Telomeres expand sphere of influence: emerging molecular impact of telomeres in non-telomeric functions

Soujanya Vinayagamurthy, Sulochana Bagri, Jean-Louis Mergny, Shantanu Chowdhury

Trends in Genetics

CellPress OPEN ACCESS

Review

Telomeres expand sphere of influence: emerging molecular impact of telomeres in non-telomeric functions

Soujanya Vinayagamurthy,[1,2,6] Sulochana Bagri,[1,2,6] Jean-Louis Mergny,[3,4] and Shantanu Chowdhury [1,2,5,*]

Although the impact of telomeres on physiology stands well established, a question remains: how do telomeres impact cellular functions at a molecular level? This is because current understanding limits the influence of telomeres to adjacent subtelomeric regions despite the wide-ranging impact of telomeres. Emerging work in two distinct aspects offers opportunities to bridge this gap. First, telomere-binding factors were found with non-telomeric functions. Second, locally induced DNA secondary structures called G-quadruplexes are notably abundant in telomeres, and gene regulatory regions genome wide. Many telomeric factors bind to G-quadruplexes for non-telomeric functions. Here we discuss a more general model of how telomeres impact the non-telomeric genome – through factors that associate at telomeres and genome wide – and influence cell-intrinsic functions, particularly aging, cancer, and pluripotency.

Highlights

Shelterin proteins are known for telomeric functions only. Emerging evidence shows non-telomeric presence and function of shelterin factors.

DNA secondary structure G-quadruplex motifs are enriched in gene promoters. Recent work shows how G-quadruplex-dependent non-telomeric binding of telomere-associated proteins regulates transcription.

Interstitial G-quadruplex motifs associate with telomeric protein in a telomere-length-dependent fashion.

Introduction

The terminal ends of linear chromosomes called telomeres comprise nucleoprotein complexes, critical for protecting telomere–telomere fusion and DNA-damage-induced genome instability [1]. Human telomeres comprise 5′-TTAGGG-3′ repeats ranging from 3 to 15 kb including ~150–300-bp 3′ single-strand overhang at the end that aids formation of the T-loop protecting telomeres from the cellular DNA-damage-repair machinery [2]. Notably, telomeric repeats were found to also form secondary structural motifs called **G-quadruplexes** (see Glossary) (also known as G4s) [3] both in solution and *in vivo* [4]. Further, telomeric G4s were reported to associate with telomeric factors that are critical for telomere homeostasis, implicating G4s in telomere protection or uncapping [5,6].

Telomeres widely influence crucial functions that impact cell and organismal biology, including cancer, pluripotency, and aging/senescence. Intriguingly, however, understanding of telomere-dependent molecular mechanisms is largely limited to the subtelomeres (~10 Mb from telomere ends). This is typically through **telomere looping**, subnuclear compartmentalization, and/or post-translational modifications (PTMs) of telomeric factors (reviewed in [7]). Therefore, how telomeres exert influence beyond subtelomeres at the molecular level – an issue possibly central to many telomere-related functions – remains poorly understood. Understanding from two somewhat distinct emerging areas might offer opportunities to bridge this gap.

First, several telomere-associated factors (TAFs) [comprising the multiprotein complexes shelterin (TRF1, TRF2, RAP1, TIN2, TPP1, and POT1) and CST (CTC1, TEN1 and STN1), the reverse transcriptase enzyme TERT along with its RNA template **TERC**, and telomeric-repeat-containing

[1]Integrative and Functional Biology Unit, CSIR Institute of Genomics and Integrative Biology, New Delhi 110025, India
[2]Academy of Scientific and Innovative Research (AcSIR), Ghaziabad 201002, India
[3]Institute of Biophysics of the CAS, v.v.i. Královopolská 135, 612 65 Brno, Czech Republic
[4]Laboratoire d'Optique et Biosciences, Ecole Polytechnique, CNRS, INSERM, Institut Polytechnique de Paris, 91128 Palaiseau, France
[5]GNR Knowledge Centre for Genome and Informatics, CSIR Institute of Genomics and Integrative Biology, New Delhi 110025, India
[6]Equal contributions

*Correspondence: shantanuc@igib.in (S. Chowdhury).

noncoding RNA (TERRA) (Figure 1)] were reported to interact with DNA and/or other regulatory factors outside telomeres. Second, the presence and genomic distribution of G4s is of interest. G4s were first reported to be present at telomeric ends in the late 1980s [8]. Later, G4s were noted to be enriched in gene regulatory regions throughout the genome [9–11], and a growing body of evidence shows that G4s are involved in gene regulation including epigenetic modifications [12]. Further, the function of TAFs like TRF2 and RAP1 was found to be dependent on association with non-telomeric G4s [13–15]. Together, these observations raise questions with important implications. Why are G4s abundant in distinct regions of the genome: telomeres and gene regulatory regions? With these in mind, we describe here recent work on how telomeres influence function at remote (non-telomeric) sites and how non-telomeric and/or telomeric G4s could be involved.

TAFs in non-telomeric roles

Non-telomeric functions of RAP1

Recent findings reveal many non-telomeric functions of TAFs (summarized in Table 1). ChIP-seq in mouse embryonic fibroblasts (MEFs) revealed RAP1 binding to both telomeric and non-telomeric sequences. More than 30 000 non-telomeric sites were identified; ~70% of these mapped in intragenic regions or in the near vicinity of coding regions, suggesting a function of non-telomeric RAP1 in gene regulation. This was further substantiated by deregulated expression of genes harboring RAP1-binding sites in RAP1-deficient MEFs [13]. Non-telomeric RAP1 was also reported to impact the senescence of human mesenchymal stem cells (HMSCs) through epigenetic regulation of *RELN*, a negative regulator of proliferation. Hypermethylation of the *RELN* promoter was noted in RAP1-deficient HMSCs, which was rescued on the expression of RAP1 exogenously (Figure 2A) [16]. In addition, RAP1-mediated regulation of *PGC1α* and *PPARα* was reported in the cellular metabolism of obesity [17,18]. Independent of DNA binding, RAP1 interacts with IκB kinase (IKK)-activated nuclear factor kappa B (NF-κB) [19] (Figure 2B).

Non-telomeric functions of TRF1 and TRF2

Besides RAP1, TRF1 and TRF2 were reported to bind to ~180 interstitial sites [13–15]. Recently, TRF2 binding was reported at thousands of non-telomeric sites [20], and TRF2-meditated regulation of several genes was observed (described in the following section) [21–25] (Figure 2C). Non-telomeric TRF2 was also shown to regulate vascularization through VEGF-A [26] and *PDGFR-β* [21], to reduce the recruitment and activation of natural killer (NK) cells [27], and to induce the recruitment of myeloid-derived suppressor cells (MDSCs) that promote tumor growth through immunosuppression [28].

Further, TRF2 induced higher-order DNA looping and chromatin compaction [29] and 3D looping of telomeres with interstitial chromatin was stabilized by TRF2–lamin association [30,31]. TRF2 silencing reduced expression of the cancer stem cell markers *Oct4*, *Sox2*, *KLF4*, and *c-Myc* [32] and computational modeling suggested direct interaction of KLF4 with TRF2 [32]. Sequestration of the RE-1 silencing factor REST in the cytoplasm by a TRF2 splice isoform results in the activation of REST target genes resulting in differentiation of neural progenitor stem cells (NPCs) [33]. The splice isoform of REST in humans, hREST4 was reported to interact with TRF2 and play a vital role in neural differentiation [34]. Similarly, TRF1 was found to be essential for **induced pluripotent stem cell (iPSC)** proliferation and directly regulated by Oct3/4 [35]. Accordingly, small-molecule inhibitors of TRF1 reduced the efficiency of reprogramming in mice [36]. Furthermore, the role of telomere length and the functions of TAFs in maintaining the pluripotent or 'stem-like' state was reviewed recently [37]. In addition, TRF1 was reported to positively regulate Aurora B's centromeric function during chromosomal segregation [38] and interaction of TRF1 with Nek2, a cell cycle regulator, induced cytokinetic failure [39].

Glossary

G4-binding ligands: small molecules that can specifically bind to G-quadruplex structures.
G-quadruplex: a DNA secondary structure formed by G-rich sequences.
Interleukin-1 receptor-1: is a cytokine receptor for IL1A, IL1B, and IL1RN. A mediator of inflammatory and immune responses.
Induced pluripotent stem cells (iPSCs): mimic ESCs. Somatic cells are induced to form pluripotent cells by expressing the pivotal genes *Oct4*, *Sox2*, *Klf4*, and *cMyc* (Yamanaka factors).
Myb domain: a helix-turn-helix DNA-binding domain present in several eukaryotic transcription factors; also found in TRF1 and TRF2 proteins.
Potential G4s (PG4s): computationally identified regions in the genome containing four consecutive guanine tetrads interspersed with 1–15 nucleotides, which can theoretically adopt a G-quadruplex structure.
PRC2: a multiprotein repressor complex. This methyltransferase methylates H3 on lysine 27 (H3K27me3).
Replicative senescence: an event in which, beyond a certain number of divisions, cells undergo irreversible cell cycle arrest.
RGG motif: arginine (R)-glycine (G)-glycine (G) amino acid repeats found in G-quadruplex-binding proteins; crucial to the G4-binding affinity of those proteins.
Telomere looping: occurs when telomeres loop back to the chromatin (~10 Mb away) inducing chromatin compaction.
Telomere position effect (TPE): repression of genes that are proximal to the telomeres.
Telomeric repeat amplification protocol (TRAP): assay determining the amount of active telomerase holoenzyme by exploiting the enzyme's ability to elongate telomeric repeats templates.
***TERC*:** the RNA component of the telomerase holoenzyme. It acts as template for telomere synthesis during DNA replication.

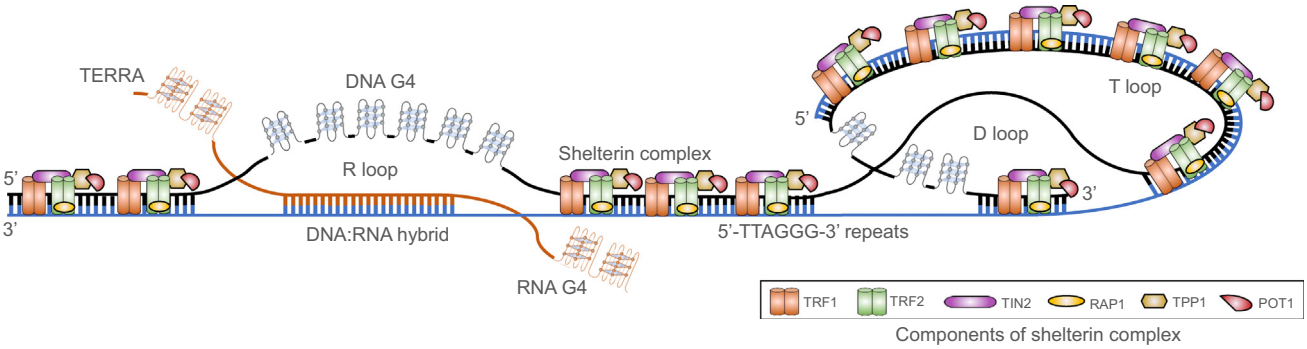

Figure 1. Structural organization of human telomere. The 3′ single-strand telomeric overhang comprising TTAGGG repeats folds back into the double-stranded telomeric DNA. This results in the T-loop and D-loop. The multiprotein shelterin complex binds telomeric DNA and protects telomere ends from the DNA-damage-repair machinery preventing telomere–telomere fusions, controls telomere homeostasis, and stabilizes the overall structure. Telomeric-repeat-containing noncoding RNA (TERRA) can form RNA G-quadruplexes and bind to telomeric DNA forming DNA–RNA hybrids and R-loops.

Non-telomeric functions of telomerase and other TAFs

Although initially only two molecules per cell of hTERT were reported [40], later this was found to be ~60–100 000 molecules per cell supporting the possibility of non-telomeric functions [41]. Importantly, transcriptional regulation of the Wnt/β-catenin genes by hTERT was shown to affect cancer progression [42]. Further, hTERT was found to interact with p65 regulating the expression of NF-κB target genes [43] and with RNA polymerase III (RNA Pol III) to enhance tRNA expression [44,45].

RNA-FISH and CHIRT-seq in mouse embryonic stem cells (mESCs) revealed TERRA-binding sites throughout the genome, implicating another TAF in non-telomeric function. TERRA was reported to regulate the expression of thousands of target genes [46] including the innate immunity genes like *STAT1*, *ISG15*, and *OAS3* in human cancer cells [47]. TERRA deletion triggered differentiation; overexpression enhanced the self-renewal of mesenchymal stem cells [16]. TERRA maintained the transcriptional landscape of ESCs in a TRF1-dependent manner [48] and aggregation of TERRA was noted in mouse medulloblastoma and proliferating neural progenitor cells [49].

The role of shelterin proteins in noncanonical functions in cancer progression/initiation was reviewed recently [50]. Interestingly, mitochondrial localization of another TAF, TIN2, (Figure 2B) and TIN2-mediated regulation of reactive oxygen species leading to HIF-1 activation were implicated in cancer [51,52]. In addition, non-telomeric function of TAFs was reported in yeast and zebrafish. In yeast, RAP1 binds to promoters and represses histone-coding genes on telomere shortening resulting in the loss of histone-mediated suppression of senescence genes [53,54]. In zebrafish, an ortholog of TRF2 was noted to regulate the expression of neurodevelopment- and aging-related genes [55,56].

Emerging impact of telomeres on non-telomeric functions

Telomere position effect (TPE)

Telomere-dependent suppression of genes, known as the **TPE**, was first noted in *Saccharomyces cerevisiae* in *URA3*, *TRP1*, *HIS3*, or *ADE2* inserted ~4.9 kb from telomeres. This was due to telomeric heterochromatinization of the adjacent regions [57]. Silencing of genes in the subtelomeric region through the TPE was widely observed in diverse species like *Trypanosoma brucei*, *Plasmodium falciparum*, *Schizosaccharomyces pombe*, *Drosophila melanogaster*, *Pneumocystis carinii*, *and Candida glabrata* [58], including human carcinoma HeLa cells [59].

Table 1. Non-telomeric functions of TAFs

TAF	Function	Cell type	Gene associated	Phenotype	Refs
TRF1	Epigenetic regulation	iPSC	*MYC*, *SOX2*, *NANOG*, *BMP*	Maintenance of stemness	[35,36,48]
TRF1	Centromeric localization of Aurora B kinase			Maintenance of microtubule–kinetochore attachment	[38]
TRF1	Nek2 directly binds and phosphorylates TRF1 leading to abnormal chromosomal alignment at metaphase	Breast cancer cells – MDA-MB231 and MCF7		Cytokinetic failure and aneuploidization	[39]
TRF2	Transcriptional regulation		*PDGFRβ*	Regulation of angiogenesis	[21]
TRF2	Transcriptional and epigenetic regulation	Patient-derived glioblastoma	*hTERT*	Suppresses reactivated human telomerase in glioblastoma	[24]
TRF2	Regulation of interleukin signaling and tumor-associated macrophage (TAM) infiltration in triple-negative breast cancer (TNBC)	Patient-derived TNBC tissue	*IL1R1*	Antitumor immune response	[25]
TRF2	Transcriptional and epigenetic regulation		*PCGF3*, *OPN4*, *P21*, *ANXA2*, *CHRM2*, *INHA*, *KCNH2*, *OBSL1*, *SAMD14*, *SMAD7*, *THRA*	Regulation of gene expression	[20,22,23]
TRF2	DNA damage repair, migration potential maintenance	Cal27 and SCC-131	*CD44*, *OCT4*, *SOX2*, *KLF4*, *c-MYC*	Maintenance of oral cancer stemness	[32]
TRF2	Association with lamin proteins	IMR90s	Genome-wide interstitial sites (ITSs)	Chromatin organization	[30,31]
TRF2	Stabilization of REST and hREST protein	SH-SY5Y, H1, and H7		Maintenance of neuronal stemness and formation of NPCs from ESCs	[33,34]
TRF2	Transcriptional regulation	Colon cancer	*HS3ST4*	Regulates NK cell density and mobilization	[27]
TRF2	Transcriptional regulation for recruitment and activation of MDSCs	B16F10 melanoma and BJcl2 41ANs	*GPC6*, *VCAN*, and *HS3ST4*	Promotes tumor growth via immune suppression	[28]
TIN2	Regulation of oxidative phosphorylation, glucose metabolism	HT-1080 and U2OS		Mitochondrial morphology	[51]
RAP1	Transcriptional regulation	MEF	*HS3ST1*, *NNMT*, *CAR6*, *IGF2*	Metabolism dysregulation in RAP1 knockout mice	[13]
RAP1	Epigenetic regulation	HMSCs	*RELN*	Senescence	[16]
RAP1	Transcriptional regulation	Mouse hepatic and adipose tissue	PCG1α, *PPARα*	Obesity	[17,18]
RAP1	Regulation of IKK-mediated phosphorylation of NF-κB	MDA-231, HeLa, A549		Tumor cell survival and invasion	[19]
TERT	Cofactor of β-catenin transcriptional complex	HeLa, mouse small intestine cells	*AXIN2*, *MYC*, *HPRT*	Cell proliferation, survival, and stemness	[42]
TERT	Interacts with p65 and regulates expression of inflammatory genes	Mouse-derived fibroblasts	NF-κB-regulated genes – *IL-6*, *IL-8*, *TNF-α*	Cell proliferation, immune regulation	[43]
TERT	Interacts with RNA Pol III	hESC-1, A2780, HCT116, P493, BLM, LOX-IMVI	Binding at rRNA- and tRNA-encoding genes in a cell-type-specific manner	Promotes cell proliferation	[44]
TERRA	Transcriptional regulation	PC-3, HBC4, MKN74	*STAT1*, *ISG15*, *OAS3*	Tumor suppression	[47]

Later intrachromosomal looping – where telomeres loop to regions ~10 Mb away – was observed in human myoblasts. This induced the repression of genes near telomeres (*ISG15*, *DSP*, and *C1S*), further extending observations on telomere-dependent gene expression, and was called TPE over long distance (TPE-OLD) [60] (Figure 3). Telomere shortening disabled the physical

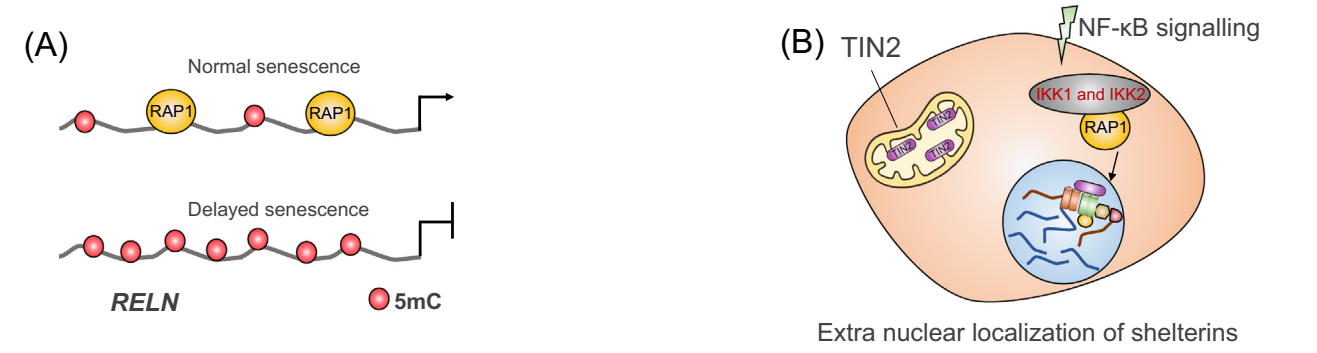

Figure 2. Telomere-associated factors (TAFs) in non-telomeric roles. (A) RAP1 positively regulates *RELN* transcription. (B) Extranuclear localization and function of TAFS: Mitochondrial localization of TIN2 causes negative regulation of oxidative phosphorylation. Cytoplasmic RAP1 regulates nuclear factor kappa B (NF-κB) signaling. TRF2 represses *CDKN1A* (p21) and *hTERT* and activates *PCGF3*, *PDGFRβ*, and *IL1R1*.

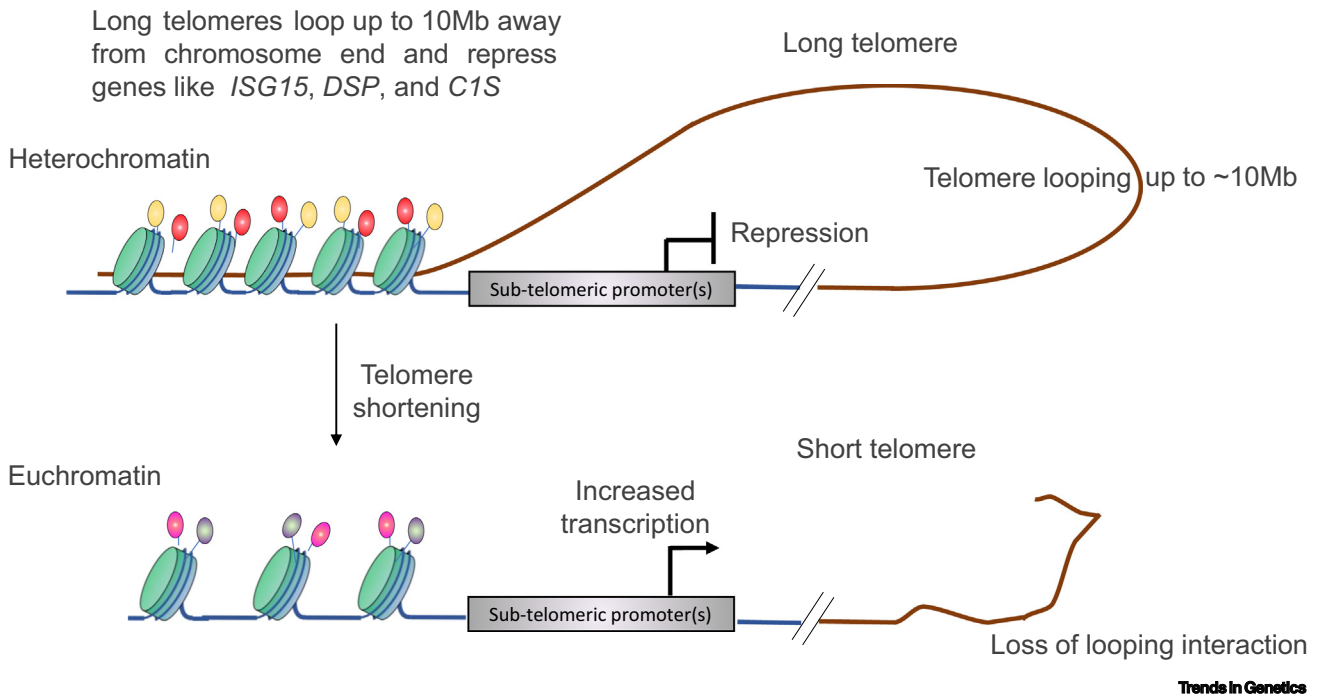

Figure 3. Telomere position effect over long distance (TPE/TPE-OLD). Long telomeres loop to subtelomeric regions. This physical association transcriptionally represses genes present in subtelomeres. Looping is lost when telomeres shorten and genes are activated.

looping of telomeres leading to activation of the genes [60]. The transcription regulation *hTERT* (located ~1.3 Mb from telomeres) was described by TPE-OLD where the chromosome 5p telomere interacts with the *hTERT* loci through chromosomal looping in human cancer and immortal cell lines [61]. Notably, TRF2 knockdown disrupted the looping interaction, suggesting the involvement of telomere-bound TRF2 in the interaction of telomeres with *hTERT* [61].

Telomere-dependent expression of genes remote from telomeres: the telomere-length-dependent partitioning (TeLPR) model

A mode of telomere-dependent gene expression distinct from TPE and TPE-OLD was recently reported. The expression of genes remote from telomeres (>60 Mb), and spread throughout the genome, altered depending on the length of telomeres [62]. Mechanistically, this was due to increased/decreased non-telomeric TRF2 binding within promoters in cells with relatively short/long telomeres, respectively. Altered TRF2 binding resulted in direct transcriptional regulation of the target gene [62]. Telomere elongation increased TRF2 occupancy at the telomeres; concomitantly, TRF2 occupancy at the non-telomeric sites was reduced. This was in cells with otherwise isogenic backgrounds where TRF2 levels remained relatively unaltered. Moreover, it was shown that promoter TRF2 occupancy induced histone modifications for either permissive (H3K4me3 and H3K4me1) or restrictive (H3K27me3) chromatin marks [62]; for instance, regulation of *p21*, the cyclin-dependent kinase [also known as cyclin-dependent kinase inhibitor 1 (CDK-1)] or the **interleukin-1 receptor-1** *IL1R1* described in the following section. Together these suggest a model where TeLPR results in differential TRF2 occupancy between telomeric versus non-telomeric sites (Figure 4). TeLPR interestingly implicates not only telomere-dependent gene expression but also other functions through TRF2, and/or other TAFs, found outside telomeres.

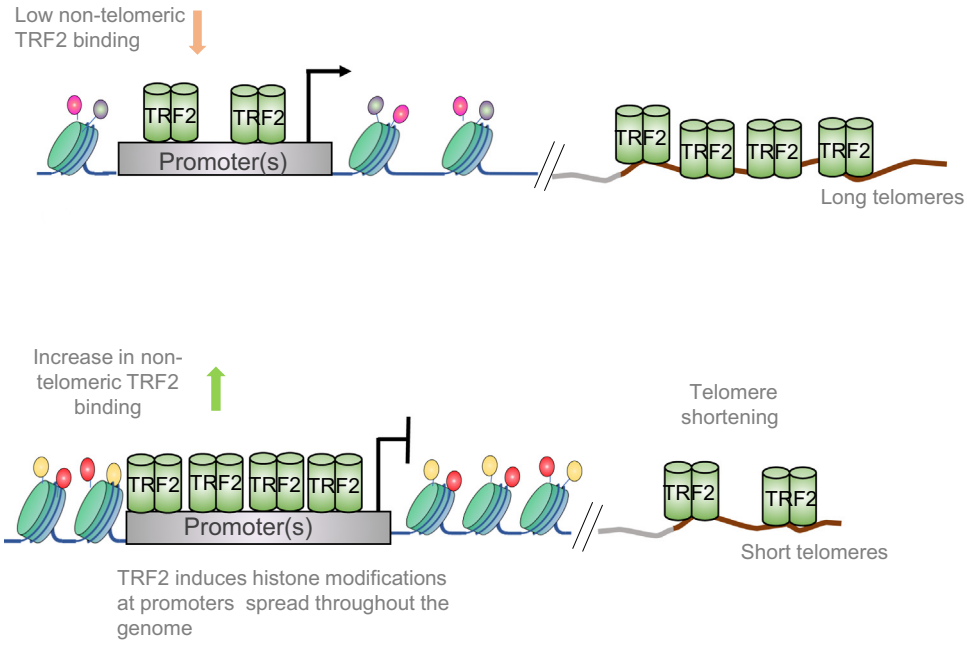

Figure 4. Telomere-length-dependent partitioning (TeLPR). Relatively long telomeres sequester TRF2 leading to less occupancy at non-telomeric sites. Shorter telomeres, by contrast, result in increased TRF2 binding to promoters. Increase/decrease in promoter-bound TRF2 affects transcription through chromatin modifications.

Independently, the association of non-telomeric RAP1 with IKK was shown to result in the activation of NF-κB [19]. It was further posited that, in cells with short telomeres, the relatively enhanced availability of non-telomeric RAP1 would give enhanced NF-κB signalling [63]. This is consistent with the TeLPR model broadening its scope to include other TAFs.

Notably, non-telomeric TRF2 sites included telomere-like guanine-rich-repeat sequences that could potentially form G4s [20], implicating interstitial G4s in the recruitment of non-telomeric TRF2. Since this would significantly contribute to the TeLPR model, later we discuss studies on telomeric and non-telomeric G4s, particularly, with regard to their role in findings implicating telomeres in functions that are outside telomeres.

G-quadruplex in telomeres

Initial analysis of synthetic guanine-rich sequences derived from the telomeres of *Tetrahymena thermophila* showed that these telomeric repeat sequences can form intramolecular G4s through non-Watson–Crick G·G base pairing [8] (see Box 1 for a detailed explanation of G4 structure). Human telomere sequences also adopted G4 structure *in vitro* [64,65]. Later, further studies showed that G4 formation is a general property for telomeric sequences from 15 different species [66]. Eukaryotic telomeres contain 3′ single-strand DNA overhangs. Approximately 150–300 nucleotides long in human telomeres, the overhang may fold back onto telomeres by strand invasion forming a lariat-like structure known as the T-loop [2]. Long single-stranded DNA comprising 5′-TTAGGG-3′ repeats has been used to mimic telomeric overhang *in vitro*, which condensed into a beaded filament-like structure stabilized by **G4-binding ligands**, suggesting G4 formation in the telomeric overhang *in vivo* [67]. Furthermore, TERRA (5′-UUAGGG-3′ repeats) has also been demonstrated to form multiple juxtaposed G4s with three-nucleotide UUA linkers [68]. TERRA hybridizes with the C-rich strand of telomeric DNA forming R-loops (Figure 1). In the ciliated protozoan *Stylonychia lemnae*, the presence of G4 *in vivo* was revealed using an antibody (Sty49) that was specific to G4s. Consistent with these, the G4-specific antibodies BG4 and 1H6 localized to the telomeres in human cells, supporting the formation of telomeric G4s in cells [69,70]. The presence of G4s in telomeres, their importance in maintaining genome stability, and targeting for cancer therapeutics have been reviewed [71,72]. TRF2 was reported to bind to telomeric G4s and mediate the interaction of TERRA with telomeres [73,74]. In ciliates, telomere-end-binding proteins regulated the formation of G4 structures *in vivo* [75]. *In vitro* observations showed unfolding of the telomeric G4 by multiple TAFs like POT1, TPP1, and the replication protein A [76,77]. In addition, heterogeneous nuclear ribonucleoprotein A1 (hnRNPA1) and unwinding protein-1 (UP1) were also noted to unwind telomeric G4s [78]. Together, these studies support the presence of telomeric G4s and their interaction with intracellular factors.

G-quadruplexes and telomerase

The discovery of the telomerase enzyme raised hopes in oncology and regenerative medicine. Telomerase is not only a tumor marker, but also a potential target, since its activity is essential for the long-term proliferation of cancer cells; most transformed cells *in vitro* can bypass progressive telomere loss and **replicative senescence** by activating telomerase. Hahn *et al.* showed that inactivation of telomerase in cancer cells was sufficient to limit the proliferation and growth of tumor cells [79]. Telomerase activity was therefore initially considered a promising target for the design of anticancer compounds.

G4 formation at the G-rich 3′ overhang of human telomeres (comprising multiple repeats of the GGGTTA hexanucleotide) generally inhibits telomerase extension *in vitro*, and this inhibition is reinforced if this G4 is stabilized by interaction with specific ligands. The G4 structure allows

Box 1. G-quadruplex: intracellular detection

G-quadruplexes (or G4s) are bookshelf-like structures formed by stacked planar guanine tetrads. A single-stranded DNA or RNA sequence comprising G-repeats can fold into the G4 through the formation of stable Hoogsteen-bonded tetrads (Figure I) [67]. As it involves single-stranded DNA, G4 formation is likely during replication or transcription when the double helix is in a relatively unwound form and/or within the single-strand overhang at the end of eukaryotic telomeres [130]. In addition, the C-rich complementary strand is known to be bound and stabilized by single-strand-binding proteins [131,132]. Based on the orientation of the folding, G4s can be parallel, antiparallel, or mixed (parallel–antiparallel), and depending on the number of strands involved G4s can be intra- or intermolecular. Detection of G4s inside cells started with computationally identified DNA sequences that contributed to PG4s in the genome [92,93]. G4-specific antibodies like sty49 and sty3 were found to bind to telomeric G4 in *Stylonychia lemnae* macronuclei [133,134]. Later, another G4 antibody, BG4, was reported to bind G4s inside human cells in regions outside telomeres supporting computational predictions [69]. In addition, intracellular fluorescent and radioactive G4-binding ligands [87] were used to detect G4s (extensively reviewed in [135]).

Intracellular G4s are likely to be predominantly protein bound [6,136]. Antibody-based intracellular detection therefore remains challenging as the antibody might be excluded by proteins that are already G4 bound. In addition, it is possible that the high affinity of the antibody promotes the formation of G4s that are otherwise not present. Although BG4-based ChIP-seq has been widely reported, BG4 was used outside cells (cell-free buffer) on sheared chromatin [137]. Therefore, the possibility of G4 formation or deformation during the isolation and purification of DNA, leading to the identification of G4s that are not intracellular, cannot be completely ruled out. With these in mind, an alternative approach used an antibody towards a protein that binds to G4s inside cells: TRF2 ChIP-seq in human cells detected thousands of TRF2-bound G4s throughout the genome [20]. Similar experiments with other G4-binding factors might be useful in revealing a larger repertoire of intracellular G4s.

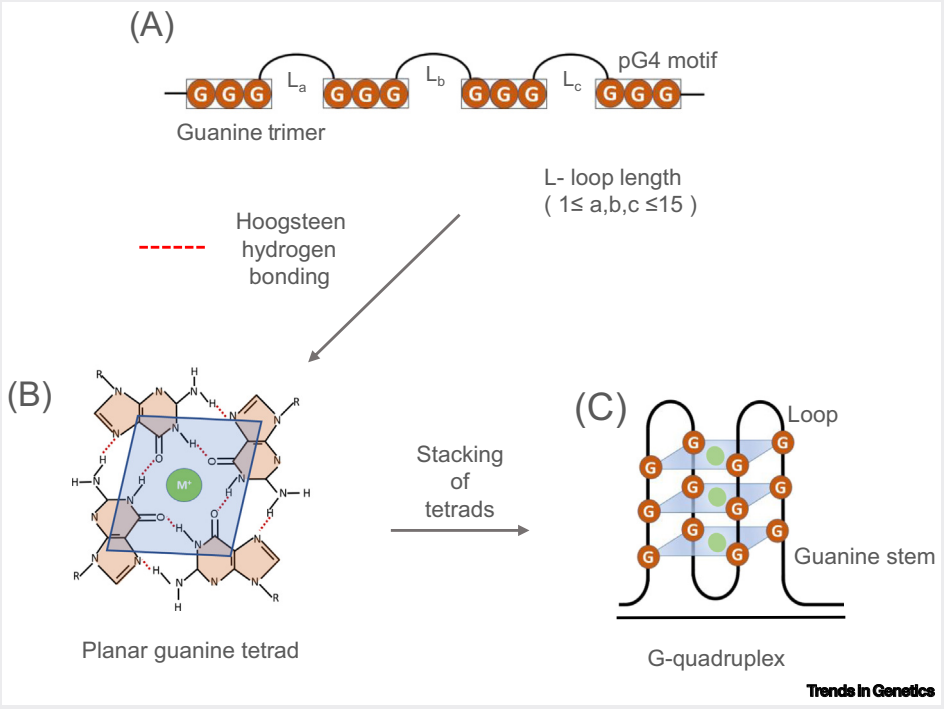

Figure I. Structure of G-quadruplex (G4). (A) Sequence pattern of potential G4 sequence with varying loop length. (B) Planar guanine tetrad formed by Hoogsteen hydrogen bonding. (C) Basic topology of the G4 structure.

the design of small molecules that recognize them selectively, which leads to indirect inhibition of telomerase [80]. This observation stimulated the quest for (telomeric) G4 ligands and a number of natural or synthetic compounds have been shown to be active in this way (for a review see [81]). There are now over 3000 different compounds known to interact with G4s [82]. Many, but not all,

of these ligands share structural characteristics such as a large, flat aromatic surface, allowing stacking on a terminal quartet, and positive charges favoring electrostatic interactions. The introduction of many positive charges may result in a high-affinity ligand, but with limited specificity, as found for core-extended naphthalene diimide (cex-NDI) derivatives, in which diethylene glycol (DEG) side chains were replaced by positively charged substituents [83]. There is therefore a trade-off to find between affinity and selectivity.

Unfortunately, although there are more than 100 clinical trials mentioning telomerase on ClinicalTrials.gov, there is still no validated cancer treatment that directly targets telomerase activity [84]. As feared from the outset, the inhibition of telomere extension by telomerase alone is probably not sufficient to lead to a significant antitumor effect; its association with other treatments is undoubtedly necessary. In addition, the antitelomerase effect of G4s is not as straightforward as initially expected.

First, quadruplexes may help to recruit telomerase rather than inhibit it [85] and only a few molecules interfere with the processivity of telomerase. G4 ligands were noted to be more effective inhibitors of telomeric DNA amplification than extension by telomerase and therefore may not be considered simple telomerase inhibitors. The **telomeric repeat amplification protocol (TRAP)** assay is inappropriate for the determination of telomerase inhibition by G4 ligands and the inhibitory effect of many G4 ligands has often been overestimated [86].

Second, many G4 sites are present in other regions of the genome, and the antiproliferative effect of some of these compounds could be explained by their action on other sites, such as the promoter of oncogenes. Compounds that exclusively target telomeric G4s are yet to be discovered. For example, a radiolabeled bisquinolinium compound was shown to bind to the terminal parts of chromosomes in metaphase spreads, but binding sites in internal chromosomal regions were also observed [87]. The polymorphic nature of telomeric G4s poses further challenges for the design of telomeric-G4-specific ligands [88]. Molecules that target the *hTERT* promoter specifically, by contrast, might be interesting, as *hTERT* repression would affect telomere maintenance. Further, perhaps, strategies that target molecules/factors in a locus-specific way (e.g., as in CRISPR-guided function) could be meaningful.

Finally, G4 ligands are also active in ALT cells, in which telomerase activity is absent; for example, the sensitivity of various osteosarcoma cell lines to a G4 ligand is the same whether these cells are ALT positive or telomerase positive, supporting the idea that G4 ligands are not simple telomerase inhibitors [89].

Non-telomeric G-quadruplexes and G-quadruplex-mediated gene regulation

Outside telomeres, guanine-repeat sequences that formed **potential G4s (PG4s)** were noted in immunoglobulin switch regions [90], in the promoter of the oncogene *c-myc* [91], and later throughout the human genome [9]. A key finding in 2006 showed that PG4s were predominant in the promoters of *Escherichia coli* genes; importantly, further analysis revealed PG4s to be evolutionarily conserved in thousands of microbes in orthologously related genes [10,92,93]. Based on these, authors proposed global G4-mediated gene regulation [10]. Furthermore, PG4 sequences were found to be conserved in rat, mouse, chimpanzee, and human promoters, underscoring the significance of G4s in gene regulation [92], which was supported by genome-wide alteration in gene expression in the presence of intracellular G4-binding ligands [94].

Notable in this regard was the description of the intracellular G4-binding protein non-metastatic factor NME2. Interaction of the *c-myc* promoter G4 with NME2 was observed to be critical for

the regulation of *c-myc*. Destabilization of the G4 gave reduced NME2 binding at the promoter resulting in altered transcription of *c-myc* in human cancer cell lines [95]. In addition to NME2, *c-myc* was reported to be regulated by nucleolin, cellular nucleic-acid-binding protein (CNBP), and poly (ADP-ribose)-polymerase-1 (PARP1) through G4 interactions [95–98]. Several other reports support G4-mediated gene regulation: NME2–G4 association in the regulation of *hTERT* [99]; interaction of PARP-1 with promoter G4s in the regulation of *c-Kit*, *KRAS*, and PARP-1 itself [100–102]; myc-associated zinc finger (MAZ) binding to the *KRAS* promoter G4 [103]; and the zinc-finger protein SP1 binding to the *c-kit* promoter [104], including computational analysis that revealed enrichment of SP1 binding sites at many promoter PG4s [105,106]. Interestingly, a recent analysis curated 77 G4-binding proteins and reported the glycine/arginine-rich **RGG motifs** to be important for interaction with G4-forming sequences [107,108]. Besides these, RecQ family helicases BLM, WRN, FANCJ, G4-resolvase-1, and Pif1 helicases were shown to unwind G4s and implicated in the facilitation of replication [109–113]. Together these findings support the genomic presence of G4s and underline the function of G4s in gene regulation in association with regulatory factors. Additionally, PG4s in 5′ untranslated regions (UTRs) in mRNA were shown to regulate alternative splicing, suggesting a role in translation control [114,115].

TRF2 binding to telomeric as well as non-telomeric sites genome wide

TRF2 comprises two DNA-binding domains: a C-terminal **Myb domain** and a RGG glycine/arginine-rich basic N-terminal domain. DNA binding by TRF2 was noted at sites/motifs that were not limited to the canonical telomeric TTAGGG repeats [116]. TRF2 ChIP-seq using human cancer cells found >20 000 TRF2 sites spread throughout the genome: A majority (38%) of the non-telomeric TRF2 sites were found to harbor G-repeats [20]. Similarly, earlier TRF2 ChIP-seq [14,15] – which recovered ~180 non-telomeric TRF2 sites due to lower read depth and/or antibody efficiency [20] – also found most binding motifs to be G-repeats. Together these show that the majority of TRF2 binding sites include G-repeats, a common feature of telomeric and non-telomeric TRF2 binding. Intriguingly, however, besides the guanine bias, there was no similarity in the compositions of the non-telomeric sequence motifs [20], suggesting that TRF2 binding might be dependent on the G4 DNA structure comprising G-repeats instead of a conserved sequence. Consistent with this, an affinity of recombinant purified TRF2 for both telomeric and non-telomeric G4s has been reported [20,24,73].

The role of TRF2 in T-loop formation at telomeres and telomere protection has been widely studied [2,117]. Of note, recent work, however, showed that T-loops can form independent of TRF2, and TRF2-mediated telomere protection is not essential in mESCs [118,119], indicating the possibility of TRF2 telomeric functions that are dependent on cell type or the developmental stage of cells.

Transcription regulation by non-telomeric TRF2 is G-quadruplex dependent

TRF2 was reported to bind to G4s in the proximal promoter of the human telomerase gene (*hTERT*) [24]. The TRF2–G4 interaction was necessary for the recruitment of the RE-1-silencing-factor REST and the canonical repressor complex **PRC2**. This induced histone H3K27 trimethylation at the *hTERT* promoter leading to epigenetic repression of *hTERT* in both cancer and normal human cells in a TRF2–G4-dependent way [24]. Furthermore, a large number of human cancers [120], particularly glioblastoma and melanoma [121,122], with two somatic guanine mutations in the *hTERT* promoter (–124G>A and –146G>A, 124 and 146 nucleotides upstream of the transcription start site of *hTERT*), were found to be clinically prevalent and causally implicated in *hTERT* reactivation. It was shown that the two mutations in G-repeats resulted in disruption of promoter G4s compromising TRF2 binding, and resulting in *hTERT* reactivation. Stabilization of the promoter G4 using G4-binding ligand or by reversing the mutation from A to

G using CRISPR in cells reinstated TRF2 binding on the *hTERT* promoter. This, in turn resulted in re-suppression of *hTERT* in glioblastoma cell lines including patient-derived primary glioblastoma cells that were treated with intracellular G4-binding ligands [24,123].

In addition, GA-binding protein α/β (GABPA) was shown to be specifically recruited to the mutant *hTERT* promoter and mediate long-range chromatin interaction between the promoter and a region 300 kb upstream (T-INT) of the promoter [124]. This enhanced *hTERT* transcription through enriched active histone marks and POL2 recruitment to the promoter. Therefore, loss of TRF2 binding, and increased GABPA recruitment, possibly sequentially, at the mutant *hTERT* promoter might result in *hTERT* reactivation. In cancer cells devoid of the *hTERT* promoter mutation, Jun-D-mediated CTCF recruitment regulated the long-range *hTERT* interaction with T-INT [125].

TRF2 binding genome wide is telomere dependent

TRF2 binding at the promoter and suppression of *p21*, crucial for the linking of cycle arrest to DNA damage, was reported to be G4 dependent [23]. Furthermore, interestingly, the regulation of *p21* through TRF2–G4 interaction was dependent on telomere length [62]. In cells with relatively long telomeres, TRF2 binding on the *p21* promoter was lower than in cells with short telomeres, with otherwise isogenous backgrounds. This resulted in relatively high expression of *p21* in cells with long telomeres. Mechanistically, TRF2-induced recruitment of the epigenetic repressor complex REST/LSD1 was attenuated in cells with longer telomeres leading to reduced H3K27me3. Together these contributed to restrictive/permissive chromatin at the *p21* promoter in a telomere-dependent fashion.

Recent work further shows the regulation of *IL1R1* expression to be telomere dependent through TRF2 binding at the *IL1R1* promoter [25]. Interaction of TRF2 with the promoter G4 was necessary for recruitment of the histone acetyl-transferase p300 leading to increased acetylation at histone K27 and the induction of *IL1R1*. In the case of cells with relatively short telomeres, TRF2 binding at the *IL1R1* promoter was greater than with long telomeres, consistent with the TeLPR model. Consequently, *IL1R1* expression was enhanced in cancer cells, with short *vis-à-vis* long telomeres resulting, interestingly, in telomere-dependent activation of interleukin-1 signaling [25].

It is of interest to note that at the *hTERT* or the *p21* promoters, TRF2 promotes repressive chromatin through interaction with REST and canonical repressor complexes [23,24]. By contrast, at the *IL1R1* promoter TRF2–p300 interaction induces active chromatin [25]. This is likely to be due to distinct PTMs of TRF2 at the respective promoters: we noted acetylation at the 293-lysine residue in the case of the *IL1R1*-promoter-bound TRF2 [25]. In the case of the *hTERT* promoter, by contrast, methylation at the TRF2 arginine-17 residue was evident (Sengupta *et al.*, unpublished). An earlier paper described non-telomeric TRF2 to be predominantly methylated *vis-à-vis* telomeric TRF2 [126]. Overall, these findings support the notion that interstitial binding of TAFs would require extensive interactions with histone modifier/other regulatory cofactors. While further work will be necessary for TRF2, and other TAFs like RAP1/TRF1, it is possible that PTMs distinguish telomeric and non-telomeric binding/function.

Importantly, these emerging findings suggest a broader interplay where TRF2 binding to telomeric and non-telomeric G4s is in line mechanistically with the TeLPR model. Furthermore, other TAFs that associate with interstitial regions might function in a similar mode.

Concluding remarks and future perspectives

In conclusion, emerging evidence increasingly shows that telomeres or TAFs can influence gene expression and other functions throughout the genome. For instance, TRF2 binds to promoters

Outstanding questions

What are the consequences of the TeLPR model – which shows how non-telomeric function is telomere dependent – in cellular events like differentiation and neoplastic progression that are closely associated with change in telomere length? Could insights from TeLPR aid in better understanding of aging-related disorders?

Do the telomeric shelterin proteins undergo PTMs for non-telomeric binding?

Can shelterin complex proteins undergo different PTMs across cell types? More importantly, do they vary in stem cells *vis-à-vis* the differentiated state?

Can DNA topology, the G-quadruplex structure in particular, be a key element in the regulation of cellular events associated with altered telomere length?

throughout the genome (e.g., genes ~60 Mb away from the nearest telomere) and regulation of the target genes was telomere dependent [62]. Telomere-dependent binding of non-telomeric TRF2 (the TeLPR model described earlier) posits signaling from the wings to the entire chromosome [127]. G4s in telomeres and the enriched presence of G4s in gene regulatory regions genome wide is notable contextually. The binding of TAFs like TRF2 to G4s, both telomeric and non-telomeric, therefore supports intrinsic genetic links between telomeres and the non-telomeric genome through G4s. Together with further work, this conceptual advance might strengthen the understanding that incremental telomeric shortening impacts biology across the genome in a continuous and gradual fashion. This therefore modifies the prevailing broad understanding that cellular triggers – for instance, replicative senescence – are activated when telomeres reach a critically short limit.

Further, as novel non-telomeric functions of TAFs are discovered, an understanding of the mechanistic details of how (and to what extent) telomeres influence non-telomeric biology would be of interest (see Outstanding questions). Here the newly released Telomere-to-Telomere (T2T) genome sequence is expected to be of significance [128]. Re-analysis of ChIP-seq with the current genomic data may reveal novel binding sites, especially comprising repetitive DNA, which was relatively poorly covered until recently and may reveal novel G4s [129]. Together these will further deepen our understanding of how telomeres impact non-telomeric functions.

Acknowledgments

This work was supported by research grants from the Wellcome Trust/DBT India Alliance Fellowship (IA/S/18/2/504021) to S.C. and by Agence Nationale de la Recherche (G4 Access) and the SYMBIT project reg. no. CZ.02.1.01/0.0/0.0/15_003/0000477 financed from the ERDF to J-L.M. We acknowledge the Council of Scientific and Industrial Research and the Academy of Scientific and Innovative Research for research and other infrastructure. S.B. and S.V. acknowledge the CSIR for research fellowships. J-L.M. thanks Lionel Guittat (LOB, Palaiseau) for interesting discussions.

Declaration of interests

No interests are declared.

References

1. O'Sullivan, R.J. and Karlseder, J. (2010) Telomeres: protecting chromosomes against genome instability. *Nat. Rev. Mol. Cell Biol.* 11, 171
2. Griffith, J.D. *et al.* (1999) Mammalian telomeres end in a large duplex loop. *Cell* 97, 503–514
3. Sundquist, W.I. and Klug, A. (1989) Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops. *Nature* 342, 825–829
4. Lipps, H.J. and Rhodes, D. (2009) G-quadruplex structures: *in vivo* evidence and function. *Trends Cell Biol.* 19, 414–422
5. Bochman, M.L. *et al.* (2012) DNA secondary structures: stability and function of G-quadruplex structures. *Nat. Rev. Genet.* 13, 770–780
6. Brázda, V. *et al.* (2014) DNA and RNA quadruplex-binding proteins. *Int. J. Mol. Sci.* 15, 17493
7. Ye, J. *et al.* (2014) Transcriptional outcome of telomere signalling. *Nat. Rev. Genet.* 15, 491–503
8. Henderson, E. *et al.* (1987) Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine-guanine base pairs. *Cell* 51, 899–908
9. Huppert, J.L. and Balasubramanian, S. (2005) Prevalence of quadruplexes in the human genome. *Nucleic Acids Res.* 33, 2908–2916
10. Rawal, P. *et al.* (2006) Genome-wide prediction of G4 DNA as regulatory motifs: role in *Escherichia coli* global regulation. *Genome Res.* 16, 644–655
11. Verma, A. *et al.* (2008) Genome-wide computational and expression analyses reveal G-quadruplex DNA motifs as conserved cis-regulatory elements in human and related species. *J. Med. Chem.* 51, 5641–5649
12. Mukherjee, A.K. *et al.* (2019) Non-duplex G-quadruplex structures emerge as mediators of epigenetic modifications. *Trends Genet.* 35, 129
13. Martinez, P. *et al.* (2010) Mammalian Rap1 controls telomere function and gene expression through binding to telomeric and extratelomeric sites. *Nat. Cell Biol.* 12, 768–780
14. Simonet, T. *et al.* (2011) The human TTAGGG repeat factors 1 and 2 bind to a subset of interstitial telomeric sequences and satellite repeats. *Cell Res.* 21, 1028–1038
15. Yang, D. *et al.* (2011) Human telomeric proteins occupy selective interstitial sites. *Cell Res.* 21, 1013–1027
16. Zhang, X. *et al.* (2019) Telomere-dependent and telomere-independent roles of RAP1 in regulating human stem cell homeostasis. *Protein Cell* 10, 649–667
17. Yeung, F. *et al.* (2013) Nontelomeric role for Rap1 in regulating metabolism and protecting against obesity. *Cell Rep.* 3, 1847
18. Martínez, P. *et al.* (2013) RAP1 protects from obesity through its extratelomeric role regulating gene expression. *Cell Rep.* 3, 2059–2074
19. Teo, H. *et al.* (2010) Telomere-independent Rap1 is an IKK adaptor and regulates NF-κB-dependent gene expression. *Nat. Cell Biol.* 12, 758–767
20. Mukherjee, A.K. *et al.* (2019) Telomere repeat-binding factor 2 binds extensively to extra-telomeric G-quadruplexes and regulates the epigenetic status of several gene promoters. *J. Biol. Chem.* 294, 17709–17722
21. El Maï, M. *et al.* (2014) The telomeric protein TRF2 regulates angiogenesis by binding and activating the PDGFRβ promoter. *Cell Rep.* 9, 1047–1060

22. Purohit, G. *et al.* (2018) Extratelomeric binding of the telomere binding protein TRF2 at the PCGF3 promoter is g-quadruplex motif-dependent. *Biochemistry* 57, 2317–2324
23. Hussain, T. *et al.* (2017) Transcription regulation of CDKN1A (p21/CIP1/WAF1) by TRF2 is epigenetically controlled through the REST repressor complex. *Sci. Rep.* 7, 11541
24. Sharma, S. *et al.* (2021) Human telomerase is directly regulated by non-telomeric TRF2-G-quadruplex interaction. *Cell Rep.* 35, 109154
25. Kishore Mukherjee, A. *et al.* (2021) Telomere-dependent interleukin-1 receptor activation promotes immune suppression in triple-negative-breast cancer. *bioRxiv* Published online December 8, 2021. https://doi.org/10.1101/2021.12.07.471419
26. Zizza, P. *et al.* (2019) TRF2 positively regulates SULF2 expression increasing VEGF-A release and activity in tumor microenvironment. *Nucleic Acids Res.* 47, 3365–3382
27. Biroccio, A. *et al.* (2013) TRF2 inhibits a cell-extrinsic pathway through which natural killer cells eliminate cancer cells. *Nat. Cell Biol.* 15, 818–828
28. Cherfils-Vicini, J. *et al.* (2019) Cancer cells induce immune escape via glycocalyx changes controlled by the telomeric protein TRF2. *EMBO J.* 38, e100012
29. Kaur, P. *et al.* (2016) Enhanced electrostatic force microscopy reveals higher-order DNA looping mediated by the telomeric protein TRF2. *Sci. Rep.* 6, 20513
30. Wood, A.M. *et al.* (2014) TRF2 and lamin A/C interact to facilitate the functional organization of chromosome ends. *Nat. Commun.* 5, 5467
31. Smith, E.D. *et al.* (2018) Interstitial telomeric loops and implications of the interaction between TRF2 and lamin A/C. *Differentiation* 102, 19–26
32. Saha, A. *et al.* (2018) Role of telomeric TRF2 in orosphere formation and CSC phenotype maintenance through efficient DNA repair pathway and its correlation with recurrence in OSCC. *Stem Cell Rev. Rep.* 14, 871–887
33. Zhang, P. *et al.* (2011) Nontelomeric splice variant of telomere repeat-binding factor 2 maintains neuronal traits by sequestering repressor element 1-silencing transcription factor. *Proc. Natl. Acad. Sci. U. S. A.* 108, 16434–16439
34. Ovando-Roche, P. *et al.* (2014) TRF2-mediated stabilization of hREST4 is critical for the differentiation and maintenance of neural progenitors. *Stem Cells* 32, 2111–2122
35. Schneider, R.P. *et al.* (2013) TRF1 is a stem cell marker and is essential for the generation of induced pluripotent stem cells. *Nat. Commun.* 4, 1946
36. Marión, R.M. *et al.* (2017) Common telomere changes during *in vivo* reprogramming and early stages of tumorigenesis. *Stem Cell Rep.* 8, 460
37. Vinayagamurthy, S. *et al.* (2020) Extra-telomeric impact of telomeres: emerging molecular connections in pluripotency or stemness. *J. Biol. Chem.* 295, 10245–10254
38. Ohishi, T. *et al.* (2014) TRF1 ensures the centromeric function of Aurora-B and proper chromosome segregation. *Mol. Cell. Biol.* 34, 2464
39. Lee, J. and Gollahon, L. (2013) Mitotic perturbations induced by Nek2 overexpression require interaction with TRF1 in breast cancer cells. *Cell Cycle* 12, 3599
40. Yi, X. *et al.* (2001) Quantitation of telomerase components and hTERT mRNA splicing patterns in immortal human cells. *Nucleic Acids Res.* 29, 4818–4825
41. Akincilar, S.C. *et al.* (2015) Quantitative assessment of telomerase components in cancer cell lines. *FEBS Lett.* 589, 974–984
42. Park, J.-I. *et al.* (2009) Telomerase modulates Wnt signalling by association with target gene chromatin. *Nature* 460, 66–72
43. Ghosh, A. *et al.* (2012) Telomerase directly regulates NF-κB-dependent transcription. *Nat. Cell Biol.* 14, 1270–1281
44. Khattar, E. *et al.* (2016) Telomerase reverse transcriptase promotes cancer cell proliferation by augmenting tRNA expression. *J. Clin. Invest.* 126, 4045–4060
45. Ozturk, M.B. *et al.* (2017) Current insights to regulation and role of telomerase in human diseases. *Antioxidants (Basel)* 6, 17
46. Chu, H.P. *et al.* (2017) TERRA RNA antagonizes ATRX and protects telomeres. *Cell* 170, 86–101.e16
47. Hirashima, K. and Seimiya, H. (2015) Telomeric repeat-containing RNA/G-quadruplex-forming sequences cause genome-wide alteration of gene expression in human cancer cells *in vivo*. *Nucleic Acids Res.* 43, 2022–2032
48. Marión, R.M. *et al.* (2019) TERRA regulate the transcriptional landscape of pluripotent cells through TRF1-dependent recruitment of PRC2. *Elife* 8, e44656
49. Deng, Z. *et al.* (2012) Formation of telomeric repeat-containing RNA (TERRA) foci in highly proliferating mouse cerebellar neuronal progenitors and medulloblastoma. *J. Cell Sci.* 125, 4383–4394
50. Akincilar, S.C. *et al.* (2021) Non-canonical roles of canonical telomere binding proteins in cancers. *Cell. Mol. Life Sci.* 78, 4235–4257
51. Chen, L.-Y. *et al.* (2012) Mitochondrial localization of telomeric protein TIN2 links telomere regulation to metabolic control. *Mol. Cell* 47, 839–850
52. Lee, J.H. *et al.* (2018) Loss of RNA-binding protein HuR facilitates cellular senescence through posttranscriptional regulation of TIN2 mRNA. *Nucleic Acids Res.* 46, 4271
53. Platt, J.M. *et al.* (2013) Rap1 relocalization contributes to the chromatin-mediated gene expression profile and pace of cell senescence. *Genes Dev.* 27, 1406–1420
54. Song, S. *et al.* (2020) Rap1-mediated nucleosome displacement can regulate gene expression in senescent cells without impacting the pace of senescence. *Aging Cell* 19, e13061
55. Wagner, K.D. *et al.* (2017) The differential spatiotemporal expression pattern of shelterin genes throughout lifespan. *Aging (Albany NY)* 9, 1219–1232
56. Ying, Y. *et al.* (2022) The non-telomeric evolutionary trajectory of TRF2 in zebrafish reveals its specific roles in neurodevelopment and aging. *Nucleic Acids Res.* 50, 2081–2095
57. Gottschling, D.E. *et al.* (1990) Position effect a *S. cerevisiae* telomeres: reversible repression of Pol II transcription. *Cell* 63, 751–762
58. Ottaviani, A. *et al.* (2008) Telomeric position effect: from the yeast paradigm to human pathologies? *Biochimie* 90, 93–107
59. Baur, J.A. *et al.* (2001) Telomere position effect in human cells. *Science* 292, 2075–2077
60. Robin, J.D. *et al.* (2014) Telomere position effect: regulation of gene expression with progressive telomere shortening over long distances. *Genes Dev.* 28, 2464–2476
61. Kim, W. *et al.* (2016) Regulation of the human telomerase gene TERT by telomere position effect-over long distances (TPE-OLD): implications for aging and cancer. *PLoS Biol.* 14, e2000016
62. Mukherjee, A.K. *et al.* (2018) Telomere length-dependent transcription and epigenetic modifications in promoters remote from telomere ends. *PLoS Genet.* 14, e1007782
63. Crabbe, L. and Karlseder, J. (2010) Mammalian Rap1 widens its impact. *Nat. Cell Biol.* 12, 733–735
64. Luu, K.N. *et al.* (2006) Structure of the human telomere in K+ solution: an intramolecular (3 + 1) G-quadruplex scaffold. *J. Am. Chem. Soc.* 128, 9963–9970
65. Phan, A.T. *et al.* (2007) Structure of two intramolecular G-quadruplexes formed by natural human telomere sequences in K^+ solution. *Nucleic Acids Res.* 35, 6517–6525
66. Tran, P.L.T. *et al.* (2011) Stability of telomeric G-quadruplexes. *Nucleic Acids Res.* 39, 3282–3294
67. Kar, A. *et al.* (2018) Long repeating $(TTAGGG)_n$ single-stranded DNA self-condenses into compact beaded filaments stabilized by G-quadruplex formation. *J. Biol. Chem.* 293, 9473–9485
68. Randall, A. and Griffith, J.D. (2009) Structure of long telomeric RNA transcripts: the G-rich RNA forms a compact repeating structure containing G-quartets. *J. Biol. Chem.* 284, 13980
69. Biffi, G. *et al.* (2013) Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.* 5, 182–186
70. Henderson, A. *et al.* (2014) Detection of G-quadruplex DNA in mammalian cells. *Nucleic Acids Res.* 42, 860–869
71. Bryan, T.M. (2020) G-quadruplexes at telomeres: friend or foe? *Molecules* 25, 3686
72. Seimiya, H. *et al.* (2021) Chemical targeting of G-quadruplexes in telomeres and beyond for molecular cancer therapeutics. *J. Antibiot.* 74, 617–628
73. Pedroso, I.M. *et al.* (2009) The effect of the TRF2 N-terminal and TRFH regions on telomeric G-quadruplex structures. *Nucleic Acids Res.* 37, 1541–1554

74. Biffi, G. *et al.* (2012) An intramolecular G-quadruplex structure is required for binding of telomeric repeat-containing RNA to the telomeric protein TRF2. *J. Am. Chem. Soc.* 134, 11974–11976
75. Paeschke, K. *et al.* (2005) Telomere end-binding proteins control the formation of G-quadruplex DNA structures *in vivo*. *Nat. Struct. Mol. Biol.* 12, 847–854
76. Hwang, H. *et al.* (2012) POT1-TPP1 regulates telomeric overhang structural dynamics. *Structure* 20, 1872–1880
77. Safa, L. *et al.* (2014) Binding polarity of RPA to telomeric sequences and influence of G-quadruplex stability. *Biochimie* 103, 80–88
78. Hudson, J.S. *et al.* (2014) Recognition and binding of human telomeric G-quadruplex DNA by unfolding protein 1. *Biochemistry* 53, 3347–3356
79. Hahn, W.C. *et al.* (1999) Inhibition of telomerase limits the growth of human cancer cells. *Nat. Med.* 5, 1164–1170
80. Sun, D. *et al.* (1997) Inhibition of human telomerase by a G-quadruplex-interactive compound. *J. Med. Chem.* 40, 2113–2116
81. Saraswati, A.P. *et al.* (2019) Raising the bar in anticancer therapy: recent advances in, and perspectives on, telomerase inhibitors. *Drug Discov. Today* 24, 1370–1388
82. Wang, Y.H. *et al.* (2022) G4LDB 2.2: a database for discovering and studying G-quadruplex and i-Motif ligands. *Nucleic Acids Res.* 50, D150–D160
83. Zuffo, M. *et al.* (2018) More is not always better: finding the right trade-off between affinity and selectivity of a G-quadruplex ligand. *Nucleic Acids Res.* 46, e115
84. Mergny, J.L. *et al.* (2021) Are telomeres and telomerase still relevant targets in oncology? *Bull. Cancer* 108, 30–38 (in French)
85. Moye, A.L. *et al.* (2015) Telomeric G-quadruplexes are a substrate and site of localization for human telomerase. *Nat. Commun.* 6, 7643
86. De Cian, A. *et al.* (2007) Reevaluation of telomerase inhibition by quadruplex ligands and their mechanisms of action. *Proc. Natl. Acad. Sci. U. S. A.* 104, 17347
87. Granotier, C. *et al.* (2005) Preferential binding of a G-quadruplex ligand to human chromosome ends. *Nucleic Acids Res.* 33, 4182–4190
88. Phan, A.T. (2010) Human telomeric G-quadruplex: structures of DNA and RNA sequences. *FEBS J.* 277, 1107–1117
89. Amato, R. *et al.* (2020) G-quadruplex stabilization fuels the ALT pathway in ALT-positive osteosarcoma cells. *Genes (Basel)* 11, 304
90. Shimizu, A. and Honjo, T. (1984) Immunoglobulin class switching. *Cell* 36, 801–803
91. Simonsson, T. *et al.* (1998) DNA tetraplex formation in the control region of c-myc. *Nucleic Acids Res.* 26, 1167–1172
92. Yadav, V.K. *et al.* (2008) QuadBase: genome-wide database of G4 DNA – occurrence and conservation in human, chimpanzee, mouse and rat promoters and 146 microbes. *Nucleic Acids Res.* 36, D381–D385
93. Dhapola, P. and Chowdhury, S. (2016) QuadBase2: web server for multiplexed guanine quadruplex mining and visualization. *Nucleic Acids Res.* 44, W277–W283
94. Verma, A. *et al.* (2009) Evidence of genome-wide G4 DNA-mediated gene expression in human cancer cells. *Nucleic Acids Res.* 37, 4194–4204
95. Thakur, R.K. *et al.* (2009) Metastases suppressor NM23-H2 interaction with G-quadruplex DNA within c-MYC promoter nuclease hypersensitive element induces c-MYC expression. *Nucleic Acids Res.* 37, 172–183
96. González, V. *et al.* (2009) Identification and characterization of nucleolin as a c-myc G-quadruplex-binding protein. *J. Biol. Chem.* 284, 23622–23635
97. Fekete, A. *et al.* (2012) The guanine-quadruplex structure in the human c-myc gene's promoter is converted into B-DNA form by the human poly(ADP-ribose)polymerase-1. *PLoS One* 7, e42690
98. Chen, S. *et al.* (2013) Mechanistic studies for the role of cellular nucleic-acid-binding protein (CNBP) in regulation of c-myc transcription. *Biochim. Biophys. Acta* 1830, 4769–4777
99. Saha, D. *et al.* (2017) Epigenetic suppression of human telomerase (hTERT) is mediated by the metastasis suppressor NME2 in a G-quadruplex-dependent fashion. *J. Biol. Chem.* 292, 15205–15215
100. Soldatenkov, V.A. *et al.* (2002) Transcriptional repression by binding of poly(ADP-ribose) polymerase to promoter sequences. *J. Biol. Chem.* 277, 665–670
101. Cogoi, S. *et al.* (2010) The KRAS promoter responds to Myc-associated zinc finger and poly(ADP-ribose) polymerase 1 proteins, which recognize a critical quadruplex-forming GA-element. *J. Biol. Chem.* 285, 22003–22016
102. Sengar, A. *et al.* (2019) Structure of a (3+1) hybrid G-quadruplex in the PARP1 promoter. *Nucleic Acids Res.* 47, 1564–1572
103. Cogoi, S. *et al.* (2013) MAZ-binding G4-decoy with locked nucleic acid and twisted intercalating nucleic acid modifications suppresses KRAS in pancreatic cancer cells and delays tumor growth in mice. *Nucleic Acids Res.* 41, 4049–4064
104. Raiber, E.-A. *et al.* (2012) A non-canonical DNA structure is a binding motif for the transcription factor SP1 *in vitro*. *Nucleic Acids Res.* 40, 1499–1508
105. Todd, A.K. and Neidle, S. (2008) The relationship of potential G-quadruplex sequences in *cis*-upstream regions of the human genome to SP1-binding elements. *Nucleic Acids Res.* 36, 2700–2704
106. Kumar, P. *et al.* (2011) Zinc-finger transcription factors are associated with guanine quadruplex motifs in human, chimpanzee, mouse and rat promoters genome-wide. *Nucleic Acids Res.* 39, 8005–8016
107. Brázda, V. *et al.* (2018) The amino acid composition of quadruplex binding proteins reveals a shared motif and predicts new potential quadruplex interactors. *Molecules* 23, 2341
108. Huang, Z.L. *et al.* (2018) Identification of G-quadruplex-binding protein from the exploration of RGG motif/G-quadruplex interactions. *J. Am. Chem. Soc.* 140, 17945–17955
109. Mohaghegh, P. *et al.* (2001) The Bloom's and Werner's syndrome proteins are DNA structure-specific helicases. *Nucleic Acids Res.* 29, 2843–2849
110. Wu, Y. *et al.* (2008) FANCJ helicase defective in Fanconia anemia and breast cancer unwinds G-quadruplex DNA to defend genomic stability. *Mol. Cell. Biol.* 28, 4116–4128
111. Creacy, S.D. *et al.* (2008) G4 resolvase 1 binds both DNA and RNA tetramolecular quadruplex with high affinity and is the major source of tetramolecular quadruplex G4-DNA and G4-RNA resolving activity in HeLa cell lysates. *J. Biol. Chem.* 283, 34626–34634
112. Huber, M.D. *et al.* (2002) G4 DNA unwinding by BLM and Sgs1p: substrate specificity and substrate-specific inhibition. *Nucleic Acids Res.* 30, 3954–3961
113. Byrd, A.K. *et al.* (2018) Pif1 helicase unfolding of G-quadruplex DNA is highly dependent on sequence and reaction conditions. *J. Biol. Chem.* 293, 17792–17802
114. Huppert, J.L. *et al.* (2008) G-quadruplexes: the beginning and end of UTRs. *Nucleic Acids Res.* 36, 6260–6268
115. Didiot, M.C. *et al.* (2008) The G-quartet containing FMRP binding site in FMR1 mRNA is a potent exonic splicing enhancer. *Nucleic Acids Res.* 36, 4902–4912
116. Fouché, N. *et al.* (2006) The basic domain of TRF2 directs binding to DNA junctions irrespective of the presence of TTAGGG repeats. *J. Biol. Chem.* 281, 37486–37495
117. Karlseder, J. *et al.* (1999) p53- and ATM-dependent apoptosis induced by telomeres lacking TRF2. *Science* 283, 1321–1325
118. Ruis, P. *et al.* (2021) TRF2-independent chromosome end protection during pluripotency. *Nature* 589, 103–109
119. Markiewicz-Potoczny, M. *et al.* (2021) TRF2-mediated telomere protection is dispensable in pluripotent stem cells. *Nature* 589, 110–115
120. Liu, T. *et al.* (2016) Cancer-specific telomerase reverse transcriptase (TERT) promoter mutations: biological and clinical implications. *Genes* 7, 38
121. Horn, S. *et al.* (2013) TERT promoter mutations in familial and sporadic melanoma. *Science* 339, 959–961
122. Killela, P.J. *et al.* (2013) TERT promoter mutations occur frequently in gliomas and a subset of tumors derived from cells with low rates of self-renewal. *Proc. Natl. Acad. Sci. U. S. A.* 110, 6021–6026

123. Sharma, S. and Chowdhury, S. (2022) Emerging mechanisms of telomerase reactivation in cancer. *Trends Cancer* 8, 632–641
124. Akıncılar, S.C. *et al.* (2016) Long-range chromatin interactions drive mutant TERT promoter activation. *Cancer Discov.* 6, 1276–1292
125. Can Akıncılar, S. *et al.* (2022) Identification of mechanism of cancer-cell-specific reactivation of hTERT offers therapeutic opportunities for blocking telomerase specifically in human colorectal cancer. *Nucleic Acids Res.* Published online June 14, 2022. https://doi.org/10.1093/NAR/GKAC479
126. Mitchell, T.R.H. and Zhu, X.D. (2014) Methylated TRF2 associates with the nuclear matrix and serves as a potential biomarker for cellular senescence. *Aging (Albany NY)* 6, 248
127. Okamoto, K. and Seimiya, H. (2019) From the wings to the center stage of chromosomes. *J. Biol. Chem.* 294, 17723–17724
128. Nurk, S. *et al.* (2022) The complete sequence of a human genome. *Science* 376, 44–53
129. Bohálová, N. *et al.* (2021) Novel G-quadruplex prone sequences emerge in the complete assembly of the human X chromosome. *Biochimie* 191, 87–90
130. Sarkies, P. *et al.* (2010) Epigenetic instability due to defective replication of structured DNA. *Mol. Cell* 40, 703–713
131. Lacroix, L. *et al.* (2000) Identification of two human nuclear proteins that recognise the cytosine-rich strand of human telomeres *in vitro*. *Nucleic Acids Res.* 28, 1564–1575
132. Kang, H.J. *et al.* (2014) The transcriptional complex between the BCL2 i-motif and hnRNP LL is a molecular switch for control of gene expression that can be modulated by small molecules. *J. Am. Chem. Soc.* 136, 4172–4185
133. Schaffitzel, C. *et al.* (2001) *In vitro* generated antibodies specific for telomeric guanine-quadruplex DNA react with *Stylonychia lemnae* macronuclei. *Proc. Natl. Acad. Sci. U. S. A.* 98, 8572–8577
134. Schaffitzel, C. *et al.* (2010) Probing telomeric G-quadruplex DNA structures in cells with in vitro generated single-chain antibody fragments. *Methods Mol. Biol.* 608, 159–181
135. Sengupta, A. *et al.* (2019) Promise of G-quadruplex structure binding ligands as epigenetic modifiers with anti-cancer effects. *Molecules* 24, 582
136. Mishra, S.K. *et al.* (2016) G4IPDB: a database for G-quadruplex structure forming nucleic acid interacting proteins. *Sci. Rep.* 6, 38144
137. Hänsel-Hertsch, R. *et al.* (2018) Genome-wide mapping of endogenous G-quadruplex DNA structures by chromatin immunoprecipitation and high-throughput sequencing. *Nat. Protoc.* 13, 551–564