Telomeric Chromatin and TERRA
Amandine Barral, Jérome Déjardin

To cite this version:
Amandine Barral, Jérome Déjardin. Telomeric Chromatin and TERRA. Journal of Molecular Biology, Elsevier, 2020, 432, pp.4244 - 4256. 10.1016/j.jmb.2020.03.003 . hal-03491403

HAL Id: hal-03491403
https://hal.archives-ouvertes.fr/hal-03491403
Submitted on 18 Jul 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License
Telomeric chromatin and TERRA

Amandine Barral and Jérome Déjardin

Institute of Human Genetics, UMR 9002 CNRS and Université de Montpellier
141, rue de la Cardonille, 34000 Montpellier, France.

Abstract:
Chromatin function in telomeres is poorly understood, but it is generally viewed as repressive. Yet, telomeric DNA sequences are transcribed into long non-coding RNAs named TERRA. As TERRA molecules mostly localize at telomeres, major research efforts have been made to understand their functions, and how TERRA transcription is regulated and affects telomere structure. This review describes the current state of knowledge about the nature of chromatin at telomeres, its functions, and the relation between chromatin structure and TERRA.
Introduction:

Telomeres are specialized structures at the extremity of chromosomes. They ensure genome stability [1-3] by preventing the recognition of chromosome ends as genuine double strand breaks and consequently their aberrant repair [4-6]. In most species, telomeric DNA is composed of short tandem repeats. In mammals, telomeres can be up to 50 kb-long, and are composed of head-to-tail repeats of the 6bp motif 5'- (TTAGGG/CCCTAA) -3'. Telomeres usually terminate with a single-stranded TTAGGG 3' -overhang that is up to 500 nt-long. In the absence of any maintenance mechanism, telomeres will shorten with cell division [7, 8]. Shelterin proteins are crucial telomere-binding proteins that are recruited through sequence-specific association with telomeric DNA. In this complex in mammals, TRF1 and TRF2 specifically bind to the double stranded TTAGGG sequence, while POT1 binds to the single-stranded overhang [9]. Shelterin proteins have two major functions at telomeres. They recruit other proteins to keep the DNA damage response in check [2-6, 10], and they control telomere length by regulating the recruitment of the telomerase holoenzyme complex [11-14]. When telomeres become critically short, shelterin proteins cannot bind sufficiently, and telomeres are not efficiently protected any longer. This situation is almost comparable to the occurrence of an irreparable intra-chromosomal double-strand break, and ultimately triggers permanent cell cycle exit [15].

Like any other genomic region, yeast and mammalian telomeres assemble into chromatin and therefore bear nucleosomes. Telomeric DNA sequences intrinsically disfavor nucleosome positioning, and this might explain some of the unusual chromatin structures observed at telomeres [16, 17]. Nonetheless the presence of nucleosomes implies interactions with dedicated machineries that support chromatin transactions, such as ATP-dependent chromatin remodellers and histone-modifying enzymes. As chromatin is a global determinant of eukaryotic DNA transactions, it should play an important role also at telomeres. In human cells, it has been reported that short telomeres have fewer nucleosomes than long telomeres, suggesting that telomere length is associated with distinct
chromatin features, like nucleosomal density. Historically, telomeric chromatin is viewed as generally heterochromatic, a type of chromatin that hinders DNA transactions. Yet the picture is even more complex because transcription also occurs at telomeric regions, an unusual feature in heterochromatin. This transcription leads to the production of long non-coding RNAs (lncRNAs) named TEIomere Repeat-containing RNA (TERRA) [18-22]. It has been suggested that TERRA plays a role in telomerase inhibition [23-25], but also in telomerase recruitment [26, 27]. TERRA metabolism also might influence telomere maintenance by recombination [28-32]. This review will describe the current state of knowledge on telomeric chromatin regulation and the potential links with TERRA and its transcription. By analogy with the well-known function of non-coding RNAs (ncRNAs) at *Schizosaccharomyces pombe* centromeric regions, or the putative role of ncRNAs in recruiting chromatin regulators, TERRA molecules are viewed as important players in telomeric heterochromatin regulation [33-39]. Telomerase RNA component, another ncRNA, also is associated with telomeres and plays an essential role in telomere elongation by telomerase [40, 41]. Currently, this ncRNA is not considered to play a role in telomeric chromatin regulation, and therefore will not be discussed in this review. Another telomere-related small RNA was described in mouse (tel s-RNA), but its function in telomere biology awaits further characterization [42].

**Telomeric chromatin**

For a long time, it was thought that telomeric regions assembled into heterochromatin because reporter genes placed near telomeres tend to be silenced. Moreover, silencing is positively influenced by the reporter gene proximity to telomeres and the telomere length. This phenomenon is known as Telomere Position Effect (TPE). TPE was extensively described in *Saccharomyces cerevisiae* [43] and then confirmed in *Schizosaccharomyces pombe* [44] and also in mammalian cell lines [45-47]. TPE is a classic ‘epigenetic’ metastable silencing phenomenon, and implies that telomere heterochromatinization is positively correlated with their length. Short telomeres are less heterochromatic, and in telomerase-negative cells, this might favour telomere lengthening by homologous recombination [30].
Whereas long telomeres favour DNA looping to install heterochromatin on distal region from telomeres and repress gene expression [48]. Nevertheless, yeast and mammalian telomeres rely on a specific DNA sequence for capping functions and thus, are not defined by an epigenetic phenomenon, as opposed to other repeated regions such as centromeres. In fact, perturbing specific epigenetic regulators destroys centromere function, while this has not been observed at telomeres. Therefore, chromatin might not be directly linked to telomere capping function and must be playing distinct roles.

*S. cerevisiae* telomeres are 250bp-long and are essentially nucleosome-free [16]. Subtelomeric regions, which are also made of distinct repeated DNA, are assembled into nucleosomes. However, the heterochromatin that forms and spreads along the subtelomeric regions depends on the binding of Rap1 (a shelterin protein) to its specific DNA motif essentially present on telomeres (*Figure 1A*) [49]. Rap1 then interacts with the Silence Information Regulation 4 (Sir4) protein that in turn recruits Sir3 and Sir2 to telomeres [49, 50]. Sir2 catalyses the removal of acetyl groups from lysine 16 of histone H4 (H4K16Ac), and leads to a hypoacetylated and repressive chromatin conformation [51, 52]. Hypoacetylation creates a binding site for Sir3 and Sir4 [53, 54] that then recruit more Sir2, providing a simple molecular explanation for hypoacetylation spreading to sub-telomeric regions and TPE enforcement in this species. Loss of Sir4 disrupts heterochromatinization and at the same time promotes telomere shortening. This is because Sir4, in addition to its hypoacetylation role, facilitates telomerase recruitment at telomeres with the help of Ku proteins [55]. Thus, *S. cerevisiae* heterochromatin partly promotes telomerase recruitment, and is at least not inhibitory to lengthening.

Similarly, in *S. pombe*, telomeres are nucleosome-free and able of TPE [44]. However, *S. pombe* heterochromatin resembles more to metazoan than to *S. cerevisiae* heterochromatin. For instance, *S. pombe* sub-telomeres harbour constitutive heterochromatin marks [56] and proteins, such as heterochromatin protein 1 (HP1), the histone methyltransferase (HMTase)
Clr4 (orthologue of the mammalian SUV39H HMTases) and the histone H9 lysine 3 trimethylation (H3K9me3) mark, all absent in *S. cerevisiae*.

In *S. pombe*, telomeric heterochromatinization also depends on a shelterin factor named Taz1 (*Figure 1B*). Taz1 tethers the SET domain-containing Clr4 to telomeres, and then Clr4 catalyses H3K9me3 deposition to neighbouring nucleosomes [57]. This histone mark in turn creates a binding site for Swi6 (an orthologue of HP1) on telomeres [58, 59]. The histone deacetylase Snf2/Hdac-containing Repressor Complex (SHREC) also is recruited by Taz1 and Swi6 to telomeres [60], and its Clr3-subunits deacetylates lysine 14 of histone H3. This contributes to the formation of repressive chromatin that spreads to sub-telomeric regions [60]. In *S. pombe*, the formation of sub-telomeric heterochromatin is promoted by the RNA interference (RNAi) machinery, as observed at centromeres [61]. Specifically, Dcr1, a subunit of the RNAi machinery, recruits Swi6 and the SHREC complex to telomeres [62]. Intriguingly, the direct involvement of the RNAi machinery in heterochromatin formation at telomeres seems to be specific to fission yeast, and is not observed at mammalian telomeres. Consistently, we never detected any RNAi pathway protein at heterochromatin purified from mammalian somatic cells [63-65].

In mammals (mouse and human cells, *Figure 1C*), the nature and function of telomeric chromatin remain unclear, and in our view, constitute a very controversial topic. Some groups reported that telomeres harbour marks of constitutive heterochromatin [38, 66-69] like most tandem repeats in the genome and consistent with TPE features [45, 47]. A more recent study showed that telomeres from most human cell lines are not heterochromatic [70] and that the heterochromatic mark H3K9me3 is enriched at telomeres only in cell lines that have activated the alternative lengthening of telomere (ALT) pathway, a recombination-based telomere lengthening mode [65, 70].

On the basis of the hypothesis that telomeres, as a source of TPE, should be heterochromatinized, it was initially proposed that telomeric heterochromatin assembly in mammalian cells is regulated by Suppressor of Variegation 3-9 homologues 1 and 2 (SUV39H1/H2) through H3K9me3 deposition on telomeres [68]. H3K9me3 acts as a landing
site for HP1α that recruits the Suppressor of Variegation 4-20 homologue 2 (SUV420H2) enzyme. SUV420H2 catalyses histone 4 lysine 20 trimethylation (H4K20me3), another mark of heterochromatin [71] with unclear functions. It was proposed that this pathway, which is very similar to what was described at pericentromeric regions, controls telomerase recruitment at telomeres, and controls the activation of telomerase-independent ALT [68, 71]. In this classical model, heterochromatin thus negatively regulates telomere lengthening. DNA methyltransferases (DNMTs) also are involved in telomere length control in mammals [72]. DNA methylation only occurs at subtelomeric regions because CpG sequences are not present in the telomere repeat motif [72-74]. As subtelomeric DNA is usually hypermethylated, it was suggested that in the mouse, this high level of DNA methylation somehow ‘compacts’ the telomeric fibre, rendering it less prone to damage and lengthening by recombination [72]. Intriguingly in human cells, loss of DNA methylation seems to have the opposite effect on telomere length compared with mouse cells [74-77]. In fact, subtelomeric chromatin is hypomethylated and telomeres are abnormally short in cells derived from patients with Immunodeficiency Centromeric instability and Facial anomalies (ICF) syndrome [75-77], who carry an inactivating mutation in DNMT3B. This shortening could be the consequence of defective telomere replication [76]. The different phenotypes observed in mouse and human cell lines might be due to the different average telomere length between the two models.

We recently analysed histone modifications at telomeres in mouse embryonic stem (ES) cells and fibroblasts [65]. We found low levels of H3K9me3 at telomeres in mouse ES cells and virtually no H3K9me3 in fibroblasts. H3K9me3 deposition mostly relies on Set domain bifurcated protein 1 (SETDB1) recruitment [65], and this heterochromatin mark is important for stimulating local nucleosome exchange, which in turn favours transcriptional processivity and recombination. Therefore, in mouse ES cells, heterochromatin indirectly promotes telomere recombination and transcription, in disagreement with the SUV39H model described above.
More recently, it has been proposed that Polycomb activities also regulate telomere biology. The Polycomb group protein EZH2 has been suggested to bind telomeres where it is able to catalyse histone H3 lysine 27 trimethylation (H3K27me3) and this is required to stabilize H3K9me3, H4K20me3 and HP1 telomeric enrichment [36]. While it has been suggested that H3K27me2/3 and H3K9me2/3 can coexist on the same native nucleosome [78], ChIP-sequencing approaches have found that H3K27me3 is largely non-overlapping with H3K9me3 throughout the genome, including at repeated DNA sequences [79-88]. The synergy between Polycomb activities and H3K9me3/HP1 [36, 89] is highly unusual, especially given that genetic screens aiming at identifying Polycomb regulators or constitutive heterochromatin activities usually retrieve distinct group of genes [79-81, 84-86]. Moreover, the role of Polycomb genes in telomere regulation remains unclear in the absence of described telomere phenotypes in Polycomb mutants. Using unbiased telomere chromatin proteomics, we failed to detect significant EZH2 binding or activity at telomeres in mouse ES cells [65]. This discrepancy might be explained by species- or tissue-specific differences, or different technical approaches, measuring relative versus absolute amounts of histone modifications.

Regardless of its nature at telomeres, heterochromatin is viewed as an important regulator of telomere length regulation. According to the classical view, telomeric heterochromatin is in a closed/condensed state that does not allow recombination and telomerase recruitment. Our recent results indicate that heterochromatin does not form on telomeres in all mouse cell types [65], consistent with the study on human cell lines [70]. However, when heterochromatin forms at telomeres (like in mouse ES and in ALT cells), it promotes rather that inhibits telomere recombination and lengthening [65], a trend globally comparable, albeit mechanistically different, to what happens in S. cerevisiae. As heterochromatin formation is often linked to the presence of ncRNAs, we discuss below TERRA production and the possible links between TERRA and local chromatin regulation.
**TERRA biogenesis**

The biological function of ncRNAs is very difficult to characterize experimentally. In the absence of any identifiable coding function, such RNAs could have a structural (e.g., ribosomal RNAs), catalytic (e.g., telomerase RNA), sequence-specific recruiting, or local protecting role. In addition, one important aspect of ncRNA function is whether it acts locally, in which case its function as a ncRNA is difficult to separate from molecular events leading to its production; or whether it acts “in trans” at long distances from its production site. ncRNA production, which involves chromatin remodelling, could have a major “in cis” function in the local chromatin regulation, regardless of the RNA that is made. Hence, we first detail below what is TERRA and how it is produced.

The first evidence that telomere sequences are transcribed came from a work performed in *Trypanosoma brucei* [90]. Since then, telomere transcription has been observed in different organisms, such as yeast [21, 22, 25], mouse [18] and humans [20], which indicates conservation during evolution, an indirect sign of functional importance.

TERRA is a single-stranded IncRNA originating from the transcription of the C-rich telomeric DNA strand. Thus, TERRA is a G-rich RNA containing 5’-UUAGGG-3’ repeats [18, 20-22, 25]. TERRA length ranges from 100bp to more than 9kb in mammals [20, 91] and by consequence long TERRA transcripts are difficult to be detect by common Northern-Blot. In Yeast, TERRA range is about 400bp in yeast [22, 25]. TERRA levels are cell-cycle regulated [29, 69, 91]. Indeed, TERRA levels peak in early G1, and decrease in late G1, reaching the lowest level in late S phase, a time that roughly corresponds to telomere replication. After the G2/M phase, TERRA levels start to increase again. RNA polymerase II catalyses transcription at telomeres [18, 25, 73, 91, 92]. Subunits of the RNA polymerase I and III complexes also have been identified during the purification of telomeric chromatin [64]; however, the biological significance of these associations is unclear and might not be linked to telomere transcription.

In all species studied so far, TERRA transcription starts in the subtelomeric region and proceeds from the centromere toward the telomere direction [20, 22, 25, 73, 92]. In S.
pombe, other telomeric lncRNAs have been identified in both sense (αARRET) and antisense directions (ARIA and ARRET) (Figure 2A) [21]. Some of these anti-sense telomeric RNAs are also present in plants [93] and in human and mouse cells [18, 20], albeit at a much lower level than TERRA. Their functions are unknown. As telomere and subtelomeric sequences are made of repetitive DNA, the mechanisms involved in RNA transcription initiation, elongation and termination are difficult to address, and the characterization of TERRA species also has led to controversial findings and models. We detail below the state of knowledge about TERRA biogenesis because it is likely that the mechanisms presiding to TERRA synthesis are linked with TERRA functions.

As TERRA is transcribed by RNA polymerase II, it contains a methyl-cap at the 5’ end and a poly-A tail at the 3’ end, like most mRNA species [18, 25, 91]. In yeast and human cells, all TERRA species are capped. In S.cerevisiae, almost all TERRA RNAs also have a poly-A tail (Figure 2B). In contrast, only 7% of human TERRA RNAs have a poly-A tail (Figure 2C) like TERRA in S.Pombe [27, 91], whereas this fraction has not been quantified in mouse cells [18]. In S. cerevisiae, poly-adenylation seems to require the action of the canonical poly-A polymerase (PAP1), but the precise molecular mechanisms of termination remain unclear because TERRA does not have the canonical poly-A signal (5’-AAUAAA-3’) normally found at the 3’ of most class II genes [25]. Indeed, in human cell lines, TERRA-poly-A(-) terminates preferentially with the 5’-UUAGG-3’ sequence, whereas TERRA-poly-A(+) terminates with 5’-UUAGGG-3’ [91], a finding which supports the idea of regulated 3’ end TERRA processing. Moreover, as the TERRA DNA template strand mostly ends with the ATC-5’ sequence [94], TERRA transcription might not process until the end of the telomere.

As is the case for most mRNA, the poly-A tail stabilizes TERRA both in yeast and in human [25, 91]. It also correlates with TERRA sub-nuclear distribution in human cells: 60% of TERRA-poly-A(-) is in the nucleoplasm, while the remaining 40% is chromatin-associated. Conversely, TERRA-poly-A(+) molecules are mostly in the nucleoplasm [91]. This suggests that chromatin-associated TERRA is not poly-adenylated, and that TERRA-poly-A(+) and TERRA-poly-A(-) might underlie distinct functions.
TERRA might also be regulated at the initiation step in the subtelomeric region. While globally heterochromatic, subtelomeres can locally bear euchromatin marks, with an enrichment in histone H3 lysine 4 trimethylation (H3K4me3) [92, 95, 96], deposited by the Mixed Lineage Leukemia protein (MLL) [97], and histone H3 lysine 27 acetylation (H3K27Ac) enrichment [95] and RNA polymerase II binding [92, 95, 96], indicative of transcriptional initiation. However, many aspects regarding the initiation mechanisms remain unclear because TERRA initiation sites lack canonical promoter sequences.

In yeast, the subtelomere sequence is made of X-elements and Y’-elements, bound by the Sir silencing complex and also Rif 1 and Rif2 (Figure 2B). While X-element sequences are strongly repressed by the Sir silencing complex and Rif1 and Rif2, the Y’ elements are weakly repressed, only by Rif 1 and Rif 2. Y’ element are enriched in H4K16Ac and harbor transcribed open reading frames and and TERRA initiation sites [19, 98].

In mammals, several groups identified TERRA production at many (if not all) telomeres [20, 33, 73, 74, 92, 99], whereas others suggested that the bulk of TERRA is produced only from one or a very limited number of (sub)telomeric regions [100-102]. These two models have fundamentally different implication on the function of TERRA. The seminal study on TERRA showed transcription from several human telomeres [20], suggesting the presence of a transcription start site (TSS) on each of them. RNA fluorescent in situ hybridization indicates that most telomeres can be found associated with TERRA but this does not prove that TERRA is transcribed from each telomere, as TERRA could be made from a limited number of loci then addressed to other telomeres. Transcription from multiple telomeres was also demonstrated by the same group [74, 99]. Two subtelomeric promoter types, Type-I and Type-II, were identified at 1Kb and 5-10 Kb from the subtelomere-telomere boundary, respectively (Figures 2C and 2D) [33, 73, 92]. Both promoter types include CCCTC-binding factor (CTCF) binding sites and CpG island elements [73, 92]. At type I promoters, CpG islands are composed of three distinct repetitive tracts of 61bp, 29bp and 37bp in length. These elements are referred as “61-29-37 repeats” and are present at 13 distinct human chromosome ends [73]. These 61-29-37 repeats are bound by RNA polymerase II
and have intrinsic promoter activity [73, 92, 95]. Moreover, the 61-29-37 repeats have high 
CpG content and are methylated by DNMT3B and DNMT1, unlike most CpG islands 
elsewhere in the genome, which usually escape DNA methylation. In fact, TERRA initiation 
at these elements is controlled by DNA methylation [73, 76]. In addition, two tandem DNA 
binding motifs are present upstream of the 61-26-37 repeats. They are bound by CTCF and 
Cohesin (Rad21) [73, 92, 96]. These two chromatin-organizing factors cooperate to promote 
TERRA correct orientation [96]. Type-II promoters were identified by RNA-seq analysis on 
ten other chromosomes, and are located 5-10 kb upstream of the telomere tracts. TERRA 
production at different telomeres has been measured also in mouse embryonic fibroblasts 
and during early developmental stages [33].

The other model proposes that TERRA molecules are produced at a single or a limited 
number of telomeres [100-102]. Then, TERRA molecules can travel (via uncharacterized 
mechanisms) through the nucleus to other telomeres and also to a subset of genes to 
regulate the local expression and chromatin composition. One implication of this model is 
that telomere transcription, and therefore the inherent local chromatin remodelling, plays a 
limited role in telomere biology, while TERRA molecules are important. The other implication 
is that the status of one telomere has the potential to govern that of all the other telomeres. It 
was proposed that in U2OS human cancer cell lines, which maintain their telomeres with 
ALT, chromosomes 20 and X are the only TERRA producers [102]. U2OS ALT telomeres 
harbor heterochromatin features [65, 70] and non-physiological TERRA expression [24]. 
Moreover these observations were not made in telomerase-positive cell lines [20]. Similarly, 
in mouse ES cells, the TERRA FISH signal forms two main dots that co-localize with the X 
chromosomes, and drive X chromosome pairing during X inactivation [101]. In Mouse 
Embryonic Fibroblasts (MEFs), the chromosome 18 telomere is the main source of TERRA 
[100]. In both cell types, TERRA signals are also found associated with other telomeres by 
FISH and with non-telomeric regions by CHIRT-seq [37, 103], suggesting a function in trans. 
Like the X telomere transcript in mouse ES cells, the chromosome 18 TERRA seems thus to 
travel to other genomic regions. It was proposed that in these cells, this interaction in trans
protects telomeres from the DNA damage response [100], therefore, presumably supporting an entirely different function from X chromosome pairing described in mouse ES cells [101]. Live-cell imaging in a human cancer cell line, using MS2 knock-ins to track endogenous TERRA from one telomere, shows TERRA molecules diffusing into the nucleus [104], sometimes co-localizing as a cluster with one telomere. This result could reflect TERRA binding to another telomere in trans, or it could reflect the actual TERRA production from the modified telomere. Similar to the mouse situation, disrupting this single telomere TERRA transcript leads to the activation of the DNA damage response throughout the genome. In a model of a single telomere-producing TERRA [100-102], addressing TERRA function by disrupting the unique TERRA promoter should be straightforward. This was attempted, but TERRA RNA production was not entirely abrogated [100, 102], complicating the interpretation and clear conclusions about TERRA roles in trans.

The distinction between the two models is critical to understand TERRA in cis and/or in trans functions. As there is currently no clear explanation about the discrepancies between laboratories, more quantitative approaches might be required to clarify this important question.

TERRA RNA functions:

It has been proposed that lncRNAs regulate genome functions [105]. Specifically, they work by interacting with chromatin-modifying enzymes and nucleosome-remodelling factors to modulate chromatin structure. For this, they might recruit chromatin factors to specific genomic regions, or they might anchor chromatin factors away from target regions. LncRNAs could also act as scaffolds to build chromatin-modifying complexes, without necessarily targeting them to specific loci. As the lncRNA TERRA is associated with telomeres, it might play an important function in telomeric chromatin regulation. TERRA cannot be genetically inactivated easily, and consequently its physiological functions remain largely unknown. Several proteins that interact with this lncRNA were identified in vitro by affinity pull-downs [37, 38, 106, 107], including TRF1 and TRF2, HP1, the Origin Recognition Complex (ORC)
[38], several heterogeneous nuclear ribonucleoproteins (HnRNPs) [107], and also various interactors that need to be further characterized to determine whether TERRA acts as a local recruiter of biologically relevant activities. Since many relevant chromatin proteins promiscuously interact with long RNA in vitro, it will be important to determine whether in vitro interactors are recruited, and if the case, whether TERRA targets specific chromatin functions to telomeres. Nevertheless, some TERRA-related regulations are starting to emerge.

In S. cerevisiae, forced TERRA expression from one telomere induces exonuclease 1-mediated shortening of that telomere without any measurable change in the length of the other telomeres [108, 109]. This key observation suggests that at least in this artificial setting, TERRA or the process of telomere transcription mostly has an effect in cis, and that an appropriate level of TERRA or transcription is required for telomere integrity.

Telomeric transcription has the potential to generate local TERRA-telomere DNA hybrids that create R-loop structures and a displaced single-stranded G-rich telomeric DNA. As these structures generate replication stress, TERRA must be cleared from telomeres during S phase to ensure complete telomere replication [29, 110] (Figure 3A). R-loop clearance is ensured by an RNase H activity. In human ALT-positive cells, RNase H1 is highly enriched at telomeres and prevents RNA-DNA hybrid accumulation, which would otherwise promote increased homologous recombination (Figure 2C) [28]. In telomerase-negative yeast, RNase H1 and RNase H2 also act to limit RNA-DNA hybrid accumulation at telomeres (Figure 2B), thus controlling their elongation through homologous recombination (Figure 3A), [29, 30]. In yeast, RNA-DNA hybrid accumulation is regulated also by the THO complex, which is normally involved in mRNA export. Indeed, inactivation of Tho2p (a THO complex subunit) leads to RNA-DNA hybrid accumulation (Figure 2C) [30-32] and exonuclease 1-dependent telomere shortening [30]. These findings indicate that two different RNA processing pathways regulate telomere RNA-DNA hybrid levels [30, 32], with an apparently stronger action for RNaseH [32], but the connection between the two pathways are not known. In human ALT-positive cells, FANCM was identified has a new regulator of telomeric R-loops
[111, 112], by limiting R-loop accumulation to control replicative stress (Figure 2C), representing another indication that TERRA processing is a highly regulated process.

In addition to forming potentially toxic local R-loop structures, TERRA might also fold into G-quadruplex structures (G4) [113, 114] which can be bound by TRF2. It seems that TRF2 is able to simultaneously bind to TERRA-RNA-G4 and to telomere-DNA-G4, forming a tri-complex, and helping TERRA association with telomeric DNA [115] (Figure 3C), providing a potential mechanism for TERRA recruitment to telomeres in trans.

Another potentially relevant TERRA function is the ability to associate with the RNA component of the telomerase holoenzyme (Figure 3B). TERRA could inhibit telomerase activity by competing for binding to the single-stranded telomeric DNA overhang, which is the normal substrate for this enzyme. This inhibition was observed in vitro [23]. Conversely, in yeast, TERRA seems to stimulate telomerase (Figure 3B) [26, 27]. It could be that TERRA positively regulates telomerase accessibility to telomeres by displacing inhibitory proteins from telomeres [116]. TERRA can also anchor hnRNPA1 proteins away from the telomere to allow proper telomere replication [35, 117].

TERRA might also regulate heterochromatin formation at telomeres because TERRA downregulation correlates with reduced heterochromatin marks at telomeres [36, 38], whereas higher TERRA levels, as in G1 phase [69], or longer TERRA correlates with enrichment of these marks [33, 69]. TERRA has been shown to bind HP1α [38, 118], suggesting a role in heterochromatin formation. This might seem counterintuitive because heterochromatin generally correlates with transcriptional silencing. The mechanisms by which TERRA correlates with heterochromatin is poorly understood, but TERRA, like several other IncRNAs, can also directly interact with different enzymes involved in heterochromatin formation [33, 34, 119-121] (Figure 3D). However, it must be noted that this interaction generally appears to be RNA sequence-independent, suggesting that locus-specific recruitment might depend on other factors or on the ability of the heterochromatin factor to identify nascent RNA and associate with it. Upon telomere deprotection, increased TERRA
production induces the recruitment of SUV39H1 to human telomeres and triggers local H3K9me3 deposition [33]. SUV39H1 N-terminus contains a chromodomain that directly interacts with TERRA [33], providing a molecular mechanism for SUV39H1 presence at deprotected telomeres. Along this line, SUV39H1 recruitment to pericentromeric regions is stabilized by interaction with local nascent RNA [122, 123], suggesting that RNA interaction might be a recruiter of molecules involved in heterochromatin activities [123], or might at least contribute to their function by stabilizing enzyme binding to chromatin [122]. Upon telomere deprotection, TERRA also promotes the recruitment of the histone demethylase LSD1 and G-quadruplex RNA binding [39, 121]. The consequence of LSD1 presence at telomeric nucleosomes was not fully explored, but it was shown that LSD1 recruitment promotes 3’ telomere overhang processing though stimulation of the double-strand break repair protein MRE11 [39].

Moreover, a protein named translocated in liposarcoma (TLS/FUS) binds to TERRA-RNA-G4 structures and to telomeric DNA, and forms a ternary complex proposed to anchor TERRA to telomeres [48, 119, 124]. Tethered TLS/FUS in turn promotes H3K9me3 deposition at telomeres by an unknown mechanism [48, 119].

Finally, on the basis of TERRA interaction with Polycomb repressive complex 2 (PRC2) subunits in vitro [120], it was proposed that TERRA recruit Polycomb activities to telomeres [36]. It should be noted that in some cases the interaction between a chromatin enzyme and RNA inhibits the enzymatic activity. This is true for LSD1 [121], and EZH2 [125], suggesting that a simple model where TERRA mediates the recruitment of enzymes such as EZH2 to work at telomeres is probably incomplete.

It was also suggested that TERRA modulates the telomeric chromatin structure through ATRX eviction from telomeres after a direct interaction with this chromatin remodelling enzyme [37]. ATRX associates with G-rich DNA sequences [126, 127], interacts with the DAXX histone chaperone [128], and promotes heterochromatin formation [129]. While the potential mechanisms were not explored, the same authors reported that TERRA knock-down led to telomerase stimulation and, counterintuitively, to telomere damage [37].
the same mouse ES cell line, we found that ATRX recruitment to telomeres largely depends on SETDB1 and does not seem to correlate negatively with TERRA [65]. The loss of ATRX in the context of heterochromatinized telomeres stimulates telomere recombination, while the loss of ATRX in the absence of heterochromatin at telomeres has no measurable impact. These effects occur without any strong change in TERRA levels, suggesting that TERRA functions are unlikely to be strictly linked to heterochromatin formation at telomeres in ES cells. However, detailed insights await more mechanistic characterizations.

**Conclusions:**

Telomeric chromatin features and telomeric transcripts actively participate in telomere stability that is required for ensuring genome stability. From yeast to human cells, telomeric chromatin has been defined as silent heterochromatin, due to TPE. Typical constitutive heterochromatin marks are enriched at chromosome ends: Sir4-Sir3-Sir2 in *S. cerevisiae* [49, 51, 52, 54], and H3K9me3 and Swi6/HP1 in *S. pombe* and mammals [57, 58, 65]. However, in most cells, telomeres are nucleosome-poor or have poorly positioned nucleosomes, a situation which generally disfavors constitutive heterochromatin formation. The organization of telomeric heterochromatin remains unclear and controversial, which is why its function is difficult to define. Mechanisms of TERRA mediated functions is challenging to address due to the telomeric RNA repeated sequence and the different producing genomic loci, and therefore we are facing major technical challenges. Moreover, in order to fully understand the interplay between TERRA and chromatin regulation, measuring TERRA steady-state levels will not be sufficient. We will need to detail the TERRA transcription mechanisms with more dedicated methods such as nuclear run-on, studying the RNA polymerase II phosphorylation state and common mechanisms linked to transcriptional elongation and termination, which are determined by chromatin. Discovering a significant function for TERRA in regulating telomeric chromatin will thus require robust and quantitative tools that await further development.
Acknowledgements

Work in the Dejardin lab is supported by grants from ARC equipe labelisée 2016, the ERC CoG METACHROM, from INCA, from the Fondation Schlumberger pour l’éducation et la recherche and from Merck (MSD). A.B. is supported by the University of Montpellier and by The Fondation pour la Recherche Medicale.

References:


**Figures:**

**Figure 1:** Telomeric heterochromatin establishment and players involved. (A) In *S.cerevisiae*, the shelterin component Rap1 recruits Sir4, Sir3, and Sir2. Sir2...
deacetylates H4K16, promoting more Sir4, Sir3, and Sir2 binding and spreading of repressive chromatin to the subtelomeric region. (B) In S. pombe, the shelterin factor Taz1 interacts with the histone methyltransferase Clr4 that trimethylates H3K9 and creates binding sites for Swi6. Swi6 recruits SHREC that deacetylates H3K14. Moreover, the RNAi machinery also contributes to Swi6 and SHREC recruitment, contributing to heterochromatin formation. (C) In mouse embryonic stem (ES) cells, SETDB1 trimethylates H3K9. This creates a binding site for HP1 and promotes recruitment of other heterochromatin factors, such as ATRX and DNA methyltransferases (DNMTs). Telomeric heterochromatin is also present at subtelomeres, where DNMTs methylate CpG motifs.

Figure 2: Telomere transcription. (A) The telomere transcriptome of S. pombe. G-rich RNAs include telomeric TERRA RNAs and sub-telomeric ARRET RNAs. C-rich RNAs include telomeric ARIA RNAs. αARRET are sub-telomeric RNAs complementary to the ARRET RNAs. (B) S. cerevisiae telomeres are transcribed into TERRA RNAs. Transcription starts from the Y'-element in the sub-telomeric region. RNase H and the THO complex destabilize RNA-DNA hybrids. (C) In human cells, telomere DNA is transcribed into TERRA RNAs. Transcription starts from Type-I or Type-II promoters. RNase H and FANCM prevents aberrant accumulation of RNA-DNA hybrids on telomeres. (D) Type-I and Type-II promoters include CTCF binding sites and CpG island promoter elements. CpG island of Type-I promoter are composed of repetitive “61-29-37 repeats” and repressed by DNMTs.

Figure 3: TERRA hypothetic functions (A) TERRA regulates telomere length. TERRA forms co-transcriptional RNA-DNA hybrids that can lead to telomere shortening by interfering with the replication machinery, or to telomere elongation, through homologous recombination. (B) TERRA controls telomerase activity either by inhibiting or by stimulating its recruitment to telomeres. (C) TERRA RNAs promote telomere integrity. TERRA is associated with telomeres, through TRF2 binding, and prevents the
DNA damage repair pathway activation and telomere degradation. (D) TERRA regulates heterochromatin formation at telomeres. TERRA interacts with different heterochromatin factors and tethers them to telomeres.
A. *S. cerevisiae*:

- Spreading
- Subtelomere
- Telomere
- RAP1
- Sir 4
- Sir 2
- Sir 3
- H4K16Ac

B. *S. pombe*:

- Subtelomere
- Spreading
- RNAi machinery
- Dcr1
- Swi6
- Swi6
- Swi6
- Swi6
- Telomere
- Clr4
- Taz1

C. Mouse ES cells:

- Subtelomere
- Telomere
- SETDB1
- ATRX
- HP1
- RNA pol II
- RNA
- DNMTs
- TERRA RNAs
- TTAG
- 5-mC
- H3K4me3
- H3K9me3
- 5-mC
- Nucleosome
**A. S. pombe:**

- **Subtelomere to Telomere:**
  - aARRET transcription
  - TERRA transcription
  - Telomere

- **TTAGGG 3'5'**
  - 7meG
  - Telomere

- **Telomere Subtelomere:**
  - TERRA RNAs
  - 7meG

**B. S. cerevisiae:**

- **Subtelomere to Telomere:**
  - Sir Silencing complex
  - Rif1/2

- **TTAGGG 3'5'**
  - 7meG
  - Telomere

- **Telomere Subtelomere:**
  - TERRA RNAs
  - 7meG

**C. Human:**

- **Subtelomere to Telomere:**
  - TERRA transcription
  - Telomere
  - TERRA RNAs
  - 7meG

- **Type I**
  - FANCM
  - RNaseH
  - Termination?

- **Type II**

**D. Type-I promoter**

- CTCF
- 61-29-37 repeats
- Cpg Island

**Type-II promoter**

- CTCF
- 61-29-37 repeats
- Cpg Island
RNA-DNA hybrids

Telomere elongation

Telomerase

7meG

Homologous recombination

Telomere shortening

Collision

Replication machinery

Telomere degradation and fusion

TRF2

TERRA-RNA-G4

TERRA interactome

Telomere-DNA-G4
TTAGGG

7meG

Histone modifications

Proteins

TERRA G-quadruplex

Chromatin modifiers

R-loop

Replication stress/Instability

Homologous recombination

5' 3'