Functions of SMUG1 and NEIL3 in telomere homeostasis

Pinelopi Kroustallaki

Thesis for the degree of Philosophiae Doctor (PhD)

Department of Clinical Molecular Biology, Akershus University hospital

Faculty of Medicine, University of Oslo

Norway

2019

AKERSHUS UNIVERSITETSSYKEHUS

Series of dissertations submitted to the
Faculty of Medicine, University of Oslo

ISBN 978-82-8377-629-4

Cover: Hanne Baadsgaard Utigard.
Print production: Reprosentralen, University of Oslo.

Educating the mind without educating the heart is no education at all.

Aristotle

Acknowledgments

The current PhD study was conducted at the Department of Clinical Molecular Biology (EpiGen) of Akershus University hospital and at the Research Institute of Internal Medicine of Oslo University Hospital, Rikshospitalet, from September 2015 to November 2019. Financial support was provided by University of Oslo and Akershus University hospital.

First and foremost I would like to express my sincere gratitude to my principal supervisor, Professor Hilde Nilsen, for the opportunity she gave me to continue working in such an exciting project and be a part of a group that was making me to evolve as a scientist, day by day. Thank you for all the positivity, patience and guidance you showed me over these years. You could always find this perfect line between inspiration and motivation that could only push me to continue doing my job better. You will always be an example to follow in my professional life and I am grateful for that.

I would like to thank my co-supervisor, Professor Magnar Bjøras, for all his time and help over the years. I would also like to thank all the co-authors and collaborators for their contribution and help in the current work.

I am grateful that I had the opportunity to work among the current and past members of the Nilsen research group as well as all the colleagues at EpiGen. Lisa L, thank you for always being there for me, both as a colleague, ready to reply to all my questions, and as a very good friend. Henok K, Panpan Y, Ying E and Sergio C thank you for making the days at the lab feel sometimes like a playground. Tanima S, Anna W, Nuriye BT and Lene A, thank you for all the fruitful scientific discussions and for creating an environment at the office that many times felt like home. Also, a very big thank you to Anna Frengen for her immense patience, time and help, making sure everything work perfectly at EpiGen.

I would like to express my gratitude to all the people in my life, outside academia that helped me through all the difficult moments of the past five years. To my friends, thank you for being by my side. Tassos, thank you for just being there for me, these years would not have been the same without you.

Finally, there are no words to express how grateful I am to Γιώργος and Κατερίνα, my parents. You have always been supporting me throughout my life, making sure I will not fall. Thank you, the last five years would have never happened without you by my side.

Oslo, December 2019

Penelope Kroustallaki

Table of Contents

Abbreviations

53BP1	p53-binding protein 1
5-caU	5-carboxyuracil
AAG	alkyladenine DNA-glycosylase
AID	activation-induced deaminase
ALT	alternative lengthening of telomeres
AP site	abasic site
APE1	AP-endonuclease 1
ATM	ataxia telangiectasia mutated
ATR	ataxia telangiectasia and Rad3 related
BER	base excision repair
CB	Cajal body
CBCA	cap-binding complex
CM	cardio myocytes
CSR	class switch recombination
DCP2	decapping mRNA 2 factor
DDR	DNA damage response
DGCR8	component DiGeorge critical region 8
DKC1	dyskerin pseudouridine synthase 1
DSBs	DNA double strand breaks
dsDNA	double-stranded DNA
FEN1	Flap endonuclease 1
G4	G-quadruplex structures
Gh	guanidinohydantoin
hmU	5-hydroxymethyl uracil
HR	homologous recombination
IFD	insertion finger domain
LigIII	DNA ligase III
LP	long-patch repair
MBD4	methyl-binding domain protein 4
MEFs	mouse mbryonic fibroblasts
MPG	methylpurine DNA-glycosylase
MRN	Mre11/Rad50/ Nbs1 complex
MUFA	monounsaturated fatty acid
MUTYH	MutY-homolog glycosylase
NAF1	nuclear assembly factor 1
NAFLD	non-alcoholic fatty liver disease
NEIL	Nei endonuclease VIII-like family
NER	nucleotide excision repair
NHEJ	non - homologous end joining
NTH1	Nth Endonuclease III-like 1
OOG1	8-oxoguanine-DNA glycosylase
PABPN1	nuclear poly(A)-binding protein
PAPD5	PAP-associated domain-containing protein’5
PARN	Poly(A)-specific ribonuclease
PD	population doublings
PML	promyelocytic leukaemia bodies
PNK	polynucleotide kinase
Pol β	DNA polymerase β
POT1	protection of telomeres 1
Rap1	Ras-related protein 1

RFC	replication factor-C
RNAPII	RNA polymerase II
RNP	ribonucleoprotein
RPA	human replication protein A
RTEL1	regulator of telomere elongation helicase 1
SHM	somatic hypermutation
SMUG1	single-strand-selective monofunctional uracil-DNA glycosylase 1
SNP	single nucleotide polymorphism
Sp	spiroiminodihydantoin
SP	short-patch repair
ssDNA	single-stranded DNA
TCAB1	telomere Cajal body protein 1
TDG	thymine-DNA-glycosylase
TERC	telomeric RNA template
TERT	telomerase reverse transcriptase
Tg	thymine glycol
TIF	telomere dysfunction-induced foci
TIN2	TRF1-interacting nuclear protein 2
TOE1	target of EGR1 protein 1
TPP1	tripeptidyl peptidase 1
TRAMP	Trf4/Air2/Mtr4p Polyadenylation complex
TRBD	telomerase RNA binding domain
TRF1	telomeric repeat factor 1
TRF2	telomeric repeat factor 2
UNG	uracil DNA-glycosylase
VSMCs	vascular smooth muscle cells,
WRN	Werner syndrome helicase
XRCC1	X-ray repair cross-complementing protein 1
XRN1	5´-3´ Exoribonuclease 1

List of papers

Paper I

Kroustallaki P, Lirussi L, Carracedo S, You P, Esbensen Y, Götz A, Jobert L, Alsøe L, Sætrom P, Gagos S, Nilsen H. SMUG1 promotes telomere maintenance through telomerase RNA end processing. Cell Rep. 2019 Aug 13;28 (7):1690-1702.

Paper II

Carracedo S, Kroustallaki P, Alsøe L, Segers F, Wang C, Bartosova Z, Bohov P, Tekin N, Esbensen Y, Kong XY, Chen L, Wennerstrøm A, Ceolotto D, Berge RK, Bruheim P, Wong G, Halvorsen B, and Nilsen H. Liver steatosis associated with telomere maintenance defects in Smug1-deficient mice. (Manuscript)

Paper III

Quiles-Jiménez A, Gregersen I, Segers FM, Skarpengland T, Kroustallaki P, Yang K; Kong XY, Lauritzen KH, Olsen MB, Karlsen TR, Nyman TA, Sagen EL, Suganthan R, Nygård S, Scheffler K, Van der Veer E, Øgaard JD, Fløisand Y, Jørgensen HF, Holven KB, Biessen EA, Nilsen H, Dahl TB, Holm S, Bennett MR, Aukrust P, Bjørås M and Halvorsen B. DNA glycosylase Neil3 controls vascular smooth muscle cell homeostasis during atherosclerosis development. (Manuscript)

Introduction

1. Telomeres-the end of chromosomes

The genome of eukaryotic organisms, have been evolved to organize into numerous linear chromosomes (Garavís et al., 2013). The switch from the ancestral circular prokaryotic to the eukaryotic genome structure as we know it, was challenged by two biological limitations. Firstly, the vulnerability of the linear chromosome's "blunt" ends and secondly, the "chromosome end-replication problem" (Garavís et al., 2013; Levy et al., 1992; Wynford-Thomas and Kipling, 1997). Activation of the DNA damage response (DDR) is immediate in the presence of exposed DNA such as the chromosome termini (Maser and DePinho, 2004), which are also endangered by DNA exonucleases (Egan and Collins, 2012). The second biological constraint, the "chromosome end-replication problem" arises from the inability of the semiconservative DNA replication machinery to fully replicate the very end of the chromosome in each round of replication (Wynford-Thomas and Kipling, 1997). The polymerase responsible for the DNA replication is incapable of copying the lagging DNA strand to completion, thus leading to trimming of the single stranded terminal part from the 5′ end of the linear chromosome (Levy et al., 1992). The above limitations led to evolution of the chromosomal termini into specialized and protected nucleoprotein structures, the so-called telomeres (Muller, 1938).

1.1 Structure of telomeric DNA

The primary structure of telomeric DNA is comprised by a highly conserved short sequence of hundreds of thousands complementary nucleotide tandem repeats (Greider, 1999). The nucleotide structure of mammalian telomeres is composed of long stretches of double-stranded TTAGGG/AATCCC (G-rich and C-rich strands) hexa-nucleotide repeats that vary in size between different organisms and species (Figure 1) (Morin, 1989). This sequence is conserved from protozoa to humans and rodents (Gomes et al., 2011; Moyzis et al., 1988; Zielke and Bodnar, 2010). The double stranded telomere terminates in a relatively short, 5'-TTAGGG-3' rich, single-stranded, overhang that extends out of the 3′-end of each linear DNA molecule (Makarov et al.,

1997). The length of the 3'-single-stranded overhang may vary from 30 to 600 nucleotides and the average length of the double-stranded mammalian telomeric repeats differ among species and even chromosomes of the same cell (Chai et al., 2005; Makarov et al., 1997; Moyzis et al., 1988). Human telomeres range between 5–20 kb whereas in several rodents telomeres extend up to 50–150 kb (Bischoff et al., 2005; Calado and Dumitriu, 2013; Moyzis et al., 1988). These unique chromosomal ends, are replicated by the semiconservative machinery and further elongated by a specialized enzyme called telomerase (Greider and Blackburn, 1989), a subject that will be further discussed later on.

The secondary structure of telomeres has been studied by electron microscopy imaging, which has revealed the organization of telomeric DNA in a lasso-like structure known as the T-loop (Griffith et al., 1999). The formation of the T-loop structure is carried out by the invasion and binding of the 3'-single-stranded G-overhang into the double stranded telomeric repeats, forming a displacement loop (D-loop) (de Lange, 2005). The size of the T-loops can reach several kb, whereas the D-loop size, is limited by the G-overhang length (Figure 1) (Greider, 1999; Griffith et al., 1999).

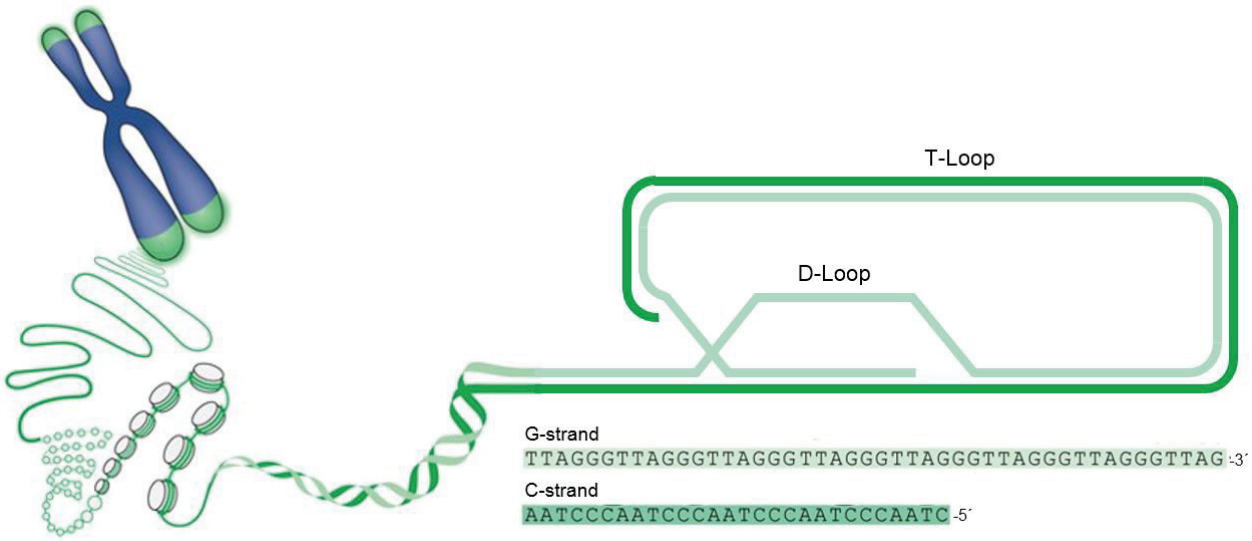


Figure 1 Secondary structure of telomeric DNA. *The ends of telomeres lead to a unique D-loop formation that protects the 3' single stranded overhang. The G-rich telomeric strand bears the TTAGGG sequence whereas the C-rich strand the complementary sequence AATCCC. Adapted from Maciejowski & de Lange, 2017.*

The mammalian T-loop structure is further capped and protected by a protein complex, known as shelterin or the telosome. The six-protein complex is comprised of three TTAGGG binding subunits (TRF1, TRF2, and POT1) and three interconnecting molecules (TIN2, TPP1, and Rap1) (Figure 2) (de Lange, 2005). TRF1 and TRF2 bind the double-stranded hexanucleotide repeats via their common Myb domain, whereas POT1 binds to the single-stranded telomeric sequence via its two Oligonucleotide/Oligosaccharide Binding (OB) fold domains (Court et al., 2005; Hwang et al., 2012). TIN2 and TPP1 are recruited by the TRF1 and TRF2 proteins, with TPP1 forming a bridge between the TRF proteins and the single-stranded binding protein, POT1. The complex is completed with the binding of Rap1 to TRF2 (Figure 2) (de Lange, 2018). Even though numerous modifications of the shelterin subunit have been reported, none seems to play an important role in the assembly of the six-subunit complex. Additionally, no DNA interaction is needed (Erdel et al., 2017). It has been reported that the components TIN2, Rap1, TRF1 and TRF2 are approximately ten times more abundant than POT1 and TPP1 (Takai et al., 2016). This suggest that the core shelterin complex might be comprised of these four abundant proteins with only some complexes containing the POT1 and TPP1 components. Cells produce enough copies of the shelterin complex to mask the full length of the telomeric sequence (de Lange, 2018). The shelterin complex is exclusively associated with the telomeric DNA throughout the cell cycle, acting as a scaffold for dynamic or transient recruitment of numerous telomere related factors. These non-shelterin proteins are essential for replication and protection of the chromosomal ends. An example of a protein group which function is not restricted to telomeres are factors of the DNA damage response (DDR) (de Lange, 2018).

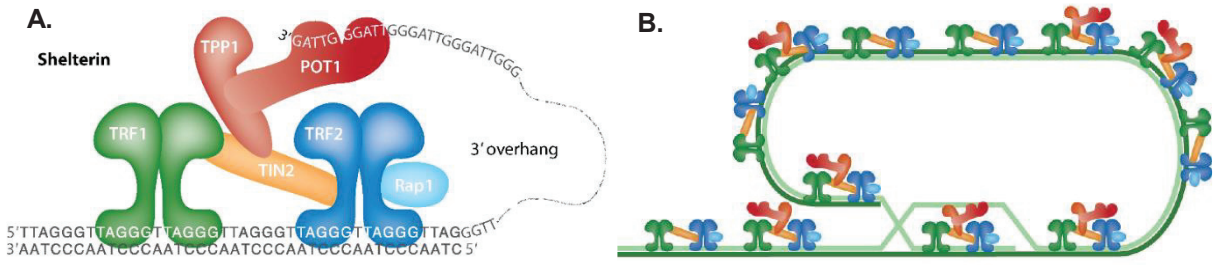


Figure 2. The shelterin complex. *The six-subunit human shelterin complex is comprised by the TRF1, TRF2, POT1, TIN2, TPP1 and RAP1. The TRF1 and TRF2 proteins bind the double stranded telomeric DNA and POT1 binds the single-stranded telomeric DNA. TPP1 interacts with TIN2 and POT1, whereas RAP1 binds TRF2. TIN2 acts as the central component, connecting TRF1, TRF2 and TPP1 (A). Shelterin binds throughout the telomeric sequence and protects the T-loop (B). Figure from de Lange., 2018, with permission from Annual review.*

1.2 Telomeres and the DNA damage machinery

At least seven DDR pathways can recognize telomeres as DNA Double Strand Breaks (DSBs) and threaten telomere integrity (Table1) (de Lange, 2018). Furthermore, the telomere's unique secondary structure makes them a target of enzymatic degradation. For that reason, telomeres evolved in a manner where they are being protected by numerous different molecular pathways, mainly targeting the initiation step of the DNA repair signaling cascade. The protein complex that carries out, directly or indirectly, the above task is shelterin. This happens either by using its own biochemical features described above or by attracting proteins that are involved in genome maintenance (de Lange, 2018). Shelterin deprived mammalian telomeres trigger ATM/ATR (ataxia telangiectasia mutated/ataxia telangiectasia and Rad3 related)-dependent DNA damage responses and are targeted either by canonical or alternative Non - Homologous End Joining (NHEJ) or Homologous Recombination (HR) (Sfeir and de Lange, 2012). Absence of particular shelterin components trigger, in turn, specific DDR pathways (Table 1). Accumulation of DDR factors, such as γ-H2AX or 53BP1, at telomeres give rise to nuclear structures called Telomere Dysfunction-Induced Foci (TIF) (Takai et al., 2003). These structures are created by critical telomere repeat shortening or by depletion of telomere protective factors (Kaul et al., 2011; Takai et al., 2003). When 53BP1 is absent, shelterin-free telomeres are immediately targeted by nucleolytic degradation pathways (Sfeir and de Lange, 2012).

DDR pathway	Shelterin subunit(s)	Defense mechanism
ATR kinase	POT1	RPA exclusion by POT1
ATM kinase	TRF2, TIN2	T-loop formation
c-NHEJ	TRF2	T-loop formation
alt-NHEJ	TRF2, TIN2, POT1	T-loop, PARP1 repression

Table 1 DDR pathways at telomeres. *A summary of the DDR pathways that potentially target telomeres and the respective shelterin subunit responsible for the defense mechanism. Adapted from, (de Lange, 2018).*

Shelterin is interacting with DDR proteins not only for DNA damage control but also for safeguarding the mitotic fidelity and to maintain telomere length homeostasis. Several factors of the DNA recombination repair machinery such as the MRN complex (Mre11/Rad50/Nbs1) interact with human shelterin to shape and maintain the T-loop (de Lange, 2005; de Lange and Petrini, 2000; Zhu et al., 2000). The DNA helicases

RTEL1 (regulator of telomere elongation helicase 1), and the Werner syndrome helicase (WRN) resolve T-loops to enable telomere replication or DNA repair at termini (Brosh, 2013; Opresko et al., 2004; Vannier et al., 2012). In mouse chromosomes, TRF2 is the only shelterin component required for biogenesis and/or maintenance of T-loops (Doksani et al., 2013), by directly interacting with ATM and blocking Ku70/Ku80 in order to suppress homologous telomeric crossovers (Karlseder et al., 2004; Ribes-Zamora et al., 2013).

Furthermore, the shelterin component, POT1 has been shown to bind and promote unfolding of telomeric G-quadruplex structures (G4) (Hwang et al., 2012; Ray et al., 2014; Wang et al., 2011). Most eukaryotic telomeres, despite the differences in their sequence, are rich in Guanines (Gomes et al., 2011). Apart from forming T-loops, this characteristic makes them efficient in creating secondary two or four-stranded parallel or anti-parallel DNA structures, known as G4. These G4 DNA structures are much more stable than double stranded DNA, making them structural barriers for DDR and, thereby, further contributing to telomere capping (Lipps and Rhodes, 2009).

1.3 Telomeres and BER

The telomeric sequence, as it has been discussed above, is susceptible to DNA damage and substrate for many different DDR pathways. Apart from replication errors during each cell division, telomeric damage can also emerge from modification of DNA bases or backbone by exogenous (e.g. ionizing radiation, mutagenic chemicals) or endogenous processes (oxidation, deamination, methylation)(Cooke et al., 2003; Iliakis et al., 2003; Poirier, 2012). DNA damage is one of the main causes of telomere dysfunction. Inappropriate recombination and aberrant chromosome end-to-end fusions give rise to genomic instability. Due to the important role that telomeres have in genome integrity, in addition to the protective role of the shelterin complex, they are closely safeguarded by molecular mechanisms that mend damaged DNA. Aside from the NHEJ and HR mechanisms, that repair double strand breaks, telomeres are subjected to repair of damaged or mismatched DNA bases. These are the main pathways: nucleotide excision repair (NER), base excision repair (BER), and DNA mismatch repair (MMR). BER is the focal point of this study.

1.3.1 Molecular mechanism of BER

BER is a highly conserved DNA repair pathway that removes and corrects small DNA damaged bases, produced by deamination, oxidation, alkylation, depurinations, as well as single-strand breaks. Five distinct steps make up the BER pathway: base recognition and removal, incision of the DNA backbone, end processing of the DNA termini, repair synthesis and ligation (Figure 3). Even though all five steps could be performed independently, the suggested biological model supports a cooperative and robust process termed as “passing the baton” (Prasad et al., 2010; Wilson and Kunkel, 2000). The negative outcomes of overexpression or upregulation of one BER protein alone, support a fine tuned, stepwise process, where the expression levels of BER factors and time of action is imperative and strictly controlled (Frosina, 2000; Leguisamo et al., 2017).The initial step in BER is carried out by a DNA glycosylase that recognize DNA lesions and which enzyme initiates the process, depends on the type of damage. Eleven of these enzymes have been identified in humans, with different substrate specificity (Table 2). They are categorized in two groups, monofunctional and bifunctional. The monofunctional DNA glycosylases include UNG (uracil DNA-glycosylase), SMUG1 (single-strand-selective monofunctional uracil DNA-glycosylase 1), TDG (thymine DNA-glycosylase), MPG or AAG (methylpurine DNA-glycosylase or alkyladenine DNA-glycosylase) and MBD4 (methyl-binding domain protein 4) (Table 2). These proteins remove damaged bases by cleaving the N-glycosyl bond between the base and the sugar, leaving an abasic site (AP site) in DNA. The AP site is then recognized by the AP-endonuclease 1 (APE1), which nicks the sugar 5´ of the abasic site, generating a strand break with 3′-hydroxyl (3′-OH) group and a 5′-deoxyribose-5-phosphate (5′-dRP) end (Figure 3) (reviewed in Wallace, 2014; Krokan and Bjøras., 2013).

The bifunctional group includes OOG1 (8-oxoguanine-DNA glycosylase), NTH1 (Nth Endonuclease III-Like 1), MUTYH (MutY-homolog glycosylase) and the NEIL 1, 2 and 3 of the NEIL (Nei Endonuclease VIII-Like) family, which all primarily recognize oxidized bases (Table 2). Bifunctional glycosylases, apart from excising the damaged base by cleaving the *N*-glycosyl bond, cleave the DNA backbone using their AP lyase activity. In the case of NTH1, OGG1 and MUTYH, a β elimination reaction is taking place generating a gap with α/β unsaturated aldehyde (3′-PUA) and 5′-phosphate (5′-

P) ends. APE1 is then recruited to generate 3′-OH termini using its 3′-phosphodiesterase activity. The NEIL family takes advantage of a β,δ-elimination reaction to remove the deoxyribose residue and generating a 3′-phosphate termini. In an APE1-independent BER sub-pathway, these termini are then removed by polynucleotide kinase (PNK) (Figure 3) (reviewed in Wallace, 2014; Krokan and Bjøras., 2013).

Enzyme	Mono-/Bifunct./ Activity	Substrate	Subcellular localization
UNG1/2	M	U, 5-FU	Mitochondria/ Nuclei
SMUG1	M	5-hmU, U, 5-FU, fU, caU	Nucleus
TDG	M	5-hmU, 5-FU, 5-CU	Nucleus
MBD4	M	CpG sites	Nucleus
MUTYH	M	8-oxoG	Mitochondria/ Nucleus
MPG (AAG)	M	3meA, 7meG, 3meG, Hx, 1A	Nucleus
OGG1	M/B (β)	8-oxoG, FapyG	Mitochondria/ Nucleus
NTHL1	B (β)	Tg, FapyG, DHU, 5-OHU, 5-OHC	Nucleus
NEIL1	B (β/δ)	Tg, FapyG, FapyA, 5-OHU, DHU, Sp, Gh	Mitochondria/ Nucleus
NEIL2	B (β/δ)	Similar to NEIL1	Mitochondria/ Nucleus
NEIL3	B (β)	FapyG, FapyA, Sp, Gh, Tg	Nucleus

Table 2. Glycosylases and their substrates. *Summary of the characteristics of the glycosylases acting on BER mechanism (Reviewed in Wllace, 2014; Krokan and Bjøras, 2013).*

Depending on several factors, BER continues further via two alternative sub-pathways, short-patch (SP) or long-patch (LP) repair. In SP-BER the removal of just one (or two) nucleotide(s) takes place, whereas in LP-BER from 2 to 13 nucleotides are excised. In SP repair, the DNA polymerase β (Pol β) is responsible for inserting a new nucleotide and with its lyase activity to remove the 5′-dRP, followed by XRCC1 and DNA ligase III (LigIII) responsible for sealing the generated nick (Figure 3). Even though the SP-pathway is considered the dominant BER mechanism, the switch to the LP-BER is believed to be decided upon several factors. Long-patch repair is primarily present in proliferative cells (Svilar et al., 2011). The cell type, availability of BER proteins and specificity of the initiating glycosylase are also determinant factors for the choice between short vs long patch repair (Fortini et al., 1999; Hanssen-Bauer et al., 2011; Narciso et al., 2007; Tichy et al., 2011). Although Polβ can be found in both pathways, some lesions such as reduced AP sites are known to be resistant to the Polβ lyase

activity. In this case, Polβ is not able to bind and remove the 5'-end and BER will be directed to the LP-pathway. Polδ/ε interacts with Polβ and together with replication factor-C (RFC) create a complex which inserts 2 to 13 nucleotides. In order for this to happen, a 5' flap structure of the displaced ''old'' DNA strand is created (Svilar et al., 2011). Flap endonuclease (FEN1) is then removing the 5′-terminal moiety and the pathway is terminated with the DNA ligase I sealing the nick (Krokan and Bjørås, 2013; Svilar et al., 2011)

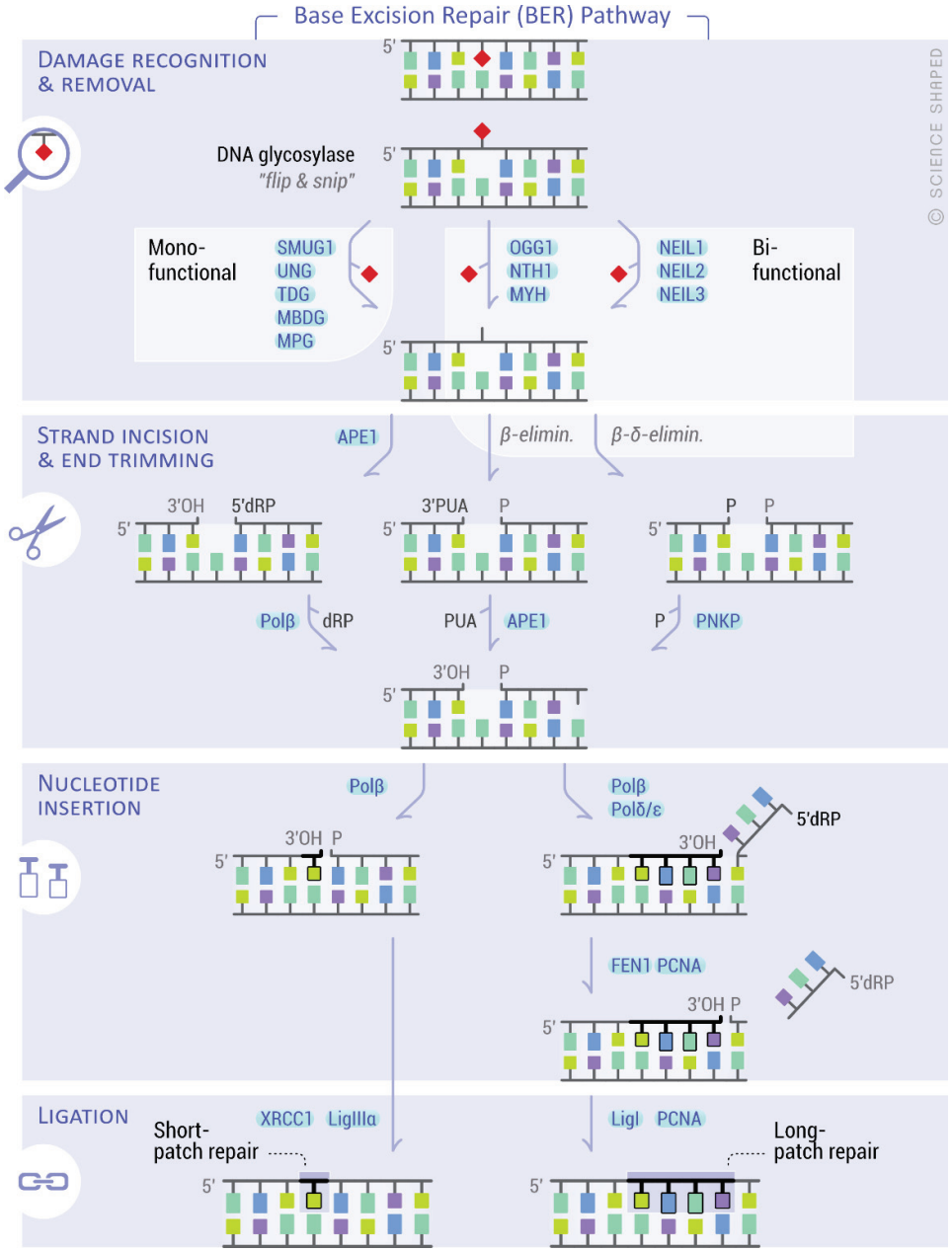


Figure 3. The Base excision repair mechanism. *Graphical representation of BER. Damage recognition and removal of the eleven DNA-glycosylases, followed by strand excision, nucleotide insertion and ligation together with the two distinct pathways of short -patch and long-patch repair. Figure by Ellen Tenstad at ScienceShaped™.*

1.3.2 BER glycosylases and Telomere homeostasis

1.3.2.1 OGG1

The classical substrate for the OGG1 glycosylase is 8-oxoG. This base damage, if not repaired, can result in a mismatch with adenine leading to a G:C to T:A transversion. Guanine is the DNA base with the lowest oxidation potential making it prone for oxidation. It has been shown, as well, that 5´ stretches of GG and GGG are more prone to oxidation than a single guanine (Hall et al., 1996; Saito et al., 1995). Considering that guanine is the most abundant DNA base in telomeric sequences it comes as no surprise that 8-oxoG lesion is the most prominent DNA damage in telomeres. It has been shown that OGG1 acts in telomeric DNA removing these guanine lesions, and thereby, actively participating in maintenance of telomere homeostasis (Wang et al., 2010). In the absence of OGG1 in MEFs, when cultivated in 20% O_2 or in the presence of oxidants, telomeres present increased levels of 8-oxoG lesions, resulting in telomere attrition. Under normoxic conditions, *Ogg1*$^{-/-}$ MEFs exhibit increased γH2Ax and XRCC1 TIF formation, showing respectively increased DSB and SSB formation at the telomeres. 53BP1 TIFs were also found at telomeres in *Ogg1*$^{-/-}$ MEFs, an additional marker of DDR activation at the telomeric region. Furthermore, Ogg1$^{-/-}$ mice presented telomere sister chromatid exchanges and a preferential loss of the lagging G-rich strand (Wang et al., 2010). Also, Fouquerel et al. have recently demonstrated that, in the absence of OGG1, human cells show chromosomal aberrations such as chromosome fusions and chromatin bridge formation. These phenotypes are only detectable upon chronic exposure to the oxidative agent $^1O_2$ (Fouquerel et al., 2019).

The exact mechanism behind the removal of telomeric 8-oxoG residues has not been elucidated and is unclear whether BER proteins team up with the shelterin complex. Opresko et al. have shown that binding efficiency of the shelterin factors TRF1 and TRF2 is significantly reduced by presence of the 8-oxoG. In particular, only a single 8-oxoG lesion per telomeric repeat decreased the amount of available telomeric sequence bound to TRF1 or TRF2 by almost 50%. Whereas, when all three telomeric G were substituted with 8-oxoG the amount of telomeric substrate bound to TRF1 and TRF2 was barely detectable (Opresko et al., 2005). However, the OGG1 incision activity was unaffected by these two shelterin components, indicating that BER

function in telomeres is not inhibited by shelterin (Rhee et al., 2011). Moreover, G-quadruplex formation that is supported to be important for replication and protection of telomeres, is disrupted by 8-oxoG lesions (Bochman et al., 2012). Hence, the removal of 8-oxoG is crucial for telomere stability.

1.3.2.2 NEIL glycosylases

Apart from the classical substrate of OGG1, 8-oxoG, several other oxidation products are found on the DNA sequence. When 8-oxoG is further oxidized, guanidinohydantoin (Gh) and spiroiminodihydantoin (Sp) are formed (Luo et al., 2000, 2001). These lesions can mispair with adenine and guanine and stall DNA polymerases (Duarte et al., 1999; Henderson et al., 2003). Furthermore, another major oxidation product results from thymine (T): thymine glycol (Tg) is a cytotoxic lesion that also obstructs replication (Ide et al., 1985). Some of the DNA glycosylases responsible for the removal of these adducts belong to the DNA endonuclease eight-like (NEIL) protein family. The members of this group are NEIL1, NEIL2 and NEIL3 and are characterized by broad substrate specificity. In particular, NEIL1 removes pyrimidine lesions such as Tg, mainly in double-stranded DNA (Bandaru et al., 2002; Dou et al., 2003). On the other hand, NEIL2 and NEIL3 show a preference on single stranded DNA, with the latter acting on Gh and Sp lesions (Dou et al., 2003; Krokeide et al., 2013; Liu et al., 2010). On a cellular level, NEIL1 is suggested to detect 5-hydroxyuracil (5-OHU) in ssDNA and stall the replisome in order to remove the lesion (Hegde et al., 2013; Rangaswamy et al., 2017), whereas NEIL3 function is not well understood. However, it is know that NEIL3 acts predominantly in highly proliferating cells, including pluripotent cells in brain and hematopoietic cells in mice, embryonic stem cells (Hildrestrand et al., 2009; Regnell et al., 2012) and cancer cells in human (Hildrestrand et al., 2009; Kauffmann et al., 2008).

As in the case of 8-oxoG, these guanine oxidation products can be also located in G-rich DNA sequences, such as G4 structures. Bioinformatic studies have shown that the G4 are present throughout the human genome, with a prevalence in gene promoter regions and telomeric sequences (Lipps and Rhodes, 2009; Todd et al., 2005). Furthermore, it is described that G4s play regulatory roles during lagging strand

replication, transcription, translation and telomeric DNA elongation (Lipps and Rhodes, 2009). *In vitro* studies have shown that oxidized DNA lesions can affect the conformation of the telomeric G4, depending on their position in the G4 structure (Zhou et al., 2015). Sp, Gh, Tg and 8-oxoG are such lesions and even though the two latter molecules, Tg or 8-oxoG, do not disrupt the formation of the intramolecular quadruplex structure, they reduce their thermostability (Vorlickova et al., 2012; Zhou et al., 2015).

As in double - and single - stranded DNA, the glycosylases known to remove the Tg, Sp and Gh lesions from the G4 structures in telomeres are NEIL1 and NEIL3. In particular, *in vitro* activity assays have shown that mNeil3 DNA-glycosylase, efficiently excises Tg lesions from G4 telomeric sequence. The same study also demonstrated that mNeil3 and NEIL1 remove Sp and Gh from telomeric G4s (Zhou et al., 2013). Based on the function and specificity of these enzymes, it is logical to assume that they play an important role in the removal of oxidized bases in the genomic DNA and its G4 structures. Zhou et al. in particular, conducted studies in human cell lines and in a mouse model in order to see the effect of the NEIL3 protein on telomere homeostasis (Zhou et al., 2017).

Considering that NEIL3 is acting on proliferating cells (Reis and Hermanson, 2012; Torisu et al., 2005), experiments on NEIL3-KD human cell lines demonstrated a role in mitosis: Even though cells entered mitosis, the duration of metaphase was extended and an increase in anaphase DNA bridges was observed (Zhou et al., 2017) Furthermore, they checked the telomeric status and observed, in the case of NEIL3 KD human cells, a 2-fold increase in telomere loss and sister chromatid fusions. Similarly, *Neil3*$^{-/-}$ MEFs compared to *Neil3*$^{+/+}$, present more telomere loss, duplications and fusions. Additional to *in vivo* studies in the mouse *Neil3*$^{-/-}$ model, experiments in patient derived primary human fibroblasts were conducted. Cells from an individual with a mutation, which abolishes the NEIL3 glycosylase activity, were assessed for telomeric aberrations. Telomere erosion of more than 2-fold was observed in the patient cells compared to WT controls, confirming that catalytic activity of NEIL3 is required at least in part, for telomere protection in humans. In addition, TIFs were observed in NEIL3 KD cells. Zhou et al postulate that the telomeric related aberrations drive the phenotypes of metaphase arrest and increase DNA bridges during anaphase. Ultimately, this leads to decreased cell proliferation in the absence of NEIL3. The same

study tried to elucidate further, how NEIL3 acts on telomeres. Using immunofluorescence and chromatin immunoprecipitation experiments, they showed in human cell lines that, during late S phase, Neil3 localization to telomeres is enhanced. This was happening in conditions of oxidative stress. Furthermore, they identified that the recruitment of NEIL3 to telomeres is dependent on the interaction between the shelterin component TRF1 and NEIL3 (Zhou et al., 2017). Taken together, NEIL3 has a crucial role to maintain telomeric integrity.

1.3.2.3 UNG

Uracil's presence in DNA is a frequent event (100 to 500 uracil residues per cell per day) that occurs either by dUMP incorporation opposite adenine (A) residues during DNA replication or by hydrolytic deamination of cytosine under physiological conditions (Brynolf et al., 1978; Lindahl, 1993; Shen et al., 1994). The deamination events if not corrected, will give rise to a stable point mutation. These mutations occur from a U:G base-pair, which result in a C to T transition, after replication. One of the DNA glycosylases mainly responsible for removing these uracils is UNG2. The C-rich strand of telomeric DNA provides a potential for enrichment of uracil. Vallabhaneni et al. showed that uracil is accumulated in *Ung*$^{-/-}$ mice. Moreover, under telomerase null conditions, the *Ung*$^{-/-}$ mice presented telomeric aberration such as fragility, telomere length increase and telomere recombination in hematopoietic cells (Vallabhaneni et al., 2015). The same study, based on *in vitro* experiments, proposed that these phenotypes could be mechanistically attributed to the lower binding affinity of POT1/TPP1 to telomeres when uracil is present (Vallabhaneni et al., 2015).

Additionally, UNG aids the modification of the immunoglobulin (Ig) genes in B cells by indirectly playing a role in the mechanisms of somatic hypermutation (SHM) and class switch recombination (CSR). In particular, both mechanisms are initiated by activation-induced deaminase (AID) which deamines cytosines into uracils, generating U:G pairs (Cortizas et al., 2016). These mismatches are then excised by UNG2 to initiate SHM and CSR. However, AID activity has as well off-target effects, and a recent study showed that telomeres are such an off-target of AID activity and the glycosylase reversing this effect is UNG2. More specifically, in splenic B cells of *Ung*$^{-/-}$ mice,

telomere loss in the C-rich telomeric strand was increased compared to those of *Ung*$^{+/+}$, whereas the double knockout of UNG and AID had phenotype mirroring the wild-type mice cells. The telomeric aberration caused by UNG deficiency are essential enough to cause AID dependent TIFs (Cortizas et al., 2016). UNG2 DNA-glycosylase, therefore, has a role in telomere maintenance either by acting as a canonical BER protein or as a repair mechanism counteracting AID off-target activity.

1.3.2.4 SMUG1

The SMUG1 glycosylase is, together with UNG2, the enzyme that removes uracil residues from DNA via the BER pathway. In mice, it acts as the main replacement of UNG and when both glycosylases are deprived, dramatic loss of uracil DNA glycosylase activity is observed (Alsoe et al., 2017; Nilsen et al., 2001). Furthermore, SMUG1 may participate in SHM and CSR but this function is more prominent in cells lacking UNG2 (Dingler et al., 2014; Kemmerich et al., 2012; Rada et al., 2002).

In addition to uracil, SMUG1 acts on substrates that are not shared with UNG. These are several pyrimidine oxidation products, such as 5-carboxyuracil (Darwanto et al., 2009), 5-formyluracil (Masaoka et al., 2003), the thymine oxidation product 5-hydroxymethyl uracil from DNA (hmU)(Boorstein et al., 2001; Kemmerich et al., 2012; Masaoka et al., 2003), as well as the deamination product xanthine on single stranded DNA (Mi et al., 2009). Mutation studies on the human SMUG1 protein revealed that the residues crucial for the hydrolysis of the N-glycosidic bond are the Asn85 and Gly87, His239 for binding of nucleic acids, whereas Met91 for the C5 substituent recognition. Additionally, Phe98 and Asn163 are important for the discrimination of pyrimidine rings (Matsubara et al., 2004). However, no connection between telomeric DNA and SMUG1 glycosylase has been reported, this subject will be further discussed in the current PhD dissertation.

2. Telomeric elongation mechanisms

The inability of the DNA replication machinery to fully polymerize the single stranded telomeres gave rise to two distinct telomere elongation mechanisms in mammals, alternative lengthening of telomeres (ALT) and telomerase-mediated maintenance (Lundblad and Blackburn, 1993). Telomerase is a ribonucleoprotein (RNP) enzyme that is evolutionary conserved among mammals. It extends telomeres by adding telomere repeats at the telomeric single-strand termini using a complementary RNA template. In humans, telomerase is inactive in most somatic cells with the exception of early embryogenesis, male germ cells, activated lymphocytes, and certain types of stem cell populations (Kim et al., 1994; Shay, 1997; Wright et al., 1996). In the absence of any lengthening mechanism, telomeres of most somatic human tissues become shorter with each division, leading to senescence. Carcinogenesis, on the other hand, is dependent on continuous cell proliferation and sustaining an active telomere lengthening mechanism is essential. Approximately 85-95% of human cancers achieve this by expression and activation of the telomerase enzyme (Shay and Bacchetti, 1997). The rest, 5-15%, maintain their telomere length by taking advantage of the ALT mechanism, which is facilitated by homologous recombination at the telomeric locus (Heaphy et al., 2011). ALT is characterized by increased homologous telomeric recombination and extreme variation of telomere length from very short, to as long as 50–60 kb. Other established ALT phenotypes are the presence of the so-called ALT associated, promyelocytic leukaemia (PML) bodies and extensive chromosomal instability (Cesare and Reddel, 2010; Neumann et al., 2013; Sakellariou et al., 2013). Co-existence of both ALT and telomerase expression has been reported in cancer cells, as well as in, embryonic and somatic stem cells (Bojovic et al., 2015; Liu et al., 2007; Wang et al., 2012; Zalzman et al., 2010). The current PhD dissertation will be focused on the telomerase related mechanism.

2.1 Telomerase structure and assembly

The active telomerase holoenzyme consists of several components and regulated by a variety of others. The main factors of the human ribonucleoprotein are the catalytic subunit human telomerase reverse transcriptase (hTERT), the RNA subunit which bears the supplementary telomeric RNA template (*hTERC*) and the Dyskerin complex (Figure 5). The assembly of the active telomerase and transfer at the telomeric loci is a multistep process which includes RNA processing of *hTERC*, proper enzymatic assembly of hTERT and the Dyskerin complex, trafficking of the holoenzyme and stimulation of activity at the telomeric substrate (Reviewed in MacNeil et al., 2016).

2.1.1 *hTERC* and H/ACA RNP complex assembly

The mature telomerase RNA component is a highly divergent transcript that varies extensively among species in structure, length and synthesis. In human, the mature *hTERC* component is a 451 nucleotide long, non-polyadenylated moiety produced by the RNA polymerase II (RNAPII). *hTERC* is comprised of the telomeric RNA template (CR1) and 5 separate conformational domains: a core pseudoknot domain (CR2/CR3), the CR4-CR5, an H/ACA box (CR6/CR8), and the CR7 domain (Figure 4) (Chen et al., 2000). Each domain provides a specific feature, essential for the viability and functionality of the telomerase enzyme. In particular, the core and CR4/CR5 domains facilitate the binding of the hTERT protein, whereas the CR7 domain and its CAB box, provides signals that help the trafficking of *hTERC (*Zhang *et al., 2011)*. The conserved H/ACA-box motif, which is located at the *hTERC* terminus, is similar to that of the H/ACA small nucleolar (sno-) and small Cajal body (CB) specific RNAs. Even though *hTERC* has a canonical H/ACA domain, it is not known to contain substrates of pseudouridylation or RNA post transcriptional modifications, and its role appears to be limited to the structure and stability of the telomerase holoenzyme (MacNeil et al., 2016). This H/ACA box forms a particular secondary structure consisting of two hairpins, joined at a single stranded H-box and ending in a single stranded 3’ ACA tail (Chen et al., 2000). This characteristic structure facilitates the interactions of several factors essential for the processing and assembly of the RNP complex (MacNeil et al., 2016).

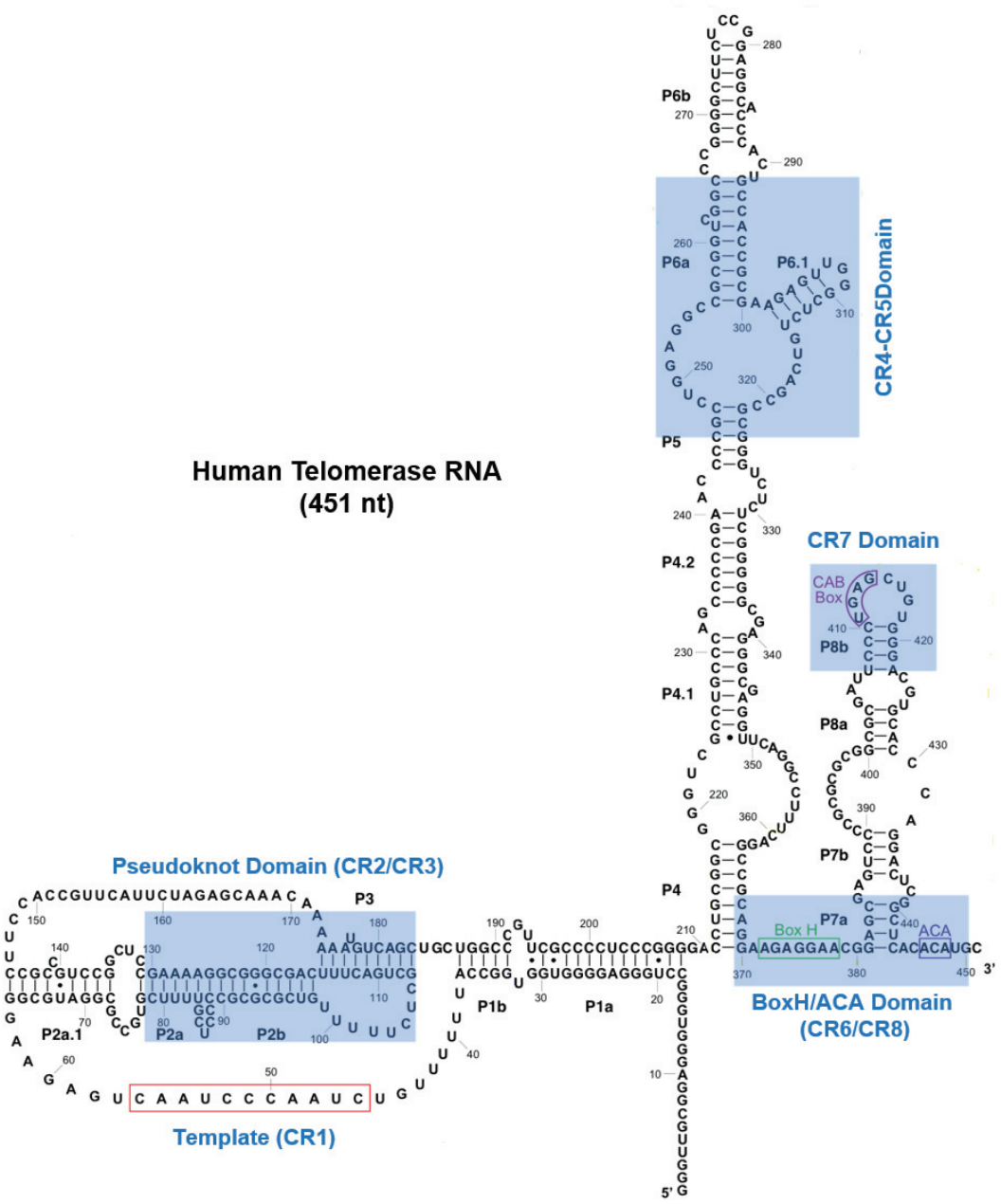


Figure 4. Secondary structure of human telomerase RNA. *Conformation of the, 451 nucleotide, telomerase RNA molecule and its main domains.*

Specifically, a pre-RNP complex, comprised of the proteins, dyskerin, NOP10, NAF1 and NHP2, is co-transcriptionally tethered to the H/ACA box before its replacement with the mature RNP complex (dyskerin, NOP10, GAR1 and NHP2) (Egan and Collins, 2012; Wang and Meier, 2004). The exact process is not well understood but a stepwise-regulated assembly of the H/ACA RNPs has been hypothesized. The HSP90-like chaperone, SHQ1, which acts as regulator of free dyskerin levels, appears to play an important role in this process. SHQ1, through transient interactions, protects dyskerin from degradation prior to H/ACA pre-RNP assembly (MacNeil et al., 2016).

Furthermore, this chaperone-like protein antagonizes RNA interactions at the PUA domain of dyskerin (Walbott et al., 2011). The proposed model supports the SHQ1-dyskerin transfer to the nucleoplasm where an RNA–dyskerin interaction displaces SHQ1, with the responsible RNA assumed to be *hTERC (Singh et al., 2015)*. The baton is then passed to the NAF1 protein, binding dyskerin and allowing the tethering of the NOP10 and NHP2 components. This is followed by the substitution of NAF1 with the GAR1 component via the formation of a NAF1-GAR1 heterodimer (Leulliot et al., 2007). The heterodimer weakens the binding affinity of NAF1 to dyskerin, allowing higher affinity binding with GAR1. This substitution is the distinguishing element between pre-RNP and the mature RNP complex (Leulliot et al., 2007; MacNeil et al., 2016).

2.1.2 hTERT scaffolding and telomerase recruitment to telomeres

The mature RNP complex is subsequently located at the subnuclear Cajal bodies (CBs) (Venteicher et al., 2009). These dynamic structures, which are found in the nucleus of proliferative cells, are comprised of several subunits, among them being the proteins coilin and WDR79 (Henriksson and Farnebo, 2015). The latter is acting as a chaperone by transferring the mature RNP from the nucleolus to the CBs, and it was later renamed Telomere Cajal body protein 1 (TCAB1) (Venteicher et al., 2009; Venteicher and Artandi, 2009). TCAB1 recognizes and binds the CAB-box of *hTERC* making the transfer to CBs feasible. It has been shown that *hTERC* is found at CBs in G1 and at the beginning of S-phase when it is still unassociated with the hTERT subunit (Tomlinson et al., 2010; Tomlinson et al., 2006). Furthermore, *hTERC* and hTERT have been found to co-localize only in S-phase, making CBs a strong candidate for the assembly site of hTERT to the mature RNP complex, where also telomerase becomes catalytically active (Tomlinson et al., 2006). Chen et al. showed that TCAB1, apart from telomerase trafficking, seems to be essential for the catalytic activity of the telomerase holoenzyme by promoting the formation of correct structural conformation of the CR4/CR5 domain. In particular, it has been shown recently that TCAB1 facilitates the proper folding of the P6.1 and P6b helices of the CR4/CR5 domain making them available to the telomerase RNA binding domain (TRBD) of TERT (Chen et al., 2018). Contrary to the previous proposed model, Lee et al. reported that the pre-RNP and not *hTERC*, associates with hTERT at the fibrillarin component of the nucleolus, prior to

the transfer of the mature RNP to the CBs (Lee et al., 2014). Thus, the exact timing and series of events is unclear but the CBs and its RNA chaperone TCAB1 seem to play an important role in the trafficking and assembly of the holoenzyme.

There are indications that TCAB1 and CBs have an additional role in the transfer of telomerase to its substrate sequence. CBs have been found to associate with telomeres, during S phase, in human cancer cells (Jady et al., 2006; Zhu et al., 2004) and silencing of TCAB1 reduces localization of *hTERC* to telomeres (Venteich (Venteicher et al., 2009). Although the exact mechanism whereby telomerase is guided to telomeres is not fully understood, the process of tethering of the holoenzyme to telomeres is better elucidated. It has been established that shelterin plays a crucial role in the recruitment of telomerase to telomeres. In particular, the OB fold domain of TPP1 (TEL) physically interacts with the TEN domain of the hTERT component, leading to telomerase recruitment (Schmidt et al., 2014; Zhong et al., 2012). Furthermore, the insertion finger domain (IFD) of hTERT is known to further facilitate the TPP1-hTERT association (Chu et al., 2016a; Chu et al., 2016b). Notably, TPP1 is not the only shelterin component that has a role in telomerase engagement to telomeres. TRF1 removal from telomeres during S phase is essential for telomerase to act on telomeric DNA. The above happens in an ATR/ATM dependent manner (Tong et al., 2015). It is proposed that, following ATM phosphorylation by ATR, TRF1 is phosphorylated by ATM causing proteasomal degradation of TRF1. This, in turn, leads to the release of the 3’ telomeric overhang, making it available for telomerase (McKerlie et al., 2012; Stiff et al., 2006; Tong et al., 2015). This model is believed to occur in both humans and mice (Lee et al., 2015).

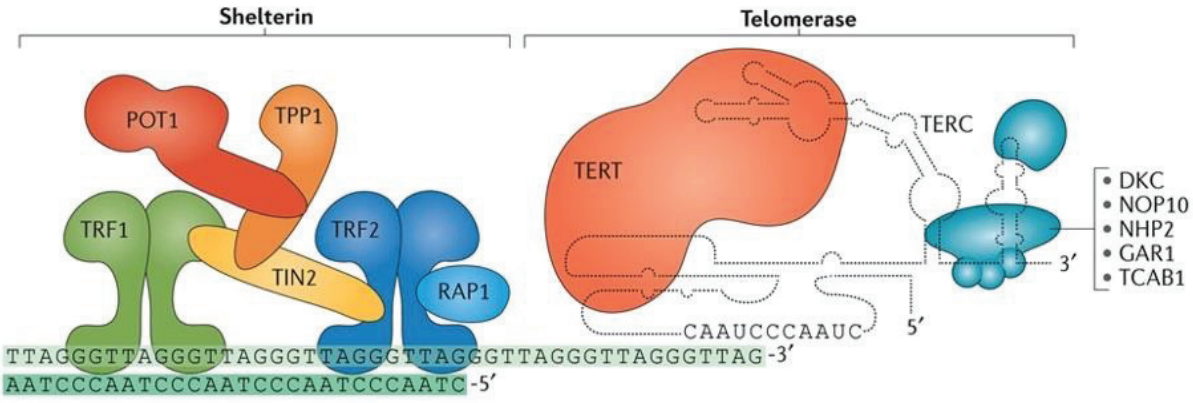


Figure 5. Telomerase action on telomeres. *Telomeres are elongated by the telomerase enzyme which is main subunits is the telomerase reverse transcriptase (TERT), telomerase RNA template component (TERC) and dyskerin complex (DKC1, NOP10, NHP2, GAR1). The TERT component catalyzes the reaction, using the complementary telomerase sequence of the TERC subunit. The dyskerin complex participates in the biogenesis and trafficking of telomerase. Figure from Maciejowski & de Lange, 2017, with permission from Springer Nature.*

2.2 *hTERC* maturation

The RNA component of telomerase, bearer of the complementary telomeric template, is a highly important molecule and its levels are tightly controlled. The mature 451nt *hTERC* moiety, positioned in the telomerase holoenzyme, is the final product of a series of events such as post-transcriptional modifications, trimming or degradation of precursor and extended RNA products in order to generate a precise 3' end. The exact mechanisms and pathways implicated in the regulation of *hTERC* are very little understood with many research groups trying to elucidate, the last years, the exact steps taken in the cells (Figure 6) (Reviewed in MacNeil et al., 2016).

It was only recently proposed that in order for transcriptional termination of *hTERC* to occur, DKC1 must bind to the H/ACA box (Roake et al., 2019). It is speculated that when this is not the case, RNAPII read-through gives rise to extended *hTERC* molecules (Nguyen et al., 2015). These extended *hTERC* molecules are then funneled to degradation, either in the cytoplasm or in the nucleus. In both cases, the extended molecules are bound to a cap-binding complex (CBCA) at the 5´terminus (MacNeil et al., 2016; Tseng et al., 2015). By export to the cytoplasm, it is proposed that the CBCA cap is removed by the decapping mRNA 2 (DCP2) factor, which is followed by degradation of *hTERC* intermediates by the 5´-3´ Exoribonuclease 1 (XRN1) (Shukla et al., 2016). However, extended *hTERC* species that are processed in the nucleus,

are targeted by the nuclear exosome targeting (NEXT) complex. The component DiGeorge critical region 8 (DGCR8) is then bound, causing the recruitment of the RNA exosome and degradation of the extended species (Figure 6) (Macias et al., 2015; Tseng et al., 2015).

In contrast to the above mechanism, shorter *hTERC* precursors that are efficiently assembled with the pre-RNP, transcription termination occurs and polyadenylation of the *hTERC* species follows (Roake et al., 2019). Several pathways seem to determine the levels and fate of the intermediate and mature products of *hTERC* and the RNA exosome is again implicated in the processing. It is not clear whether it only plays an active role in the degradation of excess species or also in the maturation and formation of the final *hTERC* product. It has been proposed that, CBCA may be recruited to the 5´- end of the *hTERC* precursor, attracting the human TRAMP complex, which is comprised of the ZCCHC7, the non-canonical poly(A) polymerase PAPD5 (TRF4-2) and the MTR4 subunit (MacNeil et al., 2016). Its *Saccharomyces cerevisiae* homolog is known to add a short stretch of oligo-adenosine to RNA 3' ends, supporting the hypothesis that this might be the fate of *hTERC* intermediates (Tseng et al., 2015). The accessibility of the 5´-end to the PARN deadenylation enzyme and the binding of the nuclear poly(A)-binding protein (PABPN1) to the polyA tail will determine if the precursor will lose its polyA tail and give rise to the mature *hTERC* or if it will be degraded by the RNA exosome (Nguyen et al., 2015; Tseng et al., 2015). Interestingly, while the hTRAMP polyA products are known to be 4-5 nt long, it was shown that almost half of the *hTERC* polyA tails associated with PABPN1 were found to be more than 15 nt long (Nguyen et al., 2015). It is believed that this longer stretches are synthesized by canonical polyadenylation polymerases (MacNeil et al., 2016). Nguyen et al. also showed that these long, extended *hTERC* species were increased upon depletion of either PABPN1 or PARN, while levels of mature *hTERC* were reduced. Similarly, depletion of the canonical poly(A) polymerases PAPα and PAPγ led to decreased mature *hTERC* (Nguyen et al., 2015). Furthermore, they reported that upon depletion of the PAPD5 component of the TRAMP complex, mature *hTERC* levels increased, which was attributed to the lack of exosome-mediated decay (MacNeil et al., 2016; Nguyen et al., 2015). Nguyen et al. concluded that canonical PAPs facilitate *hTERC* maturation whereas hTRAMP-dependent polyadenylation induce telomerase RNA decay (Nguyen et al., 2015). Opposite to that view, Tseng et al. and Roake et al.

showed that, the hTRAMP component PAPD5 is the enzyme that oligo-adenylates the shorter *hTERC* precursors and are, either, processed by PARN, giving rise to mature *hTERC*, or are degraded by the RNA exosome (Roake et al., 2019; Tseng et al., 2015). More specifically, it was proposed that the PAPD5-PARN relationship defines the maturation rate of the *hTERC* precursors. Whereas PAPD5 adenylates *hTERC* precursors, delaying their maturation, PARN deadenylates and processes them to complete mature *hTERC* species (Figure 5) (Roake et al., 2019). When PARN was absent, the mature *hTERC* molecules were still produced but at a much slower rate, which led to the view that PARN does not trim the extended products and is not essential for *hTERC* maturation, but as a facilitator of the process. The above observations led Roake et al. to propose a role for a yet non-identified 3´-5´ RNA exonuclease that performs the final trimming of *hTERC* (Roake et al., 2019). Notably, Deng et al. proposed a new function of the TOE1 deadenylase as participant of the maturation process of *hTERC*. They proposed that TOE1 deadenyates *hTERC* extended species at the CBs following their processing by PARN (Deng et al., 2019). Making TOE1 a strong candidate of the 3´-5´ RNA exonuclease proposed by Roake et al. The processing and maturation of the *hTERC* seems to be a highly complicated process, finely balanced between extension and degradation pathways, with many yet to be answered questions.

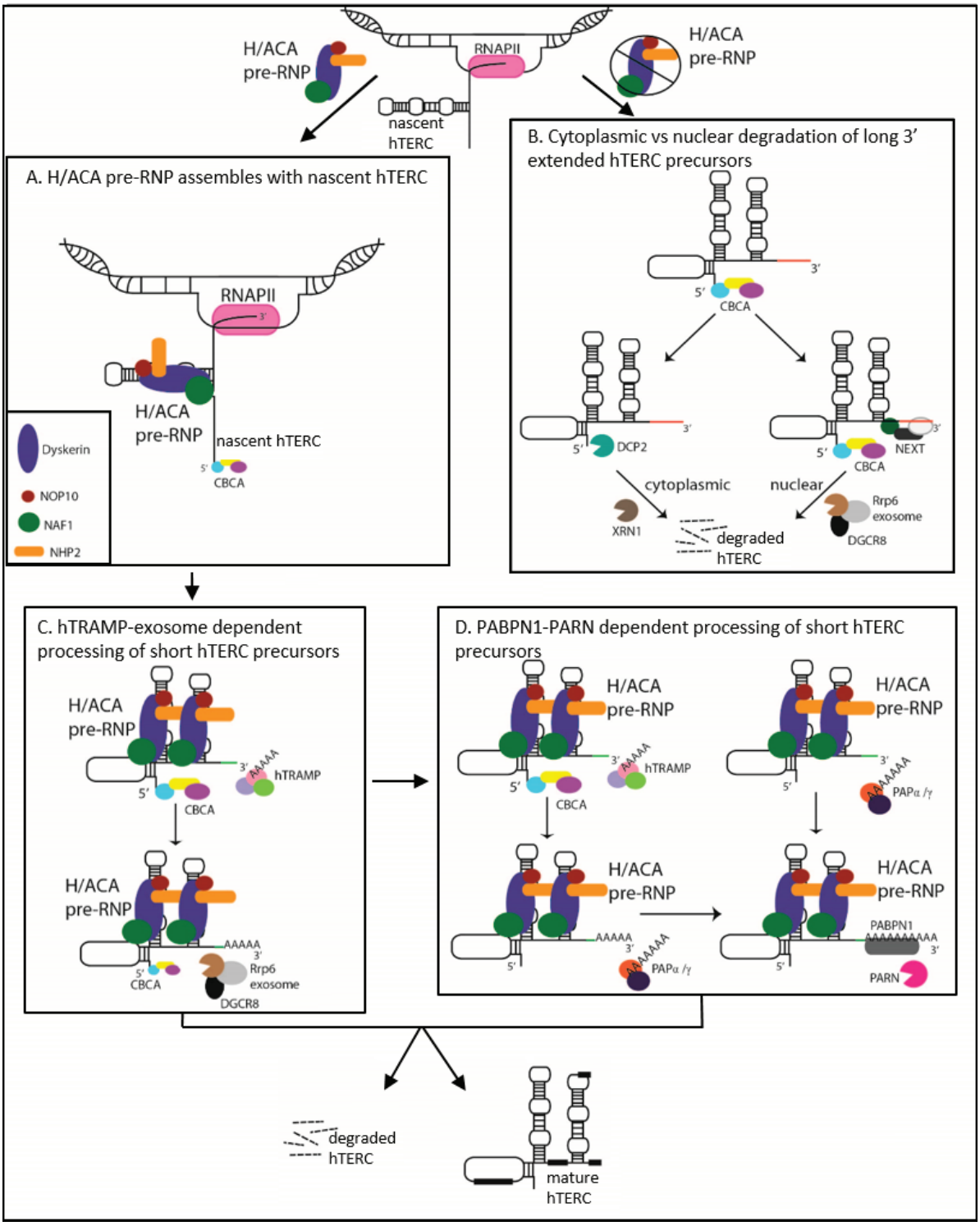


Figure 6. *hTERC* processing pathways. *Depiction of the several possible processing and maturation pathways of the mature hTERC molecule. Nascent hTERC that does not bind to the pre-RNP complex are extended due to RNA polymerase red through. They are then led to degradation either in a cytoplasmic or in a nuclear pathway (B). When nascent hTERC is assembled with the pre-RNP, transcriptional termination produces short nascent hTERC molecules which are then processed by the nuclear exosome either leading to degradation or to the production of mature hTERC. In this process, polyadenylation and deadenylation rates of the nascent molecules determines the fate of the final product (C, D). Adapted from MacNeil et al., 2016.*

3. Telomeres and ageing

The life-span of normal human cells, when grown *in culture,* is quite narrow with a defined number of population doublings (PD). This is referred to as the “Hayflick limit”. Most cells in culture, after approximately 40 to 60 PDs enter a static phase where they stop dividing. At this point, cells have reached the senescence state and undergo crisis leading them to an eventual death. The exact term of this process is replicative senescence and it was introduced three decades ago by Greider and Harley (Harley et al., 1990). They attributed this process to progressive telomere shortening. More specifically, they observed that the telomeres of diploid human cells get shorter with each cell division. Considering that most somatic cells do not bear active telomerase enzyme in order to replenish this loss, a gradual telomere erosion is unavoidable. Mechanistically, this process is attributed to the loss of shelterin components and the destabilization of the telomeric t-loop (Griffith et al., 1999). As it has been previously discussed, depletion of shelterin subunits, such as TRF2 and POT1, activates the DDR which triggers the activation of the transcription factor p53 (Beausejour et al., 2003; Takai et al., 2003; Wu et al., 2006). This in turn, results in the activation of the p21 kinase, leading to growth arrest and cellular senescence (Beausejour et al., 2003; Fischer et al., 2016). Consequently, progressive telomeric shortening is associated with cellular, tissue and organismal ageing (Armanios and Blackburn, 2012; Donate and Blasco, 2011).

Data supporting the above information come from peripheral blood cells derived from individuals with premature aging syndromes. Telomere length and proliferative capacity of these cells are significantly reduced (Ahmed and Tollefsbol, 2001; Ranganathan et al., 2001). Age-related telomere dysfunction phenotypes manifest in certain tissues which are maintained by lasting stem cell pools. The hematopoietic system and tissues, such as skin, germline, and other organs capable of regeneration, like liver, seem particularly affected in telomeropathies. Studies have shown that the telomere length of liver tissue from healthy individuals shortens with age (Aikata et al., 2000; Takubo et al., 2000). Additionally, liver phenotypes linked with aging seem to correlate with telomeric homeostasis. In telomerase null mice, progressive telomere shortening of liver tissue, caused by several generations of inbreeding, is linked to impaired liver regeneration and development of liver cirrhosis (Rudolph et al., 2000).

When senescence was induced in a mouse model where DNA repair was compromised, specifically in the liver (Alb-Xpg), the hepatocyte TIF levels were increased (Ogrodnik et al., 2017). The number of TIFs were also further elevated with increasing age of the mice. This phenotype coincided with an age-dependent fat deposition increase. Similar data were obtained from biopsies of patients with non-alcoholic fatty liver disease (NAFLD). Patients with high liver fat-content presented elevated TIF and p21 levels (Ogrodnik et al., 2017). The above, confirm the view that senescent cells are implicated in steatosis and, in turn, steatosis grade is linked with decreasing telomere length and increasing DNA damage (Aravinthan et al., 2013; Ogrodnik et al., 2017).

However, in tissues consisting of terminally differentiated, post-mitotic cells, such as neurons and heart, different rules may apply regarding tissue degeneration during aging (Anderson et al., 2019; Sapieha and Mallette, 2018). In the case of the heart tissue, even though the majority of the adult cardio myocytes (CM) are terminally differentiated, 1% of these cells have some limited proliferation capacity (Yutzey, 2017). Studies on both humans and mice have shown that CM present an increasing percentage of TIFs with advancing age (Anderson et al., 2019). Heart conditions are very tightly linked with the correct function and state of blood vessels and it is well known that cardiovascular diseases are prominent in the elderly human population. Experiments in human aorta atherosclerotic plaques showed that telomeres were shorter by almost 1kb compared to normal vessels of the same individuals (Matthews et al., 2006). Furthermore, vascular smooth muscle cells (VSMCs), from the fibrous cap of atheroma, had shorter telomeres than the ones of a healthy vascular media. These cells presented also senescence phenotypes, such as β-gal staining and increased p21 expression (Matthews et al., 2006). Another study on VSMCs of atherosclerotic plaques showed that TRF2 levels were reduced compared to VSMCs of normal aorta, while TIF levels were elevated (Wang et al., 2015). The above data support the view that telomere dysfunction is a prevalent characteristic of cardiovascular disease and aging. The correlation of telomere homeostasis and aging is indisputable, but to which extend telomere dysfunction contributes to senescence and drives the aging process is yet to be further elucidated.

4. Telomere homeostasis-Human vs Mouse

Even though they share the same telomere sequence, telomere homeostasis in humans and mice differ extensively. Mouse telomeres, like humans, protect and safeguard chromosomes and genome integrity. However, the phenotypical characteristics and the underlying causes of telomere maintenance and telomerase function defects diverge between the two species. One of the biggest differences between mouse and human telomeres is their length. Whereas the telomere length of healthy human individuals varies between 5 to 12 kb, laboratory mice can be from 5 to 10 times longer. The protein complex that bears the responsibility of defending telomeres, in mice, is again the shelterin complex. The five subunits, TRF1, TRF2, TPP1, TIN2 and Rap1 have orthologs in mice, with the exception of POT1. Due to a gene duplication event, POT1 protein has two orthologs in mice, POT1a and POT1b. Furthermore, the holoenzyme that elongates telomeres in mice is, like in humans, telomerase. Orthologs of all the main human telomerase components (*TERC*, TERT, Dyskerin complex) are found in mice. Even though the template sequence of *TERC* remains the same and the secondary structures are almost identical, the total *TERC* sequence differs in mice and humans. The processing and maturation of TERC, which has been so extensively investigated in yeast and humans the last years, is a quite unknown chapter of mice (Reviewed in Calado and Dumitriu, 2013).

A distinct difference in these two species is the correlation between age and telomere length. Even though mice have much longer telomeres, they have approximately a 30 times shorter lifespan. It has been reported that murine fibroblast do not present any significant telomere erosion before they stop dividing at around 10 to 15 population doubling (Blasco et al., 1997). Subsequently, mice do not seem to follow replicative aging as humans, rather than a culture growth arrest (Wright and Shay, 2000). In humans, telomere loss and decreased replicative capacity act as a carcinogenesis barrier, whereas mice are subjected to spontaneous transformation accompanied by polyploidy. The established model of telomerase deficient mice *mTerc*$^{-/-}$ does not show any significant telomere defects and only in later generations after repeated inter-breeding, telomeres become critically short (Blasco et al., 1997). Additionally, another group showed that tumorigenesis of first generation, telomerase deficient, mice was

closer to the wild type and it was increasing with each generation passing (Rudolph et al., 1999).

Several other murine models have been developed in order to study telomeropathies. However, the manifested phenotypes are rather different from the ones presented in humans. Mice bearing hypomorphic *Dkc1* mutation show telomere erosion only after 3 generations of inbreeding (Ruggero et al., 2003). On the other hand, human patients with hemizygous DKC1 mutations, display telomere shortening at an early age (Heiss et al., 1998). Furthermore, bone marrow aberrations in mice are mild, whereas human patients exhibit pancytopenia and dramatically decreased hematopoietic progenitor cells (Calado and Dumitriu, 2013). A phenotype, which is related to telomeropathies only in the dyskerin deficient mice and not in humans, is the impaired ribosomal function (Calado and Young, 2009; Ruggero et al., 2003). These mice show reduced levels of mature rRNA species and decreased rRNA pseudourydilation (Ruggero et al., 2003).

Another important parameter that should be taken into account is the environment that each species develops and age. Laboratory mice are observed and studied in a sterile and controlled environment with tightly regulated diet. On the contrary, humans are observed in the “wild”, exposed to environmental factors such as reactive oxygen species, ultraviolet light, inflammation and diet that may detrimentally affect telomeric homeostasis. Even though studies conducted in mice cannot be directly translated to humans, the similarities of the two models are strong enough to give a starting point for research related to telomeres.

Aims of the study

Extensive research in the field of telomeres has unveiled a role for many base excision repair proteins in maintaining the integrity of chromosome ends. Until recently their main known role was to protect telomeres from oxidative damage. More specifically, several DNA glycosylases of the BER pathway, such as UNG2 and NEIL3, act directly on telomeres and several mouse models where these enzymes are deleted, present telomere integrity phenotypes. Telomeres of mitotically active cells and tissues are dependent, mainly, on the activity of the telomerase enzyme. Until recently, there was no correlation of any BER enzyme with either the maturation and assembly of telomerase or its action at telomeres. The findings that DKC1, one of the major telomerase components, interacts with SMUG1 DNA-glycosylase gave a new angle to this protein's functions. We hypothesized that SMUG1 might be required for telomere maintenance via its ability to interact with DKC1.

Hence, the aims of this PhD dissertation were:

- Determine the function of the SMUG1 DNA-glycosylase on telomeres as a canonical BER protein.
- Define the role of SMUG1 in maturation and assembly of the telomerase enzyme.
- Determine whether the function of the NEIL3 DNA-glycosylase in telomeres promote the aging phenotype, in an atherosclerosis- prone background.

Summary of papers

Paper I: SMUG1 promotes telomere maintenance through telomerase RNA processing

Previous findings have established a role of SMUG1 in rRNA processing and a connection of the BER glycosylase with the pseudouridine synthase, DKC1. In this study, we further explored this phenotype in our mouse SMUG1-KO model and human HAP1 cells in which the SMUG1 expression was abrogated. The DKC1 distribution pattern in the *Smug1*$^{-/-}$ cells was found to deviate when compared to *Smug1*$^{+/+}$ cells. SMUG1 was also found colocalizing with coilin, a CBs protein. Hence, we show that SMUG1 is not only interacting with dyskerin but is also localized to a cellular compartment where DKC1 performs a function. Considering that DKC1 is one of the main components of the telomerase enzyme and that CBs are involved in telomerase assembly, we wondered if SMUG1 might have a role in telomere homeostasis. Indeed, *Smug1*$^{-/-}$ mice presented higher levels of fragile telomere in the C-rich strand, this phenotype was accompanied by reduced average telomere length and accumulation of telomere DNA damage in certain tissues. Telomerase status was unaffected in *Smug1*$^{-/-}$ mice, which possibly suggests that the above phenotypes are a consequence of the deprivation of the SMUG1 BER-activity in mice.

When assessing telomere homeostasis in human SMUG1-KO cell, we observed an almost 5-fold decrease of telomere length, as well as shelterin disorganization. Even though SMUG1 was found to bind in human telomeric sequence, the SMUG1 related telomeric damage found was not enough to justify the dramatic telomere erosion. When telomerase status was examined, we observed a dramatic downregulation of telomerase activity in SMUG1-KO cell, induced by limited levels of the *hTERC* RNA component. Furthermore, an increase of the extended intermediate species of the *hTERC* maturation pathway was registered upon SMUG1 depletion. These observations are in line with the role of SMUG1 in RNA metabolism. Analysis of the transcriptional *hTERC* gene body, revealed that SMUG1 associates with the actively transcribing RNAPII. When evaluating the levels of the possible substrates of SMUG1 in *hTERC*, we found that the region between CR4/CR5 domain and H box of *hTERC*, seemed to contain SMUG1 substrates. This region is also included in the DKC1 binding domain of *hTERC*. Considering the known interaction of SMUG1 with DKC1 we

examined whether SMUG1 absence affects *hTERC*/DKC1 binding and we found that in SMUG1-KO cells, the binding efficiency was downregulated by almost 50%. In humans, our data support a role of SMUG1 in the processing of the RNA telomerase component, by modifying the *hTERC* sequence, making it available for proper DKC1 binding. Whereas, in mice the BER activity of SMUG1 seems to have a more prominent role.

Paper II: Liver steatosis associated with telomere maintenance defects in SMUG1 deficient mice

Studies in human patients have shown that SMUG1 glycosylase associates with fat metabolism and non-alcoholic fatty liver disease (NAFLD). Prompted by these findings, we wanted to investigate the mechanism laying behind this observation, using our *Smug1*$^{-/-}$ mouse model. We observed increased body weight and liver fat accumulation in adult *Smug1*$^{-/-}$ mice and a lipid profile analysis uncovered higher levels of FFAs and TGs in one-year old *Smug1*$^{-/-}$ mice compared to *Smug1*$^{+/+}$. Lack of fibrosis and a significantly increased anti-inflammatory index of *Smug1*$^{-/-}$ mice suggest that SMUG1 deficiency promotes NAFLD but not NASH. NAFLD can be also considered a mitochondrial disease, for that reason we checked whether mitochondrial dysfunction was present in our model. Interestingly, *Smug1*$^{-/-}$ mice did not show impaired mitochondrial fatty acid oxidation activity or de novo lipogenesis. Analyses of adult SMUG1 deficient livers showed differential regulation of biological processes and molecular functions related to fat metabolism and steatosis. Among them was the upregulation of *cd36* and *vldlr*, which are associated with lipid influx, and suggest increased lipid uptake. Liver steatosis is as well linked to telomere erosion and it is known that presence of the SMUG1 substrates U and hmU in telomeric DNA, reduces binding of the protective shelterin proteins to telomeric sequences. For that reason, we assessed the status of telomeres in our *Smug1*$^{-/-}$ mice. We observed higher density of SMUG1 substrates accompanied by reduced average telomere length in *Smug1*$^{-/-}$ livers. As it was expected, senescence phenotypes such as lack of bone marrow cell proliferation and liver karyomegaly, were also present in *Smug1*$^{-/-}$ mice. Taken together, our data indicate a novel role of SMUG1 in regulating metabolic homeostasis of the liver.

Paper III: DNA glycosylase Neil3 controls vascular smooth muscle cell homeostasis during atherosclerosis development

The NEIL3 glycosylase has been previously linked to lipid metabolism and atherosclerosis. In this study we further investigated the role of NEIL3 in atherogenic plaque formation and the molecular mechanism behind. We observed significant increase of atherosclerotic lesion in the aortic root in *ApoE*$^{-/-}$/*Neil3*$^{-/-}$ compared to *ApoE*$^{-/-}$ mice. However, there was no effect on circulating lipids, cytokines or blood pressure upon NEIL3 absence in *ApoE*$^{-/-}$ mice. Assessing further the phenotypical characteristics of the atherosclerotic lesions, we observed an increase disorganization of the vascular smooth muscle cells (VSMC) in the media layer area of aortic roots of *ApoE*$^{-/-}$/*Neil3*$^{-/-}$ mice. These cells showed also increased proliferation upon NEIL3 deficiency (*ApoE*$^{-/-}$). Considering the proliferating and lipid-accumulating phenotypes, a phenotypic modulation of VSMCs is hinted. Additionally, experiments were carried out in primary human aortic VSMC line where NEIL3 expression was abrogated. Data derived from this cell line also confirmed that NEIL3 deficiency induces increased proliferation, phenotype switching and lipid uptake. mRNA expression analysis of NEIL3-KO VSMCs, revealed upregulation of markers associated with proliferation and phenotype switching (i.e., CD68, TGFβ, PAI-1, MMP2), driving to a lesional macrophage-like cell phenotype. In order to identify the molecular mechanism behind these phenotypes we assessed the genome integrity of our mouse model. Despite the unaltered levels of 8-oxoG in *ApoE*$^{-/-}$/*Neil3*$^{-/-}$ mice, decreased telomere length was observed which is a well know phenotype linked to atherosclerosis. Furthermore, the Akt signaling pathway, which has been extensively linked to VSMC proliferation and phenotype switching, was one of the most regulated in the mRNA-seq analyses of aortic VSMCs from *ApoE*$^{-/-}$/*Neil3*$^{-/-}$ mice. Additionally, aortic VSMCs and whole aorta showed increased Akt phosphorylation. Our data support a role of NEIL3 in atherosclerosis by regulating VSMC proliferation and phenotype switch probably promoted by the Akt signaling pathway.

Discussion

Telomeric DNA is a well know target of oxidative stress. Damage caused by oxidative stress is considered one of the primary causes of telomere erosion. This notion is supported by the fact that telomere attrition rate in cells is significantly reduced under hypoxic culture conditions or in the presence of antioxidants (von Zglinicki, 2002; Saretzki and Von Zglinicki, 2002). Damaged telomeric DNA leads to telomere dysfunction which is one of the main causes of chromosomal aberrations and, in turn, genome instability. *In vitro* and *in vivo* studies have shown that BER is one of the mechanisms that safeguard telomeres. Several glycosylases of the BER pathway have been found to be active on telomeric DNA, removing DNA lesions (Jia et al., 2015). In the current PhD thesis we investigated further the function of the SMUG1 and NEIL3 DNA-glycosylases at telomeric DNA and a novel, BER independent role of SMUG1 in telomere homeostasis.

SMUG1 associates with the Dyskerin complex and participates in RNA metabolism (Paper I)

SMUG1 DNA-glycosylase, apart from acting on double-stranded DNA and single-stranded DNA (Mi et al., 2009), was reported to act on RNA (Jobert et al., 2013). SMUG1 has incision activity on RNA substrates containing dU, which is not a natural RNA modification. Jobert et al. reported that upon SMUG1 depletion, 28S and 18S rRNAs presented increased hmU levels, establishing this modified base as present in rRNA and proposing a role for SMUG1 in rRNA quality control. Supporting this hypothesis, they further showed that SMUG1 affected the levels of the mature 18S and 5.8S rRNA species and associated *in vivo* with the 47S rRNA precursor (Jobert et al., 2013). It is known that some of these rRNA species are substrates also of the pseudouridine synthase, DKC1, and its location of action is the nucleoli and CBs (Gu et al., 2013). Jobert et al. showed that SMUG1 and DKC1 could be found colocalizing in both nucleoli and CBs (Jobert et al., 2013). Additionally, direct *in vitro* interaction was reported, not only, between SMUG1 and DKC1 but also between SMUG1 and the dyskerin complex components, GAR1 and NHP2. They further identified amino acids

in SMUG1 critical for this interaction, using mutagenesis experiments, establishing that there is direct binding between SMUG1 and the dyskerin complex (Figure 7) (Jobert et al., 2013). Considering the importance of DKC1 and the dyskerin complex in telomerase biogenesis and telomere maintenance, this interaction suggested that a role of SMUG1 would be possible in telomeric biology.

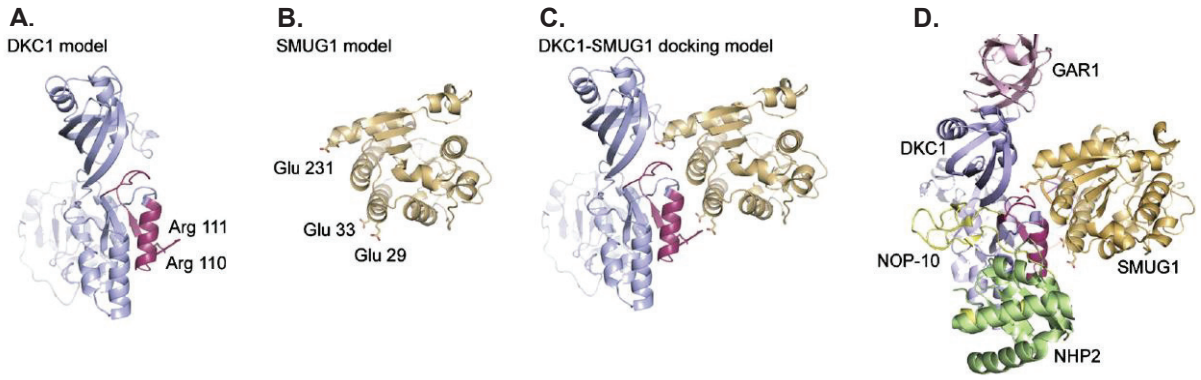


Figure 7. Structural model of the DKC1-SMUG1 interaction. *DKC1 model (blue) illustrating the amino acids interacting with SMUG1 (A). SMUG1 model illustrating the amino acids interacting with DKC1 (B). Depiction of the DKC1-SMUG1 docking model (C). Three-dimensional model of the SMUG1-DKC1 interaction, together with the DKC1 partners NOP-10 (yellow), NHP2 (green) and GAR1 (purple). Figure from Jobert et al., 2013.*

Prompted by the above findings, we wanted to further investigate whether SMUG1 affected any DKC1 related phenotypes. Using *Smug1*$^{+/+}$ and *Smug1*$^{-/-}$ MEFs we assessed the distribution pattern and expression of the DKC1 protein during S phase. In WT cells a characteristic ring shaped DKC1 structure was observed, whereas in *Smug1*$^{-/-}$ cells the number of these structures were significantly reduced. The DKC1 structural formations coincided with the round nucleoli formations, identified by BrdU staining (Zhang et al., 2007). These data indicate that the SMUG1/DKC1 interaction is crucial for the canonical distribution pattern of DKC1 in nucleoli. Apart from the established role of nucleoli in rRNA biogenesis, it has been shown that they are a site where the primary steps of the telomerase biogenesis take place. Lee et al. proposed that telomerase RNP, during early S phase, is accompanied by the protein TCAB1 in nucleoli and then transported together to the Cajal bodies during mid-S/late S phase (Lee et al., 2014). Interestingly, SMUG1 was found colocalizing with the CBs protein, coilin. Considering the interaction of SMUG1 with the dyskerin complex and association of SMUG1 with the location of telomerase biogenesis, this support the hypothesis that SMUG1 might affect telomere homeostasis.

Loss of SMUG1 leads to telomere defects in human cells

It is already established by many groups that several BER proteins act on telomeres (Jia et al., 2015). A possible role of SMUG1 in telomere biology as a BER protein is therefore not unprecedented. As it has been previously observed, unresolved damage at telomeric DNA, leads to formation of TIFs and moderate but significant telomere erosion (Kaul et al., 2011; Takai et al., 2003). In paper I we show that even though a SMUG1-telomere interaction was observed via TeloChip experiments, human cells deficient in SMUG1 did not present TIFs. Nevertheless, SMUG1-KO HAP1 cells showed an extreme reduction of telomere length accompanied by an increase of SMUG1-substrates at telomeres. Furthermore, an abnormal TRF1 and TRF2 distribution pattern was observed together with decreased POT1 and TRF2 binding. The atypical distribution pattern of the shelterin proteins might be attributed to the unavailable substrate that the short telomeres provide. However, the interference of DNA lesions with binding of shelterin components is another possible interpretation as two main substrates of SMUG1, uracil (Vallabhaneni et al., 2015) and hmU (Theruvathu et al., 2014) were reported to impair binding of TRF2 and POT1 to telomeric repeats. The presence of these lesions in telomeric DNA might also explain the more severe phenotype of POT1 binding in SMUG1 KO cells. Our data are in line with the observation of Vallabhaneni et al. that, in UNG-KO cells, POT1/TPP1 have lower binding affinity to telomeres (Vallabhaneni et al., 2015). UNG and SMUG1 share substrates and we would postulate that they both repair telomeric DNA (Alsoe et al., 2017). Thus, even though we cannot exclude that the BER activity of SMUG1 probably enhances the phenotype, our data suggest that this extreme telomere erosion is due to the dysregulation of a different pathway.

Telomerase biogenesis is dependent on SMUG1 in human cells

Encouraged by the previous observations and taking into account the correlation of SMUG1 with DKC1, we further investigated if SMUG1 is important for telomerase function. Indeed, in human SMUG1-KO cells, telomerase activity was dramatically downregulated. Several groups have shown that when any of the essential components of the telomerase enzyme is absent or downregulated telomerase levels are

dramatically affected (Heiss et al., 1998; Ibáñez-Cabellos et al., 2018). For that reason, we wanted to assess the status of the main telomerase components of our experimental model. Even though DKC1 (data not shown) and hTERT levels were not affected by SMUG1 deficiency, the levels of the RNA telomerase component, *hTERC*, were dramatically decreased. This finding is in line with the proposed role of SMUG1 in RNA metabolism (Jobert et al., 2013). In order to validate our hypothesis that *hTERC* levels were the limiting factor for telomerase activity, we transiently re-introduced both *hTERC* and hTERT in SMUG1-KO cells. While hTERT did not shift the balance of SMUG1-KO cells, a reconstitution of telomerase activity was seen when *hTERC* was overexpressed. Wanting to confirm that SMUG1 was responsible for the limited telomerase levels, we generated SMUG1 KO clones that stably re-expressed WT SMUG1. Indeed, these cells had significantly increased *hTERC* levels which was sufficient to rescue 70% to 90% the telomerase activity levels. Consecutively, telomere length in these clones was increased significantly. In order to elucidate which characteristic is responsible for the *hTERC* reduction, we assessed the effect of expressing SMUG1 mutants on telomere length. The SMUG1 H293L mutant is unable to bind nucleic acids (Matsubara et al., 2004), did not fully rescue the phenotype, even though telomere length was longer than in the SMUG1-KO cells. In order to elucidate whether the SMUG1-DKC1 interaction was essential for the canonical levels of telomerase we created a new, DKC1-binding defective SMUG1 mutant. When cells were complemented with this SMUG1 mutant (SMUG1 E29/33/231R), telomere length was mirroring the SMUG1-KO cells phenotype. These experiments supported the notion that the proposed role of SMUG1 RNA metabolism is not limited only to rRNA molecules but extends to the RNA component of telomerase as well (Jobert et al., 2013). SMUG1 seems to be essential for canonical *hTERC* levels and telomerase activity in HAP1 cells, with the DKC1/SMUG1 interaction being of great importance for this function.

SMUG1 is directly involved in co-transcriptional processing of *hTERC*

Impaired RNA polymerase II transcription would be a possible explanation for the dramatic decrease of *hTERC* levels in SMUG1-KO cells (Aalbers et al., 2012; Zhao et al., 1998). However, both transcription rate and the occupancy of RNAPII at the *hTERC* promoter and coding region were unaffected. Strikingly, SMUG1 was found to be present at the promoter and within the coding region of the *hTERC* gene, at the actively transcribed RNAPII complex. No difference in the transcription rate of the telomerase RNA component was detected upon SMUG1 depletion, suggesting that SMUG1 is not essential for RNAPII activity in the *hTERC* locus. As it has been shown before, SMUG1 can bind RNA molecules and in our study, we showed that this DNA glycosylase is binding the RNA telomerase component (Jobert et al., 2013). Since we found that SMUG1 associates with *hTERC* molecules of RNA extracted from cells we postulated that this interaction was direct, without any other protein aiding the association. Using RNA-sequencing (RNA-seq) analysis, we discovered a difference in the number of *hTERC* reads. SMUG1-KO cells presented reduced number of reads downstream of the core pseudoknot domain in the *hTERC* sequence, where DKC1 is known to bind (Chen et al., 2000). The *hTERC* molecule, as assembled in the active telomerase holoenzyme, is a product of a multistep maturation mechanism in which several other intermediate species are produced during this process (MacNeil et al., 2016). The differences we observed in our initial RNA-seq analysis, prompted us to investigate further these intermediate *hTERC* products and the possible role of SMUG1 in the h*TERC* maturation pathway. It has been shown that depletion of proteins that participate in several steps of the *hTERC* maturation process, from the DKC1-H/ACA scaffolding (Heiss et al., 1998) to polyadenylation-deadenylation steps and exosomal degradation (Shukla et al., 2016), can lead to an equilibrium imbalance of the levels of the *hTERC* molecules. Nguyen et al. suggested that extremely long 3' extended species are RNAPII read-through products (Nguyen et al. 2015). However, in the absence of SMUG1, we could not detect significantly increased extremely long *hTERC* molecules. Thus, we can deduce that, most probably, transcriptional termination is unaffected. Interestingly, short 3' extended products were increased and 3'-rapid amplification of cDNA ends (RACE)-seq experiments revealed slightly elevated levels of ployadenylated *hTERC* molecules in the SMUG1-KO cells. The balance of the

intermediate *hTERC* species was disturbed upon SMUG1 deletion even though RNAPII transcription rate was unaffected. Considering our finding that SMUG1 is present at the *hTERC* gene locus, we proposed that SMUG1 participates in co-transcriptional processing of *hTERC*.

We expected that by silencing key players of the *hTERC* maturation pathway, we would shift the balance of the intermediate species in the SMUG1-WT and KO cells (Shukla et al., 2016). As expected, when knocking down the major RNA exosome component EXOSC10, which competes with PARN for the degradation of polyA *hTERC* intermediates, *hTREC* levels were increased (Shukla et al., 2016). Strikingly, the same phenomenon was not observed in SMUG1-KO cells. This suggested that the intermediate *hTERC* molecules of SMUG1-KO cells are processed in an EXOSC10 independent manner. The 3' end processing pathway of h*TERC* is finely tuned, with a very thin line between polyadenylation and degradation. The disruption of the canonical process by silencing one or more of its components, gave results that were difficult to interpretation, perhaps a result of compensatory switches of pathway usage. Thus, we were not able to draw confident conclusions on precisely how SMUG1 deficiency imbalance this process. Nevertheless, our data indicate that even though mild *hTERC* processing defects might be present, the end processing machinery is, in essence, functional in SMUG1 deprived cells.

SMUG1 regulates the presence of *hTERC* modified bases and ensures an efficient DKC1 binding

The fact that SMUG1 is able to bind *hTERC* molecules of RNA extracted from cells and not *in vitro* transcribed, suggested that SMUG1 acts on modified bases. Considering the function of SMUG1, it is possible that the SMUG1 substrates in *hTERC* are random damaged bases. However, the SMUG1/*hTERC* interaction was unaffected when cells were grown in the presence of 5-hydroxymethyluridine, which together with deoxyU are the two known SMUG1 substrates in RNA (Jobert et al., 2013). In order to assess the frequency of SMUG1 substrates in the *hTERC* sequence, we developed a qPCR-based assay where we could examine the frequency of SMUG1 associated ribonucleotides in different *hTERC* regions. Interestingly, this method revealed that the

SMUG1 substrates are likely located in an *hTERC* region between the CR4/CR5 domain and the H box. The fact that SMUG1 substrates appear to be located in a specific region of the telomerase RNA and not homogeneously distributed throughout the molecule, is arguing against the substrate being randomly damaged RNA bases. Instead, it supports the notion that SMUG1 associates with *hTERC* modifications of a specific nature. Many years of research have unveiled a variety of post-transcriptional RNA modifications essential for the regulation of gene expression, RNA stability and protein translation (Fu et al., 2014; Wang and He, 2014). An RNA modification that can be found in both coding and noncoding RNAs is the methylation of cytosine which gives rise to the ribonucleoside 5-methylcytosine (5-mC) (Motorin et al., 2010). Enzymatic hydroxymethylation 5-mC forms 5-hydroxymethylcytosine (5-hmC) which is a molecule that has been identified as well in RNA and is proposed as a common RNA modification (Fu et al., 2014; Huber et al., 2015). When 5-hmC is subjected to deamination, hmU is produced which could be a possible substrate of SMUG1 in the *hTERC* molecule. Interestingly, two methylated sites, 5-mC at C323 and C445, have been identified in the *hTERC* region that our data indicate a SMUG1 association (Squires et al., 2012). Thus, we postulate that these *hTERC* sites represent possible SMUG1 substrates. A recent study on the function of the HuR protein further support this notion, since, similar phenotypes as seen in our SMUG1-KO model were observed upon depletion of HuR (Tang et al., 2018). More specifically, HuR was found in Cajal bodies, regulating the levels of the telomerase RNA component and affecting telomerase activity. HuR is known to post-transcriptionally regulate a number of coding and non-coding RNAs and Tang et al. proposed that HuR promotes TERT/*TERC* assembly via *TERC* methylation (Simone and Keene, 2013; Tang et al., 2018; Wang et al., 2012). Furthermore, the NEIL3 DNA-glycosylase has recently been associated to epigenetic regulation involving 5-mC and 5-hmC (Olsen et al., 2017). This new function hinted by NEIL3 and SMUG1, might represent novel roles of DNA glycosylases in epiregulation of nucleic acids regardless of whether the modification is present in DNA or in RNA.

An important aspect of the h*TERC* region that SMUG1 seems to possess is that it overlaps with the sequence that DKC1 binds (Ashbridge et al., 2009). It is well established that assembly of the pre-RNP complex, via DKC1 binding, with *hTERC* is of essential importance for the formation of the mature telomerase enzyme (MacNeil et al., 2016). Consequently, a possible explanation of the extreme telomerase related

phenotypes in SMUG1-KO cells would be that SMUG1 regulated the presence of *hTERC* modifications that are essential for DKC1 binding. Assessing the DKC1/*hTERC* interaction in our experimental model, we observed that DKC1 association to *hTERC* was decreased by almost 50%. Furthermore, when SMUG1 WT protein was re-expressed in SMUG1-KO cells, the DKC1/*hTERC* binding was fully rescued. However, the expression of SMUG1 mutants (nucleic acid- and DKC1- binding mutants) showed increased immunoprecipitation of *hTERC* compared to the SMUG1-KO cells but strongly reduced when compared to the SMUG1-WT cells (Figure 8). This increased association between DKC1 and *hTERC* in SMUG1-KO cells reconstituted with SMUG1 WT stabilized *hTERC*, leading to a partially restored telomerase activity and increased telomere length. Although we observed an increased DKC1/*hTERC* association when expressing both SMUG1 mutants, only the nucleic acid binding mutant displayed a telomere length increase, suggesting that the SMUG1/DKC1 binding is essential for the restoration of telomere defects. This data may be explained by the fact that, even though less efficiently, DKC1 is capable of scaffolding to *hTERC* through non-canonical binding positions (Ashbridge et al., 2009). The above data are in line with our hypothesis that SMUG1, by modifying h*TERC* sites, promotes DKC1 binding to the RNA telomerase component.

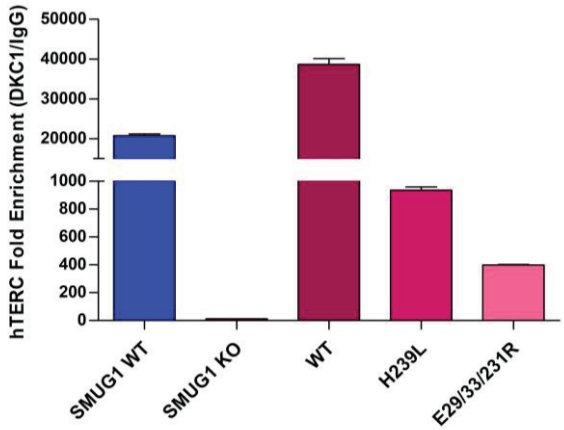


Figure 8. *qPCR showing hTERC immunoprecipitation by DKC1 in HAP1 cells and in SMUG1 KO clones stably expressing SMUG1 WT protein, SMUG1 unable to bind nucleic acids (H239L) and SMUG1 unable to bind DKC1 (E29/33/231R). Data are presented as fold enrichment above the sample specific background for hTERC.*

The molecular events that take place in the maturation process of the telomerase enzyme, and especially of its RNA molecule, are not yet completely understood. Only the last few years, several new proteins have been added to the long list of molecules that participate in this process (Shukla et al., 2016; Tang et al., 2018; Tseng et al., 2018). There might be key elements missing from the current understanding in order for us to fully comprehend precisely how SMUG1 acts in *hTERC* maturation.

mSMUG1 acts primarily as a canonical BER glycosylase at telomeres (Paper I & II)

As it has been previously discussed, in human cells, even though SMUG1 seems to act mainly in the maturation process of the *hTERC* molecule, it also acts as a canonical BER glycosylase at telomeres. The consequences of deleting SMUG1 in human cells are more prominent in the telomere elongation mechanism with a direct effect on telomere length, but when these conditions are assessed in a mouse *in vivo* system, different effects were seen. In Paper I we show that *Smug1*$^{-/-}$ MEFs do not present telomerase defects. Although SMUG1 binds to the RNA component of telomerase, *mTERC* levels and telomerase activity (data not shown and Paper II) remained unaffected. As a consequence, telomere length did not change. Even though *hTERC* and telomerase maturation pathway in humans has been extensively studied the latest years, little is known about the molecular mechanisms in mice (Mochizuki et al., 2004). The secondary structure of mouse and human *TERC* molecule are quite similar but that does not apply on the sequence (Chen et al., 2000). Considering that the results of our study points to a role of SMUG1 intimately related to the sequence and specific modifications of the *hTERC* molecule, it is not unexpected the divergence of SMUG1 actions in these two species.

Mouse knockout models of DNA glycosylase, such as OGG1 and UNG, exhibit telomere defects, initiated by DNA damage at the telomeric sequence (Opresko et al., 2005; Vallabhaneni et al., 2015). As in human cells, murine SMUG1 binds the telomeric sequence and in SMUG1-KO mice, reduction of average telomere length was seen in some tissues, accompanied by DNA damage at telomeres. In *Smug1*$^{-/-}$ MEFs, even though telomere erosion was not detected, elevated levels of fragile telomeres were observed in the C-strand. This phenotype was also present in the bone marrow cells of the SMUG1-KO mice. The presence of fragile telomeres suggested impaired replication at the C-rich strand (Sfeir et al., 2009; Özer and Hickson, 2018). The susceptibility of this specific strand could be explained by the fact that telomeric C-strand would be expected to contain more uracil or hmU which are SMUG1 substrate in DNA (Vallabhaneni et al., 2015). Additionally, AID off-target effects at the telomeric sequence might be, in part, responsible for the telomere defects presented in absence of SMUG1 both in human cells and mice (Cortizas et al., 2016; Dingler et al., 2014). It

is well known that telomeric sequence cannot be methylated, however subtelomes are abundant in CpG elements which are heavily methylated or hydroxymethylated (Fraga et al., 2005; Vera et al., 2008). These subtelomeric epigenetic markers are believed to be regulators of telomere length and associated with several biological pathways, such as cellular reprogramming and carcinogenesis (García-Cao et al., 2004; Wang et al., 2013). We could postulate that 5-hmU, created by cytosine deamination of 5-hmC, could be SMUG1 substrates in the subtelomeric DNA sequence. Due to method limitations we were not in the position to identify the exact nature of the SMUG1 lesions that give rise to the described telomere related phenotypes. Moreover, we cannot exclude the possibility of SMUG1 having a role in telomerase biogenesis in mice as well, however, the BER activity of SMUG1 seems to be the dominant function in mice in the tissues studied.

Interestingly, two independent studies associate SMUG1 with the regulation of metabolism in obesity (Gawrieh et al., 2010; Pedersen et al., 2016). In particular, a single nucleotide polymorphism (SNP) in the SMUG1 gene was one of the top-ranked SNPs found in patients presenting diabetes remission after bariatric surgery (Pedersen et al., 2016), suggesting a role for SMUG1 in metabolic homeostasis. Additionally, SMUG1 was found to be among the strongest upregulated genes in morbidly obese patients with non-alcholic fatty liver disease (NAFLD) compared to non-NAFLD control (Gawrieh et al., 2010). It is well known that DNA damage is present in NAFLD and considering the correlation of telomeric DNA damage with senescence and NAFLD (Ogrodnik et al., 2017; Seki et al., 2002), we wanted to investigate further the *in vivo* relevance of SMUG1 in this regard (Paper II). When assessing the levels of telomerase activity of the *Smug1*$^{-/-}$ mice, we did not observe any significant and consistent effect. This is in line with our previous findings regarding the impact of SMUG1 in the telomere homeostasis of our mouse model. However, telomere length was decreased and at the same time, increase of telomeric damage was observed in liver tissue of *Smug1*$^{-/-}$ mice. Additionally, increased nuclear size suggest hepatocyte senescence. Adult Smug1$^{-/-}$ animals had increased body weight and showed characteristic signs of NAFLD including fat accumulation in liver. Alterations of lipid profiles and a number of differentially expressed genes, related to fat metabolism, together suggest a mechanism that promotes fat storage in liver. These data are in line with the current notion that insufficient DNA repair induces cell senescence, possibly through

dysfunctional telomere maintenance (Hartmann et al., 2011). Taken together, the above observations indicate a novel role of SMUG1 in the onset of NAFLD, possibly due to hepatocyte cell senescence driven by impaired telomere maintenance.

NEIL3 deficiency promotes atherogenesis and is linked to telomere erosion (Paper III)

SMUG1 is not the only BER glycosylase that is linked to liver steatosis and lipid metabolism (Skarpengland et al., 2016). It has been previously shown by Skarpengland et al., that in an *ApoE*$^{-/-}$ background and high fat diet, absence of NEIL3 leads to enhanced liver steatosis. This is believed to happen due to increased triglyceride (TG) and monounsaturated fatty acid (MUFA) levels in the liver of *Neil3*$^{-/-}$*ApoE*$^{-/-}$ mice. Furthermore, these mice presented changes of the hepatic lipid metabolism pathways accompanied by increased atherosclerosis. The association of NEIL3 DNA-glycosylase with atherosclerotic plaques was further supported by the findings of increased expression of NEIL3, in human carotid plaques (Skarpengland et al., 2016). Additionally, the HUNT study results linked a specific NEIL3 variant with increased risk of myocardial infarction (Skarpengland et al., 2015). Prompted by these observations, we wanted to further evaluate the contribution of NEIL3 DNA-glycosylase in atherosclerosis development. When assessing the plaque formation of *ApoE*$^{-/-}$*/Neil3*$^{-/-}$ mice, which were fed a balanced diet (chow), we observed significantly increased atherosclerotic lesion area in the aortic root. This was happening while the systemic lipid profile remained unaffected. In order to better understand the development of atherosclerotic lesions in our model, we assessed the status of the cellular component of the atherosclerotic blood vessels, the vascular smooth muscle cells (VSMCs). VSMCs of the aorta media layer presented increased levels of disorganization. Interestingly, upon NEIL3 depletion, *ApoE*$^{-/-}$ mice exhibit increased proliferation of aortic VSMCs. This is in line with previous findings, where NEIL3 is associated with proliferation of specific cell types such as cardiac fibroblasts, neuronal stem and progenitor cells of mouse brain (Olsen et al., 2017; Sejersted et al., 2011). Increased proliferation was also observed in human aortic VSMCs upon silencing of NEIL3. Additionally, these cells presented upregulation of phenotype switching markers, among them, the CD68 macrophage marker. This come as no

surprise, since accumulation of macrophages which indicates increased lesional proliferation, has been shown to present in atherosclerotic lesions of *ApoE*$^{-/-}$/*Neil3*$^{-/-}$ mice (Olsen et al., 2017). Taken together, our findings suggest that NEIL3 deficiency may promote a phenotypical switch of VSMCs to a lesional macrophage-like cell phenotype contributing to atherosclerotic plaque progression.

Even though several phenotypes give a direct association of NEIL3 with atherosclerosis, there is no evidence for the exact molecular course of actions of the DNA glycosylase that are related to these phenotypes. Considering the established role of NEIL3 in DNA repair, the assessment of DNA damage in our model was the next step of this study. As it has been shown before (Skarpengland et al., 2016), the bulk level of oxidized DNA base lesions (measured by 8-oxoG DNA lesions) displayed no differences in liver and VSMCs of *ApoE*$^{-/-}$/*Neil3*$^{-/-}$ and *ApoE*$^{-/-}$ mice. However, we wanted to assess DNA damage in specific genomic regions that are linked to atherosclerosis, such as the telomeric sequence (Poch et al., 2004; Salpea and Humphries, 2010). This is the first study showing *in vivo* data of telomere decrease in bone marrow and liver of NEIL3 deficient young mice. This phenotype might be attributed to the formation of G4 structures that remain unresolved upon NEIL3 deficiency, leading to replication stalling and telomere erosion. The fact that this was only observed in young mice suggests that the *in vivo* system manages to balance and stabilize this phenomenon. Additionally, the telomere length erosion could be explained as an early onset senescence related phenotype which accompanies the development of atherosclerotic plaques. Further investigation of the molecular mechanism laying behind the NEIL3 association with atherosclerosis, revealed a connection with the Akt signaling pathway. This pathway is extensively linked to VSMC proliferation and phenotype switching (Fernandez-Hernando et al., 2009; Stabile et al., 2003). mRNA-seq analyses of aortic VSMCs, from *ApoE*$^{-/-}$/*Neil3*$^{-/-}$ mice, showed that the Akt pathway was one of the most regulated in our data. Furthermore, Akt phosphorylation of aortic VSMCs and whole aorta was increased upon NEIL3 deprivation of *ApoE*$^{-/--}$ mice. These data suggest that the Akt signaling pathway is implicated in the formation of atherogenic plaques upon Neil3 deficiency of atherosclerosis-prone mice. Taken together, we propose a role of NEIL3 in atherosclerosis by regulating VSMC proliferation and phenotype switch probably promoted by the Akt signaling pathway.

Conclusion and future perspectives

The current PhD dissertation is focused on the possible roles of the SMUG1 and NEIL3 DNA-glycosylases in telomeric homeostasis of human cells and mice. In our study, we unveiled a new role of SMUG1 in the processing pathway of the RNA component of telomerase. We propose that SMUG1, via modifying bases in the *hTERC* sequence, ensures proper DKC1 binding to the pre-RNP complex and thus promotes telomerase assembly. However, the line of experiments from which these conclusions were drawn were conducted in a specific human chronic myeloid leukemia cell line. Thus, we cannot exclude the possibility of the observed phenotypes to be cell type specific. This novel SMUG1 function did not appear to be present in our murine *Smug1*$^{-/-}$ model where we observed phenotypes that support a canonical BER DNA-glycosylase function. Furthermore, tissue specific telomere defects suggested a connection of the SMUG1 role in telomeric DNA, with senescence driven hepatic steatosis and impaired lipid metabolism. Apart from SMUG1, a focal point of this study was the NEIL3 DNA-glycosylase and how its function on telomeres is associated with an atherogenic background. We propose a role of NEIL3 in atherosclerosis by regulating phenotype switch and proliferation of VSMCs. In this context, the telomere erosion shown was possibly a consequence of the overall NEIL3 deficiency as a DNA repair protein.

Our data support an important role of DNA glycosylase in telomere homeostasis, by either facilitating telomerase biogenesis or by functioning as canonical BER DNA-glycosylases. The extent of their function in telomeres is yet to be discovered, since it seems that several pathways, cell types and tissues are affected when deprived from these DNA glycosylases.

For future studies, it is important to identify whether the novel SMUG1 function in *hTERC* processing is a new step in the general mechanism of *hTERC* maturation or if it is a cell type specific characteristic. Similar experiments, such as the ones conducted in the current thesis, could be designed in multiple cell lines, derived from different tissues. Additionally, the characterization of the SMUG1 substrate in the *hTERC* sequence would be essential in order to further elucidate the mechanism of *hTERC* processing. Development of specifically designed RNA-seq technology is required to answer questions regarding the *hTERC* modification state of SMUG1-KO cells, in order

to assess whether possible 5-hmU are present in *hTERC* sequence. Furthermore, it is worth mentioning that, *hTERC* shares many characteristics with several sno/sca RNA molecules. Their secondary structure is similar and the H/ACA box, is of great importance for their proper function (Cao et al., 2018). They associate with DKC1 and undergo 3' end processing with several common proteins involved (Cao et al., 2018). Considering the above, it would be interesting to study whether SMUG1 functions in maturation/biogenesis/quality control of sno/sca RNAs.

Even though our data suggest a role of SMUG1 as a canonical DNA glycosylase in telomeres, we did not observe any striking phenotypic consequence of the BER defect. We believe that this could perhaps be because of the UNG2 presence in our experimental models. Considering that UNG2 and SMUG1 have overlapping substrates, in humans and mice, it is possible that extreme telomere maintenance phenotypes would only be prominent after removing both, similar to class switch recombination (Dingler et al., 2014). UNG2/SMUG1-KO models could be an ideal tool for studying how SMUG1 acts on telomeric sequence. Furthermore, using specific human B-lymphoid cell lines or murine B-cells, we could also investigate possible off-target effect of AID in telomeres. However, currently there is no available mapping method distinguishing uracil from 5-hmU, making it hard to identify any unique functions of SMUG1 in DNA. A possible approach would be to focus on the subtelomeric regions which are known to be heavily methylated and hydroxymethylated.

In conclusion, further investigation of the roles of BER glycosylases in telomeres is needed, in order to understand the exact pathways involved in telomerase biogenesis, DNA damage response and possible epigenetic regulation in telomeres.

References

Aalbers, A.M., Kajigaya, S., van den Heuvel-Eibrink, M.M., van der Velden, V.H., Calado, R.T., and Young, N.S. (2012). Human telomere disease due to disruption of the CCAAT box of the TERC promoter. Blood *119*, 3060-3063.

Ahmed, A., and Tollefsbol, T. (2001). Telomeres and telomerase: basic science implications for aging. J Am Geriatr Soc *49*, 1105-1109.

Aikata, H., Takaishi, H., Kawakami, Y., Takahashi, S., Kitamoto, M., Nakanishi, T., Nakamura, Y., Shimamoto, F., Kajiyama, G., and Ide, T. (2000). Telomere reduction in human liver tissues with age and chronic inflammation. Experimental cell research *256*, 578-582.

Alsoe, L., Sarno, A., Carracedo, S., Domanska, D., Dingler, F., Lirussi, L., SenGupta, T., Tekin, N.B., Jobert, L., Alexandrov, L.B*., et al.* (2017). Uracil Accumulation and Mutagenesis Dominated by Cytosine Deamination in CpG Dinucleotides in Mice Lacking UNG and SMUG1. Scientific reports *7*, 7199.

Anderson, R., Lagnado, A., Maggiorani, D., Walaszczyk, A., Dookun, E., Chapman, J., Birch, J., Salmonowicz, H., Ogrodnik, M., Jurk, D*., et al.* (2019). Length-independent telomere damage drives post-mitotic cardiomyocyte senescence. The EMBO journal *38*.

Aravinthan, A., Scarpini, C., Tachtatzis, P., Verma, S., Penrhyn-Lowe, S., Harvey, R., Davies, S.E., Allison, M., Coleman, N., and Alexander, G. (2013). Hepatocyte senescence predicts progression in non-alcohol-related fatty liver disease. J Hepatol *58*, 549-556.

Armanios, M., and Blackburn, E.H. (2012). The telomere syndromes. Nat Rev Genet *13*, 693-704.

Ashbridge, B., Orte, A., Yeoman, J.A., Kirwan, M., Vulliamy, T., Dokal, I., Klenerman, D., and Balasubramanian, S. (2009). Single-molecule analysis of the human telomerase RNA.dyskerin interaction and the effect of dyskeratosis congenita mutations. Biochemistry *48*, 10858-10865.

Bandaru, V., Sunkara, S., Wallace, S.S., and Bond, J.P. (2002). A novel human DNA glycosylase that removes oxidative DNA damage and is homologous to Escherichia coli endonuclease VIII. DNA repair *1*, 517-529.

Beausejour, C.M., Krtolica, A., Galimi, F., Narita, M., Lowe, S.W., Yaswen, P., and Campisi, J. (2003). Reversal of human cellular senescence: roles of the p53 and p16 pathways. The EMBO journal *22*, 4212-4222.

Bischoff, C., Graakjaer, J., Petersen, H.C., Hjelmborg, J., Vaupel, J.W., Bohr, V., Koelvraa, S., and Christensen, K. (2005). The heritability of telomere length among the elderly and oldest-old. Twin Res Hum Genet *8*, 433-439.

Blasco, M.A., Lee, H.W., Rizen, M., Hanahan, D., DePinho, R., and Greider, C.W. (1997). Mouse models for the study of telomerase. Ciba Found Symp *211*, 160-170; discussion 170-166.

Bochman, M.L., Paeschke, K., and Zakian, V.A. (2012). DNA secondary structures: stability and function of G-quadruplex structures. Nature Reviews Genetics *13*, 770-780.

Bojovic, B., Booth, R.E., Jin, Y., Zhou, X., and Crowe, D.L. (2015). Alternative lengthening of telomeres in cancer stem cells in vivo. Oncogene *34*, 611-620.

Boorstein, R.J., Cummings, A., Jr., Marenstein, D.R., Chan, M.K., Ma, Y., Neubert, T.A., Brown, S.M., and Teebor, G.W. (2001). Definitive identification of mammalian 5-hydroxymethyluracil DNA N-glycosylase activity as SMUG1. The Journal of biological chemistry *276*, 41991-41997.

Brosh, R.M., Jr. (2013). DNA helicases involved in DNA repair and their roles in cancer. Nat Rev Cancer *13*, 542-558.

Brynolf, K., Eliasson, R., and Reichard, P. (1978). Formation of Okazaki fragments in polyoma DNA synthesis caused by misincorporation of uracil. Cell *13*, 573-580.

Calado, R.T., and Dumitriu, B. (2013). Telomere dynamics in mice and humans. Semin Hematol *50*, 165-174.

Calado, R.T., and Young, N.S. (2009). Telomere diseases. N Engl J Med *361*, 2353-2365.

Cao, T., Rajasingh, S., Samanta, S., Dawn, B., Bittel, D.C., and Rajasingh, J. (2018). Biology and clinical relevance of noncoding sno/scaRNAs. Trends Cardiovasc Med *28*, 81-90.

Cesare, A.J., and Reddel, R.R. (2010). Alternative lengthening of telomeres: models, mechanisms and implications. Nat Rev Genet *11*, 319-330.

Chai, W., Shay, J.W., and Wright, W.E. (2005). Human telomeres maintain their overhang length at senescence. Mol Cell Biol *25*, 2158-2168.

Chen, J.L., Blasco, M.A., and Greider, C.W. (2000). Secondary structure of vertebrate telomerase RNA. Cell *100*, 503-514.

Chen, L., Roake, C.M., Freund, A., Batista, P.J., Tian, S., Yin, Y.A., Gajera, C.R., Lin, S., Lee, B., Pech, M.F.*, et al.* (2018). An Activity Switch in Human Telomerase Based on RNA Conformation and Shaped by TCAB1. Cell *174*, 218-230.e213.

Chu, T.W., D'Souza, Y., and Autexier, C. (2016a). The Insertion in Fingers Domain in Human Telomerase Can Mediate Enzyme Processivity and Telomerase Recruitment to Telomeres in a TPP1-Dependent Manner. Mol Cell Biol *36*, 210-222.

Chu, T.W., MacNeil, D.E., and Autexier, C. (2016b). Multiple Mechanisms Contribute to the Cell Growth Defects Imparted by Human Telomerase Insertion in Fingers Domain Mutations Associated with Premature Aging Diseases. The Journal of biological chemistry *291*, 8374-8386.

Cooke, M.S., Evans, M.D., Dizdaroglu, M., and Lunec, J. (2003). Oxidative DNA damage: mechanisms, mutation, and disease. FASEB journal : official publication of the Federation of American Societies for Experimental Biology *17*, 1195-1214.

Cortizas, E.M., Zahn, A., Safavi, S., Reed, J.A., Vega, F., Di Noia, J.M., and Verdun, R.E. (2016). UNG protects B cells from AID-induced telomere loss. The Journal of experimental medicine *213*, 2459-2472.

Court, R., Chapman, L., Fairall, L., and Rhodes, D. (2005). How the human telomeric proteins TRF1 and TRF2 recognize telomeric DNA: a view from high-resolution crystal structures. EMBO Rep *6*, 39-45.

Darwanto, A., Theruvathu, J.A., Sowers, J.L., Rogstad, D.K., Pascal, T., Goddard, W., 3rd, and Sowers, L.C. (2009). Mechanisms of base selection by human single-stranded selective monofunctional uracil-DNA glycosylase. The Journal of biological chemistry *284*, 15835-15846.

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

de Lange, T. (2018). Shelterin-Mediated Telomere Protection. Annual review of genetics *52*, 223-247.

de Lange, T., and Petrini, J.H. (2000). A new connection at human telomeres: association of the Mre11 complex with TRF2. Cold Spring Harb Symp Quant Biol *65*, 265-273.

Deng, T., Huang, Y., Weng, K., Lin, S., Li, Y., Shi, G., Chen, Y., Huang, J., Liu, D., Ma, W.*, et al.* (2019). TOE1 acts as a 3' exonuclease for telomerase RNA and regulates telomere maintenance. Nucleic acids research *47*, 391-405.

Dingler, F.A., Kemmerich, K., Neuberger, M.S., and Rada, C. (2014). Uracil excision by endogenous SMUG1 glycosylase promotes efficient Ig class switching and impacts on A:T substitutions during somatic mutation. European journal of immunology *44*, 1925-1935.

Doksani, Y., Wu, J.Y., de Lange, T., and Zhuang, X. (2013). Super-resolution fluorescence imaging of telomeres reveals TRF2-dependent T-loop formation. Cell *155*, 345-356.

Donate, L.E., and Blasco, M.A. (2011). Telomeres in cancer and ageing. Philos Trans R Soc Lond B Biol Sci *366*, 76-84.

Dou, H., Mitra, S., and Hazra, T.K. (2003). Repair of oxidized bases in DNA bubble structures by human DNA glycosylases NEIL1 and NEIL2. The Journal of biological chemistry *278*, 49679-49684.

Duarte, V., Muller, J.G., and Burrows, C.J. (1999). Insertion of dGMP and dAMP during in vitro DNA synthesis opposite an oxidized form of 7,8-dihydro-8-oxoguanine. Nucleic acids research *27*, 496-502.

Egan, E.D., and Collins, K. (2012). Biogenesis of telomerase ribonucleoproteins. Rna *18*, 1747-1759.

Erdel, F., Kratz, K., Willcox, S., Griffith, J.D., Greene, E.C., and de Lange, T. (2017). Telomere Recognition and Assembly Mechanism of Mammalian Shelterin. Cell reports *18*, 41-53.

Fernandez-Hernando, C., Jozsef, L., Jenkins, D., Di Lorenzo, A., and Sessa, W.C. (2009). Absence of Akt1 reduces vascular smooth muscle cell migration and survival and induces features of plaque vulnerability and cardiac dysfunction during atherosclerosis. Arterioscler Thromb Vasc Biol *29*, 2033-2040.

Fischer, M., Quaas, M., Steiner, L., and Engeland, K. (2016). The p53-p21-DREAM-CDE/CHR pathway regulates G2/M cell cycle genes. Nucleic acids research *44*, 164-174.

Fortini, P., Parlanti, E., Sidorkina, O.M., Laval, J., and Dogliotti, E. (1999). The type of DNA glycosylase determines the base excision repair pathway in mammalian cells. The Journal of biological chemistry *274*, 15230-15236.

Fouquerel, E., Barnes, R.P., Uttam, S., Watkins, S.C., Bruchez, M.P., and Opresko, P.L. (2019). Targeted and Persistent 8-Oxoguanine Base Damage at Telomeres Promotes Telomere Loss and Crisis. Molecular cell *75*, 117-130.e116.

Fraga, M.F., Ballestar, E., Villar-Garea, A., Boix-Chornet, M., Espada, J., Schotta, G., Bonaldi, T., Haydon, C., Ropero, S., Petrie, K.*, et al.* (2005). Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. Nature genetics *37*, 391-400.

Frosina, G. (2000). Overexpression of enzymes that repair endogenous damage to DNA. Eur J Biochem *267*, 2135-2149.

Fu, L., Guerrero, C.R., Zhong, N., Amato, N.J., Liu, Y., Liu, S., Cai, Q., Ji, D., Jin, S.-G., Niedernhofer, L.J.*, et al.* (2014). Tet-mediated formation of 5-hydroxymethylcytosine in RNA. J Am Chem Soc *136*, 11582-11585.

Garavís, M., González, C., and Villasante, A. (2013). On the origin of the eukaryotic chromosome: the role of noncanonical DNA structures in telomere evolution. Genome biology and evolution *5*, 1142-1150.

García-Cao, M., O'Sullivan, R., Peters, A.H.F.M., Jenuwein, T., and Blasco, M.A. (2004). Epigenetic regulation of telomere length in mammalian cells by the Suv39h1 and Suv39h2 histone methyltransferases. Nature genetics *36*, 94-99.

Gawrieh, S., Baye, T.M., Carless, M., Wallace, J., Komorowski, R., Kleiner, D.E., Andris, D., Makladi, B., Cole, R., Charlton, M.*, et al.* (2010). Hepatic gene networks in morbidly obese patients with nonalcoholic fatty liver disease. Obes Surg *20*, 1698-1709.

Gomes, N.M., Ryder, O.A., Houck, M.L., Charter, S.J., Walker, W., Forsyth, N.R., Austad, S.N., Venditti, C., Pagel, M., Shay, J.W*., et al.* (2011). Comparative biology of mammalian telomeres: hypotheses on ancestral states and the roles of telomeres in longevity determination. Aging Cell *10*, 761-768.

Greider, C.W. (1999). Telomeres do D-loop-T-loop. Cell *97*, 419-422.

Greider, C.W., and Blackburn, E.H. (1989). A telomeric sequence in the RNA of Tetrahymena telomerase required for telomere repeat synthesis. Nature *337*, 331-337.

Griffith, J.D., Comeau, L., Rosenfield, S., Stansel, R.M., Bianchi, A., Moss, H., and de Lange, T. (1999). Mammalian telomeres end in a large duplex loop. Cell *97*, 503-514.

Gu, B.W., Ge, J., Fan, J.M., Bessler, M., and Mason, P.J. (2013). Slow growth and unstable ribosomal RNA lacking pseudouridine in mouse embryonic fibroblast cells expressing catalytically inactive dyskerin. FEBS letters *587*, 2112-2117.

Hall, D.B., Holmlin, R.E., and Barton, J.K. (1996). Oxidative DNA damage through long-range electron transfer. Nature *382*, 731-735.

Hanssen-Bauer, A., Solvang-Garten, K., Sundheim, O., Pena-Diaz, J., Andersen, S., Slupphaug, G., Krokan, H.E., Wilson, D.M., 3rd, Akbari, M., and Otterlei, M. (2011). XRCC1 coordinates disparate responses and multiprotein repair complexes depending on the nature and context of the DNA damage. Environmental and molecular mutagenesis *52*, 623-635.

Harley, C.B., Futcher, A.B., and Greider, C.W. (1990). Telomeres shorten during ageing of human fibroblasts. Nature *345*, 458-460.

Hartmann, D., Srivastava, U., Thaler, M., Kleinhans, K.N., N'Kontchou, G., Scheffold, A., Bauer, K., Kratzer, R.F., Kloos, N., Katz, S.F*., et al.* (2011). Telomerase gene mutations are associated with cirrhosis formation. Hepatology *53*, 1608-1617.

Heaphy, C.M., Subhawong, A.P., Hong, S.M., Goggins, M.G., Montgomery, E.A., Gabrielson, E., Netto, G.J., Epstein, J.I., Lotan, T.L., Westra, W.H*., et al.* (2011). Prevalence of the alternative lengthening of telomeres telomere maintenance mechanism in human cancer subtypes. Am J Pathol *179*, 1608-1615.

Hegde, M.L., Hegde, P.M., Bellot, L.J., Mandal, S.M., Hazra, T.K., Li, G.M., Boldogh, I., Tomkinson, A.E., and Mitra, S. (2013). Prereplicative repair of oxidized bases in the human genome is mediated by NEIL1 DNA glycosylase together with replication proteins. Proceedings of the National Academy of Sciences of the United States of America *110*, E3090-3099.

Heiss, N.S., Knight, S.W., Vulliamy, T.J., Klauck, S.M., Wiemann, S., Mason, P.J., Poustka, A., and Dokal, I. (1998). X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. Nature genetics *19*, 32-38.

Henderson, P.T., Delaney, J.C., Muller, J.G., Neeley, W.L., Tannenbaum, S.R., Burrows, C.J., and Essigmann, J.M. (2003). The hydantoin lesions formed from oxidation of 7,8-dihydro-8-oxoguanine are potent sources of replication errors in vivo. Biochemistry *42*, 9257-9262.

Henriksson, S., and Farnebo, M. (2015). On the road with WRAP53β: guardian of Cajal bodies and genome integrity. Front Genet *6*, 91-91.

Hildrestrand, G.A., Neurauter, C.G., Diep, D.B., Castellanos, C.G., Krauss, S., Bjørås, M., and Luna, L. (2009). Expression patterns of Neil3 during embryonic brain development and neoplasia. BMC Neuroscience *10*, 45.

Huber, S.M., van Delft, P., Mendil, L., Bachman, M., Smollett, K., Werner, F., Miska, E.A., and Balasubramanian, S. (2015). Formation and abundance of 5-hydroxymethylcytosine in RNA. Chembiochem *16*, 752-755.

Hwang, H., Buncher, N., Opresko, P.L., and Myong, S. (2012). POT1-TPP1 regulates telomeric overhang structural dynamics. Structure (London, England : 1993) *20*, 1872-1880.

Ibáñez-Cabellos, J.S., Pérez-Machado, G., Seco-Cervera, M., Berenguer-Pascual, E., García-Giménez, J.L., and Pallardó, F.V. (2018). Acute telomerase components depletion triggers oxidative stress as an early event previous to telomeric shortening. Redox Biol *14*, 398-408.

Ide, H., Kow, Y.W., and Wallace, S.S. (1985). Thymine glycols and urea residues in M13 DNA constitute replicative blocks in vitro. Nucleic acids research *13*, 8035-8052.

Iliakis, G., Wang, Y., Guan, J., and Wang, H. (2003). DNA damage checkpoint control in cells exposed to ionizing radiation. Oncogene *22*, 5834-5847.

Jady, B.E., Richard, P., Bertrand, E., and Kiss, T. (2006). Cell cycle-dependent recruitment of telomerase RNA and Cajal bodies to human telomeres. Mol Biol Cell *17*, 944-954.

Jia, P., Her, C., and Chai, W. (2015). DNA excision repair at telomeres. DNA repair *36*, 137-145.

Jobert, L., Skjeldam, H.K., Dalhus, B., Galashevskaya, A., Vagbo, C.B., Bjoras, M., and Nilsen, H. (2013). The human base excision repair enzyme SMUG1 directly interacts with DKC1 and contributes to RNA quality control. Molecular cell *49*, 339-345.

Karlseder, J., Hoke, K., Mirzoeva, O.K., Bakkenist, C., Kastan, M.B., Petrini, J.H.J., and de Lange, T. (2004). The telomeric protein TRF2 binds the ATM kinase and can inhibit the ATM-dependent DNA damage response. PLoS Biol *2*, E240-E240.

Kauffmann, A., Rosselli, F., Lazar, V., Winnepenninckx, V., Mansuet-Lupo, A., Dessen, P., van den Oord, J.J., Spatz, A., and Sarasin, A. (2008). High expression of DNA repair pathways is associated with metastasis in melanoma patients. Oncogene *27*, 565-573.

Kaul, Z., Cesare, A.J., Huschtscha, L.I., Neumann, A.A., and Reddel, R.R. (2011). Five dysfunctional telomeres predict onset of senescence in human cells. EMBO Rep *13*, 52-59.

Kemmerich, K., Dingler, F.A., Rada, C., and Neuberger, M.S. (2012). Germline ablation of SMUG1 DNA glycosylase causes loss of 5-hydroxymethyluracil- and UNG-backup uracil-excision activities and increases cancer predisposition of Ung-/-Msh2-/- mice. Nucleic acids research *40*, 6016-6025.

Kim, N.W., Piatyszek, M.A., Prowse, K.R., Harley, C.B., West, M.D., Ho, P.L., Coviello, G.M., Wright, W.E., Weinrich, S.L., and Shay, J.W. (1994). Specific association of human telomerase activity with immortal cells and cancer. Science (New York, NY) *266*, 2011-2015.

Krokan, H.E., and Bjørås, M. (2013). Base excision repair. Cold Spring Harb Perspect Biol *5*, a012583-a012583.

Krokeide, S.Z., Laerdahl, J.K., Salah, M., Luna, L., Cederkvist, F.H., Fleming, A.M., Burrows, C.J., Dalhus, B., and Bjoras, M. (2013). Human NEIL3 is mainly a monofunctional DNA glycosylase removing spiroimindiohydantoin and guanidinohydantoin. DNA repair *12*, 1159-1164.

Lee, J.H., Lee, Y.S., Jeong, S.A., Khadka, P., Roth, J., and Chung, I.K. (2014). Catalytically active telomerase holoenzyme is assembled in the dense fibrillar component of the nucleolus during S phase. Histochem Cell Biol *141*, 137-152.

Lee, S.S., Bohrson, C., Pike, A.M., Wheelan, S.J., and Greider, C.W. (2015). ATM Kinase Is Required for Telomere Elongation in Mouse and Human Cells. Cell reports *13*, 1623-1632.

Leguisamo, N.M., Gloria, H.C., Kalil, A.N., Martins, T.V., Azambuja, D.B., Meira, L.B., and Saffi, J. (2017). Base excision repair imbalance in colorectal cancer has prognostic value and modulates response to chemotherapy. Oncotarget *8*, 54199-54214.

Leulliot, N., Godin, K.S., Hoareau-Aveilla, C., Quevillon-Cheruel, S., Varani, G., Henry, Y., and Van Tilbeurgh, H. (2007). The box H/ACA RNP assembly factor Naf1p contains a domain homologous to Gar1p mediating its interaction with Cbf5p. J Mol Biol *371*, 1338-1353.

Levy, M.Z., Allsopp, R.C., Futcher, A.B., Greider, C.W., and Harley, C.B. (1992). Telomere end-replication problem and cell aging. J Mol Biol *225*, 951-960.

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. Nature *362*, 709-715.

Lipps, H.J., and Rhodes, D. (2009). G-quadruplex structures: in vivo evidence and function. Trends Cell Biol *19*, 414-422.

Liu, L., Bailey, S.M., Okuka, M., Munoz, P., Li, C., Zhou, L., Wu, C., Czerwiec, E., Sandler, L., Seyfang, A*., et al.* (2007). Telomere lengthening early in development. Nat Cell Biol *9*, 1436-1441.

Liu, M., Bandaru, V., Bond, J.P., Jaruga, P., Zhao, X., Christov, P.P., Burrows, C.J., Rizzo, C.J., Dizdaroglu, M., and Wallace, S.S. (2010). The mouse ortholog of NEIL3 is a functional DNA glycosylase in vitro and in vivo. Proceedings of the National Academy of Sciences of the United States of America *107*, 4925-4930.

Lundblad, V., and Blackburn, E.H. (1993). An alternative pathway for yeast telomere maintenance rescues est1- senescence. Cell *73*, 347-360.

Luo, W., Muller, J.G., Rachlin, E.M., and Burrows, C.J. (2000). Characterization of spiroiminodihydantoin as a product of one-electron oxidation of 8-Oxo-7,8-dihydroguanosine. Org Lett *2*, 613-616.

Luo, W., Muller, J.G., Rachlin, E.M., and Burrows, C.J. (2001). Characterization of hydantoin products from one-electron oxidation of 8-oxo-7,8-dihydroguanosine in a nucleoside model. Chem Res Toxicol *14*, 927-938.

Macias, S., Cordiner, R.A., Gautier, P., Plass, M., and Caceres, J.F. (2015). DGCR8 Acts as an Adaptor for the Exosome Complex to Degrade Double-Stranded Structured RNAs. Molecular cell *60*, 873-885.

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

MacNeil, D.E., Bensoussan, H.J., and Autexier, C. (2016). Telomerase Regulation from Beginning to the End. Genes *7*, 64.

Makarov, V.L., Hirose, Y., and Langmore, J.P. (1997). Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. Cell *88*, 657-666.

Masaoka, A., Matsubara, M., Tanaka, T., Terato, H., Ohyama, Y., Kubo, K., and Ide, H. (2003). Repair roles of hSMUG1 assessed by damage specificity and cellular activity. Nucleic Acids Res Suppl, 263-264.

Maser, R.S., and DePinho, R.A. (2004). Telomeres and the DNA damage response: why the fox is guarding the henhouse. DNA repair *3*, 979-988.

Matsubara, M., Tanaka, T., Terato, H., Ohmae, E., Izumi, S., Katayanagi, K., and Ide, H. (2004). Mutational analysis of the damage-recognition and catalytic mechanism of human SMUG1 DNA glycosylase. Nucleic acids research *32*, 5291-5302.

Matthews, C., Gorenne, I., Scott, S., Figg, N., Kirkpatrick, P., Ritchie, A., Goddard, M., and Bennett, M. (2006). Vascular smooth muscle cells undergo telomere-based senescence in human atherosclerosis: effects of telomerase and oxidative stress. Circ Res *99*, 156-164.

McKerlie, M., Lin, S., and Zhu, X.-D. (2012). ATM regulates proteasome-dependent subnuclear localization of TRF1, which is important for telomere maintenance. Nucleic acids research *40*, 3975-3989.

Mi, R., Dong, L., Kaulgud, T., Hackett, K.W., Dominy, B.N., and Cao, W. (2009). Insights from xanthine and uracil DNA glycosylase activities of bacterial and human SMUG1: switching SMUG1 to UDG. J Mol Biol *385*, 761-778.

Mochizuki, Y., He, J., Kulkarni, S., Bessler, M., and Mason, P.J. (2004). Mouse dyskerin mutations affect accumulation of telomerase RNA and small nucleolar RNA, telomerase activity, and ribosomal RNA processing. Proceedings of the National Academy of Sciences of the United States of America *101*, 10756-10761.

Morin, G.B. (1989). The human telomere terminal transferase enzyme is a ribonucleoprotein that synthesizes TTAGGG repeats. Cell *59*, 521-529.

Motorin, Y., Lyko, F., and Helm, M. (2010). 5-methylcytosine in RNA: detection, enzymatic formation and biological functions. Nucleic acids research *38*, 1415-1430.

Moyzis, R.K., Buckingham, J.M., Cram, L.S., Dani, M., Deaven, L.L., Jones, M.D., Meyne, J., Ratliff, R.L., and Wu, J.R. (1988). A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes. Proceedings of the National Academy of Sciences of the United States of America *85*, 6622-6626.

Muller, H. (1938). The remaking of chromosomes. Collecting Net *13*, 181–198.

Narciso, L., Fortini, P., Pajalunga, D., Franchitto, A., Liu, P., Degan, P., Frechet, M., Demple, B., Crescenzi, M., and Dogliotti, E. (2007). Terminally differentiated muscle cells are defective in base excision DNA repair and hypersensitive to oxygen injury. Proceedings of the National Academy of Sciences of the United States of America *104*, 17010-17015.

Neumann, A.A., Watson, C.M., Noble, J.R., Pickett, H.A., Tam, P.P., and Reddel, R.R. (2013). Alternative lengthening of telomeres in normal mammalian somatic cells. Genes Dev *27*, 18-23.

Nguyen, D., Grenier St-Sauveur, V., Bergeron, D., Dupuis-Sandoval, F., Scott, M.S., and Bachand, F. (2015). A Polyadenylation-Dependent 3' End Maturation Pathway Is Required for the Synthesis of the Human Telomerase RNA. Cell reports *13*, 2244-2257.

Nilsen, H., Haushalter, K.A., Robins, P., Barnes, D.E., Verdine, G.L., and Lindahl, T. (2001). Excision of deaminated cytosine from the vertebrate genome: role of the SMUG1 uracil-DNA glycosylase. The EMBO journal *20*, 4278-4286.

Ogrodnik, M., Miwa, S., Tchkonia, T., Tiniakos, D., Wilson, C.L., Lahat, A., Day, C.P., Burt, A., Palmer, A., Anstee, Q.M.*, et al.* (2017). Cellular senescence drives age-dependent hepatic steatosis. Nature communications *8*, 15691.

Olsen, M.B., Hildrestrand, G.A., Scheffler, K., Vinge, L.E., Alfsnes, K., Palibrk, V., Wang, J., Neurauter, C.G., Luna, L., Johansen, J.*, et al.* (2017). NEIL3-Dependent Regulation of Cardiac Fibroblast Proliferation Prevents Myocardial Rupture. Cell reports *18*, 82-92.

Opresko, P.L., Fan, J., Danzy, S., Wilson, D.M., 3rd, and Bohr, V.A. (2005). Oxidative damage in telomeric DNA disrupts recognition by TRF1 and TRF2. Nucleic acids research *33*, 1230-1239.

Opresko, P.L., Otterlei, M., Graakjaer, J., Bruheim, P., Dawut, L., Kolvraa, S., May, A., Seidman, M.M., and Bohr, V.A. (2004). The Werner syndrome helicase and exonuclease cooperate to resolve telomeric D loops in a manner regulated by TRF1 and TRF2. Molecular cell *14*, 763-774.

Pedersen, H.K., Gudmundsdottir, V., Pedersen, M.K., Brorsson, C., Brunak, S., and Gupta, R. (2016). Ranking factors involved in diabetes remission after bariatric surgery using machine-learning integrating clinical and genomic biomarkers. NPJ Genom Med *1*, 16035.

Poch, E., Carbonell, P., Franco, S., Diez-Juan, A., Blasco, M.A., and Andres, V. (2004). Short telomeres protect from diet-induced atherosclerosis in apolipoprotein E-null mice. FASEB

journal : official publication of the Federation of American Societies for Experimental Biology *18*, 418-420.

Poirier, M.C. (2012). Chemical-induced DNA damage and human cancer risk. Discov Med *14*, 283-288.

Prasad, R., Shock, D.D., Beard, W.A., and Wilson, S.H. (2010). Substrate channeling in mammalian base excision repair pathways: passing the baton. The Journal of biological chemistry *285*, 40479-40488.

Rada, C., Williams, G.T., Nilsen, H., Barnes, D.E., Lindahl, T., and Neuberger, M.S. (2002). Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice. Curr Biol *12*, 1748-1755.

Ranganathan, V., Heine, W.F., Ciccone, D.N., Rudolph, K.L., Wu, X., Chang, S., Hai, H., Ahearn, I.M., Livingston, D.M., Resnick, I.*, et al.* (2001). Rescue of a telomere length defect of Nijmegen breakage syndrome cells requires NBS and telomerase catalytic subunit. Curr Biol *11*, 962-966.

Rangaswamy, S., Pandey, A., Mitra, S., and Hegde, M.L. (2017). Pre-Replicative Repair of Oxidized Bases Maintains Fidelity in Mammalian Genomes: The Cowcatcher Role of NEIL1 DNA Glycosylase. Genes (Basel) *8*.

Ray, S., Bandaria, J.N., Qureshi, M.H., Yildiz, A., and Balci, H. (2014). G-quadruplex formation in telomeres enhances POT1/TPP1 protection against RPA binding. Proceedings of the National Academy of Sciences of the United States of America *111*, 2990-2995.

Regnell, C.E., Hildrestrand, G.A., Sejersted, Y., Medin, T., Moldestad, O., Rolseth, V., Krokeide, S.Z., Suganthan, R., Luna, L., Bjoras, M.*, et al.* (2012). Hippocampal adult neurogenesis is maintained by Neil3-dependent repair of oxidative DNA lesions in neural progenitor cells. Cell reports *2*, 503-510.

Reis, A., and Hermanson, O. (2012). The DNA glycosylases OGG1 and NEIL3 influence differentiation potential, proliferation, and senescence-associated signs in neural stem cells. Biochemical and biophysical research communications *423*, 621-626.

Rhee, D.B., Ghosh, A., Lu, J., Bohr, V.A., and Liu, Y. (2011). Factors that influence telomeric oxidative base damage and repair by DNA glycosylase OGG1. DNA repair *10*, 34-44.

Ribes-Zamora, A., Indiviglio, S.M., Mihalek, I., Williams, C.L., and Bertuch, A.A. (2013). TRF2 interaction with Ku heterotetramerization interface gives insight into c-NHEJ prevention at human telomeres. Cell reports *5*, 194-206.

Roake, C.M., Chen, L., Chakravarthy, A.L., Ferrell, J.E., Jr., Raffa, G.D., and Artandi, S.E. (2019). Disruption of Telomerase RNA Maturation Kinetics Precipitates Disease. Molecular cell *74*, 688-700.e683.

Rudolph, K.L., Chang, S., Lee, H.W., Blasco, M., Gottlieb, G.J., Greider, C., and DePinho, R.A. (1999). Longevity, stress response, and cancer in aging telomerase-deficient mice. Cell *96*, 701-712.

Rudolph, K.L., Chang, S., Millard, M., Schreiber-Agus, N., and DePinho, R.A. (2000). Inhibition of experimental liver cirrhosis in mice by telomerase gene delivery. Science (New York, NY) *287*, 1253-1258.

Ruggero, D., Grisendi, S., Piazza, F., Rego, E., Mari, F., Rao, P.H., Cordon-Cardo, C., and Pandolfi, P.P. (2003). Dyskeratosis congenita and cancer in mice deficient in ribosomal RNA modification. Science (New York, NY) *299*, 259-262.

Saito, I., Takayama, M., Sugiyama, H., Nakatani, K., Tsuchida, A., and Yamamoto, M. (1995). Photoinduced DNA Cleavage via Electron Transfer: Demonstration That Guanine Residues Located 5' to Guanine Are the Most Electron-Donating Sites. J Am Chem Soc *117*, 6406-6407.

Sakellariou, D., Chiourea, M., Raftopoulou, C., and Gagos, S. (2013). Alternative lengthening of telomeres: recurrent cytogenetic aberrations and chromosome stability under extreme telomere dysfunction. Neoplasia *15*, 1301-1313.

Salpea, K.D., and Humphries, S.E. (2010). Telomere length in atherosclerosis and diabetes. Atherosclerosis *209*, 35-38.

Sapieha, P., and Mallette, F.A. (2018). Cellular Senescence in Postmitotic Cells: Beyond Growth Arrest. Trends Cell Biol *28*, 595-607.

Schmidt, J.C., Dalby, A.B., and Cech, T.R. (2014). Identification of human TERT elements necessary for telomerase recruitment to telomeres. Elife *3*.

Sejersted, Y., Hildrestrand, G.A., Kunke, D., Rolseth, V., Krokeide, S.Z., Neurauter, C.G., Suganthan, R., Atneosen-Asegg, M., Fleming, A.M., Saugstad, O.D.*, et al.* (2011). Endonuclease VIII-like 3 (Neil3) DNA glycosylase promotes neurogenesis induced by hypoxia-ischemia. Proceedings of the National Academy of Sciences of the United States of America *108*, 18802-18807.

Seki, S., Kitada, T., Yamada, T., Sakaguchi, H., Nakatani, K., and Wakasa, K. (2002). In situ detection of lipid peroxidation and oxidative DNA damage in non-alcoholic fatty liver diseases. J Hepatol *37*, 56-62.

Sfeir, A., and de Lange, T. (2012). Removal of shelterin reveals the telomere end-protection problem. Science (New York, NY) *336*, 593-597.

Sfeir, A., Kosiyatrakul, S.T., Hockemeyer, D., MacRae, S.L., Karlseder, J., Schildkraut, C.L., and de Lange, T. (2009). Mammalian telomeres resemble fragile sites and require TRF1 for efficient replication. Cell *138*, 90-103.

Shay, J.W. (1997). Molecular pathogenesis of aging and cancer: are telomeres and telomerase the connection? J Clin Pathol *50*, 799-800.

Shay, J.W., and Bacchetti, S. (1997). A survey of telomerase activity in human cancer. Eur J Cancer *33*, 787-791.

Shen, J.C., Rideout, W.M., 3rd, and Jones, P.A. (1994). The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA. Nucleic acids research *22*, 972-976.

Shukla, S., Schmidt, J.C., Goldfarb, K.C., Cech, T.R., and Parker, R. (2016). Inhibition of telomerase RNA decay rescues telomerase deficiency caused by dyskerin or PARN defects. Nature structural & molecular biology *23*, 286-292.

Simone, L.E., and Keene, J.D. (2013). Mechanisms coordinating ELAV/Hu mRNA regulons. Curr Opin Genet Dev *23*, 35-43.

Singh, M., Wang, Z., Cascio, D., and Feigon, J. (2015). Structure and interactions of the CS domain of human H/ACA RNP assembly protein Shq1. J Mol Biol *427*, 807-823.

Skarpengland, T., Holm, S., Scheffler, K., Gregersen, I., Dahl, T.B., Suganthan, R., Segers, F.M., Ostlie, I., Otten, J.J., Luna, L.*, et al.* (2016). Neil3-dependent base excision repair regulates lipid metabolism and prevents atherosclerosis in Apoe-deficient mice. Scientific reports *6*, 28337.

Skarpengland, T., Laugsand, L.E., Janszky, I., Luna, L., Halvorsen, B., Platou, C.G., Wang, W., Vatten, L.J., Damas, J.K., Aukrust, P.*, et al.* (2015). Genetic variants in the DNA repair gene NEIL3 and the risk of myocardial infarction in a nested case-control study. The HUNT Study. DNA repair *28*, 21-27.

Squires, J.E., Patel, H.R., Nousch, M., Sibbritt, T., Humphreys, D.T., Parker, B.J., Suter, C.M., and Preiss, T. (2012). Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. Nucleic acids research *40*, 5023-5033.

Stabile, E., Zhou, Y.F., Saji, M., Castagna, M., Shou, M., Kinnaird, T.D., Baffour, R., Ringel, M.D., Epstein, S.E., and Fuchs, S. (2003). Akt controls vascular smooth muscle cell proliferation in vitro and in vivo by delaying G1/S exit. Circ Res *93*, 1059-1065.

Stiff, T., Walker, S.A., Cerosaletti, K., Goodarzi, A.A., Petermann, E., Concannon, P., O'Driscoll, M., and Jeggo, P.A. (2006). ATR-dependent phosphorylation and activation of ATM in response to UV treatment or replication fork stalling. The EMBO journal *25*, 5775-5782.

Svilar, D., Goellner, E.M., Almeida, K.H., and Sobol, R.W. (2011). Base excision repair and lesion-dependent subpathways for repair of oxidative DNA damage. Antioxid Redox Signal *14*, 2491-2507.

Takai, H., Jenkinson, E., Kabir, S., Babul-Hirji, R., Najm-Tehrani, N., Chitayat, D.A., Crow, Y.J., and de Lange, T. (2016). A POT1 mutation implicates defective telomere end fill-in and telomere truncations in Coats plus. Genes Dev *30*, 812-826.

Takai, H., Smogorzewska, A., and de Lange, T. (2003). DNA damage foci at dysfunctional telomeres. Curr Biol *13*, 1549-1556.

Takubo, K., Nakamura, K., Izumiyama, N., Furugori, E., Sawabe, M., Arai, T., Esaki, Y., Mafune, K., Kammori, M., Fujiwara, M*., et al.* (2000). Telomere shortening with aging in human liver. J Gerontol A Biol Sci Med Sci *55*, B533-536.

Tang, H., Wang, H., Cheng, X., Fan, X., Yang, F., Zhang, M., Chen, Y., Tian, Y., Liu, C., Shao, D*., et al.* (2018). HuR regulates telomerase activity through TERC methylation. Nature communications *9*, 2213.

Theruvathu, J.A., Darwanto, A., Hsu, C.W., and Sowers, L.C. (2014). The effect of Pot1 binding on the repair of thymine analogs in a telomeric DNA sequence. Nucleic acids research *42*, 9063-9073.

Tichy, E.D., Liang, L., Deng, L., Tischfield, J., Schwemberger, S., Babcock, G., and Stambrook, P.J. (2011). Mismatch and base excision repair proficiency in murine embryonic stem cells. DNA repair *10*, 445-451.

Todd, A.K., Johnston, M., and Neidle, S. (2005). Highly prevalent putative quadruplex sequence motifs in human DNA. Nucleic acids research *33*, 2901-2907.

Tomlinson, R.L., Li, J., Culp, B.R., Terns, R.M., and Terns, M.P. (2010). A Cajal body-independent pathway for telomerase trafficking in mice. Experimental cell research *316*, 2797-2809.

Tomlinson, R.L., Ziegler, T.D., Supakorndej, T., Terns, R.M., and Terns, M.P. (2006). Cell cycle-regulated trafficking of human telomerase to telomeres. Mol Biol Cell *17*, 955-965.

Tong, A.S., Stern, J.L., Sfeir, A., Kartawinata, M., de Lange, T., Zhu, X.D., and Bryan, T.M. (2015). ATM and ATR Signaling Regulate the Recruitment of Human Telomerase to Telomeres. Cell reports *13*, 1633-1646.

Torisu, K., Tsuchimoto, D., Ohnishi, Y., and Nakabeppu, Y. (2005). Hematopoietic tissue-specific expression of mouse Neil3 for endonuclease VIII-like protein. J Biochem *138*, 763-772.

Tseng, C.K., Wang, H.F., Burns, A.M., Schroeder, M.R., Gaspari, M., and Baumann, P. (2015). Human Telomerase RNA Processing and Quality Control. Cell reports *13*, 2232-2243.

Tseng, C.K., Wang, H.F., Schroeder, M.R., and Baumann, P. (2018). The H/ACA complex disrupts triplex in hTR precursor to permit processing by RRP6 and PARN. Nature communications *9*, 5430.

Vallabhaneni, H., Zhou, F., Maul, R.W., Sarkar, J., Yin, J., Lei, M., Harrington, L., Gearhart, P.J., and Liu, Y. (2015). Defective repair of uracil causes telomere defects in mouse hematopoietic cells. The Journal of biological chemistry *290*, 5502-5511.

Vannier, J.B., Pavicic-Kaltenbrunner, V., Petalcorin, M.I., Ding, H., and Boulton, S.J. (2012). RTEL1 dismantles T loops and counteracts telomeric G4-DNA to maintain telomere integrity. Cell *149*, 795-806.

Venteicher, A.S., Abreu, E.B., Meng, Z., McCann, K.E., Terns, R.M., Veenstra, T.D., Terns, M.P., and Artandi, S.E. (2009). A human telomerase holoenzyme protein required for Cajal body localization and telomere synthesis. Science (New York, NY) *323*, 644-648.

Venteicher, A.S., and Artandi, S.E. (2009). TCAB1: driving telomerase to Cajal bodies. Cell Cycle *8*, 1329-1331.

Vera, E., Canela, A., Fraga, M.F., Esteller, M., and Blasco, M.A. (2008). Epigenetic regulation of telomeres in human cancer. Oncogene *27*, 6817-6833.

Vorlickova, M., Tomasko, M., Sagi, A.J., Bednarova, K., and Sagi, J. (2012). 8-oxoguanine in a quadruplex of the human telomere DNA sequence. Febs j *279*, 29-39.

Walbott, H., Machado-Pinilla, R., Liger, D., Blaud, M., Rety, S., Grozdanov, P.N., Godin, K., van Tilbeurgh, H., Varani, G., Meier, U.T.*, et al.* (2011). The H/ACA RNP assembly factor SHQ1 functions as an RNA mimic. Genes Dev *25*, 2398-2408.

Wallace, S.S. (2014). Base excision repair: a critical player in many games. DNA repair *19*, 14-26.

Wang, C., and Meier, U.T. (2004). Architecture and assembly of mammalian H/ACA small nucleolar and telomerase ribonucleoproteins. The EMBO journal *23*, 1857-1867.

Wang, F., Yin, Y., Ye, X., Liu, K., Zhu, H., Wang, L., Chiourea, M., Okuka, M., Ji, G., Dan, J.*, et al.* (2012). Molecular insights into the heterogeneity of telomere reprogramming in induced pluripotent stem cells. Cell Res *22*, 757-768.

Wang, H., Nora, G.J., Ghodke, H., and Opresko, P.L. (2011). Single molecule studies of physiologically relevant telomeric tails reveal POT1 mechanism for promoting G-quadruplex unfolding. The Journal of biological chemistry *286*, 7479-7489.

Wang, J., Uryga, A.K., Reinhold, J., Figg, N., Baker, L., Finigan, A., Gray, K., Kumar, S., Clarke, M., and Bennett, M. (2015). Vascular Smooth Muscle Cell Senescence Promotes Atherosclerosis and Features of Plaque Vulnerability. Circulation *132*, 1909-1919.

Wang, T., Wu, H., Li, Y., Szulwach, K.E., Lin, L., Li, X., Chen, I.P., Goldlust, I.S., Chamberlain, S.J., Dodd, A.*, et al.* (2013). Subtelomeric hotspots of aberrant 5-hydroxymethylcytosine-mediated epigenetic modifications during reprogramming to pluripotency. Nat Cell Biol *15*, 700-711.

Wang, X., and He, C. (2014). Reading RNA methylation codes through methyl-specific binding proteins. RNA biology *11*, 669-672.

Wang, Z., Rhee, D.B., Lu, J., Bohr, C.T., Zhou, F., Vallabhaneni, H., de Souza-Pinto, N.C., and Liu, Y. (2010). Characterization of oxidative guanine damage and repair in mammalian telomeres. PLoS genetics *6*, e1000951.

Wilson, S.H., and Kunkel, T.A. (2000). Passing the baton in base excision repair. Nat Struct Biol *7*, 176-178.

Wright, W.E., Brasiskyte, D., Piatyszek, M.A., and Shay, J.W. (1996). Experimental elongation of telomeres extends the lifespan of immortal x normal cell hybrids. The EMBO journal *15*, 1734-1741.

Wright, W.E., and Shay, J.W. (2000). Telomere dynamics in cancer progression and prevention: fundamental differences in human and mouse telomere biology. Nat Med *6*, 849-851.

Wu, L., Multani, A.S., He, H., Cosme-Blanco, W., Deng, Y., Deng, J.M., Bachilo, O., Pathak, S., Tahara, H., Bailey, S.M., *et al.* (2006). Pot1 deficiency initiates DNA damage checkpoint activation and aberrant homologous recombination at telomeres. Cell *126*, 49-62.

Wynford-Thomas, D., and Kipling, D. (1997). The end-replication problem. Nature *389*, 551-551.

Yutzey, K.E. (2017). Cardiomyocyte Proliferation: Teaching an Old Dogma New Tricks. Circulation research *120*, 627-629.

Zalzman, M., Falco, G., Sharova, L.V., Nishiyama, A., Thomas, M., Lee, S.L., Stagg, C.A., Hoang, H.G., Yang, H.T., Indig, F.E., *et al.* (2010). Zscan4 regulates telomere elongation and genomic stability in ES cells. Nature *464*, 858-863.

Zhang, L.F., Huynh, K.D., and Lee, J.T. (2007). Perinucleolar targeting of the inactive X during S phase: evidence for a role in the maintenance of silencing. Cell *129*, 693-706.

Zhang, Q., Kim, N.K., and Feigon, J. (2011). Architecture of human telomerase RNA. Proceedings of the National Academy of Sciences of the United States of America *108*, 20325-20332.

Zhao, J.Q., Hoare, S.F., McFarlane, R., Muir, S., Parkinson, E.K., Black, D.M., and Keith, W.N. (1998). Cloning and characterization of human and mouse telomerase RNA gene promoter sequences. Oncogene *16*, 1345-1350.

Zhong, F.L., Batista, L.F., Freund, A., Pech, M.F., Venteicher, A.S., and Artandi, S.E. (2012). TPP1 OB-fold domain controls telomere maintenance by recruiting telomerase to chromosome ends. Cell *150*, 481-494.

Zhou, J., Chan, J., Lambele, M., Yusufzai, T., Stumpff, J., Opresko, P.L., Thali, M., and Wallace, S.S. (2017). NEIL3 Repairs Telomere Damage during S Phase to Secure Chromosome Segregation at Mitosis. Cell reports *20*, 2044-2056.

Zhou, J., Fleming, A.M., Averill, A.M., Burrows, C.J., and Wallace, S.S. (2015). The NEIL glycosylases remove oxidized guanine lesions from telomeric and promoter quadruplex DNA structures. Nucleic acids research *43*, 4039-4054.

Zhou, J., Liu, M., Fleming, A.M., Burrows, C.J., and Wallace, S.S. (2013). Neil3 and NEIL1 DNA glycosylases remove oxidative damages from quadruplex DNA and exhibit preferences for lesions in the telomeric sequence context. The Journal of biological chemistry *288*, 27263-27272.

Zhu, X.D., Kuster, B., Mann, M., Petrini, J.H., and de Lange, T. (2000). Cell-cycle-regulated association of RAD50/MRE11/NBS1 with TRF2 and human telomeres. Nature genetics *25*, 347-352.

Zhu, Y., Tomlinson, R.L., Lukowiak, A.A., Terns, R.M., and Terns, M.P. (2004). Telomerase RNA accumulates in Cajal bodies in human cancer cells. Mol Biol Cell *15*, 81-90.

Zielke, S., and Bodnar, A. (2010). Telomeres and telomerase activity in scleractinian corals and Symbiodinium spp. Biol Bull *218*, 113-121.

Özer, Ö., and Hickson, I.D. (2018). Pathways for maintenance of telomeres and common fragile sites during DNA replication stress. Open Biol *8*, 180018.

Appendix

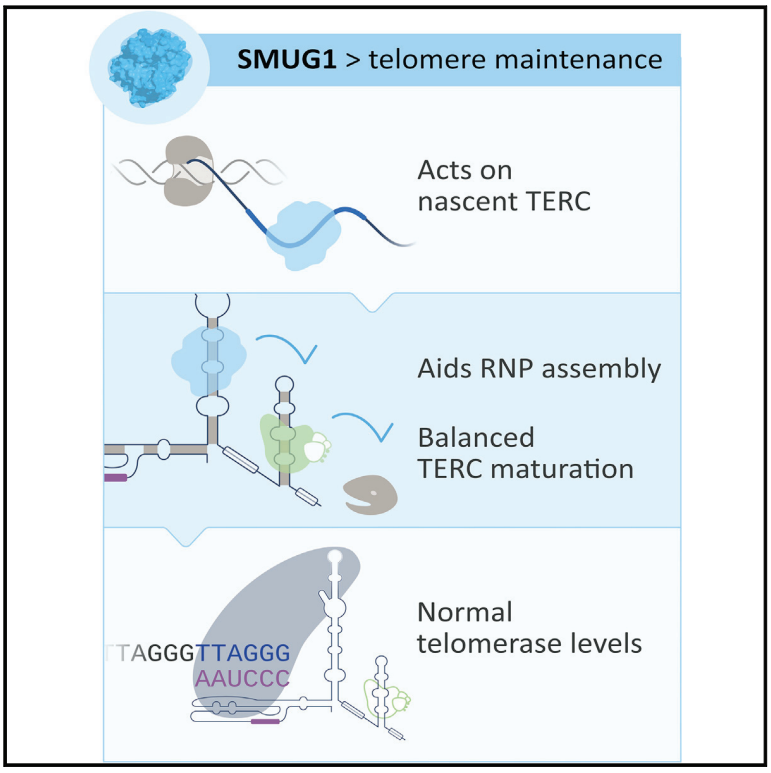
Paper I, II and III

Article

Cell Reports

SMUG1 Promotes Telomere Maintenance through Telomerase RNA Processing

Graphical Abstract



Authors

Penelope Kroustallaki, Lisa Lirussi, Sergio Carracedo, ..., Pål Sætrom, Sarantis Gagos, Hilde Nilsen

Correspondence

hilde.nilsen@medisin.uio.no

In Brief

Kroustallaki et al. show that the single-strand-selective uracil-DNA glycosylase (SMUG1) functions in telomere maintenance, by removing modified bases from telomeric DNA and also by regulating modified bases in the telomerase RNA component (hTERC). SMUG1-knockout cells accumulate hTERC containing modified bases that interfere with binding of DKC1. Consequently, SMUG1-knockout cells and mice exhibit telomere maintenance defects.

Highlights

- The SMUG1-DNA glycosylase promotes maturation of the telomeric DNA component, hTERC
- SMUG1 regulates the presence of base modifications in hTERC
- SMUG1 promotes DKC1 binding and stability of hTERC
- hTERC levels limit telomerase activity in human SMUG1-knockout cells

Kroustallaki et al., 2019, Cell Reports *28*, 1690–1702
August 13, 2019
https://doi.org/10.1016/j.celrep.2019.07.040

CellPress

OPEN ACCESS
CellPress

Cell Reports
Article

SMUG1 Promotes Telomere Maintenance through Telomerase RNA Processing

Penelope Kroustallaki,[1,2,8] Lisa Lirussi,[1,2,8] Sergio Carracedo,[1,2] Panpan You,[1,2] Q. Ying Esbensen,[1,2] Alexandra Götz,[1,2] Laure Jobert,[1,9] Lene Alsøe,[1,2] Pål Sætrom,[3,4,5,6] Sarantis Gagos,[7] and Hilde Nilsen[1,2,10,*]

[1]Department of Clinical Molecular Biology, University of Oslo, 0318 Oslo, Norway
[2]Department of Clinical Molecular Biology (EpiGen), Akershus University Hospital, 1478 Lørenskog, Norway
[3]Department of Clinical and Molecular Medicine, Norwegian University of Science and Technology, NTNU, 7491 Trondheim, Norway
[4]Department of Computer Science, Norwegian University of Science and Technology, NTNU, 7491 Trondheim, Norway
[5]Bioinformatics Core Facility-BioCore, Norwegian University of Science and Technology, NTNU, 7491 Trondheim, Norway
[6]K.G. Jebsen Center for Genetic Epidemiology, Norwegian University of Science and Technology, NTNU, 7491 Trondheim, Norway
[7]Center of Clinical, Experimental Surgery & Translational Research, Biomedical Research Foundation, Academy of Athens, Athens, Greece
[8]These authors contributed equally
[9]Present address: LifeTechnologies AS, Ullernschauseen 52, 0379 Oslo, Norway
[10]Lead Contact
*Correspondence: hilde.nilsen@medisin.uio.no
https://doi.org/10.1016/j.celrep.2019.07.040

SUMMARY

Telomerase biogenesis is a complex process where several steps remain poorly understood. Single-strand-selective uracil-DNA glycosylase (SMUG1) associates with the DKC1-containing H/ACA ribonucleoprotein complex, which is essential for telomerase biogenesis. Herein, we show that SMUG1 interacts with the telomeric RNA component (*hTERC*) and is required for co-transcriptional processing of the nascent transcript into mature *hTERC*. We demonstrate that SMUG1 regulates the presence of base modifications in *hTERC*, in a region between the CR4/CR5 domain and the H box. Increased levels of *hTERC* base modifications are accompanied by reduced DKC1 binding. Loss of SMUG1 leads to an imbalance between mature *hTERC* and its processing intermediates, leading to the accumulation of 3′-polyadenylated and 3′-extended intermediates that are degraded in an EXOSC10-independent RNA degradation pathway. Consequently, SMUG1-deprived cells exhibit telomerase deficiency, leading to impaired bone marrow proliferation in *Smug1*-knockout mice.

INTRODUCTION

Telomerase is a specialized ribonucleoprotein (RNP) complex that extends telomeric repeats at the ends of chromosomes (de Lange, 2005; Morin, 1989). The telomerase holoenzyme consists of three main subunits: the telomerase reverse transcriptase (hTERT), the telomerase RNA component (*hTERC*), and the dyskerin complex (DKC1, NHP2, NOP1, and GAR1) (Egan and Collins, 2012; Schmidt and Cech, 2015). *hTERC* is a highly structured non-coding RNA that carries the complementary template of the telomeric repeat sequence and two H/ACA domains that bind to dyskerin (Egan and Collins, 2012). The *hTERC*/dyskerin RNP complex and hTERT associate in both nucleoli and Cajal bodies (CBs) during S phase, suggesting that both these subnuclear structures are involved in the biogenesis and trafficking of the telomerase complex (Lee et al., 2014; MacNeil et al., 2016). *hTERC* biogenesis is a multistep process. First, *hTERC* is transcribed by RNA polymerase II (Pol II) to a primary transcript that can extend several hundred nucleotides downstream of the *hTERC* gene body (Nguyen et al., 2015; Tseng et al., 2015). The H/ACA complex is co-transcriptionally assembled and may mediate *hTERC* transcriptional termination and, thus, determine the length of the 3′ extension (MacNeil et al., 2016). Subsequent end-processing steps, leading to the formation of the 451-nt-long mature *hTERC*, involve polyadenylation by the Trf4/5-Air1/2-(TRAMP) complex, or the PARN/PABPN1 machinery, and processing by the nuclear exosome-targeting (NEXT) complex (Nguyen et al., 2015; Tseng et al., 2015). The balance between maturation and exosomal degradation determines the level of mature *hTERC* (MacNeil et al., 2016; Zinder and Lima, 2017). Other RNA degradation pathways might also be involved in the removal of *hTERC* intermediates (Schmidt and Cech, 2015; Zinder and Lima, 2017), and the detailed molecular mechanisms of *hTERC* maturation are not fully understood.

We recently demonstrated that the single-strand-selective uracil (SMUG1)-DNA glycosylase interacts and co-localizes with the pseudouridine synthase DKC1 (Jobert et al., 2013). DKC1 is involved in the biogenesis and maturation of several RNA classes, such as rRNA (Ge et al., 2010). SMUG1 associates with the 47S rRNA precursor, which is a major substrate of DKC1, and loss of SMUG1 leads to rRNA processing defects and accumulation of 5-hydroxymethyluridine (hmU) in rRNA. Thus, in addition to its function in DNA base excision repair (BER), SMUG1 acts in rRNA quality control (Jobert et al., 2013).

As DKC1 functions both to support telomerase biogenesis in nucleoli and CBs as a structural component of the telomerase holoenzyme (Mitchell et al., 1999; Venteicher et al., 2009; Lee et al., 2014), we tested a possible role for SMUG1 in telomere maintenance. We show that SMUG1 is present in CBs and observed a

significant decrease in telomere length in human SMUG1-knockout (KO) cells due to insufficient levels of mature *hTERC* to support telomerase activity. *hTERC* accumulated base modifications between the CR4/CR5 domain and the H box, a region important for DKC1 binding. Consistently, DCK1 was bound less efficiently in SMUG1-KO cells, leading to *hTERC* degradation. We conclude that SMUG1 promotes *hTERC* stability by regulating the presence of modified bases to allow binding of DKC1.

RESULTS

SMUG1 Influences DKC1 Localization

We previously observed that overexpression of a SMUG1 mutant unable to interact with DKC1 (E29R/E33R) affected DKC1 localization in HeLa cells (Jobert et al., 2013). To confirm that disruption of the SMUG1/DKC1 interaction surface perturbs proper localization of DKC1, we repeated these experiments in *Smug1*$^{-/-}$ mouse embryonic fibroblasts (MEFs) (Alsøe et al., 2017). To exclude any bias originating from possible small differences in cell-cycle distribution, we scored the ring-shaped structures formed by DKC1 (DKC1 circles) during S phase (Lee et al., 2014). In *Smug1*$^{-/-}$ MEFs we observed fewer DKC1 circles (white arrows) and the appearance of dense nucleolar bodies (yellow arrows) (Figures 1A and 1B). The number of DKC1 circles could not be fully restored in cells stably expressing neither wild-type SMUG1 nor a mutant that does not have DNA-glycosylase activity on RNA substrates (H241L) (Figures 1A, 1B, and S1A). Confirming our previous data, expression of the DKC1-binding mutant exacerbated the phenotype (Figures 1A and 1B). SMUG1/DKC1 interaction affects the DKC1 distribution pattern in nucleoli, which has been suggested to be the site for the biogenesis of the telomerase holoenzyme (Lee et al., 2014) before transport to the CBs (MacNeil et al., 2016). Interestingly, *in situ* proximity ligation assay (PLA) showed SMUG1-Coilin interaction in the nucleus and CBs (Figures 1C and S1B). Taken together, SMUG1 influences DKC1 organization in the nucleoli, possibly suggesting a role of SMUG1 in telomerase biogenesis.

Smug1$^{-/-}$ Mice Exhibit Telomere Maintenance Defects

As DKC1 is essential for telomerase biogenesis and regulation of telomere length (Mitchell et al., 1999), we asked whether SMUG1 functions in telomere maintenance. Telomere chromatin immunoprecipitation (TeloChIP) showed that SMUG1 associated with telomeric chromatin (Figure 2A). SMUG1 could not be detected at telomeres by telomere-specific fluorescent *in situ* hybridization (FISH) (Figure S1C), suggesting that SMUG1 transiently associates with telomeric DNA, consistent with the BER function. In tissue harvested from 3- to 4-month-old *Smug1*$^{-/-}$ mice, a significantly reduced average telomere length was found in the heart (62% reduction; $p = 0.017$) but not in the spleen and brain (Figure 2B). To assess whether accumulation of DNA base damage in the form of SMUG1 substrates occurred in telomeric DNA, we digested the genomic DNA with SMUG1 and APE1 prior to telomere length measurements by qPCR. The presence of damaged bases would be expected to reduce the amplification efficiency. In this assay, increased telomeric DNA damage was detected only in heart tissue (Figure 2B). The average telomere length was similar in *Smug1*$^{-/-}$ and isogenic wild-type MEFs (Figure 2C), but telomere-specific FISH revealed high frequencies of fragile telomeres in primary (Figures 2D and 2E) and transformed *Smug1*$^{-/-}$ MEFs (Figure S1D). Strand-specific telomere-FISH probes showed a doubling of fragile telomeres on the leading C-rich strand in primary MEFs (Figure 2E), while the G-rich strand was largely unaffected (Figure S1E). Transformed MEFs had significantly more fragile telomeric signals in both strands (Figure S1D).

The fragile telomere phenotype was even more pronounced in *Smug1*$^{-/-}$ primary bone marrow cells. Both strands showed an increase in fragile telomeres, but the increase was more pronounced on the C-rich strand where close to 4% of the telomeric signals showed fragility on the C-strand, compared to less than 1% in wild-type bone marrow cells (Figure 2F). To assess whether telomere fragility had functional relevance *in vivo*, we measured the colony-forming capacity of primary bone marrow explant cultures (Figure 2G). The colony-forming unit abilities of the erythrocyte (BFU-E) and granulocyte (CFU-G) lineages in *Smug1*$^{-/-}$ bone marrow explant cultures were reduced by 43% and 41%, respectively, in mice born from heterozygous parents (F1). As expected, based on a telomere maintenance phenotype, the colony-forming ability was further reduced in mice born from parents generated through five generations of interbreeding of homozygous *Smug1*-knockout mice (F6). In the F6 generation, the proliferative potentials of the macrophage lineage (CFU-M) and BFU-E were reduced by 40% and 61%, respectively (Figure 2G). Thus, loss of SMUG1 expression in mice led to reduced average telomere length and accumulation of telomere DNA damage in certain tissues. In addition, there was an asymmetric fragile-telomere phenotype affecting, primarily, the C-rich strand and reduced proliferative potential of primary bone marrow cells.

Dramatic Telomere Attrition in SMUG1-KO Human Cells Is Independent of BER Function

As telomere-associated phenotypes might be masked by the long telomeres of mice, we procured a human cell line in which SMUG1 expression was abrogated by a two-nucleotide deletion that introduces a premature stop codon after Asn56 of SMUG1 (Figure 3A). As expected, *SMUG1* transcription was unaffected (Figure 3B), but no SMUG1 protein could be detected using an antibody directed toward an N-terminal epitope (Figure 3C). Thus, the mutation generated a loss-of-function, or extremely hypomorphic, allele. TeloChIP analysis revealed 4-fold enrichment of telomeric DNA in HAP1 SMUG1 wild-type (WT) cells relative to SMUG1-KO cells (Figure 3D), showing that SMUG1 also associates with telomeres in human cells.

Telomeric-FISH signals were barely detectable in SMUG1-KO cells, in stark contrast to the bright signals observed in the parental cell line (Figure 3E). Scoring of fragile telomeres was therefore not possible, but the fraction of telomeric signal-free ends was increased by at least 2-fold in two independent SMUG-KO clones (Figures 3E and S2A). The dramatic telomere attrition was confirmed by telomere restriction fragment (TRF) analysis, which showed that SMUG1-KO telomeres were 2.6 kb long on average, compared to 16.6 kb in isogenic WT cells (Figure 3F). The weak signals detected in lanes loaded with genomic DNA isolated from SMUG1-KO cells, despite equal loading, were also consistent with a reduced fraction of

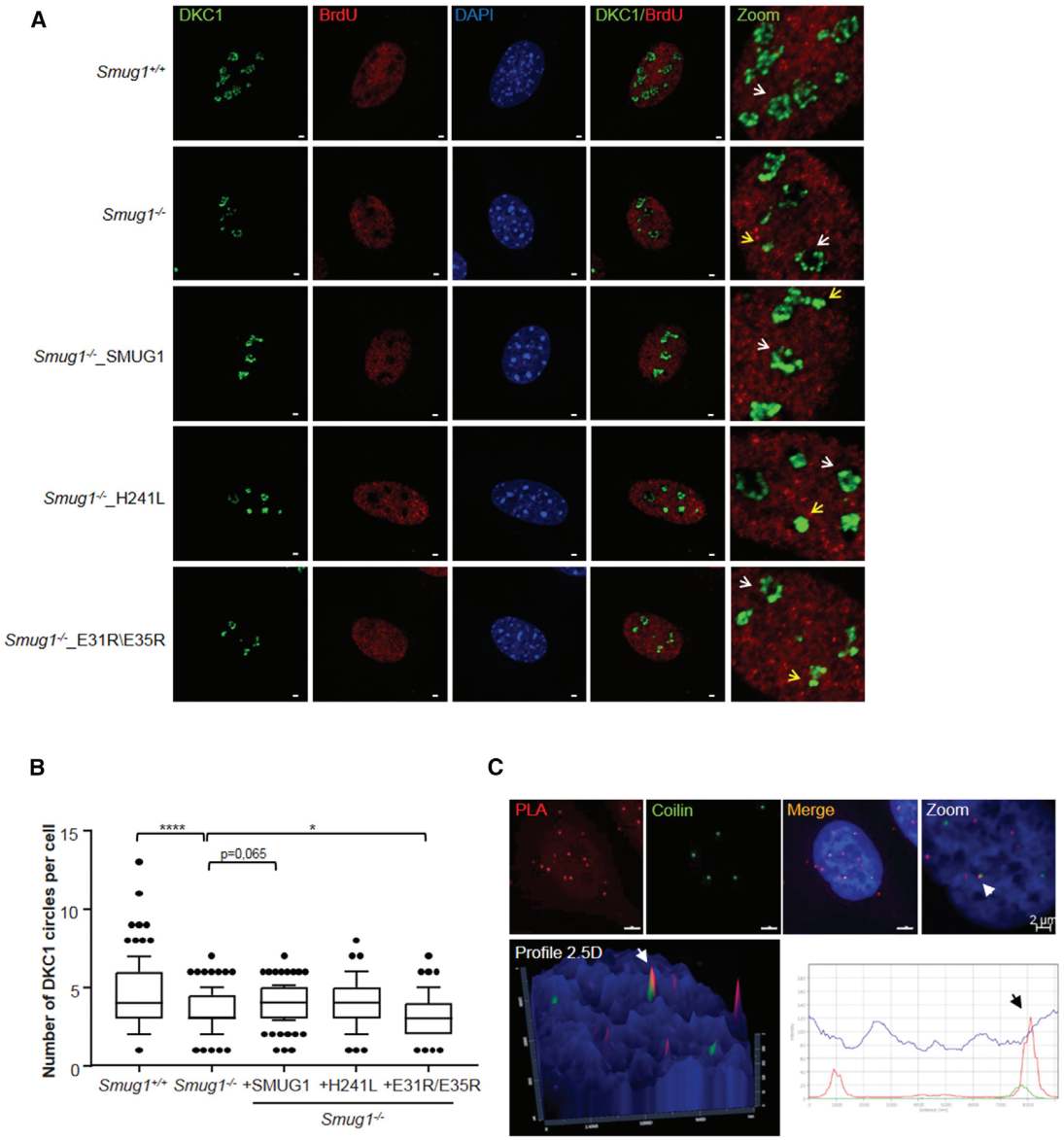


Figure 1. SMUG1 Loss Leads to DKC1 Mislocalization

(A) Localization of DKC1 in *Smug1*$^{+/+}$, *Smug1*$^{-/-}$, and *Smug1*$^{-/-}$ MEFs complemented with wild-type mouse SMUG1, or SMUG1 mutants that do not bind DNA (H241L) or DKC1 (E31R/E35R). Representative immunofluorescence (IF) images of BrdU (red) and DKC1 (green) staining are shown. DNA was labeled with 4′,6-diamidino-2-phenylindole (DAPI; blue). White and yellow arrows indicate DKC1 circles and dense bodies, respectively (scale bars, 2 μm).

(B) Boxplot showing the frequency of ring-shaped structures characteristic for DKC1 in S-phase cells, with whiskers representing the 10th and 90th percentiles; the dark line within the box represents the median. n = 100 cells, *p ≤ 0.05, and ****p ≤ 0.0001 (two-tailed Student's t test).

(C) Proximity ligation assay (PLA) showing FLAG-tagged SMUG1-Coilin interaction (red) in Cajal bodies (arrows). CBs were stained with Coilin (green). Scale bars, 1 μm.

telomeric DNA. To assess whether accumulation of DNA base damage in the form of SMUG1 substrates occurred in telomeric DNA, we added SMUG1 and APE1 to the restriction enzyme cocktail. Base damage present within the telomere restriction fragment would be expected to reduce fragment length. Indeed, the mean telomere length was further reduced in the SMUG1-KO cells (Figure 3F), whereas no change in telomere length was seen in WT cells. Thus, base damage was present in telomeres in the absence of SMUG1, reducing the average fragment length from 2.6 to 1.7 kb (Figure 3F, right).

The dramatic telomere attrition also affected the telomeric association of shelterin proteins. Both TRF1 and TRF2, which bind double-stranded telomeric DNA, exhibited a diffuse staining pattern in SMUG1-KO cells, in addition to the characteristic punctate telomere-specific staining (Figures 3G and S2B). The reduced binding of shelterin components was corroborated by

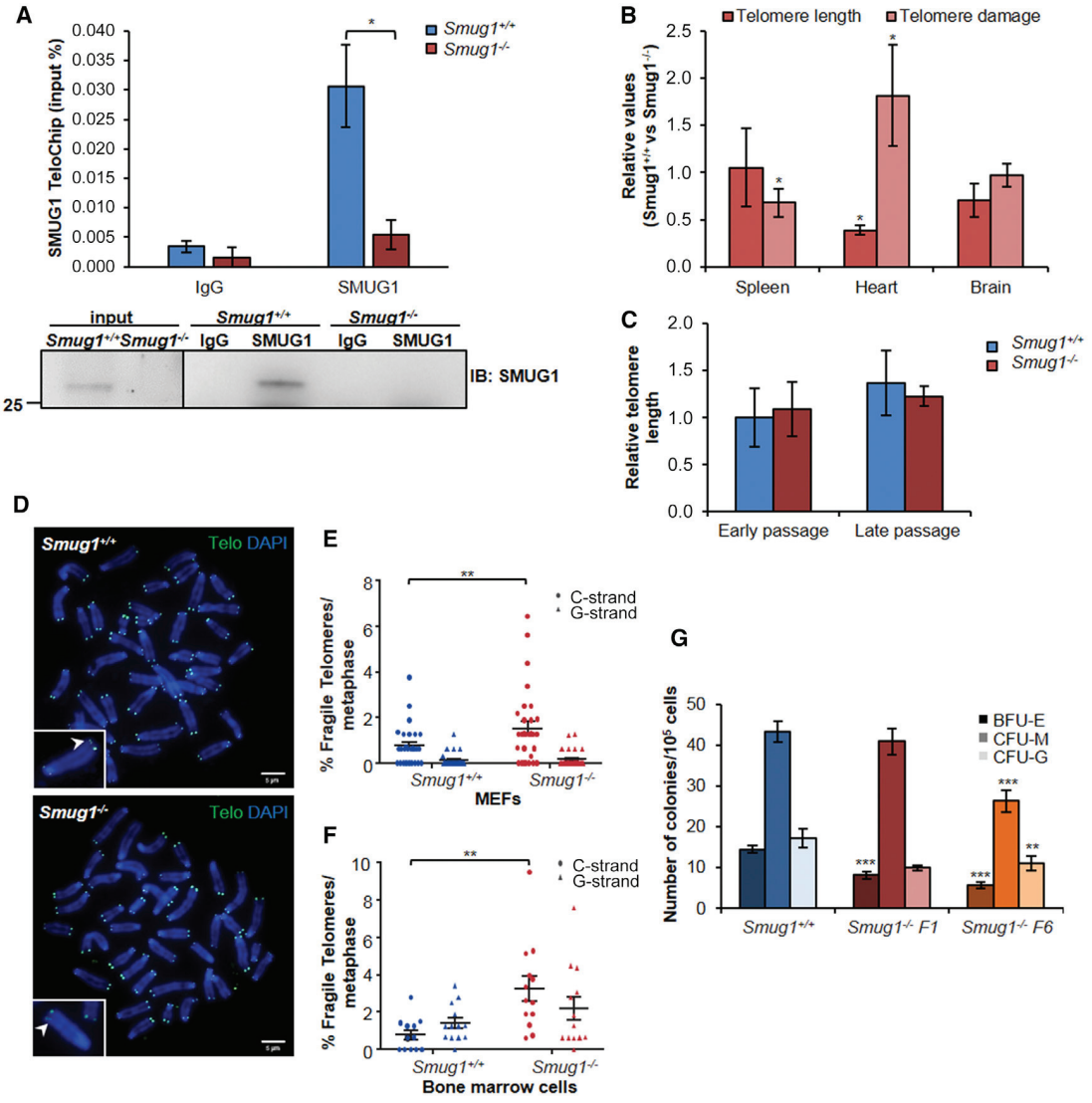


Figure 2. *Smug1*$^{-/-}$ Mice Exhibit Telomere Maintenance Defects

(A) TeloChIP followed by qPCR showing SMUG1 binding to telomere repeats in three independent MEF clones (top). Representative western blot of immunoprecipitation with IgG and SMUG1 antibodies (bottom).

(B) Telomere length and damage quantified in *Smug1*$^{-/-}$ mouse tissues (heart, spleen, and brain) by qPCR. Data are presented as fold change relative to wild-type mice.

(C) Telomere length in *Smug1*$^{+/+}$ and *Smug1*$^{-/-}$ MEF cells at early (0–10) and late (30–40) passages measured by qPCR.

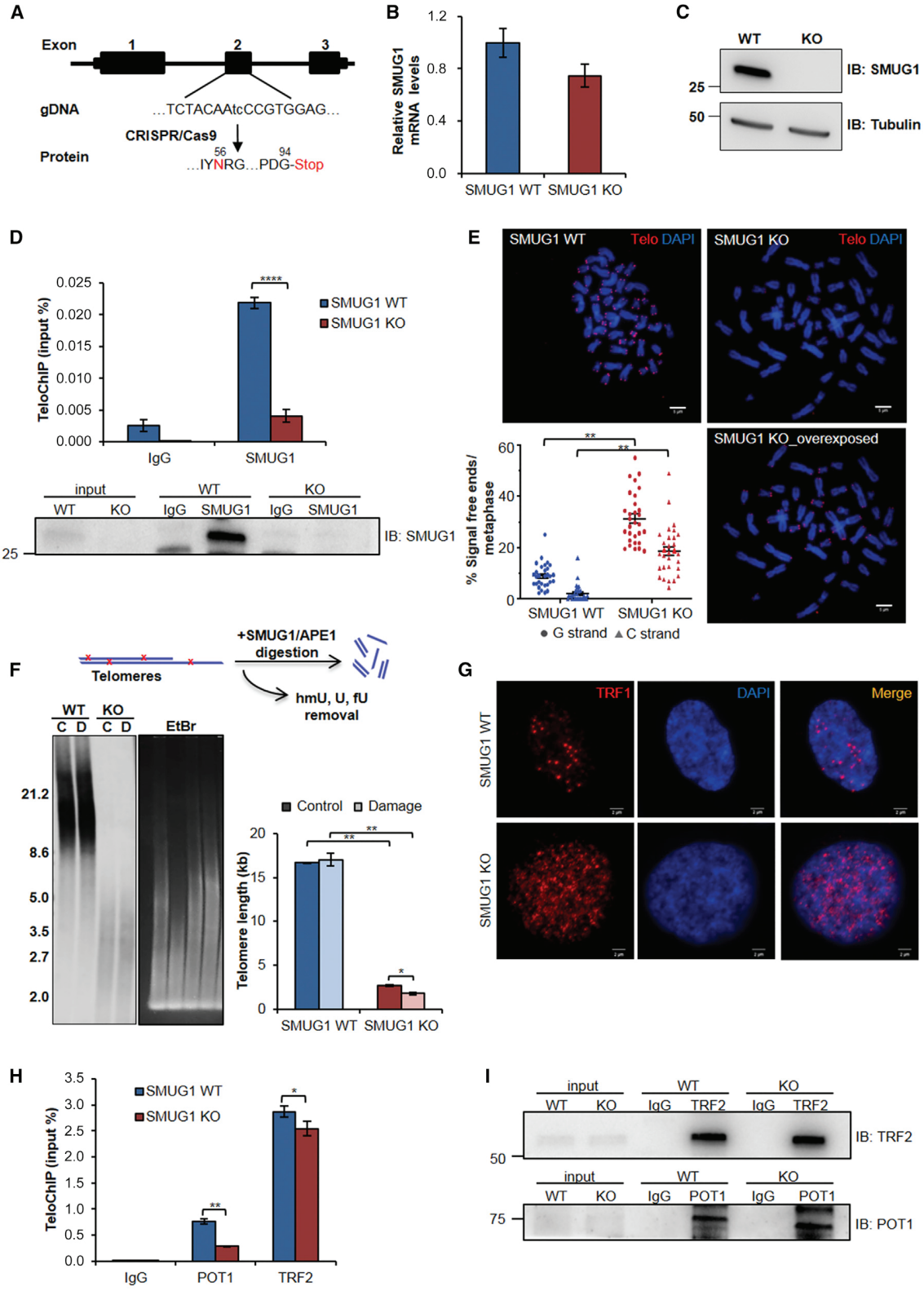
(D) Peptide nucleic acid (PNA)-FISH in MEFs metaphases. Telomeres were hybridized with a telomere-specific probe (Telo, fluorescein isothiocyanate [FITC]) and chromosomes were stained with DAPI.

(E and F) Quantification of fragile telomere ends (FTEs) in primary MEFs (E) and in primary mouse bone marrow cells (circles represent C-strand; triangles represent G-strand) (F).

(G) Colony-forming capacity toward erythrocyte (BFU-E), granulocyte (CFU-G) and macrophage (CFU-M) lineages in bone marrow from *Smug1*$^{+/+}$ and *Smug1*$^{-/-}$ mice.

(A–C and G) Data represent mean ± SEM, n = 3. (E and F) Data represent mean ± SEM. (E) 30 and (F) 15 metaphases were scored. (A–C and E–G) *p ≤ 0.05, **p ≤ 0.01, and ***p ≤ 0.001 (two-tailed Student's t test).

the reduced association of POT1 and TFR2 with SMUG1-KO telomeres, as measured by TeloChIP (Figures 3H and 3I). We did not observe co-localization of TRF1 and the telomeric C-strand probe with γH2AX, suggesting that although shorter, SMUG1-KO telomeres remained bound to and protected by shelterin (Figures S2C and S2D). In conclusion, the loss of SMUG1 results in telomere maintenance defects characterized by fragile telomeres and tissue-specific telomere erosion in mice as well as dramatic telomere attrition in human HAP1 cells.



(legend on next page)

Low *hTERC* Levels Limit Telomerase Activity in SMUG1-KO Human Cells

Prompted by the above telomeric phenotypes, we measured telomerase activity and found that SMUG1-KO cells displayed an 11-fold reduction of activity compared to the control cell line (Figure 4A). Although *hTERT* mRNA expression was somewhat higher in SMUG1-KO cells than in the control (Figure 4B), the amount of hTERT protein was unchanged (Figure 4C). In contrast, *hTERC* levels were 6-fold lower in SMUG1-KO cells than in WT cells as measured by qPCR (Figure 4B), northern blotting (Figure 4D), or RNA sequencing (RNA-seq) (Figures 4E, S3A, and S3B). Interestingly, the number of *hTERC* reads aligning downstream of the core pseudoknot domain was reduced (Figure 4E). Next, we tested whether *hTERC* levels were limiting for telomerase activity in SMUG1-KO cells by overexpressing *hTERC* or hTERT (Figures 4F and S3C). Indeed, transient expression of *hTERC* doubled telomerase activity in SMUG1-KO cells, whereas overexpressing hTERT had no effect (Figures 4F and S3C).

Similarly, *hTERC* levels and telomerase activity increased when SMUG1-WT expression was restored in two independent stable clones of SMUG1-KO cells (Figures 4G and S3D). Consistently, telomere length was also increased (Figure 4H). Overexpression of a SMUG1 mutant unable to bind nucleic acids (NABm) extended telomeres to some degree whereas overexpression of DKC1-binding mutant (DBm) mirrored SMUG1-KO cells (Figures S4D–S4F). In sum, this strongly suggests that telomere attrition in the absence of SMUG1 was caused by low telomerase activity, which was, in turn, a direct consequence of an *hTERC*-biogenesis defect in SMUG1-KO cells.

SMUG1 Is Required for Co-transcriptional Processing of *hTERC*

Since the telomerase RNA component was found to be the limiting factor for telomerase activity in SMUG1-KO cells, we next asked whether SMUG1 binds *hTERC*. In RNA-immunoprecipitation experiments (RNA-IP), we detected 20-fold enrichment of *hTERC* using an anti-SMUG1 antibody compared to the immunoglobulin G (IgG) control (Figure 5A). Similarly, SMUG1 pull-down assays using recombinant SMUG1 protein as bait confirmed that SMUG1 bound directly *hTERC* without any intermediate protein when the total RNA isolated from HAP1 cells was used as prey (Figure S4A). The presence of modified bases in *hTERC* was essential for the interaction, as shown by the inability of SMUG1 to pull-down *in vitro* transcribed *hTERC* (Figures 5B and S4B). Mature *hTERC* levels are determined by the balance between processing and degradation (Figure S4C) (MacNeil et al., 2016). To test whether SMUG1 affects the equilibrium between different *hTERC* products, we measured the levels of the 451-nt mature *hTERC* and the two processing intermediates: 3′-extended and poly(A)-*hTERC* (Figure 5C). In SMUG1-KO cells, the reduced level of mature *hTERC* was accompanied by a 2.5-fold increase in 3′-extended *hTERC* (Figures 5C and 5D) and a 1.5-fold increase in polyadenylated intermediates (Figure 5C). Thus, the absence of SMUG1 disturbed the balance between mature *hTERC* and its processing intermediates.

To further characterize these processing intermediates, we performed 3′-rapid amplification of cDNA ends (RACE)-seq experiments, which showed that the majority of reads aligning to the *hTERC* gene terminated at the expected end, although there was a small increase in fragments aligning from positions 451 to 458 (Figure 5E). Long 3′-extended RNA polymerase II readthrough products were not observed, suggesting that transcriptional termination and 3′ end processing are functional in SMUG1-KO cells. However, accumulation of 3′-extended *hTERC* with short tails (> 10 nt) was detected (Figures 5F and 5G). No dramatic differences in the poly(A) distribution could be observed (Figure 5F), but SMUG1-KO cells exhibited a 1.6-times-higher fraction of long (> 3 nt) poly(A) tails than WT cells (Figures 5G and 5H). Taken together, these data suggest that SMUG1-KO cells have mild *hTERC* processing defects but that the polyadenylation and main *hTERC* end-processing machinery are functional.

The low levels of *hTERC* in SMUG1-KO cells were not due to reduced transcription, as Pol II occupancy at two sites in the *hTERC* promoter (Aalbers et al., 2012; Zhao et al., 1998) and in the coding region was unchanged (Figures 5I, 5J, and S3F). Measurements of nascent *hTERC* kinetics also confirmed a similar transcription rate in HAP1 cells (Figure 5K). Moreover, ChIP experiments showed that SMUG1 was present together with Pol II at the *hTERC* promoter and gene body (Figure 5I). In contrast to Pol II, which was stabilized at the *hTERC* gene (Figure 5I, top), SMUG1 dissociated from chromatin after treatment with actinomycin D (ActD), suggesting that SMUG1 associates with

Figure 3. Dramatic Telomere Attrition in SMUG1-KO Human Cells Is Independent of BER Function

(A–C) Characterization of human HAP1 SMUG1-KO cells.

(A) Schematic representation of the 2-nt deletion, generating an early stop codon of the *SMUG1* gene.

(B and C) *SMUG1* mRNA levels quantified by qPCR (B) and western blot detection of the indicated proteins in HAP1 cells (C). Tubulin was used as loading control.

(D) TeloChIP followed by qPCR showing SMUG1 binding to telomere repeats (top). Enrichment of telomere sequences immunoprecipitated with the SMUG1 antibody is presented as percent of input DNA. Representative blot of immunoprecipitation with IgG and SMUG1 antibodies (bottom). IgG, negative control.

(E) PNA-FISH in metaphase spreads of HAP1 cells. Telomeres were hybridized with a telomere-specific probe (Telo, 5-Carboxytetramethylrhodamine [TAMRA]) and chromosomes were stained with DAPI. Quantification of signal-free ends is shown in the bottom left (circles represent G-strand; triangles represent C-strand).

(F) Representative southern blot of telomere restriction fragment length (TRF) assay in HAP1 cells is shown. Genomic DNA was digested with RsaI and HinfI restriction enzymes alone (Control, C) or RsaI and HinfI followed by incubation with SMUG1 and APE1 enzymes (DNA damage, D). Quantification of absolute telomere length (kb) is shown (right). Ethidium bromide staining is shown as loading control (left). An overview of the modified protocol is shown at the top.

(G) Representative IFs for TRF1 in HAP1 SMUG1-WT and SMUG1-KO cells (scale bars, 2 μm).

(H) POT1 and TRF2 binding to telomeric DNA in HAP1 cells assessed by TeloChIP followed by qPCR detection. Enrichment of the telomere-specific sequences immunoprecipitated with the indicated antibodies are presented as percent of input DNA. IgG, negative control.

(I) Representative blots of immunoprecipitation for IgG, TRF2, and POT1 antibodies.

(B, D, E, and H) Data represent mean ± SD, n = 3. (F) Data represent mean ± SD, n = 2. (E) 30 metaphases were scored. (D–F and H) *$p \leq 0.05$, **$p \leq 0.01$, and ****$p \leq 0.0001$ (two-tailed Student's t test).

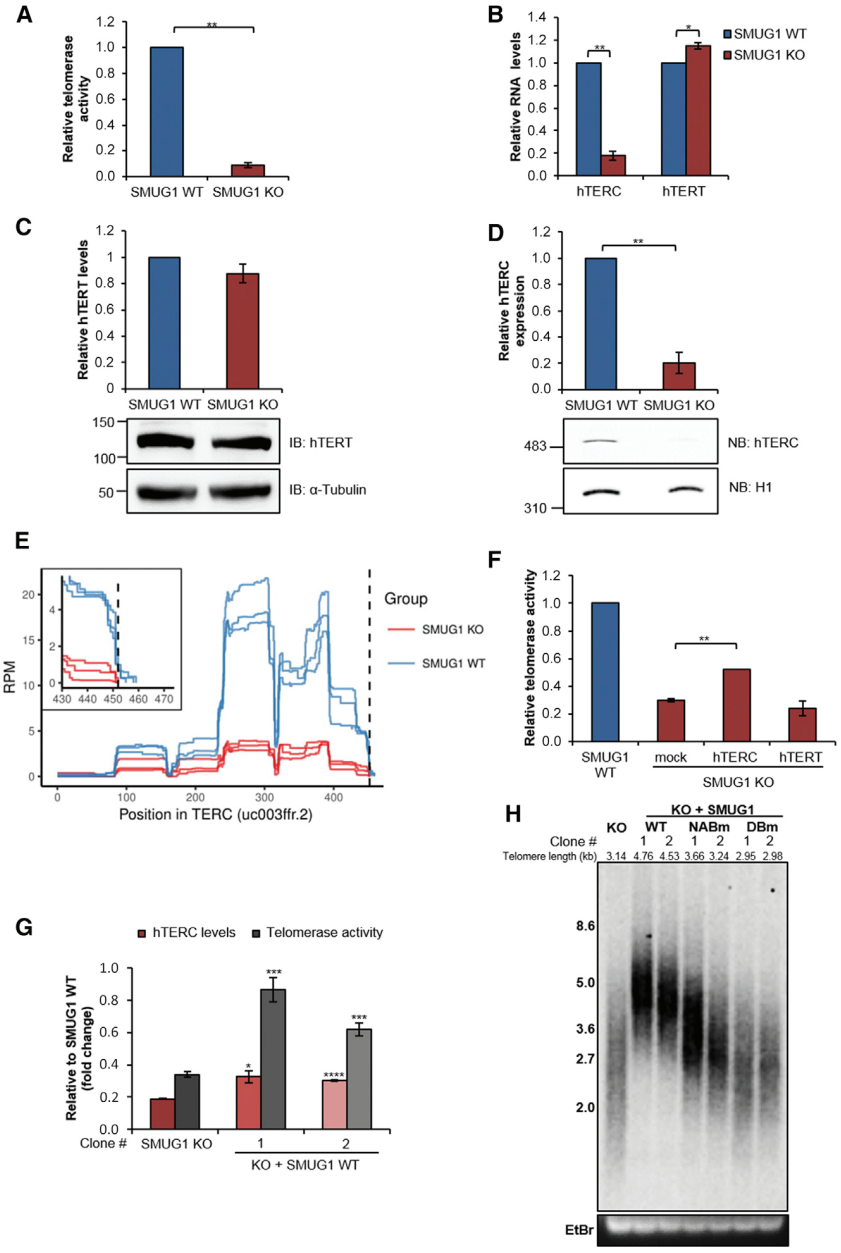


Figure 4. Low *hTERC* Levels Limit Telomerase Activity in SMUG1-KO Human Cells

(A) Telomerase activity of HAP1 cells quantified by qPCR.

(B) Relative RNA levels for *hTERC* and *hTERT* measured by qPCR.

(C) Representative western blot (WB) for hTERT (bottom) with quantification (top). Tubulin was used as loading control.

(D) Representative northern blot (NB) for *hTERC* (bottom) with quantification (top). *Histone H1* was used as loading control.

(E) Position-specific expression profiles for *hTERC*. The graphs show the normalized read depth within the *hTERC* gene locus. The vertical dashed line shows the canonical *hTERC* 3′ end as described by Moon et al. (2015); the insert shows a zoom-in of this 3′ end region.

(F) Telomerase activity quantified by droplet digital PCR (ddPCR) in HAP1 cells transfected with the indicated plasmids.

(G) Quantification of *hTERC* levels and telomerase activity measured via qPCR and ddPCR, respectively, in two SMUG1-KO clones stably expressing SMUG1 WT protein.

(H) Southern blot of the TRF assay in SMUG1-KO clones stably re-expressing WT and mutated SMUG1 unable to bind nucleic acids (NABm) and SMUG1 unable to bind DKC1 (DBm) is shown.

(A–D and G) Data represent means ± SD, n = 3. (F) Data represent means ± SD, n = 2. (A–D, F, and G) *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001, and ****p ≤ 0.0001 (two-tailed Student's t test).

SMUG1 Is Required for *hTERC* Maturation

Binding of SMUG1 to *hTERC* in RNA isolated from cells but not to *in vitro* transcribed *hTERC* strongly suggested that binding requires the presence of modified bases (Figure 5B). Thus, we asked if SMUG1 is required to remove modified or damaged *hTERC* molecules. Since the *hTERC* levels in SMUG1-KO cells were too low to allow direct detection of RNA damage by liquid chromatography-tandem mass spectrometry (LC-MS/MS) (Jobert et al., 2013), we established an assay that detected the presence of SMUG1 substrates in specific regions of the *hTERC* transcript. The assay was based on reduced amplification of transcripts that contained modified bases after incubation with SMUG1 (Figure 6A, left). Interestingly, there was no reduction of amplification efficiency upon enzyme treatment using primers that recognized the 5′ end of the *hTERC* transcript (Figure 6A, right). However, when using primers that amplified a fragment between the *hTERC* CR4/CR5 domain and the H box, we observed reduced amplification of RNA isolated from SMUG1-KO cells. The drop in amplification efficiency corresponded to a 20-fold increase in SMUG1-induced fragmentation at the *hTERC* 3′ region in RNA isolated from

the actively transcribing Pol II complex (Figure 5I, bottom). Interestingly, stabilization of Pol II at the *hTERC* gene after ActD treatment was not seen in SMUG1-KO cells (Figure 5I), suggesting that SMUG1 promotes stability of the stalled Pol II complex. Click-iT experiments showed an increased initial decay rate of *hTERC* in SMUG1-KO cells at 4 h that appeared to stabilize after 24 h (Figure 5L). No differences in *hTERC* transcription could be detected (Figures 5I and 5K), indicating that post-transcriptional mechanisms are the main cause of the observed instability of *hTERC* in SMUG1-KO cells. Taken together, our data show that SMUG1 is required for co-transcriptional processing of *hTERC* and affects its decay.

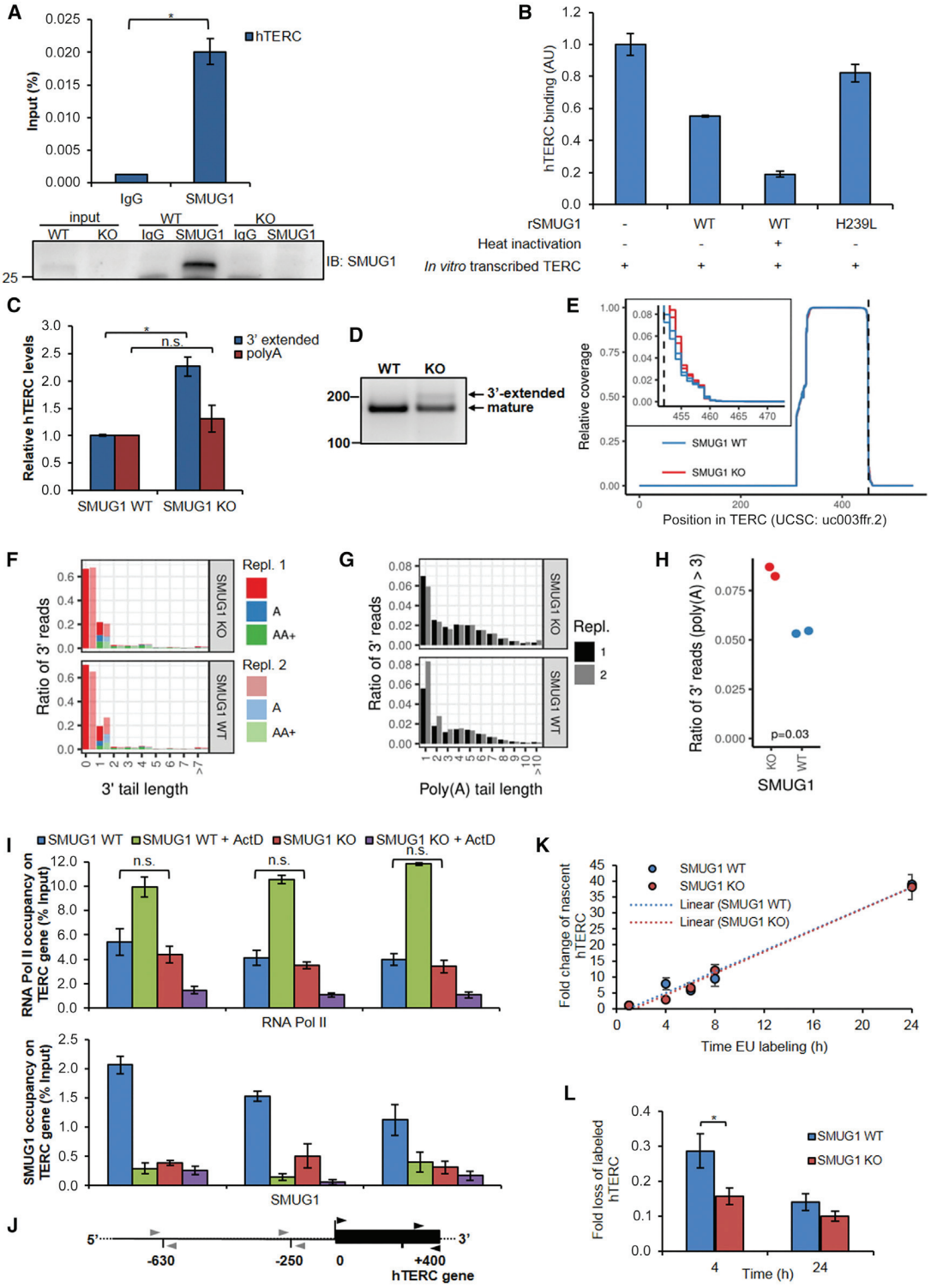


Figure 5. SMUG1 Is Required for Co-transcriptional Processing of *hTERC*

(A) RNA immunoprecipitation (RIP) of *hTERC* by SMUG1 in HAP1 cells using an antibody against SMUG1, quantified by qPCR. Data are presented as percent of input RNA (top). Representative blot of immunoprecipitation for IgG and SMUG1 antibodies (bottom). IgG, negative control.

(B) *hTERC* binding by SMUG1 WT and H239L mutant measured via qPCR after His-tag pull-down using *in vitro* transcribed *hTERC* as prey.

(C) Levels of polyadenylated and 3′-extended *hTERC* relative to mature *hTERC* as quantified by qPCR.

(legend continued on next page)

SMUG1-KO cells compared to the WT (Figure 6A, right). No effect of pretreatment with APE1 alone was detectable (data not shown), suggesting that one or more modified bases that are substrates for SMUG1 are present in *hTERC* in human cells.

RNA-IP showed that DKC1 associated less efficiently with *hTERC* in SMUG1-KO cells (Figure 6B). Taken together, this suggests that the modified base(s) interfered with DKC1 binding. Since DKC1 stabilizes *hTERC* (Shukla et al., 2016; Venteicher et al., 2009; Vulliamy et al., 2008), this reduced association might contribute to the reduced *hTERC* stability in SMUG1-KO cells.

It is also possible that the *hTERC* species containing 3′-modified bases were degraded. Indeed, increased 3′-polyadenylation of *hTERC* was observed in SMUG1-KO cells (Figure 5C). If these species are degraded, inhibition of pathways involved in *hTERC* degradation should restore *hTERC* levels. As expected, small interfering RNA (siRNA)-mediated depletion of EXOSC10, the major RNA exosome, resulted in stabilization of *hTERC* in WT cells (Figures 6C, left; Figure S5). Although no difference in *hTERC* abundance could be detected in WT cells upon PARN depletion, the expected accumulation of polyadenylated *hTERC* was observed (Figure 6C, right). Simultaneous inhibition of the EXOSC10 and PARN enzymes did not result in any synergistic effects (Figure 6C), as expected, since these enzymes act in the same pathway (Shukla et al., 2016). Interestingly, stabilization of *hTERC* was not observed in SMUG1-KO cells upon *EXOSC10* silencing (Figure 6C).

We confirmed that DKC1 affects *hTERC* processing and stability (Shukla et al., 2016), as reduced *hTERC* was observed upon the depletion of DKC1 (Figure 6D, left). Interestingly, depletion of DKC1 in SMUG1-KO cells reduced *hTERC* levels further (Figure 6D). Depletion of PARN, either alone or together with DKC1, did not stabilize *hTERC* levels, but depletion of DKC1 together with EXOSC10 introduced a small stabilizing effect in both cell lines. High accumulation of polyadenylated *hTERC* upon DKC1 inhibition (96 h) was observed in SMUG1-KO cells (Figure 6D, right), consistent with the role of DKC1 in *hTERC* stabilization. No further accumulation of polyadenylated *hTERC* could be observed upon the silencing of *EXOSC10* or *PARN* in DKC1-depleted cells (Figure 6D, right). These results suggest that SMUG1 is required to funnel *hTERC* to the exosome machinery and that *hTERC* depletion occurs via an EXOSC10-independent pathway in SMUG1-KO cells.

In conclusion, SMUG1 is required for co-transcriptional processing of *hTERC* and functions in *hTERC* biogenesis by regulating the presence of base modifications that interfere with DKC1 binding. Consequently, loss of SMUG1 leads to an imbalance between mature *hTERC* and its processing intermediates, which are degraded in an EXOSC10-independent RNA degradation pathway (Figure 7).

DISCUSSION

Here, we show that human SMUG1-DNA glycosylase is required for the maturation of *hTERC* through regulating the levels of modified bases in a region important for DKC1 binding. In the absence of SMUG1, *hTERC* molecules containing modified bases and processing intermediates accumulate, accompanied by reduced levels of mature *hTERC*. An insufficient *hTERC* level limits telomerase activity in SMUG1-KO cells, leading to severe telomere attrition.

SMUG1 is a multifunctional enzyme that acts both in BER (Nilsen et al., 2001) and in RNA processing (Jobert et al., 2013). To determine which activity is more important for telomere maintenance, we used a complementation strategy where we found that *hTERC* levels were limiting for telomerase activity in SMUG1-KO cells. Ectopic expression of SMUG1 restored *hTERC* levels and telomerase activity. As this was accompanied by increased telomere length, we conclude that the telomere maintenance defects are caused by loss of SMUG1. The SMUG1/DKC1 interaction appears essential for this function as a SMUG1 mutant that cannot bind DKC1 failed to restore telomere length. Some restoration of telomere length was seen after complementation with a SMUG1 nucleic acid binding mutant that, consequently, has low DNA-glycosylase activity on synthetic substrates *in vitro* (Matsubara et al., 2004). Because the ability to associate with DKC1 was preserved in this mutant (Jobert et al., 2013), the inability of the nucleic acid binding mutant to restore telomere length to the same extent as SMUG1 WT suggested that the DNA-glycosylase activity is also required.

A function for the DNA-glycosylase activity was further supported by the fact that binding of SMUG1 to *hTERC* appeared to require the presence of modified bases as we could only detect the association with *hTERC* isolated from cells and not *in vitro* transcribed *hTERC* that lacked modified bases. *hTERC* exhibited a greater number of modified bases in a region between the CR4/CR5 domain and the H box, where DKC1 binds. The nature of the modified base(s) in *hTERC* remains unknown as we were unable, despite some stabilization of *hTERC* after

(D) 3′-RACE products separated by agarose gel electrophoresis.
(E) Position-specific read profiles for 3′ RACE libraries. The graphs show the read depth for reads aligning to the *hTERC* gene locus, normalized to the maximum read depth within each library. The vertical dashed line shows the canonical *hTERC* 3′ end as described (Moon et al., 2015). The insert shows a zoom-in of the region immediately after the 3′ end.
(F) 3′-RACE products containing the canonical 3′ *hTERC* site (CAGGACTCGGCTCACACATGC). Reads containing the canonical 3′ *hTERC* site were classified based on the number of additional genome-matching nucleotides at their 3′ ends (x axis) and their 3′ poly(A) content (color). Cells were grouped by genotype; replicate samples are shown separately.
(G and H) Distribution (G) and ratio (H) of poly(A) tail length for 3′-RACE products containing the canonical 3′ *hTERC* site (CAGGACTCGGCTCACACATGC).
(I) Co-occupancy analysis of active RNA polymerase II (top) and SMUG1 (bottom) on the *hTERC* gene as measured by re-ChIP. ActD was used to inhibit RNA polymerase II.
(J) Schematic representation of the *hTERC* gene and its promoter region. The positions of the primers used (−630 through +400; +1 defined as the *hTERC* transcriptional start site) are indicated along the *hTERC* gene.
(K and L) Click-iT experiments showing nascent *hTERC* levels (K) and *hTERC* decay (L) over time in HAP1 cells.
(E–H) n = 2. (A–C, I, and L) Data represent means ±SEM, n = 3 *p ≤ 0.05, **p ≤ 0.01, and ****p ≤ 0.0001; ns, not significant (two-tailed Student's t test).

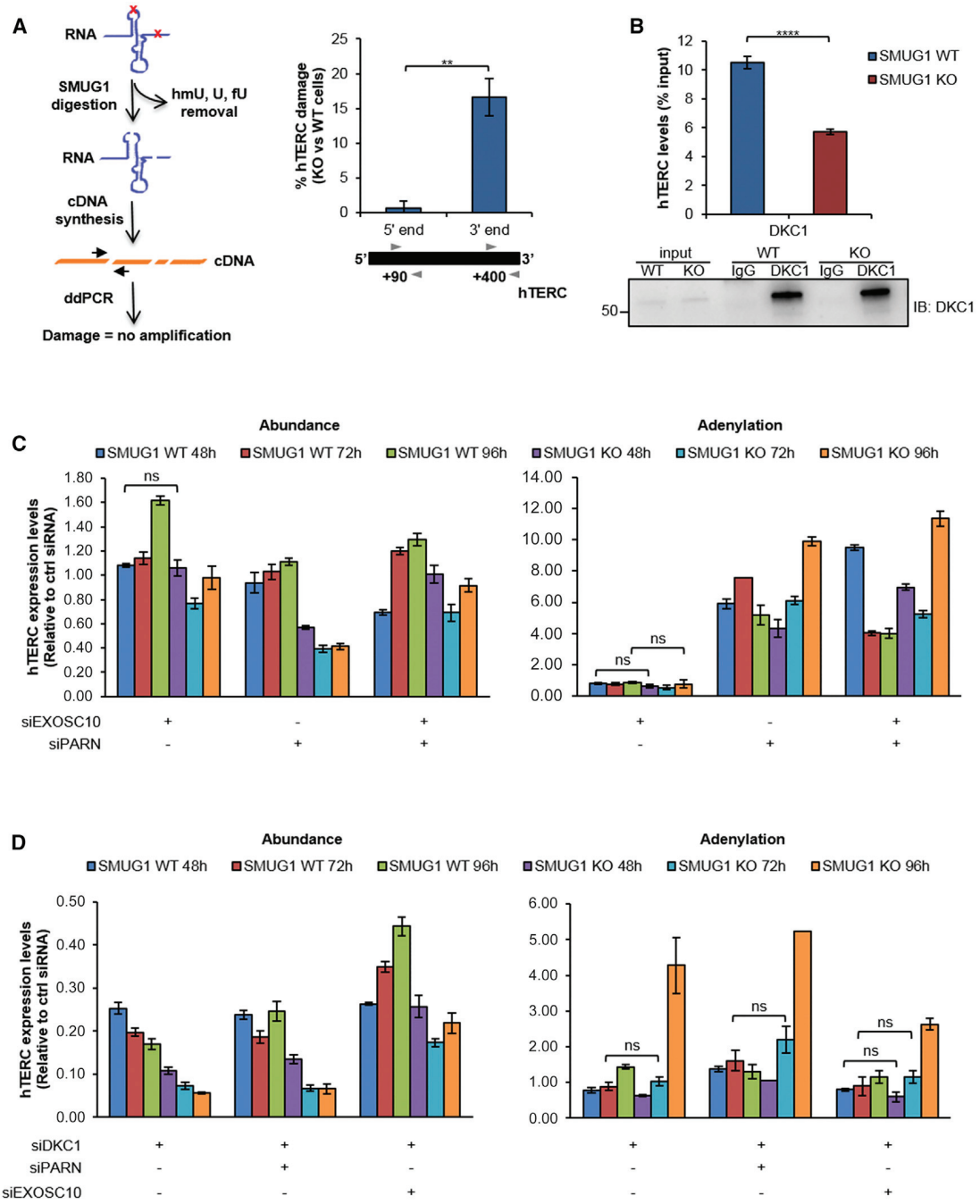


Figure 6. The Absence of SMUG1 Affects *hTERC* Degradation

(A) *hTERC* RNA damage levels at 5′ and 3′ regions as measured by ddPCR (right). RNA from HAP1 cells was digested with SMUG1 enzyme prior to cDNA synthesis (left). Primer positions are indicated along *hTERC*. An overview of the protocol is shown (left).

(B) qPCR showing *hTERC* immunoprecipitation by DKC1. Data are presented as percent of input RNA. Background signal given by IgG control was subtracted from the specific reaction (top). Representative blot of immunoprecipitation for IgG and DKC1 antibodies (bottom).

(C and D) Quantification of hTERC levels and adenylation frequency in HAP1 cells after siRNA-mediated depletion of EXOSC10 and PARN (C) or after DKC1, PARN, and EXOSC10 silencing (D) measured by qPCR. Cells were harvested at 48, 72, and 96 h after addition of siRNA. Relative abundance was measured using cDNA synthesized with random primers while adenylation frequency was estimated as the ratio of oligo(dT)-primed cDNA to random-primed cDNA.

(A–D) Data represent means ± SEM, ns, $^{**}p \leq 0.01$, $^{****}p \leq 0.0001$; ns, not significant (comparison between HAP1 SMUG1-WT and SMUG1-KO silenced for each time point, two-tailed Student's t test).

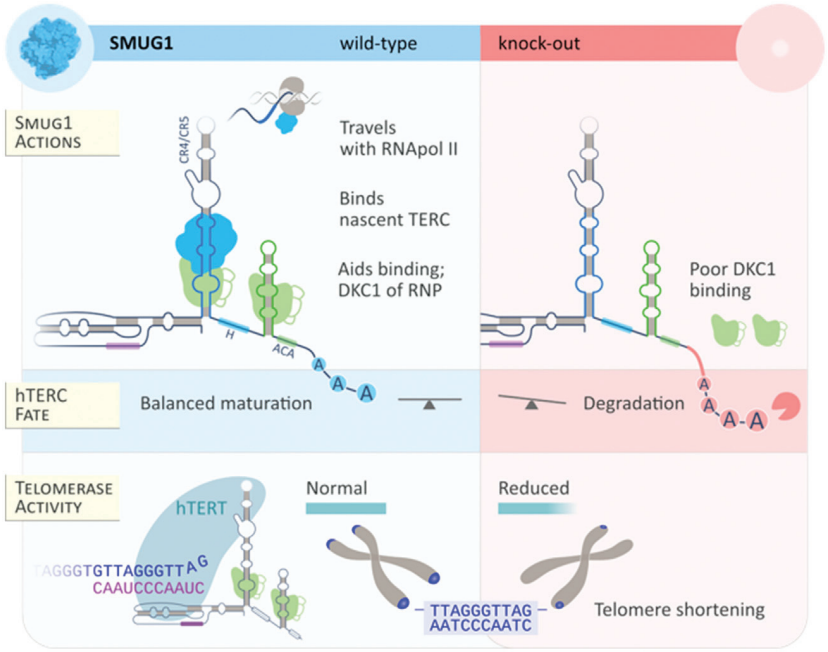


Figure 7. Working Model

Our data support a model where SMUG1 acts co-transcriptionally in *hTERC* biogenesis in a step upstream of the PARN/PABPN1 machinery by regulating the presence of modified bases in a region between CR4/CR5 and the H box. SMUG1 is required for efficient DKC1 binding and exosome-mediated degradation of these modified *hTERC* molecules. In SMUG1-KO cells the equilibrium between the mature *hTERC* and its processing intermediates is shifted toward degradation, leading to limiting amounts of *hTERC* unable to sustain telomerase activity. Figure by Ellen Tenstad/Science Shaped.

depletion of EXOSC10 and DKC1 (Figure 6D), to isolate sufficiently high amounts of *hTERC* from SMUG1-KO cells to perform lesion detection by LC-MS/MS. However, as treatment with APE1 alone did not reduce the amplification efficiency, we concluded that the base damage or modified bases accumulate in *hTERC* in SMUG1-KO are substrates for SMUG1. The likely substrate would be hmU or deoxyU, which we previously showed were substrates for SMUG1 in RNA (Jobert et al., 2013). Culturing cells in the presence of hmU did not, however, stimulate the association between SMUG1 and *hTERC* (Figure S3E). Thus, the association did not depend on exogenously induced modified RNA bases. The fact that we preferentially detected SMUG1 substrates toward the 3′ end of *hTERC*, as opposed to a uniform distribution, is more consistent with the presence of a base modification, rather than random damage. hmU has been identified in RNA (Jobert et al., 2013). Since it cannot be introduced into RNA by direct oxidation of thymine, it is likely formed by deamination of 5-hydroxymethylcytosine derived from 5-methylcytosine (Huber et al., 2015). It is tempting to speculate, therefore, that SMUG1 might regulate the presence of modified cytosines at two other methylated sites in *hTERC*, C323 and C455, which are located in the CR4/CR5 and the H box, respectively, the regions that accumulate base damage in SMUG1-KO cells.

Interestingly, the phenotype we observed in SMUG1-KO cells is reminiscent of that seen after depletion of HuR; HuR facilitated methylation of *hTERC* at C106, thereby promoting DKC1 binding to *hTERC* and assembly to hTERT (Tang et al., 2018). Our data present the possibility that SMUG1 might function as a dynamic regulator of DKC1 binding in response to base modifications in the CR4/CR5 domain, but further characterization of the specific bases requires the development of new mapping techniques. Whether dynamic modification of *hTERC* alters its secondary structure or otherwise interferes with DKC1 localization (Figure 1A) will be the focus of future studies.

hTERC levels could not be rescued by knocking down components of the RNA decay machinery. This suggests that SMUG1 is required to recruit the exosome and that other, yet-to-be-identified, degradation pathways may act in SMUG1-KO cells. Interestingly, SMUG1 stabilized the stalled RNA polymerase complex at the *hTERC* gene. The exosome is recruited co-transcriptionally, and Pol II backtracking provides a free RNA 3′ end for the core (Lemay et al., 2014). Our data do not suggest that SMUG1 affects *hTERC* transcriptional termination because long read-through molecules were not observed. The absence of large differences in the poly(A) distribution shows that the end-processing machinery is functional in SMUG1-KO cells. However, SMUG1-KO cells harbored higher levels of slightly elongated *hTERC* molecules, exhibiting base modification(s) toward their 3′ region. Hence, it is possible that SMUG1 acts co-transcriptionally to target *hTERC* containing hmU modified bases that interfere with DKC1 binding to the exosome.

Although the function of SMUG1 in *hTERC* biogenesis was the dominating phenotype in human SMUG1-KO cells, we did observe telomeric base damage (Figure 3F), which could contribute to the reduced binding of TRF2 and POT1 observed in TeloChIP experiments (Figure 3H), as the two main SMUG1 substrates (uracil [Vallabhaneni et al., 2015] and hmU [Theruvathu et al., 2014]) interfere with shelterin assembly *in vitro*. In addition, the short telomeres contain less available substrate for TRF1 and TRF2 binding, which might be the main reason for dys-localization of these proteins in SMUG1-KO cells (Figures 3G and S2B).

As observed previously in *Ogg1*$^{-/-}$ (Wang et al., 2010), *Nthl1*$^{-/-}$ (Vallabhaneni et al., 2013), and *Ung*$^{-/-}$ (Vallabhaneni et al., 2015) mice, *Smug1*$^{-/-}$ MEFs and mice show increased DNA damage in telomeres and multiple telomere defects (Figures 2 and S1D). In *Smug1*$^{-/-}$ mice, fragile telomeres and reduced proliferative capacity of bone marrow cells are seen in the presence of TERT, whereas in the other DNA-glycosylase-knockout mice, telomere maintenance defects become obvious first in the background of TERT deficiency (Vallabhaneni et al., 2015). It is possible that the function of SMUG1 in *Terc* biogenesis contributes to the apparently stronger telomere phenotype in *Smug1*$^{-/-}$ mice, but

although SMUG1 binds *Terc* in MEFs (Figure S1F), *Smug1*$^{-/-}$ MEFs did not show consistently reduced *Terc* levels (Figure S1G). Therefore, the telomere maintenance defects in MEFs appear to be mainly caused by the loss of SMUG1-dependent BER (Alsøe et al., 2017). Measurements of telomere fragility (Figures 2D, 2E, and S1D) indicated impaired replication of the C-rich telomere strand, which would be expected to contain more uracil lesions. However, as no method is available to discriminate between uracil and hmU in specific genomic regions, we do not know which SMUG1 substrate gives rise to these phenotypes in MEFs. Faithful BER of U has been shown to be needed to protect telomeres from unsolicited activation of mismatch repair at U:G pairs generated by activation-induced deaminase, leading to resection of the C-rich strand in *Ung*$^{-/-}$ B cells (Cortizas et al., 2016). We could not detect expression of AID in the HAP1 cells, but we cannot exclude the possibility that AID activation at one stage during establishment of the SMUG1-KO cell line is a cause of the extremely short telomeres in this cell line. In any case, the *hTERC* biogenesis defect prevented restoration of the telomeres.

Taken together, our data support a role of SMUG1 in telomere maintenance both as a BER enzyme and through its RNA processing function, but the extent to which these activities affect telomere maintenance differs between species and cell types. In human cells, SMUG1 acts in *hTERC* biogenesis by regulating the presence of modified bases in *hTERC* and facilitating DKC1 binding.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.celrep.2019.07.040.

ACKNOWLEDGMENTS

We thank A.E. Moen, T. Tannæs, and T. Lüders for help and technical assistance and Primo Schär for sharing reagents. This work has been funded by grants to H.N. from the Research Council of Norway (grant no. 229633). L.L. was funded by a grant from the Norwegian Cancer Society (grant no. 4501723-2015). P.K. was funded by a PhD fellowship from the University of Oslo. L.A. was funded by a grant from the South East Regional Health Authority (project no. 274901). S.C. was funded by a grant from the Norwegian Cancer Society (grant no. 4501723-2013). P.Y. was funded by a grant from the South East Regional Health Authority (project no. 2017029). The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7-PEOPLE-2013-COFUND) under grant agreement no. 609020 - Scientia Fellows to H.N. and A.G. BioCore is funded by the Faculty of Medicine at NTNU and Central Norway Regional Health Authority. P.S. is funded through grants from the Norwegian Research Council, Central Norway Regional Health Authority, and Stiftelsen Kristian Gerhard Jebsen. RNA sequencing was provided by the Genomics Core Facility (GCF), Norwegian University of Science and Technology (NTNU). GCF is funded by the Faculty of Medicine and Health Sciences at NTNU and the Central Norway Regional Health Authority. This work was initiated through a collaboration based on the COST action Cancer Control and Genomic Integrity (BM0703).

AUTHOR CONTRIBUTIONS

P.K., L.L., S.C., A.G., Q.Y.E., L.J., L.A., and P.Y. performed experiments; P.K., L.L., P.S., S.G., and H.N. analyzed data; and P.K., L.L., and H.N. wrote the manuscript. All authors edited the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 3, 2018
Revised: May 28, 2019
Accepted: July 14, 2019
Published: August 13, 2019

REFERENCES

Aalbers, A.M., Kajigaya, S., van den Huevel-Eibrink, M.M., van der Velden, V.H., Calado, R.T., and Young, N.S. (2012). Human telomere disease due to disruption of the CCAAT box of the TERC promoter. Blood *119*, 3060–3063.

Alsøe, L., Sarno, A., Carracedo, S., Domanska, D., Dingler, F., Lirussi, L., Sen-Gupta, T., Tekin, N.B., Jobert, L., Alexandrov, L.B., et al. (2017). Uracil Accumulation and Mutagenesis Dominated by Cytosine Deamination in CpG Dinucleotides in Mice Lacking UNG and SMUG1. Sci. Rep. *7*, 7199.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq–a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166–169.

Banfalvi, G. (2017). Overview of Cell Synchronization. Methods Mol. Biol. *1524*, 3–27.

Boyraz, B., Moon, D.H., Segal, M., Muosieyiri, M.Z., Aykanat, A., Tai, A.K., Cahan, P., and Agarwal, S. (2016). Posttranscriptional manipulation of TERC reverses molecular hallmarks of telomere disease. J. Clin. Invest. *126*, 3377–3382.

Cortizas, E.M., Zahn, A., Safavi, S., Reed, J.A., Vega, F., Di Noia, J.M., and Verdun, R.E. (2016). UNG protects B cells from AID-induced telomere loss. J. Exp. Med *213*, 2459–2472.

Dahl, J.A., and Collas, P. (2008). A rapid micro chromatin immunoprecipitation assay (microChIP). Nat. Protoc *3*, 1032–1045.

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

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Egan, E.D., and Collins, K. (2012). An enhanced H/ACA RNP assembly mechanism for human telomerase RNA. Mol. Cell Biol *32*, 2428–2439.

Ge, J., Crosby, S.D., Heinz, M.E., Bessler, M., and Mason, P.J. (2010). SnoRNA microarray analysis reveals changes in H/ACA and C/D RNA levels caused by dyskerin ablation in mouse liver. Biochem. J *429*, 33–41.

Göhring, J., Fulcher, N., Jacak, J., and Riha, K. (2014). TeloTool: a new tool for telomere length measurement from terminal restriction fragment analysis with improved probe intensity correction. Nucleic Acids Res. *42*, e21.

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

Huber, S.M., van Delft, P., Mendil, L., Bachman, M., Smollett, K., Werner, F., Miska, E.A., and Balasubramanian, S. (2015). Formation and abundance of 5-hydroxymethylcytosine in RNA. ChemBioChem *16*, 752–755.

Jobert, L., Skjeldam, H.K., Dalhus, B., Galashevskaya, A., Vågbø, C.B., Bjørås, M., and Nilsen, H. (2013). The human base excision repair enzyme SMUG1 directly interacts with DKC1 and contributes to RNA quality control. Mol. Cell *49*, 339–345.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Law, C.W., Chen, Y., Shi, W., and Smyth, G.K. (2014). voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol. *15*, R29.

Lee, J.H., Lee, Y.S., Jeong, S.A., Khadka, P., Roth, J., and Chung, I.K. (2014). Catalytically active telomerase holoenzyme is assembled in the dense fibrillar component of the nucleolus during S phase. Histochem. Cell Biol. *141*, 137–152.

Lemay, J.F., Larochelle, M., Marguerat, S., Atkinson, S., Bähler, J., and Bachand, F. (2014). The RNA exosome promotes transcription termination of backtracked RNA polymerase II. Nat. Struct. Mol. Biol *21*, 916–926.

Ludlow, A.T., Robin, J.D., Sayed, M., Litterst, C.M., Shelton, D.N., Shay, J.W., and Wright, W.E. (2014). Quantitative telomerase enzyme activity determination using droplet digital PCR with single cell resolution. Nucleic Acids Res. *42*, e104.

MacNeil, D.E., Bensoussan, H.J., and Autexier, C. (2016). Telomerase Regulation from Beginning to the End. Genes (Basel) *7*, E64.

Matsubara, M., Tanaka, T., Terato, H., Ohmae, E., Izumi, S., Katayanagi, K., and Ide, H. (2004). Mutational analysis of the damage-recognition and catalytic mechanism of human SMUG1 DNA glycosylase. Nucleic Acids Res. *32*, 5291–5302.

Mitchell, J.R., Wood, E., and Collins, K. (1999). A telomerase component is defective in the human disease dyskeratosis congenital. Nature *402*, 551–555.

Moon, D.H., Segal, M., Boyraz, B., Guinan, E., Hofmann, I., Cahan, P., Tai, A.K., and Agarwal, S. (2015). Poly(A)-specific ribonuclease (PARN) mediates 3′-end maturation of the telomerase RNA component. Nat. Genet. *47*, 1482–1488.

Morin, G.B. (1989). The human telomere terminal transferase enzyme is a ribonucleoprotein that synthesizes TTAGGG repeats. Cell *59*, 521–529.

Nguyen, D., Grenier St-Sauveur, V., Bergeron, D., Dupuis-Sandoval, F., Scott, M.S., and Bachand, F. (2015). A Polyadenylation-Dependent 3′ End Maturation Pathway Is Required for the Synthesis of the Human Telomerase RNA. Cell Rep. *13*, 2244–2257.

Nilsen, H., Haushalter, K.A., Robins, P., Barnes, D.E., Verdine, G.L., and Lindahl, T. (2001). Excision of deaminated cytosine from the vertebrate genome: role of the SMUG1 uracil-DNA glycosylase. EMBO J. *20*, 4278–4286.

O'Callaghan, N.J., and Fenech, M. (2011). A quantitative PCR method for measuring absolute telomere length. Biol. Proced. Online *13*, 3.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. *43*, e47.

Robinson, M.D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol *11*, R25.

Schmidt, J.C., and Cech, T.R. (2015). Human telomerase: biogenesis, trafficking, recruitment, and activation. Genes Dev *29*, 1095–1105.

Shukla, S., Schmidt, J.C., Goldfarb, K.C., Cech, T.R., and Parker, R. (2016). Inhibition of telomerase RNA decay rescues telomerase deficiency caused by dyskerin or PARN defects. Nat. Struct. Mol. Biol *23*, 286–292.

Tang, H., Wang, H., Cheng, X., Fan, X., Yang, F., Zhang, M., Chen, Y., Tian, Y., Liu, C., Shao, D., et al. (2018). HuR regulates telomerase activity through TERC methylation. Nat. Commun. *9*, 2213.

Theruvathu, J.A., Darwanto, A., Hsu, C.W., and Sowers, L.C. (2014). The effect of Pot1 binding on the repair of thymine analogs in a telomeric DNA sequence. Nucleic Acids Res *42*, 9063–9073.

Tseng, C.K., Wang, H.F., Burns, A.M., Schroeder, M.R., Gaspari, M., and Baumann, P. (2015). Human Telomerase RNA Processing and Quality Control. Cell Rep. *13*, 2232–2243.

Vallabhaneni, H., O'Callaghan, N., Sidorova, J., and Liu, Y. (2013). Defective repair of oxidative base lesions by the DNA glycosylase Nth1 associates with multiple telomere defects. PLoS Genet *9*, e1003639.

Vallabhaneni, H., Zhou, F., Maul, R.W., Sarkar, J., Yin, J., Lei, M., Harrington, L., Gearhart, P.J., and Liu, Y. (2015). Defective repair of uracil causes telomere defects in mouse hematopoietic cells. J. Biol. Chem *290*, 5502–5511.

Venteicher, A.S., Abreu, E.B., Meng, Z., McCann, K.E., Terns, R.M., Veenstra, T.D., Terns, M.P., and Artandi, S.E. (2009). A human telomerase holoenzyme protein required for Cajal body localization and telomere synthesis. Science *323*, 644–648.

Vulliamy, T., Beswick, R., Kirwan, M., Marrone, A., Digweed, M., Walne, A., and Dokal, I. (2008). Mutations in the telomerase component NHP2 cause the premature ageing syndrome dyskeratosis congenital. Proc. Natl. Acad. Sci. USA *105*, 8073–8078.

Wang, Z., Rhee, D.B., Lu, J., Bohr, C.T., Zhou, F., Vallabhaneni, H., de Souza-Pinto, N.C., and Liu, Y. (2010). Characterization of oxidative guanine damage and repair in mammalian telomeres. PLoS Genet *6*, e1000951.

Xi, L., and Cech, T.R. (2014). Inventory of telomerase components in human cells reveals multiple subpopulations of hTR and hTERT. Nucleic Acids Res. *42*, 8565–8577.

Zhao, J.Q., Hoare, S.F., McFarlane, R., Muir, S., Parkinson, E.K., Black, D.M., and Keith, W.N. (1998). Cloning and characterization of human and mouse telomerase RNA gene promoter sequences. Oncogene *16*, 1345–1350.

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

Zinder, J.C., and Lima, C.D. (2017). Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors. Genes Dev *31*, 88–100.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit anti-SMUG1	Abcam	Cat# ab192240
Rabbit anti-IgG, Isotype control	Abcam	Cat# ab172730; RRID:AB_2687931
Rabbit anti-RNA polymerase II CTD repeat YSPTSPS (phospho S5)	Abcam	Cat# ab5131; RRID:AB_449369
Rabbit anti-His tag	Cell signaling technology	Cat# 2365; RRID:AB_2115720
Rabbit anti-coilin	Cell signaling technology	Cat# 14168; RRID:AB_2798410
Rabbit anti-DKC1	Bethyl laboratories, Inc	Cat# A302-591A; RRID:AB_10554666
Rat anti-BrdU	Abcam	Cat# ab6326; RRID:AB_305426
Rabbit anti-TRF1	Gift from DeLange lab	N/A
Rabbit anti-TRF2	Gift from DeLange lab	N/A
Rabbit anti-HA tag (C29F4)	Cell signaling technology	Cat# 3724; RRID: AB_1549585
Mouse anti-phospho Histone H2A.X (Ser139), clone JBW301	Millipore	Cat# 05-636-2KL; RRID:AB_309864
Rabbit anti-TRF2	Novus Biologicals	Cat# NB110-57130; RRID:AB_844199
Rabbit anti-POT1	Abcam	Cat# ab21382; RRID:AB_777376
Rabbit anti-POT1	Abcam	Cat# ab240948
Rabbit anti-Telomerase reverse transcriptase	Abcam	Cat# ab32020; RRID:AB_778296
Rabbit anti-GAPDH	Cell signaling technology	Cat# 2118; RRID:AB_561053
Monoclonal Anti-α-Tubulin antibody produced in mouse	Sigma-Aldrich-Aldrich	Cat# T5168; RRID:AB_477579
Monoclonal ANTI-FLAG® M2 antibody produced in mouse	Sigma-Aldrich-Aldrich	Cat# F1804; RRID:AB_262044
Rabbit anti-PARN	Abcam	Cat# ab188333
Rabbit anti-EXOSC10	Abcam	Cat# ab50558; RRID:AB_869937
Goat anti-Mouse IgG (H+L) Secondary Antibody, HRP	Life Technologies	Cat# 31430; RRID:AB_228307
Goat anti-Rabbit IgG (H+L) Secondary Antibody, HRP	Life Technologies	Cat# 31460; RRID:AB_228341
Monoclonal Anti-Rabbit IgG, Native–Peroxidase antibody produced in mouse	Sigma-Aldrich-Aldrich	Cat# R3155; RRID:AB_1079117
Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 594	Life Technologies	Cat# A11012; RRID: AB_2534079
Goat anti-Rat IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 594	Life Technologies	Cat# A11007; RRID: AB_10561522
Goat anti-Rabbit IgG (H+L) Cross-Adsorbed ReadyProbes Secondary Antibody, Alexa Fluor 488	Life Technologies	Cat# R37116; RRID: AB_2556544
Goat anti-Mouse IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 488	Life Technologies	Cat# 11029; RRID: AB_2534088
Bacterial and Virus Strains		
E. coli BL21 (DE3) competent cells	Agilent technologies	Cat# 200131
Biological Samples		
Mouse: Bone marrow cells from Smug1$^{+/+}$ mice	Lab made	N/A
Mouse: Bone marrow cells from Smug1$^{-/-}$ mice	Lab made	N/A
Mouse: Liver tissue from Smug1$^{+/+}$ mice	Lab made	N/A
Mouse: Liver tissue from Smug1$^{-/-}$ mice	Lab made	N/A
Mouse: Spleen tissue from Smug1$^{+/+}$ mice	Lab made	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Mouse: Spleen tissue from Smug1$^{-/-}$ mice	Lab made	N/A
Chemicals, Peptides, and Recombinant Proteins		
Phosphate buffered saline (PBS)	Lab made	N/A
Dulbecco's Modified Eagle Medium (DMEM)	Life Technologies	Cat# 10566016
Iscove's Modified Dulbecco's Medium (IMDM)	Life Technologies	Cat# 31980048
Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F-12)	Life Technologies	Cat# 31331028
Opti-MEM I Reduced Serum Medium	Life Technologies	Cat# 31985062
Fetal Bovine Serum (FBS)	Sigma-Aldrich-Aldrich	Cat# F7524
HyClone Calf Serum	Thermo Fisher	Cat# SH30073.03
Penicillin-Streptomycin (10,000 U/mL)	Life Technologies	Cat# 15140122
MEM Non-Essential Amino Acids Solution (100X)	Life Technologies	Cat# 11140035
Lipofectamine 3000 Transfection Reagent	Life Technologies	Cat# L3000015
Fugene 6	Promega	Cat# E2691
Lipofectamine RNAiMAX Transfection Reagent	Life Technologies	Cat# 13778150
TRIzol Reagent	Life Technologies	Cat# 15596018
Puromycin dihydrochloride	Sigma-Aldrich-Aldrich	Cat# P8833
Actinomycin D	Sigma-Aldrich-Aldrich	Cat# A9415
BrdU (5-Bromo-2'-Deoxyuridine)	Life Technologies	Cat# B23151
Protease Inhibitor Cocktail	Sigma-Aldrich-Aldrich	Cat# P8340
SMARTpool: siGENOME PARN siRNA	Dharmacon	Cat# M-011348-00-0005
SMARTpool: siGENOME EXOSC10 siRNA	Dharmacon	Cat# M-010904-01-0005
Silencer® Negative Control No. 1 siRNA	Ambion	Cat# AM4611
Silencer DKC1 siRNA s4111	Ambion	Cat# 4392420
Silencer DKC1 siRNA s4112	Ambion	Cat# 4457298
ProLong Diamond Antifade Mountant with DAPI	Life Technologies	Cat# P36971
Any kD Mini-PROTEAN® TGX Precast Protein Gels, 10-well, 50 μl	Bio-Rad	Cat# 456-9034
SuperSignal West Femto Maximum Sensitivity Substrate	Life Technologies	Cat# 34095
SuperSignal West Pico PLUS Chemiluminescent Substrate	Life Technologies	Cat# 34577
RNeasy Mini Kit	QIAGEN	Cat# 74106
iScript cDNA Synthesis Kit	Bio-Rad	Cat# 1708891
iScript Select cDNA Synthesis Kit	Bio-Rad	Cat# 1708897
Power SYBR Green PCR Master Mix	Life Technologies	Cat# 4367659
SIRV-Set 3 (Iso Mix E0 / ERCC)	Lexogen	Cat# SKU: 051.01
SuperScript IV Reverse Transcriptase	Life Technologies	Cat# 18090010
hSMUG1	New England Biolabs	Cat# M0336S
Dynabeads Protein G for Immunoprecipitation	Life Technologies	Cat# 10004D
QX200 ddPCR EvaGreen Supermix	Bio-Rad	Cat# 1864033
Formaldehyde solution	Sigma-Aldrich-Aldrich	Cat# F8775
DNase I, RNase-free (1 U/μL)	Life Technologies	Cat# EN0521
DIG Easy Hyb Granules	Sigma-Aldrich-Aldrich	Cat# 11796895001
Universal miRNA Cloning Linker	New England Biolabs	Cat# S1315S
T4 RNA Ligase 2, truncated KQ	New England Biolabs	Cat# M0373S
RNA Clean & Concentrator Kit	Zymo Research	Cat# R1013
SuperScript III Reverse Transcriptase	Life Technologies	Cat# 18080093
AccuPrime Pfx SuperMix	Life Technologies	Cat# 12344040

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
mouse IL-6 recombinant protein	eBioscience	Cat# 14-8061-80
mouse stem cell factor (SCF) recombinant protein	eBioscience	Cat# 14-8341-63
Mouse Methylcellulose Complete Media	R&D systems	Cat# HSC007
Colcemid	Life Technologies	Cat# 15210040
Blocking reagent	Sigma-Aldrich-Aldrich	Cat# 11096176001
Vectashield mounting medium with DAPI	Vector labs	Cat# H-1200
Pierce BCA Protein Assay Kit	Life Technologies	Cat# 23227
DNeasy Blood & Tissue Kit	QIAGEN	Cat# 69504
APE1	New England Biolabs	Cat# M0282S
Thymidine	Sigma-Aldrich-Aldrich	Cat# T1895
Imidazole buffer Solution	Sigma-Aldrich-Aldrich	Cat# 68268
Isopropyl β-D-1-thiogalactopyranoside	Sigma-Aldrich-Aldrich	Cat# I5502
Ni-NTA agarose resin	QIAGEN	Cat# 30210
Econo-Column® Chromatography Columns	Bio-Rad	Cat# 7372522
Bio-Scale Mini Macro-Prep High S cartridge	Bio-Rad	Cat# 7324134
Ultrafree-MC Centrifugal Filter	Millipore	Cat# UFC30DV0S
HIS-Select® Nickel Magnetic Agarose beads	Sigma-Aldrich-Aldrich	Cat# H9914
SMUG1 wt recombinant protein	Lab made	N/A
SMUG1 H239L recombinant protein	Lab made	N/A
Critical Commercial Assays		
SENSE Total RNA-Seq Library Prep Kit	Lexogen	Cat# SKU: 009.24
Click-iT Nascent RNA Capture Kit, for gene expression analysis	Life Technologies	Cat# C10365
DIG Oligonucleotide 3′ End Labeling Kit, 2nd generation	Sigma-Aldrich-Aldrich	Cat# 03353575910
TeloTAGGG Telomere Length Assay	Sigma-Aldrich-Aldrich	Cat# 12209136001
TruSeq Nano DNA Low Throughput Library Prep Kit	Illumina	Cat# 20015964
MiSeq Reagent Kit v2 (500-cycles)	Illumina	Cat# MS-102-2003
Duolink® *In Situ* Red Starter Kit Mouse/Rabbit	Sigma-Aldrich-Aldrich	Cat# DUO92101
TRAPeze® Telomerase Detection Kit	Millipore	Cat# S7710
MEGAscript T7 Transcription Kit	Life Technologies	Cat# AM1333
Deposited Data		
RNA-seq data	This study	GEO: GSE116580
3′ RACE seq data	This study	GEO: GSE116580
Experimental Models: Cell Lines		
Mouse: Mouse Embryonic fibroblasts *Smug1*$^{+/+}$	Lab made	Alsøe et al., 2017
Mouse: Mouse Embryonic fibroblasts *Smug1*$^{-/-}$	Lab made	Alsøe et al., 2017
Mouse: Mouse Embryonic fibroblasts *Smug1*$^{-/-}$_SMUG1	Lab made	N/A
Mouse: Mouse Embryonic fibroblasts *Smug1*$^{-/-}$_H241L	Lab made	N/A
Mouse: Mouse Embryonic fibroblasts *Smug1*$^{-/-}$_E31R/E35R	Lab made	N/A
Human: HAP1 SMUG1-WT (control wild type cell line)	Horizon	Cat# HZGHC003300c009
Human: HAP1 SMUG1-KO	Horizon	Cat# HZGHC003300c009
Human: HAP1 SMUG1-KO_SMUG1	Lab made	N/A
Human: HAP1 SMUG1-KO_SMUG1 H239L (NABm)	Lab made	N/A
Human: HAP1 SMUG1-KO_E29R/E33R/E231R (DBm)	Lab made	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Human: HeLa	ATCC	ATCC® CCL-2
Experimental Models: Organisms/Strains		
C56BL/6 Smug1^{tm1Hln} (*Smug1*$^{-/-}$)	Lab made	Alsøe et al., 2017
Oligonucleotides		
See Table S1 for oligonucleotide information		N/A
Recombinant DNA		
Plasmid: pHH25-Smug1	Cloned and stored in lab	N/A
Plasmid: pHH25-Smug1 H241L	Cloned and stored in lab	N/A
Plasmid: pHH25-Smug1 E31R/E35R	Cloned and stored in lab	N/A
Plasmid: pHH25-SMUG1	Cloned and stored in lab	N/A
Plasmid: pHH25-SMUG1 H239L (NABm)	Cloned and stored in lab	N/A
Plasmid: pHH25-SMUG1 E29R/E31R/E231R (DBm)	Cloned and stored in lab	N/A
Plasmid: pCDNA-3xHA-hTERT	Addgene	Cat# 51631; RRID:Addgene_51637
Plasmid: pBS U3-hTR-500	Addgene	Cat# 28170; RRID:Addgene_28170
Plasmid: pETM-11-hSMUG1 wt	Cloned and stored in lab	N/A
Plasmid: pETM-11-hSMUG1 H239L	Cloned and stored in lab	N/A
Software and Algorithms		
Multi Gauge V3.1 software	Fujifilm	N/A
STAR aligner	Dobin et al., 2013	https://github.com/alexdobin/STAR; RRID:SCR_015899
htseq-count	Anders et al., 2015	https://github.com/simon-anders/htseq; RRID:SCR_011867
Voom	Law et al., 2014	http://bioinf.wehi.edu.au/voom/
Limma	Ritchie et al., 2015	https://bioconductor.org/packages/release/bioc/html/limma.html; RRID:SCR_010943
ImageJ		https://imagej.net/Welcome; RRID:SCR_003070
bowtie2	Langmead and Salzberg, 2012	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
Bedtools	Quinlan and Hall, 2010	https://github.com/arq5x/bedtools2; RRID:SCR_006646
TeloTool	Göhring et al., 2014	https://github.com/jagoehring/TeloTool
Zeiss Zen blue and black	Zeiss	https://www.zeiss.com/microscopy/en_us/products/microscope-software/zen.html; RRID:SCR_013672
GraphPad Prism	GraphPad	https://www.graphpad.com/; RRID:SCR_002798
EZLogic Integration	Bio-Rad	http://www.bio-rad.com

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and request for reagents may be directed to and will be fulfilled by the Lead Contact, Hilde Nilsen (h.l.nilsen@medisin.uio.no).

Plasmids and human and mouse lines generated in this study are stored in the lab biobank and are available under request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Animals

Wild-type and *Smug1*$^{-/-}$ C57BL/6J (male, 3- and 12-months old) mice were used for all the experiments. All mice were used straight after housing them until the appropriate experimental age. Animal maintenance, mouse handling and experimental procedures

were performed in accordance to institutional guidelines and procedures approved by the Animal Experimentation Administration in Norwegian Food Safety Authority (NFDA). The generation of *Smug1*$^{-/-}$ C57BL/6J mouse model was described previously (Alsøe et al., 2017).

Cell Lines

Primary and transformed wild-type and *Smug1*$^{-/-}$ mouse embryonic fibroblasts (MEFs) were generated from the gene-targeted *Smug1*$^{-/-}$, *Smug1*^{tm1Hln}, C57BL/6J mouse model previously established in our lab (Alsøe et al., 2017). Timed matings were set up between either wild-type or *Smug1*$^{-/-}$ mice born from heterozygous parents in order to obtain wild-type and *Smug1*$^{-/-}$ embryos. Primary mouse embryonic fibroblasts (MEFs) were generated from ED13.5 to ED14.5 mouse tissue of wild-type and *Smug1*$^{-/-}$ embryos. Limbs were removed from embryos, the tissue was chopped into small pieces and cell suspension of all embryos deriving from one female was made using a pipette. Transformed wild-type and *Smug1*$^{-/-}$ MEFs were obtained by spontaneous transformation of the primary cultures. MEFs were grown in Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12, GlutaMAX (Invitrogen) supplemented with 10% (vol/vol) fetal bovine serum (FBS, Lonza), 1x Penicillin-Streptomycin (Invitrogen) and 1x MEM non-essential amino acids (Invitrogen). Primary MEFs (Passage 1) and cells cultured for 22 continuous passages were used. The MEF genotypes were authenticated by PCR genotyping for Smug1$^{-/-}$ and SNPs for confirming C57BL/6J strain.

HAP1 cells were edited by CRISPR/Cas9 to contain a 2bp deletion in a coding exon of SMUG1. HAP1 SMUG1-WT and SMUG1-KO cells (Horizon) were maintained in culture as predominantly diploid cells in Iscove's modified Dulbecco's medium (IMDM, Life Technologies) containing 10% (vol/vol) FBS and 1x Penicillin/Streptomycin.

HeLa cells were purchased from ATCC and cultivated in Dulbecco's Modified Eagle Medium, GlutaMAX (Invitrogen) supplemented with 10% (vol/vol) fetal bovine serum (FBS, Lonza) and 1x Penicillin-Streptomycin (Invitrogen). HeLa cells were derived from female tissue while HAP1 cells were from a male cell line.

Smug1$^{-/-}$ MEF and HAP1 SMUG1-KO cells were complemented with different SMUG1 constructs cloned into the pHH25 vector. Transfection agents were FuGENE 6 (Promega) for MEF cells and Lipofectamine 3000 (Invitrogen) for HAP1 cells. Subsequently, stable cell lines re-expressing SMUG1 were selected using 2 μg/ml puromycin for MEFs and 1 μg/ml for HAP1 cells.

All the cell lines were cultured at 37°C and 5% CO_2.

METHOD DETAILS

Cell Line Treatments

HeLa cells transfected with Flag-tagged SMUG1 were synchronized at early S-phase by a double thymidine block (Banfalvi, 2017). Briefly, 24 h after seeding, cells were treated with 2 mM thymidine (Sigma-Aldrich) in complete medium for 14 h, washed twice in PBS and released in complete medium for 9 h. Then, cells were subjected to a second thymidine block (2 mM) for 14 h, washed twice in PBS and released in complete medium before fixation in PFA 4% at different time points.

To synchronize cells in S-phase, MEFs were seeded onto coverslips in a 24-well dish at 40% confluency. After 12 h, cells were serum starved for 42 h in DMEM supplemented with 0.5% FBS. Subsequently, cells were washed with 1x PBS and incubated in regular DMEM for another 24 h to re-enter the cell cycle. After 24 h, cells were given 1 h pulse with 20 μM bromodeoxyuridine (BrdU, BD Biosciences).

To inhibit RNA polymerase II, cells were treated for 2 h with 5 μg/ml of Actinomycin D (Sigma-Aldrich). For IR treatment, HAP1 cells were seeded onto coverslips in a 24-well cell culture dish and irradiated with 2 Gy. The next day, immunofluorescence experiments were conducted.

DNA and siRNA Transfections

For overexpression and siRNA experiments, HAP1 cells were seeded onto 6-well plates or 10 cm dishes and transfected 24 h later. The constructs used for *hTERT* and *hTERC* overexpression were pCDNA-3xHA-hTERT and pBS-U3-hTR-500 (Addgene). For siRNA experiments, either a scrambled control siRNA or the target-specific siRNA was used. Lipofectamine 3000 and Lipofectamine RNAiMAX (Invitrogen) transfection reagents were used as per manufacturer's indications. The target-specific siRNAs for *PARN* and *EXOSC10* were purchased from Dharmacon as siGENOME SMARTpools; scrambled control and *DKC1* siRNAs were purchased from Ambion. Cells were harvested 24 h later and siRNA treated 48, 72 or 96 h after transfection.

Immunofluorescence

Cells were seeded onto coverslips, fixed either in 70% ice cold ethanol for 10 min or with 4% paraformaldehyde at RT for 15 min and permeabilized with PBS containing 0.1% Triton X-100 at 4°C for 15 min. Cells were washed with 1x PBS and blocked for 1 h in PBS-BT solution (1x PBS, 3% BSA, 0.1% Triton X-100 and 0.05% sodium azide). Prior to blocking, antigen retrieval was necessary for BrdU labeling. Cells were incubated for 10 min on ice with 1 N HCl, 0.5% Triton X and 20 min 2 N HCl, 0.5% Triton X-100. Neutralization was followed, with 0.1 M sodium borate buffer pH 8.5. Cells were washed twice for 5 min in 1x PBS and incubated overnight at 4°C with primary antibodies in blocking solution. Cells were washed twice for 5 min in 1x PBS and secondary antibodies were added for 1 h (1:1000 dilution in PBS-BT). Cells were washed three times 5 min with 1x PBS. Coverslips were air-dried for 10 min, protected from light, and mounted onto ethanol rinsed glass slides using Prolong Diamond Antifade mounting medium containing DAPI and left

at room temperature overnight. Images were acquired using a Zeiss LSM780 confocal microscope with a 63x objective and analyzed using the Zeiss Zen Blue software.

Proximity Ligation Assay (PLA)

Synchronized Flag-tagged SMUG1 cells fixed after 4 h of release were used for Figure 1C. After fixation with PFA 4% for 20 min at RT, cells were permeabilized for 5 min in PBS 0.25% (vol/vol) Triton X-100. Cells were incubated in blocking solution (FBS 10% in TBS 0.1% [vol/vol] Tween-20) for 1 h at RT. Incubation with primary antibodies (anti-Flag (Sigma-Aldrich) and RNAPII (Abcam) or anti-Coilin (Cell Signaling Technology) diluted 1:200 in blocking solution) was carried out ON at 4°C. After three washes in PLA Washing buffer A, PLA was performed following the manufacturer's protocol. Briefly, PLUS and MINUS PLA probes were diluted 1:5 in Duolink® Antibody diluent and added to the coverslips for 1 h at 37°C. Cells were washed twice in PLA Washing buffer A and the ligation step (ligase diluted 1:40 in Ligation buffer 1x) was carried out for 30 min at 37°C followed by amplification (Polymerase diluted 1:80 in Amplification buffer 1x) for 100 min at 37°C. Coverslips were washed twice in PLA Washing buffer B for 10 min each, in PLA Washing buffer A for 1 min and counterstained for Coilin/RNAPII, incubating the cells with Alexa Fluor 488 conjugated anti-rabbit (Life Technologies) for 2 h at RT. Cells were washed twice in PLA Washing buffer A for 2 min, rinsed with PLA Washing buffer B 0.01x and mounted with Prolong Diamond Antifade mounting medium. Technical control, represented by the omission of the anti-Flag antibody, resulted in loss of PLA signal.

Antibodies

Primary antibodies used for immunofluorescence were DKC1 (1:500 and 1:50, Bethyl laboratories), BrdU (1:100, Abcam), Coilin (1:500, Cell Signaling Technology), Flag (1:200, Sigma-Aldrich), RNAPII (1:200, Abcam), TRF1 (1:500) and γH2Ax (1:500, Millipore). Secondary antibodies for immunofluorescence were purchased from Life Technologies: Alexa Fluor 594 conjugated goat-anti-rat/rabbit antibodies and Alexa Fluor 488 conjugated goat-anti-rabbit/mouse antibodies. Immunoblotting and ChIP/RIP experiments were carried out using the following antibodies: SMUG1 (1:2000 and 5 μg, Abcam), RNAPII (1:2000 and 5 mg, Abcam), PARN (1:2000, Abcam), EXOSC10 (1:2000, Abcam), HA (1:2000, Cell Signaling Technology), TRF2 (1:2000 and 5 μg, Novus Biologicals), POT1 (1:2000 and 5 μg, Abcam), DKC1 (5 μg, Bethyl Laboratories), TERT (1:2000, Abcam). Either GAPDH (1:3000, Cell Signaling Technology) or a-Tubulin (1:3000, Sigma-Aldrich) were used as loading controls.

Western Blot

Whole-cell lysates were prepared in RIPA buffer [10 mM Tris-HCl (pH 8.0), 140 mM NaCl, 1 mM EDTA (pH 8.0), 0.5 mM EGTA, 0.1% SDS (wt/vol), 0.1% sodium deoxycholate (wt/vol) and 1% Triton X-100 (vol/vol)] containing protease inhibitors (Sigma-Aldrich). Protein extracts were run on any kD Mini-PROTEAN TGX precast gel (Bio-Rad) and blotted on nitrocellulose or PVDF membranes. Blots were blocked in either 5% non-fat milk or 5% BSA dissolved in 1x PBS, 0.1% Tween-20 (blocking solution). After the incubation with the specific primary antibody, secondary antibody incubation was carried out for 1 h (1:3000 in blocking solution) at RT. Blots were developed with SuperSignal West Pico or Femto Chemiluminescent substrate (Thermo Scientific). The signals were detected with a LAS-3000 mini imaging system (FujiFilm) and quantified with Multi Gauge V3.1 software.

hTERC *In Vitro* Transcription

cDNA from HAP1 SMUG1-WT cells was used as a template to amplify *hTERC* for the *in vitro* transcription reaction. The primers used were: T7_hTERC_F, 5′-CCAAGCTTCTAATACGACTCACTATAGGGAGAGGGTTGCGGAGGGTGGGCCT-3′ and T7_hTERC_R, 5′-GCATGTGTGAGCCGAGTCCTGG-3′. PCR-amplified DNA was subsequently used as template to transcribe *hTERC* RNA *in vitro* by using the MEGAscript T7 transcription kit (Life Technologies), following the manufacturer's instructions.

Expression and Purification of SMUG1

E. coli BL21(DE3) cells harboring pETM-11-hSMUG1 WT or H239L mutants were grown in LB media with kanamycin (50 μg/ml) at 37°C until OD_{600} reached 0.6. The protein expression was induced with IPTG at 37°C for 2 h (final concentration 0.25 mM). The following procedures were performed at 4°C. Cells pelleted from 1 l culture were resuspended in 15 mL buffer A [50 mM Tris-HCl pH 8.0, 300 mM NaCl, 10 mM β-mercaptoethanol], disrupted by sonication (4 × 30 s at 60% amplitude), centrifuged (15000 rpm, 30 min), and incubated with 3 mL Ni-NTA agarose (QIAGEN) for 1 h with light agitation before loading onto an Econo-column (Bio-Rad; 2.5 × 20 cm). The column was washed with buffer A (50 ml); the fusion protein was eluted with buffer A containing 50 mM Imidazole (3 × 5 ml) and buffer A containing 150 mM Imidazole (3 × 5 ml). For hSMUG1 H239L, the eluted fractions were pooled and concentrated, glycerol was added (final concentration 25%) and stored at −20°C. For hSMUG1 WT, the eluted fractions were dialyzed against buffer B [20 mM MES pH 6.5, 50 mM NaCl, 10 mM β-mercaptoethanol] overnight. The fraction was centrifuged to remove the precipitate and concentrated to 10 ml, and loaded onto a Bio-Scale Mini Macro-Prep High S cartridge (Bio-Rad, 5 ml) onto the BioLogic DuoFlow 10 System (Bio-Rad). The column was washed with Buffer B and eluted with a linear gradient of NaCl (50-2000 mM) in buffer B. The fractions containing hSMUG1 were pooled and concentrated, glycerol was added (final concentration 25%) and stored at −20°C. In order to remove potential aggregated or degraded protein before His-tag and SMUG1 pulldown assays, diluted protein (1:10 in buffer A containing 10 mM Imidazole) was loaded onto 200 μL of pre-equilibrated Ni-NTA resin into Ultrafree-MC device (Sigma-Aldrich, 0.5 ml) and incubated for 30 min with light agitation. The resin was washed twice with buffer A containing 20 mM Imidazole (500 μl) and the protein eluted in 200 μL buffer A containing 250 mM Imidazole.

His-tag and SMUG1 Pulldown

For the His-tag and SMUG1 pulldown experiments, 0.2 nmol of either full-length Hist-tag SMUG1 WT or SMUG1 H239L mutant was added, together with total RNA isolated from HAP1 SMUG1-WT cells (300 μg per reaction) or *in vitro* transcribed *hTERC* (5 μg per reaction) to 10 μL of HIS-Select® Nickel Magnetic Agarose beads (Sigma-Aldrich) or 20 μL Protein G magnetic beads (Life Technologies) conjugated with SMUG1 antibody (Abcam). As negative control, prior performing the assay, SMUG1 WT protein was heat inactivated at 65°C for 20 min. Binding was performed in PBS completed with protease inhibitor cocktail 1x (Sigma-Aldrich) for 3 h at 4°C under rotation. The beads were washed three times in washing buffer (PBS supplemented with 1% Igepal CA-630) and resuspended in Trizol (Invitrogen) or Laemmli sample buffer for RNA isolation and western blotting, respectively.

RNA Isolation and qPCR

Total RNA was isolated with either RNeasy kit (QIAGEN) or with Trizol (Invitrogen) following the manufacturer's instructions. Reverse transcription was performed using the cDNA synthesis kits (Bio-Rad). Quantitative PCR was carried out on a QuantStudio 7 Flex detection system (Applied Biosystems) with Power SYBR green PCR master mix (Applied Biosystems). Each sample was analyzed in triplicate.

EU Incorporation, Quantification of Nascent hTERC Kinetics, and hTERC Decay

Detection of nascent *hTERC* and its decay were analyzed by using the Click-iT® Nascent RNA Capture kit (Life Technologies). 5-ethynyl Uridine (EU) (Life Technologies) was dissolved in DMSO at a concentration of 200 mM. HAP1 cells were incubated with 0.2 mM EU for different time points (1, 4, 8 and 24 h) for capturing the nascent *hTERC* or pulsed for two hours and chased for 4 and 24 h for analyzing the *hTERC* decay. Cell pellets were harvested and total RNA prepared using Trizol reagent (Invitrogen). The biotinylation, the streptavidin binding and cDNA synthesis were performed as per manufacturer's instructions. qPCR analysis was performed using the standard protocol.

RNA Sequencing and Analysis

RNA integrity was verified using the 2100 Bioanalyzer instrument. RNAs, spiked with control RNAs (Lexogen SIRV-Set 3) during cell lysis prior to RNA isolation, were submitted to the Genomic Core Facility (NTNU) for library preparation (Lexogen SENSE Total RNA-Seq Library Prep Kit) and sequencing (76 nucleotide paired end sequencing on an Illumina NextSeq 500 High Output flow cell). Sequence reads were aligned to the human genome (version GRCh38.p7) and to the control RNA sequences with the STAR aligner (Dobin et al., 2013). Read counts per gene (Encode release 84) were determined with htseq-count (Anders et al., 2015). The gene count matrix was normalized with voom (Law et al., 2014), using normalization factors computed with the TMM method (Robinson and Oshlack, 2010) from the reads aligning to the control RNAs. The normalized matrix was analyzed for differential gene expression with limma (Ritchie et al., 2015). Data is deposited in GEO: GSE116580.

hTERC RNA Damage Assay

RNA isolated from HAP1 cells was digested with 2 U of SMUG1 (New England Biolabs) at 37°C for 30 min and then retro-transcribed with SuperScript IV RT (Invitrogen) following the manufacturer's instructions. Then, the cDNA was assayed via ddPCR using Droplet Digital PCR QX system (Bio-Rad). Briefly, the cDNA was added to a 20 μL PCR mixture containing 10 μL 2x QX200 ddPCR EvaGreen Supermix (Bio-Rad) and 100 nM *hTERC* specific primers. 20 μL of PCR mixture and 70 μL Droplet generation oil for EvaGreen (Bio-Rad) were mixed. Droplets were generated using a QX100 Droplet Generator (Bio-Rad). The following PCR conditions were used: after denaturing at 95°C for 5 min, 40 cycles at 95°C for 30 s and 60°C for 1 min were followed by 1 step at 4°C for 5 min and 90°C for 5 min. Reactions were read in the QX200 Droplet Reader (Bio-Rad).

Chromatin Immunoprecipitation (ChIP) and re-ChIP

Chromatin immunoprecipitation was performed essentially as described by Dahl and Collas (2008). For re-ChIP experiments, the elution for the first immunoprecipitation (RNA polymerase II antibody) was performed in TE-SDS 0.5%, 10 mM DTT for 30 min at 37°C. 10% of the eluted chromatin was then retained as the primary ChIP. The remaining 90% was diluted 20 times in RIPA buffer (Dahl and Collas, 2008) and incubated overnight at 4°C with SMUG1 or IgG (isotype control) antibodies. Immunoprecipitates were then processed as described (Dahl and Collas, 2008). The purified DNA was analyzed via qPCR. Fold enrichment as percent of input was calculated by normalizing ChIP reactions to input DNA of the target gene.

Telomere ChIP

Telomere ChIP (TeloChIP) analysis was carried out essentially as described previously (Grolimund et al., 2013). Briefly, cells were washed twice in PBS and cross-linked in 1% formaldehyde-1x PBS for 10 min at room temperature. Glycine (pH 2.5) was added to 125 mM in order to quench the reaction before washing the cells twice with PBS. Cells were lysed in lysis buffer [1% SDS, 10 mM EDTA (pH 8.0), 50 mM Tris-HCl (pH 8.0), EDTA-free protease inhibitor complex] and incubated for 5 min at room temperature, centrifuged for 5 min at 1,500 x*g*, washed once in lysis buffer and centrifuged as above. The chromatin-enriched pellets were resuspended in lysis buffer and sonicated for 25 cycles with 30 s ON and 30 s OFF per cycle using a Bioruptor (Diagenode). The sonicated lysate was centrifuged at 4°C for 15 min at 20,000 x*g*. The supernatant was diluted in 2 volumes of ChIP dilution buffer [0.75% Triton

X-100, 10 mM EDTA (pH 8.0), 50 mM Tris-HCl (pH 8.0), 600 mM NaCl]. The lysate was incubated with POT1, TRF2, SMUG1 antibodies or normal rabbit IgG covalently coupled to Protein G dynabeads (Invitrogen). The ChIP was performed at 4°C overnight. Beads were washed once with wash buffer 1 [0.1% SDS (wt/vol), 1% Triton X-100 (vol/vol), 2 mM EDTA (pH 8.0), 20 mM Tris-HCl (pH 8.0), 300 mM NaCl], wash buffer 2 [0.1% SDS (wt/vol), 1% Triton X-100 (vol/vol), 2 mM EDTA (pH 8.0), 20 mM Tris-HCl (pH 8.0), 500 mM NaCl], wash buffer 3 [500 mM LiCl, 1% NP-40 (vol/vol), 1% Na-deoxycholate (wt/vol), 1 mM EDTA (pH 8.0), 10 mM Tris-HCl (pH 8.0)] and twice with wash buffer 4 [1 mM EDTA (pH 8.0), 10 mM Tris-HCl (pH 8.0)]. DNA–protein complexes were eluted with 2.5 bead volumes of Elution buffer [20 mM Tris–HCl (pH 7.5), 5 mM EDTA (ph 8.0), 50 mM NaCl, 20 mM Na-butyrate, 1% SDS (wt/vol), 50 μg/ml proteinase K]. DNA-protein complexes were incubated for 2 h at 68°C, 1,300 xg. The DNA was extracted with phenol–chloroform–isoamylalcohol procedure and telomeric DNA sequences were analyzed using qPCR. Data are analyzed as percentage of input for each target gene. The primer sequences were previously described (O'Callaghan and Fenech, 2011).

RNA Coimmunoprecipitation Assay

RNA immunoprecipitation assay was performed as described previously (Jobert et al., 2013) with minor modifications. Briefly, cells were washed twice in PBS and cross-linked in 1% formaldehyde-1x PBS for 10 min at room temperature. Glycine (pH 2.5) was added to 0.2 M in order to quench the reaction before washing the cells twice with ice-cold PBS. Cells were lysed in lysis buffer A [50 mM HEPES (pH 7.8), 1 mM EDTA (pH 8.0), 1% Triton X-100 (vol/vol), EDTA-free protease inhibitor complex] and sonicated for 10 cycles with 30 s ON and 30 s OFF per cycle using a Bioruptor (Diagenode). The sonicated lysate was diluted in 1 volume of lysis buffer B [50 mM HEPES (pH 7.8), 1 mM EDTA (pH 8.0), 50 mM $MgCl_2$, 10 mM $CaCl_2$, 0.4 U/μl RNaseOUT recombinant ribonuclease inhibitor (Invitrogen)]. DNA was digested with DNase I RNase free (Life Technologies) at 37°C for 15 min and digestion was stopped adding 20 mM EDTA (pH 8.0). The lysate was centrifuged at 4°C for 5 min at 20,000 x*g*. The supernatant was incubated with SMUG1, DKC1 or normal rabbit IgG covalently coupled to Protein G dynabeads (Invitrogen). The RIP was performed at 4°C overnight. Beads were washed once with Binding buffer [50 mM HEPES (pH 7.8), 20 mM EDTA (pH 8.0), 0.5% Triton X-100 (vol/vol), 25 mM $MgCl_2$, 5 mM $CaCl_2$], FA500 buffer [50 mM HEPES (pH 7.8), 1 mM EDTA (pH 8.0), 1% Triton X-100 (vol/vol), 500 mM NaCl, 0.1% Na-deoxycholate (wt/vol)], LiCl buffer [10 mM Tris-HCl (pH 7.5), 1% Triton X-100 (vol/vol), 1 mM EDTA (pH 8.0), 250 mM LiCl, 0.5% Na-deoxycholate (wt/vol)] and TES buffer [10 mM Tris-HCl (pH 7.5), 1 mM EDTA (pH 8.0), 10 mM NaCl]. RNA–protein complexes were eluted twice with 2.5 bead volumes of Elution buffer [100 mM Tris-HCl (pH 7.8), 10 mM EDTA (pH 8.0), 1% SDS (wt/vol)] for 10 min at 37°C. RNA-protein complexes and input samples were reverse-crosslinked with 200 mM NaCl for 1 h at 65°C and incubated at 42°C for 1 h after adding 20 μg proteinase K. The RNA was extracted with Trizol solution (Invitrogen) and analyzed by qPCR as percentage of input.

Northern Blot

Oligonucleotide probes were labeled using the DIG Oligonucleotide 3′ End Labeling Kit, 2nd generation (Roche) following the manufacturer's instructions.

Northern blot analyses were carried out as described previously (Xi and Cech, 2014). Briefly, 5 μg of RNA samples were mixed with equal volume of 2x formamide loading dye, heated at 95°C for 5 min and then run on a 8% polyacrylamide/7 M urea/1x TBE at 60 V for 1 h. Membranes were cross-linked at 365 nm for 20 min in a UV crosslinker (Stratalinker). The membrane was pre-hybridized in DIG Easy Hyb Granules (Roche) at 42°C for 1 h and then hybridized in DIG Easy Hyb Granules with 3′ end DIG labeled oligo probes at 42°C overnight. The membrane was washed twice in prewarmed stringent wash buffer I (2x SSC, 0.1% SDS (wt/vol)) and stringent wash buffer II (0.2x SSC, 0.1% SDS (wt/vol)) at 50°C. After that, the membrane was blocked for 1 h at room temperature with 1x blocking solution (Roche) (1x blocking solution, 1x maleic acid buffer) and then incubated for 2 h with anti-DIG-AP antibody diluted in 1x blocking solution. The membrane was washed twice in 1x washing solution (Roche) for 15 min and then incubated with the 1x detection buffer (Roche) for 2 min and the substrate solution (CDP Star, Roche) for 5 min. Signals were detected with a LAS-3000 mini imaging system (FujiFilm) and quantified with ImageJ software. Probes for *hTERC* and *H1* were previously described (Boyraz et al., 2016; Moon et al., 2015; Nguyen et al., 2015).

3′ RACE

3′ RACE experiments were carried out as described previously (Moon et al., 2015). Briefly, 600 ng of total RNA were added in a 20-μl reaction containing 5 μM of Universal miRNA Cloning Linker (New England Biolabs), 280 U of T4 RNA ligase, Truncated KQ (NEW ENGLAND BIOLABS), 25% PEG8000 and 1 μl of RNaseOUT (Life Technologies) and incubated at 25°C for 16 h. RNA was cleaned using RNA Clean and Concentrator columns (Zymo Research), digested with DNase I (Life Technologies) at RT for 15 min and retro-transcribed with 5 pmol of Universal RT primer (5′-CTACGTAACGATTGATGGTGCCTACAG-3′) using SuperScript III RT (Invitrogen). PCR amplification was performed with 5 μM of TERC_L2 and universal RT primer set using AccuPrime Pfx SuperMix (Invitrogen) following the manufacturer's instructions. PCR reactions were separated onto 2.5% agarose gel.

3′ RACE Library Preparation and Analysis

3′ RACE products were prepared for sequencing using the TruSeq Nano DNA LT Library Prep kit (Illumina), following the manufacturer's instructions. The pooled completed libraries were sequenced on an Illumina MiSeq instrument with paired-end 250-base reads using the Illumina TruSeq v2 500 Cycles kit. For data analysis, paired end reads were aligned to the *hTERC* gene sequence (UCSC gene ID uc003ffr.2) by running bowtie2 (Langmead and Salzberg, 2012) in local alignment mode. Read pairs aligning to

the *hTERC* gene were further filtered by requiring that the reads mapped to opposite strands of the *hTERC* gene and that the reads represented plausible *hTERC* RNA fragments instead of primer artifacts. Specifically, we only kept read pairs where the estimated template lengths (unsigned value of SAM field 9) was identical and > 0, and where the signed template length (SAM field 9) had opposite signs for the read pair. We used the bedtools software (Quinlan and Hall, 2010) to compute read coverage. For the poly(A) analyses, we further filtered the *hTERC*-aligning reads by only keeping the rightmost-aligning read of each read pair that also contained the sequence corresponding to the canonical *hTERC* 3′ end (CAGGACTCGGCTCACACATGC). For each such read, we removed the 3′ adaptor sequence (by using cutadapt) and the 5′ part of the read that matched the *hTERC* gene sequence (up to 8 nucleotides downstream of the canonical *hTERC* 3′ end), and only kept those with at least one adenine and at most 25% other nucleotides within the remaining sequence. The length of these remaining sequences was the estimated poly(A) tail lengths.

Bone Marrow Cell Isolation and Culture

Femurs and tibias were isolated and muscles were removed before the bones were rinsed with 70% ethanol and ice-cold HBSS under sterile conditions. For each bone, and upon separation of femurs from tibias, epiphyses were cut and bone marrow cells were flushed with 10 mL HBSS with a 26-gauge needle and passed through a 100 μm cell strainer (Corning). The inner surface of each bone was scraped with a needle. Cells were then spun at 300 x*g* for 15 min. Cell pellets were resuspended with 5 mL 1x RBC lysis buffer (eBioscience) and incubated at RT for 4 min. Lysis was stopped by adding 25 mL of ice-cold HBSS. Cells were then spun at 400 x*g* for 15 min at 4°C, resuspended in 10 mL of ice-cold HBSS and passed through a 30 μm pre-separation filter (Miltenyi Biotec). Upon cell counting, cells were spun at 300 x*g* for 15 min at 4°C and resuspended at 0.5×10^6 cells/ml in IMDM (GIBCO) supplemented with 20% fetal calf serum (Hyclone, Logan, UT) in the presence of interleukin 6 (200 ng/ml, Affymetrix eBioscience) and stem cell factor (100 ng/ml; Affymetrix eBioscience). Cells were split every two days until passage 3.

Bone marrow cell (BMC) colony formation assay was performed by using methylcellulose-based media (R&D systems) according to the manufacturer's instructions. Twentyfive thousand BMCs were seeded in triplicates for each BMC isolation at day 0, and colonies were counted at day 11. Three different mice per genotype and generation were used.

Telomeric PNA FISH

Cells were treated with 0.1 μg/ml Colcemid for 3 h prior to harvest and centrifuged at 300 x*g* for 5 min. Pellets were carefully resuspended in 7 mL of prewarmed 75 mM KCl. Tubes were placed in a 37°C water bath for 15 min and immediately centrifuged at 120 x*g* for 5 min. Pellets were resuspended in 2 mL of freshly prepared methanol:glacial acetic acid (3:1) fixative. To obtain metaphases, fixed cells were dropped onto slides and dried. Fluorescent *in situ* hybridization (FISH) was performed as earlier described (Zimmermann et al., 2014) with minor modifications. Briefly, slides were rehydrated in PBS (pH 7.0-7.5) for 5 min, fixed in 4% formaldehyde in PBS for 2 min, washed in PBS 3 × 5 min, treated at 37°C with freshly made 1 mg/ml pepsin in 10 mM glycine (pH 2), washed twice in PBS for 2 min, and dehydrated for 5 min in increasing concentrations of ethanol (70%, 95%, and 100%). After air-drying, 10 μL hybridization mix (10 mM Tris-HCl pH 7.2, 70% formamide, 0.5% blocking reagent (Roche) and probes (TelG-TAMRA or FITC-TelC, Biosynthesis) were applied. After addition of coverslips, the slides were heated on a 80°C hot plate for 3 min. Metaphases were hybridized for 2 h at RT in the dark in a humid chamber, washed two times 15 min with washing solution #1 (10 mM Tris-HCl pH 7.2, 70% formamide, 0.1% BSA), three times 5 min with washing solution #2 (0.1 M Tris-HCl pH 7.2, 0.15 M NaCl, 0.08% Tween-20), and dehydrated as described above. Slides were air-dried for 15 min and mounted with DAPI-containing mounting medium (VectaShield).

Telomerase Activity

Telomerase activity was conducted either by using the qPCR or the ddPCR method. Cells were lysed in CHAPS buffer (TRAPeze kit, Millipore) for 30 min on ice. Protein determination of lysates was conducted using a BCA method. qPCR method was performed using the TRAPeze kit according to manufacturer's instructions. The protein amount used was 500 ng. The ddPCR method was used as previously described (Ludlow et al., 2014). Briefly, 10 μL of lysate, with concentrations 30-90 ng/μL, was added to a 50 μL extension reaction containing 1x TRAP reaction buffer (10x concentration: 200 mM Tris-HCl, pH 8.3, 15 mM $MgCl_2$), 0.4 mg/ml BSA, TS telomerase extension substrate (HPLC purified, 200 nM, dNTPs (2.5 mM each) and incubated for 40 min at 25°C then held at 4°C. The ddPCR reaction was assembled containing 1x EvaGreen ddPCR Supermix v2.0 (Bio-Rad), 50 nM TS primer, 50 nM ACX primer, 10 μL of extension product on a final volume of 25 μL. The lysis-extension mixture was subsequently used for the standard ddTRAP protocol. Droplets were generated according to the manufacturer's instructions and transferred to a 96-well PCR plate. The PCR program used was 95°C for 5 min, followed by 40 cycles of 95°C for 30 s, 54°C for 30 s, 72°C for 30 s, then held at 12°C. Following PCR, the fluorescence was read on the droplet reader using the 6-Fam channel.

Telomere Length Analysis via qPCR

Genomic DNA (gDNA) was extracted from cells and tissues using the DNeasy Blood & Tissue Kit (QIAGEN). Average telomere length was measured from total gDNA samples by using the qPCR method (Anders et al., 2015) with minor modifications. The single-copy genes β-globin and 36B4 were used as references for human and mouse material, respectively. Each reaction included 5 μL 2x Power SYBR Green PCR Master mix (Applied Biosystems), 0.1 μL 10 μM telo forward and 0.9 μL 10 μM telo reverse primers, 2 μL water and 2 μL gDNA (5 ng/μL for human and 0.5 ng/μL for mouse) to yield a 10-μL reaction. Each single copy gene reaction included 5 μL 2x Power SYBR Green PCR Master mix (Applied Biosystems), 0.3 μL of 10 μM single copy gene forward and 0.7 μL 10 μM single copy

gene reverse primers, 2 μL water and 2 μL gDNA (5 ng/μL) to yield a 10-μL reaction. A QuantStudio 7 Flex Real-Time PCR System (Applied Biosystems) was used with reaction conditions of 95°C for 10 min followed by 40 cycles of data collection at 95°C for 15 s, 60°C anneal for 30 s, and 72°C extend for 45 s.

Telomeric DNA Damage Analysis

gDNA (200 ng) isolated from mice tissues was digested with 10 U of SMUG1 (New England Biolabs) in a final concentration of 10 μL, at 37°C for 30 min. 5 μL of the reaction were further digested with 10 U of APE1 (New England Biolabs), at 37°C for 30 min. The standard protocol for telomere length analysis via qPCR was then conducted, using 1.5 ng of digested gDNA. Results were normalized to the uncut gDNA qPCR reaction.

TRF Assay

Absolute telomere length analysis was carried out using the TeloTAGGG telomere length assay (Roche) according to the manufacturer's instructions with minor modifications. Briefly, 5 −10 μg of gDNA were digested with RsaI (20 U/reaction) and HinfI (20 U/reaction) restriction enzymes on a final volume of 20 μL, leaving the telomeric and subtelomeric sequence unaffected. In order to evaluate the presence of SMUG1 substrates in telomeric DNA, SMUG1 (5 U/reaction) and APE1 (1 U/reaction) enzymes (New England Biolabs) were added together with RsaI (10 U/reaction) and HinfI (10 U/reaction) on a final volume of 20 μL, using the SmartCut (New England Biolabs) digestion buffer. The digested gDNA was separated using a 0.8% standard gel electrophoresis, transferred via a semi-dry method to a Hybond XL (GE Healthcare) membrane, and hybridized using a DIG-labeled $(TTAGGG)_3$ oligonucleotide probe. Images were acquired on LAS-3000 mini imaging system (FujiFilm), and quantification was performed using the TeloTool software.

QUANTIFICATION AND STATISTICAL ANALYSIS

All quantified data are presented as mean ± s.e.m., mean ± s.d. and fold change unless stated otherwise (refer to figure legend to detailed information). Student t test or one-way ANOVA were used to assess statistical significance in GraphPad Prism 7 (GraphPad software). A $p < 0.05$ was considered as statistical significant. P values were indicated with asterisks. Replicates, statistical tests carried out and statistical significances are reported in the corresponding figure legends.

DATA AND CODE AVAILABILITY

Raw and analyzed RNA-seq and 3′ RACE-seq data have been deposited in the NCBI GEO database under accession number GEO: GSE116580.

Cell Reports, Volume *28*

Supplemental Information

SMUG1 Promotes Telomere Maintenance through Telomerase RNA Processing

Penelope Kroustallaki, Lisa Lirussi, Sergio Carracedo, Panpan You, Q. Ying Esbensen, Alexandra Götz, Laure Jobert, Lene Alsøe, Pål Sætrom, Sarantis Gagos, and Hilde Nilsen

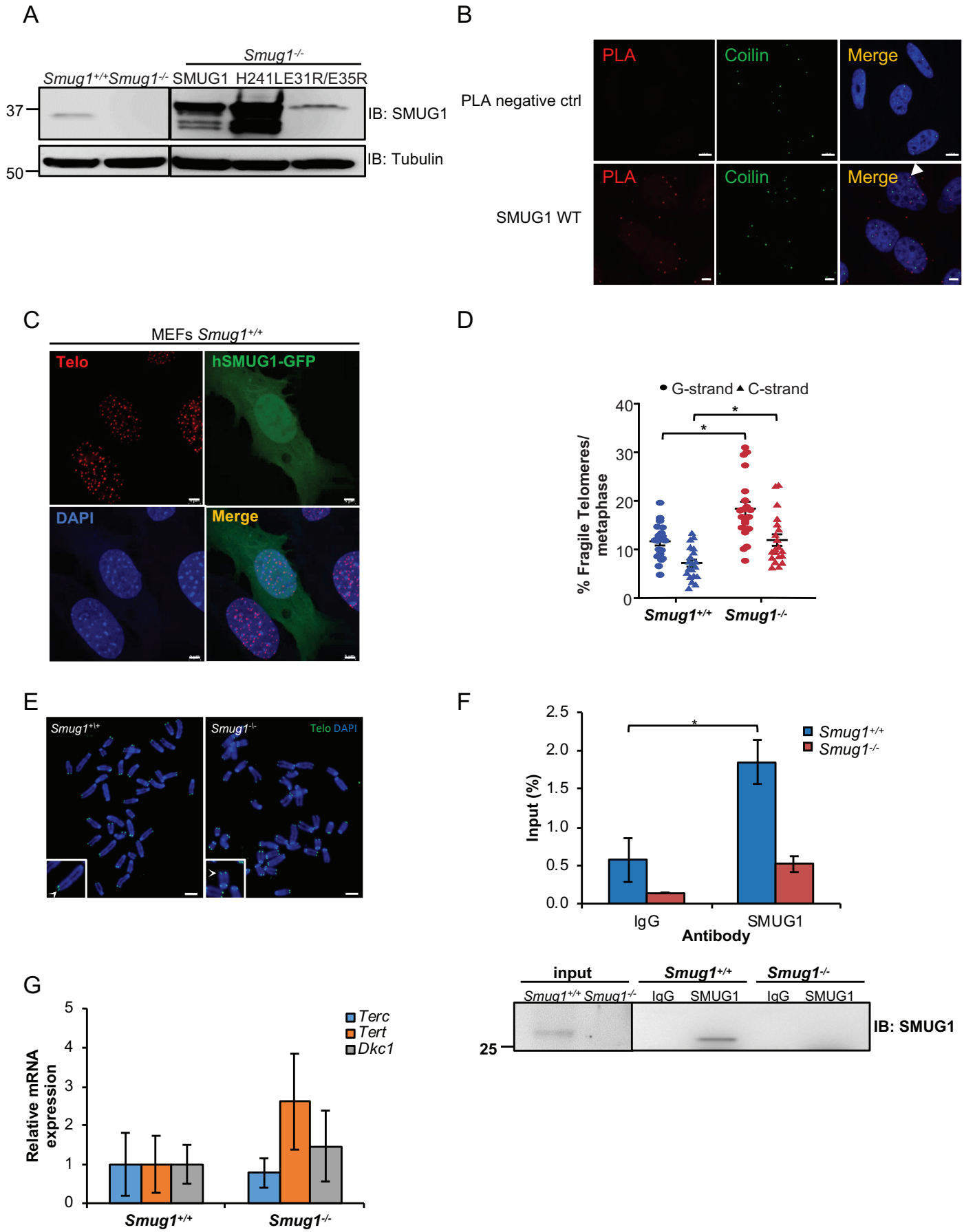


Figure S1, related to Figure 1 and 2 Telomere related phenotypes of *Smug1*$^{-/-}$ mice.
(A) Representative immunoblot indicating the expression of SMUG1 protein in stable MEF cell lines. Tubulin was used as loading control. (B) PLA images showing SMUG1 and Coilin interaction in Cajal bodies (arrow head) in HeLa cells transiently transfected with Flag-tagged SMUG1 and synchronized in S-phase. PLA negative ctrl, no Flag antibody. Scale bars, 10 μm (top) and 4 μm (bottom) (C) IF-FISH images presenting localization of SMUG1 (green) and telomeres (red) using PNA-probes in *Smug1*$^{+/+}$ MEFs (Scale bars, 5 μm). (D) Quantification of fragile telomere ends in transformed MEFs. (E) PNA-FISH in primary MEFs metaphases. Telomeres were hybridized with a telomere specific probe (Telo, FITC) and chromosomes were stained with DAPI. Scale bars, 5 μm (F) RNA immunoprecipitation of *Terc* by SMUG1 in MEF cells using antibody against SMUG1, quantified by qPCR. Data are presented as percentage of input RNA. IgG, negative control (top). Representative blots after RIP with IgG and SMUG1 antibodies (bottom). (G) Relative mRNA expression levels, in three independent clones of transformed MEFs, quantified by qPCR. (D) Data represent means ± s.e.m, n=3, 30 metaphases were scored, (D, F) *P ≤0.05, **P ≤0.01 (two-tailed Student's t-test).

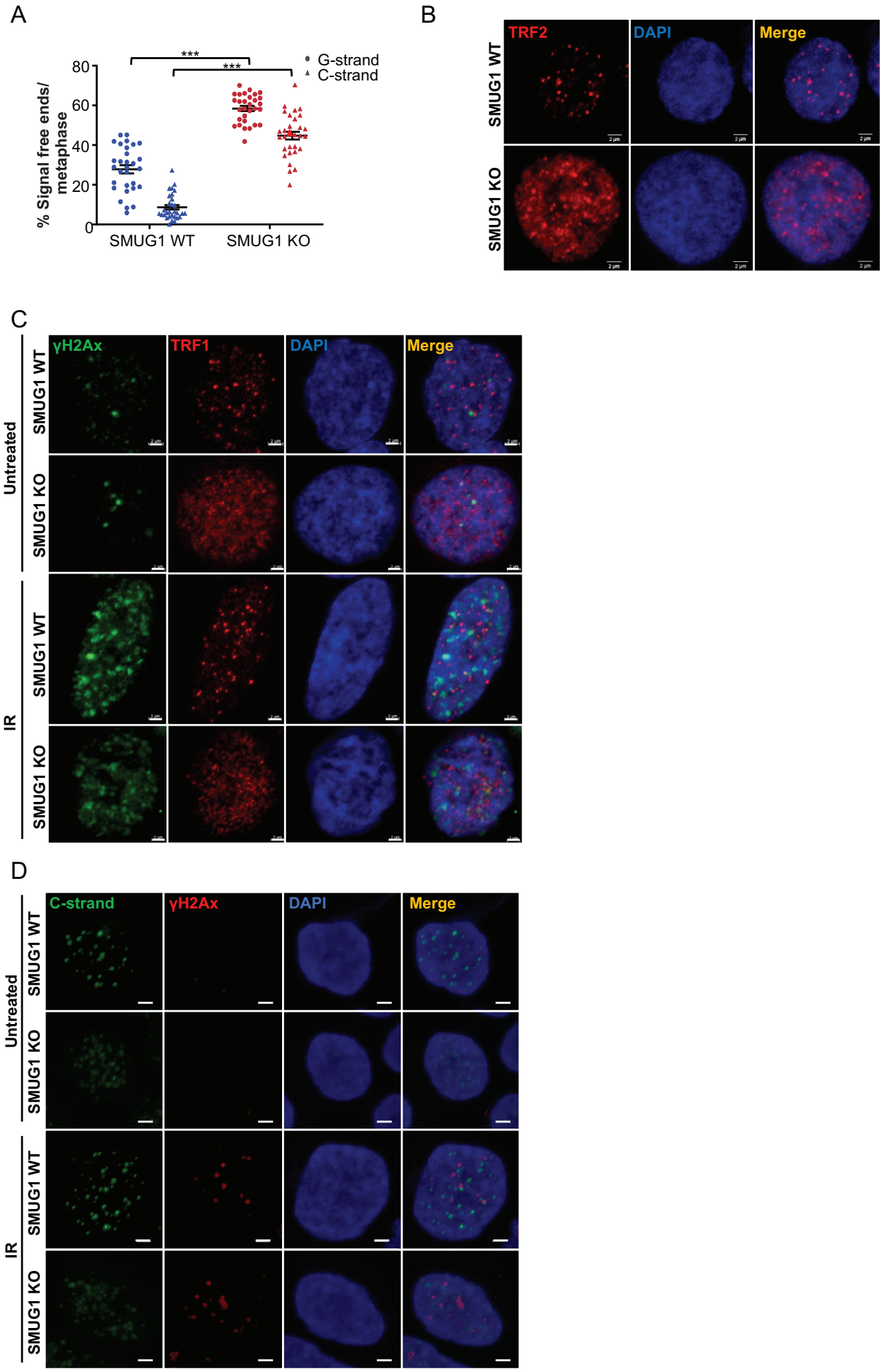


Figure S2, related to Figure 3, Telomere erosion in SMUG1-KO HAP1 cells.
(A) Quantification of signal-free chromosome ends (SFEs) in HAP1 cells is shown. (B) Representative IFs for TRF2 in HAP1 SMUG1-WT and SMUG1-KO cells (Scale bars, 2 μm). (C, D) Representative IFs for TRF1 and γH2Ax (C) and for telomeric specific probe and γH2Ax (D) in HAP1 SMUG1-WT and SMUG1-KO cells in normal culture conditions (top) and following treatment with 2 Gy ionizing irradiation (bottom) (Scale bars, 2 μm). (A) 30 metaphases were scored, ***P≤0.001 (two-tailed Student's t-test).

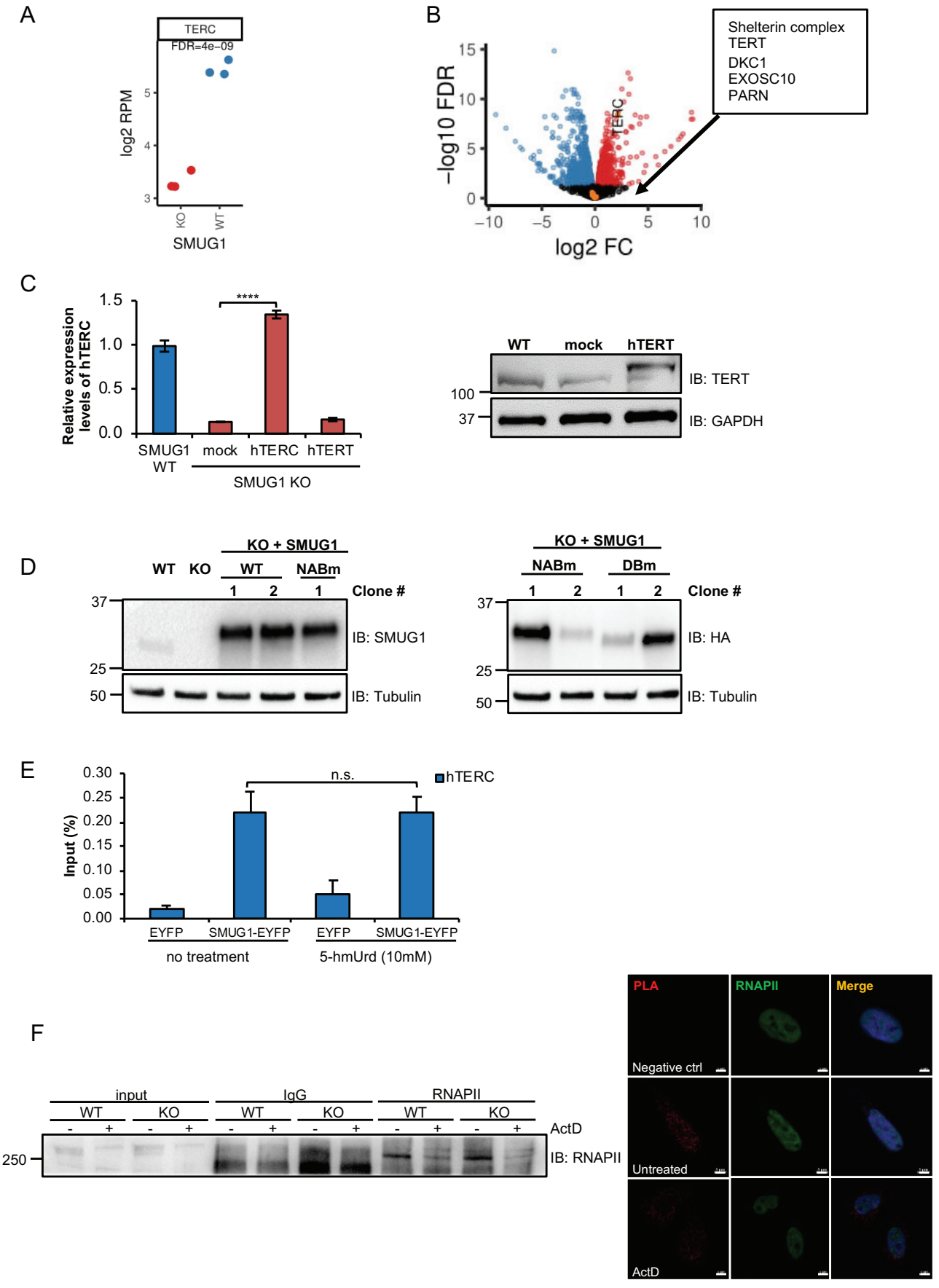


Figure S3, related to Figure 4 and 5, Telomerase related phenotypes in human cells.
(A) Expression profile for *hTERC* in SMUG1 WT and SMUG1-KO HAP1 cells. FDR values are from the comparisons between SMUG1-WT and SMUG1-KO cells. *hTERC* is reduced by 2.1 $\log_2$ fold-change in the SMUG1-KO cells. (B) Volcano plot of genes differentially expressed between SMUG1-WT and SMUG1-KO HAP1 cells (1883 genes with FDR ≤ 0.05; genes down (blue) and up (red) regulated in SMUG1-KO are shown. Genes indicated in text box (orange) were among the non-differentially expressed genes (black) (C) hTERT protein levels (right) and RNA expression levels of *hTERC* quantified by qPCR (left) in HAP1 cells transfected with the indicated plasmids. (D) Quantification of SMUG1 protein levels in two independent SMUG1-KO clones stably complemented with SMUG1 WT, nuclei acids (NABm) and DKC1 binding mutants (DBm). Tubulin was used as loading control. (E) qPCR showing *hTERC* immunoprecipitation by SMUG1-EYFP in HeLa cells grown in the absence or presence of 5-hydroxymethyluridine. Data are presented as percent of input RNA. Background signal given by EYFP control is shown. (F) Representative blots after reChIP for IgG and RNAPII antibodies (left) and PLA images for SMUG1 and RNAPII interaction (right). PLA negative ctrl, no FLAG antibody. (C, E) Data represent means ± s.d., n=2, ****P≤0.0001 (two-tailed Student's t-test).

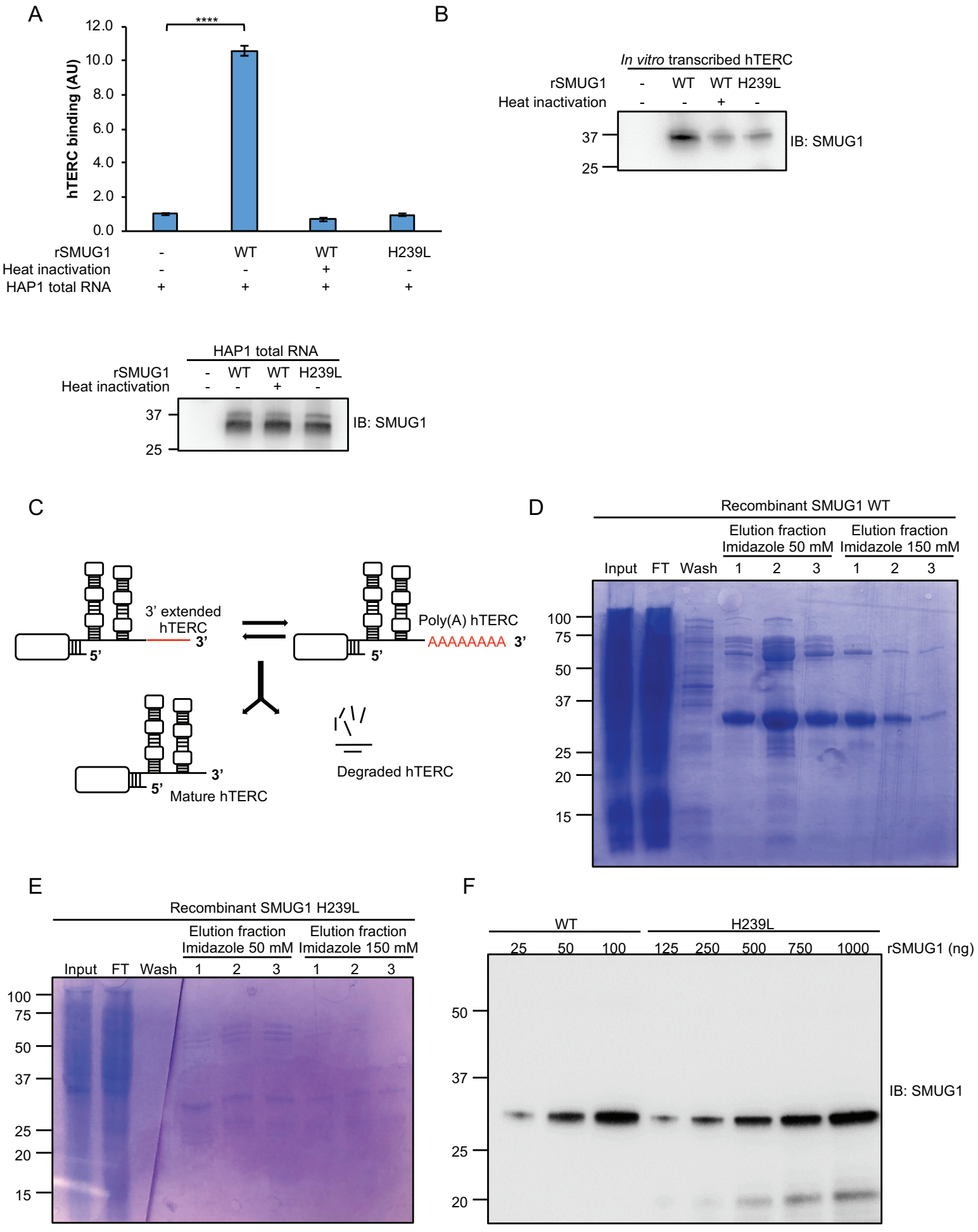


Figure S4, related to Figure 5, SMUG1/*hTERC* interaction efficiency in SMUG1 reconstituted clones.
(A) hTERC binding by SMUG1 WT and H239L mutant measured via qPCR after SMUG1 pulldown using total HAP1 RNA as prey (top). Representative immunoblot for pulldown efficiency is shown (bottom). (B) Representative immunoblot for His-tag pulldown efficiency for Figure 5B. rSMUG1, recombinant SMUG1. (C) Schematic overview of hTERC processing. (A) Data represent means ± s.d., n=3. (A) ****P≤0.0001 (two-tailed Student's t-test). (D, E) Coomassie staining for SMUG1 WT (D) and for H239L mutant (E) showing rSMUG1 purity. FT, flow-through. (F) Representative immunoblot for SMUG1 showing rSMUG1 purity of the elution fraction.

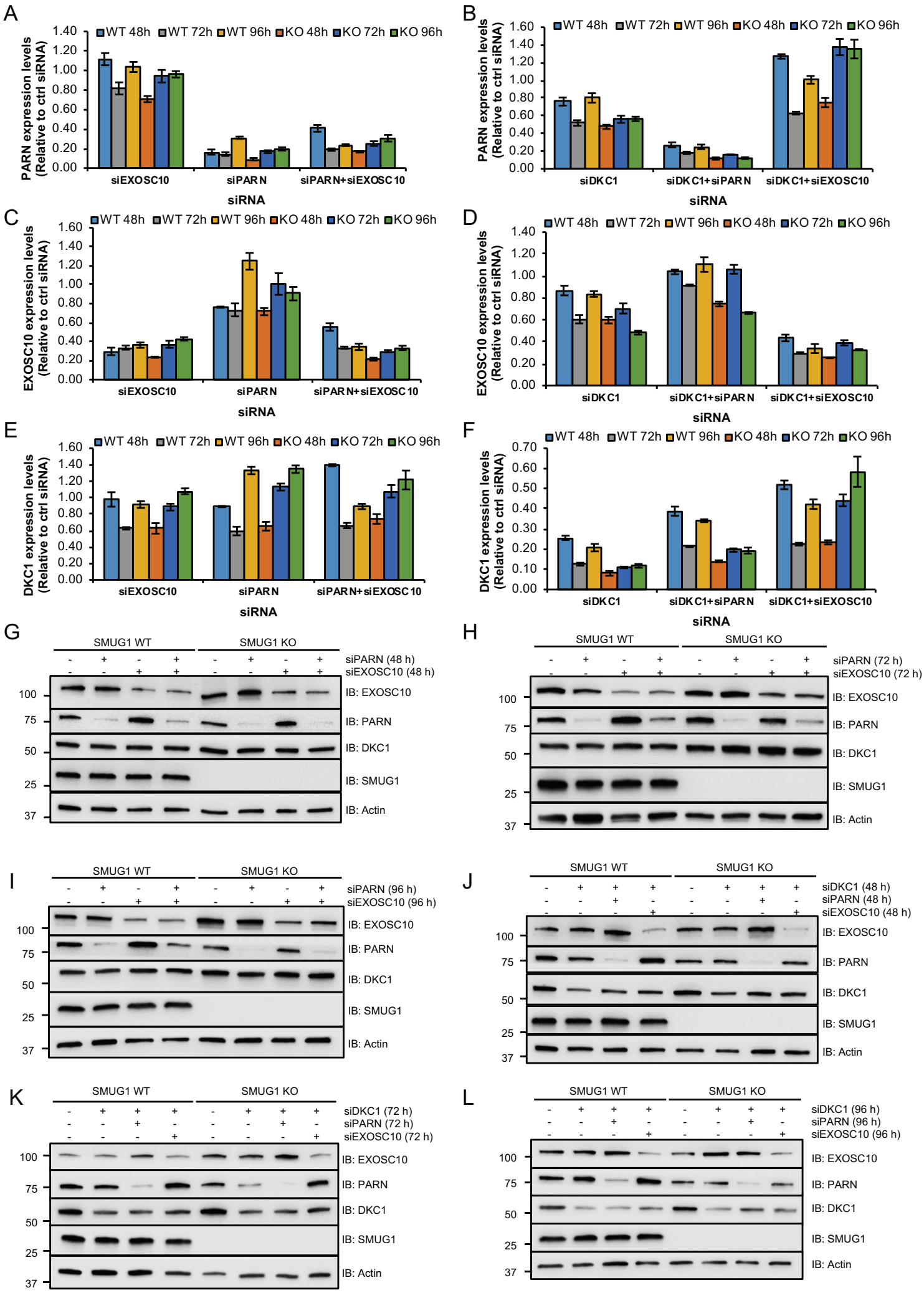


Figure S5, related to Figure 6, Efficiency of siRNA knockdown.
qPCR showing PARN, EXOSC10 (C, D) and DKC1 (E, F) mRNA levels in HAP1 cells upon silencing with the indicated siRNAs (x-axis). Data represent means ± s.d., n=3. (G, H, I, J, K, L) Representative immunoblots showing the knockdown efficiency for the indicated siRNAs at 48, 72 and 96 h.

Table S1, Oligonucleotide information, related to STAR Methods.

Oligonucleotides	Source	Identifier
GCGCGGCTTACCCTTACCCTTACCCTAACC	Ludlow et al., 2014	ACX
AATCCGTCGAGCAGAGTT	Ludlow et al., 2014	TS
GCTTCTGACACAACTGTGTTCACTAGC	O'Callaghan and Fenech, 2011	b-globin_F (human)
CACCAACTTCATCCACGTTCACC	O'Callaghan and Fenech, 2011	b-globin_R (human)
ACTGGTCTAGGACCCGAGAAG	O'Callaghan and Fenech, 2011	36B4F (mouse)
TCAATGGTGCCTCTGGAGATT	O'Callaghan and Fenech, 2011	36B4R (mouse)
CGGTTTGTTTGGGTTTGGGTTTGGGTTTGGG TTTGGGTT	O'Callaghan and Fenech, 2011	TeloF (mouse+human)
GGCTTGCCTTACCCTTACCCTTACCC TTACCCTTACCCT	O'Callaghan and Fenech, 2011	TeloR (mouse+human)
GTGCCCACAGTCAGAAGTGA	This paper	SMUG1_F (human)
TCAGCAGGAGTAAGGTTGCG	This paper	SMUG1_R (human)
CCACATCGCTCAGACACCAT	This paper	GAPDH_F (human)
CCAGGCGCCCAATACG	This paper	GAPDH_R (human)
CCTCGTGGAAGTGACATCGT	This paper	RPLP0_F (human)
CTGTCTTCCCTGGGCATCAC	This paper	RPLP0_R (human)
GAGCAAGACAGTGGGCTGG	This paper	GUSB_F (human)
TCCCATTCGCCACGACTTTG	This paper	GUSB_R (human)
CTACGTAACGATTGATGGTGCCTACAG	Moon et al., 2015	Universal_3'RACE_R
CTCTGTCAGCCGCGGGTCTCTC	Moon et al., 2015	hTERC_L3
GCGAAGAGTTGGGCTCTGTCA	Tseng et al., 2015	Mature_hTERC_F
TTCCTCTTCCTGCGGCCTGAAA	Tseng et al., 2015	Mature_hTERC_R
CTTTCAGGCCGCAGGAAGAGGAA	Tseng et al., 2015	3'ext hTERC_F
GGTGACGGATGCGCACGAT	Tseng et al., 2015	3'ext hTERC_R
GCCTTCCACCGTTCATTCTA	This paper	5'_hTERC_F
GCTGACAGAGCCCAACTCTT	This paper	5'_hTERC_R
GGAGAGTAGTCTGAATTGGGTTATGAGG	Xi and Cech, 2014	H1_1 NB probe
GCTGGCCGTGAGTCTGTTCCAAGCTCC	Xi and Cech, 2014	H1_2 NB probe
AGGCCCACCCTCCGCAACCC	Boyraz et al., 2016	hTERC_1 NB probe
CTTCTCAGTTAGGGTTAGAC	This paper	hTERC_2 NB probe
GAACGGGCCAGCAGCTGACATT	Moon et al., 2015	hTERC_3 NB probe
GGTGCACGTCCCACAGCTCA	Nguyen et al., 2015	hTERC_4 NB probe
GCATGTGTGAGCCGAGTCCTGG	Moon et al., 2015	hTERC_5 NB probe
CCGACTTTGGAGGTGCCTTC	Tseng et al., 2015	hTERC_6 NB probe
CCCAGTGCATACCCTACACC	This paper	hPARN_F
CACAGGAATCGGTCTGCTCA	This paper	hPARN_R
GTCTCTCAGGCAGCGAAGTT	This paper	hEXOSC10_F
TTCCCGGGCAGCTGTTTG	This paper	hEXOSC10_R
ATATGAGGACGGCATTGAGG	This paper	hDKC1_F
CCATGGTCGCAGGTAGAGAT	This paper	hDKC1_R
CCCTCCTTTGCCTTCCAC	This paper	hTERT_F
AGCTCCCAGGGTCCTTCTC	This paper	hTERT_R
ACCTCATCAAGACACAGCACT	This paper	hTERC_prom_-630F
CTCCTTGAGCAGAGGATGGC	This paper	hTERC_prom_-630R
AGGCCCTCTGATACCTCAAGT	This paper	hTERC_prom_-250F
GGGAAAGGCTTTGCGTCTTT	This paper	hTERC_prom_-250R
CCAAGCTTCTAATACGACTCACTATAGGGAGAGGGTTGCGG AGGGTGGGCCT	Tang et al., 2018	T7_hTERC_F
GCATGTGTGAGCCGAGTCCTGG	Tang et al., 2018	T7_hTERC_R
TAMRA-OO-TTAGGGTTAGGGTTAGGG;PNA	Zimmermann et al., 2014	TAMRA-TelG
FITC-OO-CCCTAACCCTAACCCTAA;PNA	Zimmermann et al., 2014	FITC-TelC