyadenylated fraction might simply occur due to termination of transcription within telomeric chromatin that might impede efficient read-through. The poly(A) tail of the polyadenylated fraction of TERRA is generated by the canonical poly(A) polymerase, at least in budding yeast [12]. If polyadenylation of TERRA is preceded by RNA cleavage as for mRNAs is not known.

Analyses by fluorescence in situ hybridization and cellular fractionation studies indicate that TERRA is enriched at human telomeres [10]. In addition, approximately half of TERRA is not tightly associated with chromatin. Notably, all poly(A)$^+$ TERRA was detected in a nucleoplasmic fraction and not on chromatin [24] suggesting that its functions do not involve physical interactions with telomeres. In contrast, non-polyadenylated TERRA colocalizes to a large extent with telomeres. In principle, it may be retained at telomeres through interactions with telomere binding proteins or through base-pairing with telomeric DNA,

forming so-called R-loop structures in which the TERRA:telomeric DNA RNA/DNA helix causes displacement of the telomeric strand containing 5′-TTAGGG-3′ DNA repeats (Figure 1). Several proteins have been identified to regulate TERRA association with telomeres. Using a reporter system in which TERRA was expressed ectopically with an RNA tag (PP7 stem loops) that is recognized by the bacteriophage PP7 coat protein fused to GFP, we could demonstrate that TERRA associates with telomeres post transcription, through the formation of R-loop structures [34] (Figure 2). This association is dependent on TERRA's 5′-UUAGGG-3′ repeats suggesting that the R-loops form within the telomeric repeats. However, the length and exact position of the R-loop structures has not been determined. Telomere association and R-loop formation depend on the RAD51 DNA recombinase [34] since its depletion reduces TERRA foci at telomeres (Figure 2). Interestingly, while endogenous TERRA is rare in S phase, transgenic TERRA expressed from plasmids with a heterologous promoter loses this control and efficiently associates with telomeres in S phase nuclei that were identified through pulse labeling with EdU (Figure 2). This association is also reduced upon RAD51 depletion indicating the involvement of RAD51. The presence of transgenic TERRA at telomeres during replication can explain its interference with telomere replication (see below). Also of note, though RAD51 is an essential protein, its substantial depletion by siRNA only slightly increased the number cells in G2/M and it did not substantially affect the fraction of cells in S phase during which endogenous TERRA levels are low (Figure 3).

Since RAD51 promotes homology search and strand invasion of recombining DNA molecules, the data suggested that RAD51 may also home TERRA through an analogous mechanism. In support of this model, RAD51 is bound to TERRA in cellular extracts and it binds the 5′-UUAGGG-3′ repeats of TERRA *in vitro* with high affinity. Furthermore, RAD51 catalyzes strand invasion of TERRA into plasmid DNA *in vitro* and it is required for the formation of telomeric R-loops by endogenous TERRA [34]. It will be important to determine how RAD51 distinguishes TERRA from the bulk of other nuclear RNAs it may not act upon. The 5′-UUAGGG-3′ repeats of TERRA provide a unique signature for this RNA. Also, these G-rich repeats can form G-quadruplex structures which might be recognized by RAD51. Alternatively, other TERRA-binding proteins might facilitate a preferential association of RAD51 with TERRA. BRCA2 facilitates RAD51 binding to single-stranded DNA during double strand break repair by homologous recombination. BRCA2 depletion reduced TERRA association with telomeres, but it also diminished the presence of RAD51 in the nucleus. Thus, it remains uncertain if BRCA2 promotes RAD51 binding to TERRA. Notably, these results are strikingly different from what has been seen for highly transcribed genes, which may retain R-loops from transcription and for which BRCA2 was reported to diminish R-loops through a collaboration with the TREX-2 mRNP biogenesis and export complex [35]. Interestingly, BRCA1 – which among others also promotes RAD51-mediated homologous recombination for DNA repair – is a second DNA recombination factor that was recently shown to physically interact with TERRA as well as TERRA R-loops [36]. In contrast to RAD51, however, BRCA1 counteracts telomeric R-loops [36]. The underlying mechanism remains uncertain, but it was proposed to involve BRCA1 interactions with XRN2, which is the ortholog of yeast Rat1. Rat1 is a 5′-3′ RNA exonuclease which is involved in transcription termination, among others. Rat1 degrades TERRA in *S. cerevisiae* [12]. To what extent other recombination factors may regulate TERRA at telomeres remains to be investigated. The recent findings already hint toward important differences between RNA and DNA mediated homology search. However, if TERRA strand invasion of telomeric DNA initiates at TERRA 3′ends this mechanism could explain why poly(A)$^+$ TERRA is not retained on chromatin.

While RAD51 is likely to catalyze strand invasion of TERRA into telomeric DNA, several telomere-associated proteins have also been identified to regulate strand invasion and telomere retention (Figure 1). Among the shelterin components, TRF1, TRF2 and POT1 play critical roles. The N-terminal basic domain of TRF2 can bind TERRA and stimulate R-loop formation *in vitro*

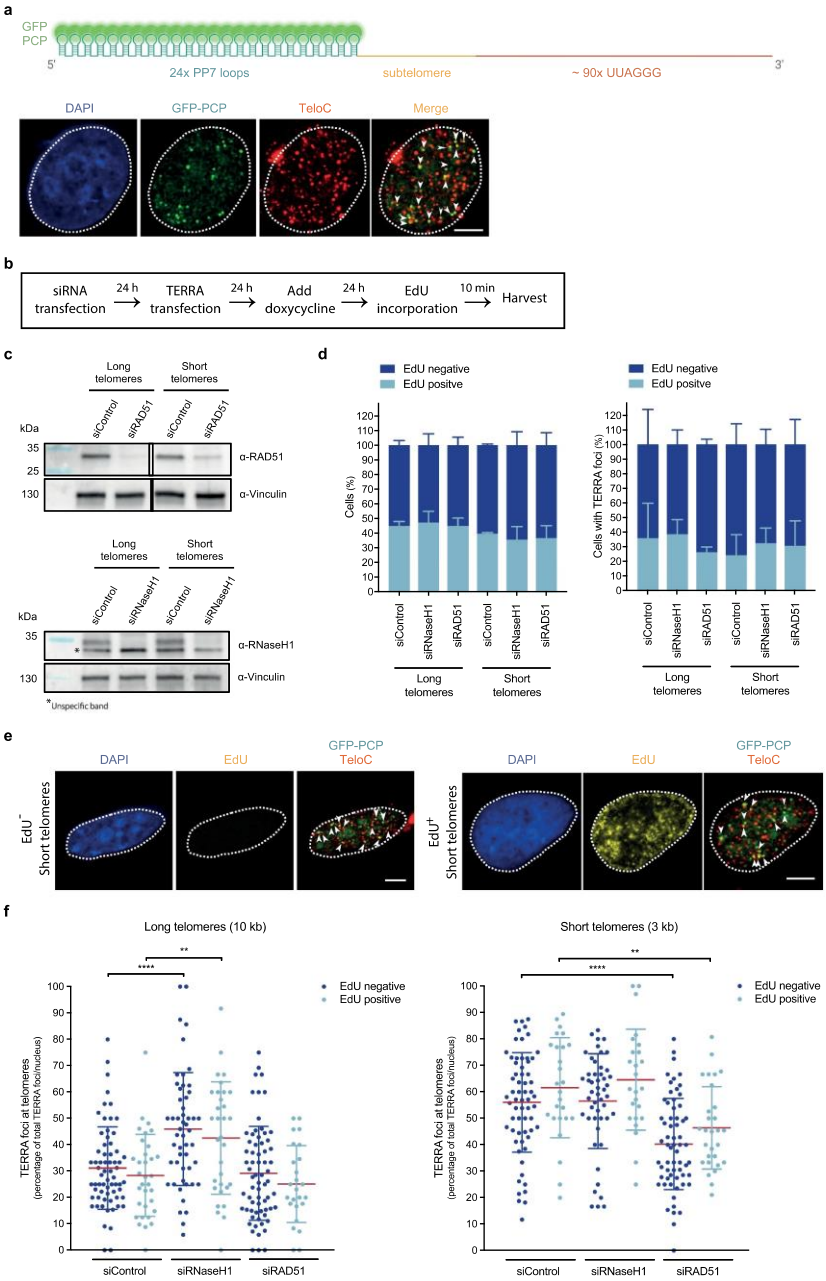


Figure 2. Transgenic TERRA associates with telomeres in and outside S phase. a) Top: Depiction of transiently expressed chimeric TERRA, comprising twenty-four *Pseudomonas aeruginosa* phage 7 (PP7) stem-loops – recognized by GFP-tagged dimerized PP7 Coat Protein (PCP) –, and a subtelomere-derived sequence, followed by UUAGGG tandem repeats. Illustration created with BioRender.

com. Bottom: Immunofluorescence of GFP (green) was employed to analyze co-localization of transiently expressed PP7-fused 15q-TERRA transcripts with telomeres (red) identified by Fluorescence in situ hybridization (FISH) (as described in ref 34). Representative images are shown and were acquired with a Leica SP8 confocal microscope. White dashed line outlines the nuclear region and was determined based on DAPI-staining. White arrowheads indicate co-localization of GFP-PCP with telomeric FISH signals. Scale bar indicates 5 µm. b) HeLa cells with long (10 kilobase average) or short (3 kb average) telomeres were transfected with siRNA pools to down-regulate RAD51 or RNase H1 mRNA levels. PP7-15q-TERRA-coding constructs were transfected and their expression was induced with doxycycline for 24 hours. Cells were then pulse-labeled with 10 µM of 5-ethynyl-2′-deoxyuridine (EdU) (Invitrogen) for 10 min and harvested. c) Western blotting was used to evaluate knockdown efficiency of RAD51 (top) or RNase H1 (bottom). Vinculin is shown as a loading control. Representative blots of three biologically-independent experiments are shown. d) Percentage of EdU-positive cells in the total population of cells (left) or cells displaying TERRA foci (right), across indicated conditions. After EdU pulse-labeling, cells were fixed in 4% paraformaldehyde for 10 min at room temperature. Anti-GFP immunofluorescence was performed as described in ref 34. After fixation of bound primary and secondary antibodies, cells were permeabilized with a detergent solution (0.1% Triton X-100, 0.02% SDS in 1x PBS) for 5 min, followed by pre-blocking with 2% bovine serum albumin (BSA) in 1x PBS for 30 min. Cells were then incubated with a click-it reaction (4 mM copper sulfate, 100 mM sodium ascorbate and 4 µM Alexa Fluor 488 Azide (Invitrogen) in 1x PBS) for 30 min in a humidity-chamber, followed by three 1x PBS washes, permeabilization with a detergent solution (indicated above) for 3 min and 4% paraformaldehyde fixation for 5 min. FISH staining was then carried out following the procedure described in ref 34. At least 460 total cells and 74 cells displaying TERRA foci were analyzed per condition, across three independent biological replicates. Data are means ± s.d. e) Representative images of EdU-negative (left) and EdU-positive (right) HeLa cells with short telomeres, obtained as described in d. Representative images were acquired with a Leica SP8 confocal microscope. White dashed line outlines the nuclear region and was determined based on DAPI-staining. EdU signal is shown, as well as GFP-PCP and TeloC merged signals. White arrowheads indicate co-localization of GFP-PCP with telomeric FISH signals. Scale bars indicate 5 µm. f) The percentage of PP7-15q-TERRA foci colocalizing with telomeric FISH signals per nucleus was assessed by GFP immunofluorescence, EdU Click-it and telomeric FISH as described in d. At least 74 cells were analyzed per condition, across three independent biological replicates. Data are means ± s.d. One-way analysis of variance (ANOVA) with Dunnett's multiple comparisons test was used, comparing all conditions with non-targeting siRNA (siControl): **$P < 0.01$; ****$P < 0.0001$. All statistical analysis was performed using GraphPad Prism. All images were processed and analyzed with Image J.

[37]. This activity is prevented *in vitro* and *in vivo* by TRF1 through its N-terminal acidic domain. Whether TRF2 promotes R-loops under physiological conditions is not known, though TRF2 depletion *per se* does not decrease R-loop formation at telomeres [34,37]. POT1 binds specifically to the single-stranded G-rich telomeric DNA but not to the corresponding 5′-UUAGGG-3′ repeats in TERRA [38]. *POT1* deletion in human cells causes rapidly dramatic telomere elongation by RAD51-mediated homologous recombination [39]. At the same time the telomeric R-loops are increased. It is unclear if TERRA R-loops are formed as an immediate consequence of POT1-loss and if they are a prerequisite for telomere recombination. Perhaps more likely, POT1-loss liberates the single-stranded telomeric DNA for RPA binding which subsequently, with assistance of BRCA2, becomes replaced by RAD51. The increased local concentration of RAD51 may facilitate the binding of TERRA by RAD51 in a succeeding step, which will trigger strand invasion and R-loop formation. The DNA strand that is displaced by the TERRA R-loop may recruit additional RPA and RAD51 in a feedforward loop to first promote RPA-dependent DNA checkpoint signaling and subsequently replacement by RAD51 to mediate homologous recombination (Figure 4).

The FANCM protein, whose mutation has been associated with Fanconi Anemia, is an ATPase associated with DNA branch migration. Its roles at telomeres have been characterized in U2OS cells [40,41], which are ALT cells (for alternative lengthening of telomeres) that use DNA recombination to maintain telomeric DNA repeats. Strikingly, FANCM depletion caused an increase in R-loops both at telomeres and elsewhere in the genome. In addition, FANCM can resolve telomeric R-loops *in vitro* using its RNA:DNA helicase activity suggesting that it directly participates in R-loop resolution [40]. NONO/SFPQ heterodimers, which are involved in various aspects of RNA metabolism in the nucleus, have also been implicated in suppressing telomeric R-loops [42]. Their depletion increased R-loop signals partially colocalizing with telomeres in nuclei of U2OS cells. Most recently, the RTEL1 helicase has been implicated in TERRA regulation [43]. RTEL1 has well-established crucial roles during telomere replication resolving T-loops in S phase as well as telomeric G-quadruplex structures [44]. The new work shows that RTEL1 also binds *in vitro* TERRA

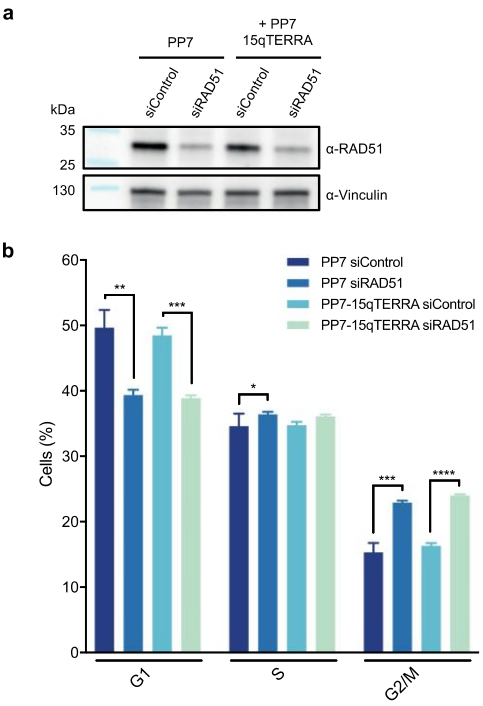


Figure 3. The fraction of cells in S phase is not strongly impacted by siRNA-mediated depletion of RAD51, with or without PP7-15q-TERRA inducible expression. a) Western blotting was used to evaluate siRNA-mediated knockdown efficiency of RAD51, with and without PP7-15q-TERRA inducible expression. Vinculin is shown as a loading control. Representative blots of three biologically-independent experiments are shown. b) Distribution of cells of indicated conditions in G1, S or G2/M cell cycle phases. To evaluate cell cycle distribution of cells with depleted RAD51, DNA content was assessed by flow cytometry analysis of fixed DAPI-stained cells. Briefly, 1 million cells/condition was harvested by trypsinization. Cells were washed twice with cold 1x PBS and resuspended in 1 ml cold 70% ethanol while vortexing. After 30 min, ethanol was aspirated and fixed cells were resuspended in 1x PBS containing 0.2 µg/ml RNase, DNase-free (Merck) and incubated at 37°C for 15 min. 1x PBS containing 2 µg/ml DAPI (BioChemica) was then added to stain DNA. Samples were processed with a BD LSR Fortessa. 20.000 events were acquired per condition, per biological replicate. Data are means ± s.d. of three independent biological replicates. Two-tailed unpaired t-tests were used to calculate P-values: *P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.0001. Flow cytometry data were analyzed with FlowJo and statistical analysis was performed using GraphPad Prism.

5′-UUAGGG-3′ repeats when adopting a G-quadruplex structure [43]. Strikingly, *RTEL1* deletion caused strong increase in TERRA levels, while TERRA association with telomeres was significantly diminished. It was proposed that RTEL1 facilitates TERRA association with chromosome ends through a stimulatory effect on telomeric R-loops, which in turn could prevent additional transcription from the R-loop containing telomeres [43]. This proposed mode of action deviates from the documented roles of RTEL1 at G-quadruplex forming DNA sequences elsewhere in the genome where RTEL1 dismantles R-loops [45]. Therefore, it will be interesting to further dissect the mechanism and also test if and to what extent the S phase-specific telomere recruitment of RTEL1 by dephosphorylated TRF2 may contribute [46]. Certainly, also other models could be at play. For example, if RTEL1 affected TERRA 3′ end formation preventing its polyadenylation, absence of RTEL1 and TERRA polyadenylation might also stabilize this RNA and trigger its dissociation from chromosome ends [24].

Proteins involved in nonsense mediated RNA decay (NMD) of mRNAs containing premature stop codons also regulate TERRA stability and R-loop formation. These proteins, though most abundant in the cytoplasm, are also present in nuclei and can be detected at telomeres by chromatin immunoprecipitation, suggesting direct roles at telomeres [10]. Depletion of the NMD proteins UPF1, UPF2, SMG1 and SMG6 increased TERRA foci at telomeres, while not markedly impacting on total TERRA levels. This suggested that these proteins displace TERRA from telomeres, though they might also degrade this RNA locally at chromosome ends. Consistent with such models, depletion of UPF1, UPF2 and SMG1 also increased association of transgenic TERRA with telomeres [34]. Live cell imaging of TERRA should allow to discriminate if NMD proteins may influence the dynamics of TERRA interaction with chromosome ends or if they are involved in its local destruction at the telomere.

Regulation of TERRA during the cell cycle and with telomere length

Telomeres change their composition during S phase of the cell cycle to mediate their replication by the semiconservative DNA replication machinery and end maintenance by telomerase.

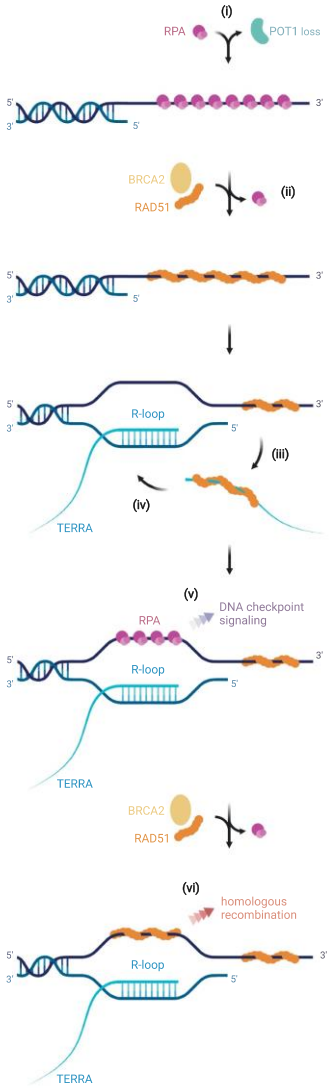


Figure 4. Possible R-loop- and RAD51-mediated telomere elongation mechanism upon loss of POT1. Loss of the ssDNA-binding shelterin POT1 at human telomeres leaves exposed the telomeric 3′ overhang – which can be bound by RPA (i). With assistance of BRCA2, RAD51 may displace and replace RPA (ii). This increased concentration of RAD51 at telomeres may facilitate binding of RAD51 to TERRA (iii), which subsequently can trigger strand invasion and R-loop formation (iv). Upon R-loop formation, the displaced single-stranded DNA strand can be bound by RPA – serving as a key platform for ATR activation (v). Assisted by BRCA2, RAD51 may then substitute RPA, mediating homologous recombination at the telomere (vi) eventually resulting in HDR-mediated telomere elongation (observed as a striking consequence of conditional deletion of *POT1* in human cells). Illustration created with BioRender.com.

Also, TERRA levels are regulated during the cell cycle. In budding yeast, TERRA levels and telomeric R-loops increase from G1 to S phase, but they drop during S phase, reaching lower levels by the time of semiconservative DNA replication of telomeres in late S phase – possibly to prevent interference with the replisome [32]. The low levels of TERRA in late S phase are a consequence of the recruitment of Rat1 nuclease to telomeres at this stage. In human cells, TERRA levels also decrease during S phase, reaching lowest levels as cells proceed from late S to G2 [24]. This cell cycle control is lost in cells carrying defects in the chromatin remodeling enzyme ATRX or the DNA methyl transferase DNMT3b (see below). However, the exact mechanisms of regulation at the transcriptional and posttranscriptional level remain to be elucidated.

TERRA has been suspected to regulate telomerase activity, and *in vitro* TERRA is a potent inhibitor of the human telomerase enzyme as it tightly binds to the RNA template [47]. *In vivo* in budding yeast, however, positive roles of TERRA for telomerase have been proposed. TERRA colocalizes with telomerase in early S phase, which then has been suggested to guide telomerase to specifically short telomeres promoting their preferential elongation [48]. Interestingly, TERRA molecules from short telomeres appeared to reassociate with the short telomeres from which they originated, which could be explained by a homology-driven search mechanism that we describe in this review for human TERRA. In fission yeast, TERRA has also been proposed to be a positive regulator of telomerase, but a positive effect of TERRA overexpression on telomere length was only observed upon treatment of cells with histone deacetylase inhibitors [49]. Finally, roles of TERRA have been postulated in human cells for promoting POT1 association with the G-rich telomeric strand after DNA replication [50]. hnRNPA1 is a major TERRA-binding protein, binding its 5′-UUAGGG-3′ repeats [51], while also having strong affinity for single-stranded 5′-TTAGGG-3′ DNA repeats [52]. As TERRA levels decline in late S phase, hnRNPA1 is liberated from TERRA – now binding the single-stranded telomeric DNA. Thus, hnRNPA1 may displace RPA from the G-rich telomeric DNA strand, which is the major single-strand DNA-binding protein associating with all single-stranded DNA for replication. As TERRA reaccumulates after S phase, hnRNPA1 may reassociate with TERRA and liberate the single-stranded G-rich telomeric strand for POT1 binding. More detailed knowledge on the regulation of TERRA expression and its interaction with telomerase, hnRNPA1 and other protein partners during the cell cycle may be required to clarify these issues.

TERRA association with chromosome ends also increases as telomeres get shorter. In budding yeast, this is due to increased transcription [48] and the absence of Rat1 nuclease and RNase H2 (a member of the Ribonuclease H family of enzymes that specifically degrade the RNA moiety in DNA:RNA hybrid structures) at short telomeres [32]. Conversely, at longer telomeres, the telomeric Rif2 recruits efficiently RNase H2 and Rat1. In human cells, telomere transcription increases at short telomeres due to a decreased presence of trimethylated H3K9, which recruits HP1 [33]. In addition, ectopically expressed TERRA is recruited to short telomeres more efficiently by the RAD51 recombinase [34]. The underlying mechanism for this has not yet been uncovered in human cells. However, RNase H1 might be involved, as its depletion increased telomere association of ectopic TERRA strongly at long but not at short telomeres [34] (Figure 2).

Consequences of R-loop formation: replication interference and stimulation of HDR

The formation of TERRA-mediated telomeric R-loops appears to be beneficial when telomeres become damaged or when they are very short. In such cases, telomeres activate DNA damage checkpoint signaling to induce cell cycle arrest or cellular senescence, they may be fused to one another by DNA repair enzymes if they completely lose capping function and become mistaken as DNA double strand breaks [31]. Alternatively, short telomeres are healed and re-elongated by telomerase or by homology directed repair. For the latter, single-stranded G-rich telomeric DNA may invade the telomeric DNA of adjacent chromosomes which is followed by elongation of the invading

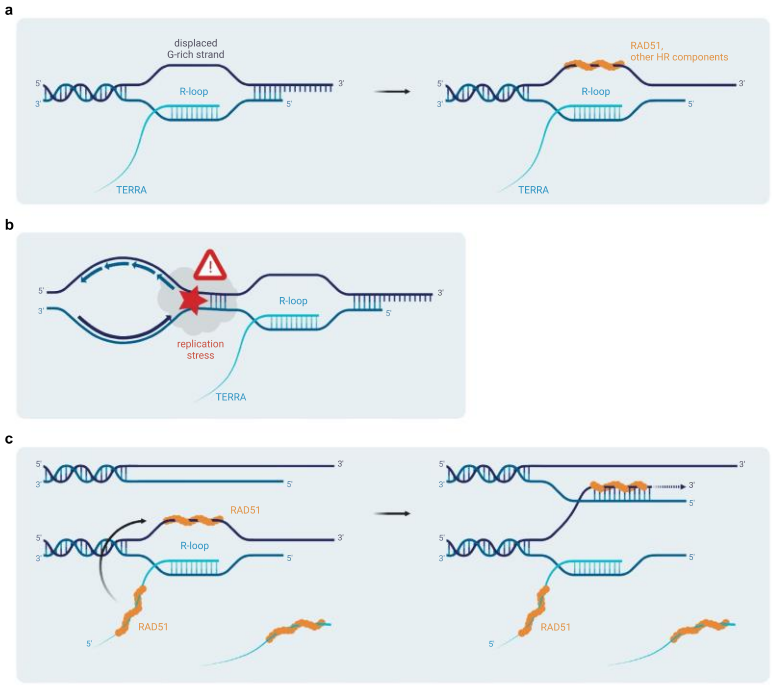


Figure 5. Possible mechanisms for TERRA R-loop-mediated stimulation of homologous recombination at telomeres. a) Upon R-loop formation, the displaced single-stranded DNA strand may facilitate the loading of the RAD51 recombinase and other recombination factors at telomeres. b) Telomeric R-loops pose a structural barrier that hinders progression of the replication machinery. Replication fork stalling may culminate in the formation of DNA double-strand breaks, which can then be repaired by homology-directed repair. c) RAD51-mediated stimulation of TERRA R-loops and association of RAD51 with the displaced DNA strand – exposed as a consequence of R-loop formation – may lead to local increased concentration of this recombinase and other recombination factors at telomeres, which in recombination-prone ALT cells may sustain telomeric homologous recombination and telomere elongation. Illustration created with BioRender.com.

3′ end by DNA polymerase δ [53,54]. This is followed by fill-in synthesis of the lagging strand, leading to telomere extension by conservative synthesis of telomeric DNA. The mechanism is referred to as break-induced replication. Recombination mediated telomere synthesis is prevalent in ALT cancer cells that lack telomerase. However, recombination mediated telomere synthesis, though partially repressed, is also at play in normal or telomerase positive mammalian cells. This notion is supported by the finding that *Brca2* deletion or Rad51 depletion in mouse embryonic fibroblasts leads to telomere damage and shortening [55] and telomere-internal double strand breaks can be repaired by homologous recombination [56]. In budding yeast, deletion of telomerase causes continuous telomere shortening. Intriguingly, TERRA R-loop formation at short telomeres promotes homology directed repair-dependent re-elongation, preventing premature onset of cellular senescence [32]. This repair of short telomeres is active in cells that have not yet engaged the ALT pathway, demonstrating the importance of homology directed repair for telomere homeostasis in normal cells.

The first evidence that TERRA R-loops can promote homology directed repair and telomere elongation came from studies in human ALT

cancer cells [57]. Coinciding with telomere recombination, TERRA expression and R-loop formation are increased in ALT cells. Telomeric R-loops cause replication stress at chromosome ends, which is a hallmark of ALT telomeres. On the other hand, inhibition of TERRA transcription was recently shown to decrease marks of DNA replication stress at telomeres, while telomere maintenance by ALT activity was impaired [58]. Interestingly, it appears that TERRA R-loops must achieve optimal levels to stimulate homology directed repair, while not completely destroying telomere integrity by overly severe replication stress [57]. RNase H1 and FANCM, which associate with ALT telomeres, play critical roles in keeping telomeric R-loops at tolerable levels [40,41,59]. FANCM depletion is selectively toxic in ALT cells. In its absence DNA replication stress at telomeres increases to unbearable levels. This stress is caused through telomeric R-loops, as RNase H1 overexpression suppressed the telomere replication stress of FANCM depleted ALT cells [40,59]. Of note, ALT cells lose the cell cycle control of TERRA expression due to mutations in chromatin remodeling protein ATRX [60]. Thus, TERRA remains high in S phase and G2 – consistent with its effects on telomere replication and elongation. Non-repressed TERRA in S phase may also, through its binding of the telomerase RNA template, reinforce telomerase repression in ALT cells.

While break-induced replication of telomeres can proceed through Rad51-dependent and independent mechanisms in yeast [61], the different synthesis pathways which operate in human ALT cells have not been fully dissected. Several studies documented roles of RAD51 in ALT cells. Depletion of RAD51 decreased the formation of ALT-associated promyelocytic leukemia bodies (APBs) in ALT cells in which telomeres gather for ALT DNA synthesis [62]. Also, chemical inhibition of RAD51 in ALT cells interfered with telomere maintenance [63] and RAD51 depletion with the frequency of telomere extension events [64]. In addition, human RAD51 facilitates long-range telomere movement and telomere clustering in ALT cells [65]. On the other hand, based on RAD51-depletion experiments, other studies described RAD51-independent mechanisms for break-induced telomere synthesis in ALT cells. A major mechanism involves RAD52, while a second ill-defined mechanism can be observed in the absence of RAD52 [54,66]. Remarkably, RAD52 can bind RNA:DNA duplexes [67,68] as well as single and double stranded DNA and thus may associate with TERRA R-loops to mediate a PCNA-RFC-Polδ-dependent pathway for conservative DNA synthesis that extends telomere length. Overall, the published work demonstrates that distinct mechanisms contribute to telomere elongation in human ALT cells. In our view, however, it may be premature to exclude crucial roles of RAD51 in a subset of the mechanisms, as this protein is required for cell viability and essential functions may have been retained in the knockdown studies.

At least three mechanisms could be envisioned of how TERRA and TERRA R-loops promote telomere recombination and elongation (Figure 5). First, the telomeric R-loops will lead to displacement of the telomeric G-rich strand which may facilitate loading of the RAD51 recombinase and other recombination proteins (Figure 5(a)). Second, R-loops can lead to replication stress and DNA double strand breaks, which can then be repaired by homology directed repair (Figure 5 (b)). Third, the increased levels of TERRA at chromosome ends through R-loops in ALT cells may not only require RAD51 – as is the case in telomerase-positive cells – but in turn also increase the local concentration of this TERRA binding recombinase to sustain DNA recombination (Figure 5(c)).

R-loops had been originally considered as toxic by-products of transcription which generate obstacles for DNA and RNA polymerases, inducing DNA damage upon replication [69,70]. Also at telomeres, R-loops do interfere with the semiconservative DNA replication machinery. The THO complex, which has been implicated in removing nascent RNA from chromatin, is present at telomeres in human cells and in budding yeast [4,71]. Deletion of THO components in *S. cerevisiae* increases telomeric R-loops, leading to replication stress at telomeres and telomere shortening [71]. As discussed above, RNase H enzymes also counteract R-loops. In their absence the accumulation of telomeric R-loops leads to telomere loss and

accelerated senescence in recombination-deficient budding yeast, supporting the notion that R-loops interfere with telomere replication [72]. Finally, depletion of NMD factors in human cells led to increased TERRA at telomeres and frequent loss of telomeric DNA [10]. Specifically, UPF1 depletion leads to inefficient replication of telomeres synthesized by leading strand synthesis, suggesting a role of this helicase in removing TERRA from leading strand telomeres for their replication [73].

Hypomorphic mutations in the *de novo* DNA methyltransferase DNMT3b cause ICF (Immunodeficiency, Centromeric instability and Facial anomalies) syndrome type I [27]. Subtelomeric DNA sequences are hypomethylated in ICF type I syndrome cells, correlating with strongly increased TERRA levels. These cells also display increased telomeric R-loops throughout the cell cycle, telomere damage, accelerated telomere shortening and premature replicative senescence, all being consistent with the interference of TERRA R-loops with telomere replication [74]. However, a direct demonstration of causative roles of DNMT3b deficiency and subtelomeric DNA demethylation in the short telomere phenotype has been cumbersome. The DNA methylation status at subtelomeres cannot be rescued in the cellular ICF models with ectopic DNMT3b, indicative of a persistent epigenetic memory [75].

Arguably, the most direct evidence for TERRA interfering with telomere replication comes from experiments in which TERRA was expressed ectopically from plasmids [34]. As discussed above, transgenic TERRA associates with telomeres post transcription through R-loops that are formed in a RAD51-dependent manner. Interference with telomere replication in human cells gives rise to so-called telomere fragility, in which the telomeric FISH signal on metaphase chromosomes is smeared or double [76]. Notably, expression of transgenic TERRA caused a significant increase of telomere fragility. This fragility was suppressed by RNase H1 overexpression or RAD51 depletion, demonstrating that it is a direct consequence of the physical presence of TERRA as R-loop at chromosome ends, even in the absence of increased RNA polymerase II and transcription at the telomere [34].

Conclusions

The finding that TERRA can associate with telomeres post transcription through the formation of R-loops was unexpected. Generally, it had been assumed that R-loops are formed during transcription mostly at the 5′ and 3′ ends of GC-skewed regions. However, in addition to TERRA, several papers reported on the formation of R-loops post transcription *in trans*. The long noncoding RNA APOLO associates with multiple loci in the genome of *Arabidopsis thaliana* post transcription through R-loops modulating gene expression through decoy of polycomb repressive complex 1 components [77]. The mechanism of R-loop formation by APOLO has remained unknown. In *S. cerevisiae*, it has been reported that Rad51p mediates hybridization of transcripts to homologous chromosomal loci distinct from their site of synthesis [78], though another paper disputed this conclusion [79]. Finally, R-loop formation is well documented for the bacterial CRISPR-Cas9 DNA endonuclease [80], in which Cas9 bends the DNA to allow guide RNA infiltration into the double helix. The mechanism of TERRA R-loop formation post transcription depends on the RAD51 recombinase which binds TERRA, and which can catalyze this reaction *in vitro*. It will be important to dissect the detailed mechanisms in the future. For instance, it is unclear how TERRA and telomere binding proteins may influence this reaction to achieve substrate specificity for TERRA and preferential invasion at short telomeres. It will also be important to determine if other factors of the recombination machinery which are required for strand invasion of recombining DNA molecules may participate in the TERRA strand invasion reaction. Finally, it will important to determine if and to what extent telomeric R-loops arise during transcription *in cis* and if RAD51 also plays a role in their establishment.

R-loops cause DNA damage and genome instability, presumably through their interference with DNA polymerases during replication. In addition, several papers reported also on beneficial roles of R-loops, for example for the repair of DNA double breaks [67,68,81]. The work on TERRA suggests that this RNA preferentially associates with short or damaged

telomeres, where it may coordinate the DNA damage response, as well as repair processes. As discussed above, excellent evidence has been provided that TERRA promotes telomere elongation by homologous recombination of short telomeres [32,57]. But does the mechanism also involve induction of replication stress and fork collapse, as a prerequisite to induce HDR? Replication stress is prevalent at telomeres in ALT, supporting the notion that these cells indeed must manage a labile balance between telomere loss by fork collapse and elongation by HDR [40,57].

Acknowledgments

We thank Eftychia Kyriacou (EPFL) and Claus Azzalin (iMM, Lisboa) for comments on the manuscript. We also acknowledge support from the Histology, BioImaging and Optics Platform and Flow Cytometry Research Core Facilities at the School of Life Sciences of EPFL.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The laboratory was supported by the Swiss National Science Foundation (SNFS grant 310030_184718), the SNFS-funded National Center of Competence in Research (NCCR) RNA and disease network (grant 182880) and the European Union's Horizon 2020 research and innovation programme under grant agreement 812829.

ORCID

Joachim Lingner http://orcid.org/0000-0002-2853-5803

References

[1] Makarov VL, Hirose Y, Langmore JP. Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. Cell. 1997;88(5):657–666.
[2] Chai W, Shay JW, Wright WE. Human telomeres maintain their overhang length at senescence. Mol Cell Biol. 2005;25(6):2158–2168.
[3] Déjardin J, Kingston RE. Purification of proteins associated with specific genomic Loci. Cell. 2009;136(1):175–186.
[4] Grolimund L, Aeby E, Hamelin R, et al. A quantitative telomeric chromatin isolation protocol identifies different telomeric states. Nat Commun. 2013;4(1):2848.
[5] Bartocci C, Diedrich JK, Ouzounov I, et al. Isolation of chromatin from dysfunctional telomeres reveals an important role for Ring1b in NHEJ-mediated chromosome fusions. Cell Rep. 2014;7(4):1320–1332.
[6] de Lange T. Shelterin-Mediated Telomere Protection. Annu Rev Genet. 2018;52(1):223–247.
[7] Baumann P, Cech TR. Pot1, the putative telomere end-binding protein in fission yeast and humans. Science. 2001;292(5519):1171–1175.
[8] Doksani Y, Wu JY, de Lange T, et al. Super-resolution fluorescence imaging of telomeres reveals TRF2-dependent T-loop formation. Cell. 2013;155:345–356.
[9] Takai KK, Kibe T, Donigian JR, et al. Telomere protection by TPP1/POT1 requires tethering to TIN2. Mol Cell. 2011;44(4):647–659.
[10] Azzalin CM, Reichenbach P, Khoriauli L, et al. Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. Science. 2007;318(5851):798–801.
[11] Schoeftner S, Blasco MA. Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. Nat Cell Biol. 2008;10(2):228–236.
[12] Luke B, Panza A, Redon S, et al. The Rat1p 5' to 3' exonuclease degrades telomeric repeat-containing RNA and promotes telomere elongation in Saccharomyces cerevisiae. Mol Cell. 2008;32(4):465–477.
[13] Bah A, Wischnewski H, Shchepachev V, et al. The telomeric transcriptome of Schizosaccharomyces pombe. Nucleic Acids Res. 2012;40(7):2995–3005.
[14] Greenwood J, Cooper JP. Non-coding telomeric and subtelomeric transcripts are differentially regulated by telomeric and heterochromatin assembly factors in fission yeast. Nucleic Acids Res. 2012;40(7):2956–2963.
[15] Vrbsky J, Akimcheva S, Watson JM, et al. siRNA-mediated methylation of Arabidopsis telomeres. PLoS Genet. 2010;6(6):e1000986.
[16] Nanavaty V, Sandhu R, Jehi SE, et al. Trypanosoma brucei RAP1 maintains telomere and subtelomere integrity by suppressing TERRA and telomeric RNA:DNA hybrids. Nucleic Acids Res. 2017;45(10):5785–5796.
[17] Tardat M, Déjardin J. Telomere chromatin establishment and its maintenance during mammalian development. Chromosoma. 2018;127(1):3–18.
[18] Tommerup H, Dousmanis A, de Lange T. Unusual chromatin in human telomeres. Mol Cell Biol. 1994;14:5777–5785.
[19] Gottschling DE, Aparicio OM, Billington BL, et al. Position effect at S. cerevisiae telomeres: reversible repression of Pol II transcription. Cell. 1990;63:751–762.
[20] Baur JA, Zou Y, Shay JW, et al. Telomere position effect in human cells. Science. 2001;292(5524):2075–2077.
[21] Porro A, Feuerhahn S, Delafontaine J, et al. Functional characterization of the TERRA transcriptome at damaged telomeres. Nat Commun. 2014;5(1):5379.

[22] Nergadze SG, Farnung BO, Wischnewski H, et al. CpG-island promoters drive transcription of human telomeres. RNA. 2009;15(12):2186–2194.
[23] Feretzaki M, Renck Nunes P, Lingner J. Expression and differential regulation of human TERRA at several chromosome ends. RNA. 2019;25(11):1470–1480.
[24] Porro A, Feuerhahn S, Reichenbach P, et al. Molecular dissection of telomeric repeat-containing RNA biogenesis unveils the presence of distinct and multiple regulatory pathways. Mol Cell Biol. 2010;30 (20):4808–4817.
[25] Saha A, Gaurav AK, Pandya UM, et al. Tb TRF suppresses the TERRA level and regulates the cell cycle-dependent TERRA foci number with a TERRA binding activity in its C-terminal Myb domain. Nucleic Acids Res. 2021;49(10):5637–5653.
[26] Iglesias N, Redon S, Pfeiffer V, et al. Subtelomeric repetitive elements determine TERRA regulation by Rap1/Rif and Rap1/Sir complexes in yeast. EMBO Rep. 2011;12(6):587–593.
[27] Yehezkel S, Segev Y, Viegas-Péquignot E, et al. Hypomethylation of subtelomeric regions in ICF syndrome is associated with abnormally short telomeres and enhanced transcription from telomeric regions. Hum Mol Genet. 2008;17(18):2776–2789.
[28] Viceconte N, Loriot A, Lona Abreu P, et al. PAR-TERRA is the main contributor to telomeric repeat-containing RNA transcripts in normal and cancer mouse cells. RNA. 2021;27(1):106–121.
[29] Chu H-P, Froberg JE, Kesner B, et al. Lee JT. PAR-TERRA directs homologous sex chromosome pairing. Nat Struct Mol Biol. 2017;24(8):620–631.
[30] Bodnar AG, Ouellette M, Frolkis M, et al. Extension of life-span by introduction of telomerase into normal human cells. Science. 1998;279(5349):349–352.
[31] Maciejowski J, de Lange T. Telomeres in cancer: tumour suppression and genome instability. Nat Rev Mol Cell Biol. 2017;18:175–186.
[32] Graf M, Bonetti D, Lockhart A, et al. Telomere length determines TERRA and R-loop regulation through the cell cycle. Cell. 2017;170(1):72–85.e14.
[33] Arnoult N, Van Beneden A, Decottignies A. Telomere length regulates TERRA levels through increased trimethylation of telomeric H3K9 and HP1α. Nat Struct Mol Biol. 2012;19(9):106–121.
[34] Feretzaki M, Pospisilova M, Valador Fernandes R, et al. RAD51-dependent recruitment of TERRA lncRNA to telomeres through R-loops. Nature. 2020;587 (7833):303–308.
[35] Bhatia V, Barroso SI, García-Rubio ML, et al. BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. Nature. 2014;511(7509):362–365.
[36] Vohhodina J, Goehring LJ, Liu B, et al. BRCA1 binds TERRA RNA and suppresses R-Loop-based telomeric DNA damage. Nat Commun. 2021;12(1):3542.
[37] Lee YW, Arora R, Wischnewski H, et al. TRF1 participates in chromosome end protection by averting TRF2-dependent telomeric R loops. Nat Struct Mol Biol. 2018;25(2):147–153.
[38] Nandakumar J, Podell ER, Cech TR. How telomeric protein POT1 avoids RNA to achieve specificity for single-stranded DNA. Proc Natl Acad Sci U S A. 2010;107(2):651–656.
[39] Glousker G, Briod A-S, Quadroni M, et al. Human shelterin protein POT 1 prevents severe telomere instability induced by homology-directed DNA repair. EMBO J. 2020;39(23):e104500.
[40] Silva B, Pentz R, Figueira AM, et al. FANCM limits ALT activity by restricting telomeric replication stress induced by deregulated BLM and R-loops. Nat Commun. 2019;10(1):2253.
[41] Lu R, O'Rourke JJ, Sobinoff AP, et al. The FANCM-BLM-TOP3A-RMI complex suppresses alternative lengthening of telomeres (ALT). Nat Commun. 2019;10(1):2252.
[42] Petti E, Buemi V, Zappone A, et al. SFPQ and NONO suppress RNA:DNA-hybrid-related telomere instability. Nat Commun. 2019;10(1):1001.
[43] Ghisays F, Garzia A, Wang H, et al. RTEL1 influences the abundance and localization of TERRA RNA. Nat Commun. 2021;12(1):3016.
[44] Vannier J-B, Pavicic-Kaltenbrunner V, Petalcorin MIR, et al. RTEL1 dismantles T loops and counteracts telomeric G4-DNA to maintain telomere integrity. Cell. 2012;149(4):795–806.
[45] Wu BR, Vogel I, Ö Ö, et al. RTEL1 suppresses G-quadruplex-associated R-loops at difficult-to-replicate loci in the human genome. Nat Struct Mol Biol. 2020;27:424–437.
[46] Sarek G, Kotsantis P, Ruis P, et al. CDK phosphorylation of TRF2 controls t-loop dynamics during the cell cycle. Nature. 2019;575(7783):523–527.
[47] Redon S, Reichenbach P, Lingner J. The non-coding RNA TERRA is a natural ligand and direct inhibitor of human telomerase. Nucleic Acids Res. 2010;38 (17):5797–5806.
[48] Cusanelli E, Romero CAP, Chartrand P. Telomeric noncoding RNA TERRA is induced by telomere shortening to nucleate telomerase molecules at short telomeres. Mol Cell. 2013;51(6):780–791.
[49] Moravec M, Wischnewski H, Bah A, et al. TERRA promotes telomerase-mediated telomere elongation in Schizosaccharomyces pombe. EMBO Rep. 2016;17 (7):999–1012.
[50] Flynn RL, Centore RC, O'Sullivan RJ, et al. TERRA and hnRNPA1 orchestrate an RPA-to-POT1 switch on telomeric single-stranded DNA. Nature. 2011;471 (7339):532–536.
[51] Redon S, Zemp I, Lingner J. A three-state model for the regulation of telomerase by TERRA and hnRNPA1. Nucleic Acids Res. 2013;41(19):9117–9128.

[52] Ding J, Hayashi MK, Zhang Y, et al. Crystal structure of the two-RRM domain of hnRNP A1 (UP1) complexed with single-stranded telomeric DNA. Genes Dev. 1999;13(9):1102–1115.
[53] Lydeard JR, Jain S, Yamaguchi M, et al. Break-induced replication and telomerase-independent telomere maintenance require Pol32. Nature. 2007;448(7155):820–823.
[54] Dilley RL, Verma P, Cho NW, et al. Break-induced telomere synthesis underlies alternative telomere maintenance. Nature. 2016;539(7627):54–58.
[55] Badie S, Escandell JM, Bouwman P, et al. BRCA2 acts as a RAD51 loader to facilitate telomere replication and capping. Nat Struct Mol Biol. 2010;17:1461–1469.
[56] Doksani Y, de Lange T. Telomere-internal double-strand breaks are repaired by homologous recombination and PARP1/Lig3-dependent end-joining. Cell Rep. 2016;17(6):1646–1656.
[57] Arora R, Lee Y, Wischnewski H, et al. RNaseH1 regulates TERRA-telomeric DNA hybrids and telomere maintenance in ALT tumour cells. Nat Commun. 2014;5(1):5220.
[58] Silva B, Arora R, Bione S, et al. TERRA transcription destabilizes telomere integrity to initiate break-induced replication in human ALT cells. Nat Commun. 2021;12(1):3760.
[59] Pan X, Chen Y, Biju B, et al. FANCM suppresses DNA replication stress at ALT telomeres by disrupting TERRA R-loops. Sci Rep. 2019;9(1):19110.
[60] Flynn RL, Cox KE, Jeitany M, et al. Alternative lengthening of telomeres renders cancer cells hypersensitive to ATR inhibitors. Science. 2015;347(6219):273–277.
[61] Chen Q, Ijpma A, Greider CW. Two survivor pathways that allow growth in the absence of telomerase are generated by distinct telomere recombination events. Mol Cell Biol. 2001;21(5):1819–1827.
[62] O'Sullivan RJ, Arnoult N, Lackner DH, et al. Rapid induction of alternative lengthening of telomeres by depletion of the histone chaperone ASF1. Nat Struct Mol Biol. 2014;21(2):167–174.
[63] Zhang T, Zhang Z, Shengzhao G, et al. Strand break-induced replication fork collapse leads to C-circles, C-overhangs and telomeric recombination. PLoS Genet. 2019;15(2):e1007925.
[64] Sobinoff AP, Allen JA, Neumann AA, et al. BLM and SLX4 play opposing roles in recombination-dependent replication at human telomeres. EMBO J. 2017;36:2907–2919.
[65] Cho NW, Dilley RL, Lampson MA, et al. Interchromosomal homology searches drive directional ALT telomere movement and synapsis. Cell. 2014;159(1):108–121.
[66] Zhang J-M, Yadav T, Ouyang J, et al. Alternative Lengthening of Telomeres through Two Distinct Break-Induced Replication Pathways. Cell Rep. 2019;26(4):955–968.e3.
[67] Yasuhara T, Kato R, Hagiwara Y, et al. Human Rad52 promotes XPG-mediated R-loop processing to initiate transcription-associated homologous recombination repair. Cell. 2018;175(2):558–570.e11.
[68] Tan J, Duan M, Yadav T, et al. An R-loop-initiated CSB-RAD52-POLD3 pathway suppresses ROS-induced telomeric DNA breaks. Nucleic Acids Res. 2020;48:1285–1300.
[69] García-Muse T, Aguilera A. R loops: from physiological to pathological roles. Cell. 2019;179(3):604–618.
[70] Marnef A, Legube G. R-loops as Janus-faced modulators of DNA repair. Nat Cell Biol. 2021;23(4):305–313.
[71] Pfeiffer V, Crittin J, Grolimund L, et al. The THO complex component Thp2 counteracts telomeric R-loops and telomere shortening. EMBO J. 2013;32(21):2861–2871.
[72] Balk B, Maicher A, Dees M, et al. Telomeric RNA-DNA hybrids affect telomere-length dynamics and senescence. Nat Struct Mol Biol. 2013;20(10):1199–1205.
[73] Chawla R, Redon S, Raftopoulou C, et al. Human UPF1 interacts with TPP1 and telomerase and sustains telomere leading-strand replication. EMBO J. 2011;30(19):4047–4058.
[74] Sagie S, Toubiana S, Hartono SR, et al. Telomeres in ICF syndrome cells are vulnerable to DNA damage due to elevated DNA:RNA hybrids. Nat Commun. 2017;8(1):14015.
[75] Toubiana S, Gagliardi M, Papa M, et al. Persistent epigenetic memory impedes rescue of the telomeric phenotype in human ICF iPSCs following DNMT3B correction. Elife. 2019;8:e47859.
[76] Sfeir A, Kosiyatrakul ST, Hockemeyer D, et al. Mammalian telomeres resemble fragile sites and require TRF1 for efficient replication. Cell. 2009;138(1):90–103.
[77] Ariel F, Lucero L, Christ A, et al. R-loop mediated trans action of the APOLO long noncoding RNA. Mol Cell. 2020;77(5):1055–1065.e4.
[78] Wahba L, Gore SK, Koshland D. The homologous recombination machinery modulates the formation of RNA–DNA hybrids and associated chromosome instability. Elife. 2013;2:e00505.
[79] Lafuente-Barquero J, García-Rubio ML, Martin-Alonso MS, et al. Harmful DNA:RNA hybrids are formed in cis and in a Rad51-independent manner. Elife. 2020;9:e56674.
[80] Jinek M, Chylinski K, Fonfara I, et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012;337(6096):816–821.
[81] Pessina F, Giavazzi F, Yin Y, et al. Functional transcription promoters at DNA double-strand breaks mediate RNA-driven phase separation of damage-response factors. Nat Cell Biol. 2019;21(10):1286–1299.

Chapter 4 Detection of TERRA R-Loops at Human Telomeres

(Chapter 11 in the book *R-Loops: Methods and Protocols*)
Galina Glousker*, Rita Valador Fernandes*, Marianna Feretzaki*, and Joachim Lingner
* These authors contributed equally
Book series Methods in Molecular Biology 2528: 159–171 (2022)
https://doi.org/10.1007/978-1-0716-2477-7_11
Reprinted with permission from Springer Nature

4.1 Abstract

"R-loops are three-stranded nucleic acid structures composed of a DNA–RNA hybrid and a displaced DNA strand. The long noncoding RNA TERRA forms R-loops at telomeres influencing the telomeric chromatin composition and impacting on telomere maintenance mechanisms by semiconservative DNA replication, homology directed DNA repair and telomerase. Here, we describe a method to detect R-loops at telomeres, which involves immunoprecipitation with the R-loop recognizing S9.6 antibody, followed by detection of telomeric DNA by either dot-blot hybridization with a radiolabeled telomeric probe, or qPCR using DNA primers that are specific for subtelomeric sequences."

4.2 Highlights

In this chapter of the book entitled "R-loops: Methods and Protocols", we compile a detailed and comprehensive protocol for the detection of DNA:RNA hybrids at telomeres by DNA:RNA immunoprecipitation (DRIP), resorting to the R-loop-recognizing S9.6 antibody. We describe two methods for the detection of immunoprecipitated R-loops: dot blot with a radiolabeled telomeric probe, and qPCR using primers amplifying subtelomeric DNA. In addition, we provide a concise scheme of the methodology for an immediate understanding of the major steps, as well as illustrative data.

4.3 Author contributions

- Contributed to the optimization of the DRIP dot blot/qPCR protocol for detection of TERRA R-loops at human telomeres.
- Prepared all figures and schemes.
- Wrote the text together with G.G. and J.L., focusing predominantly on the DRIP qPCR section.

Chapter 11

Detection of TERRA R-Loops at Human Telomeres

Galina Glousker, Rita Valador Fernandes, Marianna Feretzaki, and Joachim Lingner

Abstract

R-loops are three-stranded nucleic acid structures composed of a DNA–RNA hybrid and a displaced DNA strand. The long noncoding RNA TERRA forms R-loops at telomeres influencing the telomeric chromatin composition and impacting on telomere maintenance mechanisms by semiconservative DNA replication, homology directed DNA repair and telomerase. Here, we describe a method to detect R-loops at telomeres, which involves immunoprecipitation with the R-loop recognizing S9.6 antibody, followed by detection of telomeric DNA by either dot-blot hybridization with a radiolabeled telomeric probe, or qPCR using DNA primers that are specific for subtelomeric sequences.

Key words Telomeres, TERRA long noncoding RNA, DNA–RNA hybrids, R-loops, DRIP, Immunoprecipitation, S9.6 monoclonal antibody

1 Introduction

The detection of R-loops at telomeres is crucial to study how these structures are being formed and how they impinge on telomere biology. Of note, TERRA R-loops can form posttranscription *in trans* in dependency of the RAD51 DNA recombinase [1]. Furthermore, they are critical to promote telomere maintenance by homology directed repair which is particularly important in ALT cancer cells [2].

The detection of telomeric R-loops involves the isolation of total nucleic acids (DNA and RNA) from human cells under conditions which maintain nucleic acid double-stranded helical structures (Fig. 1a). The nucleic acid fraction is then sonicated to obtain smaller fragments with a length of approximately 100–500 base pairs (Fig. 1b). From the sonicated fraction, the DNA–RNA hybrids are immunoprecipitated using the commercially available

Galina Glousker, Rita Valador Fernandes, and Marianna Feretzaki are equally contributed.

Andrés Aguilera and Alexey Ruzov (eds.), *R-Loops: Methods and Protocols*, Methods in Molecular Biology, vol. 2528,
https://doi.org/10.1007/978-1-0716-2477-7_11,

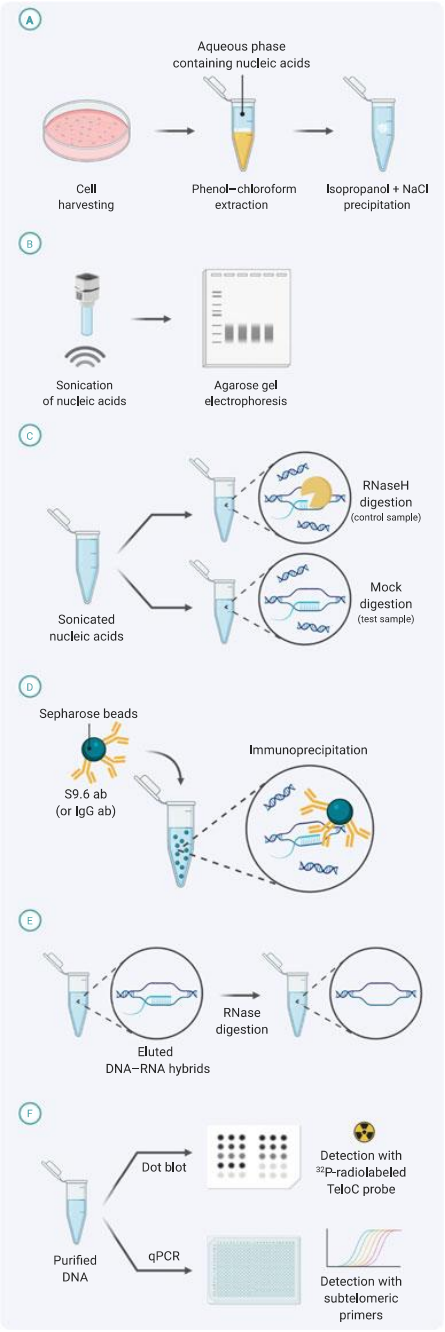


Fig. 1 Schematic depiction of a DNA–RNA immunoprecipitation, followed by detection of telomeric R-loops by dot blot or qPCR. Illustration created with BioRender.com. (**a**) Isolation of nucleic acids. (**b**) Fragmentation of nucleic acids. (**c**) RNaseH negative control. (**d**) S9.6 immunoprecipitation of DNA:RNA hybrids. (**e**) Elution and RNase digestion. (**f**) DNA analysis

S9.6 monoclonal antibody [3] which binds to R-loops (Fig. 1d). This step is often referred to as DRIP (for DNA–RNA immunoprecipitation). As a control for specificity, the nucleic acid fraction is treated in parallel prior to immunoprecipitation with the RNase H enzyme, which destroys the RNA moiety in DNA–RNA hybrid structures (Fig. 1c). Thus, this sample provides a negative control and the background signal. Finally, immunoprecipitates are eluted and treated with RNase (Fig. 1e). Telomeric DNA is then detected in the precipitated test and control samples in one of two ways (Fig. 1f). Either, the precipitated nucleic acids are spotted on a nylon membrane and telomeric DNA is detected by Southern hybridization with a radioactively labelled telomeric probe (as in Fig. 2a, b). Alternatively, telomeric DNA containing fragments are quantified by qPCR using DNA primers that recognize subtelomeric sequences residing in the vicinity of the chromosome-terminal telomeric 5′-TTAGGG-3′ repeats (as in Fig. 2c).

2 Materials

Work in RNase-free conditions. Clean all the surfaces and pipettes with RNase AWAY (Thermo Fisher Scientific) or an equivalent product.

2.1 Buffers and Solutions

1. RLN buffer: 50 mM Tris–HCl pH 8.0, 140 mM NaCl, 1.5 mM $MgCl_2$, 0.5% Nonidet P-40. Filter and store at 4 °C. Prior to use, add 100 U/mL RNasin Plus (Promega N2615) and DTT to 1 mM.
2. 10× RNase H buffer: 0.2 M HEPES–KOH pH 7.5, 0.5 M NaCl, 0.1 M $MgCl_2$, 10 mM DTT. Prepare fresh, do not filter.
3. DIP-1 buffer: 10 mM HEPES–KOH pH 7.5, 275 mM NaCl, 0.1% SDS, 1% Triton X-100, 0.1% Na-deoxycholate. Filter, store at 4 °C, short term.
4. DIP-2 buffer: 50 mM HEPES–KOH pH 7.5, 140 mM NaCl, 1 mM EDTA pH 8.0, 1% Triton X-100, 0.1% Na-deoxycholate. Filter, store at 4 °C, short term.
5. DIP-3 buffer: 50 mM HEPES–KOH pH 7.5, 500 mM NaCl, 1 mM EDTA pH 8.0, 1% Triton X-100, 0.1% Na-deoxycholate. Filter, store at 4 °C, short term.
6. DIP-4 buffer: 10 mM Tris–HCl pH 8.0, 1 mM EDTA pH 8.0, 250 mM LiCl, 1% NP-40, 1% Na-deoxycholate. Filter, store at 4 °C, short term.
7. TE buffer: 10 mM Tris–HCl pH 8.0, 1 mM EDTA pH 8.0. Filter, store at RT.
8. Elution buffer: 20 mM Tris–HCl pH 8.0, 0.1% SDS, 0.1 M $NaHCO_3$, 0.5 mM EDTA pH 8.0. Filter, store at RT. Prior to use, add RNase-DNase free (Roche) to 10 µg/mL final concentration.

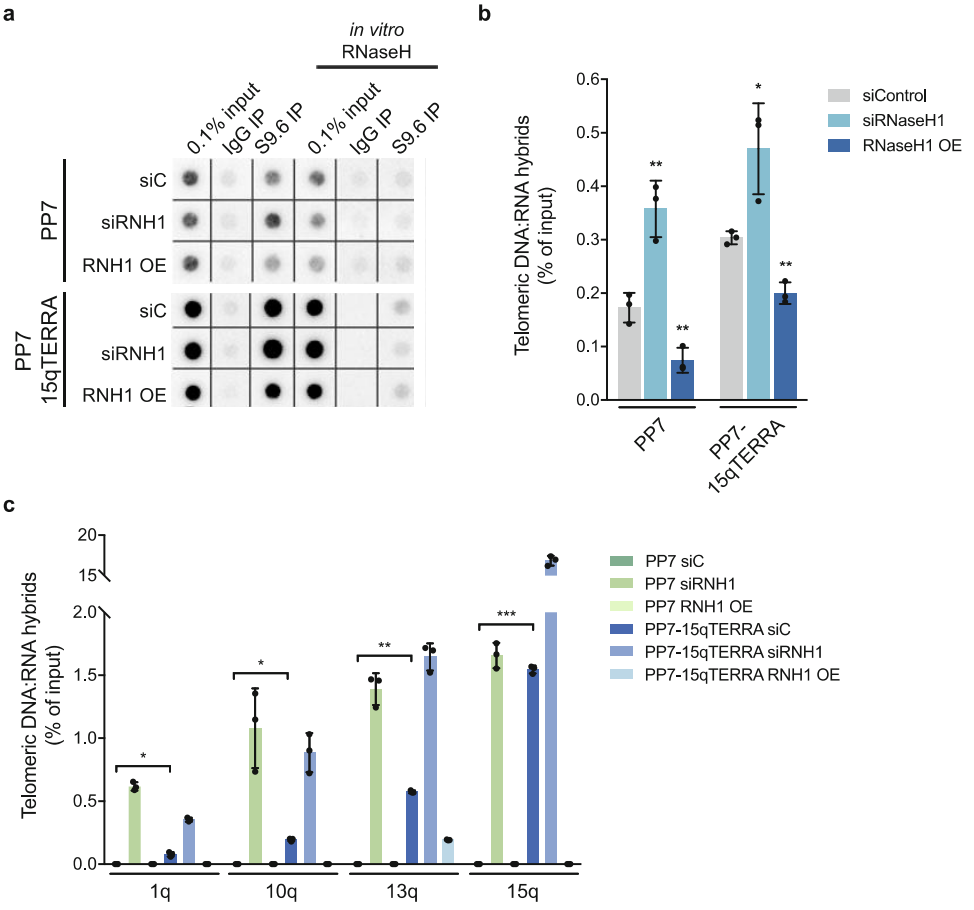


Fig. 2 Detection of telomeric R-loops by DRIP. Nucleic acids used for DRIP were isolated from HeLa cells expressing from transfected plasmids either *Pseudomonas aeruginosa* Phage 7 (PP7) RNA stem-loops or PP7-15qTERRA RNA. The cells were also transfected with nontargeting siRNA (siControl), siRNA targeting RNaseH1 (siRNaseH1) or a plasmid overexpressing RNaseH1 (RNaseH1 OE), as indicated. (**a**) DRIP samples were spotted on a nylon membrane and telomeric DNA was detected with an α-^{32}P-radiolabeled telomeric probe. The telomeric signal from PP7-15qTERRA plasmid containing cells is much stronger due to presence of telomeric repeats in the plasmid and R-loops that are formed at the site of transcription. Depletion of endogenous RNase H1 increased the DRIP signal whereas RNase H1 overexpression reduced it. In vitro RNase H treatment prior to immunoprecipitation with the S9.6 antibody reduced the signal to background levels. (**b**) Quantification of three independent biological replicates as in **a**. (**c**) DRIP samples were analyzed by qPCR and DNA was detected with primers targeting 1q, 10q, 13q, or 15q subtelomeric DNA sequences, as indicated. Three independent biological replicates are shown. Of note, expression of PP7-15qTERRA from plasmids increased the DRIP signal at telomeres 1q, 10q, and 13q which is indicative of R-loop formation at these telomeres posttranscription *in trans*. The very strong signal obtained with 15q stems mostly from the plasmid DNA which contains 15q subtelomeric sequences

9. 2× SSC buffer: 30 mM sodium citrate pH 7.0, 300 mM NaCl.
10. Denaturation buffer: 0.5 M NaOH, 1.5 M NaCl.
11. Neutralization buffer: 0.5 M Tris–HCl pH 7.0, 1.5 M NaCl.
12. Church Buffer: 0.5 M $NaHPO_4$, 1 mM EDTA, pH 8.0, 1% (weight/volume) BSA, 7% SDS.

2.2 Isolation of Nuclei and Sonication

1. 1 mL syringes.
2. 0.9 × 40 mm needles.
3. β-mercaptoethanol.
4. RA1 buffer (NucleoSpin RNA, Macherey-Nagel).
5. Phase Lock Gel Heavy 2 mL tubes (5PRIME).
6. 25:24:1 phenol–chloroform–isoamyl alcohol (pH 7.8–8.2).
7. 15 mL Falcon tubes.
8. Covaris E220 focused-ultrasonicator.
9. AFA intensifier for sonication (PN 500141).
10. Microtubes (AFA fiber Screw Cap 6 × 16 mm (Ref 520,096)).

2.3 DNA–RNA Immunoprecipitation

1. RNase H (Roche).
2. 0.5 M EDTA pH 8.0.
3. Protein G Sepharose 4 Fast Flow (Cytiva).
4. Tube rotator.
5. Table top centrifuge.
6. Anti-DNA-RNA Hybrid [S9.6] antibody (Kerafast 022715).
7. Normal mouse IgG control.
8. RNase (DNase-free).
9. NucleoSpin Gel and PCR Cleanup kit (Macherey-Nagel).
10. QIAquick PCR Purification kit (Qiagen).

2.4 Dot Blot

1. Hybond N+ membrane (GE Healthcare).
2. Bio-dot filter paper.
3. Bio-dot blot apparatus.
4. Stratalinker UV crosslinker.
5. Hybridization oven.
6. Hybridization bottles.
7. TeloC probe (PCR product generated as described [4] and labeled using RadPrime DNA Labeling System [Invitrogen]).
8. Alu probe (5′ end labeled 5′-gtgatccgcccgcctcggcctcccaaagtg-3′).
9. RadPrime DNA Labeling System (Invitrogen).
10. T4 Polynucleotide kinase (NEB).

11. α-^{32}P dCTP (10 μCi/μL; 3000 Ci/mmol).
12. γ-^{32}P ATP (10 μCi/μL; 3000 Ci/mmol).
13. Mini quick spin Oligo Column (Roche).
14. Mini quick spin DNA Column (Roche).
15. Phosphorimager screen.
16. Typhoon phosphorimager (GE).

2.5 qPCR

1. 12-channel 0.5–10 μL pipette.
2. Primers:
   (a) 1q subtelomeric primers (5′-cagcgtcgcaactcaaatg-3′; 5-′-ccctcaccctccatgagtaata-3′);
   (b) 10q subtelomeric primers (5′-gcattcctaatgcacacatgac-3′; 5′-tacccgaacctgaaccctaa-3′);
   (c) 13q subtelomeric primers (5′-gcacttgaaccctgcaatacag-3′; 5′-cctgcgcaccgagattct-3′);
   (d) 15q subtelomeric primers (5′-aaccctaaccacatgagcaacg-3′; 5′-gctgcattaaagggtccagt-3′).
3. Power SYBR Green PCR Master Mix (Applied Biosystems).
4. 384-well reaction plate.
5. Optical adhesive film.
6. QuantStudio 7 Flex Real-Time PCR System (Applied Biosystems) or equivalent.

3 Methods

Clean all surfaces and tools with RNase AWAY.

3.1 Cells Harvesting and Isolation of Nucleic Acids

1. Harvest ten million cells per condition for DRIP assay. Harvest additional one million cells for Western blot analysis if needed. Spin down the cells for 5 min at 300 × *g* at 4 °C. Keep cells on ice.
2. Wash cells with 1–2 mL cold PBS and transfer into Eppendorf microfuge tubes. Spin down again for 5 min at 300 × *g* at 4 °C.
3. Aspirate PBS. Loosen pellet thoroughly by flicking the tube. Resuspend 10 × 10^6 cells in 175 μL of cold RLN buffer containing DTT and RNaseIN, pipet up and down once, and incubate for 5 min on ice for cell lysis (*see* **Note 1**).
4. Spin down at 300 × *g*, for 2 min, 4 °C. Remove the supernatant carefully and discard.
5. Wash the nuclei containing pellet in 500 μL cold PBS. Spin down at 300 × *g*, for 2 min, 4 °C. Remove the supernatant carefully and discard.

6. Warm tubes to RT and continue to work at RT. Move to fume hood. Loosen nuclear pellet thoroughly by flicking the tube.
7. Pellet gel in Phase Lock Gel (PLG) Heavy 2 mL tubes by centrifugation at 13,000 × *g* for 1 min.
8. Prepare nuclear lysis buffer: add 10 μL of β-mercaptoethanol per 1 mL of buffer RA1 (prepare 1 mL of nuclear lysis buffer per condition).
9. Resuspend nuclei in 500 μL of nuclear lysis buffer at RT.
10. Pass the lysate 10 times up and down through a needle.
11. Transfer the lysate into the PLG tubes. Add 250 μL H_2O and 750 μL 25:24:1 phenol/chloroform/isoamyl alcohol. Mix thoroughly, but do not vortex.
12. Centrifuge at 13,000 × *g* for 5 min to separate phases.
13. Transfer the nucleic-acid-containing aqueous upper phase (NA) into a fresh 2 mL Eppendorf microfuge tube on ice.
14. Perform isopropanol precipitation: 750 μL NA + 15 μL 5 M NaCl +750 μL ice-cold isopropanol. Shake vigorously! Incubate on ice for 30 min, then spin down for 30 min at 10,000 × *g* at 4 °C (in 2 mL Eppendorf tubes) (*see* **Note 2**).
15. Wash pellet 2× with 1 mL of ice cold 70% ethanol.
16. Let the pellet air dry (*see* **Note 3**).
17. Add 135 μL H_2O to each pellet from 10 × 10^6 cells, flick to detach the pellet from the wall of the tube, incubate for approximately 1 h shaking in the cold room until nucleic acids are dissolved.
18. Take sample "A" before sonication: 3 μL.

3.2 Sonication with Covaris E220

1. Transfer each condition into a microtube for sonication (130 μL per tube). Perform sonication using the following conditions to achieve fragment length < 500 bp.

Condition	a single condition
Duty factor	10%
Peak incident power (W)	140
Cycles per burst	200
Time (seconds)	150
AFA intensifier	Yes
Temperature	4 °C
Water level	6

2. Take sample "B" after sonication: 3 µL.
3. Test sonication profile:
   (a) Add to 3 µL of sample (A and B) 7 µL of H_2O.
   (b) add 2 µL of 6× loading buffer.
   (c) Analyze on a 1% TBE-agarose gel.

You can freeze the samples in liquid nitrogen and keep them at −80 °C for a few days.

3.3 DNA–RNA Immunoprecipitation

1. Measure the concentration of nucleic acids with Nanodrop (DNA). Use the same amount of nucleic acids for all your samples (15–60 µg depending on the concentration—*see* **Note 4**).
2. Prepare fresh RNase H buffer.
3. Perform the RNase H control +/−RNase H treatment - example:
   (a) nucleic acids,
   (b) 15 µL 10× RNase H buffer
   (c) 10 µL RNase H (1 U/µL; Roche) (control) or H_2O (sample)
   (d) H_2O to 150 µL.
4. Incubate at 37 °C for 90 min.
5. Add 2 µL 0.5 M EDTA (pH 8.0) to stop the reaction.
6. During the RNase H treatment equilibrate beads (*see* **Note 5**). Use 200 µL beads per condition (equals to 400 µL of 50% slurry stock (*see* **Note 6**)).
   (a) Put the 50% Sepharose G bead stock on the wheel in the cold room for 5 min to mix it very well. Take what you need into a new tube and spin them down for 2 min at 300 × *g* at 4 °C. Remove the supernatant.
   (b) Add the same volume of DIP-1 buffer**,** wash beads on the wheel for 5 min at 4 °C. Spin down for 2 min at 300 × *g* at 4 °C, discard the supernatant. Repeat this step once.
   (c) Resuspend the beads in DIP-1 buffer to have a 50% slurry.
7. Preclear the nucleic acids. Mix in a 2 mL microfuge tube the following.
   (a) 150 µL nucleic acids after RNase H treatment.
   (b) 1270 µL DIP-1 buffer (ten-fold dilution of nucleic acids).
   (c) 80 µL of protein G beads 50% slurry.
8. Incubate for 1 h at 4 °C on the wheel. Spin down 5 min at 300 × *g* and continue with the precleared extract.

9. Put aside 1% input.
10. Assemble the IP reaction.
    (a) 600 µL of precleared extract.
    (b) 3 µg of antibody (spin down before use):
        - Tube 1: S9.6 antibody (Kerafast 022715).
        - Tube 2: mouse IgG (sc2025).
    (c) 80 µL of protein G beads 50% slurry.
11. Incubate overnight on the wheel at 4 °C.
12. Spin down for 2 min at 300 × *g* at 4 °C, discard the supernatant.
13. Wash the beads 5 min on a wheel at 4 °C with ice-cold buffers (spin for 2 min at 300 × *g* at 4 °C).
    (a) 1 mL DIP-2 buffer.
    (b) 1 mL DIP-3 buffer.
    (c) 1 mL DIP-4 buffer.
    (d) 1 mL TE.
14. Spin for 5 min at 300 × *g* at 4 °C.
15. Resuspend beads and inputs in 100 µL Elution buffer containing RNase (DNase-free).
16. Incubate for 1–2 h at 65 °C (or overnight).
17. Extract DNA using:
    (a) *If you continue with dot-blot*: Macherey-Nagel NucleoSpin Gel and PCR Cleanup kit and elute in 100 µL H_2O (Heat up H_2O to 70 °C = > Incubate for 5 min before elution).
    (b) *If you continue with qPCR*: Qiagen PCR Purification kit for qPCR and elute in 100 µl H_2O (Heat up H_2O to 70 °C = > Incubate for 5 min before elution).
18. Store DNA at −20 °C or directly continue with dot-blot or qPCR.

3.4 Dot Blot

1. Clean Dot-Blot apparatus (rinse with 0.02% SDS; spray with 70% ethanol, rinse with H_2O multiple times; dry completely).
2. Prewarm Church buffer and hybridization oven.
3. Denature DNA for 10 min at 95 °C and keep on ice for 5 min or longer.
4. Assemble the dot blot apparatus with N+ Hybond membrane soaked in 2× SSC, Whatman paper soaked in 2×SSC, then wash the slots with 200 µL 2× SSC.
5. Apply the samples. For the input (which is 1%), add 900 µL of H_2O and load 200 µL (0.2%), 100 µL (0.1%), and 50 µL (0.05%), for the IP samples load the entire sample (*see* **Note 7**).

6. Wash again the slots that have been used with 2× SSC.
7. UV-crosslink the DNA onto the membrane (do not let the membrane get completely dry), using autocrosslink as a setting.
8. Denature the membrane in 0.5 M NaOH, 1.5 M NaCl for 15 min on a shaker at room temperature.
9. Neutralize the membrane in 0.5 M Tris–HCl pH 7.0, 1.5 M NaCl for 10 min on a shaker at room temperature.
10. Prehybridize the membrane with Church Buffer for at least 1 h at 65 °C when using the telomeric probe and at 55 °C for the alu probe.
11. Label telomeric probe or alu probe.
    (a) Telomeric probe.
        - µL DNA (200 ng; telomeric PCR product (Porro et al., 2014)).
        - 20 µL H_2O.
        - Denature for 5 min at 95 °C => put on ice immediately, add the following.
        - 1 µL dATP (500 µM), 1 µL dGTP (500 µM), 1 µL dTTP (500 µM).
        - 20 µL 2,5× Random Prim buff.
        - 1 µL Klenow enzyme.
        - 5 µL α-^{32}P dCTP (3000 Ci/mmol; 10 µCi/µL).
        - H_2O to 50 µL.
        - Incubate at 37 °C for at least 10 min.
        - Purify the labeled DNA using the Mini quick spin DNA Column.
        - Denature the radiolabeled probe for 5 min at 95 °C.
        - Place on ice.
        - Add to 20–30 mL of Church buffer.
    (b) Alu probe.
        - 1 µL oligonucleotide (from a 10 µM stock).
        - 27 µL H_2O.
        - 4 µL T4 PNK buffer.
        - 2 µL T4 PNK.
        - 5 µL γ-^{32}P ATP (3000 Ci/mmol; 10 µCi/µL).
        - H_2O to 40 µL.
        - Incubate at least 1 h at 37 °C.
        - Purify using Mini quick spin Oligo Column.
        - Add to 20–30 mL of Church buffer.

12. Hybridize with radiolabeled telomeric probe or alu probe overnight.

    On the following day:

13. Rinse with 4× SSC.
14. Wash three times for 30 min at 65 °C in 1× SSC, 0.5% SDS at 65 °C for the telomeric probe or 55 °C for the alu probe.
15. Rinse in 4× SSC in a tray to remove SDS before sealing the membrane.
16. Place the membrane into a plastic folder, seal properly and expose to a phosphorimager screen (expose overnight, prolong exposure time if needed).

3.5 qPCR

1. Set up a qPCR master mix for each primer pair (all components should be kept on ice).
   (a) 5 μL Power SYBR Green (prolonged direct light exposure should be avoided).
   (b) 1 μL of each primer (forward and reverse) from a 10 μM stock.
   (c) H_2O to a total volume of 9 μL per well.
2. Prepare serial dilutions of the collected 1% input in H_2O (at least two dilutions; for example, dilution factors of 10 and 100—*see* **Note 7**).
3. Transfer all samples (diluted inputs and immunoprecipitates) and master mixes into 0.2 mL PCR tube strips, to facilitate pipetting samples with multichannel pipette.
4. Resorting to a multichannel pipette (in order to minimize pipetting-derived errors), transfer 9 μL of each primer pair master mix into each well of a 384-well reaction plate.
5. Resorting to a multichannel pipette, add 1 μL of each input dilution and each DRIP sample into each well of the reaction plate. At least two technical replicates per input dilution/sample should be carried out. For each primer pair a control reaction with no template should also be included.
6. Seal the reaction plate with optical adhesive film.
7. Briefly centrifuge the plate at 2000 × *g*.
8. Load on the qPCR instrument.
9. Perform the qPCR reaction according to manufacturer's instructions: 95 °C for 10 min, followed by 95 °C for 15 s, and annealing and extension at 60 °C for 1 min for 40 cycles, and finalized with dissociation stage for melting curve analysis.

3.5.1 qPCR Data Analysis

1. Using the Ct values obtained for the dilutions of input samples, perform a regression analysis to determine the equation of the standard curve of each input.
2. Using each input equation, calculate the corresponding S9.6 and IgG IPs as percentage of input.

4 Notes

1. Do not scale up. The efficient lysis is only possible with no more than 10×10^6 cells per 175 µL of RLN buffer. If your signal is too low and you want to increase the starting number of cells to 20×10^6, split your cells into two tubes with 10×10^6 cells in each.
2. The work in the fume hood lasts for at least 1 h.
3. You can load your Western blot during that time. Turn on Covaris in a timely manner to ensure that the water is cold and degassed when you start sonication.
4. Depending on the cell line, basal levels of telomeric RNA–DNA hybrids will vary. Adjust cell number and amount of nucleic acids accordingly (e.g., for HeLa cells 30 µg/IP are required, while for U2OS cells 7.5 µg/IP should suffice).
5. Do not use Falcon tubes to spin down the beads as the beads tend to be lost in these tubes. Use 5 mL or 2 mL tubes.
6. Start with 20% more beads than you need.
7. The input dilutions should be adjusted so that the amount of target DNA in each immunoprecipitated sample lies within the linear dynamic range of each assay.

Acknowledgments

The laboratory was supported by the Swiss National Science Foundation (SNFS grant 310030_184718), the SNFS-funded National Center of Competence in Research (NCCR) RNA and disease network (grant 182880), and the European Union's Horizon 2020 research and innovation programme under grant agreement 812829.

References

1. Feretzaki M, Pospisilova M, Valador Fernandes R et al (2020) RAD51-dependent recruitment of TERRA lncRNA to telomeres through R-loops. Nature 587:303–308
2. Silva B, Arora R, Bione S et al (2021) TERRA transcription destabilizes telomere integrity to initiate break-induced replication in human ALT cells. Nat Commun 12:3760
3. Boguslawski SJ, Smith DE, Michalak MA et al (1986) Characterization of monoclonal antibody to DNA · RNA and its application to immunodetection of hybrids. J Immunol Methods 89:123–130
4. Porro A, Feuerhahn S, Lingner J (2014) TERRA-reinforced association of LSD1 with MRE11 promotes processing of uncapped telomeres. Cell Rep 6:765–776

Chapter 5 The THO Complex counteracts TERRA R-loop-mediated telomere fragility in telomerase$^+$ cells and recombination in ALT$^+$ cells

Rita Valador Fernandes and Joachim Lingner
Manuscript under revision

5.1 Abstract

Telomeres are the nucleoprotein structures at the ends of linear chromosomes. Telomeres are transcribed into the long non-coding Telomeric Repeat-Containing RNA (TERRA), whose functions rely on its ability to associate with telomeric chromatin. The conserved THO complex (THOC) was previously identified at human telomeres. It links transcription with RNA processing, decreasing the accumulation of co-transcriptional DNA:RNA hybrids throughout the genome. Here, we explore the role of THOC at human telomeres, as a regulator of TERRA localization to chromosome ends. We show that THOC counteracts TERRA association with telomeres via R-loops formed co-transcriptionally and also post-transcriptionally, *in trans*. We demonstrate that THOC binds nucleoplasmic TERRA, and that RNaseH1 loss, which increases telomeric R-loops, promotes THOC occupancy at telomeres. Additionally, we show that THOC counteracts lagging and mainly leading strand telomere fragility, suggesting that TERRA R-loops can interfere with replication fork progression. Finally, we observed that THOC suppresses telomeric sister chromatid exchange and C-circle accumulation in ALT cancer cells, which maintain telomeres by recombination. Altogether, our findings reveal crucial roles of THOC in telomeric homeostasis through the co- and post-transcriptional regulation of TERRA R-loops.

5.2 Highlights

- The THO complex counteracts telomeric R-loops, in ALT cells.
- THOC subunits 1 and 2 counteract PP7-15qTERRA colocalization with telomeres, in the form of R-loops formed post-transcriptionally, *in trans*.
- THOC1 occupancy at the telomeric tract is increased following RNaseH1 depletion.
- THOC associates with endogenous nucleoplasmic TERRA.
- THOC prevents telomeric fragility induced by TERRA R-loops.
- THOC restrains telomeric fragility at telomeres replicated by leading and (to a lesser extent) lagging strand synthesis.
- THOC counteracts telomeric sister chromatid exchange events and C-circle accumulation in ALT cells.

5.3 Author contributions

- Performed all experiments.
- Prepared all figures, schemes and respective legends.
- Wrote the full manuscript, which was edited by J.L..

5.4 Introduction

Telomeres are the nucleoprotein structures at the termini of linear chromosomes. They are crucial for genome stability and ensure that chromosome ends are not inappropriately recognized as sites of

DNA damage. Telomere functions at mammalian telomeres are in part achieved by a telomere-specific protein complex termed shelterin. It includes the double strand (ds) DNA-binding proteins TRF1 and TRF2, POT1 which binds the single stranded telomeric overhang, TIN2 and TPP1 which connect POT1 to TRF1 and TRF2, and TRF2-bound Rap1 (de Lange, 2005).

Despite the heterochromatic features found at telomeres (Tardat & Déjardin, 2018), RNA polymerase II drives transcription of the telomeric long noncoding RNA TERRA – Telomeric Repeat containing RNA, using the C-rich strand as a template (Azzalin *et al*, 2007). In human cells, transcription of TERRA stems from promoters residing within subtelomeric sequences in most chromosome arms, and proceeds through the telomeric repetitive tract (Nergadze *et al*, 2009; Porro *et al*, 2014; Feretzaki *et al*, 2019). For that reason, TERRA is a heterogeneous class of lncRNAs, comprising chromosome arm-specific subtelomeric-derived sequences at the 5' end, followed by a variable number of UUAGGG repeats (Porro *et al*, 2010; Arnoult *et al*, 2012; Feretzaki *et al*, 2019).

Several functions regarding telomere maintenance and stability have been attributed to TERRA. Namely, TERRA has been proposed to contribute to the regulation of the heterochromatic structure of telomeres (Arnoult *et al*, 2012; Porro *et al*, 2014), particularly when telomeres are depleted of the shelterin protein TRF2 – which results in upregulation of TERRA and activation of a DNA damage response (Porro *et al,* 2014). In addition, the modulation of telomerase – the ribonucleoprotein which extends the telomeric DNA – by TERRA has previously been proposed. *In vitro*, TERRA was shown to bind the telomerase RNA template, thus acting as an inhibitor of human telomerase (Redon *et al,* 2010). On the other hand, *in vivo* experiments in budding yeast suggested that TERRA preferentially guides telomerase to short telomeres, which are then selectively elongated (Cusanelli *et al*, 2013). While the majority of cancer cells maintains telomere length by overexpressing telomerase, about 10% of cancer types compensates telomere loss by a mechanism termed alternative lengthening of telomeres (ALT), which relies on homologous recombination (HR) (Bryan *et al*, 1997). In ALT human cells, TERRA was shown to be a crucial factor involved in telomere maintenance, by promoting a finely-tuned equilibrium of replication stress required for triggering of HR by break-induced replication, allowing telomere elongation (Arora *et al*, 2014; Silva *et al*, 2019, 2021, 2022). Also in *Saccharomyces cerevisiae* cells, TERRA was shown to accumulate at critically short telomeres, where it is thought to promote replication stress, resulting in homology-directed repair and telomere elongation (Graf *et al*, 2017).

Notably, most functions attributed to TERRA in telomere maintenance were shown to depend on direct DNA:RNA interactions (hybrids), termed R-loops (Balk *et al*, 2013; Arora *et al*, 2014; Graf *et al*, 2017; Sagie *et al*, 2017; Silva *et al*, 2019). Such structures form when TERRA RNA invades the telomeric dsDNA helix, base pairing with the template strand, leaving the TTAGGG-containing DNA strand displaced. R-loops were shown to form co-transcriptionally *in cis* – as the transcription machinery progresses through the telomeric tract –, as well as post-transcriptionally *in trans* – at a telomeric locus other than the one from which the RNA originated (Feretzaki *et al*, 2020). Several factors involved in the regulation of telomeric R-loops have been identified and characterized (reviewed in (Valador Fernandes *et al*, 2021)). Among which, the DNA recombinase RAD51 was shown to be required for the formation of telomeric DNA:RNA hybrids (Feretzaki *et al*, 2020). More recently, another HR protein, RAD51AP1, has also been implicated in mediating TERRA R-loops in ALT cancer cells (Yadav *et al*, 2022; Kaminski *et al*, 2022). On the other hand, factors restraining the accumulation of telomeric R-loops include the shelterin proteins TRF1 and POT1 (Lee *et al*, 2018; Glousker *et al*, 2020), as well as RNase H enzymes (Arora *et al*, 2014) – which remove DNA:RNA hybrids through the endonucleolytic cleavage of the engaged RNA molecules.

R-loops have primarily been viewed as harmful by-products of transcription, posing an obstacle for the progression of DNA and RNA polymerases, thus jeopardizing genome stability. However, their involvement in a myriad of cellular processes has been demonstrated (García-Muse & Aguilera, 2019). At chromosome ends, R-loops are thought to contribute to DNA damage checkpoint signalling when telomeres become damaged or very short (Balk *et al*, 2013; Graf *et al*, 2017; Feretzaki *et al*, 2020). Additionally, as mentioned above, R-loops appear to be a requirement for appropriate telomerase-independent telomere elongation in ALT cancer cells. Altogether, this underscores the relevance behind

a comprehensive analysis of the regulation and impact of telomeric R-loops. In this study, we aimed at understanding the roles undertaken by the THO (Suppressor of Transcription Defects of *hpr1Δ* mutants by Overexpression) complex at human telomeres.

The THO complex was originally identified in *S. cerevisiae* as a functional unit involved in transcriptional elongation of genes with high GC content or that are particularly long, suppressing hyperrecombination events (Aguilera & Klein, 1990; Chávez *et al*, 2000, 2001). In human cells, the THO complex comprises six subunits: THOC1, -2, -3, -5, -6 and -7 (Pühringer *et al*, 2020). It facilitates transcription of a large number of human genes, and often co-localizes with splicing factors in nuclear speckles (Li *et al*, 2005; Domínguez-Sánchez *et al*, 2011; Masuda *et al*, 2005). In addition to its function during transcription, THOC plays a conserved role in RNA nuclear export, as an element of the larger TREX complex (Transcription and Export), together with the helicase and splicing factor UAP56/DDX39B and the RNA nuclear export protein Aly/ALYREF/THOC4 (along with other factors) (Fleckner *et al*, 1997; Zhang & Green, 2001; Zhou *et al*, 2000). Remarkably, THOC/TREX have also been described to take part in the regulation of DNA:RNA hybrids formed across the genome during transcription, counteracting the accumulation of R-loop-induced DNA damage (Li *et al*, 2005; Domínguez-Sánchez *et al*, 2011; Pérez-Calero *et al*, 2020). THOC/TREX factors are thought to be recruited to spliced, 5' capped RNAs during RNA biogenesis, and contribute to the adequate formation of ribonucleoprotein complexes (Masuda *et al*, 2005; Cheng *et al*, 2006; Chi *et al*, 2013; Köhler & Hurt, 2007; Carmody & Wente, 2009). Thus, THOC/TREX bridge transcription with RNA processing and nuclear export.

All THOC subunits were previously detected in purified human telomeric chromatin by mass-spectrometry (Grolimund *et al*, 2013). In yeast, several THOC subunits have also been detected at chromosome ends (Pfeiffer *et al*, 2013). THOC was previously proposed to indirectly counteract telomeric elongation, through the regulation of Rif1 mRNA (a negative regulator of telomere length in *S. cerevisiae*) (Yu *et al*, 2012). On the other hand, yeast THOC subunits Hpr1 and Thp2 were demonstrated to counteract TERRA-mediated telomeric R-loops. Genetic experiments further suggested that Thp2 protects chromosome ends from telomere shortening in an R-loop-independent but exonuclease 1-dependent manner (Pfeiffer *et al*, 2013). At human telomeres, the roles of the THO complex remained elusive.

Here, we demonstrate that THOC subunits 1 and 2 counteract R-loops formed at human telomeres. Resorting to a previously developed system to detect the nuclear localization of ectopically-expressed TERRA molecules (Feretzaki *et al*, 2020), we show that, unexpectedly, the THO complex also counteracts the association of TERRA with telomeres post-transcriptionally, *in trans*, via R-loops. In addition, we find that the THO complex is recruited to telomeres when cells are depleted of the ribonuclease RNaseH1, which specifically degrades DNA:RNA hybrids. In addition, THOC interacts with nucleoplasmic TERRA. Therefore, we propose that THOC counteracts telomeric R-loops through several mechanisms. First, THOC may prevent the association of TERRA with telomeric DNA through R-loops during transcription. Second, THOC contributes to the resolution of DNA:RNA hybrid structures that have formed post-transcriptionally. Third, through binding of TERRA in the nucleoplasm, THOC may counteract TERRA association with telomeric DNA post-transcriptionally. Finally, we observe that loss of THOC subunits results in fragility at telomeres replicated by leading strand synthesis, and also to a lesser extent by lagging strand synthesis. In addition, in ALT cells – but not in telomerase-positive HeLa cells – loss of THOC leads to an increase in the frequency of telomeric sister chromatid exchange events, as well as C-circles, likely as a consequence of accumulated R-loops at the telomeric tract in these cells. Thus, our data demonstrate that THOC promotes telomere stability through the regulation of the lncRNA TERRA.

5.5 Results

The THO complex counteracts telomeric DNA:RNA hybrids

THOC was previously shown to be necessary for appropriate expression of several genes in yeast and human cells, promoting transcriptional elongation by RNA polymerase II (Chávez *et al*, 2000; Li *et*

al, 2005; Domínguez-Sánchez *et al*, 2011; Chávez & Aguilera, 1997; Piruat & Aguilera, 1998). To understand if THOC is involved in the transcriptional regulation of TERRA affecting its expression levels, we knocked down THOC core subunits 1 and 2 in HeLa cells (Fig 1A). Of note, siRNA mediated targeting of THOC1 caused a reduction of THOC2 and vice versa, indicating that THOC subunits THOC1 and THOC2 are stabilizing one another. We evaluated total TERRA levels by RNA dot blot, probed with a ^{32}P-radiolabeled $[CCCTAA]_3$ probe and normalized the signal to 18S rRNA (Fig 1B). No significant changes were detected in UUAGGG-containing RNA levels upon depletion of THOC (Fig 1B-C). This indicates that the expression of TERRA is not directly modulated by THOC, nor is the telomeric transcriptional elongation affected by loss of THOC subunits, since a defect in TERRA transcriptional elongation would be reflected in changes in the detected UUAGGG content.

In U2OS cells, which do not use telomerase for telomere length maintenance, but resort instead to the recombination-based ALT pathway, basal TERRA levels are elevated, relative to HeLa or other telomerase-positive cell lines (Lovejoy *et al*, 2012). Nonetheless, depletion of THOC subunits in U2OS cells (Fig 1D) did not lead to any perceptible changes in TERRA levels (Fig 1E and F), similarly to what we observed in HeLa cells.

Telomeres are prone to form DNA:RNA hybrids with TERRA (Graf *et al*, 2017; Arora *et al*, 2014; Feretzaki *et al*, 2020). Since THOC has been implicated in counteracting DNA:RNA hybrids across the genome (Li *et al*, 2005; Domínguez-Sánchez *et al*, 2011), we analysed telomeric DNA:RNA hybrid levels under depletion of THOC. We implemented DNA:RNA immunoprecipitation (DRIP), which uses the R-loop-recognizing S9.6 antibody (Boguslawski *et al*, 1986), and quantified the DNA component of the hybrids by DNA dot blot, probed with a C-rich telomeric probe. Nucleic acids were *in vitro* treated with the ribonuclease RNaseH (which specifically hydrolyses the RNA moiety of DNA:RNA hybrids (Cerritelli & Crouch, 2009)) before the immunoprecipitation step, as a specificity control. Interestingly, depletion of THOC led to a perceptible (yet non statistically significant) increase in telomeric R-loop levels in HeLa cells (Fig 1G and H). In U2OS cells, basal telomeric R-loops were shown to accumulate more frequently, compared to telomerase-positive cells (Arora *et al*, 2014). When examining U2OS cells by DRIP dot blot, we observed that loss of THOC in these cells significantly increased the occurrence of telomeric R-loops by about 2-fold (Fig 1I and J). This indicates that the THO complex plays a role at human telomeres in counteracting TERRA-mediated R-loops.

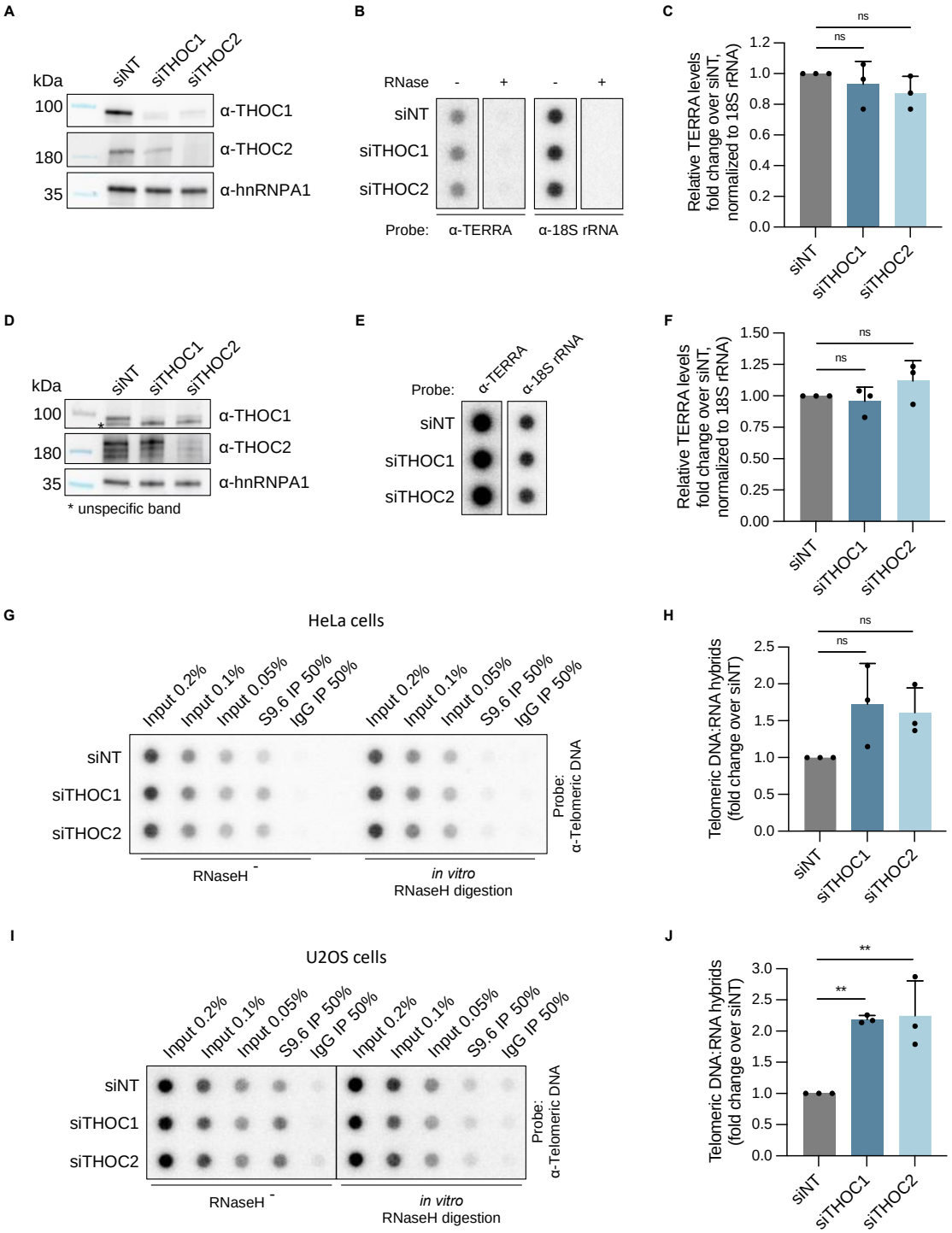


Figure 1. THOC counteracts TERRA R-loops.

A Western blot analysis of depletion efficiency with indicated siRNAs in HeLa cells with 10 kb average telomere length, used in B, C, G and H.

B RNA dot blot analysis of TERRA or 18S rRNA levels upon depletion with indicated siRNAs in HeLa cells. 4 µg total RNA were loaded per sample. RNA samples were treated with RNase (DNase-free) as a control.

C Quantification of TERRA levels (as in B), normalized to 18S rRNA levels in HeLa cells, plotted as fold change over control siNT. Data are mean ± s.d., from three independent biological replicates. One-way analysis of variance (ANOVA) with Dunnett's multiple comparisons test was applied: ns indicates non-significance ($P > 0.05$).

D Western blot analysis of depletion efficiency with indicated siRNAs in U2OS cells, used in E, F, I and J.

E RNA dot blot analysis of TERRA or 18S rRNA levels upon depletion with indicated siRNAs in U2OS cells. 2 µg total RNA were loaded per sample.

F Quantification of TERRA levels (as in E), normalized to 18S rRNA levels in U2OS cells, plotted as fold change over control siNT. Data are mean ± s.d., from three independent biological replicates. One-way analysis of variance (ANOVA) with Dunnett's multiple comparisons test was applied: ns indicates non-significance.

G DRIP assay using anti-DNA:RNA hybrid S9.6 antibody was performed in extracts from HeLa cells. *In vitro* digestion with RNaseH1 prior immunoprecipitation served as a negative control. All immunoprecipitates and input samples were treated with RNase (DNase-free) and analyzed by DNA dot blot probed with a ^{32}P-radiolabelled $[CCCTAA]_3$ probe.

H Quantification of telomeric DNA:RNA hybrids in HeLa cell extracts (as in G), as fold change over siNT. Data represent mean ± s.d., from three independent biological replicates. One-way analysis of variance (ANOVA) with Dunnett's multiple comparisons test was applied: ns indicates non-significance ($P > 0.05$).

I DRIP assay using anti-DNA:RNA hybrid S9.6 antibody was performed in extracts from U2OS cells. *In vitro* digestion with RNaseH1 prior immunoprecipitation served as a negative control. All immunoprecipitates and input samples were treated with RNase (DNase-free) and analysed by DNA dot blot probed with a ^{32}P-radiolabelled $[CCCTAA]_3$ probe.

J Quantification of telomeric DNA:RNA hybrids in U2OS cell extracts (as in I), as fold change over siNT. Data represent mean ± s.d., from three independent biological replicates. One-way analysis of variance (ANOVA) with Dunnett's multiple comparisons test was applied: ** $P \leq 0.01$.

The THO complex counteracts PP7-15qTERRA associations with telomeres formed *in trans*

While most R-loops are thought to occur only during transcription, TERRA R-loops can also form post-transcriptionally, in dependency of the RAD51 recombinase (Feretzaki *et al*, 2020). To investigate putative roles of THOC in the post-transcriptional regulation of TERRA, we used a previously developed system in which TERRA is expressed from plasmids, allowing the characterization of the subnuclear localization of ectopically transcribed TERRA-like molecules (Feretzaki *et al*, 2020). It resorts to the inducible-expression of a chimeric RNA molecule comprised of about ninety UUAGGG repeats at the 3' end, preceded by the 15q subtelomeric-derived sequence, and twenty-four *Pseudomonas aeruginosa* Phage 7 (PP7) stem-loops. The PP7 stem-loops are recognized and bound by GFP-tagged PP7-Coat Protein (PCP) proteins (Larson *et al*, 2011), thus enabling the analysis of colocalization of chimeric TERRA molecules with telomeres detected by DNA fluorescence *in situ* hybridization (FISH) (Feretzaki *et al*, 2020).

HeLa cells with an average length of 10 kb were transfected with siRNAs for depletion of THOC subunits, the TREX component DDX39B or RNaseH1. Cells were then transfected with plasmids allowing the expression of PP7-15qTERRA upon 24h doxycycline induction (Fig 2A). Notably, PP7-15qTERRA RNA molecules expressed from plasmids were found to colocalize more often with telomeres when THOC subunits 1 or 2, or RNaseH1 were depleted (Fig 2B-D). This indicates, unexpectedly, that THOC has the ability to impact PP7-15qTERRA associations with telomeres which are occurring post-transcriptionally, *in trans*. Also, in cells with shorter average telomere length (ca. 2.5 kb), where PP7-15qTERRA was previously shown to be preferentially recruited to telomeres (Feretzaki *et al*, 2020), a further increase in colocalization with telomeres was observed when THOC subunits and the TREX component DDX39B were downregulated (Supplementary fig 1A and B).

To ensure that increased colocalization prompted by loss of THOC did not stem from an increase in the nuclear PP7-15qTERRA levels, we collected total RNA from cells (10 kb average telomere length) transfected with the PP7-15qTERRA-coding plasmid and where THOC1 and -2, or RNaseH1 were downregulated (Supplementary fig 2A), and quantified TERRA by RNA dot blot (Supplementary fig 2B and C). While depletion of THOC1 or RNaseH1 did not disturb endogenous plus ectopically expressed TERRA levels, THOC2 depletion led to a slight reduction in UUAGGG-containing RNA levels (Supplementary fig 2B and C). This result was accompanied by a subtle reduction in the total number of PP7-15qTERRA foci detected per nucleus (Supplementary fig 2D). While this may reveal a different requirement of THOC by endogenous and ectopically expressed TERRA species, we consider that this should not significantly affect the analysis of TERRA association with telomeres, given that it is quantified as percentage of colocalizing TERRA foci with telomeric FISH signal, over total TERRA foci

detected per nucleus. In addition, a similar effect in PP7-15qTERRA telomeric colocalizations is observed upon depletion of THOC1 – which did not result in any perceivable changes in ectopically-expressed TERRA levels (Supplementary fig 2B and C).

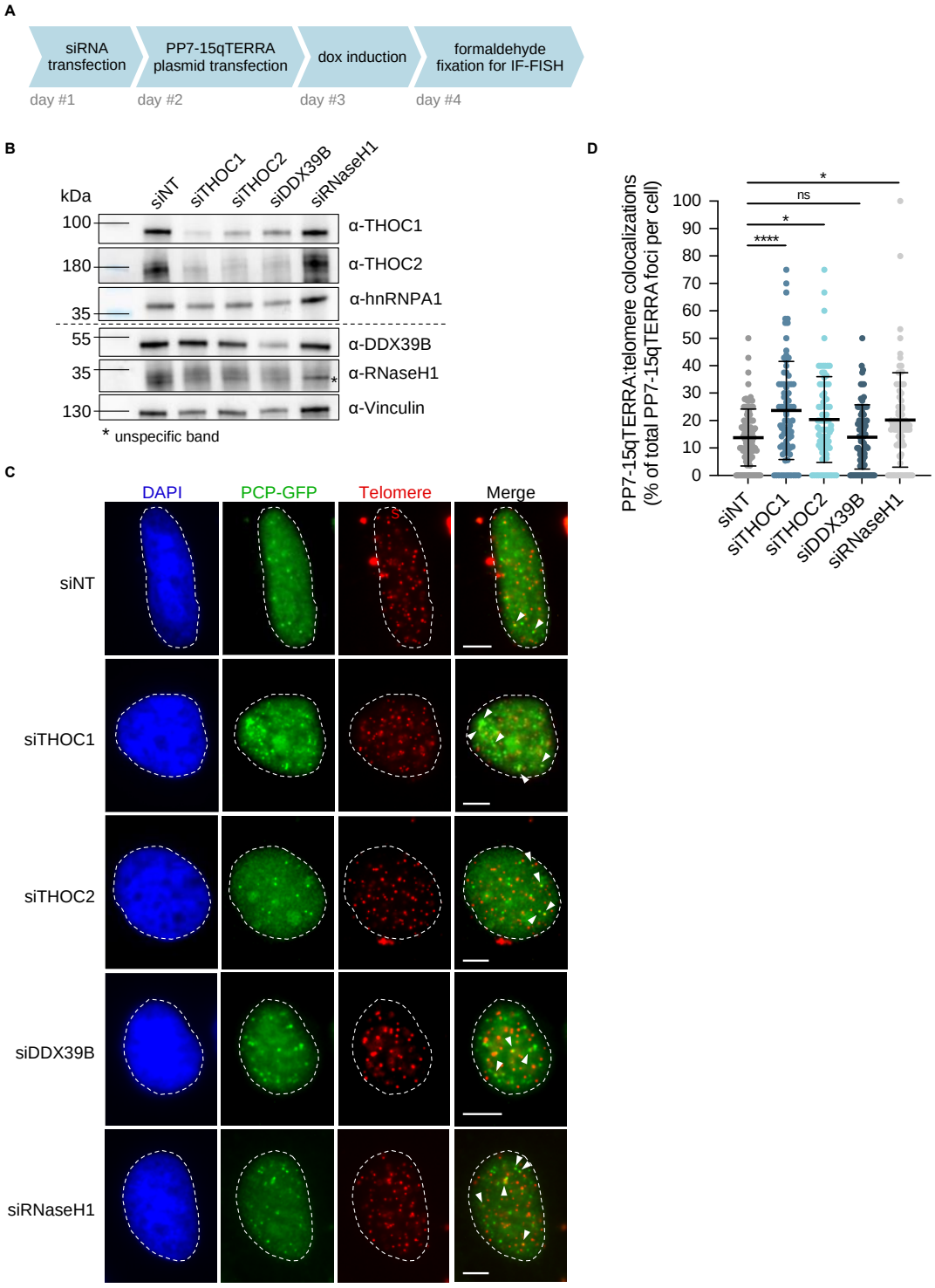


Figure 2. THOC subunits 1 and 2 counteract PP7-15qTERRA colocalization with telomeres.

A Experimental setup: HeLa cells with 10 kb average telomere length were transfected with siRNA pools, followed by PP7-15qTERRA-expressing plasmid. Cells were fixed with formaldehyde 24h post-PP715qTERRA induction and 72h after siRNA transfection.

B Western blot analysis of depletion efficiency with indicated siRNAs in HeLa cells with 10 kb average telomere length, used in C and D.

C Immunofluorescence of GFP-PCP (green) was employed to analyse co-localization of transiently expressed PP7-fused 15q-TERRA transcripts with telomeres (red) identified by fluorescence in situ hybridization (FISH). Representative images are shown for indicated conditions. White dashed lines outline the nuclear region and were determined based on DAPI-staining. White arrowheads indicate co-localization of GFP-PCP-tagged TERRA with telomeric FISH signal. Scale bars indicates 5 µm.

D Quantification of colocalization of GFP-PCP-tagged PP7-15qTERRA with telomeric FISH signal, as percentage of colocalization events over total GFP-PCP-tagged PP7-15qTERRA foci, per nucleus. At least 78 cells were analysed per condition, across three independent biological replicates. Horizontal line and error bars represent mean ± s.d.. One-way analysis of variance (ANOVA) with Dunnett's multiple comparisons test was applied: **** $P \leq 0.0001$, * $P \leq 0.05$, ns indicates non significance ($P > 0.05$).

The THOC complex counteracts R-loop-mediated PP7-15qTERRA associations with telomeres formed *in trans*

Aiming to determine the nature of PP7-15qTERRA associations with telomeres, prompted by loss of THOC subunits, we generated HeLa cells with inducible RNaseH1 overexpression by lentiviral transduction. In this experimental setup, PP7-15qTERRA expression is induced by doxycycline alongside with RNaseH1 overexpression (Fig 3A and B). Interestingly, we observed that PP7-15qTERRA:telomere colocalizations promoted by THOC deficiency are sensitive to RNaseH1 overexpression, being rescued to near basal levels (Fig 3C; Supplementary fig 2E). This suggests that THOC counteracts associations formed *in trans*, which take the form of direct DNA:RNA hybrids. Additionally, in cells with shorter telomeres, RNaseH1 overexpression could rescue the telomeric colocalizations induced by THOC siRNA transfection to the basal levels observed with a non-targeting siRNA (Supplementary fig 3A and B). However, upon RNaseH1 overexpression, colocalization levels in these cells could not be reduced to those observed in cells with longer telomere length, indicating that not all associations occur through direct DNA:RNA base-pairing.

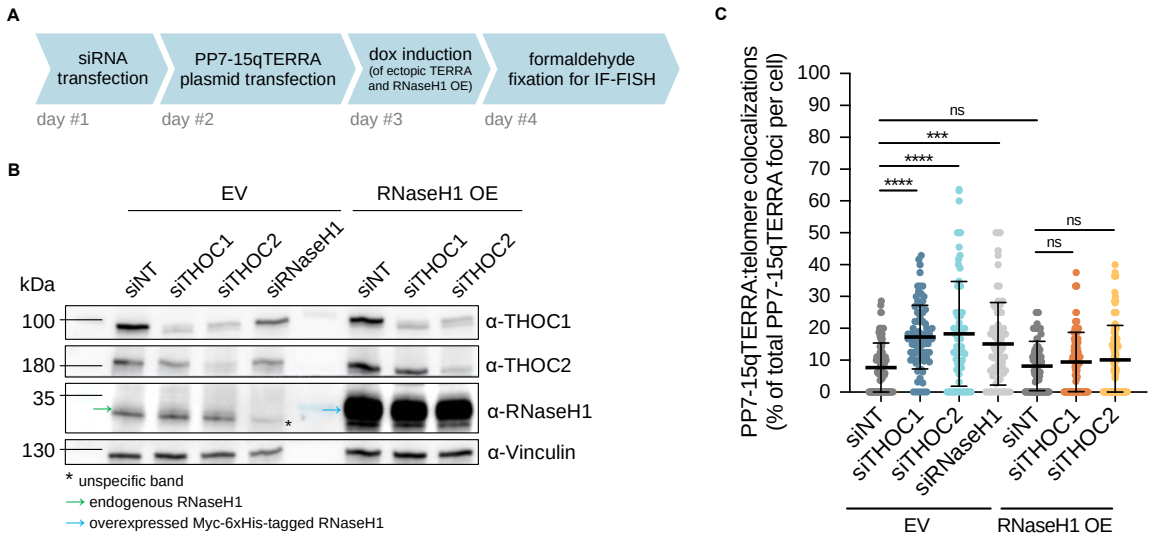


Figure 3. THOC subunits 1 and 2 counteract PP7-15qTERRA colocalization with telomeres, in the form of R-loops formed *in trans*.

A Experimental setup: HeLa cells with 10 kb average telomere length were generated by lentivirus-transduction for doxycycline-inducible overexpression of RNaseH1-Myc-6xHis (designated RNaseH1 OE) or respective empty vector (EV) control. Cells were transfected with siRNA pools followed by PP7-15qTERRA-expressing plasmid. Cells were fixed with formaldehyde 24h post-PP715qTERRA induction and 72h after siRNA transfection.

B Western blot analysis of depletion efficiency with indicated siRNAs in HeLa cells overexpressing RNaseH1 or in control cells, used in C.

C Quantification of colocalization of GFP-PCP-tagged PP7-15qTERRA with telomeric FISH signal, as percentage of colocalization events over total GFP-PCP-tagged PP7-15qTERRA foci, per nucleus. At least 74 cells were analysed per condition, across three independent biological replicates. Horizontal line and error bars represent

mean ± s.d.. Two-way analysis of variance (ANOVA) with Tukey's multiple comparisons test was applied: **** P ≤ 0.0001, *** P ≤ 0.001, ns indicates non significance (P > 0.05).

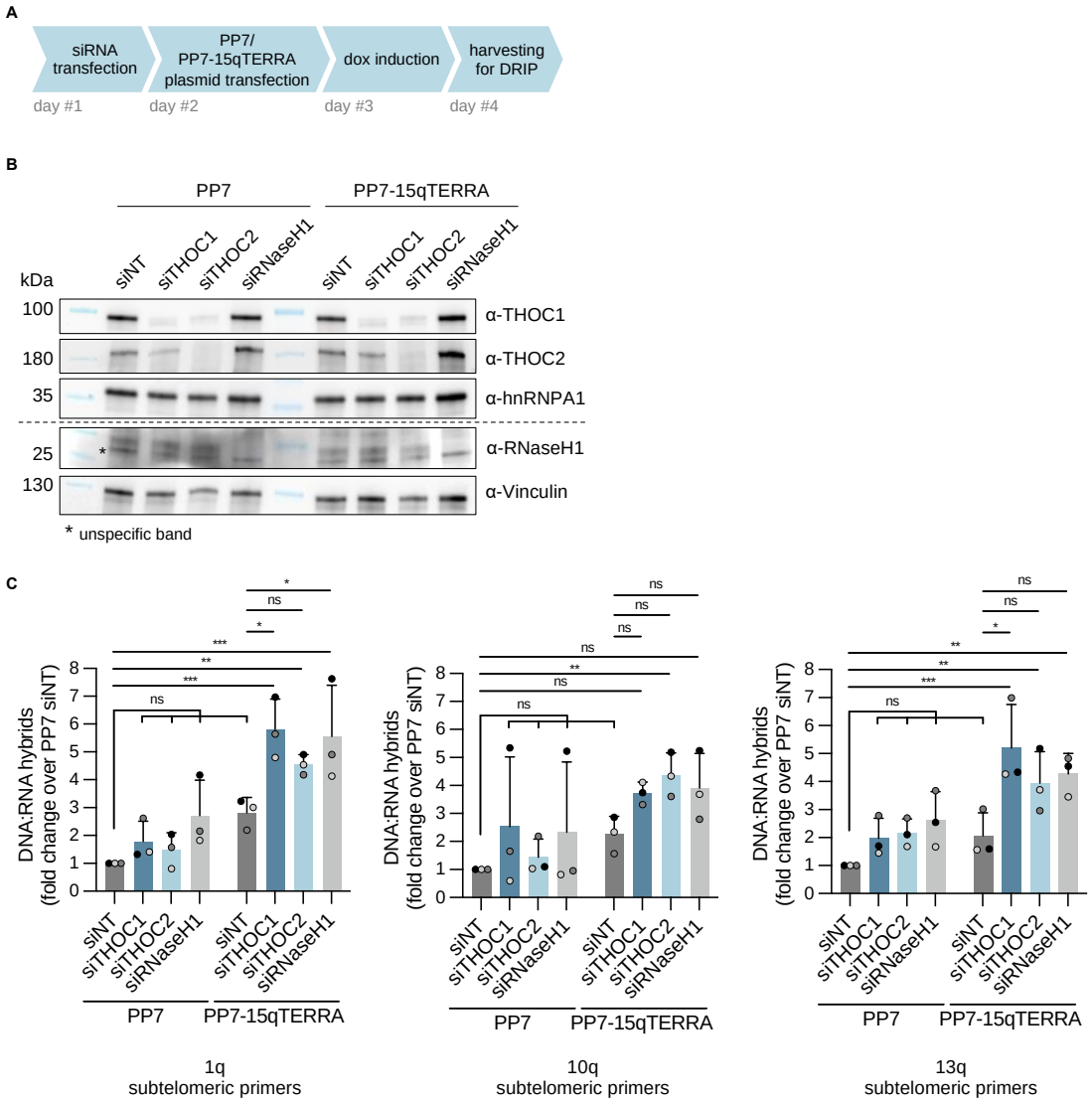


Figure 4. THOC counteracts TERRA R-loops formed *in trans*, post-transcriptionally.

A Experimental setup: HeLa cells with 10 kb average telomere length were transfected with siRNA pools, followed by PP7- or PP7-15qTERRA-expressing plasmids. Cells were collected for DNA:RNA immunoprecipitation (DRIP) 72h following siRNA transfection.

B Western blot analysis of depletion efficiency with indicated siRNAs in HeLa cells with 10 kb average telomere length, used in C.

C DRIP assay using anti-DNA:RNA hybrid S9.6 antibody was performed in extracts from HeLa cells with average 10 kb telomere length. Immunoprecipitates and input samples were treated with RNase (DNase-free), purified and analysed by qPCR with primer sets amplifying 1q, 10q or 13q subtelomeric DNA. Data are mean fold change over PP7 siNT ± s.d.. Two-way analysis of variance (ANOVA) with Dunnett's multiple comparisons test was applied: *** P ≤ 0.001, ** P ≤ 0.01, * P ≤ 0.05, ns indicates non significance (P > 0.05).

To corroborate the formation of R-loops structures at telomeres upon expression of transgenic TERRA (Fig 4A and B), we employed DRIP. In this experiment we analysed the pulled down DNA component of DNA:RNA hybrids by qPCR, using subtelomeric-specific primers residing in sequences adjacent to the telomeric TTAGGG repetitive tract, namely 1q, 10q and 13q subtelomeres. This ensures that R-loops formed at chromosome ends are being detected, as opposed to those which may form at the repetitive tract of plasmids from which PP7-15qTERRA RNAs are expressed (which would likely be

detected when analysing DRIP samples by DNA dot blot with a probe targeting telomeric repeats). While R-loops were only slightly (non-significantly) increased upon THOC depletion under endogenous TERRA expression (PP7-control expression), as well as upon TERRA overexpression (combined with control siRNA transfection), a noticeable increase of telomeric R-loops was detected when TERRA overexpression was combined with depletion of THOC subunits (Fig 4C). This is consistent with the RNaseH1-sensitive TERRA:telomeric associations induced by loss of THOC subunits 1 or 2 (Fig 3C), and suggests that THOC has the capacity to impact not only R-loops which are formed co-transcriptionally, but also those which form post-transcriptionally *in trans*.

The THO complex occupancy at the telomeric tract is increased upon loss of RNaseH1

The THO complex was previously found to be part of the human telomeric proteome (Grolimund *et al*, 2013). Thus, we wished to elucidate the basis of THOC recruitment to telomeres, in order to better understand the mechanism behind the observed effect in counteracting telomeric R-loops. For that end, we depleted RNaseH1 (Fig 5A) – which subtly elevates telomeric R-loop levels at telomeres (Fig 4C) – and evaluated THOC1 occupancy at telomeres by chromatin immunoprecipitation (ChIP) with an antibody against this THOC subunit. Pulled-down DNA was analyzed by DNA dot blot hybridization with a C-rich telomeric probe (Fig 5B). Interestingly, THOC1 occupancy at telomeres showed a near 2-fold increase when RNaseH1 was depleted, compared to control cells (Fig 5C). This suggests that upon R-loop accumulation at the telomeric tract, THOC is recruited to resolve them. Thus, THOC may assist the transcriptional machinery during elongation, and it may remove this topological obstacle from DNA also post-transcriptionally, in a timely manner, to prevent encounters with the replication machinery.

Endogenous nucleoplasmic TERRA RNA associates with the THO complex

THOC is a known RNA-binding complex (Jimeno, 2002; Masuda *et al*, 2005). To determine whether THOC subunits associate with TERRA, we carried out native RNA immunoprecipitation (RNA IP) with antibodies against THOC subunits 1 and 2 (Fig 5D; Supplementary fig 4A). As a control, hnRNPA1 was also immunoprecipitated (Fig 5D; Supplementary fig 4A), since it is a known TERRA-binding protein (Redon *et al*, 2013). Remarkably, while THOC1 is specifically immunoprecipitated by anti-THOC1 antibody, both subunits THOC1 and -2 are immunoprecipitated by anti-THOC2 antibody. Pulled-down RNA was then examined by RNA dot blot, probed with a ^{32}P-radiolabeled $[CCCTAA]_3$ probe (Fig 5E). As hnRNPA1, both THOC1 and THOC2 associate with endogenous nucleoplasmic TERRA in nuclear extracts, irrespective of telomere length (Fig 5E and F). Conversely, the very abundant 18S rRNA is not significantly bound by THOC (Supplementary fig 4B and C). In addition, we performed RNA IP in Hek293E cells transfected with shRNAs designed to downregulate THOC subunits 1 and 2 (Supplementary fig 5A and B). This validated the immunoprecipitation with anti-THOC1 and -THOC2 antibodies, since lower TERRA amounts were pulled down when compared to cells transfected with control shRNA (Supplementary fig 5C-G). Additionally, nucleic acids labelled by TERRA probe pulled down by native RNA immunoprecipitation with anti-THOC1 and -THOC2 antibodies are fully sensitive to RNase (DNase-free) digestion (Supplementary fig 5C).

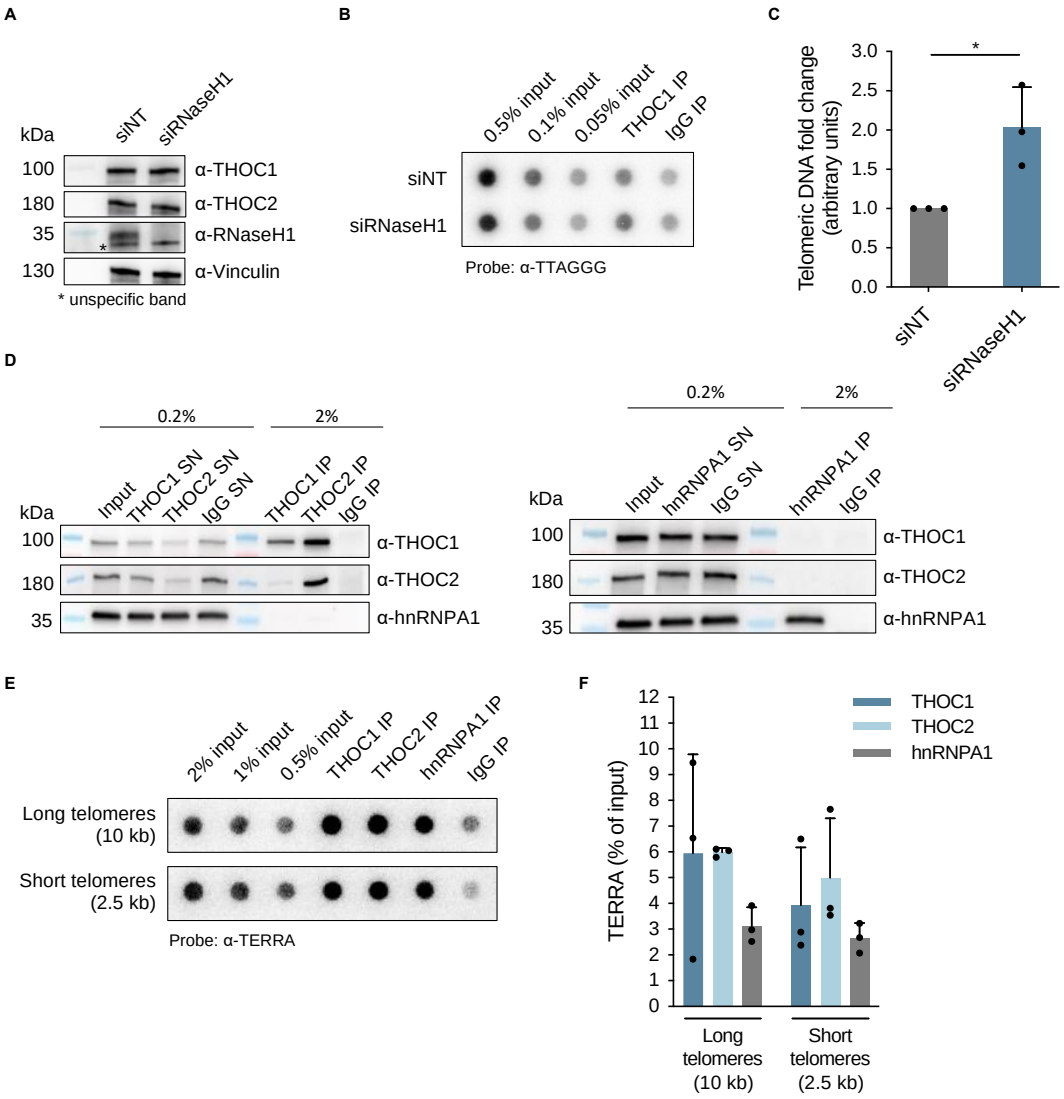


Figure 5. THOC1 occupancy at the telomeric tract is increased upon R-loop accumulation following RNaseH1 depletion; THOC associates with endogenous nucleoplasmic TERRA.

A Western blot analysis of depletion efficiency with siRNaseH1 in HeLa cells with average 10 kb telomere length, used in B and C.

B Chromatin-immunoprecipitation (ChIP) assay using anti-THOC1 antibody was performed in extracts from HeLa cells with average 10 kb telomere length, transfected with siRNA to deplete RNaseH1 or control non-targeting siRNA (siNT). Samples were analysed by DNA dot blot probed with a ^{32}P-radiolabelled $[CCCTAA]_3$ probe.

C Quantification of immunoprecipitated telomeric DNA (as in B), as fold change over siNT. The signal corresponding to immunoprecipitates using IgG antibody was subtracted from corresponding test immunoprecipitation signals as background. Data represent mean ± s.d., from three independent biological replicates. Two-tailed ratio paired t-test was applied: * $P \leq 0.05$.

D Western blot was used to evaluate the efficiency of immunoprecipitation of THOC1 and THOC2 (left hand side panel), and hnRNPA1 (right hand side panel). Samples obtained from assay with HeLa cells with average 10 kb telomere length were used in shown blots as an example.

E Native RNA-immunoprecipitation (RNA-IP) assay using anti-THOC1, anti-THOC2 and anti-hnRNPA1 antibodies was performed in extracts from HeLa cells with long (average 10 kb, top) or short (average 2.5 kb, bottom) telomere length. Samples were analysed by RNA dot blot probed with a ^{32}P-radiolabelled $[CCCTAA]_3$ probe.

F Quantification of immunoprecipitated TERRA (as in E), as percent of input. The signal corresponding to immunoprecipitates using IgG antibody was subtracted from corresponding test immunoprecipitation signals as background. Data represent mean ± s.d., from three independent biological replicates.

The THO complex counteracts TERRA R-loop-mediated telomeric fragility

Telomeres are difficult to replicate regions, given their repetitive sequence and the multiple topological hurdles encountered at the telomeric tract. These include G-quadruplex (G4) structures, that can form on the lagging strand template, as well as R-loops, which can form on the telomeric leading strand (Granotier *et al*, 2005; Balk *et al*, 2013; Arora *et al*, 2014; Sagie *et al*, 2017). Replication-dependent defects at telomeres have been shown to result in a fragile telomere phenotype observed in metaphase chromosomes. This feature is thought to originate from incomplete DNA replication or partial chromatin condensation, and is manifested as smeary or multiple telomeric FISH signals per chromosome end (Sfeir *et al*, 2009) (Fig 6). Additionally, the accumulation of telomeric G-quadruplexes and R-loop structures has previously been correlated with increased telomere fragility (Vannier *et al*, 2012; Yang *et al*, 2020; Lin *et al*, 2021; Arora *et al*, 2014; Teasley *et al*, 2015; Petti *et al*, 2019; Feretzaki *et al*, 2020).

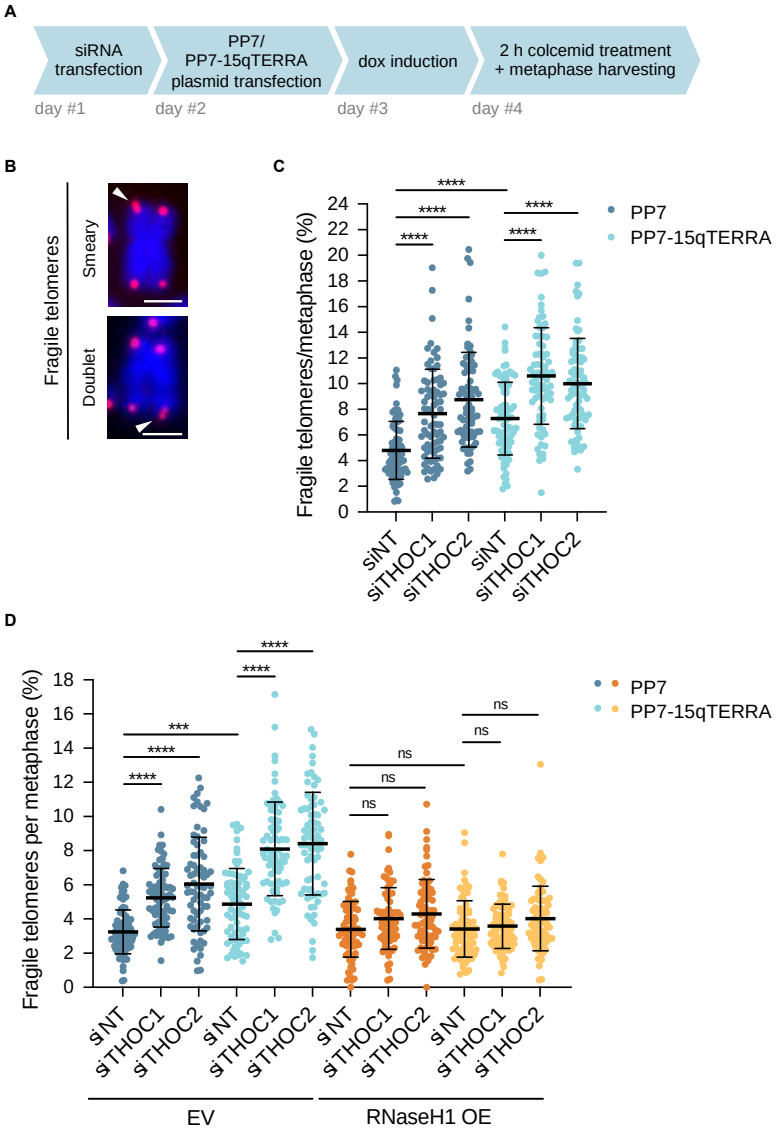


Figure 6. THOC prevents telomeric fragility induced by TERRA R-loops.

A Experimental setup: HeLa cells with 10 kb average telomere length were transfected with siRNA pools, followed by PP7-15qTERRA-expressing plasmid. Samples were collected after 2 h demecolcine treatment for enrichment of cells in metaphase.

B Telomeric FISH on metaphase spreads of HeLa cells stained with Cy3-$[CCCTAA]_3$ FISH probe (red) and DAPI (blue). White arrowheads indicate fragile telomeres, either with a smeary FISH signal (top) or multiple FISH signals (bottom) in a single chromosome arm end. Scale bars indicate 2 μm.

C Quantification of fragile telomeres, as percentage of events per metaphase spread. At least 77 metaphase spreads were analysed per condition, across three independent biological replicates. Horizontal line and error bars represent mean ± s.d.. Two-way analysis of variance (ANOVA) with Tukey's multiple comparisons test was applied: **** $P \leq 0.0001$.

D Quantification of fragile telomeres, as percentage of events per metaphase spread. Cells were lentivirus-transduced for doxycycline-inducible overexpression of RNaseH1-Myc-6xHis (designated RNaseH1 OE) or respective empty vector (EV) control. At least 71 metaphase spreads were analysed per condition, across three independent biological replicates. Horizontal line and error bars represent mean ± s.d.. Two-way analysis of variance (ANOVA) with Tukey's multiple comparisons test was applied: **** $P \leq 0.0001$; *** $P \leq 0.001$, ns indicates non significance ($P > 0.05$).

We set out to investigate whether loss of THOC subunits (Fig 6A; Supplementary fig 6A) impacts telomeric fragility. In this analysis, depletion of the shelterin component TRF1 was included as a positive control, since it has been shown to suppress telomeric fragility, namely by recruiting the G4-unwinding BLM helicase and the transcription initiation factor TFIIH, thus sustaining replication of lagging and leading strand telomeres (Sfeir *et al*, 2009; Zimmermann *et al*, 2014; Yang *et al*, 2020, 2022). Importantly, depletion of THOC subunits elevated the frequency of detected fragile telomeres (Fig 6C; Supplementary 6B). This effect was also observed upon PP7-15qTERRA expression, as previously described (Feretzaki *et al*, 2020), and further increased when combined with THOC depletion (Fig 6C; Supplementary 6B). In addition to the fragile telomere phenotype, other telomere abnormalities were quantified, including telomere loss and outsider telomeres (Supplementary Fig 6B and C), the latter being visualized as a telomeric FISH signal detached from chromosome ends (Majerska *et al*, 2018). However, no significant changes were detected in the amount of outsider or lost telomeres upon depletion of THOC subunits, under endogenous or ectopically expressed TERRA levels (Supplementary Fig 6B). As expected, TRF1 depletion elevated the frequency of fragile telomeres. Furthermore, TRF1 depletion resulted in the accumulation of outsider telomeres (Supplementary Fig 6B).

To determine the origin of this telomeric fragility phenotype, we investigated the impact of THOC in telomeric fragility in RNaseH1 overexpressing cells (Supplementary fig 6D). In these cells, fragility levels remained unaffected by depletion of THOC subunits, even when TERRA was overexpressed (Fig 6D; Supplementary 6E). This underscores the notion that the fragility phenotype prompted by loss of THOC subunits stems from telomeric R-loops, and indicates that THOC alleviates R-loop-derived telomeric fragility, thus contributing to the suppression of telomeric replication defects. Of note, abnormalities induced by loss of TRF1 could not be rescued by RNaseH1 overexpression (Supplementary fig 6E), consistently with previously acquired data in HeLa cells (Lee *et al*, 2018). This suggests that different pathways are undertaken by THOC and TRF1 to prevent telomeric aberrations.

The THO complex counteracts leading and lagging strand telomeric fragility

To better characterize the telomeric fragility observed as a consequence of R-loop accumulation, we performed chromosome-orientation telomeric FISH (CO-FISH) (Bailey *et al*, 1996). In this experiment, BrdU and BrdC are incorporated into the DNA through semi-conservative replication, during a single replication cycle (Fig 7A). Upon collection of metaphase spreads, the nascent replicated strand – which incorporated BrdU/C – is degraded by UV irradiation, followed by digestion with exonuclease III. This allows specific labelling of the parental strands with different fluorescent dyes: the chromosome end replicated by lagging strand synthesis is labelled in red, while the one replicated by leading strand synthesis in green (Fig 7A and B). Since optimal CO-FISH staining conditions were achieved in cells with long telomeres of approximately 30 kb length, telomeric FISH staining in metaphase spreads was repeated in these cells, so that a correlation could be established between telomeric FISH and CO-FISH stainings. Of note, basal fragility levels detected by FISH in these cells were higher (Fig 7C) than in cells with 10 kb telomeres (Fig 6C), consistent with the fact that telomere length correlates with the probability for the replication machinery to encounter obstacles, resulting in telomeric fragility. As expected, analysis of telomeric fragility by FISH upon depletion of THOC or TRF1 in cells with 30 kb average telomere length telomeres (Fig 7C; Supplementary fig 7A) reproduced the

trend observed in cells with 10 kb telomeres (Fig 6C; Supplementary fig 6A). Interestingly, in CO-FISH experiments, depletion of THOC subunits elevated telomeric fragility at both lagging and – most prominently – leading strand telomeres (Fig 7D). Upon TRF1 depletion, telomeric fragility also increased at both strands (Fig 7D), as previously reported (Sfeir *et al*, 2009; Lee *et al*, 2018). In addition, TRF1 depletion resulted in increased frequency of outsider telomeres, at both strands – most notably at the telomere replicated by lagging strand synthesis (Fig 7E and F) – a phenotype not commonly observed in mouse cells deleted of TRF1 (Sfeir *et al*, 2009; Yang *et al*, 2022). THOC depletion did not impact telomere loss or outsider telomere frequency (Fig 7E and F).

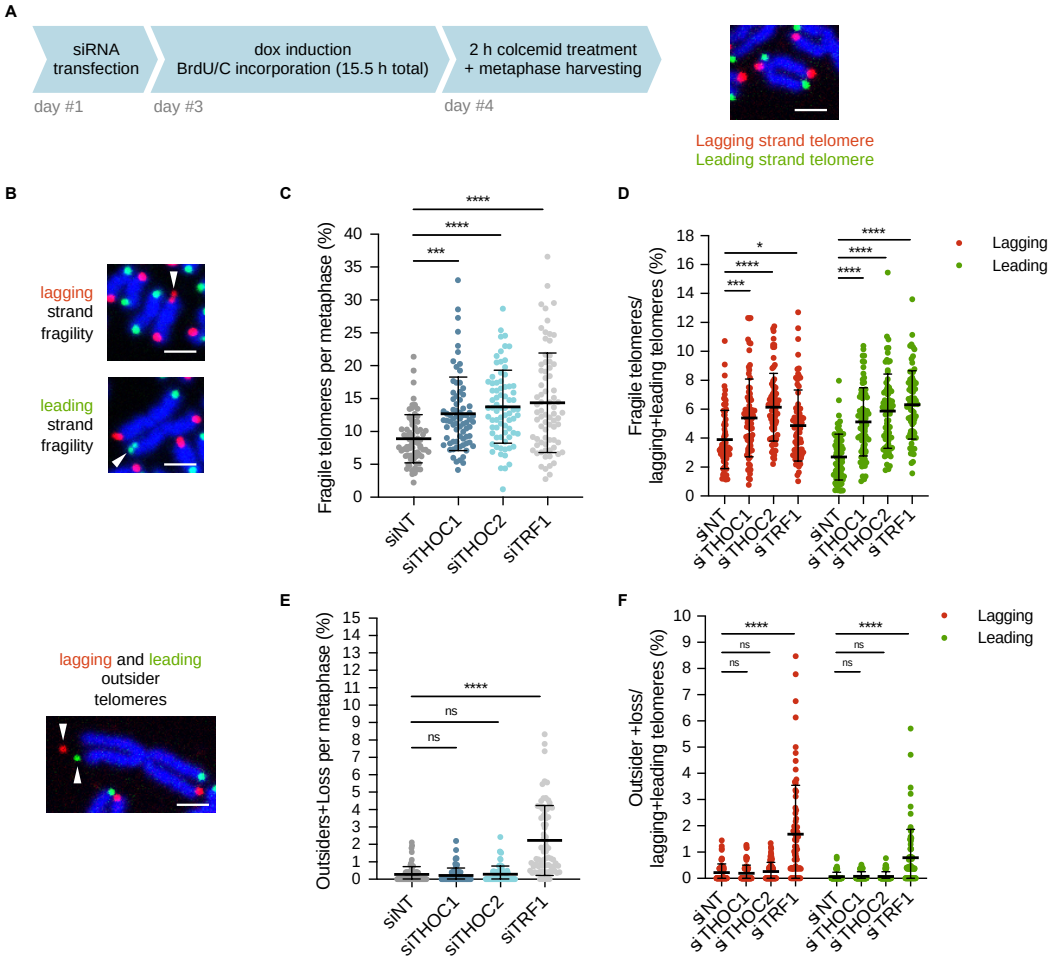


Figure 7. THOC prevents telomeric fragility at both lagging and leading strands.

A Experimental setup: HeLa cells with 30 kb average telomere length were transfected with siRNA pools, followed by BrdU/C (3:1) incorporation for a total of 15.5. Samples were collected after 2 h demecolcine treatment, for enrichment of cells in metaphase (left hand side). Representative metaphase chromosome labelled by CO-FISH, where lagging strand telomeres are labelled in red and leading strand telomeres are labelled in green (right hand side). Scale bar indicates 2 µm.

B Telomeric CO-FISH on metaphase spreads of HeLa cells stained with TYE563-TeloC LNA probe (red), FAM-TeloG LNA probe (green) and DAPI (blue). White arrowheads indicate fragile (top) or outsider (bottom) telomeres, as indicated. Scale bars indicate 2 µm.

C, E Quantification of fragile telomeres (C), or outsider and lost telomeres (E) in metaphase chromosomes stained by FISH, as percentage of events per metaphase spread (in HeLa cells with average 30 kb telomere length). 75 metaphase spreads were analysed per condition, across three independent biological replicates. Horizontal line and error bars represent mean ± s.d.. One-way analysis of variance (ANOVA) with Dunnett's multiple comparisons test was applied: **** $P \leq 0.0001$, *** $P \leq 0.001$, ns indicates non significance ($P > 0.05$).

D, F Quantification of lagging and leading strand fragile telomeres (D) or outsider and lost telomeres (F) in metaphase chromosomes stained by chromosome orientation FISH (CO-FISH), as percentage of events per metaphase spread (sum of lagging and leading strand telomeres, in HeLa cells with average 30 kb telomere

length). At least 74 metaphase spreads were analysed per condition, across three independent biological replicates. Horizontal line and error bars represent mean ± s.d.. One-way analysis of variance (ANOVA) with Dunnett's multiple comparisons test was applied: **** $P \leq 0.0001$, *** $P \leq 0.001$, * $P \leq 0.05$, ns indicates non significance ($P > 0.05$).

The THO complex counteracts telomeric sister chromatid exchange in U2OS

The CO-FISH protocol not only allows to examine telomeric strand-specific abnormalities, but to investigate telomeric recombination events with exchange of parental strand DNA. We analyzed both types of single sister chromatid exchange events (single t-sce) – resulting in double leading or double lagging telomeric signals –, as well as reciprocal sister chromatid exchange (reciprocal t-sce) (Supplementary fig 7B; Fig 8A-E). Remarkably, depletion of THOC subunits 1 and 2, or TRF1 in HeLa cells (Supplementary fig 7A) led to no perceptible changes in sister chromatid exchange events, compared to control cells (Supplementary fig 7C). Conversely, in U2OS cells, depletion of THOC1 increased the frequency of reciprocal t-sce, while depletion of THOC2 led to a mild but significant increase of leading and lagging strand single t-sce (Fig 8C-E; Supplementary fig 8A). Knock down of BLM helicase was used as a positive control for reciprocal t-sce in U2OS cells (Sarkar *et al*, 2015; Sobinoff *et al*, 2017). These data indicate that the THO complex counteracts t-sce in ALT cells.

The THO complex counteracts C-circle accumulation in ALT cells

We then analysed another hallmark of ALT cells – the accumulation of extrachromosomal telomeric DNA in the form of partially single-stranded C-rich circles, designated C-circles (Henson *et al*, 2009). For that end, we employed the phi29-mediated C-circle assay (Henson *et al*, 2009). In U2OS cells, depletion of THOC subunits led to a remarkable increase of C-circles (most prominently upon siTHOC1), as detected by non-denaturing DNA dot blot (Fig 8F, G). This trend was also perceptible upon depletion of THOC1 or THOC2 in Saos2 cells (an ALT cell line with shorter telomere length than U2OS), although to a milder non-significant extent (Fig 8F, H; Supplementary fig 8B). In contrast, in HeLa cells – with no detectable basal levels of C-circles (as an ALT^- cell line) – depletion of THO subunits was not sufficient to trigger a discernible formation of such structures (Fig 8F).

Aiming to understand whether R-loops could play a role in the C-circle accumulation observed when THOC was downregulated in ALT cells, we transiently overexpressed RNaseH1 from a plasmid in U2OS cells (Supplementary fig 8C), and evaluated C-circle levels. RNaseH1 overexpression led to a partial rescue of C-circles, when combined with depletion of THOC1 (Fig 8I, J). Of note, C-circle levels prompted by loss of THOC2 did not exhibit sensitivity to RNaseH1 overexpression. The fact that ectopically-expressed RNaseH1 protein levels were slightly affected in siTHOC2-transfected cells, compared to non-targeting siRNA (Supplementary fig 8C), may contribute to the disparity in RNAseH1-sensitivity of C-circles detected in cells depleted of THOC1 or THOC2.

Together these data point to a role of the THO complex in counteracting C-circle accumulation in ALT cells, partially through the regulation of R-loops.

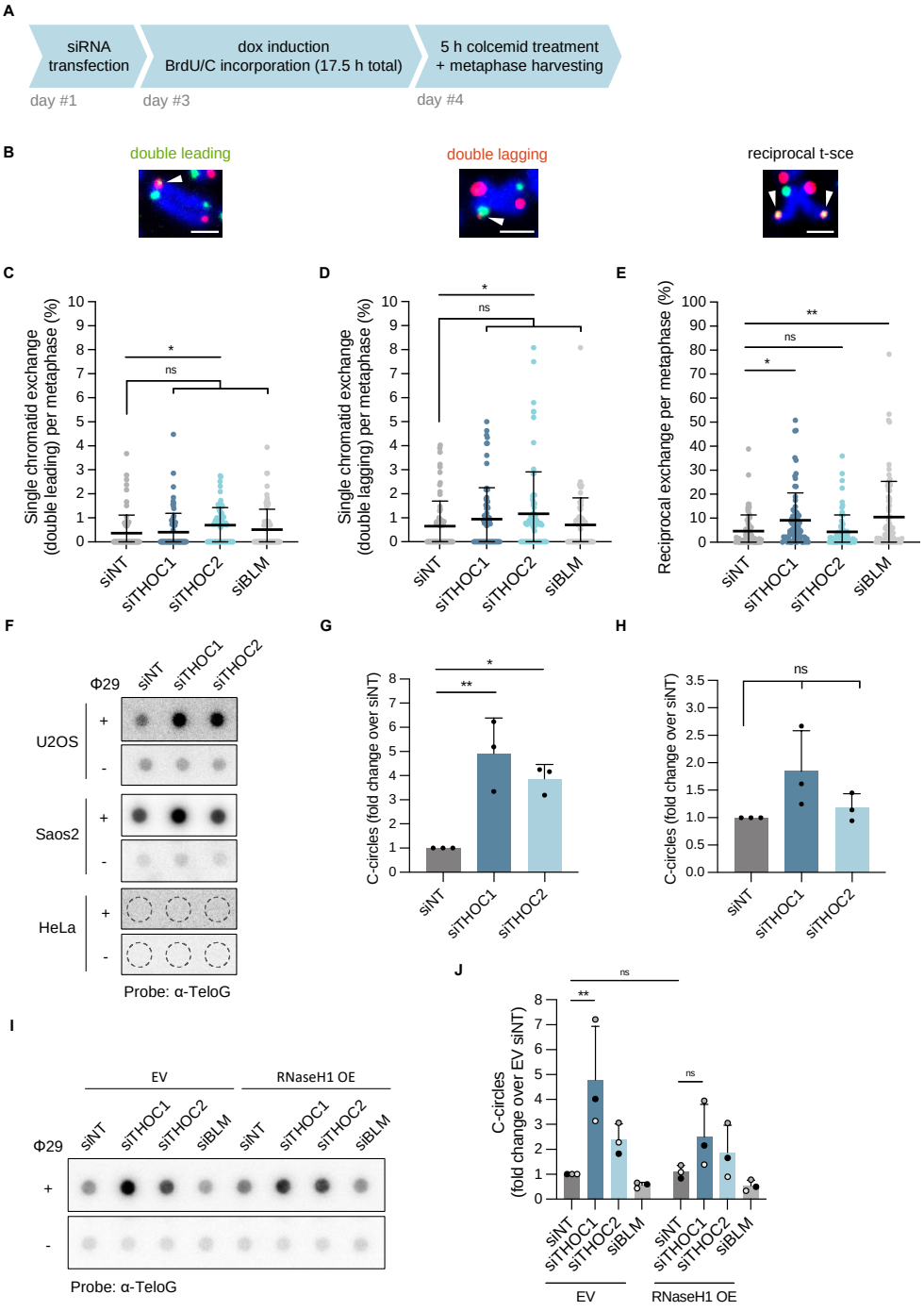


Figure 8. THOC telomeric sister chromatid exchange and C-circle accumulation in U2OS cells.

A Experimental setup: U2OS cells were transfected with siRNA pools, followed by BrdU/C (3:1) incorporation for a total of 17.5 hours. Samples were collected after 5 h demecolcine treatment, for enrichment of cells in metaphase.

B Telomeric CO-FISH on metaphase spreads of U2OS cells stained with TYE563-TeloC LNA probe (red), FAM-TeloG LNA probe (green) and DAPI (blue). White arrowheads indicate double leading (left-hand side), double lagging (centre) or reciprocal telomeric sister chromatid exchange (right-hand side), as indicated. Scale bars indicate 2 µm.

C-E Quantification of double leading telomeres (C), double lagging telomeres (D) or reciprocal telomeric sister chromatid exchange (E), as percentage of events per metaphase spread (sum of lagging and leading strand telomeres, in U2OS cells). 75 metaphase spreads were analysed per condition, across three independent biological replicates. Horizontal line and error bars represent mean ± s.d.. One-way analysis of variance (ANOVA)

with Dunnett's multiple comparisons test was applied: ** P ≤ 0.01, * P ≤ 0.05, ns indicates non significance (P > 0.05).

F Phi29 C-circle assay on 30 ng of gDNA from U2OS, Saos2 or HeLa cells, transfected with non-targeting or THOC1 or THOC2 siRNAs, as indicated. Amplification products were analysed by dot blot probed with a ^{32}P-radiolabelled C-rich telomeric probe. Phi29^{+} signals were acquired from non-denatured membrane, while Phi29^{-} signals were acquired from stripped, denatured and re-probed membrane.

G-H Quantification of C-circle assay on gDNA from U2OS cells (G) or Saos2 cells (H) (as in F), as fold change over siNT. Data represent mean ± s.d., from three independent biological replicates. One-way analysis of variance (ANOVA) with Dunnett's multiple comparisons test was applied: * P ≤ 0.05, ** P ≤ 0.01, ns indicates non significance (P > 0.05).

I Phi29 C-circle assay on 30 ng of gDNA from U2OS cells, transfected with indicated siRNAs, and with empty-vector (EV) or RNaseH1 overexpressing (OE) constructs. Amplification products were analysed by dot blot probed with a ^{32}P-radiolabelled C-rich telomeric probe. Phi29^{+} signals were acquired from non-denatured membrane, while Phi29^{-} signals were acquired from stripped, denatured and re-probed membrane.

J Quantification of C-circle assay on gDNA from U2OS cells (as in I), as fold change over EV siNT. Data represent mean ± s.d., from three independent biological replicates. Two-way analysis of variance (ANOVA) with Tukey's multiple comparisons test was applied: ** P ≤ 0.01, ns indicates non significance (P > 0.05).

5.6 Discussion

In this paper, we explore the roles of the THO complex at human telomeres. We show that THOC negatively regulates TERRA-mediated telomeric R-loops. This was most perceptible in U2OS ALT cells, which display a stronger accumulation of telomeric DNA:RNA hybrids when compared to telomerase-positive HeLa cells (Arora *et al*, 2014), rendering these structures more readily detectible by DRIP. The impact of THOC depletion in telomeric R-loops does not stem from changes in TERRA transcriptional elongation, nor TERRA expression levels, since no significant changes were detected in UUAGGG-containing RNA cellular levels upon depletion of THOC.

R-loops have previously been thought to form exclusively during transcription, where the nascent RNA strand can base pair with its template DNA *in cis* (Crossley *et al*, 2019; Lafuente-Barquero *et al*, 2020). However, studies in *Saccharomyces cerevisiae* (Wahba *et al*, 2013) and in *Arabidopsis thaliana* (Ariel *et al*, 2020) revealed that R-loops can be formed post-transcriptionally, *in trans* at genomic regions distinct from the loci where the RNA was transcribed from. Additionally, in human cells, we previously demonstrated that TERRA forms R-loops post-transcriptionally at telomeres, *in trans*, in a RAD51-mediated manner (Feretzaki *et al*, 2020). Resorting to a previously published proxy for the study of the nuclear localization of TERRA (Feretzaki *et al*, 2020), we found that THOC is able to counteract PP7-15qTERRA associations with chromosome ends, which are formed when TERRA is ectopically expressed from a plasmid. Particularly, we found that such associations take the form of DNA:RNA hybrids. This suggests that THOC is not only able to prevent the accumulation of co-transcriptionally formed R-loops, but also has the capacity to counteract R-loops formed post-transcriptionally, *in trans*.

A detailed mechanism through which THOC restrains DNA:RNA hybrids throughout the genome is not well elucidated. The discovery of an interplay between THOC and components of the Sin3 deacetylase complex suggested that THOC limits the formation of DNA:RNA hybrids indirectly, through the modulation of chromatin modifications (deacetylation of H3), which render the chromatin less accessible (Salas-Armenteros *et al*, 2017). In addition, DDX39B – a component of the RNA export TREX complex – was shown to directly resolve DNA:RNA hybrids through its helicase activity (Pérez-Calero *et al*, 2020). Here, we observed a recruitment of THOC to the telomeric tract when RNaseH1 was depleted. While RNaseH1 occupancy at telomeres was shown to be much more prominent in ALT cells compared to telomerase-positive cells (Arora *et al*, 2014), our DRIP analysis in (telomerase-positive) HeLa cells revealed a subtle but detectable increase in telomeric R-loops upon RNaseH1 depletion (even under endogenous TERRA expression only). Combined, these data indicate that THOC is recruited to telomeres when R-loops accumulate (Fig 9A). This is consistent with a function of THOC in assisting with

the progression of the elongating transcription machinery when R-loops accumulate (since THOC was previously found to interact with RNA Polymerase II (Li *et al*, 2005)). In addition, our data suggest a role of THOC in the resolution of DNA:RNA structures formed at telomeres post-transcriptionally (Fig 9A). Along with THOC subunits (Grolimund *et al*, 2013), the TREX component DDX39B was identified at human telomeres in a mass spectrometry-based approach to identify telomeric proteins in Hek293E cells (Lin *et al*, 2021). We have only observed a significant change in PP7-15qTERRA localization to telomeres in HeLa cells with ca. 2.5 kb average telomere length, upon DDX39B depletion (Supplementary fig. 1a), but not in cells with 10 kb average telomere length depleted of DDX39B. Thus, DDX39B may counteract R-loops specifically at short telomeres. It is unclear if at long telomeres THOC1/2 counteract R-loops directly or if another component is involved.

We show that THOC (subunits 1 and 2) binds nucleoplasmic endogenous TERRA in HeLa and Hek293E nuclear extracts, suggesting that THOC acts directly on TERRA. Together with the physical presence of THOC at telomeres, it suggests direct functions at chromosome ends. In human cells, THOC has been found to associate with spliced, 5' capped RNAs (Masuda *et al*, 2005; Cheng *et al*, 2006; Merz *et al*, 2007; Chi *et al*, 2013). Additionally, the m^6A modification of RNA was hypothesized to contribute to THOC recruitment (Lesbirel *et al*, 2018). Over 90% of human TERRA is not polyadenylated (Porro *et al*, 2010), and previous analysis of the 5' subtelomeric-derived region of TERRA did not reveal any splicing events undertaken during TERRA biogenesis (Porro *et al*, 2014). On the other hand, TERRA contains 7-methylguanosine 5' cap structures (Porro *et al*, 2010), and the m^6A modification was recently found to be present on the subtelomeric region of TERRA (Chen *et al*, 2022). While the basis of THOC recruitment to TERRA remains unclear, we hypothesize that loading of THOC onto nascent TERRA and steady binding to nucleoplasmic TERRA prevents invasion into telomeric DNA, and formation of co- and post-transcriptional R-loops, respectively (Fig 9A). Furthermore, our findings are consistent with a model in which THOC remains associated with nucleoplasmic TERRA following TERRA removal from DNA:RNA hybrid structures.

R-loops formed across the genome have been shown to pose an obstacle to the transcription machinery during elongation as well as an impairment to DNA replication, resulting in accumulation of DNA damage and genomic instability (Huertas & Aguilera, 2003; Wellinger *et al*, 2006; Gan *et al*, 2011; Domínguez-Sánchez *et al*, 2011). The fragile telomere phenotype is correlated with defects in the progression of the replication machinery at chromosome ends (Sfeir *et al*, 2009). In this study, we show that THOC counteracts telomeric fragility through its regulation of TERRA R-loops. Moreover, through CO-FISH staining of parental telomeric strands in metaphase chromosomes, we observed that telomeric fragility prompted by THOC deficiency is detected at telomeres replicated by lagging and most strikingly by leading strand synthesis. Leading strand telomeric fragility may be observed when factors that suppress telomeric R-loops are perturbed, while lagging strand telomeric fragility may result from accumulation of G4 structures, which have the propensity to form within the G-rich telomeric strand (Glousker & Lingner, 2021). Therefore, we can idealize a model in which loss of THOC results in build-up of telomeric R-loops, resulting in leading strand fragility. At the same time the displacement of the G-rich telomeric strand may stimulate formation and accumulation of G4s, which impair lagging strand replication (Fig 9B). Formation of G4s on the displaced strand by TERRA R-loops was recently proposed by Yadav and colleagues (Yadav *et al*, 2022).

Telomeric fragility is commonly assumed to originate from replication defects and the ensuing repair processes of damaged replication forks (Glousker & Lingner, 2021). Therefore, one can hypothesize that fragility occurs as a consequence of R-loop accumulation, which generates DNA damage at telomeres. Comparison of fragility assessed by FISH with fragility (at lagging plus leading strands) analyzed by CO-FISH in HeLa cells (with the same telomere length) depleted of THOC revealed similar fragility percentages (Fig 7C and D). This suggests that fragility elicited by accumulation of R-loops upon THOC depletion does not stem from break-induced replication alone, since this repair pathway relies on conservative DNA synthesis, and its outcome would therefore not be detected by CO-FISH staining, which involves degradation of newly synthesized DNA. In mouse cells, lagging strand fragility observed in *Blm*-deficient cells was proposed to result from G4 accumulation, followed by formation of double-

strand breaks repaired by break-induced replication coupled with alternative non-homologous end joining (Yang *et al*, 2020). Whether a similar mechanism stands behind the fragility detected by CO-FISH with our experimental setup needs to be elucidated.

TERRA levels are regulated throughout the cell cycle, reaching the lowest levels in late S phase when telomeres are replicated (Porro *et al*, 2010; Graf *et al*, 2017), presumably to reduce the frequency of R-loops which pose an obstacle to the progression of the DNA replication machinery. However, our data suggest that the repression of TERRA in S phase is not sufficient to completely eliminate telomeric R-loops, and that THOC is required to facilitate telomere replication by counteracting R-loops. Of note, the regulation of TERRA during the cell cycle is lost in ALT cancer cells (Flynn *et al*, 2015), explaining the increased telomere replication stress that is caused by TERRA in these cells (Arora *et al*, 2014; Silva *et al*, 2021). We have previously shown that transgenic PP7-15qTERRA expressed from plasmids loses cell cycle control and associates with telomeres in S phase (Valador Fernandes *et al*, 2021). Therefore, this system may recapitulate the replication stress observed in ALT cells.

Despite the observed increase in TERRA R-loops and telomeric fragility, no changes were detected in the frequency of sister chromatid exchange events in HeLa cells depleted of THOC subunits, compared to control siRNA. Thus, loss of THOC is not sufficient to induce this ALT-like hallmark phenotype in telomerase-positive cells. Additional events may be required to increase telomere recombination frequency, such as the modulation of the telomeric chromatin structure and telomerase inhibition (as observed upon depletion of the histone chaperones ASF1a+b in immortalized fibroblasts and HeLa cells (O'Sullivan *et al*, 2014)). A consensually used biomarker of cells which maintain telomere length via the ALT pathway instead of telomerase is, among others, an elevated frequency of t-sce (Londoño-Vallejo *et al*, 2004). We found that THOC counteracts single and reciprocal t-sce in U2OS cells (Fig 9B), likely by regulating the association of TERRA with telomeres via R-loops. Similarly, co-depletion of NONO and SFPQ in ALT cells has been shown to result in accumulation of telomeric R-loops, as well as an increased rate of telomeric sister chromatid exchange (Petti *et al*, 2019).

Finally, analysis of extrachromosomal telomeric DNA in the form of C-circles in ALT cells depleted of THOC subunits revealed a role for THOC in limiting the accumulation of C-circles (Fig 9B). Ectopic-expression of RNaseH1 in U2OS cells revealed that the detected increase in C-circle formation particularly prompted by loss of THOC1 is partially dependent on R-loops. The fact that RNaseH1 overexpression does not significantly impact basal C-circle levels in U2OS cells indicates that R-loops are not essential for their formation in ALT cells. In addition, our data suggest that R-loops mediate C-circle formation, as previously demonstrated upon dysregulation of other components involved in telomeric R-loop regulation (Arora *et al*, 2014; Lu *et al*, 2019; Silva *et al*, 2019; Sakellariou *et al*, 2022). R-loop-mediated C-circle formation may occur through processing of R-loops (directly through excision by endonucleases such as XPF or MUS81, and self-ligation of the C-rich DNA strand) and/or processing of DNA structures resulting from R-loop accumulation (for example following break-induced replication triggered by R-loops) (Sakellariou *et al*, 2022; Lu *et al*, 2019). Similarly to telomeric sister chromatid exchange events, no C-circle formation was detected in THOC-depleted HeLa cells.

While TERRA was shown to be a crucial trigger/amplifier of the ALT pathway, inappropriate restriction of TERRA-induced ALT activity is thought to compromise telomere integrity (Silva *et al*, 2019, 2021, 2022). Therefore, a fine regulation of TERRA and TERRA R-loops must be attained in ALT cells. Altogether, our work demonstrates that the THO complex is one of the factors that are involved in such intricate balance at telomeres.

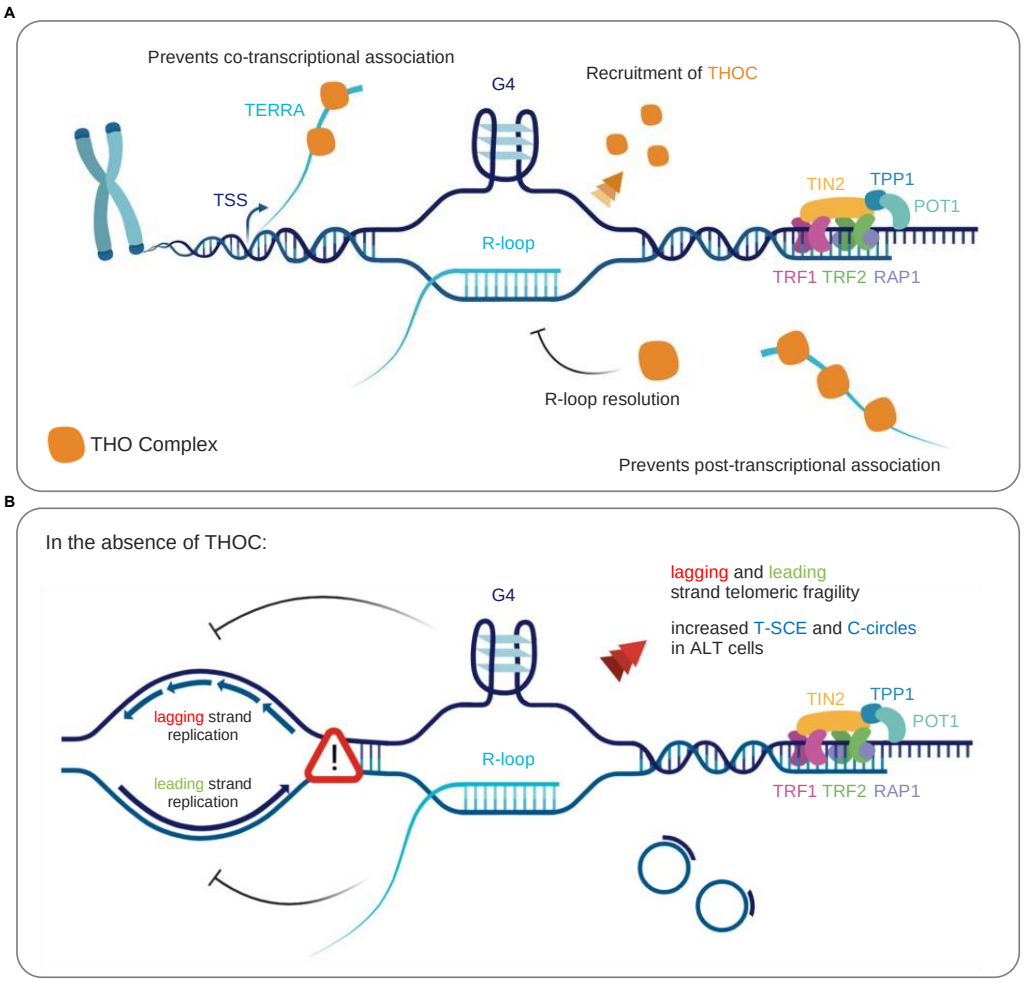


Figure 9. Working model for the role of THOC at human telomeres.

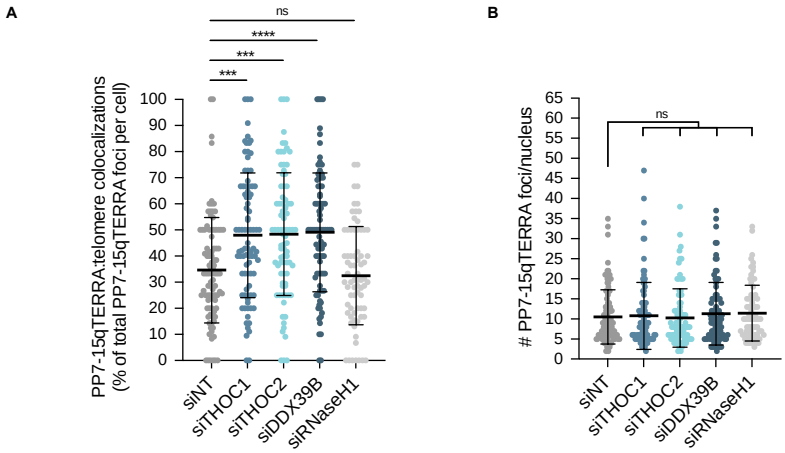
A Interplay of THOC with TERRA and putative functions of THOC at human telomeres.

B Impact of THOC deficiency in telomeric fragility, telomeric sister chromatid exchange and C-circle accumulation.

See discussion for details.

This illustration was created with BioRender.com.

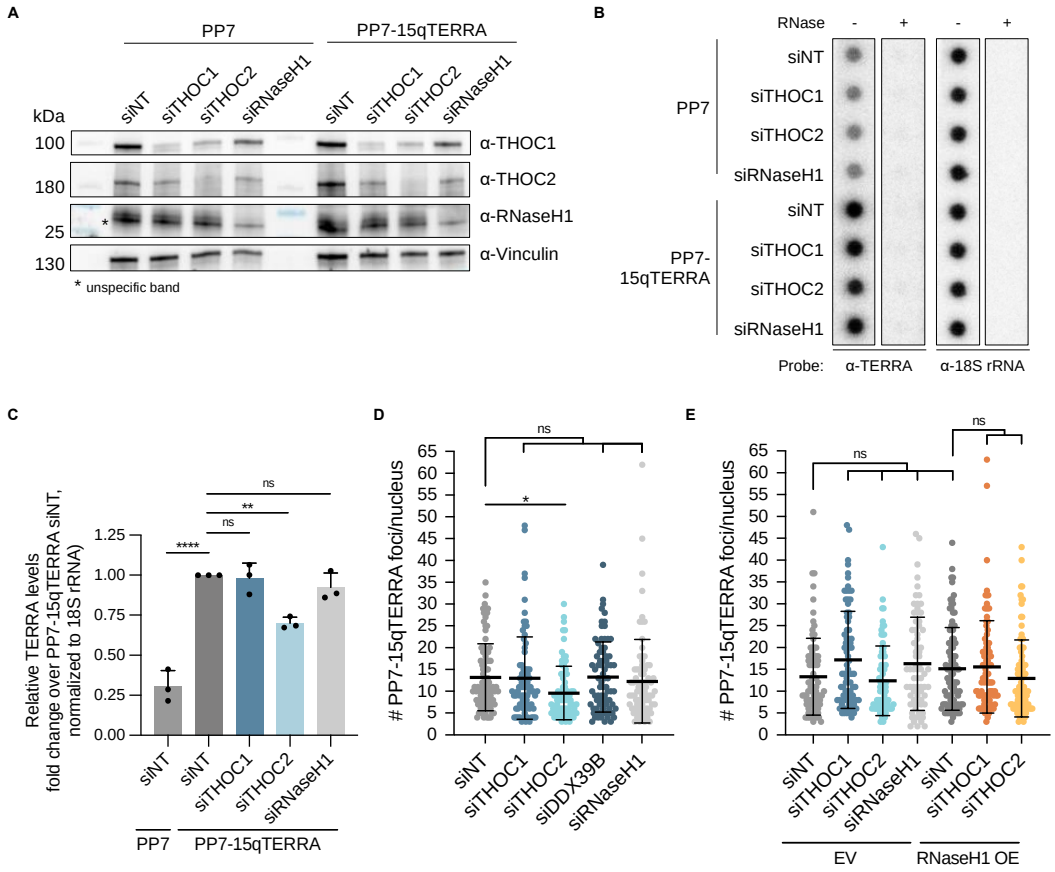
5.7 Supplementary figures and tables



Supplementary figure 1. THOC counteracts PP7-15qTERRA colocalization with telomeres in cells with short telomeres.

A Quantification of colocalization of GFP-PCP-tagged PP7-15qTERRA with telomeric FISH signal, as percentage of colocalization events over total GFP-PCP-tagged PP7-15qTERRA foci, per nucleus, in HeLa cells with ca. 2.5 kb average telomere length. At least 73 cells were analysed per condition, across three independent biological replicates. Horizontal line and error bars represent mean ± s.d.. One-way analysis of variance (ANOVA) with Dunnett's multiple comparisons test was applied: **** $P \leq 0.0001$, *** $P \leq 0.001$, ns indicates non significance ($P > 0.05$).

B Number of total PP7-15qTERRA foci per nucleus, in cells of each indicated condition used for quantification of colocalization of GFP-PCP-tagged PP7-15qTERRA with telomeric FISH signal (as in A). At least 73 cells were analysed per condition, across three independent biological replicates. Horizontal line and error bars represent mean ± s.d.. One-way analysis of variance (ANOVA) with Dunnett's multiple comparisons test was applied: ns indicates non significance ($P > 0.05$).



Supplementary figure 2. Analysis of PP7-15qTERRA levels by RNA dot blot and microscopy.

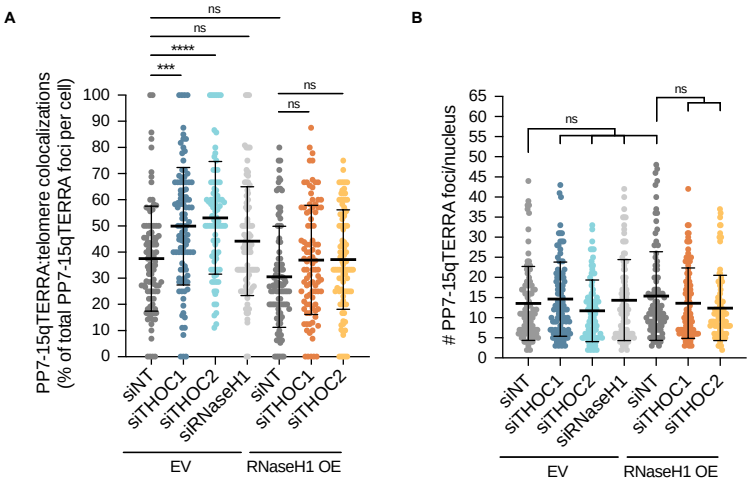
A Western blot analysis of depletion efficiency with indicated siRNAs in HeLa cells with average 10 kb telomere length.

B RNA dot blot analysis of TERRA or 18S rRNA levels upon depletion with indicated siRNAs in HeLa cells transfected with PP7- or PP7-15qTERRA-expressing plasmids. 4 µg total RNA were loaded per sample. RNA samples were treated with RNase (DNase-free) as a control.

C Quantification of TERRA levels (as in B), normalized to 18S rRNA levels in HeLa cells, plotted as fold change over PP7-15qTERRA siNT. Data are mean ± s.d.. One-way analysis of variance (ANOVA) with Dunnett's multiple comparisons test was applied: **** $P \leq 0.0001$, ** $P \leq 0.01$, ns indicates non-significance ($P > 0.05$).

D Number of total PP7-15qTERRA foci per nucleus, in HeLa cells with average 10 kb telomere length of each indicated condition, used for quantification of colocalization of GFP-PCP-tagged PP7-15qTERRA with telomeric FISH signal. At least 78 cells were analysed per condition, across three independent biological replicates. Horizontal line and error bars represent mean ± s.d.. One-way analysis of variance (ANOVA) with Dunnett's multiple comparisons test was applied: * $P \leq 0.05$, ns indicates non significance ($P > 0.05$).

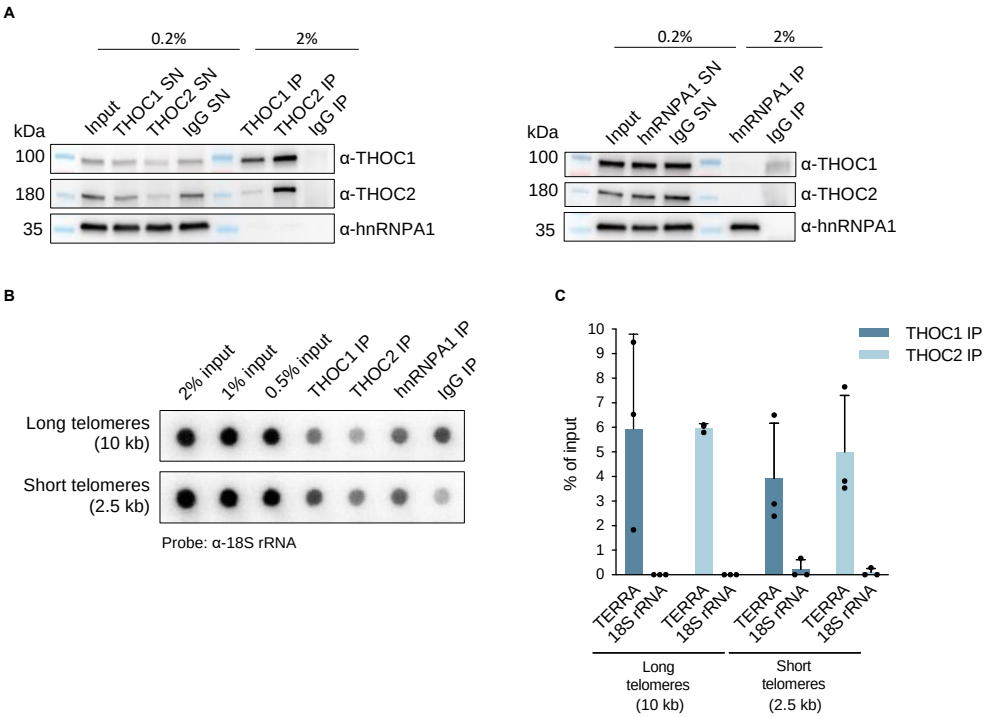
E Number of total PP7-15qTERRA foci per nucleus, in HeLa cells with average 10 kb telomere length of each indicated condition used for quantification of colocalization of GFP-PCP-tagged PP7-15qTERRA with telomeric FISH signal. At least 74 cells were analysed per condition, across three independent biological replicates. Horizontal line and error bars represent mean ± s.d.. Two-way analysis of variance (ANOVA) with Tukey's multiple comparisons test was applied: ns indicates non significance ($P > 0.05$).



Supplementary figure 3. PP7-15qTERRA telomeric colocalizations prompted by loss of THOC are partially sensitive to RNaseH1 overexpression, in cells with short telomeres.

A Quantification of colocalization of GFP-PCP-tagged PP7-15qTERRA with telomeric FISH signal, as percentage of colocalization events over total GFP-PCP-tagged PP7-15qTERRA foci, per nucleus, in HeLa cells with 2.5 kb average telomere length, overexpressing RNaseH1 or in control cells, as indicated. At least 84 cells were analysed per condition, across three independent biological replicates. Horizontal line and error bars represent mean ± s.d.. One-way analysis of variance (ANOVA) with Dunnett's multiple comparisons test was applied: **** $P \leq 0.0001$, *** $P \leq 0.001$, ns indicates non significance ($P > 0.05$).

B Number of total PP7-15qTERRA foci per nucleus, in cells of each indicated condition used for quantification of colocalization of GFP-PCP-tagged PP7-15qTERRA with telomeric FISH signal (as in A). At least 84 cells were analysed per condition, across three independent biological replicates. Horizontal line and error bars represent mean ± s.d.. One-way analysis of variance (ANOVA) with Dunnett's multiple comparisons test was applied: ns indicates non significance ($P > 0.05$).

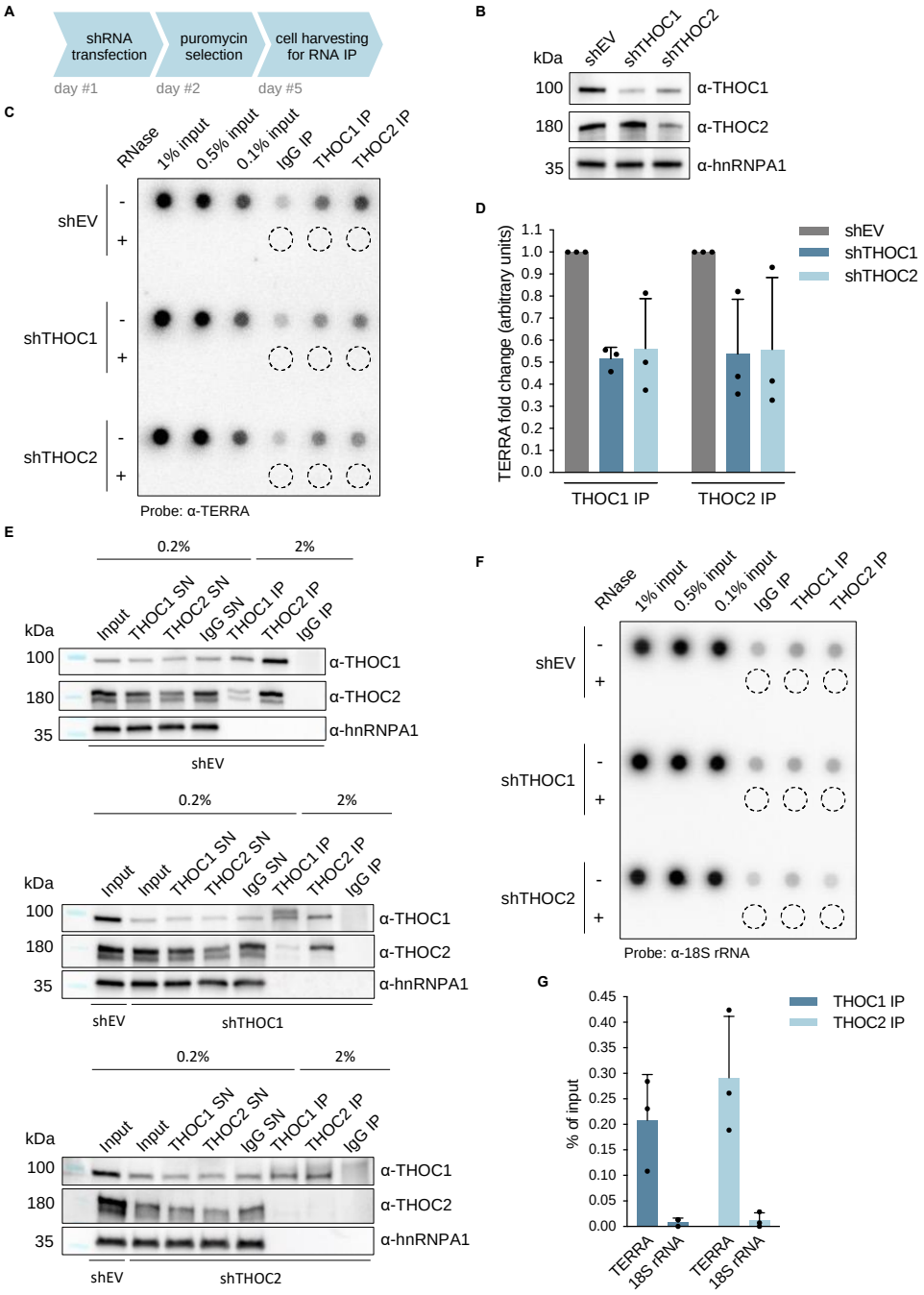


Supplementary figure 4. THOC associates with nucleoplasmic TERRA, but not the abundant 18S rRNA.

A Western blot was used to evaluate the efficiency of immunoprecipitation of THOC1 and THOC2 (left hand side panel), and hnRNPA1 (right hand side panel). Samples obtained from RNA-IP assays with HeLa cell clones with short (average 2.5 kb) were used in shown blots.

B RNA-IP samples from HeLa cells with long (average 10 kb, top) or short (average 2.5 kb, bottom) telomere length were analysed by RNA dot blot probed with a ^{32}P-radiolabelled probe complementary to the abundant 18S rRNA.

C Quantification of immunoprecipitated TERRA versus 18S rRNA (as in c), as percent of input. The signal corresponding to immunoprecipitates using IgG antibody was subtracted from corresponding test immunoprecipitation signals as background. Data represent mean ± s.d., from three independent biological replicates, for each HeLa cell clone with long (average 10 kb) or short (average 2.5 kb) telomere length.



Supplementary figure 5. THOC associates with nucleoplasmic TERRA, but not the abundant 18S rRNA, in Hek293E cells.

A Experimental setup: HEK293E cells were transfected with plasmids expressing shRNAs targeting THOC1 or THOC2 (or control shRNA designated shEV). Puromycin selection of transfected cells was performed for 72h before collecting cells for RNA-IP.

B Western blot analysis of depletion efficiency with indicated shRNAs in HEK293E cells, used in C-G.

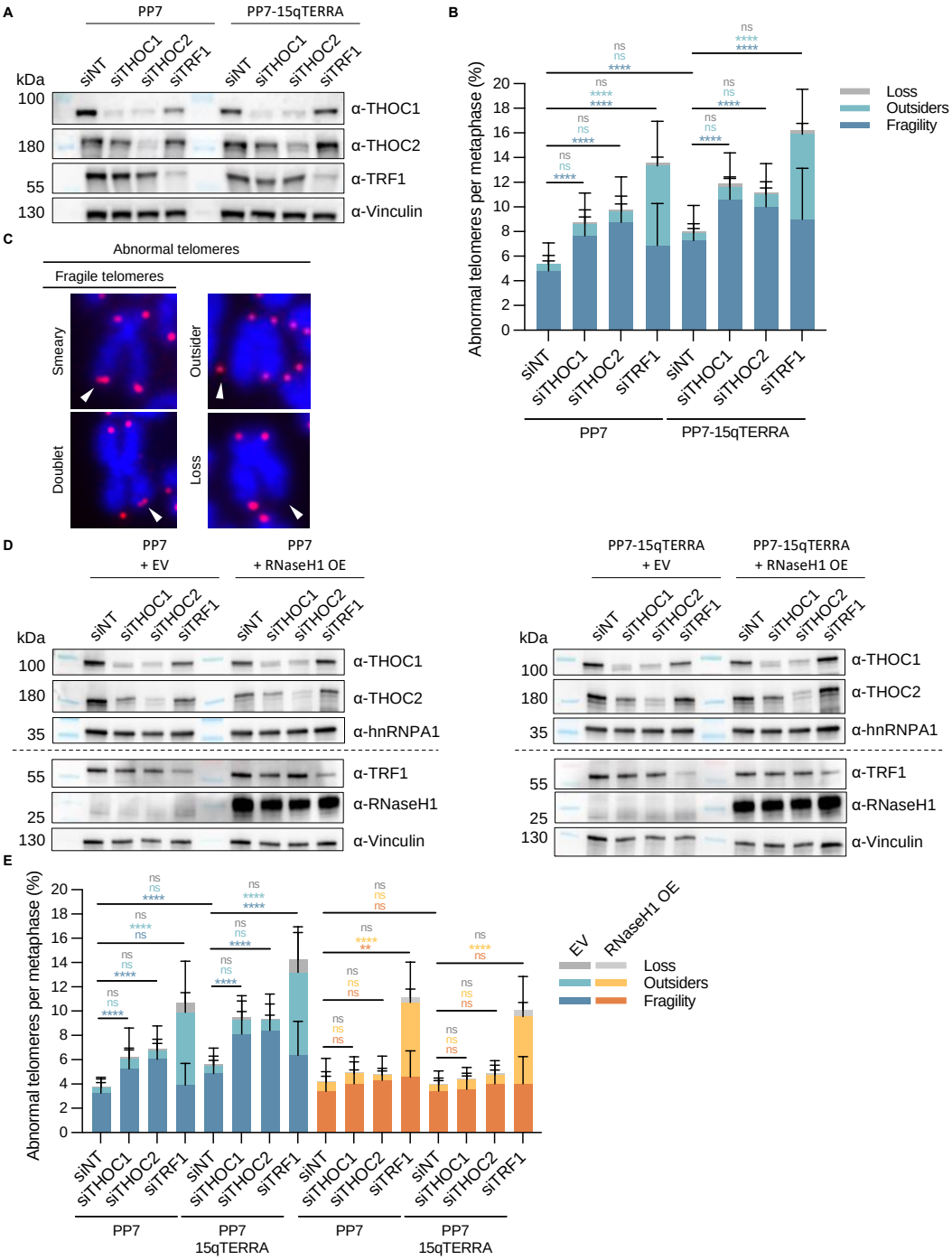
C Native RNA-immunoprecipitation (RNA-IP) assay using anti-THOC1 and anti-THOC2 antibodies was performed in extracts from HEK293E cells transfected with indicated shRNAs. Samples were analysed by RNA dot blot probed with a ^{32}P-radiolabelled $[CCCTAA]_3$ probe. Half of each IP sample was treated with RNase (DNase-free) as a control.

D Quantification of immunoprecipitated TERRA (as in C), as fold change over control shRNA (shEV). The signal corresponding to immunoprecipitates using IgG antibody was subtracted from corresponding test immunoprecipitation signals as background. Data represent mean ± s.d., from three independent biological replicates.

E Western blot was used to evaluate the efficiency of immunoprecipitation of THOC1 and THOC2 in samples obtained from RNA-IP assays with HEK293E cells transfected with indicated shRNAs.

F RNA-IP samples from HEK293E cells transfected with indicated shRNAs were analysed by RNA dot blot probed with a ^{32}P-radiolabelled probe complementary to the abundant 18S rRNA.

G Quantification of immunoprecipitated TERRA versus 18S rRNA (as in F), as percent of input. The signal corresponding to immunoprecipitates using IgG antibody was subtracted from corresponding test immunoprecipitation signals as background. Data represent mean ± s.d., from three independent biological replicates with HEK293E transfected with shEV.



Supplementary figure 6. THOC prevents telomeric fragility induced by TERRA R-loops, but has no detectable effect in telomere loss or outsider telomeres.

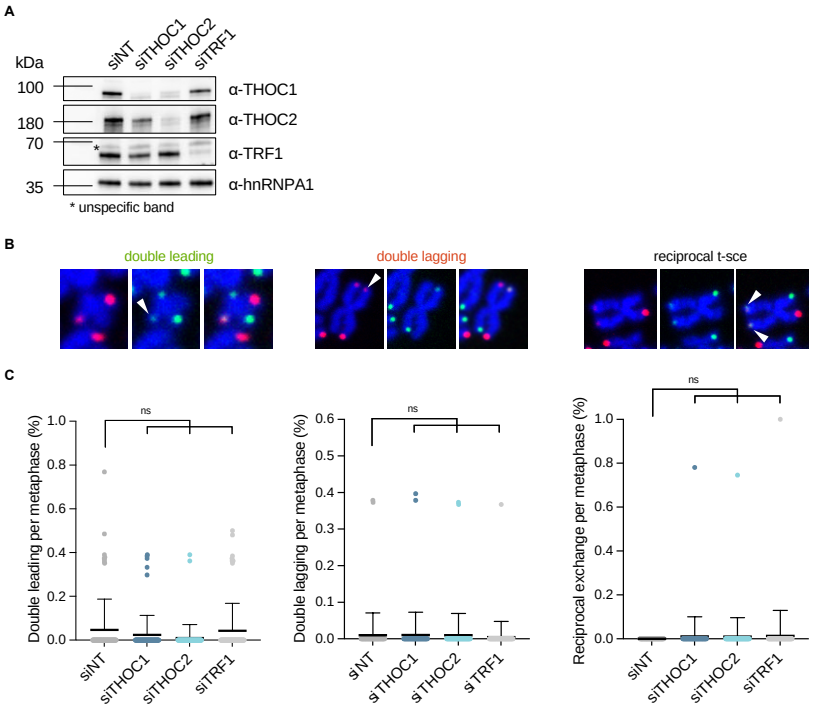
A Western blot analysis of indicated proteins in lysates collected from HeLa cells with 10 kb average telomere length, transfected with specified siRNAs and with plasmids for expression of PP7 or PP7-15qTERRA transcripts, used in B.

B Quantification of telomeric abnormalities – fragile telomeres, outsider telomeres and telomeric signal loss –, plotted as percentage of events per metaphase spread. At least 77 metaphase spreads were analysed per condition, across three independent biological replicates. Data are mean ± s.d.. Two-way analysis of variance (ANOVA) with Tukey's multiple comparisons test was applied: **** $P \leq 0.0001$, ns indicates non significance ($P > 0.05$).

C Telomeric FISH on metaphase spreads of HeLa cells stained with Cy3-$[CCCTAA]_3$ FISH probe (red) and DAPI (blue). White arrowheads indicate abnormal telomeres, either with a smeary FISH signal, multiple FISH signals, outsider telomere or loss of telomeric FISH signal in a single chromosome arm end.

D Western blot analysis of indicated proteins in lysates collected from lentivirus-transduced RNaseH1 OE or control HeLa cells with 10 kb average telomere length. Cells were transfected with indicated siRNAs and with plasmids for expression of PP7 (left panel) or PP7-15qTERRA transcripts (right panel), used in E.

E Quantification of telomeric abnormalities – fragile telomeres, outsider telomeres and telomeric signal loss – , plotted as percentage of events per metaphase spread. At least 71 metaphase spreads were analysed per condition, across three independent biological replicates. Data are mean ± s.d.. Two-way analysis of variance (ANOVA) with Tukey's multiple comparisons test was applied: **** P ≤ 0.0001, ** P ≤ 0.01, ns indicates non significance (P > 0.05).

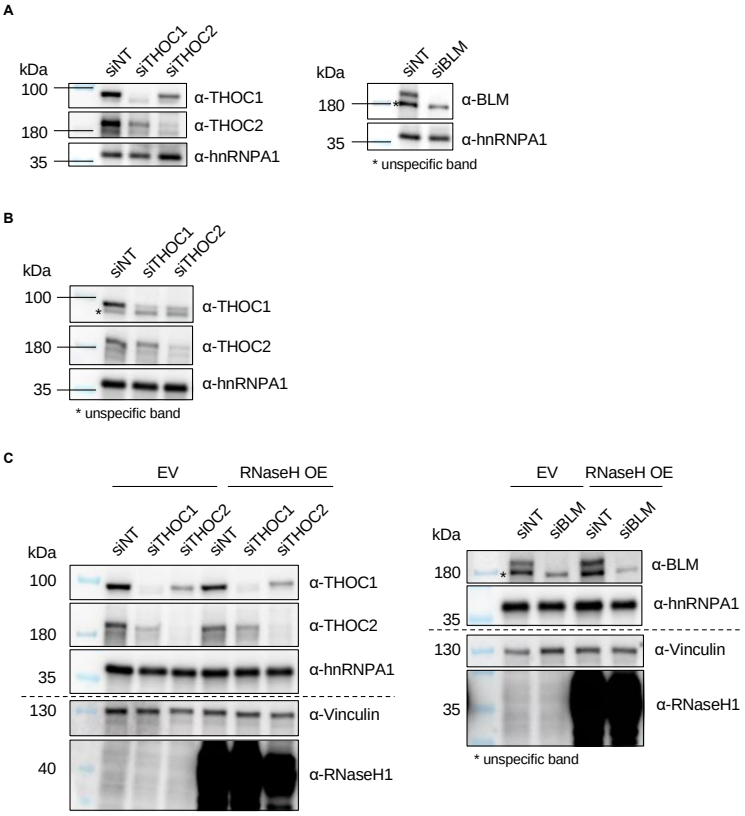


Supplementary figure 7. THOC depletion has no effect in telomeric sister chromatid exchange events in HeLa cells.

A Western blot analysis of indicated proteins in lysates collected from HeLa cells with 30 kb average telomere length, used in FISH and CO-FISH experiments, transfected with indicated siRNAs.

B Telomeric CO-FISH on metaphase spreads of HeLa cells with 30 kb average telomere length, stained with TYE563-TeloC LNA probe (red), FAM-TeloG LNA probe (green) and DAPI (blue). White arrowheads indicate double leading telomeres (left hand side), double lagging telomeres (centre) or reciprocal telomeric sister chromatid exchange (right hand side).

C Quantification of double leading telomeres (left hand side), double lagging telomeres (centre) or reciprocal telomeric sister chromatid exchange (right hand side), as percentage of events per metaphase spread (sum of lagging and leading strand telomeres, in HeLa cells with average 30 kb telomere length). At least 74 metaphase spreads were analysed per condition, across three independent biological replicates. Horizontal line and error bars represent mean ± s.d.. One-way analysis of variance (ANOVA) with Dunnett's multiple comparisons test was applied: ns indicates non significance (P > 0.05).



Supplementary figure 8. THOC depletion in U2OS and Saos2 cells.

A Western blot analysis of indicated proteins in lysates collected from U2OS cells, transfected with indicated siRNAs.

B Western blot analysis of indicated proteins in lysates collected from Saos2 cells, transfected with indicated siRNAs.

C Western blot analysis of indicated proteins in lysates collected from U2OS cells, transfected with indicated siRNAs and with empty-vector (EV) or RNaseH1 overexpressing (OE) constructs.

Supplementary table 1. siRNA sequences

siRNA	siRNA target sequences	Supplier: catalog number
siNT (non-targeting)	UAGCGACUAAACACAUCAA, UAAGGCUAUGAAGAGAUAC, AUGUAUUGGCCUGUAUUAG, AUGAACGUGAAUUGCUCAA	Dharmacon: D-001206-13
siTHOC1	GCAAUGAUCUCCUAAGAAG, GAGGAGAACAUGUAUAUUU, GGUCUUAACUUGCAGAGUC, UAACUUUCGUCGACACAUC	Dharmacon: M-019911-01
siTHOC2	GGAGAGACGUGUUCAAUAU, GAAAUAAGGCUGAUCAAUU, CAGCAUAGAUAUCGUCUGU, AAAGAACGCCGAAGUCUGA	Dharmacon: M-025006-01
siRNaseH1	GCGCAGAGCCGUAUGCAAA, GAGCUAAACAAUCGGAAGA, GCCAGGCCAUCCUUUAAAU, GACAUUCAGUGGAUGCAUG	Dharmacon: M-012595-00
siTRF1	CAAGAUAAACCUAGUGGUA, GGUGAUCCAAAUUCUCAUA, GGAAACUGGUCUAAAAUAC, GCCAGUUGAGAACGAUAUA	Dharmacon: M-010542-02
siDDX39B	ACAAUGAGCUCUUGGACUA, GGAUGAAUGUGAUAAGAUG, UUAGUGAGCUGCCUGAUGA, UAUGAGCGCUUCUCUAAAU	Dharmacon: M-003805-00
siBLM	GAGCACAUCUGUAAAUUAA, GAGAAACUCACUUCAAUAA, CAGGAUGGCUGUCAGGUUA, CUAAAUCUGUGGAGGGUUA	Dharmacon: M-007287-02

Supplementary table 2. Plasmids

Plasmid	Source	Notes
pMD2.G	Kind gift from D. Trono, EPFL	VSV-G envelope expressing plasmid, used for lentivirus production
pCMVR8.74	Kind gift from D. Trono, EPFL	Lentiviral packaging plasmid, used for lentivirus production
pTRE2_24xPP7_Puro	Feretzaki *et al*, 2020	Doxycycline inducible expression of PP7 RNA
pTRE2_24xPP7_15q subtel_90xTTAGGG_Puro	Feretzaki *et al*, 2020	Doxycycline inducible expression of PP7-15qTERRA RNA
pLenti_RNaseH1_Myc/His_Puro	Backbone: kind gift from E. Meylan, ELB	RNaseH1-Myc/His lentiviral transduction (doxycycline-inducible expression); backbone: pCW22_TREtight_Blast (selection marker exchanged with puromycin)
pcDNA6-RNAseH1-myc-His	RNaseH1 sequence amplified from construct kindly gifted by A. Straight, Stanford	Ectopic-expression of RNaseH1; backbone: pcDNA6
pshEV_Puro	Backbone: Oligoengine	Negative control shRNA; backbone: pSuper_Puro
pshTHOC1_Puro		Expression of shRNA for THOC1 depletion; backbone: pSuper_Puro
pshTHOC2_Puro		Expression of shRNA for THOC2 depletion; backbone: pSuper_Puro

Supplementary table 3. Antibodies

Antibody	Supplier: catalog number	Dilution/Amount	Application
α-THOC1	GeneTex: GTX118740	1:1,500	Western Blot
		6 µg	RNA Immunoprecipitation
		4 µg	Chromatin Immunoprecipitation
α-THOC2	Bethyl: A303-629A-T	1:10,000	Western Blot
α-THOC2	Abcam: ab129485	6 µg	RNA Immunoprecipitation
α-RNaseH1	GeneTex: GTX117624	1:1,000	Western Blot
α-TRF1	Santa Cruz: sc-6165-R	1:1,000	Western Blot
α-DDX39B	Proteintech: 14798-1-AP	1:10,000	Western Blot
α-BLM	Abcam: ab476	1:2,000	Western Blot
α-hnRNPA1	Santa Cruz: sc-32301	1:1,000	Western Blot
		6 µg	RNA Immunoprecipitation
α-Vinculin	Abcam: ab129002	1:10,000	Western Blot
α-HA	BioLegend: BGL901502	6 µg	RNA Immunoprecipitation
α-DNA:RNA hybrid (S9.6)	Kerafast: ENH001	0.1 µg/µg nucleic acids	DNA:RNA Immunoprecipitation
α-mouse IgG (H+L) HRP-conjugated	Promega: W4021	1:10,000	Western Blot
α-rabbit IgG (H+L) HRP-conjugated	Promega: W4011	1:10,000	Western Blot
α-GFP	*Homemade*	1:1,000	Immunofluorescence
α-rabbit IgG (H+L) CrossAdsorbed secondary antibody, Alexa Fluor 633	Thermo Fisher: A-21070	1:1,000	Immunofluorescence

Supplementary table 4. Oligonucleotide sequences

Purpose		Oligonucleotide sequence
1q subtelomere for DRIP-qPCR		CAGCGTCGCAACTCAAATG, CCCTCACCCTCCATGAGTAATA
10q subtelomere for DRIP-qPCR		GCATTCCTAATGCACACATGAC, TACCCGAACCTGAACCCTAA
13q subtelomere for DRIP-qPCR		GCACTTGAACCCTGCAATACAG, CCTGCGCACCGAGATTCT
Generation of pLenti_RNaseH1_ Myc/His_Puro	Replacing Blast with Puro selection marker by In-Fusion cloning	TGATAAGCTTGCCACATGACCGAGTACAAGCCCAC, AGTTAAGAATACGATTCAGGCACCGGGCTTG
	Insertion of RNaseH1 by In-Fusion cloning	GATCGCCTGGAGGTTAACATGAGCTGGCTTCTGTTCCT, ATCTTGGGTGGGTTAATTAATCAATGGTGATGGTGATGATGACCG
Generation of pshTHOC1_Puro by restriction cloning		GATCCCCGCAATGATCTCCTAAGAAGTTCAAGAGACTTCTTAGGA GATCATTGC TTTTTGGAAA, AGCTTTTCCAAAAAGCAATGATCTCCTAAGAAGTCTCTTGAACTTC TTAGGAGATCATTGCGGG
Generation of pshTHOC2_Puro by restriction cloning		GATCCCCGAAATAAGGCTGATCAATTTTCAAGAGAAATTGATCAG CCTTATTTCTTTTTGGAAA, AGCTTTTCCAAAAAGAAATAAGGCTGATCAATTTCTCTTGAAAATT GATCAGCCTTATTTCGGG

5.8 Materials and methods

Cell culture

HeLa, U2OS, Saos2 and Hek293T cell lines were cultured in Dulbecco's modified Eagle's medium (DMEM) (Gibco), supplemented with 10% fetal bovine serum (FBS) and 100 U/ml of penicillin/streptomycin. Cells were maintained in a controlled humidified atmosphere with 5% (v/v) CO_2, at 37°C. Generation of $rtTA^+$ $PCP\text{-}GFP^+$ HeLa cells for PP7-15qTERRA expression was previously described (Feretzaki *et al*, 2020). Generation of HeLa cells with ca. 30 kb average telomere length was previously described (Cristofari & Lingner, 2006). Suspension Hek293E cells were cultured in EX-CELL 293 Serum-Free Medium (Merck), supplemented with 4 mM GlutaMAX (Thermo Fisher Scientific) with 5% (v/v) CO_2, at 37°C, with constant agitation.

siRNA and plasmid transfection

180,000 cells/well were plated in 6 well plates one day before siRNA transfection. The following day, growth medium was replaced by antibiotic-free DMEM supplemented with 10% FBS, and cells were transfected with 20 pmol siRNA (see Supplementary table 1) using calcium phosphate transfection. The following day, antibiotic-supplemented growth medium was replenished, and cells were transfected with Lipofectamine 2000 (Invitrogen) using 2.5 µg plasmid DNA (see Supplementary table 2), following manufacturer's instructions. When doxycycline-inducible constructs were used, 1 µg/ml doxycycline was added to the medium 24 h after plasmid DNA transfection. Cells were harvested 72 h post-siRNA transfection and 48 h post-plasmid DNA transfection.

Lentivirus production and cell transduction

HEK293T cells were plated one day before transfection for lentivirus production. The following day, growth medium was replaced by Opti-MEM (Thermo Fisher Scientific). Cells were transfected with 1 µg pMD2.G plasmid, 3 µg pCMVR8.74 plasmid and 4 µg of lentivirus transfer plasmid (see Supplementary table 2), using Lipofectamine 2000 (Invitrogen), following manufacturer's instructions. Cells were kept in the transfection mix overnight, which was replaced by antibiotic-supplemented growth medium the next day. Supernatants containing lentiviral particles were collected on the two following days, and filtered through syringes with 0.45 µm filter units. Viruses were aliquoted and kept at -80°C. 1 ml of

lentivirus-containing medium was carefully added to each 6 cm dish containing recipient cells plated one day before transduction. Cells were split into 15 cm dishes, and selection was initiated with appropriate antibiotic 48 h post-transduction. Puromycin selection was done at 1 µg/ml final concentration.

Immunofluorescence and telomeric fluorescence *in situ* hybridization (FISH)

Cells were grown on round coverslips. Coverslips were washed twice in 1x PBS, fixed with 4% paraformaldehyde in PBS for 10 min at room temperature, and washed twice in 1x PBS. Coverslips were then incubated in detergent solution (0.1% Triton X-100, 0.02% SDS in 1x PBS) for 5 min, and incubated in 2% BSA dissolved in PBS for 10 min. Incubation with primary antibody (see Supplementary table 3) in blocking solution (10% normal goat serum in 2% BSA dissolved in PBS) was done overnight at 4°C, in a humidified chamber. The following day, three washes with 2% BSA dissolved in PBS were done for 4 min/wash, before incubation with secondary antibody (see Supplementary table 3) in blocking solution, for 30 min at room temperature, in a humidified chamber. Coverslips were then washed three times in 1x PBS, fixed with 4% paraformaldehyde in PBS for 5 min at room temperature, washed three times in 1x PBS, and dehydrated with 70%, 95% and 100% ethanol.

For FISH staining, hybridization was done with 100 nM Cy3-$[CCCTAA]_3$ PNA probe (PNA Bio) in 15 µl hybridization mix per coverslip (10 mM Tris pH 7.4, 70% formamide, 0.5% blocking reagent (Roche)), at 80°C for 3 min, followed by 3 h at room temperature, in a humidified chamber. Slides were washed twice with wash buffer 1 (10 mM Tris pH 7.4, 70% formamide) for 15 min/wash, and then washed three times with wash buffer 2 (0.1 M Tris pH 7.4, 0.15 M NaCl, 0.08% Tween-20) for 5 min/wash. DAPI was added to the second last wash at 0.1 µg/ml. Slides were dehydrated with 70%, 95% and 100% ethanol, air-dried and mounted with Vectashield.

Images were acquired with an Upright Zeiss Axioplan equipped with a 100x/1.40 oil objective.

Telomeric FISH on metaphase chromosomes

Cells were treated with 0.05 µg/ml demecolcine for 2 h before harvesting by trypsinization, resuspended in hypotonic solution (0.056 M KCl), and incubated at 37°C for 7 min. Cells were fixed in cold methanol:glacial acetic acid (3:1) solution overnight at 4°C. Fixed cells were dropped onto glass slides, incubated at 70°C for 1 min in a humidified oven, and air-dried overnight at room temperature. Slides were incubated in 4% formaldehyde in PBS for 5 min, washed three times in 1x PBS and dehydrated with 70%, 95% and 100% ethanol. Hybridization was done with 100 nM Cy3-$[CCCTAA]_3$ PNA probe (PNA Bio) in 70 µl hybridization mix (10 mM Tris pH 7.4, 70% formamide, 0.5% blocking reagent (Roche)), at 80°C for 3 min, followed by 3 h at room temperature, in a humidified chamber. Slides were washed twice with wash buffer 1 (10 mM Tris pH 7.4, 70% formamide) for 15 min/wash, and then washed three times with wash buffer 2 (0.1 M Tris pH 7.4, 0.15 M NaCl, 0.08% Tween-20) for 5 min/wash. DAPI was added to the second last wash at 0.1 µg/ml. Slides were dehydrated with 70%, 95% and 100% ethanol, air-dried and mounted with Vectashield.

Images were acquired with an Upright Zeiss Axioplan equipped with a 100x/1.40 oil objective, or with a Leica SP8 confocal microscope equipped with a 63x/1.40 oil objective and a DFC 7000 GT camera.

Chromosome Orientation (CO)-FISH on metaphase chromosomes

CO-FISH staining was performed as previously described (Lee *et al*, 2018), with some modifications. HeLa or U2OS cells were incubated with BrdU/BrdC (3:1 at a final concentration of 10 µM) for 15.5 or 17.5 h, and with 0.1 or 0.2 µg/ml demecolcine for 2 or 5 h before harvesting by trypsinization, respectively, resuspended in hypotonic solution (0.056 M KCl), and incubated at 37°C for 7 min. Cells were fixed in cold methanol:glacial acetic acid (3:1) solution overnight at 4°C. Fixed cells were dropped onto glass slides, incubated at 70°C for 1 min in a humidified oven, and air-dried overnight at room temperature. Slides were re-hydrated in 1x PBS for 5 min and treated with 250 µg/ml RNaseA (Promega) for 1 h at 37°C. Slides were washed in 1x PBS for 2 min and incubated with 10 µg/ml Hoechst 33258 (Invitrogen) in 1xPBS for 15 min at room temperature. Slides were then exposed to 365 nm UV light

using a Stratagene Stratalinker 1800 UV irradiator set to 5400 J, covered by minimal amount of 1x PBS. A brief wash in H_2O was done after irradiation, followed by treatment with 10 U/µl Exonuclease III (New England Biolabs) for 1 h at 37°C. Slides were incubated in 4% formaldehyde in PBS for 3 min, washed three times in 1x PBS, and dehydrated with 70%, 95% and 100% ethanol. Following overnight air-drying of slides, hybridizations were done sequentially using 0.22 µM TYE563- or 6-FAM-labeled LNA probes (Qiagen) (TeloC LNA probe: 5′ TYE563-CCC*TAACCC*TAACCC*TAA 3′; TeloG LNA probe: 5′ 6-FAM-T*TAGGGT*TAGGGT*TAGGG – where asterisks (*T) indicate LNA nucleotides) in 60 µl hybridization mix per slide (2x SSC, 30% formamide, 0.5% blocking reagent (Roche)), at room temperature for 2 h each. After each hybridization, slides were washed three times in 2x SCC for 3 min and dehydrated with 70%, 95% and 100% ethanol. After the last probe hybridization, DAPI was added to the second 2xSSC wash at 0.1 µg/ml. Slides were air-dried and mounted with ProLong Diamond Antifade mountant.

Images were acquired with a Leica SP8 confocal microscope equipped with a 63x/1.40 oil objective and a DFC 7000 GT camera.

Western blotting

Cells were collected by trypsinization, washed with 1x PBS and resuspended in 2x Laemmli buffer, to a final concentration of 10,000 cells/µl of Laemmli buffer. Samples were incubated at 95°C for 5 min. Proteins were separated on a 4-15% SDS-PAGE precast gel (Mini-PROTEAN TGX Gels, Bio-Rad) and transferred onto a 0.2 µm nitrocellulose membrane (Amersham). Membranes were blocked with blocking solution (3% BSA (w/v) in 1× PBS with 0.1% Tween-20) for 1 h at room temperature and incubated with primary antibodies (see Supplementary table 3) at 4°C overnight. Membranes were washed three times for 15 min/wash with 1× PBS with 0.1% Tween-20, and incubated with Horseradish Peroxidase-conjugated secondary antibodies (see Supplementary table 3) in blocking solution, for 1 h at room temperature. Membranes were then washed three times for 5 min/wash with 1× PBS with 0.1% Tween-20. ChemiGlow Chemiluminescence Substrate (Bio Techne) was used to develop the signal, which was detected by a Fusion FX imaging system (Vilber).

RNA isolation and RNA dot blot

RNA was isolated from 3 to 5 million cells with the NuceloSpin RNA kit (Macherey-Nagel). Two on-column and one in-solution rDNase digestions (Macherey-Nagel) were performed.

Purified RNA was digested with RNase (DNase-free, Roche), as a control. All samples were denatured at 65°C for 3 min and blotted onto a Hybond-XL membrane (Amersham) using a dot-blot apparatus (Bio-Rad). RNA was UV-crosslinked to the membrane. The membrane was then blocked in Church buffer (0.5 M NaHPO4, 1 mM EDTA, pH 8.0, 1% (w/v) BSA, 7% SDS) for at least 1 h at 50°C, and hybridized with a ^{32}P-radiolabeled telomeric probe in Church buffer at 50°C overnight. The membrane was washed twice in 2x SSC, 0.5% SDS and twice in 1x SSC, 0.5% SDS, for 15 min/wash. The membrane was exposed to a phosphorimager screen. Radioactive signal was detected with a Typhoon Biomolecular Imager (GE). After signal detection, the membrane was stripped by incubation with boiled 0.1x SCC, 1 % SDS at 55°C, three times for 30 min each, and blocked in Church buffer for at least 1 h at 55°C. A ^{32}P-radiolabeled 18S rRNA oligo probe was used for hybridization at 55°C overnight. The membrane was then treated as described for the telomeric probe.

RNA immunoprecipitation (RIP)

Cells were harvested by trypsinization (or collected by centrifugation in the case of suspension Hek293E cells), counted, washed with 1x PBS and placed on ice. Cells were then lysed in RLN buffer (50 mM Tris-HCl pH 8.0, 140 mM NaCl, 1.5 mM $MgCl_2$, 0.5% NP-40, 1 mM dithiothreitol (DTT)), supplemented with 400 U/ml RNasin Plus (Promega) and a protease inhibitor cocktail (cOmplete, Roche) (100 x10^6 cells/ml of RLN buffer). Lysates were homogenized with a Dounce homogenizer, incubated on a rotating wheel for 20 min at 4°C, and centrifuged at high-speed for 10 min. The supernatant was collected and pre-cleared with magnetic protein G Dynabeads (Thermo Fisher Scientific) on a rotating wheel for 1 h at 4°C (70 µl beads per 1 ml of extract). Pre-cleared extracts

equivalent to 50 x10^6 cells were used per immunoprecipitation with 6 µg of antibody (see Supplementary table 3), and incubated on a rotating wheel for 2 hours at 4°C. Input samples equivalent to 10% of each immunoprecipitation were also collected. 35 µl of magnetic protein G Dynabeads (Thermo Fisher Scientific) pre-blocked with yeast transfer RNA were then added to each immunoprecipitation sample and incubated on a rotating wheel overnight at 4°C. Samples were then washed at 4°C for 5 min/wash on a rotating wheel with RLN buffer supplemented with 6 mM EDTA pH 8.0, 0.5% NP-40 and 20 U/ml SUPERase In RNase inhibitor (Thermo Fisher Scientific). RNA was eluted from beads with 1500 rpm agitation, in 1% SDS, 5 mM EDTA pH 8.0 and 5 mM 2-mercaptoethanol, at 42°C for 30 min (for input and immunoprecipitation samples), followed by 65°C for 30 min (for immunoprecipitation samples only). The RNA was purified following the RNA clean-up protocol of the NucleoSpin RNA isolation kit (Macherey-Nagel) and eluted in 100 ul H2O. Half of each immunoprecipitation sample was digested with RNase (DNase-free, Roche)). Input samples were diluted appropriately to ensure that the amount of probed RNA in each immunoprecipitated sample lies within the linear dynamic range of the assay. RNA dot blot was done as described above.

Chromatin immunoprecipitation (ChIP)

Cells were harvested by trypsinization 72 h post-siRNA transfection, counted, and washed with cold 1x PBS. For Western blot analysis of depletions, ca. 0.5 x10^6 cells/condition were collected. For ChIP, 10 x10^6 cells/condition were crosslinked in 1 ml 1% methanol-free formaldehyde in PBS, for 15 min, on a rotating wheel, at room temperature. Formaldehyde was quenched by adding 250 mM Tris pH 8.0 in PBS, and incubating for 5 min, on a rotating wheel, at room temperature. Cells were then washed three times with cold PBS and kept at 4°C throughout the procedure until the crosslink reversal stage. Cells were washed once with 1 ml LB3 buffer (10 mM Tris pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% Na-Deoxycholate, 0.25% sodium lauroyl sarkosinate, supplemented with protease inhibitor cocktail (cOmplete, Roche)), resuspend in 1 ml LB3 buffer, transferred to sonication vials with AFA fiber (Covaris), and sonicated with a Focused-Ultrasonicator (E220, Covaris) (10% duty factor, 140 W power, 200 cycles per burst, for 20 min), to achieve fragments with <500 bp. Sonicated samples were centrifuged at 21,000 g for 15 min. Supernatants were collected and diluted 1:2 in IP dilution buffer (16.7 mM Tris pH 8.0, 1.1% Triton-X, 300 mM NaCl, 1.2 mM EDTA pH 8.0, supplemented with protease inhibitor cocktail (cOmplete, Roche)). Samples were pre-cleared with sepharose protein G beads (Cytiva) pre-blocked with yeast transfer RNA, on a rotating wheel for 1 h, at 4°C. Pre-cleared extracts equivalent to 2 x10^6 cells were used per immunoprecipitation with 4 µg of antibody (see Supplementary table 3), and 20 µl sepharose protein G beads pre-blocked with yeast transfer RNA, on a rotating wheel, at 4°C, overnight. Input samples equivalent to 10% of each immunoprecipitation were also collected. The following day, samples were washed at 4°C for 5 min/wash on a rotating wheel with buffer Wash 1 (0.1% SDS, 1% Triton, 2 mM EDTA pH 8.0, 20 mM Tris pH 8.0, 300 mM NaCl), buffer Wash 2 (0.1 % SDS, 1% Triton, 2 mM EDTA pH 8.0, 20 mM Tris pH 8.0, 500 mM NaCl), buffer Wash 3 (250 mM LiCl, 1% NP-40, 1% Na-deoxycholate, 1 mM EDTA pH 8.0, 10 mM Tris pH 8.0) and TE buffer (1 mM EDTA pH 8.0, 10 mM Tris pH 8.0). Washed beads and input samples were resuspended in crosslink reversal buffer (0.1% SDS, 0.1 M sodium bicarbonate, 0.5 mM EDTA pH 8.0, 20 mM Tris pH 8.0, supplemented with 10 µg/ml RNase (DNase-free (Roche)) at 65°C, on a rotating wheel, overnight. DNA was isolated with the NucleoSpin Gel and PCR Clean-up kit with NTB buffer (Macherey-Nagel) and eluted in 40 µl H_2O. Samples were then analyzed by dot blot (see below).

DNA:RNA immunoprecipitation (DRIP)

DRIP was performed as previously described (Glousker *et al*, 2022). Cells were harvested by trypsinization 72 h post-siRNA transfection and 48 h post-plasmid DNA transfection, counted and washed with cold 1x PBS. For Western blot analysis of depletions, ca. 0.5 x10^6 cells/condition were collected. For DRIP, 10 x10^6 cells/condition were resuspended in 175 µl of ice-cold RLN buffer (50 mM Tris-HCl pH 8.0, 140 mM NaCl, 1.5 mM $MgCl_2$, 0.5% NP-40, 1 mM dithiothreitol (DTT), and 100 U/ml RNasin Plus (Promega)), incubated on ice for 5 min, and centrifuged at 300 g for 2 min at 4°C. Nuclei

were brought to room temperature and lysed with 500 µl RA1 buffer (NucleoSpin RNA purification kit, Macherey-Nagel) containing 1% 2-mercaptoethanol, and homogenized with a syringe with a 0.9x40 mm needle. Nucleic acid extracts were then loaded in Phase Lock Gel heavy (5PRIME) tubes, mixed with 250 µl H_2O and 750 µl phenol-chloroform-isoamylalcohol (25:24:1) with a pH of 7.8 to 8.2, and centrifuged at 13,000 g for 5 min at room temperature. The aqueous phase was transferred into a new tube. 750 µl cold isopropanol and NaCl to 50 mM were added to the aqueous phase, mixed thoroughly, and incubated on ice for 30 min. Samples were centrifuged at 10,000 g for 30 min at 4°C to precipitate nucleic acids, followed by two washes with 70% cold ethanol. After air-drying, nucleic acids were dissolved in 130 µl of H2O, and sonicated with a Focused-Ultrasonicator (E220, Covaris) (10% duty factor, 140 W power, 200 cycles per burst, for 150 s, with an AFA intensifier), to achieve fragments with 100-300 bp. The concentration of fragmented nucleic acids was determined by spectrophotometry with a NanoDrop (ThermoFisher). Appropriate amount (see below) of nucleic acids was digested with 10 µl RNaseH (1 U/ µl, Roche) – as a negative control – or H_2O, in 15 µl RNaseH buffer (20 mM HEPES-KOH pH 7.5, 50 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT), in a total volume of 150 µl, and incubated at 37°C for 90 min. Digestion was stopped with 2 µl 0.5 M EDTA (pH 8.0) per sample. Samples were diluted 1:10 in DIP-1 buffer (10 mM HEPES-KOH pH 7.5, 275 mM NaCl, 0.1% Na-deoxycholate, 0.1% SDS, 1% Triton X-100) and pre-cleared with 40 µl of sepharose protein G beads (Cytiva) for 1 h, on a rotating wheel, at 4°C. 30 µg of diluted nucleic acids from HeLa cells or 12.5 µg of diluted nucleic acids from U2OS cells were used per immunoprecipitation with 3 µg of S9.6 antibody (Kerafast) or mouse IgG antibody (see Supplementary table 3), and 20 µl of sepharose protein G beads (Cytiva), and incubated on a rotating wheel at 4°C overnight. Nucleic acids equivalent to 1% of each immunoprecipitation sample were collected as input. The following day, samples were washed at 4°C for 5 min/wash on a rotating wheel with buffer DIP-2 (50 mM HEPES-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA pH 8.0, 1% Triton X-100, 0.1% Na-deoxycholate), buffer DIP-3 (50 mM HEPES-KOH pH 7.5, 500 mM NaCl, 1 mM EDTA pH 8.0, 1% Triton-X100, 0.1% Na-deoxycholate), buffer DIP-4 (10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0, 250 mM LiCl, 1% NP-40, 1% Na-deoxycholate), and TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0). Immunoprecipitation and input samples were resuspended overnight at 65°C with 100 µl elution buffer (20 mM Tris-HCl pH 8.0, 0.1% SDS, 0.1 M $NaHCO_3$, 0.5 mM EDTA pH 8.0) containing 10 µg/ml RNase (DNase-free (Roche)). DNA was isolated with the QIAquick PCR Purification kit (Qiagen) and eluted in 100 µl H_2O. Samples were then analysed by dot blot or qPCR (see below).

qPCR analysis of DRIP samples

Each qPCR reaction comprised 1 µl of purified DNA (immunoprecipitation and diluted input samples – as aforementioned), 5 µl Power SYBR Green PCR Master Mix (Thermo Fisher Scientific), 1 µM forward and reverse primers (see Supplementary table 4), and H_2O up to 10 µl total reaction volume. Each input or immunoprecipitation sample was run in technical duplicate. qPCR reactions were caried at 95°C for 10 min, followed by 95°C for 15 sec, and annealing and extension at 60°C for 1 min for 40 cycles in a QuantStudio 6 Flex Real-Time PCR system (Thermo Fisher Scientific). At least two serial dilutions (with dilution factors of 5 and 50) of each input sample were included, in order to perform a regression analysis to determine the equation of the standard curve of each input. Using each input equation, the corresponding S9.6 and IgG immunoprecipitations were calculated as percentage of input.

Dot blot analysis of CHIP and DRIP samples

Input samples (ChIP and DRIP) were diluted appropriately to ensure that the amount of probed DNA in each immunoprecipitated sample lies within the linear dynamic range of each assay. Purified DNA (diluted inputs and immunoprecipitated samples) was incubated at 95°C for 5 min and kept on ice. Samples were blotted onto a Hybond-XL membrane (Amersham) using a dot blot apparatus (Bio-Rad), and DNA was UV-crosslinked to the membrane. The membrane was denatured in 0.5 M NaOH, 1.5 M NaCl for 15 min on a shaker at room temperature, neutralized in 0.5 M Tris-Cl pH 7.0, 1.5 M NaCl for 10 min on a shaker at room temperature, and then blocked in Church buffer (0.5 M $NaHPO_4$, 1 mM EDTA, pH 8.0, 1% (w/v) BSA, 7% SDS) for at least 1 h at 65°C. The membrane was hybridized with a ^{32}P-

radiolabeled TeloC probe in Church buffer at 65°C overnight. Washes were done three times for 30 min/wash at 65°C, in 1x SSC, 0.5% SDS. The membrane was then exposed to a phosphorimager screen. Radioactive signal was detected with a Typhoon Biomolecular Imager (GE).

C-circle assay

The C-circle assay was performed as previously described (Henson et al, 2009). Briefly, genomic DNA was extracted with phenol-chloroform-isoamyl alcohol (25:24:1) and digested with HinfI and RsaI (NEB) in CutSmart buffer (NEB), overnight at 37°C. 30 ng of digested DNA were incubated with 7.5 U phi29 DNA polymerase (NEB) (or in its absence for phi29$^-$ control reaction), in the presence of dATP, dTTP and dGTP (1 mM each) at 30°C for 8 h, followed by heat-inactivation at 65°C for 20 min. Reaction products were blotted onto a Hybond-XL membrane (Amersham) using a dot blot apparatus (Bio-Rad), and DNA was UV-crosslinked to the membrane. The non-denatured membrane was blocked in Church buffer (0.5 M $NaHPO_4$, 1 mM EDTA, pH 8.0, 1% (w/v) BSA, 7% SDS) for at least 1 h at 42°C. The membrane was hybridized with a ^{32}P-radiolabeled TeloC oligo probe in Church buffer at 42°C overnight. Washes were done three times for 30 min/wash at 42°C, in 1x SSC, 0.5% SDS. The membrane was then exposed to a phosphorimager screen. Radioactive signal was detected with a Typhoon Biomolecular Imager (GE). For detection of the phi29$^-$ signal, the membrane was stripped and denatured as mentioned above, blocked in Church buffer for at least 1 h at 42°C, and re-probed with a ^{32}P-radiolabeled TeloC oligo probe in Church buffer at 42°C overnight. Washes and signal detection were done as described.

Software

Images were processed and analysed with ImageJ (2.0.0-rc-69/1.53k). Dot blots were analysed using Aida Image Analyzer (v. 5.1). Preparation of graphs and statistical analyses were performed using GraphPad Prism (v. 9.4.1 (458)). Illustrations were generated with BioRender.com.

5.9 Acknowledgements

We thank Marianna Feretzaki for sharing knowledge, Thomas Lunardi for sharing unpublished work and all members of the Lingner lab for discussions. We are also grateful to Ana Beatriz Domingues Silva and Rajika Arora (current and former members of the Azzalin lab, respectively) for help in establishing the CO-FISH protocol. We acknowledge the support provided by the Bioimaging and Optics Platform and the Gene Expression Core Facilities at the School of Life Sciences of EPFL.

5.10 Chapter references

Aguilera A & Klein HL (1990) HPR1, a Novel Yeast Gene That Prevents Intrachromosomal Excision Recombination, Shows Carboxy-Terminal Homology to the Saccharomyces cerevisiae TOP1 Gene. *MOL CELL BIOL* 10: 13

Ariel F, Lucero L, Christ A, Mammarella MF, Jegu T, Veluchamy A, Mariappan K, Latrasse D, Blein T, Liu C, *et al* (2020) R-Loop Mediated trans Action of the APOLO Long Noncoding RNA. *Molecular Cell* 77: 1055-1065.e4

Arnoult N, Van Beneden A & Decottignies A (2012) Telomere length regulates TERRA levels through increased trimethylation of telomeric H3K9 and HP1α. *Nat Struct Mol Biol* 19: 948–956

Arora R, Lee Y, Wischnewski H, Brun CM, Schwarz T & Azzalin CM (2014) RNaseH1 regulates TERRA-telomeric DNA hybrids and telomere maintenance in ALT tumour cells. *Nat Commun* 5: 5220

Azzalin CM, Reichenbach P, Khoriauli L, Giulotto E & Lingner J (2007) Telomeric Repeat–Containing RNA and RNA Surveillance Factors at Mammalian Chromosome Ends. *Science* 318: 798–801

Bailey SM, Goodwin EH, Meyne J & Cornforth MN (1996) CO-FISH reveals inversions associated with isochromosome formation. *Mutagenesis* 11: 139–144

Balk B, Maicher A, Dees M, Klermund J, Luke-Glaser S, Bender K & Luke B (2013) Telomeric RNA-DNA hybrids affect telomere-length dynamics and senescence. *Nat Struct Mol Biol* 20: 1199–1205

Boguslawski SJ, Smith DE, Michalak MA, Mickelson KE, Yehle CO, Patterson WL & Carrico RJ (1986) Characterization of monoclonal antibody to DNA · RNA and its application to immunodetection of hybrids. *Journal of Immunological Methods* 89: 123–130

Bryan TM, Englezou A, Dalla-Pozza L, Dunham MA & Reddel RR (1997) Evidence for an alternative mechanism for maintaining telomere length in human tumors and tumor-derived cell lines. *Nat Med* 3: 1271–1274

Carmody SR & Wente SR (2009) mRNA nuclear export at a glance. *Journal of Cell Science* 122: 1933–1937

Cerritelli SM & Crouch RJ (2009) Ribonuclease H: the enzymes in eukaryotes. *The FEBS Journal* 276: 1494–1505

Chávez S & Aguilera A (1997) The yeast *HPR1* gene has a functional role in transcriptional elongation that uncovers a novel source of genome instability. *Genes Dev* 11: 3459–3470

Chávez S, Beilharz T, Rondón AG, Erdjument-Bromage H, Tempst P, Svejstrup JQ, Lithgow T & Aguilera A (2000) A protein complex containing Tho2, Hpr1, Mft1 and a novel protein, Thp2, connects transcription elongation with mitotic recombination in Saccharomyces cerevisiae. *EMBO J* 19: 5824–5834

Chávez S, García-Rubio M, Prado F & Aguilera A (2001) Hpr1 is preferentially required for transcription of either long or G+C-rich DNA sequences in Saccharomyces cerevisiae. *Mol Cell Biol* 21: 7054–7064

Chen L, Zhang C, Ma W, Huang J, Zhao Y & Liu H (2022) METTL3-mediated m6A modification stabilizes TERRA and maintains telomere stability. *Nucleic Acids Research*: gkac1027

Cheng H, Dufu K, Lee C-S, Hsu JL, Dias A & Reed R (2006) Human mRNA Export Machinery Recruited to the 5′ End of mRNA. *Cell* 127: 1389–1400

Chi B, Wang Q, Wu G, Tan M, Wang L, Shi M, Chang X & Cheng H (2013) Aly and THO are required for assembly of the human TREX complex and association of TREX components with the spliced mRNA. *Nucleic Acids Research* 41: 1294–1306

Cristofari G & Lingner J (2006) Telomere length homeostasis requires that telomerase levels are limiting. *The EMBO Journal* 25: 565–574

Crossley MP, Bocek M & Cimprich KA (2019) R-Loops as Cellular Regulators and Genomic Threats. *Molecular Cell* 73: 398–411

Cusanelli E, Romero CAP & Chartrand P (2013) Telomeric Noncoding RNA TERRA Is Induced by Telomere Shortening to Nucleate Telomerase Molecules at Short Telomeres. *Molecular Cell* 51: 780–791

Domínguez-Sánchez MS, Barroso S, Gómez-González B, Luna R & Aguilera A (2011) Genome Instability and Transcription Elongation Impairment in Human Cells Depleted of THO/TREX. *PLoS Genet* 7: e1002386

Feretzaki M, Pospisilova M, Valador Fernandes R, Lunardi T, Krejci L & Lingner J (2020) RAD51-dependent recruitment of TERRA lncRNA to telomeres through R-loops. *Nature* 587: 303–308

Feretzaki M, Renck Nunes P & Lingner J (2019) Expression and differential regulation of human TERRA at several chromosome ends. *RNA* 25: 1470–1480

Fleckner J, Zhang M, Valcárcel J & Green MR (1997) U2AF65 recruits a novel human DEAD box protein required for the U2 snRNP-branchpoint interaction. *Genes Dev* 11: 1864–1872

Flynn RL, Cox KE, Jeitany M, Wakimoto H, Bryll AR, Ganem NJ, Bersani F, Pineda JR, Suvà ML, Benes CH, *et al* (2015) Alternative lengthening of telomeres renders cancer cells hypersensitive to ATR inhibitors. *Science* 347: 273–277

Gan W, Guan Z, Liu J, Gui T, Shen K, Manley JL & Li X (2011) R-loop-mediated genomic instability is caused by impairment of replication fork progression. *Genes Dev* 25: 2041–2056

García-Muse T & Aguilera A (2019) R Loops: From Physiological to Pathological Roles. *Cell* 179: 604–618

Glousker G, Briod A-S, Quadroni M & Lingner J (2020) Human shelterin protein POT1 prevents severe telomere instability induced by homology-directed DNA repair. *The EMBO Journal* 39: e104500

Glousker G, Fernandes RV, Feretzaki M & Lingner J (2022) Detection of TERRA R-Loops at Human Telomeres. In *R-Loops*, Aguilera A & Ruzov A (eds) pp 159–171. New York, NY: Springer US

Glousker G & Lingner J (2021) Challenging endings: How telomeres prevent fragility. *BioEssays* 43: 2100157

Graf M, Bonetti D, Lockhart A, Serhal K, Kellner V, Maicher A, Jolivet P, Teixeira MT & Luke B (2017) Telomere Length Determines TERRA and R-Loop Regulation through the Cell Cycle. *Cell* 170: 72-85.e14

Granotier C, Pennarun G, Riou L, Hoffschir F, Gauthier LR, De Cian A, Gomez D, Mandine E, Riou J-F, Mergny J-L, *et al* (2005) Preferential binding of a G-quadruplex ligand to human chromosome ends. *Nucleic Acids Res* 33: 4182–4190

Grolimund L, Aeby E, Hamelin R, Armand F, Chiappe D, Moniatte M & Lingner J (2013) A quantitative telomeric chromatin isolation protocol identifies different telomeric states. *Nat Commun* 4: 2848

Henson JD, Cao Y, Huschtscha LI, Chang AC, Au AYM, Pickett HA & Reddel RR (2009) DNA C-circles are specific and quantifiable markers of alternative-lengthening-of-telomeres activity. *Nat Biotechnol* 27: 1181–1185

Huertas P & Aguilera A (2003) Cotranscriptionally Formed DNA:RNA Hybrids Mediate Transcription Elongation Impairment and Transcription-Associated Recombination. *Molecular Cell* 12: 711–721

Jimeno S (2002) The yeast THO complex and mRNA export factors link RNA metabolism with transcription and genome instability. *The EMBO Journal* 21: 3526–3535

Kaminski N, Wondisford AR, Kwon Y, Lynskey ML, Bhargava R, Barroso-González J, García-Expósito L, He B, Xu M, Mellacheruvu D, *et al* (2022) RAD51AP1 regulates ALT-HDR through chromatin-directed homeostasis of TERRA. *Molecular Cell* 82: 4001-4017.e7

Köhler A & Hurt E (2007) Exporting RNA from the nucleus to the cytoplasm. *Nat Rev Mol Cell Biol* 8: 761–773

Lafuente-Barquero J, García-Rubio ML, Martin-Alonso MS, Gómez-González B & Aguilera A (2020) Harmful DNA:RNA hybrids are formed in cis and in a Rad51-independent manner. *eLife* 9: e56674

de Lange T (2005) Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes Dev* 19: 2100–2110

Larson DR, Zenklusen D, Wu B, Chao JA & Singer RH (2011) Real-Time Observation of Transcription Initiation and Elongation on an Endogenous Yeast Gene. *Science* 332: 475–478

Lee YW, Arora R, Wischnewski H & Azzalin CM (2018) TRF1 participates in chromosome end protection by averting TRF2-dependent telomeric R loops. *Nat Struct Mol Biol* 25: 147–153

Lesbirel S, Viphakone N, Parker M, Parker J, Heath C, Sudbery I & Wilson SA (2018) The m6A-methylase complex recruits TREX and regulates mRNA export. *Sci Rep* 8: 13827

Li Y, Wang X, Zhang X & Goodrich DW (2005) Human hHpr1/p84/Thoc1 Regulates Transcriptional Elongation and Physically Links RNA Polymerase II and RNA Processing Factors. *Mol Cell Biol* 25: 4023–4033

Lin C-YG, Näger AC, Lunardi T, Vančevska A, Lossaint G & Lingner J (2021) The human telomeric proteome during telomere replication. *Nucleic Acids Research* 49: 12119–12135

Londoño-Vallejo JA, Der-Sarkissian H, Cazes L, Bacchetti S & Reddel RR (2004) Alternative Lengthening of Telomeres Is Characterized by High Rates of Telomeric Exchange. *Cancer Research* 64: 2324–2327

Lovejoy CA, Li W, Reisenweber S, Thongthip S, Bruno J, Lange T de, De S, Petrini JHJ, Sung PA, 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 Genetics* 8: e1002772

Lu R, O'Rourke JJ, Sobinoff AP, Allen JAM, Nelson CB, Tomlinson CG, Lee M, Reddel RR, Deans AJ & Pickett HA (2019) The FANCM-BLM-TOP3A-RMI complex suppresses alternative lengthening of telomeres (ALT). *Nat Commun* 10: 2252

Majerska J, Feretzaki M, Glousker G & Lingner J (2018) Transformation-induced stress at telomeres is counteracted through changes in the telomeric proteome including SAMHD1. *Life Sci Alliance* 1: e201800121

Masuda S, Das R, Cheng H, Hurt E, Dorman N & Reed R (2005) Recruitment of the human TREX complex to mRNA during splicing. *Genes Dev* 19: 1512–1517

Merz C, Urlaub H, Will CL & Lührmann R (2007) Protein composition of human mRNPs spliced in vitro and differential requirements for mRNP protein recruitment. *RNA* 13: 116–128

Nergadze SG, Farnung BO, Wischnewski H, Khoriauli L, Vitelli V, Chawla R, Giulotto E & Azzalin CM (2009) CpG-island promoters drive transcription of human telomeres. *RNA* 15: 2186–2194

O'Sullivan RJ, Arnoult N, Lackner DH, Oganesian L, Haggblom C, Corpet A, Almouzni G & Karlseder J (2014) Rapid induction of alternative lengthening of telomeres by depletion of the histone chaperone ASF1. *Nat Struct Mol Biol* 21: 167–174

Pérez-Calero C, Bayona-Feliu A, Xue X, Barroso SI, Muñoz S, González-Basallote VM, Sung P & Aguilera A (2020) UAP56/DDX39B is a major cotranscriptional RNA–DNA helicase that unwinds harmful R loops genome-wide. *Genes Dev* 34: 898–912

Petti E, Buemi V, Zappone A, Schillaci O, Broccia PV, Dinami R, Matteoni S, Benetti R & Schoeftner S (2019) SFPQ and NONO suppress RNA:DNA-hybrid-related telomere instability. *Nat Commun* 10: 1001

Pfeiffer V, Crittin J, Grolimund L & Lingner J (2013) The THO complex component Thp2 counteracts telomeric R-loops and telomere shortening. *EMBO J* 32: 2861–2871

Piruat JI & Aguilera A (1998) A novel yeast gene, THO2, is involved in RNA pol II transcription and provides new evidence for transcriptional elongation-associated recombination. *The EMBO Journal* 17: 4859–4872

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

Porro A, Feuerhahn S, Reichenbach P & 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

Pühringer T, Hohmann U, Fin L, Pacheco-Fiallos B, Schellhaas U, Brennecke J & Plaschka C (2020) Structure of the human core transcription-export complex reveals a hub for multivalent interactions. *eLife* 9: e61503

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

Redon S, Zemp I & Lingner J (2013) A three-state model for the regulation of telomerase by TERRA and hnRNPA1. *Nucleic Acids Research* 41: 9117–9128

Sagie S, Toubiana S, Hartono SR, Katzir H, Tzur-Gilat A, Havazelet S, Francastel C, Velasco G, Chédin F & Selig S (2017) Telomeres in ICF syndrome cells are vulnerable to DNA damage due to elevated DNA:RNA hybrids. *Nat Commun* 8: 14015

Sakellariou D, Bak ST, Isik E, Barroso SI, Porro A, Aguilera A, Bartek J, Janscak P & Peña-Diaz J (2022) MutSβ regulates G4-associated telomeric R-loops to maintain telomere integrity in ALT cancer cells. *Cell Reports* 39: 110602

Salas-Armenteros I, Pérez-Calero C, Bayona-Feliu A, Tumini E, Luna R & Aguilera A (2017) Human THO–Sin3A interaction reveals new mechanisms to prevent R-loops that cause genome instability. *EMBO J* 36: 3532–3547

Sarkar J, Wan B, Yin J, Vallabhaneni H, Horvath K, Kulikowicz T, Bohr VA, Zhang Y, Lei M & Liu Y (2015) SLX4 contributes to telomere preservation and regulated processing of telomeric joint molecule intermediates. *Nucleic Acids Res* 43: 5912–5923

Sfeir A, Kosiyatrakul ST, Hockemeyer D, MacRae SL, Karlseder J, Schildkraut CL & de Lange T (2009) Mammalian Telomeres Resemble Fragile Sites and Require TRF1 for Efficient Replication. *Cell* 138: 90–103

Silva B, Arora R & Azzalin CM (2022) The alternative lengthening of telomeres mechanism jeopardizes telomere integrity if not properly restricted. *Proceedings of the National Academy of Sciences* 119: e2208669119

Silva B, Arora R, Bione S & Azzalin CM (2021) TERRA transcription destabilizes telomere integrity to initiate break-induced replication in human ALT cells. *Nat Commun* 12: 3760

Silva B, Pentz R, Figueira AM, Arora R, Lee YW, Hodson C, Wischnewski H, Deans AJ & Azzalin CM (2019) FANCM limits ALT activity by restricting telomeric replication stress induced by deregulated BLM and R-loops. *Nat Commun* 10: 2253

Sobinoff AP, Allen JA, Neumann AA, Yang SF, Walsh ME, Henson JD, Reddel RR & Pickett HA (2017) BLM and SLX4 play opposing roles in recombination-dependent replication at human telomeres. *EMBO J* 36: 2907–2919

Tardat M & Déjardin J (2018) Telomere chromatin establishment and its maintenance during mammalian development. *Chromosoma* 127: 3–18

Teasley DC, Parajuli S, Nguyen M, Moore HR, Alspach E, Lock YJ, Honaker Y, Saharia A, Piwnica-Worms H & Stewart SA (2015) Flap Endonuclease 1 Limits Telomere Fragility on the Leading Strand. *Journal of Biological Chemistry* 290: 15133–15145

Valador Fernandes R, Feretzaki M & Lingner J (2021) The makings of TERRA R-loops at chromosome ends. *Cell Cycle* 20: 1745–1759

Vannier J-B, Pavicic-Kaltenbrunner V, Petalcorin MIR, Ding H & Boulton SJ (2012) RTEL1 Dismantles T Loops and Counteracts Telomeric G4-DNA to Maintain Telomere Integrity. *Cell* 149: 795–806

Wahba L, Gore SK & Koshland D (2013) The homologous recombination machinery modulates the formation of RNA–DNA hybrids and associated chromosome instability. *eLife* 2: e00505

Wellinger RE, Prado F & Aguilera A (2006) Replication Fork Progression Is Impaired by Transcription in Hyperrecombinant Yeast Cells Lacking a Functional THO Complex. *Mol Cell Biol* 26: 3327–3334

Yadav T, Zhang J-M, Ouyang J, Leung W, Simoneau A & Zou L (2022) TERRA and RAD51AP1 promote alternative lengthening of telomeres through an R- to D-loop switch. *Molecular Cell* 82: 3985-4000.e4

Yang Z, Sharma K & de Lange T (2022) TRF1 uses a noncanonical function of TFIIH to promote telomere replication. *Genes Dev*: genesdev;gad.349975.122v1

Yang Z, Takai KK, Lovejoy CA & Lange T de (2020) Break-induced replication promotes fragile telomere formation. *Genes Dev* 34: 1392–1405

Yu T-Y, Wang C-Y & Lin J-J (2012) Depleting Components of the THO Complex Causes Increased Telomere Length by Reducing the Expression of the Telomere-Associated Protein Rif1p. *PLoS ONE* 7: e33498

Zhang M & Green MR (2001) Identification and characterization of yUAP/Sub2p, a yeast homolog of the essential human pre-mRNA splicing factor hUAP56. *Genes Dev* 15: 30–35

Zhou Z, Luo M, Straesser K, Katahira J, Hurt E & Reed R (2000) The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans. *Nature* 407: 401–405

Zimmermann M, Kibe T, Kabir S & Lange T de (2014) TRF1 negotiates TTAGGG repeat-associated replication problems by recruiting the BLM helicase and the TPP1/POT1 repressor of ATR signaling. *Genes Dev* 28: 2477–2491

Chapter 6 Conclusion and future perspectives

The telomeric RNA TERRA is a crucial telomere component, involved in the regulation of telomere stability and cellular lifespan. TERRA is capable of directly hybridizing telomeric DNA, forming R-loops – which consist of a three-stranded nucleic acid structure, comprising a DNA:RNA hybrid and a displaced DNA strand. The abundance of TERRA R-loops has previously been shown to have a major impact in telomeric integrity, underscoring the relevance of a comprehensive investigation into their formation and accumulation at chromosome ends, and the possible outcomes of their dysregulation.

In chapter 2 a reporter system was developed for the study of ectopically-expressed TERRA-like RNA molecules. This system enables the detailed analysis of the association of TERRA with telomeres. It revealed that the recruitment of TERRA to chromosome ends can occur post-transcriptionally *in trans*, requiring the UUAGGG repetitive tract of TERRA. In addition, the work presented in chapter 2 demonstrated that TERRA R-loops form preferentially at short telomeres and are mediated by the DNA recombinase RAD51, which also has the ability to directly bind TERRA RNA. It will be of great interest to dedicate efforts in the identification and characterization of the requirements of a putative RAD51-dependent general mechanism that facilitates the association of other long non-coding RNAs to chromatin. Furthermore, adapting the PP7-fused TERRA system to live cell imaging will allow to describe TERRA nuclear dynamics.

The work presented in chapter 3 is based on the findings described in chapter 2. It theorizes on the consequences of RAD51-mediated R-loop formation and accumulation at chromosome ends, namely in the stimulation of homologous recombination at telomeres. A detailed analysis of the possible involvement of other factors implicated in homology-directed repair in TERRA R-loop regulation would be of great significance in the study of telomere maintenance.

The detection of TERRA R-loops is central to study the formation and impact of these structures in telomere homeostasis. Different techniques have previously been proposed for the detection of R-loops formed across the genome (reviewed in (García-Muse & Aguilera, 2019)). A straight-forward method that was primarily used in the study of DNA:RNA hybrids is the analysis of sensitivity of isolated nucleic acids to RNaseH (Huertas & Aguilera, 2003). Other common approaches include the detection of hybrids with inactivated RNaseH or the RNAseH1 hybrid-binding domain, followed by IF or ChIP (Ginno *et al*, 2012; Bhatia *et al*, 2014; Chen *et al*, 2017). The DNA:RNA hybrid-recognizing S9.6 antibody (Boguslawski *et al*, 1986) has also frequently been employed in IF and immunoprecipitation experiments. Of note, the S9.6 antibody is able to recognize double stranded RNAs, therefore treatment with RNaseH is necessary to determine the specificity of the detected structures (Smolka *et al*, 2021). Furthermore, sodium bisulfite treatment or activation-induced cytidine deaminase (AID)-derived DNA mutagenesis followed by sequencing have been used for the mapping of R-loops (Yu *et al*, 2003; Gómez-González & Aguilera, 2007). The protocol presented in chapter 4 provides a practical tool for the methodical and controlled assessment of R-loops formed at telomeres, using the S9.6 antibody in an immunoprecipitation-based approach.

In chapter 5 we investigate the roles of the THO complex at human telomeres. We demonstrate that THOC contributes to the negative regulation of TERRA R-loops. While this effect was better perceived using U2OS ALT cells, a subtle non-significant increase in telomeric hybrids was observed in telomerase-positive HeLa cells, transfected with siRNAs for THOC depletion. Since full deletion of THOC components could not be employed, a partial down regulation may still allow for THOC to fulfil its function, justifying the tenuous effect observed. In addition, the possibility of THOC subcomplexes to act as functional units has not yet been fully elucidated. Nevertheless, depletion of THOC1 caused a reduction in THOC2 protein levels and vice versa, which is consistent with an interdependency of the multiple THOC/TREX subunits to sustain THOC/TREX complexes (Chi *et al*, 2013). Furthermore, considering the numerous proteins which have previously been identified as regulators of TERRA R-loops and the detrimental effects identified thus far as an outcome of the accumulation of TERRA R-loops (see introduction), it is conceivable that different redundant pathways are involved in their suppression at chromosome ends.

Therefore, co-depletion of THOC subunits with other factors implicated in the control of TERRA R-loops may allow to better manipulate the levels of these structures and thoroughly assess their consequences. Of note, previous DRIP experiments performed in the Lingner lab using shRNAs to deplete THOC1 in telomerase-positive cells have failed to detect a significant change in telomeric R-loops (Grolimund, 2013). It should be noted that prolonged shRNA-mediated knock down of THOC1 – 5 days – yielded a very striking percentage of cell death, which may have compromised the assessment of any telomeric phenotype, as it may select for cells undergoing partial rescue of the effects derived from loss of THOC (Grolimund, 2013).

Resorting to RNA immunoprecipitation experiments, we demonstrate that THOC associates with nucleoplasmic TERRA. In addition, ongoing *in vitro* experiments have validated the binding of THOC1 to a TERRA-like UUAGGG-containing oligonucleotide (unpublished data acquired by the Lingner lab member Thomas Lunardi). It remains unclear whether the presence of THOC at human telomeres is mediated by TERRA RNA (directly hybridizing telomeric DNA or interacting with other telomeric proteins). Previous QTIP experiments – further validated by ChIP – have readily detected THOC subunits at telomeres (Grolimund *et al*, 2013). However, the use of an alternative protocol to determine the telomeric chromatin composition – termed PICh – has not identified THOC components (Déjardin & Kingston, 2009). Importantly, in contrast to QTIP, the PICh method includes an RNase digestion step, which may result in the loss of TERRA-mediated telomere-interacting factors – as may be the case for THOC components. A modified ChIP protocol including an RNase treatment (including RNaseH which would degrade TERRA molecules engaged in R-loops) could elucidate this point.

The exact mechanism through which THOC counteracts the accumulation of R-loops at telomeres is not yet clearly disclosed. Previously proposed models for the function of THOC in the modulation of R-loops formed across the genome suggest an interplay between THOC and chromatin modifiers, leading to decreased chromatin condensation, and thus increased propensity for DNA:RNA hybrid formation. However, analysis of AcH3 cellular levels in lysates from cells depleted of THOC1/2 has not succeeded in reproducing the increase in global histone acetylation levels demonstrated by Salas-Armenteros and colleagues (data not shown) (Salas-Armenteros *et al*, 2017). Additionally, the use of the histone acetyltransferase inhibitor anacardic acid has failed to rescue the increase in PP7-15qTERRA colocalization with telomeres promoted by loss of THOC (data not shown). Therefore, further experiments will be required to determine whether the modulation of chromatin accessibility is required for THOC to restrain R-loop formation at telomeres. Consistently with a model depicting THOC/TREX as a factor involved in the resolution of R-loops (as previously proposed (Pérez-Calero *et al*, 2020), depletion of RNaseH1 – which leads to an accumulation of TERRA R-loops – led to an elevated occupancy of telomeres by THOC subunit 1. Moreover, a QTIP experiment aimed at identifying the telomeric proteome under menadione treatment – which generates reactive oxygen species and thus induces oxidative stress – has detected an enrichment of THOC/TREX components at telomeres, comparing to untreated cells (unpublished data acquired by the Lingner lab member Trang Nguyen). Interestingly, menadione treatment correlates with an increase of TERRA R-loops (unpublished data acquired by the Lingner lab member Trang Nguyen). This observation is consistent with a recruitment of THOC to telomeres to preserve telomere integrity when R-loops pose a threat. *In vitro* experiments may assist in the elucidation of the detailed mechanism involving THOC in the management of telomeric R-loops. Additionally, understanding a putative role of THOC in the response to oxidative damage could open new lines of investigation for therapeutic purposes.

Employing telomeric FISH and CO-FISH, we have demonstrated that loss of THOC induces R-loop-dependent telomere fragility – a distinct indicator of replication defects at telomeres. Moreover, the fragility promoted by THOC depletion was observed not only at telomeres replicated by leading strand synthesis – as often reported upon dysregulation of factors involved in counteracting TERRA R-loops – but also, (to a lesser extent) at telomeres replicated by lagging strand synthesis. While the accumulation of telomeric R-loops has previously been correlated with leading strand fragility, the occurrence of lagging strand fragility under depletion of THOC was unexpected. Displacement of the G-rich telomeric strand upon the formation of R-loops may fuel the formation and stabilization of G4s, which in turn

may impair lagging strand replication. In addition, a direct role of THOC in the regulation of G4s has not yet been explored. Examining whether a build-up of telomeric G4 structures occurs at chromosome ends when THOC subunits are depleted – possibly resorting to G4-specific antibodies – may uncover the origin of the observed increase in lagging strand fragility.

Telomeric CO-FISH staining has also revealed that depletion of THOC subunits elevates the frequency of sister chromatid exchange events quantified specifically in U2OS ALT cells – while it is not sufficient to trigger this ALT hallmark in telomerase-positive HeLa cells. It would be interesting to explore if this effect of THOC depletion occurs in other ALT cell lines, and whether it depends on the formation of R-loops at telomeres.

Finally, we found that the THO complex counteracts C-circle accumulation in ALT cells, partially by regulating R-loops. While C-circle formation does not seem to require telomeric R-loops, our data reinforces the notion that R-loops can contribute/mediate to the genesis of C-circles. Given that the generation of C-circles and the relevance of its presence in the cell is not fully understood, it will be interesting to understand whether R-loops are involved in C-circle formation in a direct manner – being excised, circularized and ligated –, or indirect way – by stimulating break-induced replication at telomeres.

Future experiments intended at evaluating whether other ALT-characteristic features are amplified when THOC is depleted in ALT cells – such as telomere extension by BIR in APBs – would further contribute to the characterization of the function of THOC/TREX components and TERRA in the maintenance of human telomeres.

References

Aguilera A (2005) mRNA processing and genomic instability. *Nat Struct Mol Biol* 12: 737–738

Aguilera A & García-Muse T (2012) R Loops: From Transcription Byproducts to Threats to Genome Stability. *Mol Cell* 46: 115–124

Aguilera A & Klein HL (1988) Genetic control of intrachromosomal recombination in Saccharomyces cerevisiae. I. Isolation and genetic characterization of hyper-recombination mutations. *Genetics* 119: 779–790

Aguilera A & Klein HL (1990) HPR1, a Novel Yeast Gene That Prevents Intrachromosomal Excision Recombination, Shows Carboxy-Terminal Homology to the Saccharomyces cerevisiae TOP1 Gene. *MOL CELL BIOL* 10: 13

d'Alcontres MS, Palacios JA, Mejias D & Blasco MA (2014) TopoIIα prevents telomere fragility and formation of ultra thin DNA bridges during mitosis through TRF1-dependent binding to telomeres. *Cell Cycle Georget Tex* 13: 1463–1481

Ariel F, Lucero L, Christ A, Mammarella MF, Jegu T, Veluchamy A, Mariappan K, Latrasse D, Blein T, Liu C, *et al* (2020) R-Loop Mediated trans Action of the APOLO Long Noncoding RNA. *Mol Cell* 77: 1055-1065.e4

Arnoult N, Van Beneden A & Decottignies A (2012) Telomere length regulates TERRA levels through increased trimethylation of telomeric H3K9 and HP1α. *Nat Struct Mol Biol* 19: 948–956

Arora R, Lee Y, Wischnewski H, Brun CM, Schwarz T & Azzalin CM (2014) RNaseH1 regulates TERRA-telomeric DNA hybrids and telomere maintenance in ALT tumour cells. *Nat Commun* 5: 5220

Azzalin CM, Reichenbach P, Khoriauli L, Giulotto E & 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 & Azzalin CM (2012) The telomeric transcriptome of Schizosaccharomyces pombe. *Nucleic Acids Res* 40: 2995–3005

Balk B, Maicher A, Dees M, Klermund J, Luke-Glaser S, Bender K & Luke B (2013) Telomeric RNA-DNA hybrids affect telomere-length dynamics and senescence. *Nat Struct Mol Biol* 20: 1199–1205

Baumann P & Cech TR (2001) Pot1, the putative telomere end-binding protein in fission yeast and humans. *Science* 292: 1171–1175

Baur JA, Zou Y, Shay JW & Wright WE (2001) Telomere Position Effect in Human Cells. *Science* 292: 2075–2077

Bentin T, Cherny D, Larsen HJ & Nielsen PE (2005) Transcription arrest caused by long nascent RNA chains. *Biochim Biophys Acta BBA - Gene Struct Expr* 1727: 97–105

Bermejo R, Doksani Y, Capra T, Katou Y-M, Tanaka H, Shirahige K & Foiani M (2007) Top1- and Top2-mediated topological transitions at replication forks ensure fork progression and stability and prevent DNA damage checkpoint activation. *Genes Dev* 21: 1921–1936

Bhatia V, Barroso SI, García-Rubio ML, Tumini E, Herrera-Moyano E & Aguilera A (2014) BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. *Nature* 511: 362–365

Bilaud T, Brun C, Ancelin K, Koering CE, Laroche T & Gilson E (1997) Telomeric localization of TRF2, a novel human telobox protein. *Nat Genet* 17: 236–239

Blackburn EH & Gall JG (1978) A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in Tetrahymena. *J Mol Biol* 120: 33–53

Blasco MA (2007) The epigenetic regulation of mammalian telomeres. *Nat Rev Genet* 8: 299–309

Boguslawski SJ, Smith DE, Michalak MA, Mickelson KE, Yehle CO, Patterson WL & Carrico RJ (1986) Characterization of monoclonal antibody to DNA · RNA and its application to immunodetection of hybrids. *J Immunol Methods* 89: 123–130

Brickner JR, Garzon JL & Cimprich KA (2022) Walking a tightrope: The complex balancing act of R-loops in genome stability. *Mol Cell* 82: 2267–2297

Broccoli D, Smogorzewska A, Chong L & de Lange T (1997) Human telomeres contain two distinct Myb-related proteins, TRF1 and TRF2. *Nat Genet* 17: 231–235

Bryan TM, Englezou A, Gupta J, Bacchetti S & Reddel RR (1995) Telomere elongation in immortal human cells without detectable telomerase activity. *EMBO J* 14: 4240–4248

Cai S, Bai Y, Wang H, Zhao Z, Ding X, Zhang H, Zhang X, Liu Y, Jia Y, Li Y, *et al* (2020) Knockdown of THOC1 reduces the proliferation of hepatocellular carcinoma and increases the sensitivity to cisplatin. *J Exp Clin Cancer Res* 39: 135

Carmody SR & Wente SR (2009) mRNA nuclear export at a glance. *J Cell Sci* 122: 1933–1937

Celli GB & de Lange T (2005) DNA processing is not required for ATM-mediated telomere damage response after TRF2 deletion. *Nat Cell Biol* 7: 712–718

Cerritelli SM & Crouch RJ (2009) Ribonuclease H: the enzymes in eukaryotes: Ribonucleases H of eukaryotes. *FEBS J* 276: 1494–1505

Chang M, French-Cornay D, Fan H, Klein H, Denis CL & Jaehning JA (1999) A Complex Containing RNA Polymerase II, Paf1p, Cdc73p, Hpr1p, and Ccr4p Plays a Role in Protein Kinase C Signaling. *Mol Cell Biol* 19: 1056–1067

Chávez S & Aguilera A (1997) The yeast *HPR1* gene has a functional role in transcriptional elongation that uncovers a novel source of genome instability. *Genes Dev* 11: 3459–3470

Chávez S, Beilharz T, Rondón AG, Erdjument-Bromage H, Tempst P, Svejstrup JQ, Lithgow T & Aguilera A (2000) A protein complex containing Tho2, Hpr1, Mft1 and a novel protein, Thp2, connects transcription elongation with mitotic recombination in Saccharomyces cerevisiae. *EMBO J* 19: 5824–5834

Chávez S, García-Rubio M, Prado F & Aguilera A (2001) Hpr1 is preferentially required for transcription of either long or G+C-rich DNA sequences in Saccharomyces cerevisiae. *Mol Cell Biol* 21: 7054–7064

Chawla R, Redon S, Raftopoulou C, Wischnewski H, Gagos S & Azzalin CM (2011) Human UPF1 interacts with TPP1 and telomerase and sustains telomere leading-strand replication. *EMBO J* 30: 4047–4058

Chen L, Chen J-Y, Zhang X, Gu Y, Xiao R, Shao C, Tang P, Qian H, Luo D, Li H, *et al* (2017) R-ChIP Using Inactive RNase H Reveals Dynamic Coupling of R-loops with Transcriptional Pausing at Gene Promoters. *Molecular Cell* 68: 745-757.e5

Chen L, Zhang C, Ma W, Huang J, Zhao Y & Liu H (2022) METTL3-mediated m6A modification stabilizes TERRA and maintains telomere stability. *Nucleic Acids Res*: gkac1027

Cheng H, Dufu K, Lee C-S, Hsu JL, Dias A & Reed R (2006) Human mRNA Export Machinery Recruited to the 5′ End of mRNA. *Cell* 127: 1389–1400

Chi B, Wang Q, Wu G, Tan M, Wang L, Shi M, Chang X & Cheng H (2013) Aly and THO are required for assembly of the human TREX complex and association of TREX components with the spliced mRNA. *Nucleic Acids Res* 41: 1294–1306

Chong L, van Steensel B, Broccoli D, Erdjument-Bromage H, Hanish J, Tempst P & de Lange T (1995) A human telomeric protein. *Science* 270: 1663–1667

Cong Y-S, Wright WE & Shay JW (2002) Human Telomerase and Its Regulation. *Microbiol Mol Biol Rev* 66: 407

Crossley MP, Bocek M & Cimprich KA (2019) R-Loops as Cellular Regulators and Genomic Threats. *Mol Cell* 73: 398–411

Cusanelli E, Romero CAP & 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 & Chai W (2010) Molecular steps of G-overhang generation at human telomeres and its function in chromosome end protection. *EMBO J* 29: 2788–2801

Déjardin J & Kingston RE (2009) Purification of Proteins Associated with Specific Genomic Loci. *Cell* 136: 175–186

Denchi EL & de Lange T (2007) Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1. *Nature* 448: 1068–1071

Deng Z, Wang Z, Stong N, Plasschaert R, Moczan A, Chen H-S, Hu S, Wikramasinghe P, Davuluri RV, Bartolomei MS, *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

Dilley RL, Verma P, Cho NW, Winters HD, Wondisford AR & Greenberg RA (2016) Break-induced telomere synthesis underlies alternative telomere maintenance. *Nature* 539: 54–58

Doksani Y, Wu JY, de Lange T & Zhuang X (2013) Super-Resolution Fluorescence Imaging of Telomeres Reveals TRF2-Dependent T-loop Formation. *Cell* 155: 345–356

Domínguez-Sánchez MS, Barroso S, Gómez-González B, Luna R & Aguilera A (2011) Genome Instability and Transcription Elongation Impairment in Human Cells Depleted of THO/TREX. *PLoS Genet* 7: e1002386

Drosopoulos WC, Kosiyatrakul ST, Yan Z, Calderano SG & Schildkraut CL (2012) Human telomeres replicate using chromosome-specific, rather than universal, replication programs. *J Cell Biol* 197: 253–266

Dunham MA, Neumann AA, Fasching CL & Reddel RR (2000) Telomere maintenance by recombination in human cells. *Nat Genet* 26: 447–450

Duquette ML, Handa P, Vincent JA, Taylor AF & Maizels N (2004) Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Genes Dev* 18: 1618–1629

Durfee T, Mancini MA, Jones D, Elledge SJ & Lee WH (1994) The amino-terminal region of the retinoblastoma gene product binds a novel nuclear matrix protein that co-localizes to centers for RNA processing. *J Cell Biol* 127: 609–622

Episkopou H, Draskovic I, Van Beneden A, Tilman G, Mattiussi M, Gobin M, Arnoult N, Londoño-Vallejo A & Decottignies A (2014) Alternative Lengthening of Telomeres is characterized by reduced compaction of telomeric chromatin. *Nucleic Acids Res* 42: 4391–4405

Fagagna F d'Adda di, Reaper PM, Clay-Farrace L, Fiegler H, Carr P, von Zglinicki T, Saretzki G, Carter NP & Jackson SP (2003) A DNA damage checkpoint response in telomere-initiated senescence. *Nature* 426: 194–198

Fan HY, Cheng KK & Klein HL (1996) Mutations in the RNA polymerase II transcription machinery suppress the hyperrecombination mutant hpr1 delta of Saccharomyces cerevisiae. *Genetics* 142: 749–759

Feng J, Funk WD, Wang S-S, Weinrich SL, Avilion AA, Chiu C-P, Adams RR, Chang E, Allsopp RC, Yu J, *et al* (1995) The RNA Component of Human Telomerase. *Science* 269: 1236–1241

Feretzaki M, Pospisilova M, Valador Fernandes R, Lunardi T, Krejci L & Lingner J (2020) RAD51-dependent recruitment of TERRA lncRNA to telomeres through R-loops. *Nature* 587: 303–308

Feretzaki M, Renck Nunes P & Lingner J (2019) Expression and differential regulation of human TERRA at several chromosome ends. *RNA* 25: 1470–1480

Fleckner J, Zhang M, Valcárcel J & Green MR (1997) U2AF65 recruits a novel human DEAD box protein required for the U2 snRNP-branchpoint interaction. *Genes Dev* 11: 1864–1872

Flynn RL, Centore RC, O'Sullivan RJ, Rai R, Tse A, Songyang Z, Chang S, Karlseder J & Zou L (2011) TERRA and hnRNPA1 orchestrate an RPA-to-POT1 switch on telomeric single-stranded DNA. *Nature* 471: 532–536

Flynn RL, Cox KE, Jeitany M, Wakimoto H, Bryll AR, Ganem NJ, Bersani F, Pineda JR, Suvà ML, Benes CH, *et al* (2015) Alternative lengthening of telomeres renders cancer cells hypersensitive to ATR inhibitors. *Science* 347: 273–277

Gan W, Guan Z, Liu J, Gui T, Shen K, Manley JL & Li X (2011a) R-loop-mediated genomic instability is caused by impairment of replication fork progression. *Genes Dev* 25: 2041–2056

Gan W, Guan Z, Liu J, Gui T, Shen K, Manley JL & Li X (2011b) R-loop-mediated genomic instability is caused by impairment of replication fork progression. *Genes Dev* 25: 2041–2056

García-Muse T & Aguilera A (2019) R Loops: From Physiological to Pathological Roles. *Cell* 179: 604–618
García-Pichardo D, Cañas JC, García-Rubio ML, Gómez-González B, Rondón AG & Aguilera A (2017) Histone Mutants Separate R Loop Formation from Genome Instability Induction. *Mol Cell* 66: 597-609.e5
García-Rubio M, Aguilera P, Lafuente-Barquero J, Ruiz JF, Simon M-N, Geli V, Rondón AG & Aguilera A (2018) Yra1-bound RNA–DNA hybrids cause orientation-independent transcription–replication collisions and telomere instability. *Genes Dev* 32: 965–977
Gavaldá S, Santos-Pereira JM, García-Rubio ML, Luna R & Aguilera A (2016) Excess of Yra1 RNA-Binding Factor Causes Transcription-Dependent Genome Instability, Replication Impairment and Telomere Shortening. *PLOS Genet* 12: e1005966
Ghisays F, Garzia A, Wang H, Canasto-Chibuque C, Hohl M, Savage SA, Tuschl T & Petrini JHJ (2021) RTEL1 influences the abundance and localization of TERRA RNA. *Nat Commun* 12: 3016
Ginno PA, Lott PL, Christensen HC, Korf I & Chédin F (2012) R-Loop Formation Is a Distinctive Characteristic of Unmethylated Human CpG Island Promoters. *Molecular Cell* 45: 814–825
Glousker G, Briod A-S, Quadroni M & Lingner J (2020) Human shelterin protein POT1 prevents severe telomere instability induced by homology-directed DNA repair. *EMBO J* 39: e104500
Glousker G, Fernandes RV, Feretzaki M & Lingner J (2022) Detection of TERRA R-Loops at Human Telomeres. In *R-Loops*, Aguilera A & Ruzov A (eds) pp 159–171. New York, NY: Springer US
Glousker G & Lingner J (2021) Challenging endings: How telomeres prevent fragility. *BioEssays* 43: 2100157
Glover TW, Wilson TE & Arlt MF (2017) Fragile sites in cancer: more than meets the eye. *Nat Rev Cancer* 17: 489–501
Gómez-González B & Aguilera A (2007) Activation-induced cytidine deaminase action is strongly stimulated by mutations of the THO complex. *Proceedings of the National Academy of Sciences* 104: 8409–8414
Gómez-González B, García-Rubio M, Bermejo R, Gaillard H, Shirahige K, Marín A, Foiani M & Aguilera A (2011) Genome-wide function of THO/TREX in active genes prevents R-loop-dependent replication obstacles. *EMBO J* 30: 3106–3119
Gonzalez JB, Exposito LG, Hoang SM, Lynskey ML, Roncaioli JL, Ghosh A, Wallace CT, Modesti M, Bernstein KA, Sarkar SN, *et al* (2019) RAD51AP1 is an essential mediator of Alternative Lengthening of Telomeres. *Mol Cell* 76: 11-26.e7
Gottschling DE, Aparicio OM, Billington BL & Zakian VA (1990) Position effect at S. cerevisiae telomeres: Reversible repression of Pol II transcription. *Cell* 63: 751–762
Graf M, Bonetti D, Lockhart A, Serhal K, Kellner V, Maicher A, Jolivet P, Teixeira MT & Luke B (2017) Telomere Length Determines TERRA and R-Loop Regulation through the Cell Cycle. *Cell* 170: 72-85.e14
Granotier C, Pennarun G, Riou L, Hoffschir F, Gauthier LR, De Cian A, Gomez D, Mandine E, Riou J-F, Mergny J-L, *et al* (2005) Preferential binding of a G-quadruplex ligand to human chromosome ends. *Nucleic Acids Res* 33: 4182–4190
Greenwood J & Cooper JP (2012) Non-coding telomeric and subtelomeric transcripts are differentially regulated by telomeric and heterochromatin assembly factors in fission yeast. *Nucleic Acids Res* 40: 2956–2963
Greider CW & Blackburn EH (1985) Identification of a specific telomere terminal transferase activity in tetrahymena extracts. *Cell* 43: 405–413
Greider CW & Blackburn EH (1987) The telomere terminal transferase of tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity. *Cell* 51: 887–898
Griffith JD, Comeau L, Rosenfield S, Stansel RM, Bianchi A, Moss H & de Lange T (1999) Mammalian Telomeres End in a Large Duplex Loop. *Cell* 97: 503–514
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 & Lingner J (2013) A quantitative telomeric chromatin isolation protocol identifies different telomeric states. *Nat Commun* 4: 2848

Guh C-Y, Shen H-J, Chen LW, Chiu P-C, Liao I-H, Lo C-C, Chen Y, Hsieh Y-H, Chang T-C, Yen C-P, *et al* (2022) XPF activates break-induced telomere synthesis. *Nat Commun* 13: 5781

Guo S, Hakimi M-A, Baillat D, Chen X, Farber MJ, Klein-Szanto AJP, Cooch NS, Godwin AK & Shiekhattar R (2005) Linking transcriptional elongation and messenger RNA export to metastatic breast cancers. *Cancer Res* 65: 3011–3016

Hardy CF, Sussel L & Shore D (1992) A RAP1-interacting protein involved in transcriptional silencing and telomere length regulation. *Genes Dev* 6: 801–814

Hastie ND, Dempster M, Dunlop MG, Thompson AM, Green DK & Allshire RC (1990) Telomere reduction in human colorectal carcinoma and with ageing. *Nature* 346: 866–868

Hayflick L (1965) The limited in vitro lifetime of human diploid cell strains. *Exp Cell Res* 37: 614–636

Heaphy CM, Subhawong AP, Hong S-M, Goggins MG, Montgomery EA, Gabrielson E, Netto GJ, Epstein JI, Lotan TL, Westra WH, *et al* (2011a) Prevalence of the Alternative Lengthening of Telomeres Telomere Maintenance Mechanism in Human Cancer Subtypes. *Am J Pathol* 179: 1608–1615

Heaphy CM, de Wilde RF, Jiao Y, Klein AP, Edil BH, Shi C, Bettegowda C, Rodriguez FJ, Eberhart CG, Hebbar S, *et al* (2011b) Altered Telomeres in Tumors with ATRX and DAXX Mutations. *Science* 333: 425–425

Henderson ER & Blackburn EH (1989) An overhanging 3' terminus is a conserved feature of telomeres. *Mol Cell Biol* 9: 345–348

Henson JD, Hannay JA, McCarthy SW, Royds JA, Yeager TR, Robinson RA, Wharton SB, Jellinek DA, Arbuckle SM, Yoo J, *et al* (2005) A Robust Assay for Alternative Lengthening of Telomeres in Tumors Shows the Significance of Alternative Lengthening of Telomeres in Sarcomas and Astrocytomas. *Clin Cancer Res* 11: 217–225

Hodson C, Twest S van, Dylewska M, O'Rourke JJ, Tan W, Murphy VJ, Walia M, Abbouche L, Nieminuszczy J, Dunn E, *et al* (2022) Branchpoint translocation by fork remodelers as a general mechanism of R-loop removal. *Cell Rep* 41

Houghtaling BR, Cuttonaro L, Chang W & Smith S (2004) A dynamic molecular link between the telomere length regulator TRF1 and the chromosome end protector TRF2. *Curr Biol CB* 14: 1621–1631

Huertas P & Aguilera A (2003) Cotranscriptionally Formed DNA:RNA Hybrids Mediate Transcription Elongation Impairment and Transcription-Associated Recombination. *Mol Cell* 12: 711–721

Huffman KE, Levene SD, Tesmer VM, Shay JW & Wright WE (2000) Telomere Shortening Is Proportional to the Size of the G-rich Telomeric 3′-Overhang*. *J Biol Chem* 275: 19719–19722

Iglesias N & Stutz F (2008) Regulation of mRNP dynamics along the export pathway. *FEBS Lett* 582: 1987–1996

Jimeno S, Rondón AG, Luna R & Aguilera A (2002) The yeast THO complex and mRNA export factors link RNA metabolism with transcription and genome instability. *EMBO J* 21: 3526–3535

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA & Charpentier E (2012) A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* 337: 816–821

Kaminski N, Wondisford AR, Kwon Y, Lynskey ML, Bhargava R, Barroso-González J, García-Expósito L, He B, Xu M, Mellacheruvu D, *et al* (2022) RAD51AP1 regulates ALT-HDR through chromatin-directed homeostasis of TERRA. *Mol Cell* 82: 4001-4017.e7

Kelleher C, Kurth I & Lingner J (2005) Human Protection of Telomeres 1 (POT1) Is a Negative Regulator of Telomerase Activity In Vitro. *Mol Cell Biol* 25: 808–818

Kim NW, Piatyszek MA, Prowse KR, Harley CB, West MD, Ho PLC, Coviello GM, Wright WE, Weinrich SL & Shay JW (1994) Specific Association of Human Telomerase Activity with Immortal Cells and Cancer. *Science* 266: 2011–2015

Kim SH, Kaminker P & Campisi J (1999) TIN2, a new regulator of telomere length in human cells. *Nat Genet* 23: 405–412

Klobutcher LA, Swanton MT, Donini P & Prescott DM (1981) All gene-sized DNA molecules in four species of hypotrichs have the same terminal sequence and an unusual 3' terminus. *Proc Natl Acad Sci U S A* 78: 3015–3019

Köhler A & Hurt E (2007) Exporting RNA from the nucleus to the cytoplasm. *Nat Rev Mol Cell Biol* 8: 761–773

König P, Fairall L & Rhodes D (1998) Sequence-specific DNA recognition by the Myb-like domain of the human telomere binding protein TRF1: A model for the protein-DNA complex. *Nucleic Acids Res* 26: 1731–1740

Lafuente-Barquero J, García-Rubio ML, Martin-Alonso MS, Gómez-González B & Aguilera A (2020) Harmful DNA:RNA hybrids are formed in cis and in a Rad51-independent manner. *eLife* 9: e56674

Lahue E, Heckathorn J, Meyer Z, Smith J & Wolfe C (2005) TheSaccharomyces cerevisiae Sub2 protein suppresses heterochromatic silencing at telomeres and subtelomeric genes. *Yeast* 22: 537–551

de Lange T (2005) Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes Dev* 19: 2100–2110

de Lange T, Shiue L, Myers RM, Cox DR, Naylor SL, Killery AM & Varmus HE (1990) Structure and variability of human chromosome ends. *Mol Cell Biol* 10: 518–527

Lee YW, Arora R, Wischnewski H & Azzalin CM (2018) TRF1 participates in chromosome end protection by averting TRF2-dependent telomeric R loops. *Nat Struct Mol Biol* 25: 147–153

Lei EP, Krebber H & Silver PA (2001) Messenger RNAs are recruited for nuclear export during transcription. *Genes Dev* 15: 1771–1782

Leman AR, Dheekollu J, Deng Z, Lee SW, Das MM, Lieberman PM & Noguchi E (2012) Timeless preserves telomere length by promoting efficient DNA replication through human telomeres. *Cell Cycle Georget Tex* 11: 2337–2347

Lerner LK & Sale JE (2019) Replication of G Quadruplex DNA. *Genes* 10: 95

Lesbirel S, Viphakone N, Parker M, Parker J, Heath C, Sudbery I & Wilson SA (2018) The m6A-methylase complex recruits TREX and regulates mRNA export. *Sci Rep* 8: 13827

Lewis PW, Elsaesser SJ, Noh K-M, Stadler SC & Allis CD (2010) Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. *Proc Natl Acad Sci* 107: 14075–14080

Li B, Oestreich S & Lange T de (2000) Identification of Human Rap1: Implications for Telomere Evolution. *Cell* 101: 471–483

Li Y, Lin AW, Zhang X, Wang Y, Wang X & Goodrich DW (2007) Cancer Cells and Normal Cells Differ in Their Requirements for *Thoc1*. *Cancer Res* 67: 6657–6664

Li Y, Wang X, Zhang X & Goodrich DW (2005) Human hHpr1/p84/Thoc1 Regulates Transcriptional Elongation and Physically Links RNA Polymerase II and RNA Processing Factors. *Mol Cell Biol* 25: 4023–4033

Libri D, Dower K, Boulay J, Thomsen R, Rosbash M & Jensen TH (2002) Interactions between mRNA Export Commitment, 3′-End Quality Control, and Nuclear Degradation. 22: 13

Libri D, Graziani N, Saguez C & Boulay J (2001) Multiple roles for the yeast *SUB2* /yUAP56 gene in splicing. *Genes Dev* 15: 36–41

Lin C-YG, Näger AC, Lunardi T, Vančevska A, Lossaint G & Lingner J (2021) The human telomeric proteome during telomere replication. *Nucleic Acids Res* 49: 12119–12135

Lingner J, Hughes TR, Shevchenko A, Mann M, Lundblad V & Cech TR (1997) Reverse Transcriptase Motifs in the Catalytic Subunit of Telomerase. *Science* 276: 561–567

Liu D, O'Connor MS, Qin J & Songyang Z (2004a) Telosome, a mammalian telomere-associated complex formed by multiple telomeric proteins. *J Biol Chem* 279: 51338–51342

Liu D, Safari A, O'Connor MS, Chan DW, Laegeler A, Qin J & Songyang Z (2004b) PTOP interacts with POT1 and regulates its localization to telomeres. *Nat Cell Biol* 6: 673–680

Loayza D & De Lange T (2003) POT1 as a terminal transducer of TRF1 telomere length control. *Nature* 423: 1013–1018

Londoño-Vallejo JA, Der-Sarkissian H, Cazes L, Bacchetti S & Reddel RR (2004) Alternative Lengthening of Telomeres Is Characterized by High Rates of Telomeric Exchange. *Cancer Res* 64: 2324–2327

Lovejoy CA, Li W, Reisenweber S, Thongthip S, Bruno J, Lange T de, De S, Petrini JHJ, Sung PA, 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

Lu R & Pickett HA (2022) Telomeric replication stress: the beginning and the end for alternative lengthening of telomeres cancers. *Open Biol* 12: 220011

Luke B, Panza A, Redon S, Iglesias N, Li Z & 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

Lundblad V & Szostak JW (1989) A mutant with a defect in telomere elongation leads to senescence in yeast. *Cell* 57: 633–643

Luo M-J, Zhou Z, Magni K, Christoforides C, Rappsilber J, Mann M & Reed R (2001) Pre-mRNA splicing and mRNA export linked by direct interactions between UAP56 and Aly. *Nature* 413: 644–647

Maciejowski J & de Lange T (2017) Telomeres in cancer: tumour suppression and genome instability. *Nat Rev Mol Cell Biol* 18: 175–186

MacNeill S (2012) Composition and Dynamics of the Eukaryotic Replisome: A Brief Overview. In *The Eukaryotic Replisome: a Guide to Protein Structure and Function*, MacNeill S (ed) pp 1–17. Dordrecht: Springer Netherlands

Makarov VL, Hirose Y & Langmore JP (1997) Long G Tails at Both Ends of Human Chromosomes Suggest a C Strand Degradation Mechanism for Telomere Shortening. *Cell* 88: 657–666

Masuda S, Das R, Cheng H, Hurt E, Dorman N & Reed R (2005) Recruitment of the human TREX complex to mRNA during splicing. *Genes Dev* 19: 1512–1517

McClintock B (1939) The Behavior in Successive Nuclear Divisions of a Chromosome Broken at Meiosis. *Proc Natl Acad Sci U S A* 25: 405–416

McClintock B (1941) The Stability of Broken Ends of Chromosomes in Zea Mays. *Genetics* 26: 234–282

McElligott R & Wellinger RJ (1997) The terminal DNA structure of mammalian chromosomes. *EMBO J* 16: 3705–3714

Merz C, Urlaub H, Will CL & Lührmann R (2007) Protein composition of human mRNPs spliced in vitro and differential requirements for mRNP protein recruitment. *RNA* 13: 116–128

Meselson M & Stahl F (1958) The replication of DNA in Escherichia coli. *Proc Natl Acad Sci U S A* 44: 671–682

Miglietta G, Russo M & Capranico G (2020) G-quadruplex–R-loop interactions and the mechanism of anticancer G-quadruplex binders. *Nucleic Acids Res* 48: 11942

Misino S, Busch A, Wagner CB, Bento F & Luke B (2022) TERRA increases at short telomeres in yeast survivors and regulates survivor associated senescence (SAS). *Nucleic Acids Research* 50: 12829–12843

Moravec M, Wischnewski H, Bah A, Hu Y, Liu N, Lafranchi L, King MC & Azzalin CM (2016) TERRA promotes telomerase-mediated telomere elongation in Schizosaccharomyces pombe. *EMBO Rep* 17: 999–1012

Moyzis RK, Buckingham JM, Cram LS, Dani M, Deaven LL, Jones MD, Meyne J, Ratliff RL & Wu JR (1988) A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes. *Proc Natl Acad Sci* 85: 6622–6626

Muller HJ (1938) The remaking of chromosomes. *The Collecting Net* 13: 181–195

Mun SH, Oh B, Lee MJ, Bae S, Yang Y & Park-Min K-H (2022) THOC5 regulates human osteoclastogenesis. *Eur J Cell Biol* 101: 151248

Nakamura TM, Morin GB, Chapman KB, Weinrich SL, Andrews WH, Lingner J, Harley CB & Cech TR (1997) Telomerase Catalytic Subunit Homologs from Fission Yeast and Human. *Science* 277: 955–959

Nanavaty V, Sandhu R, Jehi SE, Pandya UM & Li B (2017) Trypanosoma brucei RAP1 maintains telomere and subtelomere integrity by suppressing TERRA and telomeric RNA:DNA hybrids. *Nucleic Acids Res* 45: 5785–5796

Nergadze SG, Farnung BO, Wischnewski H, Khoriauli L, Vitelli V, Chawla R, Giulotto E & Azzalin CM (2009) CpG-island promoters drive transcription of human telomeres. *RNA* 15: 2186–2194

Nguyen HD, Yadav T, Giri S, Saez B, Graubert TA & Zou L (2017) Functions of Replication Protein A as a Sensor of R Loops and a Regulator of RNaseH1. *Molecular Cell* 65: 832-847.e4

Ogino H, Nakabayashi K, Suzuki M, Takahashi E, Fujii M, Suzuki T & Ayusawa D (1998) Release of Telomeric DNA from Chromosomes in Immortal Human Cells Lacking Telomerase Activity. *Biochem Biophys Res Commun* 248: 223–227

Ohki R & Ishikawa F (2004) Telomere-bound TRF1 and TRF2 stall the replication fork at telomeric repeats. *Nucleic Acids Res* 32: 1627–1637

Okamoto K, Bartocci C, Ouzounov I, Diedrich JK, Yates III JR & Denchi EL (2013) A two-step mechanism for TRF2-mediated chromosome-end protection. *Nature* 494: 502–505

Olovnikov AM (1973) A theory of marginotomy: The incomplete copying of template margin in enzymic synthesis of polynucleotides and biological significance of the phenomenon. *J Theor Biol* 41: 181–190

Olson E, Nievera CJ, Klimovich V, Fanning E & Wu X (2006) RPA2 is a direct downstream target for ATR to regulate the S-phase checkpoint. *J Biol Chem* 281: 39517–39533

O'Sullivan RJ, Arnoult N, Lackner DH, Oganesian L, Haggblom C, Corpet A, Almouzni G & Karlseder J (2014) Rapid induction of alternative lengthening of telomeres by depletion of the histone chaperone ASF1. *Nat Struct Mol Biol* 21: 167–174

Peña Á, Gewartowski K, Mroczek S, Cuéllar J, Szykowska A, Prokop A, Czarnocki-Cieciura M, Piwowarski J, Tous C, Aguilera A, *et al* (2012) Architecture and nucleic acids recognition mechanism of the THO complex, an mRNP assembly factor: Structure and function of the THO complex. *EMBO J* 31: 1605–1616

Pérez-Calero C, Bayona-Feliu A, Xue X, Barroso SI, Muñoz S, González-Basallote VM, Sung P & Aguilera A (2020) UAP56/DDX39B is a major cotranscriptional RNA–DNA helicase that unwinds harmful R loops genome-wide. *Genes Dev* 34: 898–912

Pérez-Martínez L, Öztürk M, Butter F & Luke B (2020) Npl3 stabilizes R-loops at telomeres to prevent accelerated replicative senescence. *EMBO Rep* 21: e49087

Petti E, Buemi V, Zappone A, Schillaci O, Broccia PV, Dinami R, Matteoni S, Benetti R & Schoeftner S (2019) SFPQ and NONO suppress RNA:DNA-hybrid-related telomere instability. *Nat Commun* 10: 1001

Pfeiffer V, Crittin J, Grolimund L & Lingner J (2013) The THO complex component Thp2 counteracts telomeric R-loops and telomere shortening. *EMBO J* 32: 2861–2871

Pickett HA & Reddel RR (2015) Molecular mechanisms of activity and derepression of alternative lengthening of telomeres. *Nat Struct Mol Biol* 22: 875–880

Piruat JI & Aguilera A (1998) A novel yeast gene, THO2, is involved in RNA pol II transcription and provides new evidence for transcriptional elongation-associated recombination. *EMBO J* 17: 4859–4872

Pitzonka L, Ullas S, Chinnam M, Povinelli BJ, Fisher DT, Golding M, Appenheimer MM, Nemeth MJ, Evans S & Goodrich DW (2014) The Thoc1 Encoded Ribonucleoprotein Is Required for Myeloid Progenitor Cell Homeostasis in the Adult Mouse. *PLoS ONE* 9: e97628

Pitzonka L, Wang X, Ullas S, Wolff DW, Wang Y & Goodrich DW (2013) The THO Ribonucleoprotein Complex Is Required for Stem Cell Homeostasis in the Adult Mouse Small Intestine. *Mol Cell Biol* 33: 3505–3514

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

Porro A, Feuerhahn S, Reichenbach P & 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

Portman DS, O'Connor JP & Dreyfuss G (1997) YRA1, an essential Saccharomyces cerevisiae gene, encodes a novel nuclear protein with RNA annealing activity. *RNA N Y N* 3: 527–537

Preker PJ & Guthrie C (2006) Autoregulation of the mRNA export factor Yra1p requires inefficient splicing of its pre-mRNA. *RNA* 12: 994–1006

Pühringer T, Hohmann U, Fin L, Pacheco-Fiallos B, Schellhaas U, Brennecke J & Plaschka C (2020) Structure of the human core transcription-export complex reveals a hub for multivalent interactions. *eLife* 9: e61503

Rai R, Zheng H, He H, Luo Y, Multani A, Carpenter PB & Chang S (2010) The function of classical and alternative non-homologous end-joining pathways in the fusion of dysfunctional telomeres. *EMBO J* 29: 2598–2610

Reaban ME, Lebowitz J & Griffin JA (1994) Transcription induces the formation of a stable RNA.DNA hybrid in the immunoglobulin alpha switch region. *J Biol Chem* 269: 21850–21857

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

Redon S, Zemp I & Lingner J (2013) A three-state model for the regulation of telomerase by TERRA and hnRNPA1. *Nucleic Acids Res* 41: 9117–9128

Rehwinkel J, Herold A, Gari K, Köcher T, Rode M, Ciccarelli FL, Wilm M & Izaurralde E (2004) Genome-wide analysis of mRNAs regulated by the THO complex in Drosophila melanogaster. *Nat Struct Mol Biol* 11: 558–566

Ribes-Zamora A, Indiviglio SM, Mihalek I, Williams CL & Bertuch AA (2013) TRF2 Interaction with Ku Heterotetramerization Interface Gives Insight into c-NHEJ Prevention at Human Telomeres. *Cell Rep* 5: 194–206

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

Rodríguez-Navarro S, Sträßer K & Hurt E (2002) An intron in the *YRA1* gene is required to control Yra1 protein expression and mRNA export in yeast. *EMBO Rep* 3: 438–442

Rondón AG, Jimeno S, García-Rubio M & Aguilera A (2003) Molecular Evidence That the Eukaryotic THO/TREX Complex Is Required for Efficient Transcription Elongation. *J Biol Chem* 278: 39037–39043

Roumelioti F-M, Sotiriou SK, Katsini V, Chiourea M, Halazonetis TD & Gagos S (2016) Alternative lengthening of human telomeres is a conservative DNA replication process with features of break-induced replication. *EMBO Rep* 17: 1731–1737

Sagie S, Toubiana S, Hartono SR, Katzir H, Tzur-Gilat A, Havazelet S, Francastel C, Velasco G, Chédin F & Selig S (2017) Telomeres in ICF syndrome cells are vulnerable to DNA damage due to elevated DNA:RNA hybrids. *Nat Commun* 8: 14015

Saguez C, Schmid M, Olesen JR, Ghazy MAE-H, Qu X, Poulsen MB, Nasser T, Moore C & Jensen TH (2008) Nuclear mRNA Surveillance in THO/sub2 Mutants Is Triggered by Inefficient Polyadenylation. *Mol Cell* 31: 91–103

Salas-Armenteros I, Pérez-Calero C, Bayona-Feliu A, Tumini E, Luna R & Aguilera A (2017) Human THO–Sin3A interaction reveals new mechanisms to prevent R-loops that cause genome instability. *EMBO J* 36: 3532–3547

San Martin-Alonso M, Soler-Oliva ME, García-Rubio M, García-Muse T & Aguilera A (2021) Harmful R-loops are prevented via different cell cycle-specific mechanisms. *Nat Commun* 12: 4451

Sarek G, Kotsantis P, Ruis P, Van Ly D, Margalef P, Borel V, Zheng X-F, Flynn HR, Snijders AP, Chowdhury D, *et al* (2019) CDK phosphorylation of TRF2 controls t-loop dynamics during the cell cycle. *Nature* 575: 523–527

Sarek G, Vannier J-B, Panier S, Petrini JHJ & Boulton SJ (2015) TRF2 Recruits RTEL1 to Telomeres in S Phase to Promote T-Loop Unwinding. *Mol Cell* 57: 622–635

Schoeftner S & Blasco MA (2008) Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. *Nat Cell Biol* 10: 228–236

Schwartzentruber J, Korshunov A, Liu X-Y, Jones DTW, Pfaff E, Jacob K, Sturm D, Fontebasso AM, Quang D-AK, Tönjes M, *et al* (2012) Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature* 482: 226–231

Sfeir A, Kosiyatrakul ST, Hockemeyer D, MacRae SL, Karlseder J, Schildkraut CL & de Lange T (2009) Mammalian Telomeres Resemble Fragile Sites and Require TRF1 for Efficient Replication. *Cell* 138: 90–103

Sharma S & Chowdhury S (2022) Emerging mechanisms of telomerase reactivation in cancer. *Trends Cancer* 8: 632–641

Shatkin AJ (1976) Capping of eucaryotic mRNAs. *Cell* 9: 645–653

Shay JW & Bacchetti S (1997) A survey of telomerase activity in human cancer. *Eur J Cancer* 33: 787–791

Shen J, Zhang L & Zhao R (2007) Biochemical Characterization of the ATPase and Helicase Activity of UAP56, an Essential Pre-mRNA Splicing and mRNA Export Factor. *J Biol Chem* 282: 22544–22550

Silva B, Arora R & Azzalin CM (2022) The alternative lengthening of telomeres mechanism jeopardizes telomere integrity if not properly restricted. *Proc Natl Acad Sci* 119: e2208669119

Silva B, Arora R, Bione S & Azzalin CM (2021) TERRA transcription destabilizes telomere integrity to initiate break-induced replication in human ALT cells. *Nat Commun* 12: 3760

Silva B, Pentz R, Figueira AM, Arora R, Lee YW, Hodson C, Wischnewski H, Deans AJ & Azzalin CM (2019) FANCM limits ALT activity by restricting telomeric replication stress induced by deregulated BLM and R-loops. *Nat Commun* 10: 2253

Smolka JA, Sanz LA, Hartono SR & Chédin F (2021) Recognition of RNA by the S9.6 antibody creates pervasive artifacts when imaging RNA:DNA hybrids. *Journal of Cell Biology* 220: e202004079

Sobinoff AP, Allen JA, Neumann AA, Yang SF, Walsh ME, Henson JD, Reddel RR & Pickett HA (2017) BLM and SLX4 play opposing roles in recombination-dependent replication at human telomeres. *EMBO J* 36: 2907–2919

Söding J, Biegert A & Lupas AN (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* 33: W244–W248

Sollier J, Stork CT, García-Rubio ML, Paulsen RD, Aguilera A & Cimprich KA (2014) Transcription-coupled nucleotide excision repair factors promote R-loop-induced genome instability. *Mol Cell* 56: 777–785

Statello L, Guo C-J, Chen L-L & Huarte M (2021) Gene regulation by long non-coding RNAs and its biological functions. *Nat Rev Mol Cell Biol* 22: 96–118

van Steensel B, Smogorzewska A & de Lange T (1998) TRF2 Protects Human Telomeres from End-to-End Fusions. *Cell* 92: 401–413

Sträßer K & Hurt E (2000) Yra1p, a conserved nuclear RNA-binding protein, interacts directly with Mex67p and is required for mRNA export. *EMBO J* 19: 410–420

Sträßer K & Hurt E (2001) Splicing factor Sub2p is required for nuclear mRNA export through its interaction with Yra1p. *Nature* 413: 648–652

Sträßer K, Masuda S, Mason P, Pfannstiel J, Oppizzi M, Rodriguez-Navarro S, Rondón AG, Aguilera A, Struhl K, Reed R, *et al* (2002) TREX is a conserved complex coupling transcription with messenger RNA export. *Nature* 417: 304–308

Svendsen JM, Smogorzewska A, Sowa ME, O'Connell BC, Gygi SP, Elledge SJ & Harper JW (2009) Mammalian BTBD12/SLX4 assembles a Holliday junction resolvase and is required for DNA repair. *Cell* 138: 63–77

Szostak JW & Blackburn EH (1982) Cloning yeast telomeres on linear plasmid vectors. *Cell* 29: 245–255

Takai H, Smogorzewska A & de Lange T (2003) DNA Damage Foci at Dysfunctional Telomeres. *Curr Biol* 13: 1549–1556

Takai KK, Kibe T, Donigian JR, Frescas D & de Lange T (2011) Telomere protection by TPP1/POT1 requires tethering to TIN2. *Mol Cell* 44: 647–659

Taniguchi I & Ohno M (2008) ATP-Dependent Recruitment of Export Factor Aly/REF onto Intronless mRNAs by RNA Helicase UAP56. *Mol Cell Biol* 28: 601–608

Tardat M & Déjardin J (2018) Telomere chromatin establishment and its maintenance during mammalian development. *Chromosoma* 127: 3–18

Teasley DC, Parajuli S, Nguyen M, Moore HR, Alspach E, Lock YJ, Honaker Y, Saharia A, Piwnica-Worms H & Stewart SA (2015) Flap Endonuclease 1 Limits Telomere Fragility on the Leading Strand. *J Biol Chem* 290: 15133–15145

Tokutake Y, Matsumoto T, Watanabe T, Maeda S, Tahara H, Sakamoto S, Niida H, Sugimoto M, Ide T & Furuichi Y (1998) Extra-Chromosomal Telomere Repeat DNA in Telomerase-Negative Immortalized Cell Lines. *Biochem Biophys Res Commun* 247: 765–772

Tomlinson RL, Ziegler TD, Supakorndej T, Terns RM & Terns MP (2006) Cell cycle-regulated trafficking of human telomerase to telomeres. *Mol Biol Cell* 17: 955–965

Tous C & Aguilera A (2007) Impairment of transcription elongation by R-loops in vitro. *Biochem Biophys Res Commun* 360: 428–432

Valador Fernandes R, Feretzaki M & Lingner J (2021) The makings of TERRA R-loops at chromosome ends. *Cell Cycle* 20: 1745–1759

Vannier J-B, Pavicic-Kaltenbrunner V, Petalcorin MIR, Ding H & Boulton SJ (2012) RTEL1 Dismantles T Loops and Counteracts Telomeric G4-DNA to Maintain Telomere Integrity. *Cell* 149: 795–806

Vohhodina J, Goehring LJ, Liu B, Kong Q, Botchkarev VV, Huynh M, Liu Z, Abderazzaq FO, Clark AP, Ficarro SB, *et al* (2021) BRCA1 binds TERRA RNA and suppresses R-Loop-based telomeric DNA damage. *Nat Commun* 12: 3542

Voynov V, Verstrepen KJ, Jansen A, Runner VM, Buratowski S & Fink GR (2006) Genes with internal repeats require the THO complex for transcription. *Proc Natl Acad Sci* 103: 14423–14428

Vrbsky J, Akimcheva S, Watson JM, Turner TL, Daxinger L, Vyskot B, Aufsatz W & Riha K (2010) siRNA–Mediated Methylation of Arabidopsis Telomeres. *PLOS Genet* 6: e1000986

Wahba L, Gore SK & Koshland D (2013) The homologous recombination machinery modulates the formation of RNA–DNA hybrids and associated chromosome instability. *eLife* 2: e00505

Wan J, Zou S, Hu M, Zhu R, Xu J, Jiao Y & Fan S (2014) Thoc1 inhibits cell growth via induction of cell cycle arrest and apoptosis in lung cancer cells. *Mol Med Rep* 9: 2321–2327

Wang L, Miao Y-L, Zheng X, Lackford B, Zhou B, Han L, Yao C, Ward JM, Burkholder A, Lipchina I, *et al* (2013) The THO Complex Regulates Pluripotency Gene mRNA Export and Controls Embryonic Stem Cell Self-Renewal and Somatic Cell Reprogramming. *Cell Stem Cell* 13: 676–690

Wang X, Chang Y, Li Y, Zhang X & Goodrich DW (2006) Thoc1/Hpr1/p84 Is Essential for Early Embryonic Development in the Mouse. *Mol Cell Biol* 26: 4362–4367

Wang X, Chinnam M, Wang J, Wang Y, Zhang X, Marcon E, Moens P & Goodrich DW (2009) *Thoc1* Deficiency Compromises Gene Expression Necessary for Normal Testis Development in the Mouse. *Mol Cell Biol* 29: 2794–2803

Watson JD (1972) Origin of Concatemeric T7DNA. *Nature New Biol* 239: 197–201

Watson JD & Crick FHC (1953) Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* 171: 737–738

Wellinger RE, Prado F & Aguilera A (2006) Replication Fork Progression Is Impaired by Transcription in Hyperrecombinant Yeast Cells Lacking a Functional THO Complex. *Mol Cell Biol* 26: 3327–3334

Wellinger RJ, Wolf AJ & Zakian VA (1993) Saccharomyces telomeres acquire single-strand TG1–3 tails late in S phase. *Cell* 72: 51–60

West RW, Kruger B, Thomas S, Ma J & Milgrom E (2000) RLR1 (THO2), required for expressing lacZ fusions in yeast, is conserved from yeast to humans and is a suppressor of SIN4. *Gene* 243: 195–205

Wu L & Hickson ID (2003) The Bloom's syndrome helicase suppresses crossing over during homologous recombination. *Nature* 426: 870–874

Wu L, Multani AS, He H, Cosme-Blanco W, Deng Y, Deng JM, Bachilo O, Pathak S, Tahara H, Bailey SM, *et al* (2006) Pot1 Deficiency Initiates DNA Damage Checkpoint Activation and Aberrant Homologous Recombination at Telomeres. *Cell* 126: 49–62

Wu P, Takai H & 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

Xu B & Clayton DA (1996) RNA-DNA hybrid formation at the human mitochondrial heavy-strand origin ceases at replication start sites: an implication for RNA-DNA hybrids serving as primers. *EMBO J* 15: 3135–3143

Yadav T, Zhang J-M, Ouyang J, Leung W, Simoneau A & Zou L (2022) TERRA and RAD51AP1 promote alternative lengthening of telomeres through an R- to D-loop switch. *Mol Cell* 82: 3985-4000.e4

Yang Z, Sharma K & de Lange T (2022) TRF1 uses a noncanonical function of TFIIH to promote telomere replication. *Genes Dev*: genesdev;gad.349975.122v1

Ye J, Lenain C, Bauwens S, Rizzo A, Saint-Léger A, Poulet A, Benarroch D, Magdinier F, Morere J, Amiard S, *et al* (2010) TRF2 and Apollo Cooperate with Topoisomerase 2α to Protect Human Telomeres from Replicative Damage. *Cell* 142: 230–242

Ye JZ-S, Hockemeyer D, Krutchinsky AN, Loayza D, Hooper SM, Chait BT & de Lange T (2004) POT1-interacting protein PIP1: a telomere length regulator that recruits POT1 to the TIN2/TRF1 complex. *Genes Dev* 18: 1649–1654

Yeager TR, Neumann AA, Englezou A, Huschtscha LI, Noble JR & Reddel RR (1999) Telomerase-negative Immortalized Human Cells Contain a Novel Type of Promyelocytic Leukemia (PML) Body1. *Cancer Res* 59: 4175–4179

Yehezkel S, Segev Y, Viegas-Péquignot E, Skorecki K & 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

Yu G-L, Bradley JD, Attardi LD & Blackburn EH (1990) In vivo alteration of telomere sequences and senescence caused by mutated Tetrahymena telomerase RNAs. *Nature* 344: 126–132

Yu K, Chedin F, Hsieh C-L, Wilson TE & Lieber MR (2003) R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nat Immunol* 4: 442–451

Yu T-Y, Wang C-Y & Lin J-J (2012) Depleting Components of the THO Complex Causes Increased Telomere Length by Reducing the Expression of the Telomere-Associated Protein Rif1p. *PLoS ONE* 7: e33498

Zenklusen D, Vinciguerra P, Wyss J-C & Stutz F (2002) Stable mRNP Formation and Export Require Cotranscriptional Recruitment of the mRNA Export Factors Yra1p and Sub2p by Hpr1p. *Mol Cell Biol* 22: 8241–8253

Zhang J-M, Yadav T, Ouyang J, Lan L & Zou L (2019) Alternative Lengthening of Telomeres through Two Distinct Break-Induced Replication Pathways. *Cell Rep* 26: 955-968.e3

Zhang M & Green MR (2001) Identification and characterization of yUAP/Sub2p, a yeast homolog of the essential human pre-mRNA splicing factor hUAP56. *Genes Dev* 15: 30–35

Zhong Z, Shiue L, Kaplan S & de Lange T (1992) A mammalian factor that binds telomeric TTAGGG repeats in vitro. *Mol Cell Biol* 12: 4834–4843

Zhou Z, Luo M, Straesser K, Katahira J, Hurt E & Reed R (2000) The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans. *Nature* 407: 401–405

Zimmermann M, Kibe T, Kabir S & Lange T de (2014) TRF1 negotiates TTAGGG repeat-associated replication problems by recruiting the BLM helicase and the TPP1/POT1 repressor of ATR signaling. *Genes Dev* 28: 2477–2491

Curriculum vitae

Rita Valador Fernandes

PhD candidate in Molecular Life Sciences

14.Dec.1994 - Portugal

rvalador.fernandes@gmail.com

+41 76 247 12 94

linkedin.com/in/ritavfernandes

Av. Victor Ruffy, 23
Lausanne - Switzerland

EDUCATION

2018 – present | Lausanne, Switzerland

Doctoral degree in Molecular Life Sciences

École Polytechnique Fédérale de Lausanne (EPFL)
Swiss Institute for Experimental Cancer Research (ISREC)
Thesis Director: Professor Joachim Lingner
Regulation and impact of TERRA R-loops at human telomeres

2015 – 2017 | Stockholm, Sweden

Master's degree in Molecular Biology and Biomedicine

2nd year - **Karolinska Institutet - Cancer Centrum Karolinska (CCK)**
Erasmus Plus Mobility Program
Master's thesis project (grade: 20 out of 20)
Dr. Dan Grandér and Professor Katja Pokrovskaja's research group, supervised by Dr. Linda Viðarsdóttir
Study of the Long Non-Coding RNA C/EBPB-AS in Melanoma
Additional contribution to following projects:
Study of the lncRNAs CREB-AS, LRIF-AS and WT1-AS in Cancer
Study of the lncRNA FDG5-AS in Cancer and Autophagy

Lisbon, Portugal

1st year – **Nova University of Lisbon (UNL)** - School of Sciences and Technology (FCT)
Final weighted average: 19 out of 20
Courses: guia.unl.pt/en;
Optional courses: Biosensors, Biomaterials, Bionanotechnology, Molecular Immunology, Molecular Basis of Disease

2012-2015 | Lisbon, Portugal

Bachelor's degree in Cell and Molecular Biology

Nova University of Lisbon (UNL) - School of Sciences and Technology (FCT)
Final weighted average: 16 out of 20
Bachelor's internship - Dr. Rita Sobral's research group
Contribution of the Peptidoglycan Amidation for Antibiotic Resistance Mechanism in *Staphylococcus aureus*
Courses: guia.unl.pt/en;
Optional courses: Immunology, Molecular Toxicology, Forensic Chemistry

2009-2012 | Sintra, Portugal

Highschool – Sciences and Technology

Escola Secundária Santa Maria - Sintra
Final weighted average: 17.6 out of 20.0

DISTINCTIONS

- **Erasmus+ Traineeship Grant** – Issued by European Commission
- **Certificate of Merit "Best Student of Master's Degree in Molecular Genetics and Biomedicine of 2016/2017"**
- **Middle School/Highschool student merit award** – 9th, 10th and 11th grades

Rita Valador Fernandes

PhD candidate in Molecular Life Sciences

14.Dec.1994 - Portugal

SCIENTIFIC PUBLICATIONS/MANUSCRIPTS

The THO Complex counteracts TERRA R-loop-mediated telomere fragility in telomerase+ cells and recombination in ALT+ cells
Valador Fernandes R, Lingner J.
Manuscript under revision

Detection of TERRA R-Loops at Human Telomeres
Glousker G*, Valador Fernandes R*, Feretzaki M*, Lingner J. (*equal contribution)
Methods in Molecular Biology 2022

The makings of TERRA R-loops at chromosome ends
Valador Fernandes R, Feretzaki M, Lingner J.
Cell Cycle 2021

RAD51-dependent recruitment of TERRA lncRNA to telomeres through R-loops
Feretzaki M, Pospisilova M, Valador Fernandes R, Lunardi T, Krejci L, Lingner J.
Nature 2020

Silencing of CEBPB-AS1 modulates CEBPB expression and resensitizes BRAF-inhibitor resistant melanoma cells to vemurafenib
Vidarsdottir L, Valador Fernandes R, Zachariadis V, Das I, Edsbäcker E, Sigvaldadottir I, Azimi A, Höiom V, Hansson J, Grandér D, Egyházi Brage S, Pokrovskaja Tamm K.
Melanoma Research 2020

TECHNICAL/PERSONAL SKILLS

Fluent in **English** and **Portuguese;** Intermediate understanding of **Spanish** and **French**
English language assessment through Erasmus Plus Mobility program:
C2 (Proficient User – Mastery/ Proficiency) *Common European Framework of Reference for Languages*

Wet lab skills:

- Chromatin immunoprecipitation (ChIP), RNA immunoprecipitation (RNA IP), DNA:RNA immunoprecipitation (DRIP), molecular cloning, mutagenesis PCR, mammalian cell culture (including transfection, CRISPR-Cas9 cell engineering and viral transduction), bacterial cell culture (including transformation), subcellular fractionation, cytogenetics, immunofluorescence, fluorescent in situ hybridization (FISH), Western/Northern/Southern/Dot blot, SDS-PAGE, Bradford assay, pulse field gel electrophoresis, radioactive probe labelling, RT-quantitative PCR, flow cytometry (FACS) spectrophotometry, luminescence-based cell viability assays

Software skills:

- ImageJ – Fiji; Prism – GraphPad; SnapGene; Benchling; Management Software SLIMS; Adobe Illustrator CC; Adobe Photoshop CC; Microsoft Office Word, Excel and PowerPoint

+350 hours of **teaching** and **supervising** students

Additional certifications

- Certified course in **"Drug Discovery" University of California San Diego University** (Coursera) - 2023
- Certified course in **"Research Integrity" EMBO** - 2021
- Certified course in **"Design and Interpretation of Clinical Trials" Johns Hopkins University** (Coursera) - 2020

PARTICIPATION IN CONFERENCES/SEMINARS

- **EMBO meeting** "Telomere function and evolution in health and disease" (**seminar**) – Portugal - 2022
- **NCCR RNA & Disease Retreat** (**seminar**) – Switzerland - 2020 and 2022
- **Cold Spring Harbor meeting** on "Telomeres and Telomerase" (**poster presentation**) – USA - 2019
- **Austrian Swiss RNA Meeting** (**poster presentation**) – Austria - 2018