Telomere-regulating Genes and the Telomere Interactome in Familial Cancers

Authors:
Carla Daniela Robles-Espinoza¹, Martin del Castillo Velasco-Herrera¹, Nicholas K. Hayward² and David J. Adams¹

Affiliations:
¹Experimental Cancer Genetics, Wellcome Trust Sanger Institute, Hinxton, UK
²Oncogenomics Laboratory, QIMR Berghofer Medical Research Institute, Herston, Brisbane, Queensland, Australia

Corresponding author:
Carla Daniela Robles-Espinoza, Experimental Cancer Genetics, Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambs., UK. CB10 1SA. Telephone: +44 1223 834244, Fax +44 1223 494919, E-mail: cdre@sanger.ac.uk

Financial support:
C.D.R.-E., M.d.C.V.H. and D.J.A. were supported by Cancer Research UK and the Wellcome Trust (WT098051). C.D.R.-E. was also supported by the Consejo Nacional de Ciencia y Tecnología of Mexico. N.K.H. was supported by a fellowship from the National Health and Medical Research Council of Australia (NHMRC).

Running title:
Telomere-regulating genes in familial cancers

Keywords:
Telomeres, telomerase, shelterin, germline variation, cancer predisposition

Conflicts of interest:
The authors declare no conflicts of interest.

Word count: 6,362 (including figure legends)
Number of figures and tables: 2 figures in main text, 3 tables in supplementary material
Abstract

Telomeres are repetitive sequence structures at the ends of linear chromosomes that consist of double-stranded DNA repeats followed by a short single-stranded DNA (ssDNA) protrusion. Telomeres need to be replicated each cell cycle and protected from DNA-processing enzymes, tasks that cells execute using specialized protein complexes such as telomerase (TERT), which aids in telomere maintenance and replication, and the shelterin complex, which protects chromosome ends. These complexes are also able to interact with a variety of other proteins, referred to as the telomere interactome, in order to fulfil their biological functions and control signaling cascades originating from telomeres. Given their essential role in genomic maintenance and cell cycle control, germline mutations in telomere-regulating proteins and their interacting partners have been found to underlie a variety of diseases and cancer-predisposition syndromes. These syndromes can be characterized by progressively shortening telomeres, in which carriers can present with organ failure due to stem cell senescence among other characteristics, or can also present with long or unprotected telomeres, providing an alternative route for cancer formation. This review, summarizes the critical roles that telomere-regulating proteins play in cell cycle control and cell fate and explores the current knowledge on different cancer-predisposing conditions that have been linked to germline defects in these proteins and their interacting partners.

Introduction

Telomeres are specialized structures at the ends of linear chromosomes that consist of arrays of repeated nucleotide sequences. Telomeres play an essential role in regulating genomic stability by allowing the cell to distinguish between chromosome ends and double-stand DNA breaks, a function that is controlled by adopting characteristic chromosomic structures and by complexes of telomere-binding proteins. Protein complexes also control telomere length, access of the DNA repair machinery, and telomere end-protection, which is also known as “capping” (1-3). In humans
and other vertebrates, telomeres comprise 9-15 kilobases of double-stranded TTAGGGG repeats, followed by a 50-300-nucleotide protrusion of G-rich ssDNA, known as the G-strand overhang (3-5).

Tight regulation of telomere length is essential for normal cell function. In the absence of telomere length maintenance mechanisms, the ends of chromosomes are shortened with each round of cell division due to the inability of the replication machinery to fully synthesize the 5’ end of the lagging strand (6), as well as the need to enzymatically generate the G-strand overhang (7-9). In order to counteract these effects, cells use a specialized enzyme complex called telomerase that is able to add TTAGGG repeats to the ends of chromosomes (10). In most somatic cells, telomerase is under stringent control and expressed only transiently and at low levels, the exception being cells that are undergoing rapid proliferation such as bone marrow precursors cells, as well as embryonic or adult stem cells (11-13). However, the level of telomerase expression in somatic tissues is insufficient to sustain indefinite cell proliferation, and so telomeres progressively erode with each cell division, ultimately resulting in cell senescence or death (reviewed in (14)).

In addition to carefully regulating the length of telomeres, cells also need to distinguish telomere ends from DNA breaks that need to be repaired. Chromosome ends can adopt alternative structures that probably aid in this function. For example, the G-strand overhang is able to form G-quadruplexes, which are stacks of planar arrays of four G’s hydrogen-bonded by Hoogsteen base pairs (15, 16). This single-stranded overhang is also able to invade the double-stranded region of the telomere, forming what are referred to as a D- and T-loops and shielding the ssDNA from nucleases and DNA-processing enzymes (2, 17, 18) (Figure 1a). The shelterin complex, a large macromolecular structure composed of six proteins, also contributes to this function by binding to telomeres and inhibiting DNA repair, whilst also contributing to telomere length control by regulating access of telomerase to telomeres (2, 19) (Figure 1a). Alterations to telomere structure, or
the loss of shelterin, results in defects in sister chromatid exchange, telomere fusions, polyploidization, telomere length dysregulation and other chromosomal abnormalities (2, 5, 20). The structures of telomerase and shelterin, as well as their roles in telomere protection and length regulation, are discussed in more detail in the next section.

Telomerase and shelterin, with key roles in the regulation of telomere length and end-protection, have altered expression or are affected by somatic mutations in cancers (21-23), which confers upon malignant cells the ability to bypass senescence whilst promoting genomic instability. Many cancers display increased telomerase activity or have activated the alternative lengthening of telomeres (ALT) pathway leading to sustained telomere maintenance (reviewed in (24)). Further, ectopic expression of telomerase has been shown to elicit replicative immortality in certain cell types, a hallmark of cancer (25, 26). It has also been suggested that shelterin complex mutations might facilitate the acquisition of somatic aberrations, therefore driving cancer progression through accelerated tumor evolution (23, 27).

Here we discuss what is known about cancer predisposition syndromes driven by germline variants in telomere-regulating proteins.

Proteins involved in telomere maintenance

In this section, we briefly discuss the proteins and complexes involved in telomere maintenance.

**Telomerase.** This protein complex is formed by two core subunits: the catalytic telomerase reverse transcriptase (encoded by *TERT*) and the telomerase RNA component (*TERC*, also known as *TR*) (17, 28) (Figure 1b). Several additional accessory factors are necessary for telomerase assembly and activation, such as the auxiliary protein dyskerin (encoded by *DKC1*), the localization protein
TCAB1 (encoded by *WRAP53*), EST1A (encoded by *SMG6*), which might aid in telomerase recruitment (29), and Pontin and Reptin (encoded by *RUVBL1* and *RUVBL2*, respectively) (15).

The template for DNA synthesis is provided by *TERC*, a 451 nucleotide-long RNA subunit that can bind accessory proteins as well as TERT to ensure its own stability and telomerase biogenesis or localization (15, 30). The highly conserved TERT subunit provides the catalytic site for the addition of telomeric repeats to the end of chromosomes, and its amino acid sequence contains domains for binding *TERC* and for forming a closed ring-like structure that can bind the DNA-RNA duplex (15, 31). Dyskerin is an essential protein that associates with *TERC* and is required for its correct processing, as well as for telomerase activity and stability, and is found in a complex with both TERT and *TERC* when the catalytic enzyme is purified from cell lines (32, 33). Dyskerin can also bind the accessory proteins NOP10, NHP2 and GAR1 to form a subcomplex that is necessary for *TERC* accumulation *in vivo* (Reviewed in (12)) (Figure 1b). If proteins such as TCAB1 or EST1A are defective or abnormally expressed, then telomerase function is impaired due to mislocalization of *TERC* to the nucleoli or by telomere structural abnormalities (29, 34, 35).

In normal somatic cells, telomerase is assembled in Cajal bodies, which are specialized RNA-processing sub-organelles in the cell nucleus (36). It is then shipped to telomeres by TCAB1 and modified into an active conformation by Reptin and Pontin, by which point it is able to start adding nucleotides to the ends of telomeres (5, 37, 38). Telomerase is active during the S-phase of the cell cycle, and it is thought that it targets telomeres for elongation in a random fashion, despite some studies showing that it is recruited preferentially to the shortest telomeres (39-41). Regardless, telomerase expression is tightly regulated, as its unintended activity, or inactivity, can lead to tumor predisposition or stem cell exhaustion, respectively (42, 43).
Given the importance of the telomerase complex in telomere length maintenance, defects in any of its subunits or in proteins that assist its assembly and transport have been associated with premature aging, stem cell depletion, and predisposition to cancer and other syndromes (15, 44).

**The shelterin complex.** Despite being formed by only six proteins, shelterin displays a wide range of functions that include telomere length maintenance, protection from DNA repair mechanisms, and the regulation of signaling cascades from telomeres (5). It is composed of six core proteins: telomeric repeat-binding factors 1 and 2 (TERF1 and TERF2, also known as TRF1 and TRF2), TERF1-interacting protein 2 (TINF2, also known as TIN2), adrenocortical dysplasia protein homolog (ACD, also known as TPP1, TINT1, PTOP and PIP1), TERF2-interacting protein 1 (TERF2IP, also known as RAP1) and protection of telomeres 1 (POT1) (2) (Figure 1a). Although these proteins are fast-evolving, with the architecture of the complex being different in organisms such as ciliates and yeast when compared to mammals, the overall functionality of shelterin is highly conserved (45, 46).

TERF1 and TERF2 are double-stranded DNA-binding proteins that recognize telomeric repeats with high affinity upon homodimerization, while POT1, the most evolutionarily conserved member of shelterin, can specifically recognize telomeric ssDNA (2, 47-49). Therefore, the presence of several TTAGGG-binding domains in the complex gives shelterin its exquisite specificity for telomeric sequence. ACD binds to POT1, recruiting POT1 to telomeres and enhancing its recognition ability, whereas TERF2IP localizes to telomeres via its interaction with TERF2, and might have a role in the distribution of telomeres in post-mitotic nuclear assembly (2, 50-52). TINF2 is able to bind TERF1, TERF2 and the ACD/POT1 subcomplex, therefore bringing all the shelterin components together (15, 53) (Figure 1a).
The importance of the shelterin complex is evidenced by the fact that null mutations in the majority of its components result in embryonic lethality in mice (54-57). One of the most important functions of shelterin is to protect chromosome ends from DNA repair nucleases, which it achieves by inhibiting six DNA damage signaling pathways: ATM- and ATR-signaling, classical non-homologous end joining (NHEJ), alternative NHEJ, homologous recombination, and resection (20). Shelterin also has an important role in regulating telomere structure and length. TERF1 and TERF2 have DNA remodeling activities, being able to bend DNA and contributing to T-loop formation, whereas POT1 is able to regulate telomere length by contributing to the nucleolytic processing of the telomeric 5’ end and control telomerase access to the end of chromosomes (2, 47, 58, 59). The absence of functional shelterin therefore leads to polyploidization, fragile telomeres and sister chromatid exchanges, among other chromosomal aberrations (20).

The telomere interactome. The telomere interactome consists of the telomerase and the shelterin complexes along with an extended network of interacting partners, which aid in regulating diverse signaling cascades originating from chromosome ends (60). In this section, we discuss the different molecular processes that have been linked to the telomere interactome, as well as other non-telomeric functions for telomeric proteins relevant for cancer formation.

DNA repair signaling pathways and cell cycle control. Components of the shelterin complex, as mentioned above, are known to interact with DNA repair signaling pathways to control cell fate when telomeres are damaged. TERF1 co-immunoprecipitates with ATM, and is phosphorylated by it in vivo upon DNA damage, which prevents the cell from entering mitosis or apoptosis, and therefore results in a reduction in radiation sensitivity (61). It can also interact with RTEL1, a protein that aids in telomere elongation, to suppress telomere fragility possibly by unwinding G-quadruplexes (62-64). TERF1 can also bind EB1 (encoded by MAPRE1) and tankyrase (encoded by
TNKS1), conferring it a role in microtubule polymerization and telomere elongation in a cell cycle-dependent manner (65-67) (Figure 1a).

TERF2 can recruit the ERCC1/XPF (encoded by ERCC4) heterodimer to the telomeric complex, which helps protect against telomere recombination with interstitial telomere-related sequences, and at the same time prevent NHEJ by blocking its access to the G-strand overhang (68). It is also able to recruit the MRE11A-RAD50-NBN (MRN) complex to telomeres, where it acts to detect the presence of uncapped telomeres (69, 70), and can also bind the Werner (WRN) and Bloom (BLM) helicases to stimulate their activity (71, 72). Other identified TERF2 binding partners are Apollo (encoded by DCLRE1B), Ku70 (encoded by XRCC6), PARP1/2 (Reviewed in (73)), and the chromatin regulator UBR5 (74) (Figure 1a).

Other members of the shelterin complex also interact with proteins involved in cell cycle regulation and genomic stability. For example, TINF2 can interact with the E3 ubiquitin ligase SIAH2, which regulates its stability allowing for dynamic control of the shelterin complex during the cell cycle (75). It can also bind CBX3 to maintain sister telomere cohesion and regulate telomere length (76). TERF2IP can interact with the nuclear envelope protein SUN1, which has a critical role in cell cycle progression by tethering telomeres to the nuclear membrane after mitosis (52), and with Ku80 (encoded by XRCC5) (77), a protein required for double-strand DNA repair. POT1 can physically interact with the WRN and BLM helicases, stimulating their activity and allowing efficient displacement of D-loops and G-quadruplexes (78, 79). ACD can physically bind the helicase and RNA surveillance protein UPF1 to prevent telomere instability and prevent activation of the DNA damage response (80), and can also interact with OBFC1 to regulate telomere length (81) (Figure 1a).
**Other non-telomeric functions of telomeric proteins.** TERF2IP has been found to have additional roles outside telomere maintenance. It has been shown that it can associate with the inhibitor of the NF-κB kinase (IKK) complex, thus positively regulating the NF-κB signaling pathway (82). Additionally, it can bind to extratelomeric chromosomal regions and possibly interact with other factors, thereby participating in subtelomeric gene silencing, metabolism control and interstitial genomic stability (83).

The catalytic subunit of telomerase also has other ‘non-canonical’ functions. TERT can act as a transcriptional modulator of the Wnt-β-catenin signaling pathway by associating with SMARCA4, a member of the SWI/SNF chromatin-remodeling complex, and can also physically occupy Wnt-dependent promoters (84). Also, in addition to using *TERC* as its RNA subunit, TERT is also able to bind the RNA component of mitochondrial RNA processing endoribonuclease (*RMRP*). The TERT-*RMRP* complex is then able to process *RMRP* into small interfering RNAs (siRNAs) to modulate its own levels, and it remains possible that TERT-*RMRP* is able to regulate the expression of other genes via the generation of siRNAs (85).

The association of telomeric proteins with these diverse factors provides a glimpse of the immensely complex interaction network that surrounds the telomere maintenance machinery, and its inherent connection to cell cycle regulation and DNA repair. The fact that some of the abovementioned proteins have ‘non-canonical’ roles outside of their role in safeguarding telomere length and end-protection has significantly extended our understanding of other biological processes, such as the regulation of important developmental pathways, mitochondrial function and cancer.
Telomere syndromes: variants in telomere-regulating genes and cancer predisposition

The contribution of acquired alterations in telomere-regulating genes and their interacting partners to cancer development is well documented. For example, telomerase has been found activated in hepatocellular carcinoma (86), melanoma (87), glioma and other cancers with low self-renewal rates (88), as well as other tissues (89), as it helps cells acquire replicative immortality. Moreover, the shelterin complex member POT1 has been reported to be somatically inactivated in chronic lymphocytic leukemia (23), where it causes telomere de-protection and length extension. DNA repair genes, such as XRCC5, XRCC6 and PARP1 are also frequent targets of somatic alteration in B cell lymphomas (90), and the NF-κB signaling pathway is frequently activated in breast cancer (91) and lymphomas (92), where it contributes to cell survival. It remains possible, given the important roles of these DNA repair and NF-κB pathway proteins in telomere maintenance (93-96), that patients that have somatically-acquired mutations in them display telomeric abnormalities. However, germline mutations in these genes also underlie an important proportion of hereditary cancer predisposition syndromes (Supplementary Tables 1-3). In this section we discuss the different disorders attributed to malfunction of components of the telomere interactome, including hereditary telomere dysfunction.

Telomere shortening syndromes. Progressive shortening of telomeres is seen in patients carrying germline mutations inactivating telomerase or telomere maintenance mechanisms, as their cells are unable to counteract telomere attrition after each cell division (41). In these individuals, telomeres might quickly reach their critical minimum size in highly proliferating tissues such as the bone marrow, and organ failure might ensue as stem cells senesce. However, in spite of this seemingly cancer-protecting effect, these individuals are also cancer-prone. The mechanism by which this
happens is largely unexplained, although the fusion of chromosome ends, possibly by NHEJ, is potentially involved, as it has been observed in mutation carriers and in Terc−/− mice (97).

Although telomere-shortening syndromes are precipitated via a common mechanism, their clinical presentations can be quite variable. This is, in part, due to a phenomenon known as genetic anticipation (41). Telomere length is a heritable trait, and because telomere shortening also occurs in germ cells, offspring of mutation carriers inherit chromosomes with shorter telomeres, and thus some patients will present earlier with a more severe phenotype than others (98). This is thought to be the reason behind the observation that, within a pedigree, the disease can evolve from generation to generation, becoming increasingly more severe. It might also be the case that offspring that inherit a wild-type copy of the disease-causing gene sometimes present with mild symptoms, as they will have inherited shorter telomeres. However, although this observation has been made in mice, it remains unclear whether this applies to humans (41, 42). It is worth mentioning that progressive telomere shortening is one of only two known molecular mechanism for genetic anticipation, the other one being the trinucleotide repeat expansion that underlies conditions such as Huntington’s disease and myotonic dystrophy (41).

Germline mutations in several of the telomere maintenance genes cause dyskeratosis congenita (DC), a rare multisystem disorder characterized by cutaneous abnormalities such as aberrant skin pigmentation, mucosal leukoplakia and nail dystrophy, in addition to bone marrow failure and cancer predisposition (99). About 10% of DC patients develop cancer, usually in the third decade of life, much younger than in the general population (100). Carcinomas of the bronchus, colon, larynx, pancreas, esophagus, skin and tongue, as well as leukemias and lymphomas are common in these patients (99). It has been estimated that DC patients have a 11-fold increase in cancer incidence compared to the general population, and that their risk, and cancer susceptibility spectrum, are similar to those of Fanconi anemia (FA) patients (100).
As DC is caused by mutations in several genes, it presents with different modes of inheritance. The autosomal dominant form, referred to as AD-DC, can be caused by mutations in \( TERT, \) \( TERC \), and \( TINF2 \), and it has been shown that mutations in these genes can act as either haploinsufficient or dominant negative alleles (reviewed in (101)) (Figure 2a-c). Variants in TINF2 have been shown to affect the CBX3/SIAH2 binding motif, affecting sister telomere cohesion (76). The autosomal recessive form (AR-DC) can be caused by variants in TCAB1, as well as in the genes coding for the telomerase accessory proteins NOP10 and NHP2, and are predicted to cause telomerase loss of function (41) (Figure 2d-f). The X-linked form of DC, caused by mutations in dyskerin (Figure 2g), is the most common and most severe, with more than 40% of patients developing bone marrow failure by the age of 10 years (102).

In addition to classic DC, other telomere shortening syndromes have been described in the literature. Hoyeraal-Hreidarsson syndrome (HHS) is recognized as a clinically severe form of DC presenting with developmental delay, microcephaly and immunodeficiency, in addition to the characteristics mentioned above. It is caused by rare variants in dyskerin (reviewed in (101)), \( TERT \) and \( TINF2 \) (collected in (103)), recessive mutations in \( RTEL1 \) (104), and has also been associated with a rare variant in \( TERC \) (105) (Figure 2a-c,g,h). Like individuals with DC, HHS patients also seem to have increased chromosomal instability, although they usually die in their first decade of life due to progressive bone marrow failure (104). Other conditions with overlapping phenotypes are Revesz and Coats plus syndromes, in which patients present with exudative retinopathy, intracranial calcifications, and neurological and bone marrow defects (41). In these disorders, germline mutations have been found in telomeric proteins such as TINF2 (106) and telomere maintenance component 1 (CTC1), which lead to extensive telomere de-protection (107) (Figure 2c,i). Although their cells may show markers of genomic instability such as spontaneous DNA
damage, anaphase bridges and sister telomere losses (104, 108), cancer is not commonly seen in patients with these more severe syndromes, as they succumb to other complications at a young age.

Aplastic anemia (AA) is a bone marrow disorder, characterized by pancytopenia (low counts of red blood cells, white blood cells and platelets). It can be acquired upon exposure to radiation or toxic chemicals. An inherited form has been associated with mutations in the shelterin complex members TERF1 (an intronic variant) and TINF2, and in the telomerase components TERT and TERC (101, 109, 110) (Figure 2a-c). Mutations in these genes are found as heterozygous variants, and are thought to result in haploinsufficiency (101). Patients with AA resulting from mutations in the abovementioned genes have shorter telomeres than controls (111) and may develop cancer, especially leukemias and lymphomas (112). There are a number of related conditions, such as FA, paroxysmal nocturnal hemoglobinuria (PNH) and myelodysplastic syndromes (MDS). All of these may be characterized by blood and bone marrow abnormalities as well as significantly shorter telomeres (113-115), and can be associated with germline mutations in proteins part of the telomere interactome such as XPF in the case of FA (116) and in TERC or its promoter in the case of MDS and PNH (101, 117) (Figure 2b). FA patients often progress towards MDS or acute myeloid leukemia (AML), which has also been associated with rare germline mutations in TERT (113, 118). MDS are sometimes considered a form of cancer, as they are clonal diseases arising from a single cell, and indeed, about 30% of MDS patients develop AML (101). Recently, an unexpected high incidence of diverse cancers, including pancreatic, gastric, prostatic and lymphoma, has been seen in a small cohort of PNH patients (119). This observation suggests that this condition might predispose to the development of neoplasias, although more studies with larger patient cohorts are necessary in order to establish the link definitively.

Germline mutations in other genes part of the telomere interactome can also cause a different spectrum of short telomere syndromes. The Nijmegen breakage syndrome (NBS) is characterized
by hypersensitivity to ionizing radiation, immunodeficiency and a strong predisposition to malignancy. It is caused by recessive mutations in the NBN gene, rendering a defective protein that causes impaired ATM phosphorylation, a delayed cell cycle arrest and an accumulation of somatic mutations (78). In accordance with the role of the MRN complex at telomeres, cells isolated from NBS patients display accelerated telomere attrition and an increased number of telomere fusions. Another closely related condition is ataxia telangiectasia (A-T), which is defined by cerebellar degeneration in addition to all NBS characteristics mentioned above (120). A-T is caused by deleterious germline variants in ATM, leading to a defective response to DNA damage and impaired cell cycle control. A-T cells have accelerated telomere shortening, genomic instability and altered telomere-nuclear matrix interactions (78). NBS and A-T patients are strongly predisposed to cancer, and they develop predominantly lymphomas and B- and T-cell leukemias, although solid tumors have also been observed (120). Rare variants in other members of the MRN complex also give rise to cancer-predisposing syndromes, for example, deleterious mutations in MRE11 cause a similar disorder to A-T (121), and some germline variants in NBN and RAD50 have been found associated with childhood acute lymphoblastic leukemia (122, 123).

The Bloom (BS) and Werner (WS) syndromes are caused by autosomal recessive mutations inactivating two interacting partners of TERF2 and POT1, the DNA helicases BLM and WRN, respectively. BS can present with male infertility, skin pigmentation aberrations and a predisposition to malignancy (78). WS is characterized by premature aging, short stature, endocrine disorders and an elevated frequency of cancer (124). WS patients present with an excess of cancers, especially thyroid carcinomas, osteosarcomas, melanomas, meningiomas and leukemias, and cancer is one of the main causes of death in these patients (124). BS patients have an elevated incidence of lymphomas, leukemias, gastrointestinal tract neoplasias, genital and urinary tract tumors and cutaneous neoplasias, and often present with multiple malignancies (125). Cells lacking functional
WRN display catastrophic telomere loss and an increased frequency of chromosomal fusions, and cells from BS patients have increased chromosomal breakage and sister-chromatid exchanges (78).

There are other telomere-shortening syndromes that manifest later in life. Idiopathic pulmonary fibrosis (IPF) is a chronic and progressive lung disease characterized by usual interstitial pneumonia that manifests in middle-aged to older patients (126). It has been estimated that 8-15% of familial cases carry haploinsufficient mutations in \textit{TERT} or \textit{TERC}, and because IPF affects as many as 63 people per 100,000 in the United States alone, it is the most common manifestation of telomerase gene mutation carriers (41, 101, 126) (Figure 2a,b). In one family, it has also been associated with a mutation in dyskerin (127) (Figure 2g). IPF cases have a higher incidence of lung cancer according to several studies, and it has been found to be one of the common causes of death in these patients (for a review see reference (126)). Individuals that carry germline mutations in telomerase components may also present for the first time with adult-onset liver cirrhosis (41, 128), which often progresses to hepatocellular carcinoma.

As discussed above, mutations in the same genes can give rise to a broad range of phenotypes. For example, rare variants in telomerase components can cause DC, AA or IPF, and mutations in \textit{TINF2} can result in DC, AA or Revesz syndrome (Figure 2a-c, Supplementary Table 1). How this happens is largely unexplained, though the phenomenon of genetic anticipation, which includes factors such as inherited short telomeres, may play an important role. It is noteworthy however, that these syndromes have largely overlapping characteristics that will often present concurrently in the same individual; for example, IPF patients with telomerase mutations are at an increased risk of developing bone marrow failure, and AA patients are conversely more prone to developing IPF (41, 129). Nonetheless, it remains to be shown if factors such as the positions where the mutations lie within \textit{TERC} or the telomeric proteins contribute to the spectrum of telomere shortening syndromes.
**Germline mutations can underlie long telomeres and cancer predisposition.** Long telomeres might also be a risk factor for cancer development. For example, it has been observed that both breast cancer cases and women at high genetic risk for developing the disease have longer telomeres than controls, with telomere length displaying a positive correlation with risk (130, 131). In fact, BRCA1 and BRCA2 mutation carriers have recently been found to have longer telomeres than controls (132), and longer telomeres were also associated with a worse prognosis in a subset of breast cancer patients with advanced disease (131). Other cancers in which longer telomeres have been associated with increased risk are non-Hodgkin lymphoma (133), melanoma (134) and lung cancer in both smokers (135) and non-smokers (136). Interestingly, longer telomeres have also been associated with risk of colorectal cancer, although only in young individuals, as risk is associated with shorter telomere length in older individuals (137).

The complex relationship between telomere length and cancer risk might seem paradoxical at first, as both long and short telomeres are associated with an increased cancer incidence. However, there are other factors likely to influence this risk. For example, longer telomeres in younger individuals might be predictive of a malfunction in the telomere maintenance machinery (137), which might lead to an activation of telomerase and therefore a longer cell lifespan, allowing a higher accumulation of somatic mutations. Conversely, as discussed in the previous section, chromosomes with short telomeres are prone to genomic instability and chromosomal rearrangements, and because telomeres shorten with age, older individuals might be at an increased risk to develop malignancy-associated chromosomal lesions. The fact that telomere lengths outside an optimal range results in cancer predisposition suggests that there is a delicate balance to be struck for tissue homeostasis.
The genomic variation underlying telomere length has been investigated in an unbiased manner by genome-wide association studies (GWAS). Variants in the \textit{TERC}, \textit{TERT}, \textit{RTEL1}, \textit{OBFC1} and \textit{CTCI} loci have been found to influence mean telomere length, in addition to several other genomic regions (138, 139), observations replicated in numerous studies. In some cases, such as for \textit{CTCI} and \textit{TERC}, it has been suggested by \textit{in vitro} experiments or analysis of patterns of genome-wide expression data, that the variants associated with longer telomeres are also correlated with changes in gene expression (138-140), and that this in turn may increase telomerase expression. However, to our knowledge, no specific germline genetic changes had been associated with an increased telomere length in humans until recently.

Recently, two studies identified rare, germline variants in \textit{POT1} predisposing to the development of familial melanoma (141, 142) (Figure 2j). Carrier individuals in these cohorts had significantly longer and more fragile telomeres than controls, and in some cases developed not only melanoma but also cancer in other tissues. Some of the variants identified abolish the binding of POT1 to ssDNA, and thus it is possible that carriers are predisposed to malignancy via telomere uncapping and a more permissive extension of chromosome ends. However, the biological mechanism underlying the strong cancer predisposition observed in carriers requires further investigation.

\textit{Other cancer-predisposition syndromes arising from mutations in the telomere interactome.} There are other cancer-predisposition conditions caused by mutations in telomere-interacting components that do not have a clear effect on telomere length (or in which telomere length has not been measured thus far). For example, a germline mutation in the promoter of \textit{TERT} was recently discovered to co-segregate with melanoma in a 14-case German pedigree (143). This rare variant creates an E-twenty six/ternary complex factor (Ets/TCF)-binding motif, increasing telomerase expression as shown by luciferase reporter assays. At present it is unclear if this effect is translated to TERT protein levels or if it has any effect on telomere length, but similar \textit{TERT} promoter
mutations have been associated with poor outcome parameters such as increased Breslow thickness and tumor ulceration (144). Following these reports, several other studies have found other somatically-acquired mutations in the promoter of *TERT* that also increase *TERT* expression. However, in at least one of these studies, these mutations were associated with shorter telomeres, contrary to what would be expected given the role of TERT in telomere length maintenance (145). This seemingly counterintuitive result could be due to these mutations arising late in the tumor evolutionary history, helping cells survive in the late stages when extensive chromosomal instability has taken place (145). Therefore, the presence of activating *TERT* promoter mutations and increased telomerase expression in a tumor is not predictive of longer telomere length. It would be expected that individuals harbouring *TERT* activating germline mutations have increased telomerase expression levels and augmented telomere length. However, it remains to be determined whether the additional ‘non canonical’ roles of TERT affect the relationship between telomerase expression levels and telomere length.

Defects in one of the roles of TERT outside telomere maintenance, affecting the TERT-*RMRP* complex, have been found to underlie cartilage-hair hypoplasia (CHH) (146). CHH patients present with short stature, hair anomaly, immunodeficiency and predisposition to malignancy. The causal mutations, found in both the promoter and the transcribed region of *RMRP*, are predicted to decrease *RMRP* levels and modify ribosomal processing leading to altered cytokine signaling and cell cycle progression (146). CHH patients have a tendency to develop cancer, principally lymphoma, although basal cell carcinoma has also been observed (147, 148).

As previously discussed, TERT can also associate with SMARCA4, a member of the SWI/SNF chromatin-remodelling complex, and regulate Wnt-dependent promoters (84). Recently, germline deleterious mutations in the *SMARCA4* gene were found in all available affected members of four families prone to ovarian small cell carcinoma, and loss of heterozygosity or second somatic
SMARCA4 mutations were observed in the tumors from these patients (149). Interestingly, TERT promoter mutations associated with increased TERT mRNA expression and longer telomere length have also been observed in clear cell ovary carcinomas (150). These TERT mutations are mutually exclusive with loss of expression of ARID1A, another member of the SWI/SNF complex that includes SMARCA4, and the most common alteration seen in ovarian cancer (150). These data might suggest that mutations affecting the SWI/SNF complex members and TERT mRNA levels act on the same biological pathway, although this hypothesis requires further investigation.

Germline mutations affecting other proteins that interact with members of the shelterin complex and that increase cancer risk have been described. For example, a GWAS and a subsequent independent replication found a SNP in SIAH2, encoding an interacting partner of TINF2, to be associated with ER-positive breast cancer in Japanese and Chinese cohorts (151, 152). Rare germline variants in the TERF2IP-interacting protein Ku80 and the TERF2-interacting partner PARP1 have also been found in diffuse large B-cell lymphomas, in which DNA repair genes have been found to be selectively inactivated possibly leading to increased survival in cells that have suffered extensive DNA damage (90). Ku70 polymorphisms have also been reported to be associated with gastric cancer (153), and PARP1 polymorphisms with melanoma (154).

Moreover, efforts are currently on going with the aim of identifying novel telomere-interacting proteins, which then might provide valuable biological insight into the mechanisms of cancer development. For example, a recently described chromatin purification method identified proteins such as BRIP1, NRIP1 and HMBOX1 bound to telomeres maintained by the ALT pathway (155). HMBOX1 has also recently been found to bind both telomeres and telomerase, contributing to telomere elongation (156). All these proteins have been linked to cancer predisposition or differentiation (157-159), and thus their newly discovered association with telomeres might provide mechanistic clues to their role in cancer. Another recent RNA pull-down experiment identified 115
proteins that bind specifically to telomeric repeat-containing RNA (TERRA), the result of telomere transcription (160). Among these proteins are HMGA1, HMGB1, HNRNPA1 and HNRNPM, all of which have previously been implicated in cancer development (161-164).

Summary and concluding remarks

Cells need a way to tell apart the ends of linear chromosomes from accidental DNA breakage, as failing to get this distinction right can predispose an organism to the development of cancer due to the unnecessary “repair” of chromosome ends. Telomeres serve this purpose by adopting alternative structures such as G-quadruplexes and D- and T-loops, which help shield the ends of chromosomes from nucleases and DNA-processing enzymes. Dedicated complexes, such as telomerase and shelterin, aid in telomere protection and maintenance functions and also regulate signaling cascades that originate from telomeres by intervening in a number of biological pathways. These pathways intrinsically link telomere status with cell cycle control, and contribute to determine cell fate when DNA damage takes place.

In cancer, somatic mutations in the telomere proteins and their interaction partners occur frequently, as these allow the bypass of senescence or death checkpoints and the subsequent accumulation of potentially advantageous mutations. In humans, rare germline variants that impair the function of telomere proteins have been found to underlie a wide spectrum of diseases and cancer-predisposition syndromes. The majority of the conditions described in the literature have progressive shortening of telomeres as a common characteristic, but they can present with distinct phenotypes, different ages of onset and varying severity. This heterogeneity in clinical presentation can be attributed to the phenomenon of genetic anticipation, and possibly other causes such as individual genetic make-up and lifestyle choices. Despite the fact that cells from individuals with impaired telomerase activity show signs of early entry into senescence and can therefore lead to
organ failure, these cells are also prone to genomic rearrangements, such as telomeric fusions and sister chromatid exchanges, and thus predispose their host to neoplasia.

Long telomeres are also a risk factor for cancer development; for example, it has been observed that in some cases, breast cancer, lymphoma and melanoma patients have longer telomeres than controls. The intricate relationship between telomere length and cancer predisposition, which includes cell replicative ability and protection from chromosomal instability, explains the need for tight, cell cycle-dependent regulation of telomere proteins. Progress has been made over recent years to elucidate the functions of these proteins, and important roles have been found in DNA repair pathway inhibition, cell fate decision upon DNA damage, facilitation of DNA replication and even non-telomeric roles such as regulation of gene silencing and metabolism control.

These discoveries have allowed us to appreciate the myriad of biological pathways that influence telomeric functions and the effects of particular inherited and acquired mutations in telomere genes. Current research, exploring novel genes that participate in telomere-regulating functions, the mechanisms by which the telomeric proteins contribute to biological cell cycle progression and their influence on telomere dysregulation should extend our understanding of the biology of cancer predisposition and hopefully, in the future, aid in clinical decision-making and patient management.

Figures

Figure 1. Telomere-interacting proteins and their roles. a) The shelterin complex, composed of six proteins (left), and several of its associated proteins are depicted with the biological functions they play in the cell (right). For simplicity, TERF1 and TERF2 homodimers are drawn as single boxes. The alternative T- and D-loop telomere structures are shown. Tnk: Tankyrase, Ub: ubiquitination, P: phosphorylation, ADPr: ADP ribosylation. b) The members of the telomerase
complex are shown bound to the G-strand overhang with a simplified structure for TERC, which provides the template for repeat addition.

Figure 2. **Germline mutations in telomere-interacting proteins found in familial cancer-predisposition syndromes.** The mutations depicted in panels a-c have been compiled from references (101) and (103). Protein structure has been taken from Ensembl release 75 (165) unless specified otherwise.  


**c) Mutations in TINF2.** Structure taken from references (2) and (75). K280X: DC, HHS and Revesz syndrome, K280Rfs: DC and Revesz syndrome, R282S: DC and Revesz syndrome, R282C: DC and AA, R282H: DC, HHS and Revesz syndrome, P283S: DC and HHS, T284Hfs: DC and AA.  

**d) Mutations in TCAB1.** Protein structure taken from (166). Mutations taken from reference (34).  

**e) Mutations in NOP10 were taken from reference (103).**  

**f) Mutations in NHP2 were taken from reference (103).**  

**g) Mutations in dyskerin, encoded by DKC1, were taken from references (103) and (127).** P10L: DC and HHS, I38T: DC and HHS, T49M: DC and HHS, T66A: DC and HHS, T67I: DC and HHS, H68Q: DC and HHS, S121G: DC and HHS, R158W: DC and HHS, S304N: DC and HHS, K314R: DC and HHS, A353V: DC and HHS, A386T: DC and HHS.  

**h) Mutations in RTEL1.** Mutations have been taken from references (62), (104), (167) and (168). E591D: DC and HHS, R974X: DC and HHS, R986X: DC and HHS.  

**i) Mutations in CTC1.** Structure taken from reference (169) and mutations from (107).  

**j) Mutations in POT1.** Structure and mutations taken from references (141, 142).
References


Figure 1
**a**

**TERT**

- Dyskeratosis congenita
- Aplastic anemia
- Acute myeloid leukemia
- Idiopathic pulmonary fibrosis
- Several disorders (see legend)

**b**

**TERC**

- Dyskeratosis congenita
- Hoyeraal-Hreidarsson
- Aplastic anemia
- Myelodysplastic syndromes
- Idiopathic pulmonary fibrosis
- Several disorders (see legend)

**c**

**TINF2**

- Dyskeratosis congenita
- Aplastic anemia
- Acute myeloid leukemia
- Idiopathic pulmonary fibrosis
- Several disorders (see legend)
**d**

TCAB1

- P-rich
- G-rich
- WD40

**e**

NOP10

- H/ACA RNP
- Dyskeratosis congenita

**g**

Dyskerin

- Pseudouridylate synthase (TruB)
- PUA
- Dyskeratosis congenita
- Hoyeraal-Hreidarsson
- Idiopathic pulmonary fibrosis
- Several disorders (see legend)
Molecular Cancer Research

Telomere-regulating Genes and the Telomere Interactome in Familial Cancers

Carla Daniela Robles-Espinoza, Martin del Castillo Velasco-Herrera, Nicholas K. Hayward, et al.

Mol Cancer Res  Published OnlineFirst September 22, 2014.

Updated version
Access the most recent version of this article at:
doi:10.1158/1541-7786.MCR-14-0305

Supplementary Material
Access the most recent supplemental material at:
http://mcr.aacrjournals.org/content/suppl/2014/09/23/1541-7786.MCR-14-0305.DC1

Author Manuscript
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.