UNIVERSITÀ DEGLI STUDI DI TRIESTE

XXXI CICLO DEL DOTTORATO DI RICERCA IN BIOMEDICINA MOLECOLARE

Role of SFPQ function in RNA:DNA hybrids related genomic instability in telomeres

Settore scientifico disciplinare: **BIO/11**

DOTTORANDA
ANTONINA ZAPPONE

COORDINATORE
PROF. GERMANA MERONI

SUPERVISORE DI TESI
PROF. STEFAN SCHOEFTNER

ANNO ACCADEMICO: 2017/2018

ABSTRACT

In vertebrates, the UUAGGG telomere repeat containing long non-coding RNA TERRA is prone to form RNA:DNA hybrids at telomeres resulting the formation of R-loop structures, replicative stress and telomere instability. RNA:DNA hybrids represent a threat to genomic stability, but also contribute to recombination based alternative lengthening of telomeres (ALT).

My PhD thesis aimed to identifie novel TERRA interactors involved in regulation of RNA:DNA hybrids (project 1) and to obtain new insights into the molecular function of TERRA interactrors in telomere regulation by identifying novel proteins (project 2).

Here, we identify the TERRA binding proteins SFPQ as novel regulators of RNA:DNA hybrid related telomere instability. NONO and SFPQ locate at telomeres and have a common role in suppressing RNA:DNA hybrids and replication defects at telomeres. SFPQ acts as a barrier for homologous recombination at telomeres, thereby impacting on telomere length homeostasis. Our data identifies the RNA binding proteins SFPQ as novel regulators at telomeres that collaborate to ensure telomere integrity by suppressing telomere fragility and homologous recombination triggered by RNA:DNA hybrids.

Because of lack of enzymatic activity in SFPQ, we performed a mass spectrometry analysis and we found DAXX as novel SFPQ interacting protein. After the identification od SFPQ interacting domain with DAXX, we carried out preliminary experiment to evaluate the role od DAXX in DNA damage activation, in induction of replication defects and telomere dysfunction.

1. INTRODUCTION

1.1 Telomeres

1.1.1 Telomere structure and functions

Telomeres (from Greek telos (end) and meros (part)) are heterochromatic nucleoprotein complex sited at the end of eukaryotic chromosomes that are composed by thousands of repeats of a specific sequence. In vertebrates, telomere length is species specific, being approximately 10 -15 kb in humans and 50 kb in inbred strains of Mus musculus (De Lange et al. 1990; Palm and De Lange 2008) (Table 1).

Species	Telomere length	Telomere sequence	Reference
Ciliates			
Protozoan (*T. thermophila*)	120–420 bp	T_2G_4	[6]
Yeast			
Baker's yeast (*S. cerevisiae*)	200–300 bp	$TG_{2-3}(TG)_{1-6}$	[76]
Vertebrates			
Humans	5–15 kb	T_2AG_3	[59]
Mice	Up to 150 kb	T_2AG_3	[39]
Rats	20–100 kb	T_2AG_3	[17]
Birds	5–20 kb	T_2AG_3	[37]
Invertebrate			
Ants	9–13 kb	T_2AG_2	[50]
Plants			
Thale cress (*A. thaliana*)	2–5 kb	T_3AG_3	[68]

Table 1. Telomere length and telomere sequence in different species (Oeseburg H. et al. 2010)

In 1930s, two geneticists Herman J. Muller and Barbara McClintock provided the first evidence for telomeres thus showing that the end of irradiated chromosomes, unlike the rest of the genome, did not present alterations such as deletions and insertions (Muller H.J. et al. 1938). Moreover, the rupture of chromosomes induced by X-ray irradiation, resulted in the fusion of their ends without involving telomeres, giving rise to dicentric chromosomes (McClintock B. et al. 1939). These pioneering studies were crucial to demostrate the important role of telomeres for chromosome end protection and genome stability.

Telomeric DNA is double-stranded with a G-rich, single stranded overhang at the 3'-OH end. The 3' overhang of mammalian telomeres varies between 50–500 nt, which is considerably longer than the protrusion of most other eukaryotes. (Hemann and Greider 1999). This single strand has two important functions: i) it serves as a primer for telomere extension by telomerase; ii) it forms a

peculiar structure, known as T-loop. This configuration can hide the free 3'-OH end avoiding its recognitionas a double-strand break. Telomeric chromatin shares several characteristics with pericentromeric heterochromatin, such as the ability to silence nearby genes. This phenomenon known as "telomere position effect" or TPE was first described in *Drosophila melanogaster* but it is conserved also in yeast and mammals (Gottschling et al. 1990; Levis, Hazelrigg, and Rubin 1985). Differently to yeast, that only show nucleosomes in subtelomeric regions, vertebrate telomeres and subtelomeres have a classic nucleosomal organization with a slightly altered spacing (Makarov et al. 1993; Tommerup, Dousmanis, and de Lange 1994). Vertebrate telomeric chromatin is under-acetylated and shows characteristic features of other repetitive heterochromatic elements.

Despite the absence of CpG island, vertebrate telomeres are hyper-condensed because of specific histone post-translational modifications (PTMs), such as trimethylation of lysine 9 of histone H3 (H3K9me3), trimethylation of lysine 20 of histone H4 (H4K20me3), hypoacetylation of histone H3 and H4, and HP1 proteins (Schoeftner and Blasco 2010).

In addition to their protective function, telomeres are involved in several important cellular processes including mitosis, meiosis, aging and carcinogenesis. Telomeres ensure the faithful segregation of sister chromatids into daughter cells by playing an important role in tethering the chromosomes to the nuclear envelope and helping to position the chromosomes and promote homologous recognition and pairing (Calderón et al. 2014; Ohishi et al. 2014; Viera et al. 2015).

1.1.2 The Shelterin complex protects telomeres from signalling DNA damage

Telomere DNA is associated with a multi-unit protein complex known as telosome or Shelterin.

The Shelterin complex is composed of six proteins: TRF1 and TRF2 (Telomeric repeats Binding Factor 1 and 2), RAP1 (Repressor/Activator Protein1), TIN2 (TRF1 Interacting Nuclear factor 2), POT1 (Protector of Telomeres 1) and TPP1 (POT1-TIN2 organizing protein) (De Lange 2005) (Figure 1.1).

TRF1 and TRF2 share a common domain structure consisting of the TRF homology (TRFH) domain and a C-terminal SANT/Myb DNA-binding domain, which are connected through a flexible hinge domain (Broccoli et al. 1997; Y. Chen et al. 2008; Court et al. 2005).These proteins cooperate forming homodimers or tetramers and bind the double stranded telomere repeats.

TRF1 and TRF2 are linked by TIN2 which helps the stabilization of these proteins on the telomeric DNA. Moreover, TIN2 in crucial for the assembly of a complete shelterin complex and it links TIPP1/POT1 heterodimer to TRF1 and TRF2 (Frescas and De Lange 2014; K. K. Takai et al. 2017).

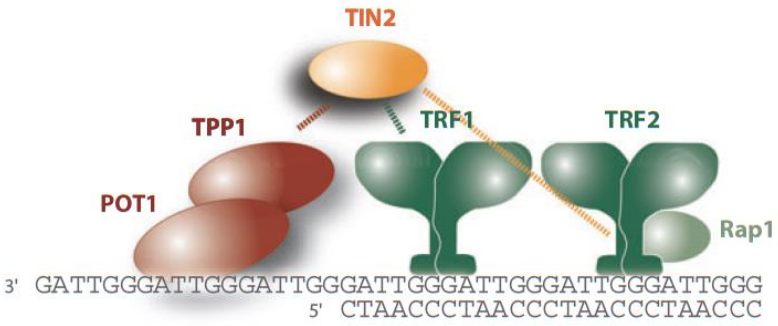


Figure 1.1. Shelterin complex. Schematic representation of six component of Shelterin complex (taken from Doksani Y. and de Lange T. 2014)

TRF2 is critical for T-loop formation (Super-resolution fluorescence imaging of telomeres reveals TRF2-dependent t-loop formation Ylli Doksani,#1 John Y. Wu,#2,3 Titia de Lange,1,* and Xiaowei Zhuang3,4,*). Telomeres normally terminate with a 3' single stranded-DNA overhang that invades into the double-stranded telomeric DNA by pairing with the C-strand, thus displacing the G-strand, forming a D-loop (Displacement Loop).

The strand invasion takes place at a distance from the physical end of telomeres and results in a large T-loop structure that function as protective cap at chromosome ends thereby protecting telomeres from cellular activities such as DNA damage checkpoints, DNA repair enzymes, degradation and end-to-end fusions (de Lange, 2004).

TRF2 is also implicated in the repression of DNA damage pathway blocking activation of ATM kinase (Doksani et al. 2013; Palm and De Lange 2008), while the inhibition of ATR kinase is mediated by TRF1 and POT1 (Figure 1.2).

The ATM pathway responds primarily to double strand breaks (DSBs), whereas ATR activation requires the formation of ssDNA. Both kinases phosphorylate histone H2AX on Serine 139 in a large chromatin domain surrounding the site of damage and promote the local accumulation of other DNA damage response factors (MDC1, 53BP1, the Mre11 complex, etc.), resulting in cytologically detectable foci that encompass hundreds of kb.

ATM and ATR phosphorylate two nucleoplasmic effector kinases, Chk1 and Chk2, that can block cell cycle progression. Chk1 and Chk2 also cooperate with ATM and ATR to activate p53, which further inhibits cell cycle progression through induction of the Cdk inhibitor p21.

TRF2 blocks the action of ATM to the telomere according to two models. In the first proposed model, TRF2 maintains a T-loop structure, hiding chromosome ends from Mre11 and ATM action. The second model proposes a direct inhibition of ATM signaling cascade. The damage structure is recognized but the downstream pathway is blocked. In favor of this model, overexpression of TRF2 can dampen the

activation of ATM kinase, even at non-telomeric sites of DNA damage (Lazzerini-Denchi and Sfeir 2016).

When TRF1 is deleted from mouse cells, replication frequently stalls near the subtelomeric/telomeric junction. These replication problems lead to the so-called fragile telomere phenotype. This phenotype is recognized as multiple telomeric fluorescence in situ hybridization (FISH) signals that appear at single chromatid ends, giving telomeres a broken or incompletely condensed appearance (Martínez et al. 2009; Sfeir et al. 2009).

The telomeric replication stress induced upon TRF1 deletion results in the activation of the ATR kinase, as evidenced by the ATR-dependent appearance of γH2AX and 53BP1 at telomeres (Martínez et al. 2009; Sfeir et al. 2009). These telomere dysfunction-induced foci (TIFs) (H. Takai, Smogorzewska, and De Lange 2003) only form in cells that have progressed through S phase without TRF1 (Sfeir et al. 2009), arguing that ATR activation is a consequence of the replication problems and not due to general telomere deprotection.

POT1 depletion causes the same phenotype of TRF1 depletion (Denchi and De Lange 2007). The ability of POT1 to repress the ATR kinase signaling cascade is dependent on its association with TPP1. (Denchi and De Lange 2007; Hockemeyer et al. 2007). TPP1 functions to recruit POT1 to telomeres and improves its ability to bind to single-stranded DNA *in vitro* (F. Wang et al. 2007). Both features of TPP1 are proposed to be relevant in the repression of ATR. Inhibition of TPP1 gives rise to a DNA damage response at telomeres that is indistinguishable from the response to POT1 deletion (Denchi and De Lange 2007; F. Wang et al. 2007).

The model for ATR signal repression by telomeres proposes that POT1 /TPP1 prevents the binding of RPA to Single-stranded telomere DNA (F. Wang et al. 2007).

Shelterin components have been shown to be involved in telomere length maintenance. Mutations in TRF1 or TRF2 promote telomere elongation, while overexpression of these proteins lead to telomere shortening (Hug and Lingner 2006). RAP1 interacts only with TRF2 and it was demonstrated that acts as a negative regulator of telomere length (Frank et al. 2015).

Suppression of TPP1 by RNAi or the disruption of the TPP1–POT1 interaction results in telomere lengthening, accompanied by loss of the POT1 at telomere.

Recent evidences suggest that the proteins TTP1 (Nandakumar and Cech 2013) and TIN 2 (Frank et al. 2015) play a role in the recruitment of the reverse transcriptase complex telomerase at telomere. Other proteins, as POT1, are negative regulator of telomere length suggesting that shelterin subunits may have a function in counting telomeric repeats to regulate telomerase activity (Smogorzewska and de Lange 2004).

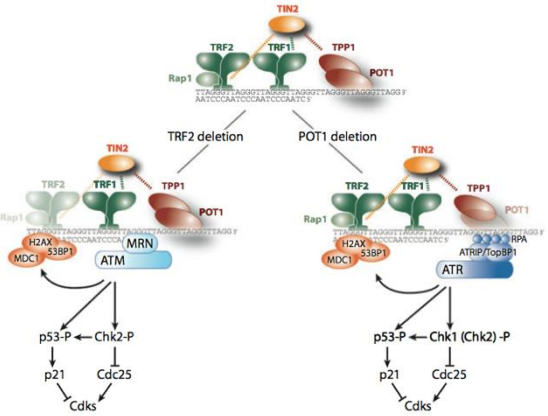


Figure 1.2. The DNA damage response at telomeres is regulated by the Shelterin complex .
TRF2 deletion induces the activation of ATM pathway due to the unprotected dsDNA at chromosome end; whereas POT1 deletion leads to RPA recruitment at 3' overhang eliciting the activation of ATR pathway (adapted from Palm W. and de Lange L. 2008)

In addition to Shelterin, other protein complexes were shown to be located at telomere repeats. For example, CST complex consisting of conserved telomere protection component 1 (CTC1), suppressor of cdc thirteen 1 (STN1) and telomeric pathway with STN1 (TEN1) (Rice and Skordalakes 2016).

CST localizes specifically to the single-stranded telomeric DNA, including the telomeric overhang, where it is involved in chromosome end capping and telomere length regulation (L. Y. Chen, Redon, and Lingner 2012; Nakaoka et al. 2012; Wellinger 2009). However, there are increasing evidences that Stn1-Ten1, CST sub-complex has additional, non-telomere related functions.

Moreover, telomeres contain factors involved in DNA damage response such as the Ku70/86 heterodimer and the MRE11/RAD50/NBS1 (MRN) complex (Dejardin and Kingsto 2012; Nittis et al. 2010), nucleases (Structure-specific endonuclease subunit (SLX4), Apollo), helicases (RecQ helicases Werner (WRN) and Bloom (BLM)) (Du et al. 2004; Lillard-Wetherell et al. 2004) and chromatin modifiers, such as α-thalassemia/mental retardation syndrome X-linked (ATRX).

1.2 Regulation of telomere maintenance

Each cell division causes telomere shortening due to the inability of DNA polymerase to fully synthesize the ends of linear chromosomes (Olovnikov 1973) . This process leads to the progressive loss of 200 nucleotides until telomeres reach a threshold length that activates DNA damage, apoptosis and senescence (d'Adda di Fagagna et al. 2003).

However, a subset of cell types, such as germinal cells, stem cells and cancer cells are able to maintain telomere length due to the activation of mechanisms known as Telomere Maintenance Mechanisms

(TMMs). Two TMMs are known: telomerase-mediated telomere maintenance and alternative lengthening of telomere (ALT).

1.2.1 Replicative senescence

Conventional DNA polymerases replicate DNA only in the 5' to 3'direction and require short RNA primers for initiation. Hence, while synthesis of the leading strand in the 5'-3' direction results in complete strand replication, synthesis of the lagging strand results in DNA loss at the 5'end due to the removal of the terminal RNA primer. Thus, the incapability of the DNA replication machinery to fully synthesize the lagging strand determines a progressive shortening of chromosome ends upon repeated mitotic cell division, a phenomenon known as "end replication problem".

Based on these findings, Olovnikov proposed a link between telomeric DNA shortening and the Hayflick limit (Olovnikov 1973) which refers induction of senescence in primary human cells that were cultured in vitro to the maximum cell divisions numbers (Hayflick and Moorhead 1961). This state, named as "replicative senescence",is associated with the appearance of DNA damage at chromosome ends (Levy et al. 1992). For this reason, telomere shortening has been proposed as one of the molecular hallmarks of aging (López-Otín et al. 2013).

In this contest, the loss of telomeric repetitions leads to a dislocation of the shelterin proteins and a consequent destabilization of the t-loop conformation (Griffith et al. 1999). The lack of a structural organization that protects chromosome ends leads to the activation of the DNA repair machinery at telomeres, recognized as DNA double-strand breaks. Moreover critical telomere shortening leads to telomere dysfunction with consequent chromosome fusion and genomic instability (Capper et al. 2007).

Senescence represents a physiological response that cells must overcome in order to divide indefinitely and develop into tumors. Almost 90% of human cancer cells acquire unlimited proliferative potential through reactivation of telomerase (Artandi and Depinho 2010; N. W. Kim et al. 1994), while the remaining 10% maintains telomere length by a recombination-mediated process termed ALT (Neumann et al. 2013; T M Bryan 1995).

1.2.2 Telomere length maintenance by telomerase

In the mid-1980s, Elisabeth Blackburn and Carol Greider demonstrated the existence of an enzymatic activity within human cell extracts that added tandem hexanucleotides to natural chromosome ends: this finally led to the discovery of telomerase (Greider and Blackburn 1985).

Telomerase is a reverse transcriptase (RNA dependent polymerase) composed by different subunits and additional associated proteins (Figure 1.3) Two components are fundamental for its catalytic activity: the reverse transcriptase (TERT in human and in mouse) and the RNA component (TERC in mouse, hTR in human) which serves as a template to synthesize telomere DNA repeats.

Mass spectrometry analysis revealed that telomerase contains two molecules each of TERT, hTR and a pseudouridylase H/ACA ribonucleoprotein protein, dyskerin (D. Fu and Collins 2007). Dyskerin is essential for the stabilization of telomerase complex (Greider and Blackburn 1985). Other accessory proteins are required for telomerase assembly: two ATPases, reptin and pontin (Venteicher et al. 2008) -and three additional H/ACA RNP proteins - NOP10, NHP2, and GAR1 (Egan and Collins 2012). An additional protein, NAF1, may be transiently involved in telomerase assembly. NAF1 interact with dyskerin and escorts it to the nascent H/ACA RNA before being replaced by GAR1 to yield mature H/ACA RNPs in Cajal bodies and nucleoli (Darzacq et al. 2006).

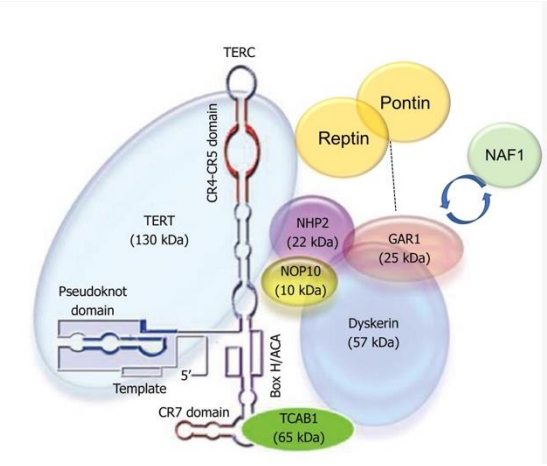


Figure 1.3. Telomerase complex. Telomerase is a large RNP. Catalytic subunit TERT and RNA component TERC with associated telomerase factors : TCAB1, dyskerin, its three associated proteins NHP2, NOP10 and GAR, Repentin, Pontin and NAF1 are shown in the figure (adapted from Ly 2011)

The gene hTERT of ~37 kb is present in the human genome as a single copy sequence in the chromosome 5p15.33 (Bryce et al. 2000) and synthetizes a relatively large protein (127 kDa). Both, bioinformatics and mutational studies have collectively established that hTERT contains four main structural elements: i) the telomerase N-terminal (TEN) domain that contains conserved DNA and RNA-binding domains, ii) a central catalytic reverse transcriptase (RT) domain; iii) TR-binding domain (TRBD), and iv) a short C-terminal extension (CTE) (Blackburn and Collins 2011).

hTR has been isolated from ciliates, yeasts and mammals and is essential for telomerase activity, even if it is not considered a limiting factor (Blasco et al. 1995; Cairney and Keith 2008). In vertebrates, RNA of telomerase has divergences in size and sequence and has an extension between

382 and 559 nucleotides (Tsao, Wu, and Lin 1998). In humans, hTR is 451 nucleotides long and contains a sequence of 11 bp (5'-CUAACCCUAAC-3') that serves as template for reverse transcription by telomerase (Morin 1989).

Telomerase-dependent telomere elongation is a three steps mechanism that normally occurs in S phase and includes: binding of telomerase to telomeric 3' ssDNA; addition of 50-60 nucleotides to the 3' overhang of telomeres and successive translocation and repositioning of the 3' overhang of telomere with the 3' template boundary of hTR to start a new round of telomere elongation (Schmidt J.C. and Cech T.R. 2015) (Figure 1.4). Once telomerase has completed its function, DNA primase synthesizes an RNA primer at the newly synthesized lagging strand, DNA polymerase elongates the Okazaki fragments joined by DNA ligase to form a long continuous DNA strand. A short region at the 3' end will remain single stranded.

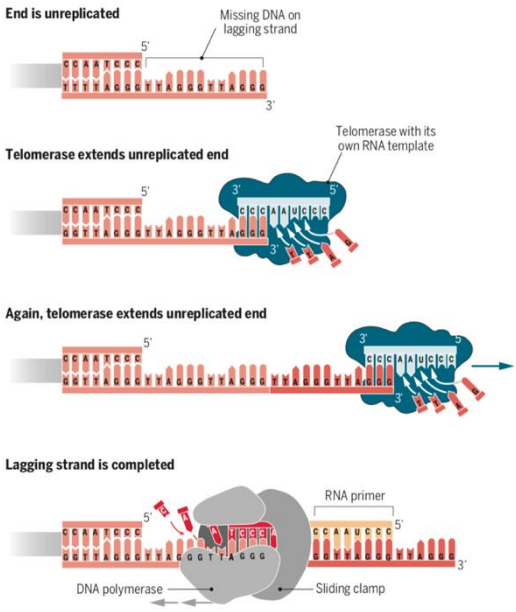


Figure 1.4. Extension of chromosome ends by telomerase which catalyzes synthesis of telomeric repeat. Telomerase translocates in the 5' to 3' direction on the chromosome and adds more telomeric repeats. DNA polymerase and other associated enzymes fill in the leading strand using lagging strand synthesis. Telomere is restored to appropriate length. (Figure takenf from (Blackburn, Epel, and Lin 2015))

The majority of human somatic cells, except for hematopoietic and germ cells, do not show telomerase activity resulting in a progressive telomere shortening. Telomerase activity is tightly regulated during normal growth and development in order to avoid an unlimited proliferation that sets the stage for malignancy (Wright et al. 1996). Telomerase activity has been observed in several human fetal tissues such as muscle, lung and skin, suggesting that this enzyme is active during development and is repressed in adult tissues (Chatziantonou et al. 2001).

Telomerase is detected in approximately 90% of all malignant tumors and it may predict poor outcome, thus making telomerase both a highly attractive biomarker and target for the development of mechanism-based cancer diagnostics, prognostics and therapeutics (Akincilar, Unal, and Tergaonkar 2016).

The activity and expression of telomerase is tightly regulated at transcriptional, post-transcriprional and epigenetis levels. hTERT promoter is regulated by the c-Myc oncogene, which recognizes E-boxes within the promoter (Greenberg et al. 1999), by estrogen receptor and Sp1 (Kyo et al. 2008) and by Wnt/β-catenin pathway (Zhang et al. 2012). Moreover, hTERT promoter is linked to a large CpG island that is involved in the epigenetic regulation of hTERT expression (Sui et al. 2013). Finally, post-transcriptional regulation of hTERT expression can occur through alternative splicing that affect the enzymatic activity of the hTERT protein.

1.2.2.1 Alternative Lengthening of Telomeres

It has been demonstrated that 10% of human cancer cells are able to maintain their telomeres through a mechanism based on homologous recombination. ALT mechanism is the prevalent mechanism in tumors of mesenchymal origin including osteosarcoma, astrocytoma, and liposarcoma and tumors from Li-fraumeni syndrome patients (Bryan et al. 1997; Anthony J. Cesare and Reddel 2010).

Telomeres of ALT cells retain canonical attributes: the presence of duplex TTAGGG repeats with single-stranded terminal overhangs of the G-rich strand (G-tails), the presence of the shelterin complex and other telomere-associated proteins, and the ability to form t-loops.

In addition to these features, ALT cells show peculiar characteristics linked to the activation of homologous recombination. The most striking ALT feature is the loss of telomeric DNA from telomeres that remains as extrachromosomal telomeric DNA forming double-strand telomeric circles (A J Cesare and Griffith 2004; R. C. Wang, Smogorzewska, and De Lange 2004) or single-strand telomeric DNA named C-circle or G-circle depending on whether it is the C-rich or G-rich strand, respectively.

C-circles are generated by nucleolytic degradation of the G-rich strand of t-circles and it was estimated that there are approximately 1,000 of C-circles per cell. ALT cells also contain G-circles, but these are 100-fold less abundant (Henson et al. 2009).

A quantitative assay showed that there are 750-fold more C-circles in ALT cells than in telomerase positive cell lines or non-immortalized cell strains.

Assaying C-circle levels may be a useful screen for chemicals that inhibit ALT activity.

Moreover, C-circles were detected in blood samples from patients with ALT-positive osteosarcomas. This promising result suggests that assaying C-circle levels might have utility as a blood test for diagnosing ALT positive tumours or for monitoring effectiveness of their treatment. (Harrington and Pucci 2018). Furthermore, when ALT was inhibited, most C-circles disappeared within 24 hours (Henson et al. 2009).

It is also possible to find linear double-stranded DNA and very high molecular weight 't-complex' DNA that is likely to contain abnormal, highly branched structures (A. Nabetani and Ishikawa 2009)

In contrast to human telomerase-positive cells, ALT cells are characterized by very heterogeneous telomere length that varies from 3 kb to more than 50 kb (De Vitis, Berardinelli, and Sgura 2018) due to the unbalanced homologous recombination.

There are other features that identify ALT cells as an increased occurrence of telomere sister chromatid exchange (T-SCE). In ALT cells telomeric DNA and associated binding proteins are found in promyelocytic leukaemia nuclear bodies (PML bodies) defined as ALT-associated PML bodies (APBs) It is widely demonstrated that APBs are formed by PML-NB components like PML, SP100, SUMO and telomere repeat-associated proteins such as TRF1, TRF2, POT1 or RAP1 (Yeager et al. 1999). Furthermore, APBs host factors are involved in DNA damage response and repair, such as RAD9, hHUS1, hRAD1, hRAD17, MRN complex consisting of NBS1, MRE11 and RAD50, the phosphorylated histone variant H2A.X (γH2A.X), the RecQ-like helicase BML, heterochromatin protein HP1 or the structural maintenance of chromosome complex (SMC5/6), which includes the SUMO E3 ligase MMS21. In line with telomere recombination APBs contain proteins involved in homologous recombination such as the endonuclease MUS81, replication protein A (RPA), RAD51 and RAD52, breast cancer susceptibility protein 1 (BRCA1) (Akira Nabetani and Ishikawa 2011).

Due to the special protein content, APBs promote replication and/or homologous recombination-mediated extension of telomeres (Chung, Leonhardt, and Rippe 2011).

1.2.2.2 Alternative Lengthening of Telomeres mechanism

Two mechanisms for telomere elongation by ALT have been proposed that are not mutually exclusive (Figure 1.5).

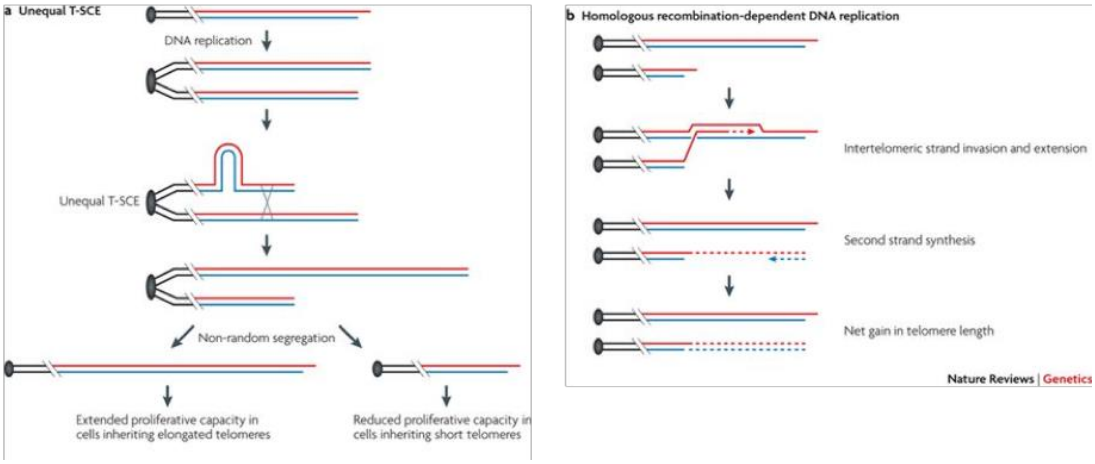


Fig. 1.5. Two model of ALT mechanism: Two distinct models propose to explain ALT pathway a) this model is based on unequal telomere sister chromatid exchange (T-SCE). It can lead to an heterogeneus poèilation of telomeres. b) The second model propose recombination-mediated synthesis of new telomeris DNA using the sequence of an adjacent telomere from another chromosome (adapted from Pickett H.A. and Reddel R.R. 2015)

The first mechanism is based on the hypothesis that unequal telomere sister chromatid exchanges (T-SCEs) can give rise to a daughter cell that has a lengthened telomere and a prolonged proliferative capacity and another daughter cell with a shortened telomere and decreased proliferative capacity It is currently unknown whether such a mechanism for segregation of telomeres exists (Muntoni and Reddel 2005; Akira Nabetani and Ishikawa 2011).

According to the second hypothesis, ALT results from the recombination-mediated synthesis of new telomeric DNA using an existing telomeric sequence from an adjacent telomere as a copy template. This proposed model is supported by the observation that a DNA tag inserted in a telomere was copied into an adjacent one and was also duplicated in its original location in human ALT cells, but not in telomerase-positive cells (Muntoni et al. 2009; Neumann et al. 2013).

On the mechanistic level, a putative four steps model has been proposed for the initiation of recombination (Pickett and Reddel 2015) (Figure 1.6). The single-stranded G-overhang at the telomere terminus can invade homologous DNA forming an HR intermediate structure (step 1). The strand invasion step is followed by the template directed synthesis of telomeric DNA that could be operated by polymerase δ or ζ, given their role in HR (step2) (Sharma et al. 2012). The third step consists in the processing of HR intermediate products that must occur before chromosome segregation. A fourth step of filling in of the complementary strand has been hypothesized, but is still unknown whether it occurs. The DNA template for new telomeric DNA synthesis by ALT may be a sister chromatid (Figure 1.6 a) or telomeres could copy themselves looping out or looping back on themselves (Figure 1.6 b-c). Recent evidences also identified a long-range movement driven by a

homology-directed searching mechanism that allows telomere clustering, providing distant copy templates for ALT activity (Figure 1.d) (Cho et al. 2014).

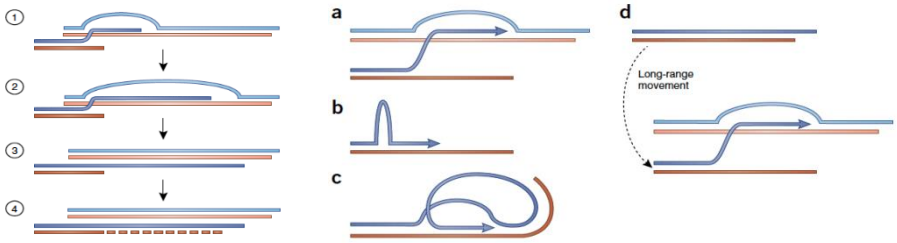


Fig. 1.6. ALT mechanism: Putative four steps model of ALT mechanism (left panel). Different possible DNA copy templates in alternatie lengthening of telomere mechanism (right panel). Use of a nearby copy template in the sister-chromatid telomere (a) or in the same telomere by looping out (b) or by t-loop formation (c) is depicted. (d) A distant telomere may be used as a copy template after long-range movement (adapted from Pickett H.A. and Reddel R.R. 2015).

Telomere dysfunction is associated with genetic instability and tumorigenesis or senescence in mitotically dividing cells. Two large, highly conserved protein kinases play a central role in sensing the DNA damage response (DDR) mechanism in human cells: ataxia telangiectasia mutated (ATM) and ataxia telangiectasia related (ATR) kinases, which have different, but partly overlapping functions (Guo et al. 2007).

ATM and ATR phosphorylate a variety of downstream target proteins, affecting checkpoint-mediated cell cycle arrest, DNA repair, apoptosis, and senescence (Matsuoka et al., 2007). A short telomere–induced DDR is manifested through the rapid accumulation of phosphorylated ATM, 53BP1, γ-H2AX, the MRN (Mre11–Rad50–Nbs1) complex at telomeres. Paradoxically, several important DDR proteinsalso localize to telomeres in the absence of DNA damage and have a role in telomere protection from DDRs, telomere length maintenance, and telomere replication. These proteins ATM, ATR, MRN, Ku86, ERCC1, PARP2, BLM (Bloom syndrome protein), and WRN (Werner syndrome protein) (Blasco, 2005), demonstrate a critical interplay between DDR proteins and telomere metabolism in eukaryotic cells. Moreover, ATM and ATR activate the tumor suppressor p53, through its phosphorylation. This leads to the activation of p21, a kinase inhibitor that inhibits two cyclin-dependent kinase, Cyclin E and A-Cdk2, interfering with cell cycle progression (Fig. 14) (Maser and DePhino, 2004).

ATM is a nuclear protein kinase with a catalytic domain similar to the one of phosphatidylinositol 3-kinases (PI 3-kinases) (Shiloh, 2003). This protein was discovered as a mutated protein in patients with ataxia-telagiectasia (A-T), a sever genetic disorder characterized by cerebellar degeneration,

neuromotor dysfunction, chromosomal instability, immune system defects, cancer predisposition, and acute sensitivity to ionizing radiations. Importantly, the cells of this patients present shorter telomeres compared to the normal cells (Metcalfe et al. 1996). This data was also confirmed through ATM knock out mice (Hande et al. 2001). It was then investigated if ATM play a central role in the regulation of telomere length. The main results obtained indicate that ATM and ATR are not needed to recruit telomerase neither or maintain telomere length, however their absence leads to an increased fragility and telomeric recombination (McNees et al. 2010). On the other hand, the ATR-deficient Seckel mouse strain was used to examine the function of ATR in telomerase recruitment and telomere function (Murga et al. 2009), concluding that ATR has an important role in the suppression of telomere fragility and recombination.

1.3 The DNA damage repair at telomeres

One of the most important functions of telomeres is to inhibit the activation of ATM and ATR pathways at the ends of chromosomes, in order to prevent DNA damage signalling. This function is carried out by shelterin components TRF1, TRF2 and POT1 (section 1.1.2).

Shelterin components play a key role not only in suppressing the DNA damage response but also in protecting the natural chromosome ends from inappropriate repair. In human cells, double strand breaks are repaired through two main systems: non-homologous end joining (NHEJ) and homologous recombination (HR) (Palm and de Lange, 2008).

1.3.1 Non-homologous end joining (NHEJ)

Telomere dysfunction are repaired mainly through NHEJ mechanism in which the endonuclease ERCC1-XPF degrades the 3' overhang and subsequently DNA ligase IV joins the breaks, providing suitable substrates to form dicentric chromosomes or chromosomal fusions (Celli and de Lange, 2005). TRF2 is crucial to prevent the activation of NHEJ. In cells that lack TRF2, the majority of telomeres is processed by NHEJ (Van Steensel et al. 1998). However, the mechanism by which TRF2 prevents the NHEJ activity on telomeres is unclear. Two main models have been proposed. the first one is that the presence of t-loop prevents the recruitment of ku70/80 heterodimer, thus cells lacking TRF2 shows chromosomal fusions (Celli et al. 2008). The second hypothesis is that the presence of TRF2/Rap1 at telomeres blocks the NHEJ machinery access (Bae and Baumann, 2007).

1.3.2 Homology-Directed Repair (HDR)

Three major types of HR outcomes at telomeres have been described (Figure 1.7): t-loop, telomere sister chromatid exchange (T-SCE) and recombination between a telomere and interstitial telomeric DNA.

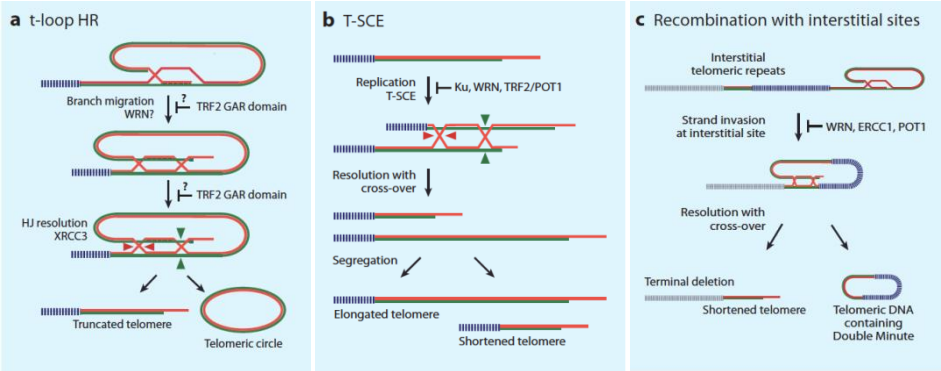


Figure 1.7. Homology-directed repair mechanisms: in mammalian cells. Homology-directed repair exists in three different forms. a) T-loop homologous recombination, b) Telomere sister chromatid exchange and c) Recombination with interstitial sites (adapted from Palm W. and de Lange T. 2008)

1.3.2.1 T-loop homologous recombination (T-loop HR)

The T-loop structure results from the single stranded 3' overhang of telomeres invasion into telomeric double-stranded DNA that generates the so-called D-loop, resembling an early step HR. The Shelterin protein TRF2 is the main protein involved in T-loop formation (Doksani Y. et al. 2013). In fact, TRF2 is able to induce topological changes in telomeres (Amiard S. et al. 2007; Poulet A. et al. 2009), but it can also bind Holliday junctions (HJ) with its amino-terminal domain in vitro, stabilizing the strand-invasion process (Fouche N. et al. 2006; Poulet A. et al. 2009). T-loop homologous recombination has been observed in cells overexpressing a TRF2 mutant lacking the amino-terminal GAR domain (Wang R.C. et al. 2004). These cells show telomere shortening and accumulation of circular telomeric DNA (T-circles) formed by the resolution of HJ generated by branch migration at the strand invasion point of the t-loop. This mechanism is dependent on XRCC3 resolvase and WRN DNA helicase that is responsible for branch migration in "t-loop HR" (Liu Y. et al. 2004; Li B. et al. 2008; Nora G.J. et al. 2010). These evidences demonstrate that Shelterin components and in particular the GAR domain of TRF2 is important to repress HR event at telomeres.

1.3.2.2 Telomeric sister chromatid exchange (T-SCE)

After replication, telomeres of sister chromatids can exchange telomeric sequences in a process known as Telomeric Sister Chromatid Exchange (T-SCE). T-SCE can be visualized in metaphase chromosomes by Chromosome Orientation Fluorescent in Situ Hybridization (CO-FISH) (Palm W. and de Lange T. 2008). (Hagelstrom R.T. et al. 2010). CO-FISH is a strand-specific FISH that involves selective removal of newly replicated strands from DNA of metaphase chromosomes resulting in single-stranded target DNA. The use of two different single stranded telomeric probes (a G-rich and a C-rich probe) guarantees a strand-specific hybridization that discriminates between telomeres produced by leading versus lagging strand synthesis (for details see Results Figure 4.1.8, Material and Methods par. 3.11) (Bailey S.M. et al. 2004). A merged signal of the two probes at the same chromosome end indicates that T-SCE occurred.

T-SCE causes an unequal exchange of telomeres, resulting in the inheritance of shortened telomeres in daughter cells. Since telomerase negative cells show shorter replicative life span, T-SCE has to be strictly controlled (Doksani Y. and de Lange T. 2014).

The Ku70/80 heterodimer has been demonstrated to act as a crucial repressor of HR at telomeres (Celli G.B. et al. 2006; Wang Y. et al. 2009). In addition, components of the shelterin complex repress the exchanges between sister telomeres. The highest frequency of T-SCE has been observed in cells lacking both TRF2 and Ku70 (15-20% of the chromosome ends) (Celli G.B et al. 2006). Single deletion of TRF2 or Ku70 is not sufficient to induce T-SCE, suggesting that the two proteins have a redundant role in suppressing recombination. Increased T-SCE frequency has also been observed in other settings such as in Ku 70/80-deficient mouse cells that also lack Rap1 or both Pot1a and Pot1b (Palm W. et al. 2009; Sfeir A. et al. 2010). How Rap1 or POT1 repress HR has not yet been explained. Rap1/Ku70 double-knockout cells don't show DNA damage signaling at telomeres and undergo HR (Sfeir A. et al. 2010). In addition, the helicase WRN or RTEL1 can repress T-SCE. In fact, murine cells with severely shortened telomeres due to telomerase deficiency show high levels of T-SCE in the absence of WRN. In a similar way, RTEL1 deficient mouse cells or human cells with mutated RTEL1 show high levels of T-SCE (Laud P.R. et al. 2005; Sarek G. et al. 2015). In addition, increased T-SCE has also been observed in mouse cells that lack telomeric heterochromatin components such as the Suv39h1 and Suv39h2 histon methyltransferases (HTMases), DNA methyl transferases (DNMTases) or retinoblastoma family proteins (Garcia-Cao M. et al. 2004; Gonzalo S. et al. 2005; Gonzalo S. et al. 2006).

Altogether these data indicate a major role of shelterin components and DNA helicases in suppressing TSCE. This suggests that an unprotected state of telomeres or the accumulation of secondary DNA structures at telomeres favour telomeric recombination.

1.3.2.3 Homologous recombination with interstitial sites

Telomeres can partake in a third homology-directed repair pathway that involves internal portions of TTAGGG repeats. This pathway can be detrimental for telomeres as it can lead to terminal deletions of all sequence located distant to the interstitial TTAGGG repeats, thereby producing an extra-chromosomal element called double-minute chromosome (TDMs). These TDMs can contain telomeric sequences and have a variable size depending on the location of the interstitial sequences. TDMs have been frequently observed in immortalized mouse embryo fibroblast lacking ERCC1, an endonuclease usually recruited to telomere by the direct interaction with TRF2. On the molecular level ERCC1 prevents recombination of the telomere terminus by promoting the cleavage of the strand-invaded intermediates (Zhu X.D. et al 2003).

In addition, another contributor is the WRN helicase that counteracts TDMs formation in presence of short telomere (Laud P.R. et al. 2005). Finally, reduction of POT1a/b is strictly correlated with a high frequency of TDMs (He H. et al. 2006; Wu L. et al. 2006)

1.4 Physiological relevance of RNA:DNA by hybrids

1.4.1 R-Loops formation and functions

R-Loops are three-stranded nucleic acid structures, consisting of RNA:DNA hybrid and a displaced single-stranded DNA (Santos-Pereira and Aguilera 2015). R-loops were discovered over 40 years ago in bacteria and later were identified in several organisms ranging from yeast to humans (REF.).

R-loops tend to form in regions rich in cytosine on the template strand and guanine template on the non-template strand. An in silico analysis of the human genome has identified more than 250000 putative R-loop forming sequences (Wongsurawat et al. 2012).

To date, the mechanism that leads to the formation of these structures in vivo is unclear, but there are three proposed models (Figure 1.8). The first model postulates that the newly synthesized RNA strand, emerging from the transcription machinery, reanneals with the homologous sequence of the DNA duplex displacing the non-template strand. The G nucleotides in the non-template strand can frequently interact with each other, thus forming secondary structures known as G-quadruplex.

In the second model pairing between DNA and RNA occurs in cis. In this case, the nascent RNA does not detach from the template DNA, but it remains firmly bound to form RNA:DNA hybrid (Westover, Bushnell, and Kornberg 2004).

In the third model, R-loops are formed in trans, post transcriptionally, upon invading of RNA into DNA duplexes with homologous sequence at a different locus from where the RNA was originally transcribed (Toriumi, Tsukahara, and Hanai 2013).

1.4.2 Physiological roles of RNA:DNA hybrids

The R loops cover significant role in diverse biological processes. R- loops promote replication od plasmids in bacteria as well as mitochondrial DNA in human cells (Xu, Baoji; Clayton 1996). RNA:DNA hybrids promote immunoglobulin class switch recombination, which contributes to the antibody isotype diversity in activated B cells (Yu et al. 2003). R-loops can form on many genes in yeast and human cells and have been implicated in regulation of gene expression (Chan et al. 2014; Ginno et al. 2012). R-loops can repress transcription and promote transcriptional termination (Huertas and Aguilera 2003; Skourti-Stathaki, Proudfoot, and Gromak 2011; Tous and Aguilera 2007). Furthermore, R-loops are associated with epigenetic mechanisms such as DNA methylation and post-translational histone modifications (Castellano-Pozo et al. 2013; Ginno et al. 2012; Nakama et al. 2012; Yang et al. 2014)

1.4.3 Molecular factors involved in R-Loops Formation

Several mechanisms have been shown to control the formation of RNA: DNA hybrids.

The first complex, whose importance has been discovered in the formation of R-loops in human cells, is the THO/TRE complex. The THO/TREX complex is responsible for packaging of pre-mRNA with RNA-binding proteins and mRNA export.

This complex consists of four nuclear proteins (Hrp1, Tho2, Mft1 and Thp2) and is associated with the transcription-export complex (TREX), containing Tex1 and the mRNA export factors Sub2 and Yra1 (Chavez et al. 2000; Strasser et al. 2002). The physical interaction of THO complex with TREX directly links mRNA packaging with RNA export. Mutations affecting THO/TREX have been shown to affect transcriptional elongation, proper mRNA export and recombination (Strasser et al. 2002; Rondon et al. 2003). A distinctive phenotype of yeast THO mutants (firstly identified in the *HRP1* and *THO2* genes) is their transcription-associated hyper- recombination phenotype (Aguilera and Klein, 1990; Piruat and Aguilera, 1998) and this is directly associated with R-loop formation (Huertas and Aguilera, 2003).

DNA Topoisomerase I (DNA-TopI) is an evolutionarily conserved factor that suppresses R-loop formation. This enzyme resolves the negative torsion created upstream of RNA polymerase II in order to prevent annealing of nascent RNA with the DNA template (Sordet et al. 2009; Tuduri et al. 2009). In the absence of TopI, negative supercoiling accumulates behind elongating Pol II to promote opening of DNA, which in turn facilitates reannealing between nascent RNA and DNA template strand and subsequent R-loop formation (El Hage et al. 2010; Hraiky, Raymond, and Drolet 2000; Tuduri et al. 2009).

RNase H enzymes have been shown to cleave the RNA component of RNA:DNA hybrids (Stein and Hausen, 1969; Hausen and Stein, 1970; Cerritelli and Crouch, 2009). Most organisms encodes two types of RNase H. Eukaryotic RNase H1 consists of a single protein that uses its N-terminal domain to bind RNA:DNA hybrids (hence named hybrid binding domain-HBD). The C-terminal domain containing the RNase H active site is responsible for RNA cleavage (Cerritelli and Crouch, 2009). RNase H2 is composed of three subunits with RNase 2A as the catalytic subunit. RNase H1 and H2 have different *in vivo* substrates due to their differences in hybrid hydrolysis (REF.). RNases H have also been implicated in nuclear and mitochondrial DNA replication by removing RNA primers (Cerritelli and Crouch, 2009).

Sen1 helicase acts to remove R-loops and to prevent genomic instability by R-loop mediated DNA damage (Mischo et al. 2011). Two recent studies reveal a very interesting connection between transcription termination factors such as Sen1, 3' end processing factors such as Pcf11 and Rna15 ((Mischo et al. 2011; Stirling et al. 2012) and R-loop mediated genomic instability. This suggests a link between genomic instability and 3' end processing/termination factors

In line with this, in higher vertebrates the ASF/SF2, a serine-arginine-rich (SR) protein that regulates the first steps of splicing, interconnects between pre-mRNA processing and genomic instability (Li and Manley, 2005). Li and Manley demonstrated in chicken DT40 cells and human HeLa cells that depletion of ASF/SF2 causes the nascent transcript to form R-loops which in turn promote DNA rearrangements mediated by double-strand breaks (DSBs) (Li and Manley, 2005). In essence, the model proposed is that ASF/SF2 is co-transcriptionally loaded into the nascent pre-mRNA via the phosphorylated CTD of Pol II not only to promote splicing but also to prevent R-loop formation and subsequent genomic instability (Li and Manley, 2005; Aguilera, 2005).

Finally, RAD51 which promotes strand exchange by forming nucleoprotein filaments during homologous recombination, can result in reduced R-loop formation, suggesting that Rad51 likely promotes R-loops formation in *trans* (Wahba, Gore, and Koshland 2013).

1.4.4 R-Loop and DNA Damage

Defects in factors involved in the prevention or resolution of R-loops can lead to the activation of DNA damage pathway and genomic instability. The first experimental evidences derive from the study of yeast mutants in THO/TREX complex, that exhibit a hyper recombination phenotype. Overexpression of RNase H in these mutants causes suppression of the hype-recombinant phenotype, indicating R-loop as a source of DNA damage (Huertas and Aguilera 2003).

Genome-wide screen was used to identify numerous transcription, splicing or mRNA processing factors that suppress accumulation of DNA damage or genome instability in an RNase H-dependent manner (Paulsen et al. 2009; Stirling et al. 2012). Three different mechanisms have been proposed that link R-loop and damage to DNA have been proposed.

The unpaired DNA strand resulting from R-loop formation is susceptible to spontaneous deamination of dC to dU, finally leading to double-strand breaks and recombination. Activation-induced cytidine deaminase (AID) enhances this phenomenon by converting dC into dU residues in single-stranded DNA (U. Basu et al. 2011). This modification makes the DNA calls the action of the base excision repair enzyme uracil DNA glycosylase that excises the uracil base to create an abasic site, which can be processed by apurinic/apyrimidinic endonucleases to create a SSB.

Other structural elements susceptible to activation in the DNA damage pathway are the flaps that form at either end of the R-loop. In human cells, XPF and XPG (pigmentary xeroderma types F and G), two flap endonucleases involved in nucleotide excision repair (NER), can process into DSBs.

Because the formation of DSBs requires specific factors involved in the repair of the transcribed strand, as opposed to those involved in global genomic repair, it is possible to conclude that transcription-coupled NER pathway (TC-NER) that is able to activate DSBs and induce genomic instability (Sollier et al. 2014).

An additional possible scenario suggests that transcriptional R-loops induce genomic instability by interfering with DNA replication (Aguilera, 2002; Houlard et al. 2011; Gan et al. 2011; Aguilera and Garcia-Muse, 2012). This transcription-dependent replication fork collisions have been shown to induce DNA breaks (Prado and Aguilera, 2005; Gottipati et al. 2008; Boubakri et al. 2010). R-loop induced replication blockage could occur by unrepaired DNA lesions in the unpaired ssDNA strand derived from R-loop formation or by RNA/DNA hybrids themselves. Such blocked replication forks could then generate DNA lesions and DSBs in the newly synthesised DNA. These would induce recombination-mediated repair which in turn could lead to chromosome rearrangements and genomic instability (Aguilera and Garcia-Muse, 2012).

1.5 Complex nucleic structures formed at telomere

1.5.1 Telomere repeat containing RNA and R-loops at telomeres

For a long time, telomeres have been considered to be transcriptionally silent. Although telomeres and subtelomeres carry repressive chromatin marks, transcriptional activity has been detected in vertebrates. Telomeric repeat-containing RNA (TERRA) is a long non coding RNA transcribed by RNA polymerase II from the subtelomeric regions, which contain DNA methylation sensitive promoters,

towards the end of the chromosomes. TERRA abundance and its localization are regulated in a cell-cycle-dependent manner, with low level of TERRA occurring in late S phase when telomeres are replicated (Porro et al. 2010).

TERRA molecules are composed of subtelomeres sequences and of varying sizes of UUAGGG repeat sequences. TERRA is heterogeneous in length, ranging from 100 nt to 9 kb in mammalian cells (Azzalin et al. 2007; Schoeftner and Blasco 2008), but it can reach up to 100kb in rodents (De Lange 2009).

As a product of RNA polymerase II, all human and most yeast TERRA 5' ends contain a 7-methylguanosine (m^7G) cap structure; meanwhile only 7% of human TERRA molecules are polyadenylated (Azzalin and Lingner 2008; Feuerhahn et al. 2010). Polyadenylation is a crucial determinant for TERRA localization: polyA- TERRA preferentially associates with telomeric chromatin, while polyA+ TERRA molecules localize to the nucleoplasm (Porro et al. 2010).

Recent studies reveal that TERRA transcribed from a single telomere can associated with multiple chromosome ends in mouse cells. This suggest that TERRA can act in *cis*, maintaining a contact with telomere template, but also in *trans* (De Silanes et al. 2014).

Association of TERRA with telomeres is regulated by proteins of different complexes. The members of nonsense-mediated RNA decay (NMD) such as the RNA/DNA helicase and ATPase UPF1 and the RNA endonuclease SMG6, are considered negative regulators (Azzalin et al. 2007); other proteins such us TRF1 and TRF2, Histone H3K9me3, and MORF4L2, a member of the NuA2 histone acetyltransferase complex, and other heterochromatin marks are essential for TERRA interaction with telomeres.

TERRA interacts with TRF1 and TRF2 shelterin components through TRF2 amino-terminal GAR domain and the carboxy-terminal myb domain (Deng Z. et al. 2009). In addition, TERRA interacts with telomeric heterochromatin components such as Suv39h1 and H3K9me3, HP1, subunits of the origin recognition complex (ORC) and MORF4L2, a member of the NuA2 histone acetyltransferase complex, thereby contributing to heterochromatin formation at chromosome ends (Deng Z. et al. 2009; de Silanes I.L. et al. 2010; Scheibe M. et al. 2013; Porro A. et al. 2014). Alterations of the heterochromatic status of telomeres directly impact on TERRA expression. Indeed, DNA methyl-transferase enzymes DNMT1 and DNMT3b repress TERRA transcription through CpG methylation (Figure 1.13) (Nergadze S.G. et al. 2009). In addition, Suv39h1 H3K9 histone methyltransferase and the H3K9me3-binding protein HP1α negatively regulate TERRA expression (Figure 1.13) (Arnoult N. et al. 2012). On the contrary, the chromatin organizing factor CTCF and the cohesin Rad21 (radiationsensitive 21) act as positive regulators of TERRA transcription in a particular subset of TERRA promoters which comprise a third repetitive element of 61bp tandem repeats bound by CTCF- Rad21 (Figure 1.13) (Deng Z. et al. 2012). Based on the finding that TERRA interacts with several components of telomeric

heterochromatin, a role of TERRA in the control of chromatin status at telomeres has been suggested. In particular, TERRA has been demonstrated to stabilize the interaction between TRF2 GAR domain and the origin recognition complex (ORC1), thus promoting heterochromatin formation and transcriptional silencing at telomere (Sasaki T. and Gilbert D.M. 2007; Deng Z. et al. 2009). In line with this, a negative feedback loop has been proposed for TERRA transcription regulation in which longer TERRA molecules repress their own transcription upon telomere elongation by increasing trimethylation of H3K9 through recruitment of more histone methyltransferases and/or histone deacetylases (Arnoult N. et al. 2012).

TERRA levels are regulated during cell cycle. In human cells, TERRA levels are higher in G1, decreases in S-phase and start increasing again in G2 (Porro A. et al. 2010). During the cell cycle, single-stranded telomeric DNA is bound by POT1/TPP1, but in S-phase ss-DNA may be bound by RPA to allow semiconservative DNA replication. After telomere replication, RPA needs to be removed from telomeric ssDNA to prevent persistent DNA damage activation. TERRA has been demonstrated to orchestrate the switch between RPA and POT1 at telomeres after replication. In particular, TERRA directly interacts in vitro with hnRNPA1 that can displace RPA from telomeric ssDNA (de Silanes L.I et al. 2014; Flynn R.L. et al 2011). An additional role of TERRA in assisting DNA replication is mediate by its interaction with the origin replication complex (ORC) that has a major role in pre-replication complex assembly in addition to heterochromatin formation (Deng Z. et al. 2009).

It has been reported that TERRA inhibits telomerase activity in vitro by binding to both the telomerase RNA (TR) moiety (Schoeftner and Blasco 2008) and the telomerase reverse transcriptase (TERT) polypeptide independently of hTR (Redon, Reichenbach, and Lingner 2010).

Supporting this fact, TERRA was found to precipitate with telomerase in yest and human cell extract, indicating that the TERRA-mediated function in telomerase regulation might be evolutionary conserved (Cusanelli, Romero, and Chartrand 2013; Redon, Reichenbach, and Lingner 2010). Indeed, elegant in vitro experiments with TERRA-like oligonucleotides revealed the unique ability of TERRA to inhibit mammalian telomerase (Redon, Reichenbach, and Lingner 2010; Schoeftner and Blasco 2008).

1.5.2 TERRA and telomeric R-Loop formation

Due to its high GC content and presence of transcriptional activity, telomeric DNA repeats are proposed to be prone to form secondary structures (G-quadruplex) and RNA:DNA hybrids.

TERRA was shown to remain base-paired with telomere DNA leading to the formation of R-loops (Balk et al. 2013; Chawla and Azzalin 2009).

R loops formation hinders DNA replication, causing the replication-fork stalling and collapse. Double-strand breaks generated by collapsed replication forks and associated nucleolytic processing could therefore take into account for the telomere dysfunction phenotypes that are associated with impaired TERRA displacement or increased TERRA transcription (Gómez-González et al. 2011).

Thus, cells can regulate R-loops formation via activation of different enzymes such as RNase H1 that degrades the RNA part of telomeric RNA:DNA hybrids (Aguilera and García-Muse 2012), helicases, which unwind RNA:DNA hybrid structures (Paeschke et al. 2013) and factors involved in RNA biogenesis such as the THO/TREX protein complex (Pfeiffer et al. 2013).

Luke at al. have demonstrated that TERRA, through the formation of telomeric RNA:DNA hybrids (or Telomere R-Loops, TRLs), is able to regulate telomere-length dynamics in presenescent cells.

In ALT cancer cells the elevated TERRA transcription and the reduced compaction of telomeric chromatin, leads to an increase in TRLs that are instrumental in facilitating recombination at telomeres.

Several studies have been demonstrated an important role for ATRX (ATP-dependent helicase ATRX, X-linked helicase II) in the management of TRLs in ALT tumors cells. ATRX belongs to SWI/SNF family of chromatin remodelers that form a complex with the chaperone protein DAXX (Death Domain Associated Protein), and deposits the variant Histone H3.3 at telomeric regions (Lewis et al. 2010). ATRX is frequently mutated in ALT tumors and this is linked with the accumulation of TERRA at telomeres in G2-M phase of the cell cycle. ATRX deals with the maintenance of telomeres through different mechanisms including cell cycle regulation of TERRA, RPA removal from telomeres following replication to avoid recombinogenic structure and C-circle excision (Chu et al. 2017). In addition, ATRX promotes telomere chromatin cohesion dissolutionto reduce HR between sister chromatids following the generation of TRLs and disassembles aberrant secondary structures such as G-quadruplex (Ramamoorthy and Smith 2015; Clynes et al. 2015).

Together this suggests that telomeres of ALT take advantage of R-loops to promote telomere maintenance by HR. In this scenario, the regulation of the TERRA:DNA hybrid management has high relevance for telomerase negative cancer cell

2. AIM OF THE THESIS

TERRA, a long non coding RNA trascribed from telomeres, has a fundamental role in telomere homeostasis. In this study we aimed to explore new TERRA/Telomere interaction proteins in understand their role in telomere regulation in cancer cells.

In particular, my PhD thesis aimed:

- To identify novel TERRA interactors involved in telomere regulation and homeostasis (project 1)

- To study the role of TERRA interactors in RNA:DNA hybrids formation at telomeres (project 2)

3. MATERIALS AND METHODS

3.1 Cell lines and culture

Cell lines used were obtained from ATCC and have not been cultured for longer than 6 months. In this project U-2 OS (osteosarcoma) cells were cultured in low glucose Dulbecco's modified Eagle's (DMEM) medium (Lonza) with 10% fetal bovine serum (Gibco), 1% L- glutamine (Gibco), 1% penicillin/streptomycin (Gibco) .

U-2 OS cells transduced with the retroviral vector pLPC or pLPC-FLAGNONO vectors were selected with puromycin (1µg/ml). These cells transfected with linearized pLPC (control) vector or with a combination of linearized pLPC and pCMVmyc-SFPQ-WT (addgene) vectors were selected using puromycin (1µg/ml). Treatments: hydroxyurea (HU) (Sigma), 5 mM for 6 hours; 5-Aza-2'deoxycytidine (Sigma) 10µM for 72 hours.

3.2 siRNA and plasmids transient transfection

For siRNAs transfection RNAi-MAX Lipofectamine (Invitrogen) was used according to the manufacturer's suggestions. siRNAs listed in Supplementary Table 1 have been transfected at a final concentration of 30 nM for 72 hours. For plasmids transfections Lipofectamine 2000 (Invitrogen) was used according to the manufacturer's suggestions.

siRNA or RNA-oligonucleotides used to transiently transfect cells are the following:

Human, ON-TARGETplus non-targeting siRNA#1 (Dharmacon)

ON-TARGETplus smartpool TRF1 siRNAs (Dharmacon)

ON-TARGETplus smartpool SFPQ siRNAs (Dharmacon)

ON-TARGETplus smartpool RAD51 siRNAs (Dharmacon)

miRIDIAN microRNA mimic negative control siRNA (Dharmacon)

3.3 Retroviral transduction of human cells

HEK 293GP (Gag and Pol) were used as packaging cell line. The solution containing DNA and $CaCl_2$ (2,5 M $CaCl_2$ sterilized by filtration) for a 10cm dish was prepared adding sequentially:

H_2O up to 450 ul

50ul $CaCl_2$

5-10 g of specific plasmid DNA.

In a different tube 500 µl of 2X Hepes Buffer Saline were added and the DNA solution was transferred drop by drop to the 2X HBS tube and mixed by introducing air bubbles and incubated for 20-30

minutes in order to allow the efficient formation of DNA-CaPO4 precipitates. After the incubation, this solution was added to cells and incubated at 37 °C. for 6 hours, after which the medium was changed and cells were incubated for additionally 48 hours. Subsequently, the supernatant containing the virus was collected, filtered and mixed with 1ml of fetal bovine serum (FBS) and 8 μg/ml of polybrene. This solution was used to infect selected cells. The day after, medium was changed. After 24 hours the specific antibiotic to select transfected cells was added.

3.4 Generation of SFPQ mutants

SFPQ mutants have been generated using the QuikChange II XL Site-Directed Mutagenesis Kit (Agilent). To insert the two mutations within the SFPQ sequence, we designed a pair of primers (forward and reverse) for each insertion. The forward primer is complementary to the wild-type sequence of SFPQ in the insertion region. It was designed with the addition of the nucleotides in order to create the desired mutation t in the correct position (Figure N). The reverse primer, was designed to complement the forward primer. During the PCR cycles the primers amplify the vector in opposite directions thus obtaining the complete sequence of SFPQ containing the insertion.
The protocol was performed according to the manufacturer's suggestions and the specific primers are listed below:

Oligo name	Sequence
Myc-SFPQ Ins 448-449	AGGGAGGGAGGGGCAAACCACACCCAGGCCC
Myc-SFPQ Ins 442-443	GGGCCTGGGTGTGGTTTGCCCCTCCCTCCCT

3.5 Protein extracts and Western blotting

The whole-cell lysates were prepared using a modified RIPA buffer (20 mmol/L Tris-HCl (pH 7.5) 350 mmol/L NaCl, 1 mmol/L Na_2EDTA, 1 mmol/L EGTA, 1% NP-40, 1% sodium deoxycholate, 2.5 mmol/L sodium pyrophosphate, 1 mmol/L b-glycerophosphate, 1 mmol/L Na_3VO_4, 1 mg/mL leupeptin).
In order to prepare nuclear extracts, cells were resuspended in Buffer 1 (20 mmol/L Hepes-KOH (pH 7.9), 10 mmol/L KCl, 1.5 mmol/L $MgCl_2$, 1 mmol/L EDTA, 1 mmol/L EGTA, 0.2% NP40) and incubated for 10 minutes on ice. Nuclei were pelleted and lysed in Buffer 2 (20 mmol/L Hepes-KOH (pH7.9), 350 mmol/L NaCl, 1.5 mmol/L $MgCl_2$, 10 mmol/L KCl, 10% glycerol, 1 mmol/L dithiothreitol).
Both whole cell lysates and nuclear extracts have been additioned with complete protease inhibitor (Roche) and sonicated. After centrifugation, supernatants have been recovered and used for Western blotting according to standard procedures.

The primary antibodies used are listed in the antibodies table 3.24.

Membranes were incubated with the specific secondary antibodies bound to the HRP enzyme (horseradish peroxidase-conjugated antibody) (GE Healthcare) and the levels of protein expression were detected by chemiluminescence using the ECL system (GE Healthcare) with subsequent exposure on autoradiography film (GE Healthcare).

3.6 Immunofluorescence

Cells were washed with 1x PBS and fixed in 4% paraformaldehyde (PFA) for 15 minutes.

Subsequently cells were washed with 1x PBS, treated with citrate buffer (0.1% sodium citrate, 0.05% Triton X-100) for 10 minutes at room temperature and washed with 1x PBS, 0,1% Tween-20. Cells were blocked for 1 hour in 3% BSA (1X PBS, 0.1% Tween-20) and incubated with primary antibodies (check 2.23 antibodies table) diluted in blocking solution for 2h at room temperature in a wet chamber. Cells were washed three times in 0.3% BSA (1X PBS, 0.1% Tween-20) for 5 minutes and incubated with Alexa-fluor secondary antibodies (Invitrogen) (1:500) diluted in washing solution for 1h in a wet chamber. After incubation, slides were washed three times in 0.3% BSA (1X PBS, 0.1% Tween-20) for 10 minutes. Afterwards, DAPI (Vector Laboratories) was added in order to stain the nuclei and slides were incubated for 5 minutes. Slides were mounted with Vectashield.

Images were captured using classic immunofluorescence (Leica DM4000B) microscope. For classic immunofluorescence analysis, the number of co-localizations per focal plane was counted. Quantitative immunofluorescence analysis were performed using ImageJ. The Student t test was used to calculate the statistical significance.

3.7 Interphase Telomere Q-FISH

Cells were briefly washed in 1X PBS, fixed with 4% formaldehyde (PFA) for 2 minutes at RT and washed three times for 5 minutes with 1X PBS. Slides were dehydrated with three successive washes in 70%, 90%, and 100% ethanol for 5 minutes each and let to airdry for 20 minutes at RT. Afterwards, PNA-telomere probe (Table 1) were added to each slide and cover with a coverslip, the slides were denaturated at 80 °C for 3 minutes exactly.

The slides were incubated for 2 hours at RT in the dark in a wet chamber. After the incubation, slides were washed in mild agitation twice for 15 minutes with FISH solution (Formamide 70%, Tris pH 7.2 10mM, BSA 0.1%, H2O) at RT, followed by three washes of 5 minutes in 1X TBS with 0,01% of Tween-20 at RT. Nuclei were stained with DAPI included in a fourth wash.

The slides were then dehydrated with 3 washes in 70%, 90%, 100% of ethanol, left to airdry for 20 minutes and mounted with Vectashield. Images were captured using Leica DM4000B microscope.

Telomere signals intensity of interphase nuclei was analyzed using TFL-TELO software. The Student t test was used to calculate the statistical significance.

Stock	PNA-Telomere Probe (1 slide)	Final Concentration
1M Tris pH 7.2	0.25 μl	10 mM
Buffer MgCl pH 7.0 (25 mM MgCl, 9 mM citric acid, 82 mM Na2HPO4)	2.14 μl	
Deionized Formamide	17.5 μl	70%
probe 25 μg/ml	0.5 μl	0.5 μg /ml
Blocking Reagent 10%	1.25 μl	0.5%
H_2O	3.36 μl	

Table 1. PNA-telomere probe solution

3.8 Immunofluorescence combined with RNA FISH

Cell were permeabilized with cytobuffer as described for RNA-FISH and then subjected to the immunofluorescence protocol. Used primary and secondary antibodies are indicated Supplementary table 2-3. After the washing step for the secondary antibody, cells were fixed at room temperature for 2 minutes with 4% paraformaldehyde (PFA), dehydrated and RNA-FISH was performed. Co-localization of foci from FISH and immunostainings were quantified by visual inspection.

3.9 Chromosome Orientation FISH (CO-FISH)

Confluent experimental cells were subcultured in the presence of 5'-bromo-2'-deoxyuridine (BrdU, 10μM; Sigma) and cultivated for 18-20 hours. Colcemid was added during the final period of BrdU treatment (U-2 OS, 0.2μg/m 3 hours; H1299, 1μg/ml for 2 hours). Cells were recovered and metaphases prepared as previously described 54. CO-FISH was performed as previously described (Bailey et al. 2001; Benetti, García-Cao, and Blasco 2007)using first a (CCCTAA)3 probe labeled with Cy3 and then a second (TTAGGG)3 probe labeled with FITC. Samples were processed as described for DNA-FISH.

3.10 ChIP assay and telomere dot-blots

Cells were fixed adding formaldehyde to culture medium to a final concentration of 1% and incubated for 15 min at RT in mild agitation. The addition of glycine to a final concentration of 0.125 M stopped the cross-linking. Crosslinked cells were washed twice with cold 1X PBS, scraped and lysed for 10 minutes at 4 °C in a lysis buffer containing 1% SDS, 50 mM Tris-HCl (pH 8.0), 10 mM EDTA and supplemented with protease inhibitors.

Lysates were sonicated to obtain chromatin fragments >1 kb and centrifuged for 15 min in a microcentrifuge at room temperature. Chromatin was diluted 1:10 with 1.1% Triton X-100, 0.01% SDS, 1.2 mM EDTA, 167 mM NaCl and 16.7 mM Tris-HCl (pH 8.0) containing protease inhibitors and precleared with Protein A or Protein G agarose beads

(Santa Cruz). Chromatin was incubated with 3,5 µg of specific antibodies (check 3.23 antibodies table) or rabbit/mouse IgG as negative control rocking overnight at 4°C. Protein A or Protein G agarose beads were then added and the incubation continued for 1 h. Immunoprecipitated pellets were washed with 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl (pH 8.0) and 150 mM NaCl (one wash); 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl (pH 8.0) and 500 mM NaCl (one wash); 0.25 M LiCl, 1% NP-40, 1% sodium deoxycholate, 1 mM EDTA and 10 mM Tris-HCl, pH 8.0 (one wash); and 10 mM Tris-HCl (pH 8.0) and 1 mM EDTA (two washes). Chromatin was eluted from the beads twice by incubation with 250 µl of 1% SDS/ 50mM $NaHCO_3$ for 15 min rocking at room temperature. After adding 20 µl of 5 M NaCl, crosslinks were reversed by incubation for 4 h at 65 °C. Samples were supplemented with 20 µl of 1 M Tris-HCl (pH 6.5), 10 µl of 0.5 M EDTA, 20 µg of RNase A and 40 µg of proteinase K and incubated for 1 h at 45°C. DNA was recovered by phenol-chloroform extraction, ethanol precipitation and spotted into a dot-blot with a Hybond N+ membrane which was hybridized with a telomeric probe obtained from a plasmid containing 1.6 kb of TTAGGG repeats. For total DNA samples, aliquots corresponding to a 1:10 dilution of the amount of lysate used in the immunoprecipitations were processed along with the rest of the samples during the crosslink reversal step.

3.11 Mass-spectrometry analysis

Cell lysates were prepared with ice-cold lysis buffer containing 50 mM Tris-HCl, pH 8, 150 mM NaCl, 5mM EDTA, 1mM DTT, 1mM PMSF, 1 mM beta-glycerophosphate, 1mM Na3VO4, 5 mM Na2F, and protease inhibitor cocktail (Sigma) plus 5% glycerol and 1% NP-40. After centrifugation and preclearing, lysates were incubated at 4 °C with specific antibodies or rabbit IgG as negative control (check 2.23 antibodies table). After 2 hours, protein-A agarose-beads (Santa Cruz) were added to

each IP and incubated overnight. The resin was then washed and bound proteins were eluted in sample buffer 2X.

The immunoprecipitates were separated by SDS-PAGE and a mass spectrometry-compatible silver staining (SilverQuest™ Kit, Invitrogen) or Coomassie Brilliant blue R-250 (42660, Sigma) was used to detect proteins on gel. Selected lanes were excised from the gels, de-stained and washed twice in 50% ACN with 50 mM ammonium bicarbonate, dehydrated in 100% ACN. In-gel reduction was performed by incubating lanes in a 10 mM DTT, 100 mM ammonium bicarbonate solution for 30 min at 56°C and alkylation in 55 mM iodoacetamide, 100 mM ammonium bicarbonate for 20 min at room temperature. Protein samples were dehydrated again in 100% ACN and digested by rehydrating gel pieces in 50 mM ammonium bicarbonate containing 4 ng/uL of trypsin (#V5111, Promega) overnight at 37°C. Tryptic peptides were desalted and concentrated using ZipTip mC18 pipet tips (Millipore) and co-eluted onto the MALDI target in 1 µL of α-cyano-4-hydroxycinnamic acid matrix (5 mg/mL in 50% ACN,0.1% TFA).

Mass spectra were acquired over a mass range of 800–4000 m/z (Nd:YAG laser at 355 nm, 40 Shots/Sub-Spectrum for 2,000 Total Shots/Spectrum) by reflectron positive mode on an Applied Biosystems 4,800 Proteomics Analyzer mass spectrometer (Applied Biosystems) and calibrated using a standard mixture (Mass Standards Kit, AB SCIEX). MS/MS spectra were acquired in positive mode (Nd:YAG laser at 355 nm, 40 Shots/Sub-Spectrum for 4,000 Total Shots/Spectrum) and MS/MS calibration was achieved by using the default calibration method.

Protein identifications were performed with the ProteinPilotTM software (version 2.0.1; Applied Biosystems) using the ParagonTM algorithm as the search engine. Each MS/MS spectrum was searched against an Uniprot/SwissProt database of mouse. The search parameters allowed for cysteine modification by iodoacetamide and biological modifications programmed in the algorithm (i.e., phosphorylations, semitryptic fragments, etc.). The detected protein threshold (ProtScore) in the software was set to 1.3 to achieve 95% confidence interval.

The analysis of Protein-Protein Interaction (PPI) network was performed by querying the STRING database version 10.5, through the Cytoscape STRING app (Szklarczyk et al., NAR 2017). Only the STRING experimentally verified PPIs, showing a confidence threshold >= 0.4, were brought forward in the analysis. Functional enrichment for Gene Ontology (GO) terms, included in the Biological Process (BP), Cellular Component (CC) and Molecular Function (MF) domains, KEGG Pathways, Pfam and InterPro classification systems, were performed by the Cytoscape STRING Enrichment app. Finally, Cytoscape clusterMaker app was used for clustering analysis by applying the MCL algorithm (Enright et al., NAR 2002) to the PPI network, with the inflation parameter set to 2. The PPI networks were visualized within Cytoscape and after manually modified.

3.12 Northern blot

To perform Northern blot analysis total RNA (10 µg) was loaded onto 1.2% formaldehyde agarose gels and separated by electrophoresis. RNA was then transferred to Nylon membranes overnight and blots were incubated overnight rocking at 65°C with a probe of 450 bp labelled with ^{32}P using the Amersham Rediprime II DNA Labeling System (GE Healthcare) that specifically hybridizes to telomerase RNA component, obtained from a vector containing the gene encoding for hTR. Afterwards membranes were washed twice with Wash Buffer (20 mM $NaPO_4$ pH 7.2, 1% SDS) rocking for 20 minutes at 65°C.

Subsequently, the levels of hTR expression were detected by exposure on autoradiography film (GE Healthcare). GAPDH was used as a loading control.

3.13 DNA:RNA immunoprecipitation (DRIP)

The purified R-loops obtained from U-2 OS cells transiently transfected with siRNAS for DAXX, SFPQ ad control, were incubated in 300 uL of interaction buffer (Tris 10 mM pH 8.0, NaCl 140 mM, 0,5% Triton). 100 uL were kept as Input (fraction IN, Figure 2B). The remaining 200 uL , 100 uL were incubated directily with 4 ug of S9.6 antibody coupled to 40 uL of magnetic beads at 4°C with agitation, the other 100 uL were treated with RNase H1 and thet incubated with 4 ug of S9.6 antibody coupled to 40 uL of magnetic beads at 4°C. After an hour incubation, the beads were collected and the unbound fraction (FT) was kept. The beads were washed 3 times with 1 mL of interaction buffer (5' incubation each time) and then resuspended in 200 uL of incubation buffer. After 5' incubation, the 200 uL were collected and kept as fraction W4. The beads were resuspended in 200 uL of elution buffer (50 mM Tris pH 8.0, 10 mM EDTA, 0,5% SDS) contaning 50 ug of Proteinase K and incubated at 55°C. After 30', the beads were eliminated and the supernatant recovered (IP fraction). The IN, FT, W4 and IP fractions were then purified using standard phenol/chloroform and ethanol precipitation techniques and resuspended in 22 uL of Tris 10 mM pH 8,0. 6 uL were used for dot blot analysis with the S9.6 (DNA:RNA hybrids), J2 (RNA:RNA hybrids) or ab27156 (dsDNA) antibodies. 12 uL were run on a 0,8% agarose TBE 1X gel. The different fractions were then diluted 1/10,000 for dot blot analysis

3.14 Antibodies table

Antibodies	company	WB	IF	IP	Chip
Mouse anti–Actin	Sigma, A2228	1:10000			
Rabbit anti-pATR (S428)	Cell Signalling, 2853	1:1000	1:50		
Mouse anti-DNA-RNA Hybrid (S9.6)	Kerafast, ENH001		1:150		
Mouse anti-FUS/TSL (4H11)	Santa Cruz, sc-47711	1:1000			
Mouse anti- γH2AX (ser139) (clone JBW301)	Millipore, 06-70	1:2000	1:200		
Rabbit anti-global histone H3	Abcam, ab1791	1:10000			3.5 μg
Goat anti-Lamin A/C (clone N-18)	Santa Cruz, sc-6215	1:1000			
Rabbit anti-NONO	A300-582A, Bethyl lab	1:5000	1:200	2 μg	3.5 μg
Mouse anti-p53 (DO-1)	Santa Cruz, sc-126	1:1000			
Rabbit PML (H-238)	Santa Cruz, sc-5621		1:100		
Rabbit anti-SFPQ/SFPQ	A301-321A, Bethyl lab	1:5000		2 μg	3.5 μg
Mouse anti-SFPQ (clone B92)	Sigma, P2860	1:5000	1:150		
Rabbit anti-phospho RPA32(S33)	A300-246A, Bethyl Lab		1:800		
Rabbit anti-TRF1(N-19)	Santa Cruz, sc-6165-R		1:100		
Mouse anti-TRF2 (4A794)	Millipore, 05-521	1:500	1:200		3.5 μg
Rabbit anti-RAD51 (H-92)	Santa Cruz, sc-8349		1:200		
Rabbit anti-ATRX		1:200	1:200		
Rabbit anti-DAXX		1:200	1:200		
Rabbit antiH3.3					

4. RESULTS

4.1 Project 1: Identification of novel TERRA interactors that control mammalian telomeres

4.1.1 Novel TERRA interaction proteins

In order to better understand the role of TERRA at telomeres and identify novel TERRA interacting proteins, a TERRA RNA Pull down was performed. The experiment was carried out using Embryonic mouse stem cells that maintain telomere length both through telomerase-dependent mechanisms and *via* independent telomere maintenance pathways (mESC).

Biotinylated RNA synthetic oligo containing $[UUAGGG]_6$ repeats (biotin-r-$[UUAGGG]_6$) or control oligonucleotides were incubated with nuclear extract of mESC and then recovered using strepavidin-coated agarose beads (Figure 4.1.1 A). Eluates from TERRA and control RNA Pull down were run on a polyacrylamide gel. Silver staining showed specific enrichment of 13 proteins in biotin-r-$[UUAGGG]_6$ eluates (Figure 4.1.1 B, bands labeled with *) that were subsequently identified by mass spectrometry (Proteomics unit, Istituto Regina Elena IRE/IFO). This analysis revealed several reported TERRA interacting proteins, including a series of heterogeneus nuclear ribonucleoproteins (hnRNPs) (De Silanes, D'Alcontres, and Blasco 2010) and Fused in Sarcoma/TranSlocated in Liposarcoma (FUS/TSL) (Takahama et al. 2013). Importantly, besides the known interactors, two TERRA interacting proteins have been identified: polypyrimidine tract-binding protein-associated splicing factor (SFPQ/SFPQ) and the nuclear RNA-binding protein 54 kDa (NONO in human, NONO in mouse). We also found an enrichment of DAZAP1 (DAZ associated protein 1), an RNA-binding protein involved in spermatogenesis that is not subject of this study (Hsu et al. 2008).

Interaction of NONO and SFPQ with TERRA were confirmed by western blot analysis using eluates obtained from TERRA pull-down experiments using extracts from mESCs (Figure 4.1.1 C) and human H1299 non-small cell lung carcinoma cells (Figure 3.1.1D). As expected, actin was not detected in biotin-r[UUAGGG]6 or biotin-r[EGFP] eluates. We conclude that SFPQ and NONO are novel TERRA interacting proteins.

NONO and SFPQ function is reported to depend on the obligatory formation of homo- or heterodimers of both proteins[43]. In line with this, immunoprecipitation experiments using extracts from telomerase negative U-2 OS osteosarcoma cells that maintain telomeres *via* the recombination based "alternative lengthening of telomeres" (ALT) pathway, confirm a direct interaction between NONO and SFPQ in the cell model systems used in this study (Fig 4.1.1 E).

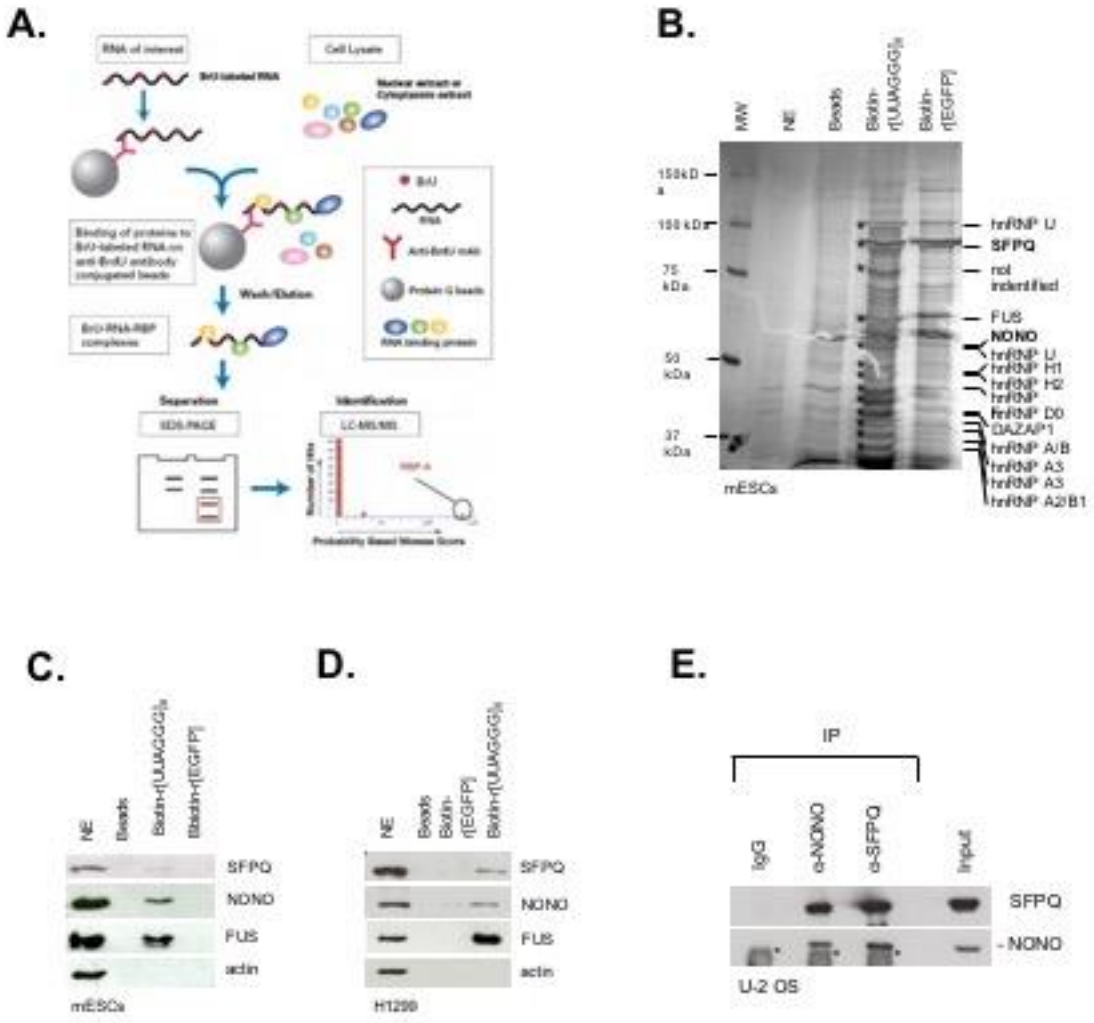


Figure 4.1.1. PTB-associated splicing factor (SFPQ) and nuclear RNA-binding protein 54 kDa (NONO) interact with TERRA. A. RNA pull-down assay. Nuclear extracts from mESCs were incubated with biotinylated r[UUAGGG]$_6$ or biotinylated r[EGFP] RNA oligonucleotides. Complexes were recovered using streptavidine beads, washed, eluted and subjected to SDS-PAGE followed by Silver staining and mass spectrometry analysis. B. Silver staining of SDS-PAGE gels after mono-dimensional electrophoresis of eluates obtained from RNA-pull down experiments using biotinylated r[UUAGGG]6, biotinylated r[EGFP] RNA oligonucleotides or empty beads. Candidate TERRA interacting proteins identified by mass spectrometry were indicated. Pull-down experiments identify SFPQ and NONO as novel TERRA interacting protein. C-D. Western blotting of RNA pull-down eluates using specific anti-NONO and anti-SFPQ antibodies confirms binding specificity of NONO and SFPQ for UUAGGG RNA repeats in mESCs (C) and H1299 cells (D). FUS western blotting was used as positive control and confirms specificity of pull-down. Actin was used as a loading control. E. NONO and SFPQ form a complex as demonstrated by anti-NONO and anti-SFPQ immunoprecipitation experiments. Control Immunoglobulins (IgG) were used.

4.1.2 SFPQ and NONO structure and functions

4.1.2.1 SFPQ and NONO structure

SFPQ and NONO are two proteins belong to the conserved family of "Drosophila behavior / human splicing" (DBHS) proteins. This family of proteins includes three members: SFPQ, NONO and the

paraspekle protein component 1 (PSPC1), remarkably, the latter protein has not been detected in eluates from TERRA RNA pull down experiments.

These proteins have a conserved compound structure N-Terminal RNA Recognition motifs (RRMs) NonA/paraspeckes domain (NOPS) and C-terminal coiled-coil. NOPS and RRMs domain are involved in the formation of homodimers and heterodimers among the DBHS proteins (Passon et al. 2012).

The DBHS proteins show nuclear localization but are enriched in subnuclear bodies termed "paraspeckes" on chromatin or DNA damage foci (Shav-Tal and Zipori 2002b).

Paraspeckles are subnuclear bodies formed by the association of the NEAT1 (nuclear-enriched abundant transcript 1) long non-coding RNA (lncRNA) with DBHS proteins (Bond and Fox 2009). A major function of Paraspekles consists in gene regulation by nuclear retention of mRNAs, a mechanism involved in cellular processes such as stress responses, viral infection and circadian rhythm (Fox and Lamond 2010).

4.1.2.2 Role os SFPQ and NONO in RNA metabolism, transcription regulation and DNA repair

Initial studies demonstrated a role for SFPQ and NONO in splicing (A. Basu et al. 1997; Patton et al. 1991). However, subsequent studies have linked SFPQ and NONO with a wide range of cellular activities such as RNA metabolism, transcriptional regulation and DNA repair (X. Dong et al. 2007; Xuesen Dong et al. 2005; Montes et al. 2012; Yadav et al. 2014).

SFPQ was first identified in a stable complex with polypyrimidine tract-binding protein (PTB), required for pre-mRNA splicing (Patton et al. 1993). Subsequent proteomic and biochemical studies demonstrated a role of SFPQ in the second catalytic step of splicing (exon joining) of some pre-mRNA substrates, thus suggesting a role of SFPQ as alternative splicing regulator (X. D. Fu and Ares 2014). Accordingly, recent works have demonstrated the role of SFPQ in repressing or promoting exon inclusion in the final mRNA of some specific pre-mRNAs (Cho et al. 2014; K. K. Kim et al. 2011; Ray et al. 2011). In addition to a direct regulation of the spliceosome, SFPQ may impact on splicing through its effect on transcription and 3'-end processing Actually, SFPQ and NONO have been identified in a complex containing the U1A protein (distinct from the U1snRNP) important to couple splicing and polyadenylation at suboptimal polyadenylation sites (Knott, Bond, and Fox 2016). Furthermore SFPQ/NONO recruits the exonuclease XRN2 to facilitate pre-mRNA 3'-end processing and transcription termination (Jaafar et al. 2017). In addition to splicing and 3'-end processing, the SFPQ/NONO complex is involved in permitting nuclear retention of selected mRNAs in paraspeckles. The signal for nuclear retention is established by ADAR (Adenosine De-Aminase RNAspecific), RNA-editing enzymes that catalyze the deamination of adenosines (A) to inosines (I) in the mRNAs to be

retained. The SFPQ/NONO complex has high affinity for hyperdited RNAs and retains these RNAs within paraspeckles, thus preventing their export to the cytoplasm (Zhang Z. and Carmichael G.G. 2001; Chen L.L. and Carmichael G.G. 2009). In conclusion, SFPQ/NONO complex is able to regulate gene expression impacting at least on three coupled steps of RNA processing: splicing, 3'-end processing and nuclear export. Interestingly, SFPQ and NONO are also implicated in DNA-mediated nuclear processes such as transcription and DNA repair.

As transcription regulators, SFPQ and NONO regulate both repression and activation. The transcriptional repression of DBHS proteins appears largely driven and dependent on SFPQ, both in a homodimer, or heterodimer context. Through the recruitment of epigenetic regulators, SFPQ / NONO can act on hormone receptors such as the thyroid and retinoid X receptors (Mathur, Tucker, and Samuels 2001), or in complex with the steroidogenic factor 1 (SF-1), to repress the human CYP17 gene or genes involved in circadian rhythms (Hurley, Loros, and Dunlap 2016; Sewer et al. 2002).

The transcriptional activation of DBHS proteins seems to be guided by NONO and, in many cases, involves the processing of the nascent RNA transcription, which probably presents itself as a scaffold for DBHS Association and interaction with transcriptional machinery.

In addition, the SFPQ/NONO complex specifically binds to the RNA polymerase II (Pol II) C-terminal domain and enhance transcription by forming a bridge between Pol II transcription machinery and other splicing or polyadenylation factors (Emili A. et al. 2002).

DBHS proteins are implicated in double-stranded break (DSB) repair where they assist in homology directed repair or nonhomologous end joining (NHEJ).

Experiments in vitro demonstrated that SFPQ alone binds to ssDNA and dsDNA to promote D-loop formation between ssDNA and dsDNA and shows DNA pairing activity of complementary ssDNAs (Akhmedov A.T. and Lopez B.S. 2000). In addition, SFPQ binds directly to DSBs through its N-terminal region (RGG domain and prolin-rich domain) (Ha K. et al. 2011; Morozumi Y. et al. 2009; Rajesh C. et al. 2011). The same region is also involved in the interaction with RAD51D, a member of RAD51 family that plays an integral role in preserving genomic stability by homologous recombination-directed repair (Rajesh C. et al 2011). Notably, SFPQ activates strand invasion and stimulates the repair activity of RAD51 complex, thus promoting HR (Akhmedov A.T. and Lopez B.S. 2000; Morozumi Y. et al. 2009). Consistent with this, deficiency of SFPQ leads to sister chromatid cohesion defects, chromosome instability and increased sensitivity to DNA damaging agents (Rajesh C. et al 2011). Collectively, the DBHS proteins promote end joining of homologous DNA by direct interaction with DNA ends and recruitment/stabilization of a preligation complex (125). The role of NONO in DNA damage response is supported by a NONO knockout mouse model (Li S. et al. 2014). Murine embryonic fibroblast (MEFs) derived from wild-type and knockout of NONO, show the same growth rates and cell cycle

distribution in absence of genotoxic stress, but show increased sensitivity to radiation. Remarkably, NONO knockout correlates with increased level of the DHBS protein family member PSPC1, that appears to compensate the lack of NONO by interacting with SFPQ. Cells depleted for both NONO and PSPC1 become markedly radiosensitive and show delayed DNA DSB repair (Li S. et al. 2014). Along these lines of evidence, a very recent work demonstrates a role of NONO in triggering the intra-S-phase checkpoint in response to UV-induced DNA damage. In particular, depletion of NONO in HeLa cells reduces the chromatin loading of TOBP1 and prevents full activation of ATR upon UV-induced DNA damage. Despite DNA damage, NONO-depleted cells continue to synthesize DNA and fail to block new origin firing (Alfano L. et al. 2015).

In conclusion, NONO and SFPQ are examples of how splicing and gene regulation factors can act as gatekeepers of genome integrity (Naro C. et al. 2015).

4.1.3 SFPQ and NONO localize to telomere

In order to understand if SFPQ and NONO act as telomeric regulatorwe studied their location at the telomere.

Confocal microscopy revealed a significant fraction of nuclear restricted SFPQ and NONO foci co-localizing with telomere repeat binding factor 2 (TRF2) or telomere repeat binding factor 1 (TRF1) in U-2 OS interphase cells (Figure 4.1.4 A-C).

To obtain independent evidence for the enrichment of NONO and SFPQ at telomeres, we performed telomeric chromatin immunoprecipitation (ChIP) experiments using specific antibodies against NONO and SFPQ (Figure 4.1.4 D-E). Mouse/rabbit immunoglobulins from preimmunesera (IgG M, IgG R), TRF2 and H3 antibody have been used, respectively ,as negative and positive control. DNA was purified from immunoprecipitated chromatin, transferred to a nitrocellulose membrane and subjected to Southern blotting using a radioactive labelled telomere probe. We show that NONO and SFPQ are components of telomere chromatin, thus supporting *data* from confocal microscopy. We also excluded binding to AluY repeats. As expected, TRF2 and the TERRA interactor FUS are enriched at telomeric repeats. Together, these results provide evidence for the binding of NONO and SFPQ to telomeres.

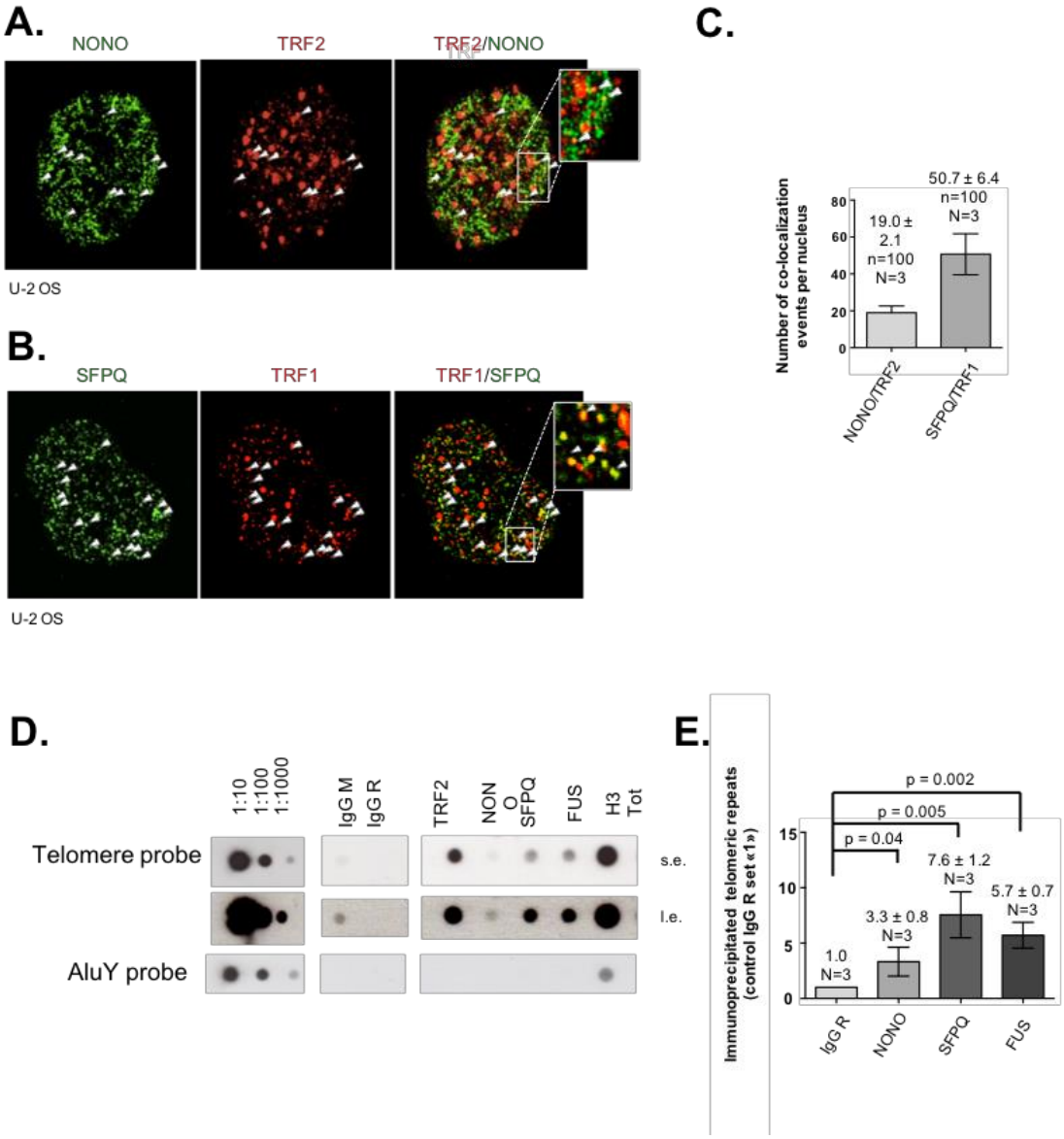


Figure 4.1.4. SFPQ and NONO localize to telomeres. A.-B. Representative image of co-localization events (arrowheads) between NONO and SFPQ and the shelterin proteins TRF2 and TRF1, respectively. C. Quantification of (A) and (B). Mean number of NONO-TRF2 and SFPQ-TRF2 co-localization events per nucleus is shown. Error bars indicate standard deviation. n= number of analyzed nuclei. D. Chromatin immunoprecipitation experiments (ChIP) using U-2 OS cells and mouse anti-TRF2, rabbit anti-histone H3, rabbit anti-FUS, rabbit anti-NONO and rabbit anti-SFPQ antibodies. Mouse and rabbit control IgGs (IgG M/IgG R) were used as negative control. Serial dilutions of chromatin extract (input) prepared from U-2 OS cells were loaded. Left panel, representative images. E. quantification of three independent ChIP experiments, average enrichment of telomeric repeats is indicated; s.e.: short exposure; l.e.: long exposure**.** N, number of independent experiments, whiskers indicate standard deviation; a student t-test was used to calculate p-values.

Our *data* show a specific interaction of SFPQ to telomere repeats. We therefore decided to focus our experiments on the role of SFPQ in regulating telomere function.

4.1.4 SFPQ suppress TERRA accumulation at telomere preventing the formation of RNA:DNA hybrids

4.1.4.1 SFPQ control TERRA abundance at telomeres

To understand the role of SFPQ in the management of TERRA, we performed experiments in U-2 OS cell line. By performing a TERRA RNA-FISH experiments in U-2 OS ALT cells depleted for SFPQ, we observed a significantly increase in the number of focal TERRA signal per nucleus (Figure 4.1.5.1 A,B). Experimental cells were divided in four categories based on number of TERRA foci, with intervals of 0-3, 3-8, 8-12, >12 TERRA foci per nucleus. This analysis revealed that the proportion of cells with high TERRA foci number (>12) was increased in SFPQ depleted cells (figure 4.1.5.1 C). In addition, TERRA RNA-FISH signal intensity was quantified. We found that depletion of SFPQ resulted in a significant increase of intensity TERRA signals (Figure 4.1.5.1 D). However, northern blotting and quantitative RT-PCR using primers that amplify subtelomeric portions of TERRA originating from chromosomes 1 and 21 or chromosomes 2, 10 and 13 revealed that RNAi mediated depletion of SFPQ does not impact on total TERRA levels (Figure 4.1.5.1 E-G).

These results suggest that SFPQ does not impact on overall TERRA expression levels, but rather prevent TERRA accumulation at telomeres in U2-OS cells.

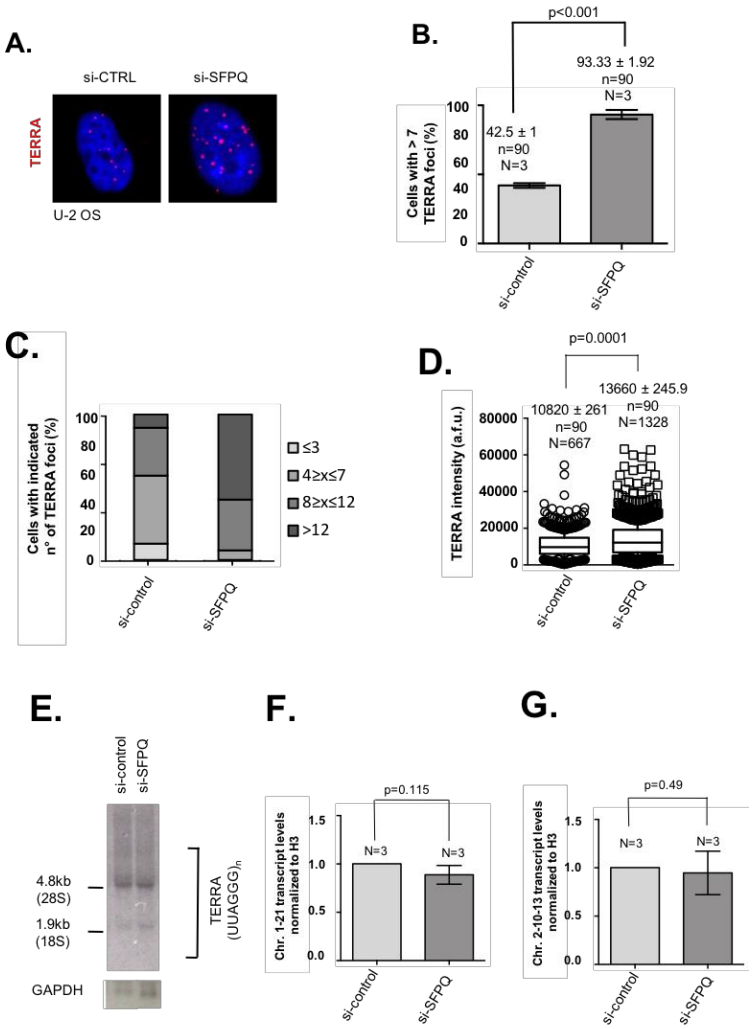


Figure 4.1.5.1 SFPQ regulates association of TERRA with telomeric chromatin. A. Representative images of TERRA RNA-FISH performed in U-2 OS cells transiently transfected with SFPQ specific siRNAs or with a non-targeting control siRNA **B.** Quantification of the percentage of cells with more than seven TERRA foci per nucleus are indicated. **C**. Distribution of TERRA foci number per nucleus of experimental U-2 OS cells **D.** Quantification of TERRA signals intensity. **E.** Classical TERRA northern blotting of RNA extracts prepared from SFPQ siRNA transfected cells. SFPQ knock-down does not alter TERRA expression or TERRA size distribution. GAPDH probe was used as loading control. **F-G.** SFPQ knock-down does not alter TERRA transcription. Levels of TERRA transcripts arising from chromosomes 1, 21 (F) or chromosomes 2, 10, 13 (G) were quantified using qRT-pcr. TERRA expression levels were normalized against histone 3 (H3) levels. Arbitrary fluorescence units (a.f.u.) are shown. Box blots: middle line represents median, and the box extends from the 25th to 75th percentiles. The whiskers mark the 10th and 90th percentiles. p values were calculated using a Mann-Whitney test. Median a.f.u. values and standard deviation are indicated; n=number of analyzed nuclei, N= number of analyzed TERRA signals. Means (bars) and SDs (error bars) are reported. N=number of independent experiments. n= number of analyzed nuclei. A student t-test was used to calculate statistical significance; p-values are shown.

4.1.4.2 SFPQ prevent RNA:DNA hybrid formation at telomeres

To have further information about the consequence of TERRA accumulation at telomeres, we analyzed the formation of γH2AX foci, a marker of DNA damage, at telomeres. We combined TERRA RNA-FISH with a specific antibody that recognize γH2AX.

Interestingly, we found that SFPQ-depleted cells show an increased percentage of colocalization of TERRA foci with γH2AX compared to control cells (Figure 4.1.5.2a A-B). Telomere damage was also verified by anti-γH2AX immuno-telomere DNA FISH at metaphase chromosomes of U-2 OS cells depleted for SFPQ (Figure 4.1.5.2a C-D).

Together, these results suggest that loss of SFPQ leads to altered TERRA homeostasis and promotes the formation of DNA damage at telomeres. Recent studies showed that RNA:DNA hybrid and R-loop formation at telomeres is linked with increased abundance of TERRA at telomeric chromatin, without altering total TERRA lncRNA levels (Arora et al. 2014; Teasley et al. 2015).

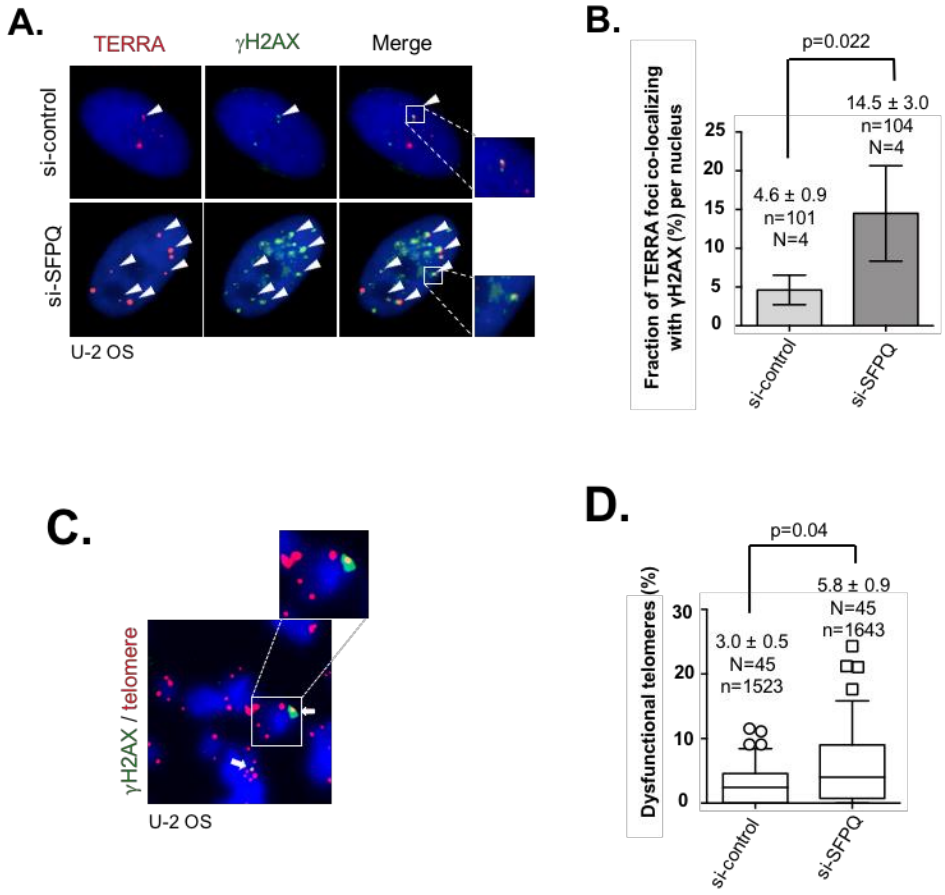


Figure 4.1.452a. SFPQ prevent DNA damage activation at telomeres. A. U2-OS cells were transfected with SFPQ specific siRNAs and TERRA RNA-FISH combined with anti-γH2AX staining was performed. Representative images are shown. **B.** Depletion of SFPQ significantly increases the percentage of TERRA foci that colocalize with γH2AX per nucleus. n=number of analyzed nuclei. Means (bars) and SDs (error bars) are reported. A student t-test was used to calculate statistical significance; p-values are shown. **C.** U-2 OS cells were transfected with NONO or SFPQ specific siRNAs, blocked in metaphase using colcemide (1ug/ml, 2h) and then collected. Metaphase spreads were prepared by cytospin. Staining with anti-γH2AX combined with telomeric DNA FISH was performed. Representative images are shown. **D**. Knock-down of SFPQ increased the percentage of dysfunctional telomeres. Each dot represents the percentage of dysfunctional telomeres (telomeres that co-localize with γH2AX) of one metaphase. Metaphases from three independent experiments were analyzed. At least 1000 chromosome ends were analyzed for each condition. Red bars indicate means. N= number of analyzed metaphases. p-values were computed using Mann-Whitney test.

Based on our previous result that demonstrate TERRA accumulation at telomere repeats in SFPQ silencing condition, we hypotisized that SFPQ may have a role in TERRA RNA-telomeric DNA hybrids formation.

We performed co-immunocytochemistry using specific anti-TRF1 antibodies and affinity purified S9.6 monoclonal antibodies that detect RNA:DNA hybrids in a sequence independent manner. Confocal microscopy revealed that RNAi mediated depletion of SFPQ significantly increased co-localization frequencies between RNA:DNA hybrids and TRF1 foci in U-2 OS (Figure 4.1.5.2b A-B). This result was re-capitulated in RNAseH1 knock-down cells that have been recently reported to show increased abundance of telomeric RNA:DNA hybrids (Figure 4.1.5.2b A-B) (Arora et al. 2014). RNAi mediated depletion of TRF1 from U-2 OS cells resulted in a complete loss of immunostaining for TRF1, thus excluding cross reactivity of the anti-TRF1 antibody in our experimental setup (Figure 4.1.5.2bC). Together, these resultss indicate that SFPQ has a role in suppressing TERRA:telomere RNA:DNA hybrid formation in telomerase negative cancer cells.

The schematic model in Figure 4.1.4.2b, shows how TERRA transcript remains annealed with its template DNA (C-strand) forming an RNA:DNA hybrid structure. The non-template strand (G-strand) remains unpaired and exposed as ssDNA. These structures are named R-loops and are particularly favoured by G-rich telomeric strand formation of secondary structures (G4- quadruplexes). R-loops are demonstrated to be a barrier for DNA replication machinery progression, thus causing stalling of replication forks. This situation leads to the increase of free single stranded telomeric DNA bound by RPA32pSer33 produced at stalled replication fork, but also by the exposition of the G-strand telomere that remains unpaired. RNaseH1 digests the RNA part of the RNA:DNA hybrids, thus favouring the re-annealing of telomeric dsDNA. S9.6 antibody recognizes RNA:DNA hybrids structures: DNA pol, DNA polymerase; RNA pol II, RNA polymerase II.

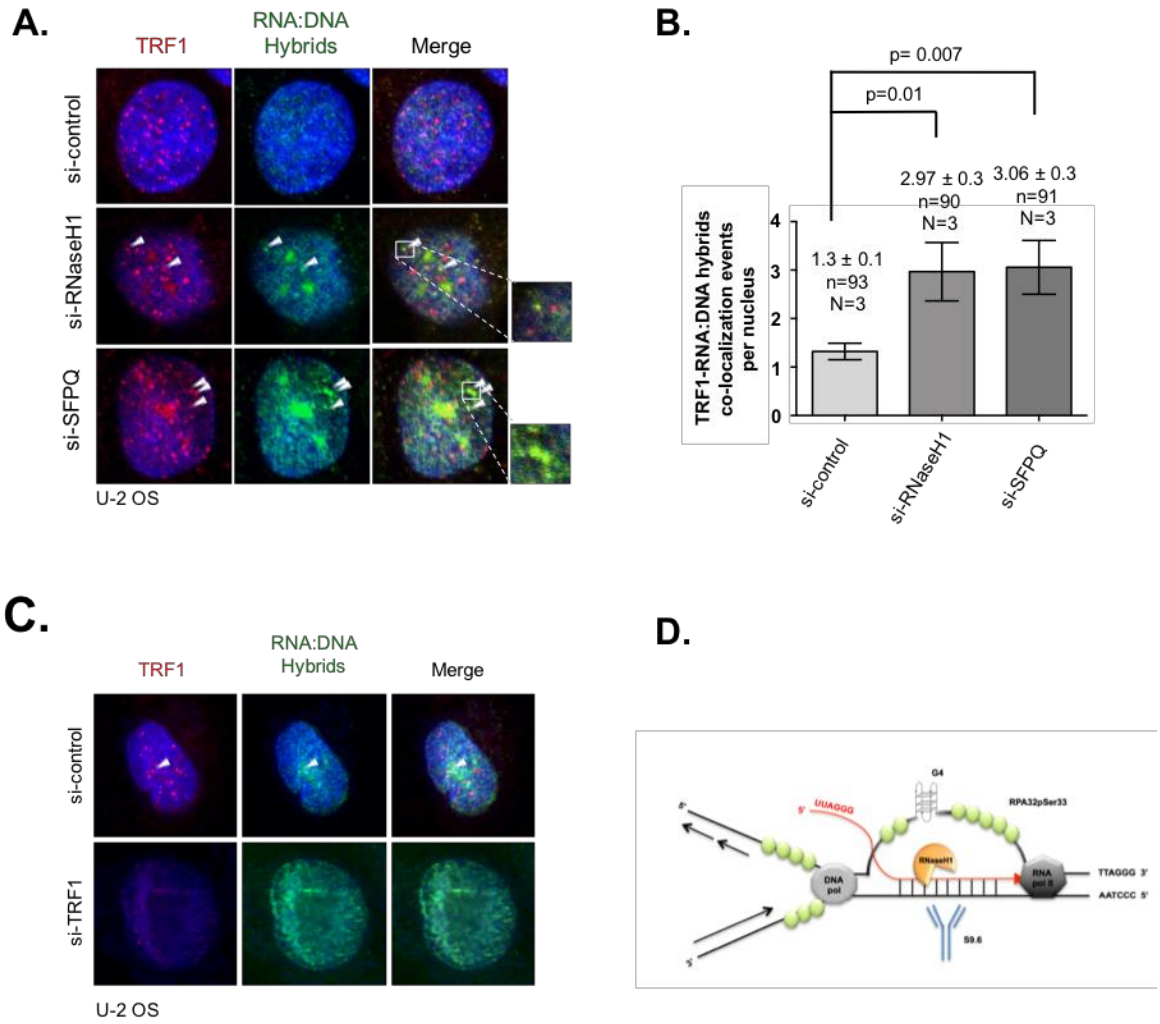


Figure 4.1.4.2b. SFPQ suppress formation of telomeric RNA:DNA hybrids. A. U2-OS cells were transfected with SFPQ and RNase H1 specific siRNAs and immunofluoresce experiments with anti-TRF1 and anti-RNA:DNA hybrids (S9.6) antibodies were performed. **B.** Depletion of SFPQ increases the number of RNA:DNA hybrids that co-localize with TRF1. Similar effect is induced by knockdown of RNase H1 that eliminates stretches of RNA from RNA:DNA hybrids. N= number of independent experiments; n=number of analyzed nuclei. Means (bars) and SDs (error bars) are reported. A student t-test was used to calculate statistical significance; p-values are shown. **C.** TRF1 siRNA was used as control of TRF1 antibody specificity. Representative images are shown. **D.** Schematic model of the impact caused by loss of SFPQ at telomeres and TERRA-telomere hybrids formation.

4.1.5 Loss of SFPQ causes replication defects at telomeres

RNA-DNA hybrids pose a threat to genome stability by causing replication fork stalling and may trigger the formation of double-stranded breaks (Arora et al. 2014).

To have further evidence of a role of SFPQ in suppressing RNA:DNA hybrids, we investigated the consequences of hybrid formations at telomeres. Induction of replication defects at telomeres is linked with the phosphorylation of ATR and the phosphorylation of Serine 33 of the 32kDa subunit of the Replication protein A (RPA32pSer33), both surrogate markers for replication stress[17]. As expected, induction of replicative stress by treatment of U-2 OS with hydroxyurea cause an increase of global RPA32pSer33 staining (Figure 4.1.6 C). In line with increased RNA:DNA hybrid formation, we found that percentage of TERRA foci that co-localize with RPA32pSer33 strongly increases upon SFPQ silencing (Figure 4.1.6 A-B).

To confirm this result, we performed co-immunofluorescence experiments (co-IF) using TRF2 antibody in combination with RPA32pSer33 (Figure 4.1.6 C-D). As expected, a concomitant HU treatment of U2-OS cells leads to a vast accumulation of RPA32pSer33 in foci dispersed across the entire nucleus (Figure 4.1.6 C). In line with RNA-FISH *data*, we found that depletion of SFPQ correlates with increased RPA32pSer33 recruitment at telomere (Figure 4.1.6 D).

Replicative stress is reported to trigger the phosphorylation of Ataxia telangiectasia and Rad3 related (ATR). In line with this, we found an efficient phosphorylation of ATR at telomeres in SFPQ knock-down U-2-OS (Figure 4.1.5 E, F). Together, these data indicate that SFPQ is a TERRA interacting protein with a crucial role in preventing RNA:DNA hybrid accumulation and R-loop related replication defects at telomere repeat sequences.

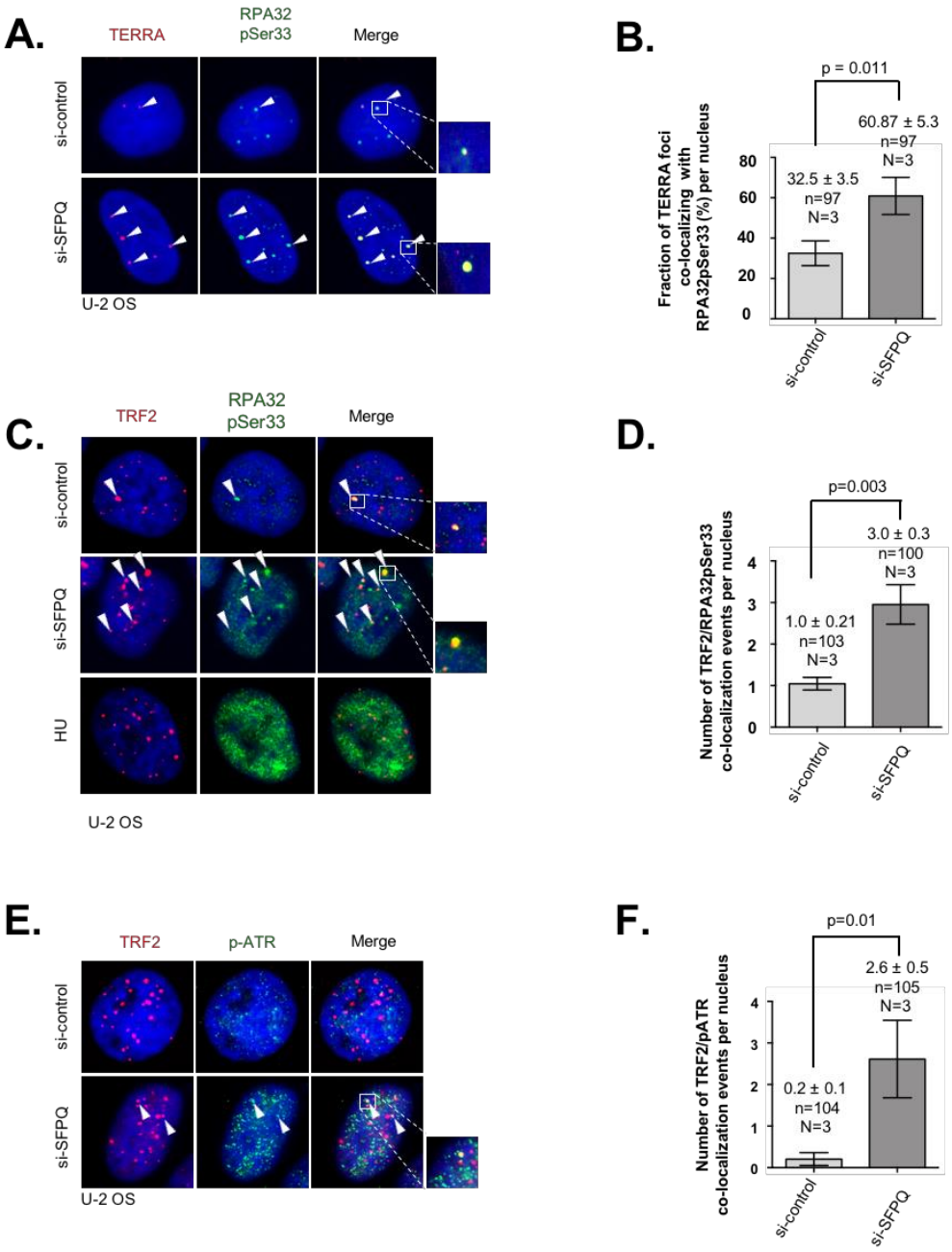


Figure 4.1.6. SFPQ prevents replicative defects at telomeres. A. Representative images of combined immunofluorescence using anti-anti-RPA32Ser33 antibodies with TERRA RNA Fish in U-2 OS cells transfected with indicated siRNAs. Arrowheads indicate co-localization events.**B** Depletion of SFPQ significantly increases the percentage of TERRA foci that colocalize with RPA32pSer33 per nucleus. **C.** Immunofluorescence with anti-TRF2 and anti-RPA32pSer33 antibodies was performed in U2-OS cells depleted for SFPQ **D**. Representative images are shown. Depletion of SFPQ significantly increases the number of TRF2-RPA32pSer33 colocalizations per nucleus. **E.** Representative images of combined immunofluorescence using anti-TRF2 and anti-pATR antibodies (S428) in U-2 OS cells transfected with indicated siRNAs. Arrowheads indicate co-localization events. **F.** Quantification of TRF2/pATR co-localization events per nucleus. Cells treated for 6 h with 5 mM hydroxyurea (HU) were used as controls for RPA32pSer33 activation. N=number of independent experiments; n=number of analyzed nuclei. Means (bars) and SDs (error bars) are reported. A student t-test was used to calculate statistical significance; p-values are shown.

4.1.6 LOSS of SFPQ regulate telomere fragility and homologous recombination at telomere

4.1.6.1 Fragility

Increase of R-Loops leads to fragility of the telomeric leading (C-rich) strand in vertebrate cells (Arora et al. 2014; Nadel et al. 2015; Wahba et al. 2011). Moreover, RNA:DNA hybrids is related to replication fork stalling and consequent breakage of the single strand DNA that remains unpaired. Leading to the formation of substrates for Homologous recombination at yeast and human telomeres.

In order to understand the direct importance of SFPQ for telomere integrity, we performed telomere chromosome orientation DNA-FISH (CO-FISH) experiment in SFPQ loss and gain of function cells. In this method BrdU incorporation during S-Phase and subsequent enzymatic digestion of the neo-synthesized DNA strand after UV treatment allows the selective detection of the TTAGGG containing lagging telomere strand or the CCCTAA repeat containing leading telomere strand using differentially labelled, telomere strand specific DNA-FISH probes (Figure 4.1.7.1 A). This method allows to detect telomere-strand specific aberrations, but also telomere recombination events (Figure 4.1.7.2 A).

Fragile telomeres appeare as multimeric telomere signals as shown in Figure 4.1.7.1 A.

SFPQ knockdown strongly reduces basal fragility of the telomeric leading strand telomere and does not impact on lagging strand fragility (Figure 4.1.7.1 B).

Differently, stable over-expression of SFPQ in U2-OS cells does not have a significant impact on telomere fragility of neither the lagging or the leading strand (Figure 4.1.7.1 C). This suggest that endogenous SFPQ expression levels are sufficient to keep telomere fragility at sustainable levels. Western blotting experiments confirmed the NONO and SFPQ silencing/overexpression efficiency (Figure 4.1.7.1 B-C).

Together, these results suggest the role of SFPQ as a novel regulator of telomere fragility.

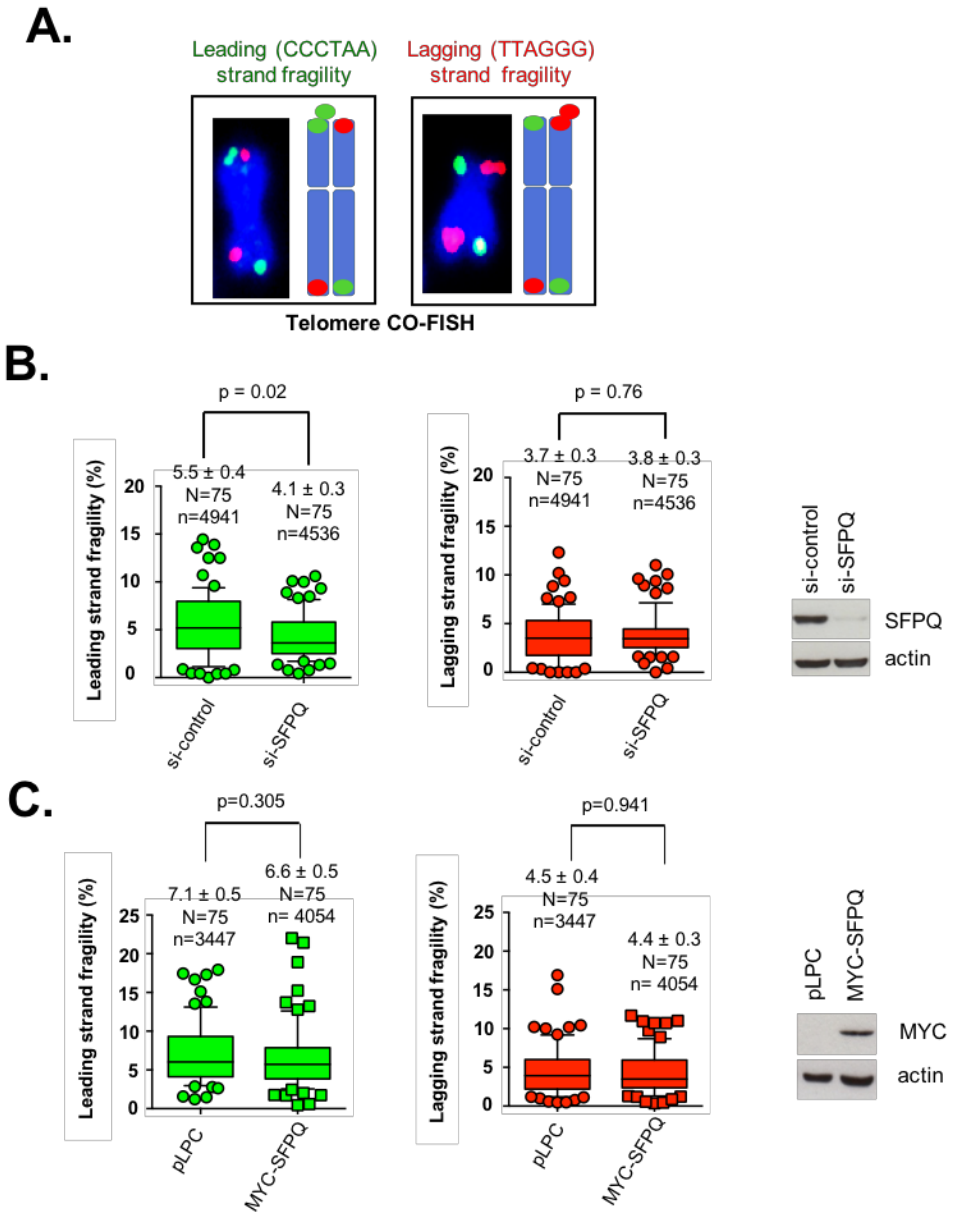


Figure 4.1.7.1. SFPQ regulate telomere fragility in U2-OS cells. A. Cells were transfected with SFPQ or control siRNAs and after 72 hours CO-FISH analysis was performed **B.** Quantification of leading strand fragility (left panel) and lagging strand fragility (central panel) in U-2 OS cells transfected with SFPQ specific siRNA or a non-targeting control siRNA. Right panel, western blotting showing SFPQ knock-down efficiency. **C**. Quantification of leading strand fragility (left panel) and lagging strand fragility (central panel) in U-2 OS cells that stably overexpress myc-tagged SFPQ or an empty vector. Right panel, western blotting showing expression of ectopic myc-SFPQ. For quantifications in panels B.-C., data points represent the percentage of fragile telomeres per metaphase spread. Only chromosome ends with detectable telomeres were considered for analysis. Metaphases from three independent experiments were analyzed. Box blot diagrams B-E: middle line represents median; boxes extend from the 25th to 75th percentiles. The whiskers mark the 10th and 90th percentiles. p values were calculated using a Mann-Whitney test. Median fragility values and standard deviation are indicated; n=number of analyzed telomere repeat signals, N= number of analyzed metaphase spreads.

4.1.6.2 SFPQ suppress recombination at telomeres

To have further information about the role of SFPQ at telomere, we studyed recombination based on telomere sister chromatin exchange, a reported downstream result of RNA:DNA hybrid related R-loop formation at eukaryotic telomeres (Arora et al. 2014; Balk et al. 2013).

A merged signal (yellow) of the two telomeric probes used for CO-FISH indicates that a recombination event occurred between telomeric sister chromatids (Telomeric sister chromatid exchange or T-SCE) (Figure 4.1.7.2a A). We found that depletion of SFPQ induces two fold of increase in T-SCE frequencyof U2-OS cell lines (Figure 4.1.7.2a B). These results demonstrate that SFPQ has a critical relevance in suppressing homologous recombination at telomeres.

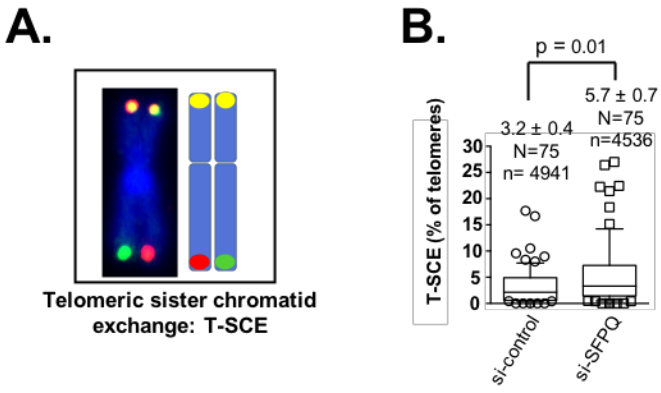


Figure 4.1.7.2a SFPQ complex suppresses homologous recombination at telomeres. CO-FISH analysis was performed in cells transfected with SFPQ siRNAs (A-E). **A**. Representative images of metaphase spreads; normal telomeres in si-control condition and recombinant telomeres in si-SFPQ cells are shown. **B.** T-SCE frequency was measured as the percentage of T-SCE events per number of chromosome ends in one metaphase. Average of T-SCE frequency of metaphases from three independent experiments is represented. At least 1000 chromosome ends were analyzed for each condition. Box blot diagrams B: middle line represents median; boxes extend from the 25th to 75th percentiles. The whiskers mark the 10th and 90th percentiles. p values were calculated using a Mann-Whitney test. Median T-SCE/fragility values and standard deviation are indicated; n=number of analyzed telomere repeat signals, N= number of analyzed metaphase spreads.

ALT cells are characterized by elevated levels of telomere recombination events and the appearance of ALT-associatedpromyelicytic leukaemia bodies (APBs).

To demonstrate the importance of SFPQ in suppressing homologous recombination, showed in the CO-FISH analysis, we performed co-immunostaining TRF2 and anti Promyelocytic Leukaemia (PML) antibodies.

In line with increased T-SCE frequency, loss of SFPQ correlates with an increased TRF2-PML number of co-localizations events (Number of APBs per nucleus) (Figure 4.1.7.2b A-B). We also observed a global increase in the number of APBs per nucleus. Together, these results indicate that SFPQ impacts

on two distinct features of ALT cells: T-SCE frequency and number of APBs. We further found that depletion of SFPQ knockdown results in an increase of RAD51 loading at telomeres, indicative for abundance of recombinogenic DNA substrates (Figure 4.1.7.2b C-D).

In addition, we performed a TERRA RNA-FISH in combination with PML antibody showingan increase colocalization between TERRA and PML bodies in SFPQ depleted cells.

Together these *data* suggest that SFPQ is important to suppress homologous recombination of telomeric DNA in ALT-associated promyelocytic leukaemia body (APBs).

Altogether, these results demonstrate that SFPQ acts as a suppressor of telomere recombination in U-2 OS cells.

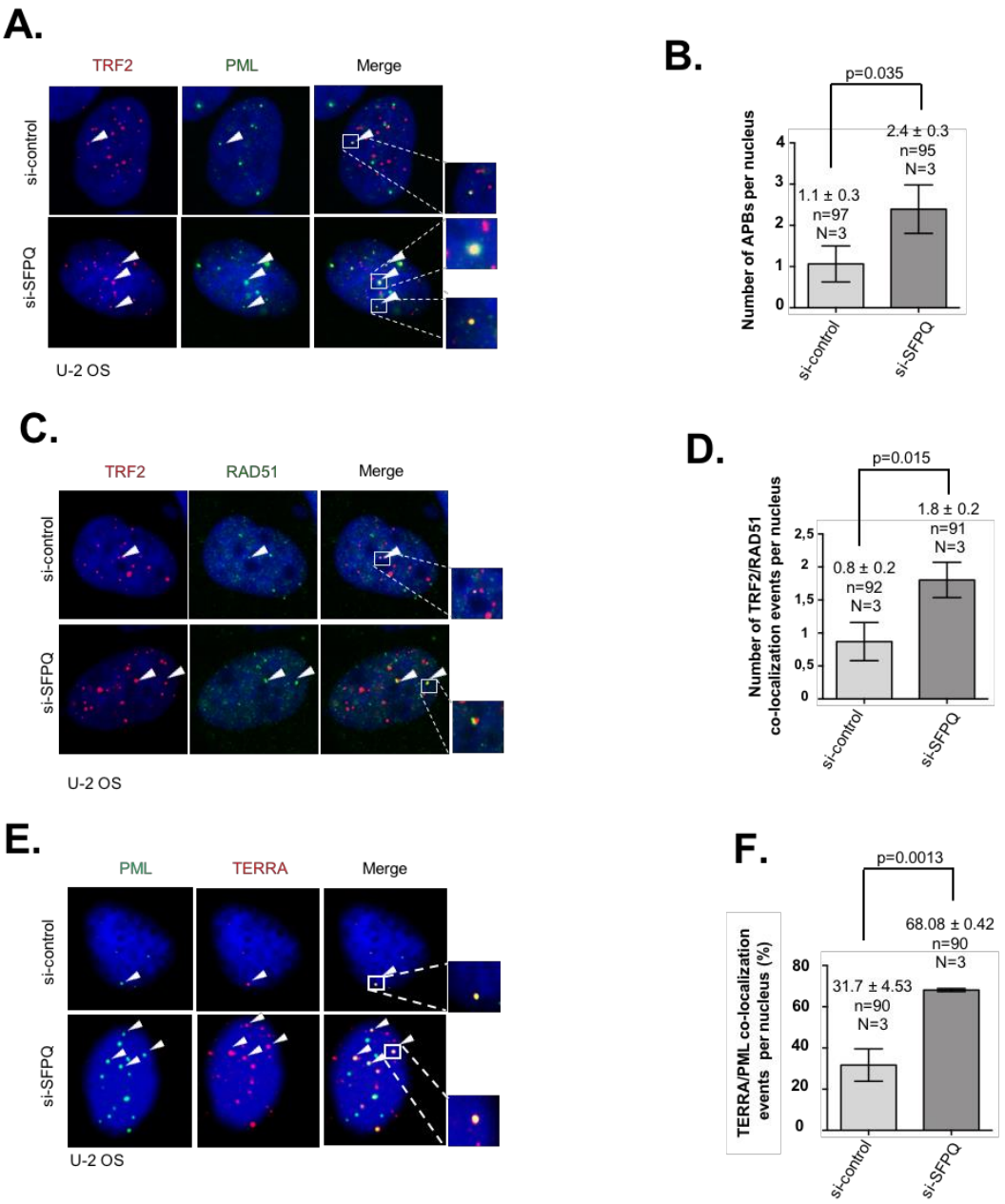


Figure 4.1.7.2b SFPQ complex suppresses recombination processes in ALT-associated promyelocytic leukaemia body (APBs). A. U2-OS cells were transfected with SFPQ specific siRNAs and immunofluorescence experiments were performed. Representative images anti-TRF2 and anti-PML co-immunofluorescence to detect APBs in U-2 OS cells transfected with indicated siRNAs. **B.** Quantification of the number of APBs per nucleus of data shown in A. **C.** Representative images anti-TRF2 and anti-RAD51 co-immunofluorescence in U-2 OS cells transfected with indicated siRNAs. **D.** Quantification of RAD51-TRF2 co-localization events per nucleus of data shown in C. **E.** U2-OS cells were transfected with SFPQ specific siRNAs and immunofluorescence experiments were performed. Representative images TERRA RNA FIsh and anti-PML co-immunofluorescence to detect APBs in U-2 OS cells transfected with indicated siRNAs. **B.** Quantification of the number of APBs per nucleus of data shown in E. Box blot diagrams B, D, F: means (bars) and standard deviation (error bars) are reported. N=number of independent experiments. n= number of analyzed nuclei. A student t-test was used to calculate statistical significance; p-values are shown.

4.1.7 SFPQ regulate telomere length homeostasis in telomerase negative cancer cells

Our data have shown that the loss of SFPQ causes an increase of RNA-DNA hybrids at the telomere with consequent increase in replication defects and telomeric recombination. The impairment of pathways related to homologous recombination result in an alteration of telomere length homeostasis in ALT cells (Clynes et al. 2015). Wetherefore investigated the impact of depletion of SFPQ on telomere length homeostasis in telomerase negative U-2 OS ALT cells.

Quantitative telomere DNA-FISH analysis was performed on SFPQ knockdown U-2 OS. We observed that transient loss of SFPQ leads to an increase of 17% in telomere length in a 3 days experimental period. This is in line with the dramatically increase of T-SCE frequency under these conditions.

Altogether, our data highlight a role for SFPQ in controlling RNA:DNA hybrid related telomere instability and telomere length homeostasis in telomerase negative cancer cells.

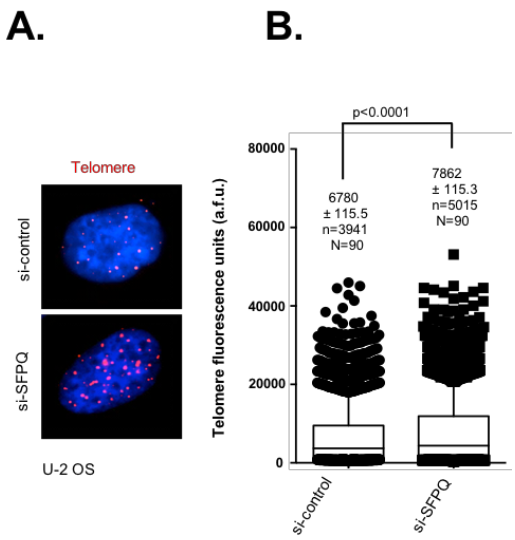


Figure 4.1.8. SFPQ regulate telomere length in human cancer cells. A. U2-OS cells were transfected with control or SFPQ siRNAs. After 72 hours, telomere length measurements were performed by quantitative telomere DNA FISH on interphase cells. Representative images are shown **B.** Telomere fluorescence intensity analyzed for each telomere of (A). Box blot diagrams B: middle line represents median arbitrary fluorescence units (a.f.u.), boxes extend from the 25th to 75th percentiles. The whiskers mark the 10th and 90th percentiles. p values were calculated using a Mann-Whitney test. Median a.f.u. values with standard deviation are indicated; n=number of analyzed telomere repeat signals, N= number of analyzed interphase nuclei. N=number of analyzed nuclei; n= total signals analyzed; Arbitrary fluorescence units (a.f.u.) are shown.

4.1.8 Working model

In the first part of my thesis, we discovered SFPQ and NONO as two novel TERRA interactors that work as a acomplex at chromosome ends in this study.

We focalized our attention on SFPQ. We propose that this protein prevents RNA:DNA hybrids formation at telomeres stabilizing telomere structure.

Loss of SFPQ leads to accumulation of TERRA foci that correspond to RNA:DNA hybrids structure and consequent displacement of non-template strand (G-strand) that remains unpaired and exposed as ssDNA bound by RPA32pSer33 (Figure 4.1.7). R-loops are prone to form secondary structures, such as G-quadruplex on G-rich strand, and might be source of fragility.

ssDNA lesions and DNA breaks fuel recombination processes at telomeres mediated by RAD51.

This leads to increase T-SCE rates and increased telomere length in recombination proficient ALT cells.

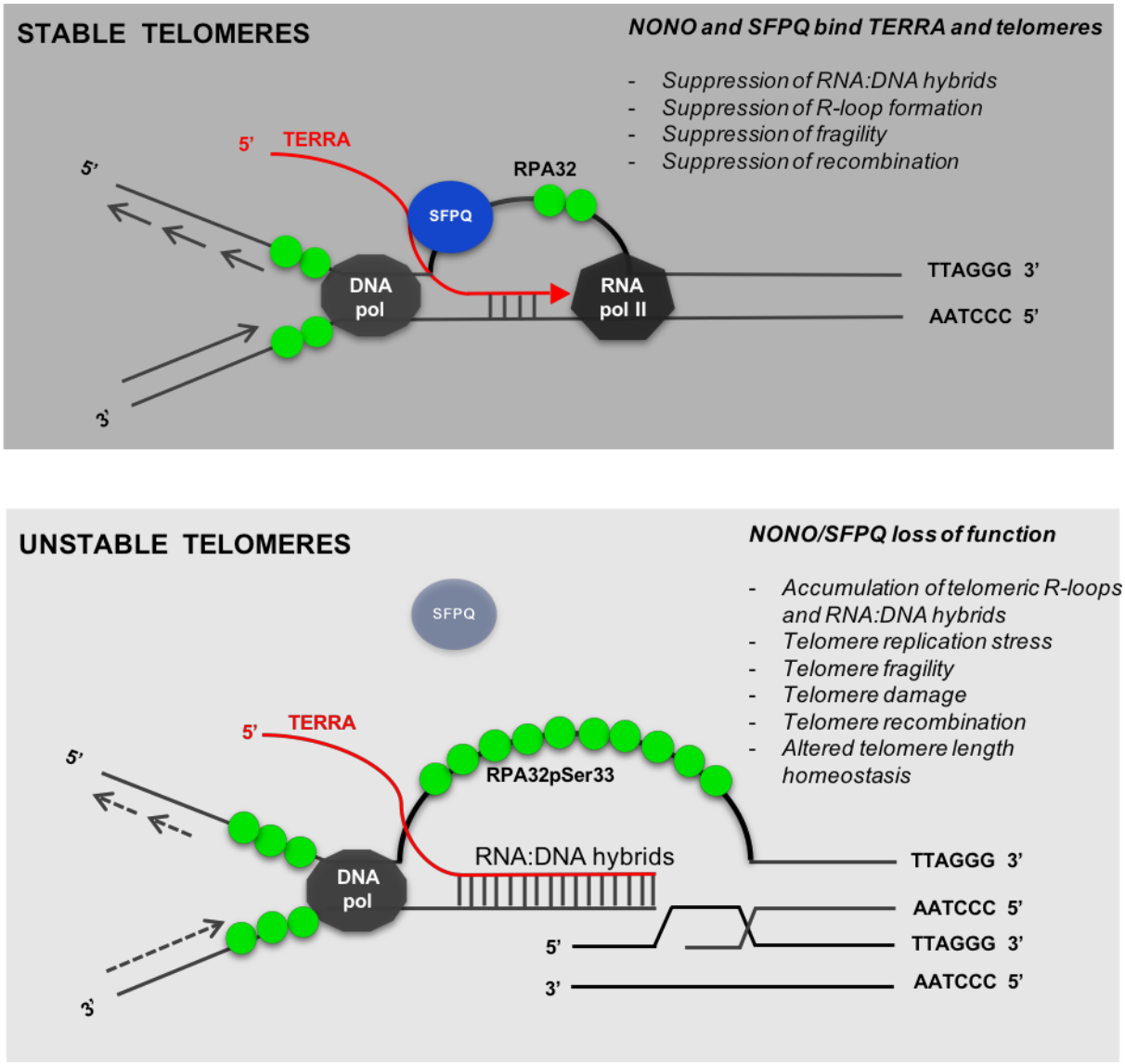


Figure 4.1.9 Working model. A. In normal conditions SFPQ interact with telomeric chromatin and bind to nascent TERRA transcript, thus allowing proper TERRA displacement from the template strand (C-strand telomere). In this way, SFPQ/NONO complex prevents stalling of DNA replication forks and consequent fragility as well as the formation of exposed ssDNA or DNA breaks that can trigger recombination processes at telomeres. B. Loss of SFPQ leads to formation of TERRA-RNA: telomeric-DNA hybrids structures and consequent displacement of the non-template strand (G-strand) that remains unpaired and exposed as ssDNA bound by RPA32pSer33. These structures named R-loops are particularly favoured by the propensity of G-rich telomeric strand to form secondary structures (G4-quadruplexes). The exposed ssDNA of R-loop is unstable and prone to DNA lesions. In addition, R-loops form a barrier to DNA replication machinery progression, thus causing stalling of replication forks and consequent DNA breaks. ssDNA lesions and DNA breaks might fuel recombination processes at telomeres mediated by RAD51. DNA pol, DNA polymerase; RNA pol II, RNA polymeraseII.

4.2 Project 2: Identification and functional analysis of novel SFPQ interactors

4.2.1 SFPQ interaction with the histone chaperone DAXX

4.2.1.1 Identification os SFPQ interacting protein

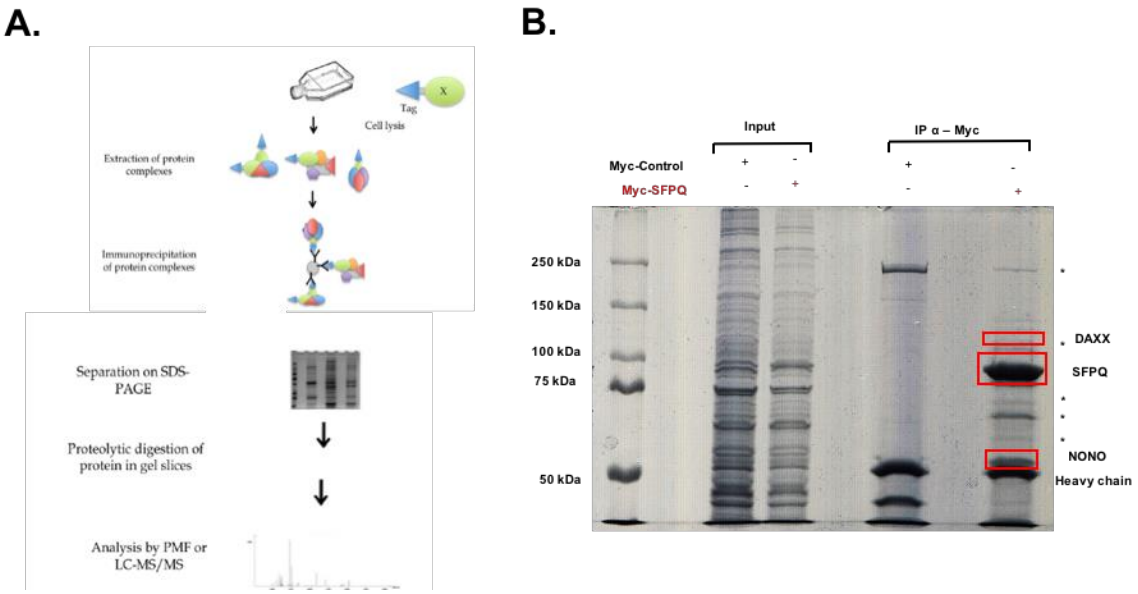
The results obtained in the first project evidence a role of SFPQ in the management of RNA:DNA hybrids at telomeres. The lack of a catalytic function of SFPQ suggest that SFPQ might act as recruitment factor for enzymes that control RNA:DNA hybrids load at telomeres.

In order to identify novel interactors of SFPQ, H1299 cells were transiently transfected with a control vector and a vector encoding a Myc-tagged version of SFPQ. 48 hours post transfection cells were subjected to immunoprecipitation (IP) assay using an anti-Myc specific antibody. In particular, whole cell extracts were incubated with a specific anti-Myc antibody conjugated to magnetic beads. Bound protein complexes were eluted and subjected to gel electrophoresis and Coomassie staining (Figure 4.2.1 B). Proteins specifically eluted from anti-Myc beads were cut from the gel and subjected to mass spectrometry analysis (Schematic representation Figure 4.2.1 A).

As expected, we identified NONO as major SFPQ interacting protein. Moreover, we found the histon chaperone DAXX in anti-Myc SFPQ eluates. Given the importance of DAXX as regulator of chromatin structure at telomeres and its role in ALT functions, we focused our reserch on SFPQ-DAXX interaction. To get a better insight into SFPQ interactome, we used identifical peptides to generate an interaction network focusing on experimentally validation DAXX and SFPQ interacting protein. ATRX was included in the network.

The ATRX/DAXX complex is enriched at telomeric repeats where it has been shown to regulateH3.3 deposition (Goldberg et al. 2010; Law et al. 2010; Lewis et al. 2010; Wong 2010) (Fig. 2A). ATRX depletion impairs the heterochromatic state of the telomeres, because of a reduction in H3.3 incorporation, which leads to telomere destabilization (Wong 2010; Watson et al. 2013; Ramamoorthy and Smith 2015; Watson et al. 2015).

Dysfunction in the ATRX-DAXX complex leads to the formation of non-B DNA structures like these G4-DNA structure that creates a recombinogenic nucleoprotein structure for homologous recombination (HR) at the telomeres and thereby facilitates the ALT development (Law et al. 2010). A recent study suggests that ATRX does not appear to possess G4-DNA unwinding activity and that it might overcome these impediments indirectly, perhaps by favoring the histone H3.3 deposition to maintain DNA in a B-form conformation or by promoting a fork bypass via template switching (Clynes et al. 2014, 2015).



4.2.1.1 Mass spectrometry sample preparation. H1299 cells were transfected with a Myc-tagged SFPQ and an Immunoprecipitation was performed. A. schematic representation of Immunopr4cipitation procedure. B. Comassie staining of SDS-PAGE gel after mono-dimentional electrophoresis of eluates obtained from anti-Myc immunoprecipitation performed in U2-OS cell transiently transfected with a Myc-tagged SFPQ. Candidate SPFQ interacting proteins analysed by mass spectrometry analysis are indicated with *.

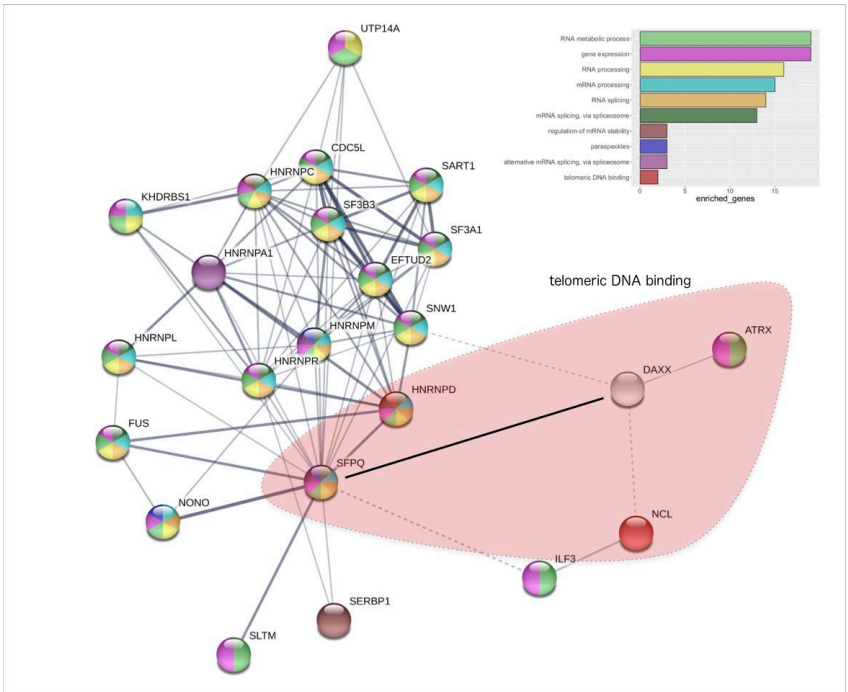
4.2.2 Protein-protein network analysis

Two hundred and sixty-eight unique proteins were identified by the mass spectroscopy of SFPQ co-precipitation eluates, from H1299 cells transfected with myc-tagged-SFPQ.

The list of gene symbols from proteomics analysis was then subjected to STRING version 10.5 (http://string-db.org/) for protein-protein interaction (PPI) network analysis. The PPI networks were obtained on the basis of the 'experimental evidence' (p-value <1.06e-16) for identified peptides.

We focused our attention on DAXX (Death domain associated protein), a histone chaperone that mediates the incorporation of the histone variant H3.3 into telomeric and pericentromeric chromatin (Drané et al. 2010; Lewis et al. 2010). DAXX can work in an independent manner or in cooperation with ATRX (Alpha Thalassemia/Mental Retardation Syndrome X-Linked), a chromatin remodelling factor. In the absence of ATRX or DAXX, loss of H3.3 nucleosome assembly leads to defective chromatin compaction and heterochromatin silencing at telomeres. (Ahmad and Henikoff 2002; Tagami et al. 2004). This creates a permissive environment for the activation of ALT. In addition, the presence of G-quadruplex structures in an ATRX null tumour cells, leads to replication fork stalling and collapse, providing a substrate for homologous recombination and maintenance of telomeres length through ALT pathway. (Clynes et al. 2015).

A subset of direct eighteen SFPQ- and two DAXX-interactors, identified in mass spectrometry experiment, was then used to construct the SFPQ-DAXX network of experimentally verified interactions. The ATRX protein, not present in mass spectrometry output, was also included in the network analysis, given its cole connection to DAXX. Thus, we obtained a network of 23 direct interacting proteins (SFPQ-DAXX network) (Table 1). In order to understand the involvement of these proteins in the biological pathways, we performed the functional enrichment for Gene Ontology (GO) terms, including the Biological Process (BP) and Molecular Function (MF) domains. The most relevant pathways in the SFPQ-DAXX network were highlighted in the Figure 4.2.2, including RNA processing, RNA spicing and but also telomeric DNA binding. According to these results, we propose that interaction involved in the telomeric DNA binding pathway



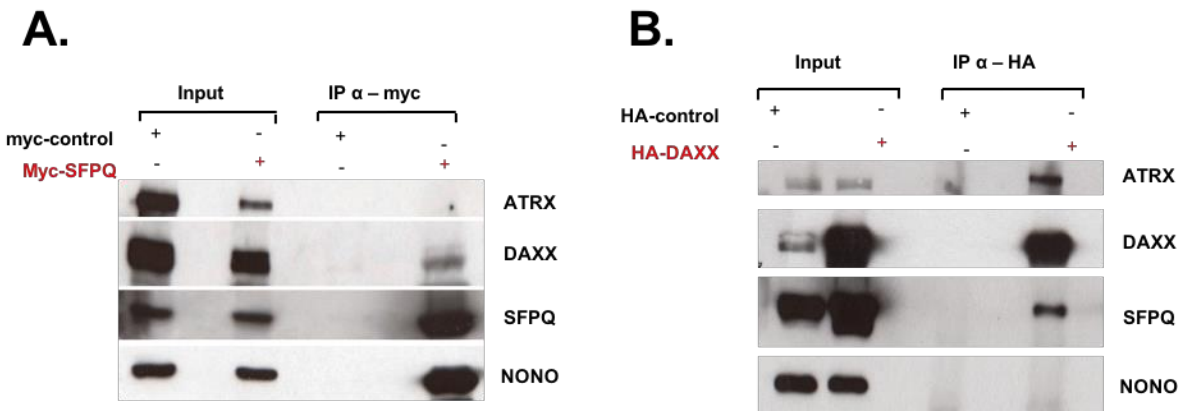
Protein	Interactors
SFPQ	UTP14A, SERBP1, SF3B3, CDC5L, EFTUD2, HNRNPR, HNRNPD, HNRNPC, HNRNPM, NONO, SART1, SNW1, HNRNPA1, FUS, HNRNPL, KHDRBS1, SF3A1, SLTM
DAXX	ILF3, NCL
ATRX	

4.2.2 SFPQ-DAXX-interactor network **Upper panel.** The SFPQ-DAXX network (23 interacting proteins). Nodes showing the same colour belongs to the same GO term. Dashed-lines correspond to Inter-term edges. The possible SFPQ-DAXX interaction involved in the telomeric DNA binding pathway (highlighted in red) are indicated with a black line. **Bottom panel.** SFPQ and DAXX direct interactors

4.2.3 Validation of SFPQ-DAXX interaction

In order to test the specificity of SFPQ/NONO-DAXX interaction, separate co-immunoprecipitation experiment followed by immunoblotting were performed. H1299 cells were transiently co-transfected with Myc-tagged SFPQ and HA-tagged DAXX. Co-immunoprecipitation experiments were carried out 48 hours post transfection using specific antibodies against Myc and HA target protein. The anti-Myc immunoprecipitation reveals the presence of DAXX in the eluate of Myc-SFPQ CO-immunoprecipitation (Figure 4.2.2 A). The detection of NONO was used as positive control of the experiment. The absence of ATRX in Myc-SFPQ immunoprecipitation may be due to a problem of binding between the antibody and the immunoprecipitated complex. Another hypothesis may be that DAXX interacts with SFPQ independently from ATRX.

Next, we performed an anti-HA immunoprecipitation. Western blot confirmed the interaction between DAXX and SFPQ (Figure 4.2.2 B). The interaction between DAXX and ATRX was used as positive control of the experiment. Altogether, these data strongly support the idea that NONO, SFPQ, DAXX and ATRX may form a functional relevant protein complex in U2-OS cells. The absence of ATRX in anti-Myc SFPQ eluate may suggest that only a subset of SFPQ, DAXX, NONO and forms a stable complex with ATRX.



4.2.2 SFPQ-DAXX interaction H1299 cells were transiently transfected with Myc control, Myc-SFPQ and HA-contro, HA-DAXX respectively**.** 48 hours after transfection, transfected cells were lysed and immunoprecipitated by anti-Myc and anti-HA antibody **A.** Co-immunoprecipitation was carried out by incubating the cleared cell lysate with Myc magnetic beads and subsequent immunoblotting of immunoprecipitates with anti-SFPQ antibody, anti-NONO antibody, anti-DAXX antibody and anti-ATRX antibody respectively. The western blot shows the new interaction between DAXX and SFPQ. We detected also NONO that was used as positive control. The obtained results confirm the bioinformatic analysis. **B.** HA-DAXX immunoprecipitation was performed in order to confirm the interaction between DAXX and SFPQ. We confirmed the intraction between DAXX and SFPQ. Detection of ATRX was used as positive control.

4.2.4 Characterization of domains of interaction between SFPQ and DAXX

4.2.4.1 Protein domains of SFPQ

SFPQ consists of 707 aminoacids and has a molecular weigth of 76 kDa. SFPQ is characterized by seven distinct domains. RGG box: the first 27 amino acids is particularly enriched in arginine and glycine residues, including multiple trimeric RGG repeats. RGG boxes have been shown to mediate interaction with DNA and , in particular, it's crucial for the association with G-quartet structures. RGG box is involved in protein- protein interaction, as well as in controlling protein localization (Thandapani et al. 2013) In PSF, the N-terminal RGG box is essential for RNAs cleavage and polyadenylation, but is dispensable for interaction with the RNA Polymerase II CTD, and has only a minor effect on splicing (Emanuel Rosonina et al. 2005).

Proline-rich: The ∼ 200 amino acids following the N-terminal RGG box are characterized by an enrichment of proline and glutamine residues. This domain consist of two parts: a ∼50-amino acid stretch where all but 6 amino acids are proline or glutamine followed by an additional ∼150 amino acids where over a third are prolines (P). Based on the understanding of short proline-rich domains within signaling proteins, it is likely that the proline-rich region of SFPQ plays an important role in mediating protein–protein interactions, particularly with non-DBHS proteins. As evidenced in a work of Rosonina E. et al., deletion of the proline-rich domain abolished the ability of SFPQ to associate with the strong transcriptional enhancer VP16 (Emanuel Rosonina et al. 2005).

PR Linker made up of ∼33 amino acids and is located between the proline-rich domain and RNA Recognition Motif 1. It is the less studied domain, but experimental data show that it translational modifications (PTMs) that could regulate the interaction and activity of SFPQ (Lukong, Huot, and Richard 2009). The RNA Recognition Motifs are frequently involved in RNA recognition and binding SFPQ, two tandem RRMs are present roughly in the middle of the primary sequence of the protein, where they are separated from each other by a seven-amino acid alanine-rich linker (FATHAAA). RRMs are characterized by highly conserved βαββαβ fold that consists of a four-stranded anti-parallel beta sheet packed against two alpha helices.

Recent studies have revealed that the more C-terminal RRM (RRM2) contains the primary RNA-binding activity of SFPQ, while the more N-terminal RRM (RRM1) has little ability to bind RNA on its own. Notably, RRMs domains can participate in protein–protein interactions (Clery et al. 2013; Maris, Dominguez, and Allain 2005). NOPs is the complete name of this domain is NONA/ParaSpeckle domain, for its homology and structure with the NONO/PSPC1 dimer. NOPS domain makes contacts with RRM2 and the coiled-coil domain (Passon et al. 2012). Moreover, the NOPS domain is described as a protein–protein interaction domain that is essential for the formation of functional dimers in the cell.

Coiled-coil domain: consist of ∼60–100 amino acids, the entire coiled-coil domain is necessary for targeting of the DBHS proteins to paraspeckles, while only half of this domain is needed for dimerization.

The C-terminal ∼100 amino acids does not notable homology to the other DBHS proteins. A residual of 7 aminoacids of the C-Terminal domain works together with a portion of the coiled-coil region (AA 547–574) to mediate the nuclear localization of SFPQ (Dye and Patton 2001). No additional striking sequence or structural feature has been proposed for the C-terminal. This domain is moderately enriched for glycine (∼30%) suggesting flexibility of the region.

4.2.4.2 A deletion analysis reveals the SFPQ domain that interacts with DAXX

After having experimentally verified the interaction of DAXX with SFPQ, we have focused our attention on identifying of the domains involved in this interaction.

We used vector encoding a Myc-tagged full length version of SFPQ or a series of SFPQ deletion constructs (ΔRGG, ΔRRM1, ΔRRM2, ΔN, ΔP/Q, Δ+/-) (Figure 4.2.4.2 A)

U-2 OS were transiently transfected using Myc-SFPQ full length or delition constucts in combination with HA-DAXX and Flag-NONO. After 48 hours, cells were collected and whole cell extracts were used to performed anti-MYC immunoprecipitation experiments. The immunoprecipitated complexes were analysed by Western bloting.

The full length version of MYC SFPQ interacts with NONO and DAXX (Figure 4.2.4.2 A-B). Importantly, binding of Flag-NONO and HA-DAXX to SFPQ, is reduced in cells transiently transfected with SFPQ constructs lacking the RRM1 or RRM2 domains (Figure 4.2.4.2a). Interestingly, the ΔN mutant showed defect in binding of DAXX without interefering with SFPQ-NONO interaction. We did not observe changes in interaction with the other mutants.

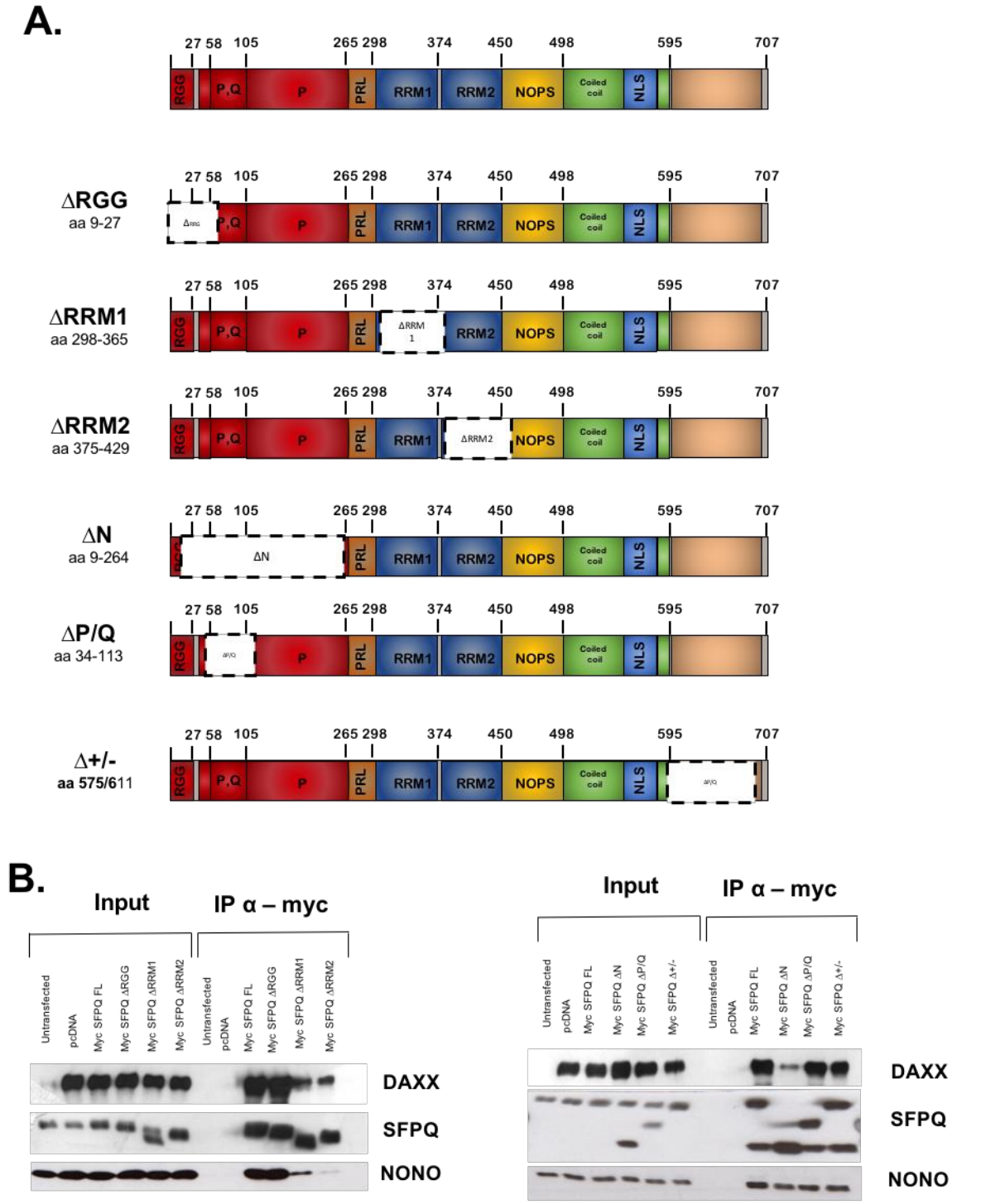


Figure 4.2.4.2 a A) Schematic representation of the anti-Myc Western blot deleted constructs **B)** Full length myc-tagged SFPQ and deletion mutant vectors were transiently overexpressed in combination with HA-DAXX and Flag-NONO in U2-OS cells. Anti-Myc western blots were performed using SFPQ, DAXX and NONO specific antibodies. show a decrease of the interaction between DAXX and constructs lacking of N-term, RRM1 and RRM2 domains. Untrasfeceted sample was loades as negative control for Myc-tagged proteins.

The results obtained from anti-Myc-SFPQ CO-immunoprecipitation have been confirmed by anti-HA DAXX IP experiments. We transiently transfected the U2OS with HA-DAXX full length and myc-tagged vector encoding the full length form or SFPQ deletion constructs. Anti-HA DAXX was performed and western blot analysis clearly show a reduction in the interaction between DAXX and SFPQ lacking of N-term, RRM1 and RRM2. These results confirm what we observed in the anti-Myc CO-IPs.

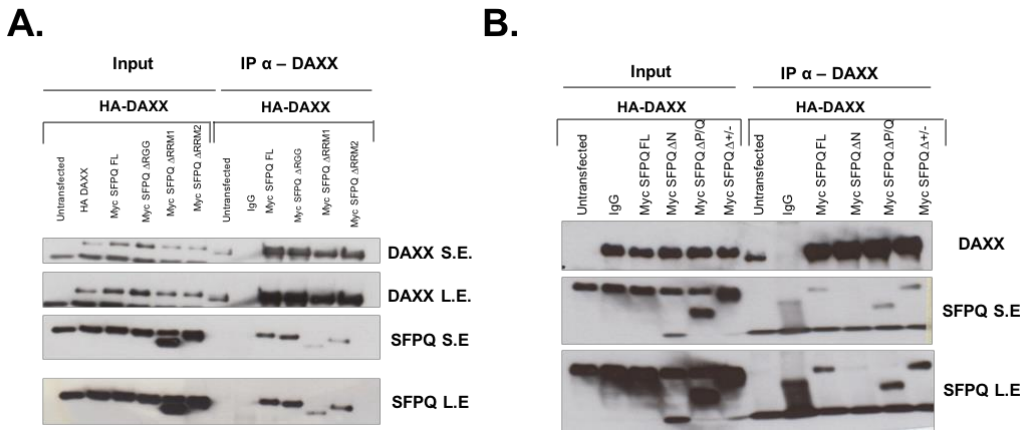


Figure 4.2.4.2 B Full length HA-tagged DAXX, full length Myc-tagged SFPQ and deletion mutant vectors were transiently overexpressed in U2-OS cells. Western blots show a decrease of the interaction between DAXX and constructs lacking of N-term, RRM1 and RRM2 domains

CO-immunoprecipitation experiments have shown that the fundamental domain for the link between SFPQ and DAXX is the N-terminal domain of SFPQ. This domain consists of three sub-domains: RGG box, PR domain and P domain. Among the various transfected mutants, we also used SFPQ constructs deleted for RGG and P/Q domains, which showed no changes in the SFPQ-DAXX binding. We, therefore, can conclude that the fundamental domain for the interaction between SFPQ and DAXX is located inthe P domain.

4.2.5 SFPQ mutations in Osteosarcoma

Osteosarcomas are primary malignant tumors of bone tissue, which have complex karyotypes characterized by many structural and numerical aberrations. A high-throughput sequencing study of 123 samples of osteosarcoma was observed to detect mutations in driver genes of this type of tumor (Kovac et al. 2015).

From this study, among the various mutated genes, two mutations have emerged concerning SFPQ:

SFPQ:NM_005066:exon1:c.448_449insAGC:p.T150delinsKP

SFPQ:NM_005066:exon1:c.442_443insAGCG:p.G148fs

The first mutation is a triplet (AGC) insertion between the nucleotide 448 and 449. This mutation occurs in P domain and causes the loss of a threonine in position 150 of the amino acid sequence of SFPQ and an acquisition of a lysine and a proline. The second mutation consists in the insertion of four (AGCG) nucleotides between the nucleotide 442 and 443. This insertion leads to the variation of the reading frame, starting from the amino acid 148 in the P domain. Thus, many premature stop codons are formed, the first one (TAA) is found in position 200 of the amino acid sequence of SFPQ.

In order to have more information about the iimpact of these SFPQ mutated protein in patients, we wanted to investigate whether these mutations impact on the interaction between SFPQ and DAXX. An in vitro site-directed mutagenesis kit allowed us to create SFPQ mutants. U-2 OS cells were transiently transfected with Myc-tagged full length version of SFPQ and the new mutated versions. Whole cells extracts were incubated with anti-Myc tag antibody conjugate with magnetic beads. The eluted were loaded on SDS-PAGE gel. From western blot analysis emerged that SFPQ mutation don't affected the interaction between and SFPQ and DAXX. As shown in the Figure 4.2.4, endogenous SFPQ is presence in all co-immunoprecipitation eluates. The presence of endogenous SFPQ could be explained by the fact that the proteins of the DBHS family can form homodimers. Therefore, endogenous SFPQ may have been linked to the exogenous mutated protein. It may also be that mutation SFPQ may still bind DAXX, however impair DAXX function.

Future experiments involve the use of cells that stably express mutated versions of Myc SFPQ. Depletion of endogenous SFPQ in these stable cells line can give us more information on the real interaction between the SFPQ mutants and DAXX and the impact on telomere functions.

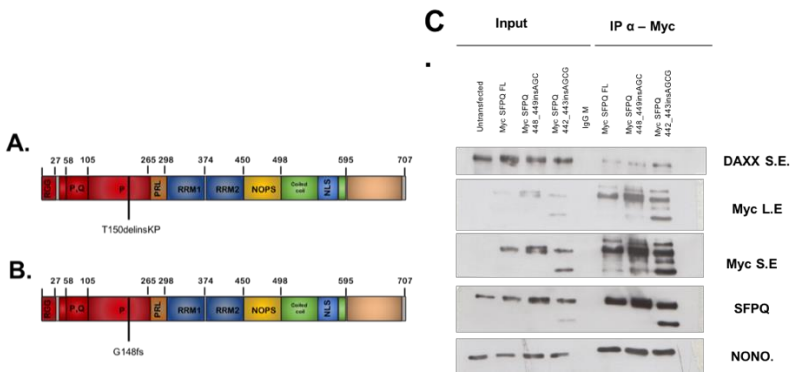


Figure 4.2.5 SFPQ patient mutations A-B schematic representation of SFPQ patient mutations. C. U-2 OS cells were transiently transfected with anti-MycSFPQ full length and Myc SFPQ mutated constructs. Western blot analysis was performed using specific antibodies for NONO, SFPQ, Myc and DAXX. pc

4.2.6 Depletion of SFPQ leads to a dislocation of DAXX from telomere

A recent study showed that the ATRX/DAXX complex is enriched in PML bodies (Ivanauskiene et al. 2014). PML bodies frequently associate with telomeres and other proteins in ALT positive cells, forming APBs (ALT-associated promyelocytic leukaemia bodies) that are considered to be the sites of homologous recombination mediated-telomere elongation. Remarkably, in the absence of PML, ATRX/DAXX complex loses its H3.3 loading ability (Ivanauskiene et al. 2014).

In order to have more detailed information about a possible SFPQ dependent localization of DAXX, we performed immunofluorescence experiments in U2-OS telomerase-negative cells depleted for SFPQ. We did not observed a change in DAXX/PML colocalization frequency upon depletion of SFPQ (Figure 4.2.5.1 A,B) . However, lack of SFPQ caused a more dispersed DAXX staining. We hypothesized that, although a significant fraction of DAXX still remained localized to PML bodies, a large portion of DAXX is dispersed throughout the nucleus.

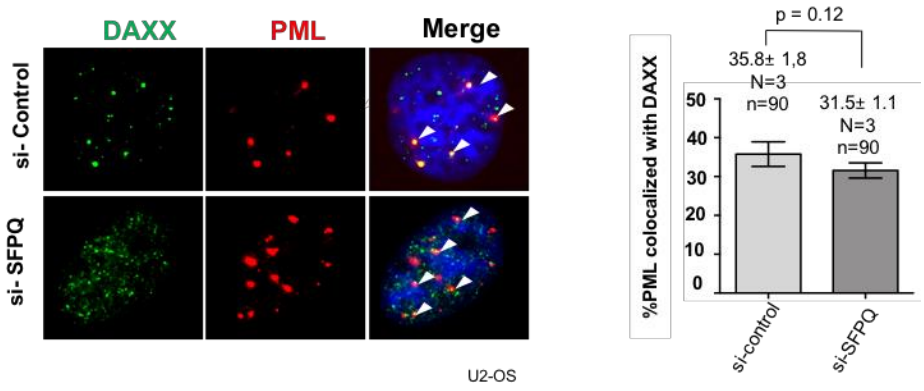
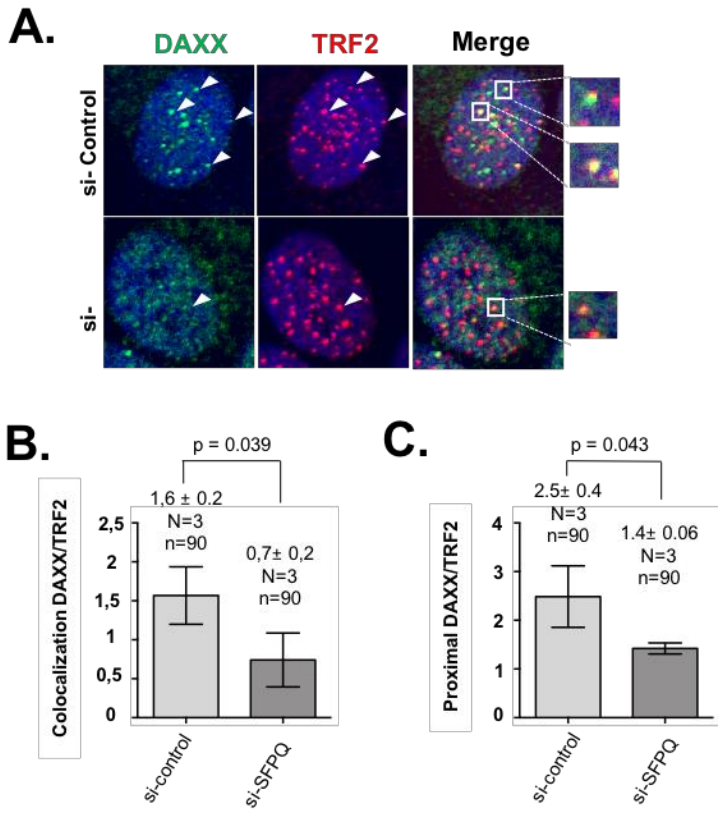


Figure 4.2.5.1 Co-immunofluorescence using PML and DAXX specific antibody in U-2 OS cells transiently depleted for SFPQ Right panel. Quantification of the percentage of colocalization between PML and DAXX foci. **Left panel.** Representative images anti-DAXX and anti-PML co-immunofluorescence in U-2 OS cells transfected with indicated siRNAs. in panels right mean values and standard deviations are reported, error bars indicated standard deviation. N=number of independent experiments. n= number of analyzed nuclei. A student t-test was used to calculate statistical significance; p-values are shown

The next step was to investigate as to whether SFPQ has a role in the linking of DAXX to the telomere

To address whether the binding of DAXX to telomere required SFPQ, we performed immunofluorescence experiments in U2-OS telomerase-negative cells depleted for SFPQ. The colocalizing spots and the proximal spots of DAXX signals with TRF2 were evaluated

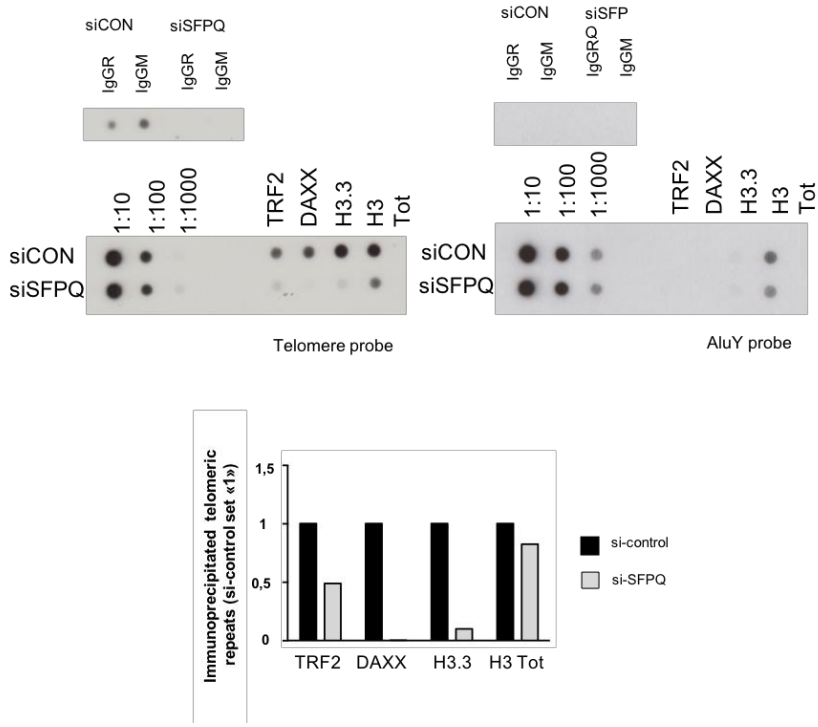
We observed that SFPQ depletion causes a reduction both in colocalization and in proximal telomere spots.



4.2.5.2 Loss of SFPQ causes the dispersion of DAXX A. Representative image of TRF2 combined with anti-DAXX antibody. **B.** Quantification of colocalization of DAXX with telomeric protein TRF2 C. quantification of proximal DAXX to TRF2. For quantifications in panels Band C mean values and standard deviations are reported, error bars indicated standard deviation. N=number of independent experiments. n= number of analyzed nuclei. A student t-test was used to calculate statistical significance; p-values are shown

This suggests that SFPQ is required for localization of DAXX to telomere, but also telomere proximal DNA regions.

To confirm these data, we have performed telomeric ChIP using specific antibodies agaist DAXX and H3.3 (Figure 4.2.5.1) Mouse/rabbit immunoglobulins from preimmunesera (IgG M, IgG R), TRF2 H3 antibody have been used as negative and positive control, respectively. (Figure 4.2.5 A). We also excluded binding to AluY repeats. As expected, TRF2 is enriched at telomeric repeats. These results are indicative of the role of SFPQ in the recruitment of DAXX to the telomere.



4.2.5.3 Telomeric ChIP Chromatin immunoprecipitation experiments (ChIP) using U-2 OS cells and mouse anti-TRF2, rabbit anti-histone H3, rabbit anti-H3.3 and anti-DAXX antibodies. Mouse and rabbit control IgGs (IgG M/IgG R) were used as negative control. Serial dilutions of chromatin extract (input) prepared from U-2 OS cells were loaded. Left panel, representative images.

4.2.7 Altered telomere homeostasis and genomic instability in cells depleted for novel SFPQ interactors

Results emerged from Immunofluerescence and ChIP experiments, has further strengthened our conviction that DAXX needs SFPQ to be recluted at telomere and solve his role as chromatin regulator.

To assess whether DAXX depletion resulted in the same phenotype observed in SFPQ knockdown conditions, we decided to carry out preliminary experiments to confirmed and evaluated the role of DAXX in DNA damage activation, in induction of replication defects and telomere dysfunction (He et al. 2015; Voon and Wong 2015)

4.2.8 Loss of DAXX impact on DNA damage activation and replicative stress

To evaluated the activation of DNA damage pathway at telomeres, we performed DNA FISH in combination with anti-γH2AX antibody in U-2 OS cells depleted for DAXX. We observed a significat increase of colocalization of γH2AX with Telo signals. Co immunofluorescence of RPA32pSer33 and

TRF2 revealed an increase in replication defect formation observed as an increased recruitment of of RPA32pSer33 at telomeres in DAXX depletion condition.

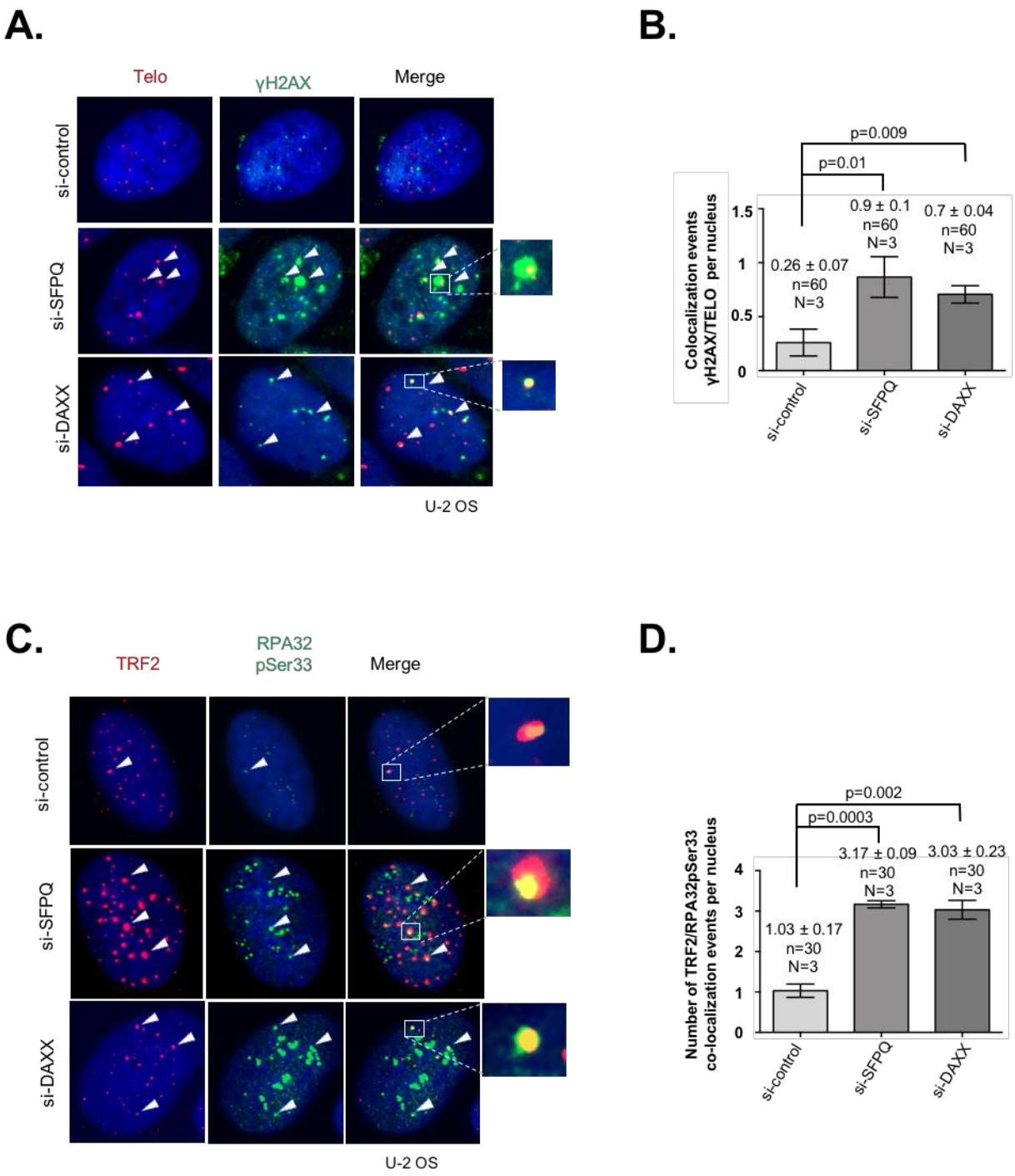


Figure 4.1.6. DAXX prevents DNA damage and replicative defects at telomeres. A. Representative images of combined immunofluorescence using anti- anti- γH2AX antibodies with Q Fish in U-2 OS cells transfected with indicated siRNAs. Arrowheads indicate co-localization events. **B** Depletion of SFPQ and DAXX significantly increases the colocalization events between γH2AX and Telo probe **C.** Immunofluorescence with anti-TRF2 and anti-RPA32pSer33 antibodies was performed in U2-OS cells depleted for SFPQ **D**. Representative images are shown. Depletion of SFPQ and7OR NONO significantly increases the number of TRF2-RPA32pSer33 colocalizations per nucleus. **E.** Representative images of combined immunofluorescence using anti-TRF2 and anti-pATR antibodies (S428) in U-2 OS cells transfected with indicated siRNAs. Arrowheads indicate co-localization events. . N=number of independent experiments; n=number of analyzed nuclei. Means (bars) and SDs (error bars) are reported. A student t-test was used to calculate statistical significance; p-values are shown.

4.2.8.1 DAXX depletion leads to RNA:DNA hybrids formation

We next tested wheter loss of DAXX reproduces the RNA:DNAhybrids phenotypeof siSFPQ cells. Performed a co immunofluorence experiment combining TRF1 and anti-S9.6 antibodies. We obserbed a significant increase of hybrids at telomere both in siDAXX and siSPFQ condition when compared to control cells.

This experiment was supported by DRIP assay. U-2 OS transiently transfected with control, SFPQ and DAXX specific siRNAs, were collected 72 hours after transfection. Whole cell extracts were split: half were directly used for immunoprecipitation, half were treated with RNase H (negative control).After DNA extraction, chromatin was incubated with S 9.6 antibody.Mouse immunoglobulins from preimmunesera (IgG M), have been used as negative. DNA was purified from immunoprecipitated chromatin, transferred to a nitrocellulose membrane and subjected to Southern blotting using a radioactive labelled telomere probe. We show that depletion of DAXX and SFPQ induce the increase of hybrids to the telomereare, thus supporting previous data from confocal microscopy.

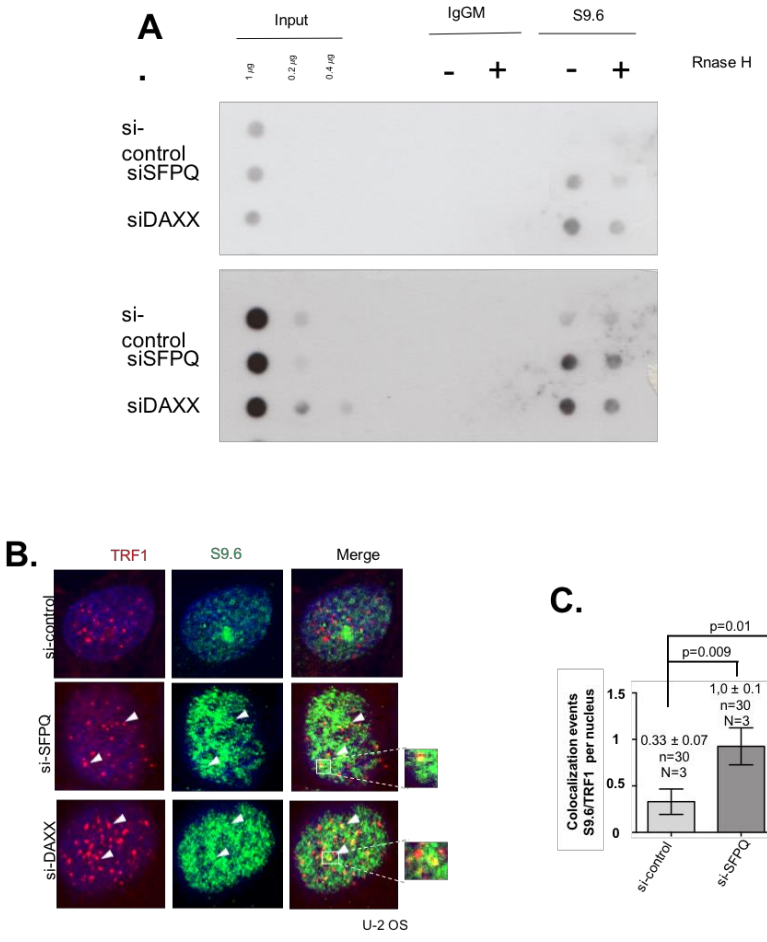


Figure 4.1.4.2b. DAXX suppress formation of telomeric RNA:DNA hybrids.A. DRIP assay show an increase of RNA;DNA hybrids in SFPQ and DAXX knock down condition **B**. U2-OS cells were transfected with DAXX and RNase H1 specific siRNAs and immunofluoresce

experiments with anti-TRF1 and anti-RNA:DNA hybrids (S9.6) antibodies were performed. **C** Depletion of SFPQ increases the number of RNA:DNA hybrids that co-localize with TRF1. Similar effect is induced by knockdown of RNase H1 that eliminates stretches of RNA from RNA:DNA hybrids. N= number of independent experiments; n=number of analyzed nuclei. Means (bars) and SDs (error bars) are reported. A student t-test was used to calculate statistical significance; p-values are shown.

4.2.8.2 DAXX as regulator of telomere strand fragily and suppressor of recombination

We found that the depletion of DAXX can lead to an accumulation of DNA damage and replicative stress markers at telomeres and we observed an increase in the formation of telomere hybrids. In this scenario, we wanted to understand the importance of DAXX for telomere integrity. We performed CO-FISH experiment in U-2 OS cells under SFPQ and DAXX knock down conditions. As shown in the first part of the thesis, U-2 OS cells depleted for SFPQ show a reduction in leading strand fragility and do not have any impact on lagging strand fragility (see section 4.1.7). As expected, DAXX knockdown cells show a significat reduction in leading strand fragility. DAXX, as well as SFPQ, have a role in suppress telomere fragility

Recombination events in depleted cells for DAXX were next evaluated. Is has been observed an increase of T-SCE frequency. Several studies have shown that DAXX is Important in maintaining telomere chromatin. loss of DAXX causes telomeric deprotation leading to the formation of hybrids and possibility of formation of secondary DNA structures (G-Quadruplex).

this leads to the stall of the replication fork and creates a substrate for activating the homologous recombination pathway.

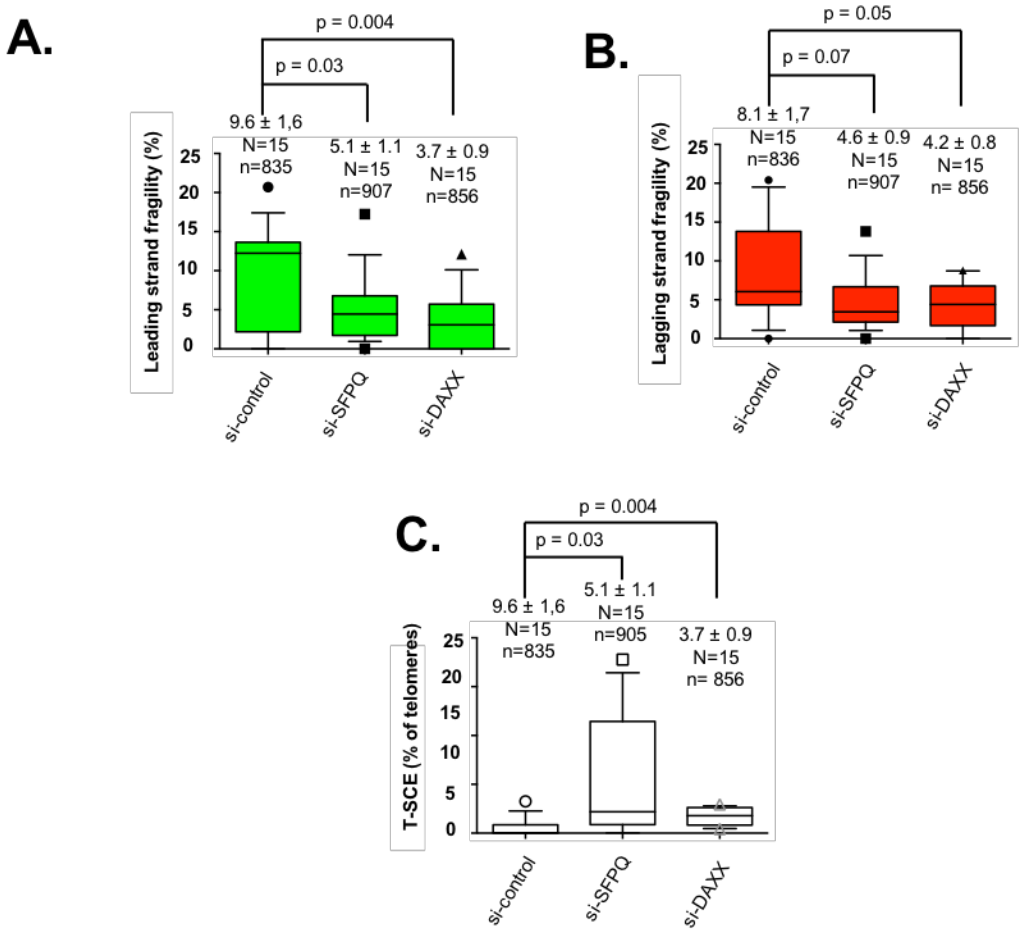


Figure 4.1.7.1. DAXX regulate telomere fragility and suppresses homologous recombinationat telomeres. Cells were transfected with SFPQ , DAXX control siRNAs and after 72 hours CO-FISH analysis was performed **A-B.** Quantification of leading strand fragility (A) and lagging strand fragility (B) in U-2 OS cells transfected with DAXX, SFPQ specific siRNA or a non-targeting control siRNA. **C.** -SCE frequency was measured as the percentage of T-SCE events per number of chromosome ends in one metaphase.Metaphases from three independent experiments were analyzed. Box blot diagrams B-E: middle line represents median; boxes extend from the 25th to 75th percentiles. The whiskers mark the 10th and 90th percentiles. p values were calculated using a Mann-Whitney test. Median fragility values and standard deviation are indicated; n=number of analyzed telomere repeat signals, N= number of analyzed metaphase spreads

The variation of telomere length is linked to homologous recombination, for this reason we have performed DNA fish experiment to evaluate the role of daxx in telomere length homeostasis. We have observed an increase in telomere length both in SFPQ and in DAXX depleted cells.

This data was in line with the increase of T-SCE frequency under these conditions.

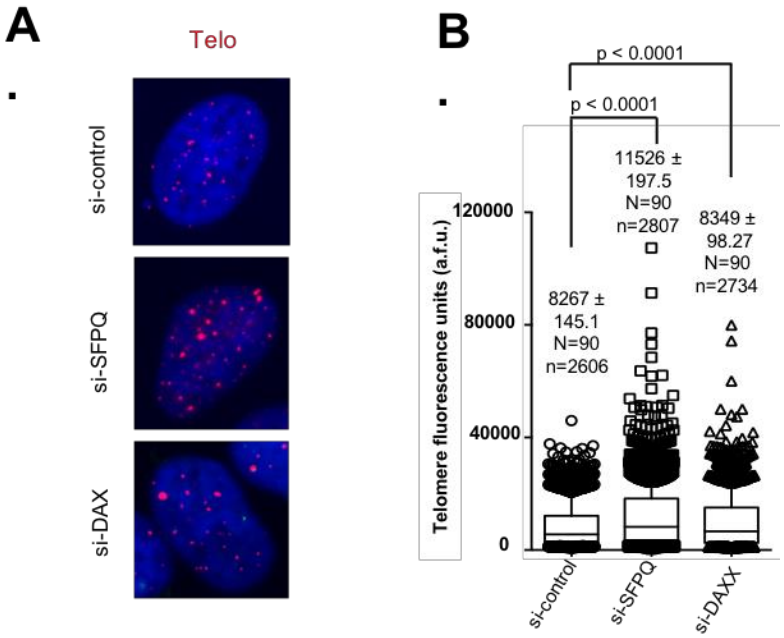


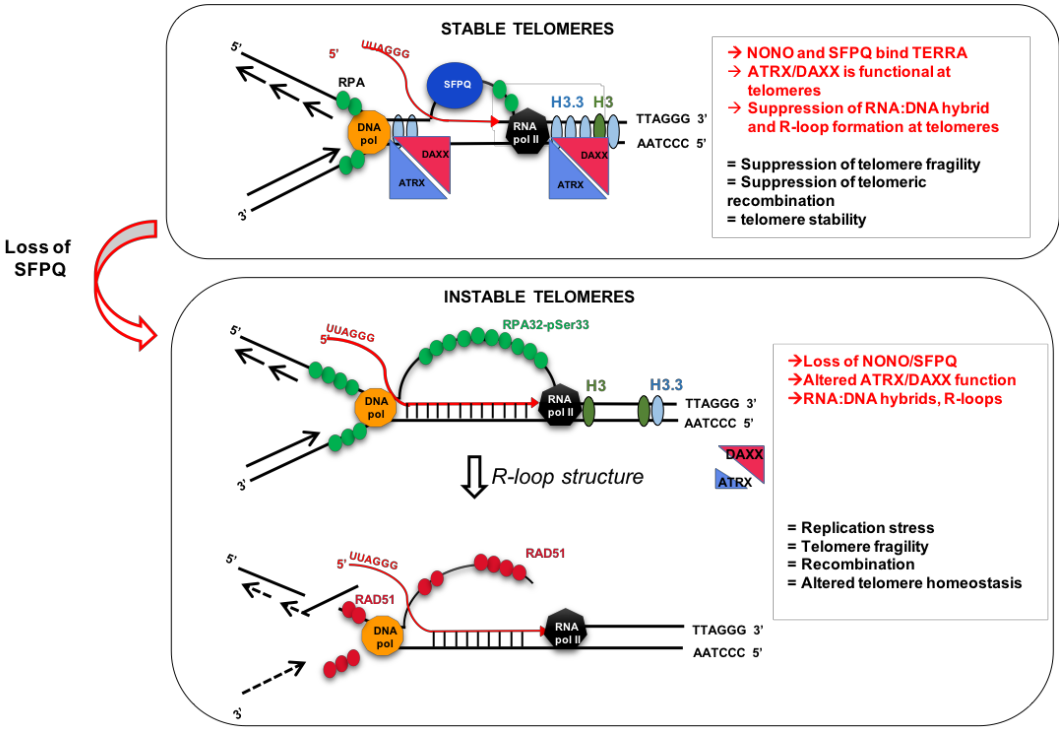
Figure 4.1.8. DAXX regulate telomere length in human cancer cells. A. U2-OS cells were transfected with control, SFPQ and DAXX siRNAs. After 72 hours, telomere length measurements were performed by quantitative telomere DNA FISH on interphase cells. Representative images are shown **B.** Telomere fluorescence intensity analyzed for each telomere of (A). Box blot diagrams B: middle line represents median arbitrary fluorescence units (a.f.u.), boxes extend from the 25th to 75th percentiles. The whiskers mark the 10th and 90th percentiles. p values were calculated using a Mann-Whitney test. Median a.f.u. values with standard deviation are indicated; n=number of analyzed telomere repeat signals, N= number of analyzed interphase nuclei. N=number of analyzed nuclei; n= total signals analyzed; Arbitrary fluorescence units (a.f.u.) are shown

4.2.9 WORKING MODEL

In the second part of my thesis, we discovered DAXX as a novel SFPQ interacting protein

We propose that SFPQ in involved in the recruitment of DAXX at telomere and togheter can prevents RNA:DNA hybrids formation at telomeres stabilizing telomere structure.

Loss of DAXX leads to accumulation of RNA:DNA hybrids structure and consequent displacement of non-template strand (G-strand) that remains unpaired and exposed as ssDNA bound by RPA32pSer33 (Figure 4.2.9). This lead to an accumulation of DNA damage marker γH2AX at telomere as index of telomeric fragility associated withT-SCE rates and increased telomere length in recombination proficient ALT cells.



5. Discussion

Telomeres are nucleoprotein structures located at the ends of eukaryotic chromosomes that ensure genome integrity. Despite their heterochromatic structure, telomeres are transcribed by RNA polymerase II giving rise to the long noncoding RNAs TERRA (telomeric repeat-containing RNA).

TERRA physically interacts with telomeric chromatin and plays crucial roles in the regulation of telomere homeostasis and telomere function, telomere length regulation, DNA replication, telomere recombination, DNA damage response and chromatin structure. Recent studies indicate that TERRA transcripts can form DNA-RNA hybrids at chromosome ends that can promote telomere instability.

Genomic instability is a hallmark of human cancer formation and progression. Persistent RNA:DNA-hybrids created during transcription causes the progressive displacement of the un-transcribed DNA strand, creating so called “R-loop” structures. R-loops are prone to breakage and mutagenesis, provoke replication stress, drive chromatin condensation and are highly recombinogenic. Thus, R-loops and RNA:DNA-hybrids are anticipated to represent an important source for genomic instability in human cancer. RNA:DNA-hybrid and R-loop formation is a frequent event across the genome, however the mechanisms of R-loop management and the direct contribution of R-loops to genomic instability and cancer progression is poorly understood.

Using a TERRA pull-down approach we identified NONO and SFPQ as novel TERRA interacting proteins that have a central role in controlling telomere function by suppressing RNA:DNA hybrids formation at telomeric repeats in cancer cells. Loss of function of these novel TERRA interacting proteins result in telomere replication stress, telomere recombination and telomere instability.

The purpose of this work is to shed light on the management of the RNA:DNA hybrids using the telomere as a study model. In this regard we have investigated new TERRA interacting proteins with possible involvement in the homeostasis of the formation of hybrids and regulation of telomere functions.

5.1 Project 1: SFPQ complex safeguards telomere stability by preventing TERRA-telomere hybrids

Long non-coding RNAs (lncRNAs) have been demonstrated to play a central role in controlling genome function by diverse mechanisms. LncRNAs impact on chromatin structure through the recruitment of chromatin remodeling factors, regulating protein activity and acting as a scaffold to target enzymatic activities to their sites of action (Morris K.V. et al. 2014). The telomeric repeat containing non-coding RNA (TERRA) performs all these activities at telomeres, thus regulating telomere homeostasis in a very dynamic manner.

In particular, RNA association with telomeric heterochromatin can target accessory factors to chromosome ends via RNA-protein interactions (Deng Z. et al. 2009; Arnoult N. et al. 2012; Porro A. et al. 2014; de Silanes I.L. et al. 2010; Takahama K. et al. 2013; Sheibe M. et al. 2013).

Performing a TERRA RNA pull down followed by mass spectrometry, we identified two novel TERRA interacting proteins: Splicing factor proline- and glutamine- rich (SFPQ) and the Non-POU domain-containing octamer-binding protein (NONO). SFPQ and NONO are members of the Drosophila Behavior Human Splicing (DBHS) family proteins and can form homo- and heterodimers and have a pleiotropic function in RNA metabolism, gene regulation and DNA damage response (Shav-Tal Y. and Zipori D. 2002; Bond C.S. and Fox A.H. 2009; Passon D.M. et al. 2012). Initially identified as factors involved in splicing and transcription regulation, to date SFPQ and NONO have been demonstrated to act as a multifunctional regulator complex of RNA and DNA metabolism, including an important role in DNA damage response and repair (Patton J.G et al. 1991; Hallier M. et al. 1996; Basu A. et al. 1997; Salton M. et al. 2010; Rajesh C. et al. 2011; Li S. et al. 2014; Alfano L. et al. 2015; Yarosh C.A. et al. 2015).

We found that SFPQ and NONO associate with telomeric heterochromatin as determined by ChIP and immunofluorescence.

Depletion of SFPQ induced a substantial increase in the number of TERRA foci per nucleus. Northern blotting and quantitative RT-PCR show that the observed accumulation of TERRA in the nuclei of SFPQ-depleted cells is not due to the increased steady state TERRA levels. Together this indicates a role of SFPQ in regulating TERRA at chromosome ends rather than TERRA metabolism. Alterations of TERRA abundance at telomeres has been linked to telomere dysfunction and DNA damage response at telomeres (de Silanes L.I. et al. 2014; Porro A. et al. 2014). We found that depletion of SFPQ leads to an increase localization of DNA damage marker γH2AX with nuclear TERRA foci and increased frequency of dysfunctional telomeres at metaphase chromosomes. Our results show that SFPQ are important to prevent DNA damage response activation at telomeres in U2-OS human cancer lines. This underlines a general role of SFPQ in telomere protection of telomerase negative ALT cells (U2-OS) cancer cells.

Summarizing, our results indicate that SFPQ prevent DNA damage activation at telomeres, thus indicating a role in stabilizing telomere function.

Telomere dysfunction has been reported to be the consequence of critical telomere shortening, loss of shelterin components or defective telomere replication (Palm W. and de Lange T. 2008; Martínez P. et al. 2009; Zimmermann M. et al. 2014). ATR activation in SFPQ knockdown cells suggests the formation of single-stranded DNA (ssDNA) ruptures at telomeres. ATR is activated by the presence of single-stranded DNA (ssDNA) coated by replication protein A (RPA) often generated at stalled

replication forks that are susceptible to DNA breaks (Zhou B.B. and Elledge S.J. 2000; Jazayeri A. et al. 2006). We hypothesized that aberrant accumulation of TERRA transcripts at telomeres might impede progression of DNA replication fork resulting in ssDNA damage. This was confirmed by increased phosphorylated RPA32, a marker of replicative stress at telomeres in SFPQ-depleted cells. Increased number of co-localizations of TERRA foci with phosphorylated RPA32 in U2-OS cells depleted for SFPQ suggest that replication defects are associated with the accumulation of TERRA at telomeres.

R-loops are structures triggered by RNA:DNA hybrids resulting in the displacement of the non-template strand. RNA:DNA hybrids are prone to be formed in genomic regions with negative supercoiling and high G-content (Roy D. and Lieber M.R. 2009). In addition, the presence of nicks or G-quadruplexes in the displaced ssDNA favors R-loop formation (Duquette M.L. et al. 2004; Roy D. et al. 2010). The existence of telomeric R-loops has been demonstrated in both yeast and human cells (Balk B. et al. 2013; Pfeiffer V. et al. 2013; Arora R. et al. 2014). Importantly, we demonstrate that depletion of SFPQ increases the accumulation of TERRA-telomere hybrids in U2-OS cells. The associated effects of replication damage can be rescued by RNase H1 overexpression. This demonstrates that telomeric replicative stress induced by depletion of SFPQ is mediated by TERRA-telomere hybrids formation. Mechanisms that control RNA:DNA hybrids formation regulate R-loops abundance. These include factors involved in RNA biogenesis and metabolism such as THO complex, the splicing factor SRSF1 and exoribonucleases exosome component 3 and 10 (EXOSC3 and EXOSC10) as well as DNA-RNA helicases such as senataxin (SETX) and topoisomerase 1 (TOP1) that resolve the negative supercoiling behind RNA pol II progression (Tuduri S. et al. 2009; Gomez-Gonzalez B. et al. 2011; Santos-Pereira J. M. and Aguilera A. 2015). In addition, resolution of R-loops is mediated by degradation activity of RNase H1 (Wahba L. et al. 2011). SFPQ does not have enzymatic activity. Thus, we propose that SFPQ may recruit factors with enzymatic activities that resolve RNA:DNA hybrids.

R-loops are a source of genomic instability and chromosome fragility by different mechanisms. The displaced single strand DNA of R-loops structures can act as a substrate to DNA-damaging agents, deaminases (activation-induced cytidine deaminase) and repair enzymes (base excision repair), leading to DNA lesions and nicks (Yu K. et al. 2003). In addition, R-loops can interfere with DNA replication by blocking the progression of replication forks, thus generating DNA lesions and DSBs (Gan W. et al. 2011; Aguilera A. and Garcia-Muse T. 2012; Skourti-Stathaki K. et al. 2014). Importantly, telomeric RNA:DNA hybrids causing replicative stress at telomeres impacts on telomere fragility as demonstrated by the fact that overexpression of RNase H1 was shown to reduce telomere fragility of ALT positive cells (Arora R. et al. 2014).

CO-FISH analysis revealed that SFPQ acts as barrier to homologous recombination at telomeres. Loss of SFPQ drives RNA:DNA hybrid formation, impaired replication and results in dramatically increased recombination frequencies in the context of ATR phosphorylation at telomeres of telomerase positive and ALT cancer cells. This effect is paralleled by a reduction of telomere fragility, leading to the interesting speculation that homologous recombination may help to rescue telomere fragility. Our data show that SFPQ functions as barrier to homologous recombination at telomeres. Phosphorylation of ATR in SFPQ knock-down cells is in line with a recent report that demonstrated that activation of ATR promotes homologous recombination (Buisson et al. 2017).

Telomeric hybrids have been demonstrated to allow telomere length maintenance in ALT positive cells by promoting telomeric homologous recombination (Arora R. et al. 2014; Yu T.Y. et al. 2014). In line with this, our Q-FISH results demonstrate that depletion of SFPQ increases telomere length in U2-OS cells. Increased telomere length induced by SFPQ depletion appears as a direct consequence of increased ALT activity. Taken together these data suggest that SFPQ is important to regulate ALT activity and telomere length, thus preventing telomere instability in telomerase–negative U2-OS cells.

In conclusion, our work identified a new TERRA interacting complex composed by SFPQ that acts as novel regulator of telomere homeostasis. Importantly, we demonstrated that the SFPQ is essential to suppress TERRA-telomere hybrids formation thereby preventing telomere fragility and telomere recombination (via SFPQ). This represents a new molecular mechanism that is essential to safeguard telomere stability in human cancer cells.

Telomere maintenance and genomic instability represent two hallmarks of cancer.

Importantly, our data demonstrate that the SFPQ strongly impacts on telomere length homeostasis, fragility and recombination, thus ensuring telomere stability. Genomic instability induced by replication stress is considered an early step in cancer formation (Gailard H. et al. 2015). High levels of SFPQ might protect cells from telomeric replication stress, fragility and uncontrolled recombination, thus promoting cancer cell proliferation, allowing tumor progression. In contrast, low levels of SFPQ may increase cell sentitivity to replication defects and DNA damaging agents. In this scenario, tumor cells with low SFPQ might be more susceptible to inhibitors of DNA synthesis, including PCNA (proliferating cell nuclear antigen) and BLM (Bloom helicase) inhibitors or G-quadruplexes stabilizer agents (Punchihewa C. et al. 2012; Nguyen G.H. et al. 2013; Zimmer J. et al. 2016). Interestingly, our complex acquires a great importance in the context of ALT tumors where telomere length maintenance depends on the activation of telomere recombination. ALT telomeres require a balance of pro- and anti-recombinogenic signals. In fact, inhibition of recombination would lead to loss of telomere sequence and block of cell proliferation.

New mechanistic insights into the regulation of ALT activity by SFPQ may highlight vulnerabilities of ALT cells in order to develop therapeutic strategies.

5.2 Project 2: Identification of novel SFPQ interactors

SFPQ is a member of DBHS family associated with nuclear paraspeckles (Fox et al. 2002). It is a multifunctional protein (Shav-Tal and Zipori 2002a) originally isolated from the spliceosome, known to be involved in pre-mRNA splicing (E. Rosonina et al. 2005), transcriptional repression and DNA repair.

In order to better characterize and understand the specific mechanism of action of SFPQ, we performed Co-Immunoprecipitation experiments, paired with Mass-Spectrometry analysis, leading us to identify novel SFPQ interactors.

DAXX is a histone chaperone mediates the incorporation of the histone variant H3.3 into telomeric and pericentromeric chromatin. This protein can work in an independent manner or in cooperation with a protein known as Alpha Thalassemia/Mental Retardation Syndrome X-Linked (ATRX), known to be a chromatin remodeling factor.

We have created a network with a subset of proteins identified in mass spectrometry (direct eighteen SFPQ- and two DAXX-interactors). The ATRX protein, not present in mass spectrometry output, was also included in the network analysis (23 interacting proteins). The ATRX/DAXX complex is enriched at telomeric repeats where it has been shown to regulateH3.3 deposition (Goldberg et al. 2010; Law et al. 2010; Lewis et al. 2010; Wong 2010) (Fig. 2A). ATRX depletion impairs the heterochromatic state of the telomeres, because of a reduction in H3.3 incorporation, which leads to telomere destabilization (Wong 2010; Watson et al. 2013; Ramamoorthy and Smith 2015; Watson et al. 2015). This predicted interaction has been confirmed experimentally by co immunoprecipitation experiments performed in U-2 OS cells transiently transfected with Myc-tagged SFPQ and HA-tagged DAXX (Figure 4.2.2). Then, we moved to the identification of the domain of SFPQ able to interact with DAXX. It is known that the structure of SFPQ is conserved and contain a N-teminal domain formed by three subdomains: RGG, RGG box; P,Q, proline/glutamine-rich and P, proline-rich domain subdomain (P,Q) (Lee et al., 2015; Patton et al., 1993; Peng et al., 2002; Shav-Tal and Zipori, 2002; Urban et al., 2000); tandem RNA recognition motif domains (RRM1 and RRM2), a family-specific NOPS (Nono, Pspc1, and Sfpq) domain and a coiled-coil region (Passon et al., 2012). Several CO-IPs were performed using deleted constructs of SFPQ, in order to better characterize the domain of interaction between SFPQ and DAXX (Figures 4.2.3.2.1 and 4.2.3.2.2). CO-IP experiments have shown that the fundamental domain for the link between SFPQ and DAXX is the N-terminal domain. This domain consists of three sub-domains: RGG box, PR domain and P domain. Among the various

transfected mutants, we used SFPQ constructs lacking of RGG and P/Q domains which showed no changes in the SFPQ-DAXX binding. Therefore, we can conclude that the fundamental domain for the interaction between SFPQ and DAXX is the P domain. A study of exome sequencing on patients with osteosarcoma revealed 2 mutations in SFPQ (Kovac et al. 2015). These mutations, insertions of three (AGC) and four (AGCG) nucleotides respectively, are located in the P domain. The first mutation causes the loss of a threonine in position 150 of the amino acid sequence of SFPQ and an acquisition of a lysine and a proline. The second mutation leads to the variation of the reading frame, starting from the amino acid 148 in the P domain. Thus, many premature stop codons are formed, the first one (TAA) is found in position 200 of the amino acid sequence of SFPQ. In order to deeply understand the impact of these SFPQ mutated protein in patients with osteosarcoma, we investigated whether these mutations impair interaction between SFPQ and DAXX.

We performed Co-immunoprecipitations in U-2 OS transiently transfected with Myc-tagged SFPQ full length version and Myc-tagged SFPQ mutant constructs. Western blot experiments showed that SFPQ mutation does not affect the interaction between SFPQ and DAXX. The presence of endogenous SFPQ could be due to the ability of DBHS proteins in forming homodimers. Therefore, endogenous SFPQ could be linked to the exogenous mutated protein.

Further experiments involving the use of cells stably expressing the mutated versions of Myc SFPQ will give us is more information on the real interaction between the SFPQ mutants and DAXX.

Recent study showed that the ATRX/DAXX complex localizes in PML bodies (Ivanauskiene et al. 2014). PML bodies frequently associate with telomeres in ALT positive cells forming APBs (ALT-associated promyelocytic leukaemia bodies) that are considered to be the sites of Homologous Recombination mediated-telomere elongation. Remarkably, in the absence of PML, ATRX/DAXX complex loses its H3.3 loading ability. Delbarre E. et al identified a propose model for the organization of chromatin domains involving a PML-ATRX-H3.3 axis. In wilde type cells, PML interacts with H3K9me3-rich domains where H3.3 can be incorporated, yet at low level, by HIRA, DAXX, or ATRX. FRAP and ChIP analyses suggest that PML restricts access of H3.3 to regions of low transcriptional activity, maybe due to the low nucleosome turnover (Huang and Zhu 2014). In the absence of PML, deposition of H3.3 in PADs may reflect higher nucleosome turnover.

The ATRX/DAXX complex loses its H3.3 loading ability. This could be due to the need ofATRX/DAXX to associate with PML in PADs in order to function as an H3.3 chaperone. Alternatively, ATRX association with chromatin is decreased due to the lower H3K9me3 level in these regions. In contrast, HIRA retains its H3.3 loading capacity in PADs in the absence of PML, leading additional support to the idea of a gap-filling process delivering H3.3 to regions of altered chromatin integrity (Ray-Gallet et al. 2011; Adam et al. 2013; Ivanauskiene et al. 2014). The loss of PML results in a remodeling of

the chromatin state of PADs through a shift from H3K9 toward H3K27me3, maybe as a way to preserve heterochromatin integrity.

The ATRX/DAXX complex is enriched at telomeric repeats where it has been shown to regulateH3.3 deposition (Goldberg et al. 2010; Law et al. 2010; Lewis et al. 2010; Wong 2010) (Fig. 2A). ATRX depletion impairs the heterochromatic state of the telomeres due toa reduction in H3.3 incorporation, which leads to telomere destabilization (Wong 2010; Watson et al. 2013; Ramamoorthy and Smith 2015; Watson et al. 2015).

Dysfunction in the ATRX-DAXX complex leads to the formation of non-B DNA structures like these G4-DNA structure that creates a recombinogenic nucleoprotein structure for homologous recombination (HR) at the telomeres and thereby facilitates the ALT development (Law et al. 2010).

In order to understand if SFPQ could affect DAXX localization, we performed a co immunofluorescence experiments in U-2 OS depleted for SFPQ. We do not observe changes in colocalization between DAXX and PML. However, depletion of SFPQ caused a more dispersed DAXX staining. We hypothesized that, although a significant fraction of DAXX still remained localized to PML bodies, a large portion of DAXX is dispersed throughout the nucleus. Whereas, co immunofluorescence has been shown a reduction of DAXX from telomeres and telomere proximal region. Telomeric ChIP experiments, using specific antibodies against DAXX and histone variant H3.3, reveal a reduction of DAXX binding to telomeric sequences. Therefore, the lack of SFPQ affects the binding of DAXX to the telomere without causing significant dislocation of DAXX from PML bodies. This might lead us to speculate that a possible SFPQ function is to recruit DAXX to the telomere in order to maintain the correct chromatin conformation.

We have knockdown DAXX has been knocked down in U2-OS cells and classic telomere features such as telomere fragility (CO-FISH), telomere sister chromatid exchange (T-SCE, recombination), telomere length, telomere damage (Telomere-γH2AX-immunoFISH) and telomere replicative defects (Telomere-RPAp32Ser33) have been investigated. As expected, we found that the depletion of DAXX can reproduce the phenotype observed in lack of SFPQ.

Recent studies have shown that DAXX binds SUV39H1, major histone methyltransferases for H3K9. Repressive epigenetic marks such as H3K9me3 are highly enriched on LTRs and telomeres. It is possible that H3K9me3 enrichment may prime telomeres for the recruitment of DAXX/ ATRX, similar to what occurs at pericentric heterochromatin (Kourmouli et al., 2005). Therefore, DAXX appears to play an active role in maintaining H3K9me3 at telomeres. Depletion of DAXX resulted in the loss of telomeric localization of both ATRX and H3K9me3. (He et al. 2015)

In line with this, we found that SFPQ knock-down in U-2 OS triggers increased T-SCE frequencies thus resulting telomere elongation.

Together, this identifies SFPQ and DAXX as regulators of telomere length homeostasis by controlling RNA:DNA hybrid abundance at telomeres. Given the relevance of RNA:DNA hybrids in recombination and telomere length homeostasis, SFPQ and DAXX are expected to be of special relevance in human tumors that maintain telomeres length via the ALT pathway. In line with this, SFPQ mutations have been recently discovered by exome sequencing in human osteosarcoma, a tumor type that is reported to use recombination based ALT as predominant pathway for telomere maintenance[53]. We hypothesize that NONO and SFPQ act as platforms to recruit factors that resolve RNA:DNA hybrid structures thereby suppressing fragility and recombination at telomeres (Fig. 7). Understanding the function of the entire SFPQ/DAXX complex at telomeres is expected to open new avenues in understanding the suppression of the ALT pathway in telomerase negative tumors and activation of the ALT pathway in telomerase negative cells. These insights are expected to be highly relevant for potential telomere based anti-cancer therapies.

6. Bibliography

Aguilera, Andrés, and Tatiana García-Muse. 2012. "R Loops: From Transcription Byproducts to Threats to Genome Stability." *Molecular Cell*.

Ahmad, Kami, and Steven Henikoff. 2002. "The Histone Variant H3.3 Marks Active Chromatin by Replication-Independent Nucleosome Assembly." *Molecular Cell*.

Akincilar, Semih Can, Bilal Unal, and Vinay Tergaonkar. 2016. "Reactivation of Telomerase in Cancer." *Cellular and Molecular Life Sciences*.

Arora, Rajika et al. 2014. "RNaseH1 Regulates TERRA-Telomeric DNA Hybrids and Telomere Maintenance in ALT Tumour Cells." *Nature Communications* 5.

Artandi, Steven E, and Ronald A Depinho. 2010. "Telomeres and Telomerase in Cancer." *Carcinogenesis* 31(1): 9–18. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3003493/pdf/bgp268.pdf (September 9, 2018).

Azzalin, Claus M., and Joachim Lingner. 2008. "Telomeres: The Silence Is Broken." *Cell Cycle* 7(9): 1161–65.

Azzalin, Claus M et al. 2007. "And RNA Surveillance Factors At." *Science (New York, N.Y.)* 318(November): 798–801.

Bailey, S. M. et al. 2001. "Strand-Specific Postreplicative Processing of Mammalian Telomeres." *Science*.

Balk, Bettina et al. 2013. "Telomeric RNA-DNA Hybrids Affect Telomere-Length Dynamics and Senescence." *Nature Structural and Molecular Biology*.

Basu, A, B Dong, A R Krainer, and C C Howe. 1997. "The Intracisternal A-Particle Proximal Enhancer-Binding Protein Activates Transcription and Is Identical to the RNA- and DNA-Binding Protein P54nrb/NonO." *Molecular and Cellular Biology*.

Basu, Uttiya et al. 2011. "The RNA Exosome Targets the AID Cytidine Deaminase to Both Strands of Transcribed Duplex DNA Substrates." *Cell*.

Benetti, Roberta, Marta García-Cao, and María A. Blasco. 2007. "Telomere Length Regulates the Epigenetic Status of Mammalian Telomeres and Subtelomeres." *Nature Genetics*.

Blackburn, Elizabeth H., and Kathleen Collins. 2011. "Telomerase: An RNP Enzyme Synthesizes DNA." *Cold Spring Harbor Perspectives in Biology*.

Blackburn, Elizabeth H., Elissa S. Epel, and Jue Lin. 2015. "Human Telomere Biology: A Contributory and Interactive Factor in Aging, Disease Risks, and Protection." *Science*.

Blasco, M A, W Funk, B Villeponteau, and C W Greider. 1995. "Functional Characterization and Developmental Regulation of Mouse Telomerase RNA." *Science*.

Bond, Charles S., and Archa H. Fox. 2009. "Paraspeckles: Nuclear Bodies Built on Long Noncoding RNA." *Journal of Cell Biology*.

Broccoli, Dominique, Agata Smogorzewska, Laura Chong, and Titia de Lange. 1997. "Human Telomeres Contain Two Distinct Myb-Related Proteins, TRF1 and TRF2." *Nature Genetics*.

Bryan, Tracy M. et al. 1997. "Evidence for an Alternative Mechanism for Maintaining Telomere Length in Human Tumors and Tumor-Derived Cell Lines." *Nature Medicine*.

Bryce, L a et al. 2000. "Mapping of the Gene for the Human Telomerase Reverse Transcriptase, HTERT, to Chromosome 5p15.33 by Fluorescence in Situ Hybridization." *Neoplasia (New York, N.Y.)* 2(3): 197–201. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1507564&tool=pmcentrez&rendertype=abstract.

Buisson, Rémi et al. 2017. "Coupling of Homologous Recombination and the Checkpoint by ATR." *Molecular Cell*.

Cairney, C. J., and W. N. Keith. 2008. “Telomerase Redefined: Integrated Regulation of HTR and HTERT for Telomere Maintenance and Telomerase Activity.” *Biochimie*.

Calderón, María Del Carmen, María Dolores Rey, Adoración Cabrera, and Pilar Prieto. 2014. “The Subtelomeric Region Is Important for Chromosome Recognition and Pairing during Meiosis.” *Scientific Reports* 4: 1–6.

Capper, Rebecca et al. 2007. “The Nature of Telomere Fusion and a Definition of the Critical Telomere Length in Human Cells.” *Genes and Development* 21(19): 2495–2508.

Castellano-Pozo, Maikel et al. 2013. “R Loops Are Linked to Histone H3 S10 Phosphorylation and Chromatin Condensation.” *Molecular Cell*.

Cesare, A J, and J D Griffith. 2004. “Telomeric DNA in ALT Cells Is Characterized by Free Telomeric Circles and Heterogeneous T-Loops.” *Mol Cell Biol*.

Cesare, Anthony J., and Roger R. Reddel. 2010. “Alternative Lengthening of Telomeres: Models, Mechanisms and Implications.” *Nature Reviews Genetics*.

Chan, Yujia A. et al. 2014. “Genome-Wide Profiling of Yeast DNA:RNA Hybrid Prone Sites with DRIP-Chip.” *PLoS Genetics*.

Chawla, R., and C. M. Azzalin. 2009. “The Telomeric Transcriptome and SMG Proteins at the Crossroads.” *Cytogenetic and Genome Research* 122(3–4): 194–201.

Chen, Liuh Yow, Sophie Redon, and Joachim Lingner. 2012. “The Human CST Complex Is a Terminator of Telomerase Activity.” *Nature*.

Chen, Yong et al. 2008. “A Shared Docking Motif in TRF1 and TRF2 Used for Differential Recruitment of Telomeric Proteins.” *Science*.

Cho, Nam Woo, Robert L. Dilley, Michael A. Lampson, and Roger A. Greenberg. 2014. “Interchromosomal Homology Searches Drive Directional ALT Telomere Movement and Synapsis.” *Cell*.

Chu, Hsueh Ping et al. 2017. “TERRA RNA Antagonizes ATRX and Protects Telomeres.” *Cell*.

Chung, I., H. Leonhardt, and K. Rippe. 2011. “De Novo Assembly of a PML Nuclear Subcompartment Occurs through Multiple Pathways and Induces Telomere Elongation.” *Journal of Cell Science* 124(21): 3603–18. http://jcs.biologists.org/cgi/doi/10.1242/jcs.084681.

Clery, A. et al. 2013. “Isolated Pseudo-RNA-Recognition Motifs of SR Proteins Can Regulate Splicing Using a Noncanonical Mode of RNA Recognition.” *Proceedings of the National Academy of Sciences*.

Clynes, David et al. 2014. “ATRX Dysfunction Induces Replication Defects in Primary Mouse Cells.” *PLoS ONE*.

———. 2015. “Suppression of the Alternative Lengthening of Telomere Pathway by the Chromatin Remodelling Factor ATRX.” *Nature Communications*.

Court, Robert, Lynda Chapman, Louise Fairall, and Daniela Rhodes. 2005. “How the Human Telomeric Proteins TRF1 and TRF2 Recognize Telomeric DNA: A View from High-Resolution Crystal Structures.” *EMBO Reports*.

Cusanelli, Emilio, Carmina Angelica Perez Romero, and Pascal Chartrand. 2013. “Telomeric Noncoding RNA TERRA Is Induced by Telomere Shortening to Nucleate Telomerase Molecules at Short Telomeres.” *Molecular Cell*.

d’Adda di Fagagna, Fabrizio et al. 2003. “A DNA Damage Checkpoint Response in Telomere-Associated Senescence.” *Nature* 426(6963): 194–98.

Darzacq, Xavier et al. 2006. “Stepwise RNP Assembly at the Site of H/ACA RNA Transcription in Human Cells.” *Journal of Cell Biology* 173(2): 207–18.

Dejardin, Jerome, and Robert Kingsto. 2012. “Purification of Proteins Associated with Specific Genomic Loci.”

NIH Public Access 136(1): 175–86.

Denchi, Eros Lazzerini, and Titia De Lange. 2007. "Protection of Telomeres through Independent Control of ATM and ATR by TRF2 and POT1." *Nature* 448(7157): 1068–71. https://www.nature.com/articles/nature06065.pdf (September 9, 2018).

Doksani, Ylli, John Y. Wu, Titia de Lange, and Xiaowei Zhuang. 2013. "Super-Resolution Fluorescence Imaging of Telomeres Reveals TRF2-Dependent T-Loop Formation." *Cell* 155(2): 345–56. http://www.ncbi.nlm.nih.gov/pubmed/24120135 (September 9, 2018).

Dong, X. et al. 2007. "Transcriptional Activity of Androgen Receptor Is Modulated by Two RNA Splicing Factors, PSF and P54nrb." *Molecular and Cellular Biology*.

Dong, Xuesen, Oksana Shylnova, John R.G. Challis, and Stephen J. Lye. 2005. "Identification and Characterization of the Protein-Associated Splicing Factor as a Negative Co-Regulator of the Progesterone Receptor." *Journal of Biological Chemistry*.

Drané, Pascal et al. 2010. "The Death-Associated Protein DAXX Is a Novel Histone Chaperone Involved in the Replication-Independent Deposition of H3.3." *Genes and Development*.

Du, X. et al. 2004. "Telomere Shortening Exposes Functions for the Mouse Werner and Bloom Syndrome Genes." *Molecular and Cellular Biology* 24(19): 8437–46. http://mcb.asm.org/cgi/doi/10.1128/MCB.24.19.8437-8446.2004.

Dye, B. T., and J. G. Patton. 2001. "An RNA Recognition Motif (RRM) Is Required for the Localization of PTB-Associated Splicing Factor (PSF) to Subnuclear Speckles." *Experimental Cell Research*.

Egan, Emily D., and Kathleen Collins. 2012. "Biogenesis of Telomerase Ribonucleoproteins." *Rna* 18(10): 1747–59.

Feuerhahn, Sascha et al. 2010. "TERRA Biogenesis, Turnover and Implications for Function." *FEBS Letters* 584(17): 3812–18. http://dx.doi.org/10.1016/j.febslet.2010.07.032.

Fox, Archa H. et al. 2002. "Paraspeckles: A Novel Nuclear Domain." *Current Biology*.

Fox, Archa H., and Angus I. Lamond. 2010. "Paraspeckles." *Cold Spring Harbor perspectives in biology*.

Frank, Amanda K. et al. 2015. "The Shelterin TIN2 Subunit Mediates Recruitment of Telomerase to Telomeres." *PLoS Genetics* 11(7): 1–19.

Frescas, David, and Titia De Lange. 2014. "Binding of TPP1 Protein to TIN2 Protein Is Required for POT1a,b Protein-Mediated Telomere Protection." *Journal of Biological Chemistry* 289(35): 24180–87. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4148849/pdf/zbc24180.pdf (September 9, 2018).

Fu, Dragony, and Kathleen Collins. 2007. "Purification of Human Telomerase Complexes Identifies Factors Involved in Telomerase Biogenesis and Telomere Length Regulation." *Molecular Cell*.

Fu, Xiang Dong, and Manuel Ares. 2014. "Context-Dependent Control of Alternative Splicing by RNA-Binding Proteins." *Nature Reviews Genetics*.

Ginno, Paul A. et al. 2012. "R-Loop Formation Is a Distinctive Characteristic of Unmethylated Human CpG Island Promoters." *Molecular Cell*.

Gómez-González, Belén et al. 2011. "Genome-Wide Function of THO/TREX in Active Genes Prevents R-Loop-Dependent Replication Obstacles." *EMBO Journal* 30(15): 3106–19.

Gottschling, Daniel E., Oscar M. Aparicio, Barbara L. Billington, and Virginia A. Zakian. 1990. "Position Effect at S. Cerevisiae Telomeres: Reversible Repression of Pol II Transcription." *Cell* 63(4): 751–62.

Greider, Carol W, and Elizabeth H Blackburn. 1985. 43 Cell *Identification of a Specific Telomere Terminal Transferase Activity in Tetrahymena Extracts*. https://www.cell.com/cell/pdf/0092-8674(85)90170-9.pdf?_returnURL=https%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2F0092867485901709%3Fshowall%3Dtrue (September 9, 2018).

Griffith, Jack D. et al. 1999. "Mammalian Telomeres End in a Large Duplex Loop." *Cell* 97(4): 503–14.

El Hage, Aziz, Sarah L. French, Ann L. Beyer, and David Tollervey. 2010. "Loss of Topoisomerase I Leads to R-Loop-Mediated Transcriptional Blocks during Ribosomal RNA Synthesis." *Genes and Development*.

Harrington, Lea, and Fabio Pucci. 2018. "*In Medio Stat Virtus* : Unanticipated Consequences of Telomere Dysequilibrium." *Philosophical Transactions of the Royal Society B: Biological Sciences* 373(1741): 20160444. http://www.ncbi.nlm.nih.gov/pubmed/29335368 (August 30, 2018).

Hayflick, L., and P.S. Moorhead. 1961. "The Serial Cultivation of Human Diploid Cell Strains." *Experimental Cell Research* 25(3): 585–621. http://linkinghub.elsevier.com/retrieve/pii/0014482761901926 (September 9, 2018).

He, Quanyuan et al. 2015. "The Daxx/Atrx Complex Protects Tandem Repetitive Elements during DNA Hypomethylation by Promoting H3K9 Trimethylation." *Cell Stem Cell*.

Hemann, Michael T, and Carol W Greider. 1999. 27 Nucleic Acids Research *G-Strand Overhangs on Telomeres in Telomerase-Deficient Mouse Cells*. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC148662/pdf/273964.pdf (September 9, 2018).

Henson, Jeremy D. et al. 2009. "DNA C-Circles Are Specific and Quantifiable Markers of Alternative-Lengthening-of-Telomeres Activity." *Nature Biotechnology*.

Hockemeyer, Dirk et al. 2007. "Telomere Protection by Mammalian Pot1 Requires Interaction with Tpp1." *Nature Structural and Molecular Biology*.

Hraiky, Chadi, Marc André Raymond, and Marc Drolet. 2000. "RNase H Overproduction Corrects a Defect at the Level of Transcription Elongation during RRNA Synthesis in the Absence of DNA Topoisomerase I in Escherichia Coli." *Journal of Biological Chemistry*.

Hsu, Lea Chia Ling et al. 2008. "DAZAP1, an HnRNP Protein, Is Required for Normal Growth and Spermatogenesis in Mice." *RNA*.

Huertas, Pablo, and Andrés Aguilera. 2003. "Cotranscriptionally Formed DNA:RNA Hybrids Mediate Transcription Elongation Impairment and Transcription-Associated Recombination." *Molecular Cell*.

Hug, Nele, and Joachim Lingner. 2006. "Telomere Length Homeostasis." *Chromosoma* 115(6): 413–25.

Hurley, Jennifer M., Jennifer J. Loros, and Jay C. Dunlap. 2016. "Circadian Oscillators: Around the Transcription–Translation Feedback Loop and on to Output." *Trends in Biochemical Sciences*.

Ivanauskiene, Kristina et al. 2014. "The PML-Associated Protein DEK Regulates the Balance of H3.3 Loading on Chromatin and Is Important for Telomere Integrity." *Genome Research*.

Jaafar, Lahcen, Zhentian Li, Shuyi Li, and William S. Dynan. 2017. "SFPQ•NONO and XLF Function Separately and Together to Promote DNA Double-Strand Break Repair via Canonical Nonhomologous End Joining." *Nucleic Acids Research*.

Kim, Kee K., Yong C. Kim, Robert S. Adelstein, and Sachiyo Kawamoto. 2011. "Fox-3 and PSF Interact to Activate Neural Cell-Specific Alternative Splicing." *Nucleic Acids Research*.

Kim, Nam W. et al. 1994. "Specific Association of Human Telomerase Activity with Immortal Cells and Cancer." *Science*.

Knott, Gavin J., Charles S. Bond, and Archa H. Fox. 2016. "The DBHS Proteins SFPQ, NONO and PSPC1: A Multipurpose Molecular Scaffold." *Nucleic Acids Research*.

Kovac, Michal et al. 2015. "Exome Sequencing of Osteosarcoma Reveals Mutation Signatures Reminiscent of BRCA Deficiency." *Nature Communications*.

De Lange, Titia et al. 1990. MOLECULAR AND CELLULAR BIOLOGY *Structure and Variability of Human Chromosome Ends*. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC360828/pdf/molcellb00038-0094.pdf (September 9, 2018).

———. 2005. “Shelterin: The Protein Complex That Shapes and Safeguards Human Telomeres.” *Genes and Development* 19(18): 2100–2110.

———. 2009. “How Telomeres Solve the End-Protection Problem.” *Science*.

Law, Martin J. et al. 2010. “ATR-X Syndrome Protein Targets Tandem Repeats and Influences Allele-Specific Expression in a Size-Dependent Manner.” *Cell*.

Lazzerini-Denchi, Eros, and Agnel Sfeir. 2016. “Stop Pulling My Strings-What Telomeres Taught Us about the DNA Damage Response.” *Nature Reviews Molecular Cell Biology*.

Levis, Robert, Tulle Hazelrigg, and Gerald M. Rubin. 1985. “Effects of Genomic Position on the Expression of Transduced Copies of the White Gene of Drosophila.” *Science* 229(4713): 558–61.

Levy, Michael Z. et al. 1992. “Telomere End-Replication Problem and Cell Aging.” *Journal of Molecular Biology*.

Lewis, P. W. et al. 2010. “Daxx Is an H3.3-Specific Histone Chaperone and Cooperates with ATRX in Replication-Independent Chromatin Assembly at Telomeres.” *Proceedings of the National Academy of Sciences*.

Lillard-Wetherell, Kate et al. 2004. “Association and Regulation of the BLM Helicase by the Telomere Proteins TRF1 and TRF2.” *Human Molecular Genetics* 13(17): 1919–32.

López-Otín, Carlos et al. 2013. “The Hallmarks of Aging.” *Cell*.

Lukong, Kiven E., Marc Étienne Huot, and Stéphane Richard. 2009. “BRK Phosphorylates PSF Promoting Its Cytoplasmic Localization and Cell Cycle Arrest.” *Cellular Signalling*.

Makarov, Vladimir L., Serguei Lejnine, Jirair Bedoyan, and John P. Langmore. 1993. “Nucleosomal Organization of Telomere-Specific Chromatin in Rat.” *Cell* 73(4): 775–87.

Maris, Christophe, Cyril Dominguez, and Frédéric H.T. Allain. 2005. “The RNA Recognition Motif, a Plastic RNA-Binding Platform to Regulate Post-Transcriptional Gene Expression.” *FEBS Journal*.

Martínez, Paula et al. 2009. “Increased Telomere Fragility and Fusions Resulting from TRF1 Deficiency Lead to Degenerative Pathologies and Increased Cancer in Mice.” *Genes and Development*.

Mathur, Mukul, Philip W Tucker, and Herbert H Samuels. 2001. “PSF Is a Novel Corepressor That Mediates Its Effect through Sin3A and the DNA Binding Domain of Nuclear Hormone Receptors.” *Molecular and cellular biology*.

Montes, Marta, Soraya Becerra, Miguel Sánchez-Álvarez, and Carlos Suñé. 2012. “Functional Coupling of Transcription and Splicing.” *Gene*.

Morin, Gregg B. 1989. “The Human Telomere Terminal Transferase Enzyme Is a Ribonucleoprotein That Synthesizes TTAGGG Repeats.” *Cell*.

Muntoni, Alessandra, Axel A. Neumann, Mark Hills, and Roger R. Reddel. 2009. “Telomere Elongation Involves Intra-Molecular DNA Replication in Cells Utilizing Alternative Lengthening of Telomeres.” *Human Molecular Genetics* 18(6): 1017–27.

Muntoni, Alessandra, and Roger R. Reddel. 2005. “The First Molecular Details of ALT in Human Tumor Cells.” *Human Molecular Genetics*.

Nabetani, A., and F. Ishikawa. 2009. “Unusual Telomeric DNAs in Human Telomerase-Negative Immortalized Cells.” *Molecular and Cellular Biology* 29(3): 703–13. http://mcb.asm.org/cgi/doi/10.1128/MCB.00603-08.

Nabetani, Akira, and Fuyuki Ishikawa. 2011. “Alternative Lengthening of Telomeres Pathway: Recombination-Mediated Telomere Maintenance Mechanism in Human Cells.” *Journal of Biochemistry*.

Nadel, Julie et al. 2015. “RNA:DNA Hybrids in the Human Genome Have Distinctive Nucleotide Characteristics,

Chromatin Composition, and Transcriptional Relationships." *Epigenetics and Chromatin*.

Nakama, Mina et al. 2012. "DNA-RNA Hybrid Formation Mediates RNAi-Directed Heterochromatin Formation." *Genes to Cells*.

Nakaoka, Hidenori, Atsuya Nishiyama, Motoki Saito, and Fuyuki Ishikawa. 2012. "Xenopus Laevis Ctc1-Stn1-Ten1 (XCST) Protein Complex Is Involved in Priming DNA Synthesis on Single-Stranded DNA Template in Xenopus Egg Extract." *Journal of Biological Chemistry*.

Nandakumar, Jayakrishnan, and Thomas R. Cech. 2013. "Finding the End: Recruitment of Telomerase to Telomeres." *Nature Reviews Molecular Cell Biology* 14(2): 69–82.

Neumann, Axel A. et al. 2013. "Alternative Lengthening of Telomeres in Normal Mammalian Somatic Cells." *Genes and Development* 27(1): 18–23.

Nittis, Thalia et al. 2010. "Revealing Novel Telomere Proteins Using *in Vivo* Cross-Linking, Tandem Affinity Purification, and Label-Free Quantitative LC-FTICR-MS." *Molecular & Cellular Proteomics* 9(6): 1144–56. http://www.mcponline.org/lookup/doi/10.1074/mcp.M900490-MCP200.

Ohishi, T., Y. Muramatsu, H. Yoshida, and H. Seimiya. 2014. "TRF1 Ensures the Centromeric Function of Aurora-B and Proper Chromosome Segregation." *Molecular and Cellular Biology* 34(13): 2464–78. http://mcb.asm.org/cgi/doi/10.1128/MCB.00161-14.

Olovnikov, A. M. 1973. "A Theory of Marginotomy. The Incomplete Copying of Template Margin in Enzymic Synthesis of Polynucleotides and Biological Significance of the Phenomenon." *Journal of Theoretical Biology* 41(1): 181–90. http://linkinghub.elsevier.com/retrieve/pii/0022519373901987 (September 8, 2018).

Paeschke, Katrin et al. 2013. "Pif1 Family Helicases Suppress Genome Instability at G-Quadruplex Motifs." *Nature*.

Palm, Wilhelm, and Titia De Lange. 2008. "How Shelterin Protects Mammalian Telomeres." www.annualreviews.org (September 9, 2018).

Passon, D. M. et al. 2012. "Structure of the Heterodimer of Human NONO and Paraspeckle Protein Component 1 and Analysis of Its Role in Subnuclear Body Formation." *Proceedings of the National Academy of Sciences*.

Patton, James G. et al. 1993. "Cloning and Characterization of PSF, a Novel Pre-MRNA Splicing Factor." *Genes and Development*.

Patton, James G., Stephen A. Mayer, Paul Tempst, and Bernardo Nadal-Ginard. 1991. "Characterization and Molecular Cloning of Polypyrimidine Tract-Binding Protein: A Component of a Complex Necessary for Pre-MRNA Splicing." *Genes and Development*.

Paulsen, Renee D. et al. 2009. "A Genome-Wide SiRNA Screen Reveals Diverse Cellular Processes and Pathways That Mediate Genome Stability." *Molecular Cell*.

Pfeiffer, Verena, Jérôme Crittin, Larissa Grolimund, and Joachim Lingner. 2013. "The THO Complex Component Thp2 Counteracts Telomeric R-Loops and Telomere Shortening." *EMBO Journal*.

Pickett, Hilda A., and Roger R. Reddel. 2015. "Molecular Mechanisms of Activity and Derepression of Alternative Lengthening of Telomeres." *Nature Structural and Molecular Biology*.

Porro, A., S. Feuerhahn, P. Reichenbach, and J. Lingner. 2010. "Molecular Dissection of Telomeric Repeat-Containing RNA Biogenesis Unveils the Presence of Distinct and Multiple Regulatory Pathways." *Molecular and Cellular Biology* 30(20): 4808–17. http://mcb.asm.org/cgi/doi/10.1128/MCB.00460-10.

Ramamoorthy, Mahesh, and Susan Smith. 2015. "Loss of ATRX Suppresses Resolution of Telomere Cohesion to Control Recombination in ALT Cancer Cells." *Cancer Cell*.

Ray, Payal et al. 2011. "PSF Suppresses Tau Exon 10 Inclusion by Interacting with a Stem-Loop Structure

Downstream of Exon 10." In *Journal of Molecular Neuroscience*,.

Redon, Sophie, Patrick Reichenbach, and Joachim Lingner. 2010. "The Non-Coding RNA TERRA Is a Natural Ligand and Direct Inhibitor of Human Telomerase." *Nucleic Acids Research*.

Rice, Cory, and Emmanuel Skordalakes. 2016. "Structure and Function of the Telomeric CST Complex." *Computational and Structural Biotechnology Journal* 14: 161–67. http://dx.doi.org/10.1016/j.csbj.2016.04.002.

Rosonina, E. et al. 2005. "Role for PSF in Mediating Transcriptional Activator-Dependent Stimulation of Pre-MRNA Processing In Vivo." *Molecular and Cellular Biology*.

Rosonina, Emanuel et al. 2005. "Role for PSF in Mediating Transcriptional Activator-Dependent Stimulation of Pre-MRNA Processing in Vivo." *Molecular and cellular biology*.

Santos-Pereira, José M., and Andrés Aguilera. 2015. "R Loops: New Modulators of Genome Dynamics and Function." *Nature Reviews Genetics*.

Schoeftner, Stefan, and Maria A. Blasco. 2008. "Developmentally Regulated Transcription of Mammalian Telomeres by DNA-Dependent RNA Polymerase II." *Nature Cell Biology* 10(2): 228–36.

———. 2010. "Chromatin Regulation and Non-Coding RNAs at Mammalian Telomeres." *Seminars in Cell and Developmental Biology* 21(2): 186–93. http://dx.doi.org/10.1016/j.semcdb.2009.09.015.

Sewer, Marion B. et al. 2002. "Transcriptional Activation of Human CYP17 in H295R Adrenocortical Cells Depends on Complex Formation among P54nrb/NonO, Protein-Associated Splicing Factor, and SF-1, a Complex That Also Participates in Repression of Transcription." *Endocrinology*.

Sfeir, Agnel et al. 2009. "Mammalian Telomeres Resemble Fragile Sites and Require TRF1 for Efficient Replication." *Cell*.

Sharma, Shilpy et al. 2012. "REV1 and Polymerase ζ Facilitate Homologous Recombination Repair." *Nucleic Acids Research*.

Shav-Tal, Yaron, and Dov Zipori. 2002a. "PSF and P54(Nrb)/NonO--Multi-Functional Nuclear Proteins." *FEBS Lett*.

———. 2002b. "PSF and P54nrb/NonO - Multi-Functional Nuclear Proteins." *FEBS Letters*.

De Silanes, Isabel López et al. 2014. "Identification of TERRA Locus Unveils a Telomere Protection Role through Association to Nearly All Chromosomes." *Nature Communications* 5: 1–13.

De Silanes, Isabel López, Martina Stagno D'Alcontres, and Maria A. Blasco. 2010. "TERRA Transcripts Are Bound by a Complex Array of RNA-Binding Proteins." *Nature Communications*.

Skourti-Stathaki, Konstantina, Nicholas J. Proudfoot, and Natalia Gromak. 2011. "Human Senataxin Resolves RNA/DNA Hybrids Formed at Transcriptional Pause Sites to Promote Xrn2-Dependent Termination." *Molecular Cell*.

Smogorzewska, Agata, and Titia de Lange. 2004. "Regulation of Telomerase by Telomeric Proteins." *Annual Review of Biochemistry* 73(1): 177–208. http://www.annualreviews.org/doi/10.1146/annurev.biochem.73.071403.160049.

Sollier, Julie et al. 2014. "Transcription-Coupled Nucleotide Excision Repair Factors Promote R-Loop-Induced Genome Instability." *Molecular Cell*.

Sordet, Olivier et al. 2009. "Ataxia Telangiectasia Mutated Activation by Transcription- and Topoisomerase I-Induced DNA Double-Strand Breaks." *EMBO Reports*.

Stirling, Peter C. et al. 2012. "R-Loop-Mediated Genome Instability in MRNA Cleavage and Polyadenylation Mutants." *Genes and Development*.

T M Bryan, A Englezou J Gupta S Bacchetti R R Reddel. 1995. "Telomere Elongation in Immortal Human Cells

without Detectable Telomerase Activity." *The EMBO Journal*.

Tagami, Hideaki, Dominique Ray-Gallet, Geneviève Almouzni, and Yoshihiro Nakatani. 2004. "Histone H3.1 and H3.3 Complexes Mediate Nucleosome Assembly Pathways Dependent or Independent of DNA Synthesis." *Cell*.

Takahama, Kentaro et al. 2013. "Regulation of Telomere Length by G-Quadruplex Telomere DNA- and TERRA-Binding Protein TLS/FUS." *Chemistry and Biology* 20(3): 341–50. http://dx.doi.org/10.1016/j.chembiol.2013.02.013.

Takai, Hiroyuki, Agata Smogorzewska, and Titia De Lange. 2003. "DNA Damage Foci at Dysfunctional Telomeres." *Current Biology*.

Takai, Kaori K. et al. 2017. "Erratum: Telomere Protection by TPP1/POT1 Requires Tethering to TIN2 (Molecular Cell (2011) 44(4) (647–659)(S109727651100815X)(10.1016/j.Molcel.2011.08.043))." *Molecular Cell* 67(1): 162. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3222871/pdf/nihms334966.pdf (September 9, 2018).

Teasley, Daniel C. et al. 2015. "Flap Endonuclease 1 Limits Telomere Fragility on the Leading Strand." *Journal of Biological Chemistry*.

Thandapani, Palaniraja, Timothy R. O'Connor, Timothy L. Bailey, and Stéphane Richard. 2013. "Defining the RGG/RG Motif." *Molecular Cell* 50(5): 613–23.

Tommerup, H, a Dousmanis, and T de Lange. 1994. "Unusual Chromatin in Human Telomeres." *Molecular and cellular biology* 14(9): 5777–85. http://www.ncbi.nlm.nih.gov/pubmed/23687383.

Toriumi, Kazuya, Takuma Tsukahara, and Ryo Hanai. 2013. "R-Loop Formation in Trans at an AGGAG Repeat." *Journal of Nucleic Acids*.

Tous, Cristina, and Andrés Aguilera. 2007. "Impairment of Transcription Elongation by R-Loops in Vitro." *Biochemical and Biophysical Research Communications*.

Tsao, Der An, Cheng Wen Wu, and Young Sun Lin. 1998. "Molecular Cloning of Bovine Telomerase RNA." *Gene*.

Tuduri, Sandie et al. 2009. "Topoisomerase I Suppresses Genomic Instability by Preventing Interference between Replication and Transcription." *Nature Cell Biology*.

Venteicher, Andrew S. et al. 2008. "Identification of ATPases Pontin and Reptin as Telomerase Components Essential for Holoenzyme Assembly." *Cell* 132(6): 945–57.

Viera, Alberto et al. 2015. "CDK2 Regulates Nuclear Envelope Protein Dynamics and Telomere Attachment in Mouse Meiotic Prophase." *Journal of cell science* 128(1): 88–99. http://www.ncbi.nlm.nih.gov/pubmed/25380821.

De Vitis, Marco, Francesco Berardinelli, and Antonella Sgura. 2018. "Molecular Sciences Telomere Length Maintenance in Cancer: At the Crossroad between Telomerase and Alternative Lengthening of Telomeres (ALT)." www.mdpi.com/journal/ijms (August 30, 2018).

Voon, Hsiao P.J., and Lee H. Wong. 2015. "New Players in Heterochromatin Silencing: Histone Variant H3.3 and the ATRX/DAXX Chaperone." *Nucleic Acids Research*.

Wahba, Lamia, Jeremy D. Amon, Douglas Koshland, and Milena Vuica-Ross. 2011. "RNase H and Multiple RNA Biogenesis Factors Cooperate to Prevent RNA:DNA Hybrids from Generating Genome Instability." *Molecular Cell*.

Wahba, Lamia, Steven K. Gore, and Douglas Koshland. 2013. "The Homologous Recombination Machinery Modulates the Formation of RNA-DNA Hybrids and Associated Chromosome Instability." *eLife*.

Wang, Feng et al. 2007. "The POT1-TPP1 Telomere Complex Is a Telomerase Processivity Factor." *Nature*.

Wang, Richard C., Agata Smogorzewska, and Titia De Lange. 2004. "Homologous Recombination Generates

T-Loop-Sized Deletions at Human Telomeres." *Cell* 119(3): 355–68.

Wellinger, Raymund J. 2009. "The CST Complex and Telomere Maintenance: The Exception Becomes the Rule." *Molecular Cell*.

Westover, Kenneth D., David A. Bushnell, and Roger D. Kornberg. 2004. "Structural Basis of Transcription: Nucleotide Selection by Rotation in the RNA Polymerase II Active Center." *Cell*.

Wongsurawat, Thidathip, Piroon Jenjaroenpun, Chee Keong Kwoh, and Vladimir Kuznetsov. 2012. "Quantitative Model of R-Loop Forming Structures Reveals a Novel Level of RNA-DNA Interactome Complexity." *Nucleic Acids Research*.

Xu, Baoji; Clayton, David A. 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." *The EMBO Journal*.

Yadav, Sharda P. et al. 2014. "The Transcription-Splicing Protein NonO/P54nrband Three NonO-Interacting Proteins Bind to Distal Enhancer Region and Augment Rhodopsin Expression." *Human Molecular Genetics*.

Yang, Yanzhong et al. 2014. "Arginine Methylation Facilitates the Recruitment of TOP3B to Chromatin to Prevent R Loop Accumulation." *Molecular Cell*.

Yeager, Thomas R. et al. 1999. "Telomerase-Negative Immortalized Human Cells Contain a Novel Type of Promyelocytic Leukemia (PML) Body." *Cancer Research* 59(17): 4175–79.

Yu, Kefei et al. 2003. "R-Loops at Immunoglobulin Class Switch Regions in the Chromosomes of Stimulated B Cells." *Nature Immunology*.