Shelterin complex gene: Prognosis and therapeutic vulnerability in cancer

Vikas Kumar Bhari, Durgesh Kumar, Surendra Kumar, Rajeev Mishra

**A R T I C L E   I N F O**

**Keywords:**
Telomere
DNA damage
hTERT
Shelterin complex

**A B S T R A C T**

Telomere encompasses a (TTAGGG)n tandem repeats, and its dysfunction has emerged as the epicenter of driving carcinogenesis by promoting genetic instability. Indeed, they play an essential role in stabilizing chromosomes and therefore protecting them from end-to-end fusion and DNA degradation. Telomere length homeostasis is regulated by several key players including shelterin complex genes, telomerase, and various other regulators.

Targeting these regulatory players can be a good approach to combat cancer as telomere length is increasingly correlated with cancer initiation and progression. In this review, we have aimed to describe the telomere length regulator’s role in prognostic significance and important drug targets in breast cancer.

Moreover, we also assessed alteration in telomeric function by various telomere length regulators and compares this to the regulatory mechanisms that can be associated with clinical biomarkers in cancer. Using publicly available software we summarized mutational and CpG island prediction analysis of the TERT gene breast cancer patient database. Studies have reported that the TERT gene has prognostic significance in breast cancer progression however mechanistic approaches are not defined yet. Interestingly, we reported using the UCSC Xena web-based tool, we confirmed a positive correlation of shelterin complex genes TERF1 and TERF2 in recurrent free survival, indicating the critical role of these genes in breast cancer prognosis. Moreover, the epigenetic landscape of DNA damage repair genes in different breast cancer subtypes also being analyzed using the UCSC Xena database.

Together, these datasets provide a comprehensive resource for shelterin complex gene profiles and define epigenetic landscapes of DNA damage repair genes which reveals the key role of shelterin complex genes in breast cancer with the potential to identify novel and actionable targets for treatment.

**1. Introduction**

Human biology revolves around one term i.e., Deoxyribonucleic acid (DNA), and underpinning the role of DNA, in today’s era of science, therefore, it becomes obvious that its protection is critical for our survival. However, the discovery and function of telomeres were hypothesized even way before the DNA was identified as genetic material. While conducting radiation experiment on Drosophila chromosomes, Moorhead described that normal human cells have finite life span (later this phenomenon is known as Hayflick limit) [3]. Around the same time (1940) Barbara McClintock while working with Zea Mays (where dicentric chromosomes were produced with high frequency) reported that the absence of special end structures of chromosomes led to chromosome instability a detrimental factor for cell survival, congregated on the same conclusion drawn by Muller [2]. Both scientists were awarded Nobel prizes in physiology or Medicine (Hermann Muller in 1946 and Barbara McClintock in 1983) but not for a discovery made in the telomere biology field [3], however, these pioneering studies led the strong foundation for future telomere research [1,2]. Hayflick and Paul Moorhead described that normal human cells have finite life span when cultured *in vitro* (later this phenomenon is known as Hayflick limit or replicative senescence) which was contrary to earlier belief where it has been demonstrated that all cells have infinite life span [4]. Telomere shortening has been assumed to act as a “mitotic clock” that limit the cell cycle number [5] and critically short telomere eventually triggered...
cellular senescence [6]. More interestingly, biological aging can be correlated by telomere length as aged persons had shorter telomeres than younger ones [7,8]. Initially, telomere related discoveries were conducted in ciliated protozoa Tetrahymena thermophila, which contain abundant short linear chromosome and therefore provided a unique model system where the sequencing of telomeric DNA was conducted, leading to the discovery of the molecular basis of telomere [9]. Later it has been found that telomeric DNA is the repetitive sequence of TTAGGG in humans [10,11] as well as all vertebrates [12] and variation in tandem repeats i.e (TTAGGG)n, led to variation in telomere length from chromosome to chromosome. In humans, this unique model system where the sequencing of telomeric DNA was conducted, leading to the discovery of the molecular basis of telomere [9]. Telomere related discoveries were correlated by telomere length as aged persons had shorter telomeres than younger ones [7,8]. Initially, telomere related discoveries were conducted, leading to the discovery of the molecular basis of telomere [9].

V.K. Bhari et al.

Later it has been found that telomeric DNA is the repetitive sequence of TTAGGG in humans [10,11] as well as all vertebrates [12] and variation in tandem repeats i.e (TTAGGG)n, led to variation in tandem repeats i.e (TTAGGG)n, led to variation in telomere length from chromosome to chromosome. In humans, this variation ranges from 3 to 20 kilobase pairs. DNA polymerase needs a labile primer to initiate unidirectional 5′–3′ synthesis to copied full DNA at the 3′ end during DNA replication. Therefore, DNA polymerase unable to synthesize full DNA at the end (3′ end) during DNA replication, and for this reason, this has been termed as an end-replication problem [13,14]. To understand how cancerous cell bypasses this end-replication problem, Greider and Blackburn identified that the presence of telomerase enzyme ensures the cell to bypass telomeric shortening problem [15]. The significance of telomere was demonstrated in many ways, which include coverage for the coding DNA at the end, therefore avoiding loss of genetic information in linear DNA. Moreover, telomere also acts as a cap to protect chromosomas from being recognized as Double-strand breaks (DSBs) avoiding its degradation [16]. The telomerase enzyme is a large multisubunit ribonucleoprotein complex, which adds telomeric DNA repeats at the end of chromosomess. In human, telomerase activity is driven by a complex of two major components: human telomerase reverse transcriptase (hTERT) encoded by the TERT gene which acts as a functional catalytic protein having reverse transcriptase capacity, and the RNA component (also known as human telomerase RNA component: hTERRA or hTR) which acts as a template for reverse transcription [17,18]. Both these components are located at chromosomal region 5p15.33 and 3q26 respectively in humans [19]. This enzyme has been demonstrated that it’s components are largely conserved across the species from mammals to vertebrates, therefore demonstrating its critical evolutionary role in biology [20]. Compelling evidence has also accumulated that human telomerase promoter mutations are also strongly correlated with poor disease outcomes [21–23]. Additionally, regulation of telomere length and telomerase enzyme expression also controlled through transcriptional, post-transcriptional, and epigenetic modifications, and therefore gaining insight into these mechanisms will not only provide novel biomarkers and but also help in early detection of disease, identifying novel prognostic markers and therapeutics [24]. Epigenetic marks like trimethylation of histones of certain amino acids also regulate the telomere length, along with some others such as Telomeric repeat-containing RNA (TERRA) [25,26]. Recently, studies related to protein structural resolution of telomere such as a recent cryo-electron microscopy study [27], could be the key finding for targeting human telomerase by small-molecules in days to come.

2. Role of telomere length dynamics in cancer progression and development

Cancer is an umbrella term for a large group of diseases that can be characterized by many capabilities leading to many complications in the human body. These capabilities are best described in one of the highly cited review articles entitled “hallmarks of cancer” by Hanahan and Weinberg in 2000 and 2011 [20]. These hallmarks provide a solid foundation for understanding the biology of cancer which is relevant to today’s era as well. A highly complementary hallmark, by which cancer cells can acquire the capability for nourishing proliferative signaling to escape growth suppression and adept for indefinite cell division and growth (immortality) [28]. During cell division, there is a reduction of 50–200 bp of the nucleotide sequence, which is reduced until an acute edge is reached and normal cells undergo replicative senescence [29]. However, cancer cell escapes this replicative senescence crisis and maintain stable but usually shortened telomere lengths by dictating steady level telomerase expression for continued cell growth, which allow them to divide indefinitely [30,31]. Interestingly, in human somatic cells, TERT expression is silenced, however, it’s expressed in ~90% of human tumors [32,33] demonstrating the critical role of this enzyme in maintaining cell proliferating capacity of cells, an important hallmark of cancer cells. Due to the critical role of telomerase in cancers various mechanism has been elucidated for hTERT activation, which includes genetic (promoter mutations), amplification, pre-mRNA alternative splicing, epigenetic modification (promoter methylation and miRNAs) and telomere position effect [34,35]. The telomere is also been explored for therapeutic benefit as short telomeres can deliver deleterious effects against tumor cells as they have proliferative characteristics [36]. Telomerase is a good target for anti-cancer drugs as it expresses more in tumor cells than any other marker. Oligonucleotide based drug like Imetelstat (GRN163) is already popular and efficient and small synthetic drug BIBR1532 is one of the most potent inhibitors of human telomerase [37]. BRACO19 and telomestatin are G-quadruplex stabilizing ligands that inhibit the telomerase to bind to the telomere DNA [38]. An anti-telomere drug such as T-oligo is a well-developed drug used to inhibit telomere from various DNA damage responses (DDR) [39]. The development of anti-cancer drugs requires a proper understanding of processes and mechanisms which regulate telomere and its length. Thus, targeting telomerase enzyme and its access to telomeric end thereby regulating telomere homeostasis may be considered a promising approach for cancer therapeutics [40–42].

3. Canonical regulation of telomere length by telomerase enzyme: TERT

Telomerase enzyme was first reported in the early 1980s, however, its genes for RNA component (TER or TERC) was cloned in 1995 [17] and its core protein component (also called TERT) was cloned in 1997 [43,44]. Cryo-EM structure of telomerase revealed that TERT protein contains four domains: Telomerase essential N-terminal (TEN) domain, TERT RNA binding domain (TRBD), Reverse transcriptase domain (RT), and a C-terminal extension (CTE) domain [27]. In most organisms, the telomerase enzyme is a predominant factor for telomere elongation. Initial discoveries involved the origin of the “protein counting model” in S. cerevisiae, which stated that telomere bound proteins inhibit the telomerase access to the telomere end. According to this model, the longer the telomere, have the abundance of telomeric binding proteins, (which play an inhibitory role for telomerase recruitment) hence more the repressive effect on the telomerase on the other hand short telomeres have a less strong repressive effect and hence telomerase can elongate them. This model further gets acceptance in the mammalian system as knockdown of telomere binding proteins resulted in elongated telomeres [45,46]. However, this model was unable to explain how telomere-bound proteins can halt telomerase access from long distances. Later “replication fork” model was better able to explain this problem which stated that telomerase must be deposited at the end of the telomere to elongate it by traveling across the telomere length through the replication fork [47]. Telomere bound proteins and probably nucleosome exerts a negative effect on telomerase movement on replication fork thereby dissociating it and lowering the probability of telomerase reaching the end, where it has to perform its catalytic function [47]. Therefore, short telomeres have a high probability of telomerase to be associated with the telomere and perform its catalytic function to elongate the telomere. Accumulating evidence suggests that G-strand overhang availability is also a major factor for promoting telomerase activity which is positively regulated by POT1-TPP1 heterodimer by inhibiting ATR kinase [48]. Hence, it has been revealed that loss of telomere end protection, telomere length increases aberrantly due to the unregulated access of telomerase to the telomere [45,49,50].

Interestingly, among two important components of telomerase
enzyme i.e. TERC and TERT, all human tissues express TERC, while most human adults do not express TERT [51]. TERT gene was silenced in a tissue specific manner during early human fetal development while constitutively expressing telomerase [52]. Apart from its canonical role in telomere length regulation, compelling evidence has been accumulated the essential role of TERT, in cellular immortalization and malignant transformation by stabilizing telomere length and erasing the senescence barrier [53,54]. TERT expression/telomerase reactivation remains a ubiquitous process, which can be detectable in more than 90% of human primary cancers [55], constituting it’s one of the important major hallmarks of cancer [56], however its precise mechanism of activation is still unclear [53]. Here we listed some of the mechanisms of TERT regulation in cancers.

3.1. TERT gene mutations

The TERT gene (42 kb long) comprises 16 exons and 15 introns located on the short arm of chromosome 5 (5p.15:33) [18]. TERT promoter is rich with various transcription factor binding motifs such as Sp1 and oncogene MYC. The most prominent form of mutation of the TERT promoter is C > T mutation at −146 or −124 upstream from the start codon (ATG) which is termed as C228T and C250T respectively. Both these mutations are associated with the generation of ETS (E-twenty-six) transcription factors binding motifs (CCGGAA) [57,58]. Family members of ETS transcription factors such as GA binding proteins GABPA and GABPB1 are specifically recruited are main drivers for the mutated promoter activation by promoting TERT transcription and telomerase activation [59]. This mutant promoter displays active chromatin marks such as H3K4me2/3 while wild type promoter remains epigenetically silenced suggested by the H3K27me3 [60]. Studies suggested that TERT promoter mutations are negatively correlated with telomere length [61] leading to genetic instability. To explore whether TERT promoter mutations can be correlated with prevalence and prognostic significance in human breast cancer, we analyzed the COSMIC database (https://cancer.sanger.ac.uk/cosmic/gene/analysis?ln=TERT) [62]. Analyzed results described information which includes nonsense, missense, synonymous, substitution, deletion frame, and insertion frameshift mutations. We found that TERT mutations are largely missense and non-sense in Breast cancer (BC) samples (Fig. 1A). BC had 36.09% G > A and 35.56% C > T mutation in TERT coding strand (Fig. 1B).

We also analyzed compare the frequency of TERT gene alterations in 9344 samples using the cbioPortal online tool (https://www.cbioportal.org) [63,64]. The summary graph of TERT gene alterations (shown in different colors) in individual breast cancer (total 9 studies) is shown in Fig. 2. Obtained results indicate more than 8% TERT mutation in BC patients. The distribution of the hTERT mutations is shown in the inset (Fig. 2).

3.2. TERT amplification and TERT rearrangements

TERT is also activated by several genomic rearrangements that affect the TERT locus. It brings enhancers in close contact with the coding part of the gene. Studies reveal that these rearrangements involve direct overlapping between super-enhancers and the locus of the gene which results in significant chromatin remodeling and subsequent activation of TERT transcription [65,66]. MYCN amplification and ATRX mutation are exclusive during the process of TERT rearrangement observed mainly in high-grade neuroblastoma [66]. Genetic instability is one of the important hallmarks of cancer as it favors cancer cells to accumulate a large number of mutations. Gene amplification refers to the differential increase in a specific portion of the gene and can be considered as an important event in oncogenesis which imparts cancer cells to achieve overexpression of some oncogenes [67]. Recent reports have documented that TERT gene amplification was observed in 4% of tumors in...

![Fig. 1. COSMIC database analysis of TERT mutation distributions and its types A. TERT mutations showed in the pie chart illustrate mostly missense and nonsense in BC. B. TERT coding strands largely included G > A (36.09%) and C > T (35.56%) mutations. TERT, Human telomerase reverse transcriptase; BC, breast cancer.](image-url)
the systematic analysis of the TERT gene [68]. Our STRING interaction analysis (https://string-db.org/) revealed that TERT protein involved in many molecular and biological process and has high-confidence interactions with Dyskerin pseudouridine synthase 1 (DKC1), MYC oncogene, XRCC5 (X-Ray Repair Cross Complementing 5), XRCC6 (X-Ray Repair Cross Complementing 6), ATP-dependent DNA Helicase PIF1 (PIF1), telomerase-binding protein EST1A (SMG6), TERF1-interacting telomerase inhibitor 1 (PINX1), NHP2 ribonucleoprotein (NHP2) and SWI/SNF related, matrix associated, actin-dependent regulator of chromatin, subfamily A, member 4 (SMARCA4) and serine-threonine protein kinase (AKT1) (Fig. 3A).

3.3. Epigenetic regulation of TERT promoter

CpG sites are usually the hotspots for the DNA methylation to occur, the process mediated by DNA methyltransferase enzyme which catalyzes the addition of methyl groups (-CH3) at the 5th carbon of the cytosine base. Using CpG island prediction software (https://www.urogene.org/methprimer/) [69], we found that the TERT promoter is rich in CpG island (Fig. 3B). Due to this frequent presence of CpG islands on gene promoters, DNA methylation unarguably plays an important part in gene expression [70]. The TERT gene contains around 52 CpG sites within its promoter and this region is identified as TERT hypermethylated oncological region (THOR). Hypermethylation of THOR led to TERT up-regulation which is an important epigenetic mechanism for malignant transformation [71] Hypomethylation is also observed to initiate the binding of transcriptional activators such as c-MYC on the contrary hypermethylation led to inhibit the binding of repressor proteins like CTCF and Wilms tumor 1 (WT1) to the promoter region of the TERT [34,71,72]. All these findings conclude that dynamic changes in methylation patterns are an important regulator of TERT transcription.

4. Role of shelterin complex in telomere length regulation

Telomere protects the chromosomal ends to be identified as Double-Stranded Breaks (DSBs) but the telomere itself needs to be protected from DDR [73,74]. So, to protect the protector of telomeric DNA, there lies the most important complex associated with telomere which is known as shelterin complex. It consists of (Fig. 4A) six major proteins-Telomeric repeat-binding factor 1 (TERF1), TERF2, Telomeric repeat-binding factor 2 (TERF2), TRF1-interacting nuclear protein 2 (TIN-2), Protection of telomeres 1 (POT1), Repressor and Activator Protein 1 (Rap-1) and TPP1 or Adrenocortical dysplasia protein homolog (ACD). Shelterin complex is also known as telosome due to its specific role in protecting telomere.

Fig. 2. TERT mutation analysis and visualization using cBioportal database. Accumulated alterations of TERT in BC are summarized in the graph in individual studies deposited in the database. Furthermore, the distribution of these alterations are shown in the inset, black arrows showing nonsense mutations while others showing missense mutation. TERT, telomerase reverse transcriptase; BC, breast cancer.

Fig. 3. Bioinformatics analysis of CpG island prediction on TERT gene promoter and interaction network using the STRING database. (A) TERT, and its strongly interacting protein interaction network derived from the STRING database. In a biological network, a node is any biological molecule and an edge indicates the interaction between the two nodes. (B) CpG island presence near the promoter of a gene is strongly associated with methylation mark recruitment. Our bioinformatics analysis provided the information that there is a heavy abundance of CpG island near the TERT gene promoter. As the TERT promoter contains a site for various transcription factors, these CpG islands may play a role in the alteration and regulation of TERT transcription. The location of the CpG islands is colored in light blue. TERT, Human Telomerase reverse transcriptase. TERT, human telomerase reverse transcriptase. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)
4.1. TERF-1

TERF-1 is a homodimer protein binding directly to the duplex region of the telomere. Both TERF1 and TERF2 contain a TERF homology (TRFH) domain which helps in homodimerization and a Myb domain which is highly specific to the telomeric DNA [75]. There are many theories reported in the past as to how TERF-1 regulates the telomere length. Studies have reported that in the presence of TERF-1, telomere length is decreased [49, 50]. With the help of the chromatin immunoprecipitation (ChIP) technique, Loayza and de Lange reported that TERF-1 regulates telomere length via POT1, either directly blocks the telomerase binding or it stabilizes the t-loop structure of telomere, which means telomerase cannot bind to 3' end [49]. This means TERF-1 does not interact with telomerase directly. Tankyrase-1 overexpression is the most effective method to remove TERF-1 from the telomere [76]. Tankyrase-1 is an enzyme of the Poly (ADP-Ribose) polymerase (PARP) superfamily. PARPs generate the ADP-ribose polymers on the glutamic acid residues of the protein acceptor, which in this case is TERF-1, which results in inhibition of TERF-1 binding to the telomere [77]. Overexpression of the tankyrase enzyme in the nucleus of human cells has shown to remove the TERF-1 from telomeres [76].

4.2. TERF-2

As TERF-1 and TERF-2 are almost similar to each other, TERF-2 is also a homodimer protein and binds to the ds telomeric DNA. It is demonstrated that it helps in the formation of a secondary structure known as the t-loop in which 3’ ss overhang invades into the double-stranded DNA sequence [78]. This remodeling of telomere into t-loop help remove threats of non-homologous end joining (NHEJ) and ataxia telangiectasia mutated kinase (ATM) signaling [79, 80]. It also acts as a negative regulator of the telomere length like TERF-1 as t-loop formation discourages telomerase binding to the 3’ end [50]. TERF-2 has also a role in the protection of telomere from DNA damage signaling by inhibiting the autophosphorylation property of ATM kinase [81].

4.3. TIN-2

TIN2 acts as a linchpin of the shelterin complex. It acts as the bridge between telomeric dsDNA and single-stranded DNA (ssDNA) by connecting TERF1 and TERF-2 to TPP1/POT1 heterodimer, thus stabilizing the shelterin complex [73]. As previously reported that tankyrase-1 PARsylates the TERF-1 and inhibit its binding to the telomere, TIN2 is reported to prevent this action of tankyrase-1 on TERF-1, thereby acting as a negative regulator of telomere length in this case [45, 77]. TIN2 mutation that leads to dyskeratosis congenital were shown to decrease the telomere length without any interference with the telomere protection which gives us the clear idea that TIN2 mutations somehow has inhibited the activity of telomerase [82].
4.4. POT1 and TPP1 - homologs of TEBPα and TEBPβ of oxytricha nova

POT1 and TPP1 together bind to the 3’ ss overhang as a stable heterodimer. TEBPα and TEBPβ were the first discovered proteins bound on the single-stranded telomeric DNA in the ciliate Oxytricha nova [83]. It is reported that TPP1 and POT1 are homologs of TEBPα and TEBPβ [84–86]. TPP1 is a very diverse protein in the sense that it interacts with ss DNA binding protein POT1 to form a heterodimer as well as to TIN2 to connect the ds DNA binding proteins (TERF1 & TERF2) to POT1 [45,87]. This heterodimer also prevents ataxia telangiectasia and Rad3 related (ATR) dependent DDR to act on telomere [88]. Both these events overall increase the stability of the shelterin complex as well as a telomere.

N terminal OB domain of TPP1 contains TPP1 glutamate (E) and leucine (L)-rich (TEL) patch which is also reported to regulate the telomere length by interacting with the telomerase-N-terminal (TEN) domain and as a processivity factor for telomerase during telomerase extension by improving template translocation and decreasing primer dissociation [89,90].

4.5. TERF2-Rap1

In mammals, Rap1 is a highly conserved protein present in the shelterin complex [73]. It is recruited by TERF-2 and they form a heterodimer that is essential in protecting telomere from homology-directed repair (HDR) pathways (SLX4 and PARP1) [91]. The absence of this heterodimer is reported to result in telomere shortening and telomere-free chromosome fusion by homology-directed repair (HDR) pathway [91].

5. Differential mRNA expression analysis of shelterin complex genes and its correlation with prognostic potential in breast cancer

We analyzed the mRNA expression of shelterin complex genes in tumor and normal cells, by using the University of California, Santa Cruz (UCSC) XENA platform (https://xenabrowser.net/) [92]. We selected the TCGA TARGET GTEx study that enables us to compare the expression using the TCGA database of tumor tissue and GTEx database for normal tissue samples in breast cancer patients. Taking this cohort study, we filtered down the data to breast cancer and checked the mRNA expression of shelterin complex genes. RSEM expected count dataset was used for the analysis and values were downloaded as log2 (expected_count +1). The Box plot was designed using GraphPad Prism software (8.4.2) and P < 0.05, were considered as significant. Analyzed data suggested that all the shelterin complex genes upregulated in breast cancer (Fig. 4B) compare to normal samples. To the best of our knowledge, no study reported the role of shelterin complex genes in breast cancer, however, mutation/or altered expression in shelterin components at telomeres have been described in various types of cancers [26]. To identify whether these genes might play a role in breast cancer prognosis or survival we used Kaplan-Meier plotter analysis (https://kmplot.com) and assessed the prognostic significance of the shelterin complex genes in all BC patients [93]. The relationship between survival and gene expression level of individual shelterin complex genes were demonstrated in Fig. 5.

Among the six sheltering complex genes we analyzed, high TERF1 and TERF2 expressions were significantly associated with a poor prognosis (TERF1, HR 1.16 [95% CI: 1.04–1.3], p = 0.0062; TERF2, HR 1.16 [95% CI: 1.1–1.36], p = 0.056; Fig. 5A and B Table 1). However, POT1 and TIN2 mRNA high expression were correlated with significantly better relapse-free survival (RFS) for all BC patients, POT1 (HR 0.85 [95% CI: 0.76–0.94], p = 0.0027; TIN2, HR 0.83 [95% CI: 0.74–0.92], p = 0.0006 (Fig. 5D and F) and Table 1. In addition, the mRNA expression levels of TPP1, and RAP1 were not correlated with RFS (Fig. 5C and E). Our analysis identifies TERF1 and TERF2 may have prognostic potential in breast cancer. This is correlated with a recent study that shows that TERF-1 inhibition using combinatorial therapy can reduce stemness in glioblastoma xenograft models [94].

In conclusion, our data analysis showed that overexpression of TERF1 and TERF2 mRNA is correlated to a poor prognosis for all BC patients and may be a useful tool for prognosis. The possible responsiveness to these shelterin complex genes expressed by breast cancer cells can be the consequence of activation of DNA damage and suppression of DDR pathways. DNA damage causes genetic instability, and

---

Fig. 5. Survival analysis of shelterin complex genes in Breast cancer patients using Kaplan-Meier plotter. Survival curves of (A) TERF1 (Affymetrix ID:203448_s_at); (B) TERF2 (Affymetrix IDs: 1553485_s_at); (C) Rap1 (Affymetrix IDs:201174_s_at); (D) POT1 (Affymetrix IDs:204353_s_at); (E) TPP1 (Affymetrix IDs:204617_s_at); (F) TIN2 (Affymetrix ID:220052_s_at). Abbreviations: TERF1, telomeric repeat-binding factor 1; TERF2.
the accumulation of mutation can be increased during DNA damage. It causes the activation of DDR which repairs the DNA by recognizing it as a DSB and its result can be apoptosis or senescence [95]. Activation of two major kinase – ATR and ATM kinases initiate the process of DDR. It is followed by the DNA damage foci formation characterized by the presence of phosphorylated histone H2A.X and at last cell cycle is arrested when cell cycle checkpoints are alarmed such as p21 and CDK inhibitor [96]. Telomeric DNA is sequestered by the shelterin complex and prevents it from being recognized by DNA damage signaling [97]. Association of DDR proteins such as ATM, 53BP1, and MRE11 with uncapped telomere is observed resulting in cell-cycle arrest or apoptosis [98]. TERF2 and POT1 are also reported to block ATM and ATR kinase respectively [74]. ATM and ATR are also reported to help in the assembly of telomerase components which is essential in telomerase function [99].

6. The aberrant epigenetic landscape of DNA repair genes in the basal subtype of breast cancer

Deficiency in the DNA repair pathway is previously reported to result in tumor development [100]. Therefore, it is possible that tumor functions by inactivating the effectiveness of the DNA repair system by epigenetic gene inactivation which directly affects the DNA repair genes. To investigate the epigenetic landscape of DNA repair genes in breast cancer subtypes, we used the “TCGA Breast Cancer (BRCA)” study in the UCSC XENAG browser. Interestingly, we found out that some of the DNA repair genes showed a contrasting pattern in only basal-like breast cancer subtype. Some of the CpG islands of PARP1, XRCC5 (coding gene for protein Ku80), and PRKDC (coding gene for protein DNA-PK) were either hypomethylated or hypermethylated only in basal-like breast cancer subtype (Fig. 6). DNA-PK, in association with Ku70/Ku80 heterodimer protein, functions in recombination and DNA double-stranded break repair. Epigenetic inactivation or activation of these DNA repair genes can be exploited using targeted therapies.

7. Advancements in clinical therapeutics of anti-cancer drugs—clinical implications of drugs targeting telomere length regulatory factors

How the ends of chromosomes (telomeres) are maintained for the perpetual proliferation of cancer cells cancer cell immortality is still an area of extensive research. Induction of telomerase enzyme is one of the most prominent phenomena occurring in 85–90% of all malignant tumors and thus telomerase inhibitors are an attractive tool for cancer therapeutics [33,101]. However, in the rest 10–15% of cancers — which include some cancers which do have a poor prognosis as well as poor outcome — the telomere elongation is achieved through a mechanism known as alternative lengthening of telomeres (ALT) [102,103]. ALT to adversely impact the outcome of patients having soft tissue sarcoma, mesenchymal, and neuroepithelial tumor [103]. Interestingly, maintenance of the telomere length mechanism and its targeting became more complex as there has been evidence that both the mechanism can coexist within a single tumor [104]. This dynamic nature of tumor in maintaining telomere length suggested that telomerase targeted therapies may require accurate telomere lengthening mechanism characterization before treatment to get a maximum therapeutic benefit.

**Table 1**

Datasets of the shelterin complex genes in breast cancer (Kaplan-Meier plotter).

| Gene | Affymetrix ID | No. of cases | Cut off value | HR   | 95% CI | P-Value |
|------|---------------|--------------|---------------|------|--------|---------|
| TERF-1 | 203448_s_at | 3951 | 700 | 1.16 | 1.04–1.3 | 0.0062 |
| TERF-2 | 1555185_s_at | 1764 | 44 | 1.16 | 1.1–1.36 | 0.056 |
| Rap1/TERF2IP | 201174_s_at | 3951 | 1339 | 1.03 | 0.93–1.15 | 0.56 |
| POT1 | 204353_s_at | 3951 | 422 | 0.85 | 0.76–0.94 | 0.0027 |
| ACD/TPP1 | 204617_s_at | 3951 | 414 | 1.06 | 0.95–1.18 | 0.30 |
| TIN2 | 220052_s_at | 3951 | 585 | 0.83 | 0.74–0.92 | 0.0006 |

**Abbreviations:** HR, hazard ratio; CI, confidence level; TERF1, telomeric repeat-binding factor 1; TERF2, telomeric repeat-binding factor 2; Rap1/TERF2IP, telomeric-repeating binding factor 2-interacting factor; POT1, protection of telomeres protein 1; TPP1/ACD; Adrenocortical dysplasia protein; TIN2/TIN2F, TERF1-interacting nuclear factor 2.
Altogether, the concept of hindering the telomere length regulating mechanism in cancer cells drug targeting telomere length mechanism is an emerging target for cancer therapeutics. Currently, there are a variety of antitelomerase drugs like small molecule drugs (BBR1532), antisense oligonucleotides (ASO- GRN163L), ribozyme, and dominant-negative molecules like DN-TERT are in the trial/developmental phase. Here we are summarizing some telomerase targeting drug molecules which are currently in clinics or the developmental phase.

7.1. Imetelstat

Imetelstat is an FDA approved anti-telomerase drug that is a potent competitive inhibitor of the enzymatic activity of TERT. It is a lipid-based conjugate of oligonucleotide GRN163, thus named GRN163L, lipid components helping in an increase in the bioavailability [105]. Gathering clues from a pilot study for myelofibrosis [106], a recent clinical study in intermediate-2 (int-2), or high-risk myelofibrosis patients who are relapsed or refractory to Janus kinase inhibitor (JAKi) demonstrated favorable clinical activity with an acceptable safety profile [107].

Telomerase inhibition by imetelstat has been shown in mouse xenograft models and cell culture systems like breast [108,109], lungs [109,110], and bladder [111]. In addition to these hematological malignancies like multiple myeloma and lymphoma cancer cells are also being studied extensively for its antitelomerase activity [112].

7.2. MST-312

Epigallocatechin gallate (EGCG) chemical derivative known as MST-312 is reported to possess telomerase inhibitory (TI) activity and can be seen as a prototype of anticancer drugs [113,114]. MST-312 has a lower effective dose requirement and drug resistance and chemically more stable than its parent i.e., EGCG [114]. The mechanism of action of this drug is associated with the downregulation of several genes such as TNF-α, c-Myc, IL-6, and Bcl-2 and related pathways of these genes.

MST-312 has been shown to decrease telomerase activity and promoting telomere dysfunction leading to growth arrest in breast cancer cells [115]. According to reports, MST-312 is believed to express that its effects differ based on exposure time. Short term treatment of 90 days led to telomere attrition and senescence [114]. Studies reported that U-266 cells (myeloma cells), when treated with MST-312, can induce acute growth arrest and apoptosis indicating that this drug may have the potential to treat multiple myeloma (MM) which is described to be an incurable disease [116]. Long term exposure (130 days) of MST-312 resulted in TERT overexpression in breast cancer cells [117]. Similar results were also reported in another study where MST-312 can induce resistance by the selection of long telomere cells in vitro [113]. Therefore, the drawback of this drug is resistance induction due to long term treatment according to the limited available knowledge must be acknowledged.

7.3. BBR1532

BBR (2-[E]-3-naphthalen-2-yl-but-2-enoylamino]-benzoic acid) is a small and non-nucleosidic anti-telomerase drug that is widely used in kinetic, biochemical, and in vivo studies of telomerase function due to its selective nature [118]. It has been shown that BBR1532 binds to the novel FVYL motif of the thumb domain of telomerase and through its binding, it prevents telomerase complex assembly formation [119]. Owing to this property, BBR1532 is reported to be a therapeutic agent that suppresses telomerase activity in breast cancer [120] and glioblastoma [121], acute lymphoblastic leukemia, and Kras mutant NSCLC [122]. Moreover, in combination with doxorubicin, BBR1532 is shown encouraging results to treat acute lymphoblastic leukemia (ALL) [123].

7.4. G-Quadruplex stabilizing ligands

The human genome contains a guanine rich sequence that is capable of forming a four-stranded quadruplex structure which is known as G-quadruplex [124,125]. The presence of these G-quadruplex structures on G-rich telomeric DNA at the chromosomal ends restricts the telomerase access to the telomere and therefore hinders its activity [38]. Small molecules called G4 ligands can facilitate the production of G-quadruplex structures [126,127]. Telomere shortening is induced by G4 ligands which stabilize G-quadruplex on telomeric DNA [128-130]. A study revealed that ligands such as BRACO-19 and RHP84 may require monovalent cations like central K+ to stabilize the G-quadruplex structure [131].

7.5. BRACO-19

BRACO-19 is an important ligand that is proved to stabilize and promote the formation of a G-quadruplex structure on telomere [132]. It is evident by the previous studies that BRACO-19 inhibits the catalytic property of telomerase, induce telomere shortening, and enable telomerase positive tumor cells to senescence [133]. Using the glioblastoma cancer model, it was recently shown that BRACO carries that cytoplasmic transport of telomerase and consequent reduction in telomerase activity. Additionally, the release of telomere-binding proteins was also seen in this study [134]. Therefore, it is obvious that these dysfunctional telomeres activate the DDR and undergo senescence. Recently, with the help of molecular dynamics simulations with the latest AMBER force field, detailed structural information was gathered about the BRACO-19 [135].

7.6. Telomestatin

It is one of the most popular and well-explored drugs and has been reported as a potent telomerase inhibitor [136]. It is a G-quadruplex stabilizer isolated from Streptomyces anulatus 3533-SV4. In a study conducted by Tahara and others, it was found that telomestatin not only affects telomerase but also dissociates TERF2 protein from telomere and induces loss of 3’ single-stranded overhang [137]. Telomere shortening and the apoptotic effect is observed in many cancer types and thus ultimately results in the induction of anti-cancer phenotypes [138-140].

7.7. T-oligo

T-oligo is an oligonucleotide resembling the 3’ overhang of the telomere. T-oligo is reported to activate DDR (p53, transcription factor E2F1, ATM) and anti-cancer responses (apoptosis, senescence, and differentiation) in malignant cells [39,141]. T-oligo is also reported to inhibit growth and induce apoptosis in prostate, ovarian, and breast cancer cells, melanoma and lymphoma [39,141-143].

8. Conclusion

Telomere length as a clinical marker is getting enough attention today in the scientific world. There is a need to explore this field to find new ways to diagnose cancer and treat it. Players of telomere length modulation are specific and interconnecting in function and there is a need to find a common link between them to understand tumor progression. Understanding telomere length homeostasis as well as identifying the key players affecting this homeostasis in cancer can help to unravel new therapeutic opportunities.

Authors’ contributions

RM and VKB wrote and edit the manuscript; RM, VKB, DK and SK contributed to writing and provided experimental results and information. The author(s) read and approved the final manuscript.
Declaration of competing interest

None.

Acknowledgment

The work was supported in part by Start-up Research Grant (SRG/2019/001049), from the Science and Engineering Research Board (SERB), New Delhi, India, and Enhanced Seed grants EF/2018-19/QF04-11 (to R.M.) from Manipal University Jaipur, Rajasthan, India is gratefully acknowledged.

References

[1] H.J. Muller, LH. Hershkowitz, Concerning the healing of chromosome ends produced by breakage in Drosophila melanogaster, Am. Nat. 88 (1954) 177–208, https://doi.org/10.1086/281830.
[2] B. McClintock, The stability of broken ends of chromosomes in Zea mays, Genetics 26 (1941) 234–282, http://www.ncbi.nlm.nih.gov/pubmed/17247004%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=P MC1290127.
[3] V.A. Zakian, The ends have arrived, Cell 139 (2009) 1038
[4] L. Hayflick, P.S. Moorhead, The serial cultivation of human diploid cell strains, Exp. Cell Res. 25 (1961) 585–621, https://doi.org/10.1016/0141-0216(61)90035-3.
[5] A.M. Olovnikov, Telomeres, telomerase, and aging: origin of the theory, Exp. Cell Res. 25 (1961) 585–621, https://doi.org/10.1016/0014-4827(61)90035-3.
[6] E.H. Blackburn, C.B. Harley, A.B. Futcher, C.W. Greider, Telomeres shorten during ageing of human fibroblasts, Nature 346 (1990) 866–868, https://doi.org/10.1038/346866a0.
[7] C.B. Harley, A.B. Futcher, C.W. Greider, Telomeres shorten during ageing of human fibroblasts, Nature 345 (1990) 458–460, https://doi.org/10.1038/345458a0.
[8] E.H. Blackburn, J.G. Gall, A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in Tetrhyymena, J. Mol. Biol. 120 (1978) 33–53, https://doi.org/10.1016/0022-2836(78)90294-2.
[9] V.K. Bhari et al..
