Review

CellPress

Telomere functions grounding on *TERRA firma*

Claus M. Azzalin[1] and Joachim Lingner[2]

[1] Institute of Biochemistry, Eidgenössische Technische Hochschule Zürich (ETHZ), 8093 Zürich, Switzerland
[2] Swiss Institute for Experimental Cancer Research (ISREC), School of Life Sciences (SV), Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland

Long noncoding telomeric repeat-containing RNAs – TERRAs – are transcribed in a regulated manner from telomeres throughout eukaryotes. TERRA molecules consist of chromosome end-specific subtelomeric sequences and telomeric repeats at their 3′ ends. Recent work suggests that TERRA sustains several important functions at chromosome ends. TERRA can regulate telomere length through modulation of exonuclease 1 and telomerase, it may promote recruitment of chromatin modifiers to damaged telomeres and thereby enable DNA end-processing, and it may promote telomere protein composition changes during cell cycle progression. Furthermore, telomere transcription regulates chromosome-end mobility within the nucleus. We review how TERRA, by regulated expression and by providing a molecular scaffold for various protein enzymes, can support a large variety of vital functions.

TERRA in telomere biology

Telomeres are nucleoprotein structures capping the physical ends of linear eukaryotic chromosomes. They consist of telomeric repeat DNA, a large number of specialized proteins, and RNA [1,2]. If sufficiently long, and assembled with the correct set of telomeric proteins, telomeres protect chromosome ends from degradation and DNA repair activities that physiologically seal chromosomal DNA breaks. Telomeres are dynamic because they change composition and function during the cell cycle, and possibly during development. Semiconservative DNA replication is unable to fully copy the ends of linear chromosomes, leading to continuous shortening of the ends of telomeres with each round of DNA replication. In addition, nucleolytic end-processing contributes to telomere shortening. A telomere-lengthening enzyme known as telomerase (see Glossary) adds back telomeric repeat DNA that was lost; however, in humans telomerase is only found in a limited number of cell types. Therefore, telomere shortening can occur with each division in the absence of telomerase, and this plays a major role in regulating cellular lifespan. Short telomeres elicit permanent DNA damage responses that trigger cellular senescence, a powerful anticancer barrier that suppresses unlimited proliferation in many cells of the human soma that lack telomerase.

Telomeres carry features of heterochromatin in that they contain HP1 (heterochromatin protein 1) proteins and heterochromatin-typical histone marks including histone 3 trimethylated at lysine 9 (H3K9me3) and H4K20me3 [3]. Thus, genes placed experimentally near telomeres are transcriptionally silenced in a variegated fashion [4,5]. Nonetheless, telomeres were discovered to be transcribed into long noncoding (lnc) RNAs termed TERRA [6,7]. Telomere transcription is a widely conserved eukaryotic feature because it has now been reported in numerous phyla. Since its discovery, TERRA has been associated with a large variety of apparently disconnected telomere functions and different telomeric states. We discuss how telomere transcription, being regulated during the cell cycle and induced upon telomere damage, may trigger transitions between alternative telomeric states, and how TERRA may sustain a variety of telomere functions by providing a scaffold for diverse enzymes. We also discuss how defective TERRA biogenesis and regulation can impair the physiological roles of telomeres during DNA replication and end-protection and can be associated with human disease.

TERRA biogenesis

The C-rich telomeric strand provides the template for TERRA transcription starting from subtelomeric regions.

Glossary

Alternative lengthening of telomeres (ALT) pathway: telomere lengthening mechanism that relies on DNA recombination instead of telomerase. ALT is active in 10–15% of human tumors including sarcomas, gastric carcinomas, central nervous system malignancies, and bladder carcinomas.
DNA methyltransferases: enzymes that catalyze transfer of methyl groups to cytosine or adenine bases in DNA.
Heterochromatin: complex of DNA, protein, and RNA which is tightly packaged and in which gene expression is generally repressed.
Nonhomologous end-joining: DNA repair pathway in which the two ends of a DNA ds break are ligated together.
Nonsense-mediated RNA decay (NMD): RNA quality-control pathway that targets for degradation aberrant mRNAs carrying premature stop codons.
R-loops: nucleic acid structures in which RNA invades ds DNA by base-pairing with one of the DNA strands while the other DNA strand is displaced to form a loop.
RNase H: endonuclease that degrades RNA that is base-paired with DNA.
Shelterin: a group of proteins (six in mammals: TRF1, TRF2, POT1, TPP1, TIN2, Rap1) that are assembled on telomeres and cap the ends of chromosomes.
Telomerase: cellular reverse transcriptase that uses an internal RNA template for the synthesis of short telomeric DNA repeats. Telomerase counteracts telomere shortening that occurs due to the inability of conventional DNA polymerases to fully copy chromosome ends.

Corresponding authors: Azzalin, C.M. (claus.azzalin@bc.biol.ethz.ch); Lingner, J. (joachim.lingner@epfl.ch).
Keywords: TERRA; long noncoding RNA; telomeres; chromatin.

0962-8924/
 http://dx.doi.org/10.1016/j.tcb.2014.08.007

Box 1. The telomeric transcriptome

Together with TERRA, the ARIA, ARRET and αARRET noncoding RNA species are transcribed from chromosome ends in *Schizosaccharomyces pombe* [8,9] constituting the telomeric transcriptome in this organism (Figure I). ARIA comprises mostly or exclusively C-rich telomeric repeats. Inactivation of RNA-dependent RNA polymerase activities did not affect ARIA cellular levels [8], ruling out the possibility that ARIA is generated using TERRA as a template and suggesting that telomeric DNA repeats *per se* can function as transcription start-sites for RNA polymerases. ARIA species are readily detected in plants, but are largely suppressed in mammals [53]. ARRET and αARRET are two complementary subtelomeric lncRNAs. ARRET exists also in budding yeast [10], while the existence of αARRET remains to be tested outside *S. pombe*. Telomeric and subtelomeric RNAs are regulated through partially overlapping pathways. For example, the *S. pombe* telomeric factors Taz1 and Rap1 suppress all telomeric lncRNA species. By contrast, inactivation of the RNAi pathway leads to stabilization of only ARRET, leaving TERRA and ARIA apparently unaffected [8,9].

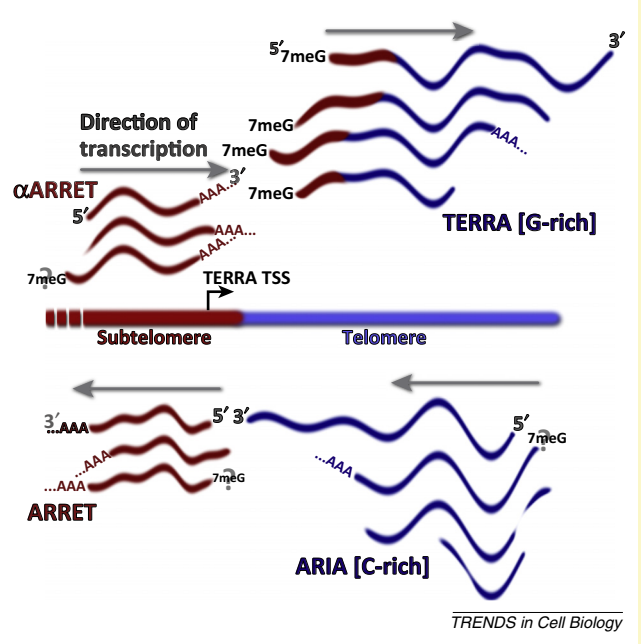

Figure I. The telomeric transcriptome. Schematic representation of RNA species generated by transcription of eukaryotic chromosome ends. RNA and DNA sequences of telomeric or subtelomeric origin are indicated in blue or red, respectively. AAA... indicates poly(A) tails present at the 3′ end of a fraction of TERRA and ARIA, and of all ARRET and αARRET species. 7meG indicates 7-methyl-guanosine caps present at TERRA 5′ ends, while their presence on other RNA species has not been tested yet (question marks). Grey arrows indicate the direction of transcription. TERRA transcription start-site (TSS) is also indicated. The totality of depicted telomeric transcripts have so far only been identified in *S. pombe*.

Consequently, individual TERRA molecules start with a subtelomeric RNA tract followed by a variable number of telomeric G-rich repeats (5′-UUAGGG-3′ in vertebrates). In addition to TERRA, other lncRNA species originate from chromosome ends and form, together with TERRA, the telomeric transcriptome (Box 1). In fission yeast, where these species have been comprehensively characterized, chromosome ends are transcribed into TERRA and the antisense RNA ARIA (named from *aria*, Italian for 'air', the complement to 'earth', *terra*), as well as into two complementary subtelomeric RNAs dubbed ARRET and αARRET [8,9]. Although TERRA has been the major focus of research, it will be important to better characterize all different members of the telomeric transcriptome and dissect whether and how they crosstalk.

The main RNA polymerase synthesizing TERRA is RNA polymerase II (RNAPII), as first shown in mammalian cells using RNAPII inhibitors [7], and successively confirmed in budding yeast using thermosensitive RNAPII alleles [10]. Similarly, inactivation of the RNAPII subunit Rpb7 led to a reduction of lncRNAs originating from chromosome ends in fission yeast [9]. Further confirming the involvement of RNAPII in TERRA biogenesis, components of the RNAPII holoenzyme were shown to associate with telomeres in yeasts and human cells [9–11]. Human and *Saccharomyces cerevisiae* TERRA 5′ ends carry 7-methyl-guanosine cap structures and approximately 10% of TERRA species are polyadenylated at their 3′ ends in humans and fission yeast, whereas all TERRA molecules are polyadenylated in budding yeast [6,9,10,12]. TERRA polyadenylation influences its stability and association with chromatin (Figure 1).

In yeast and humans, discrete TERRA transcription start-sites have been identified in the subtelomeric region of several chromosome ends [11,13]. However, the lengths of the telomeric repeat tracts of TERRA are heterogeneous. Telomere elongation in HeLa and HT1080 cancer cells upon overexpression of telomerase led to TERRA elongation, indicating that large portions of the telomeric tract can be transcribed [14]. Similarly, TERRA, specifically transcribed from an inducible promoter inserted immediately upstream of a unique telomere in HeLa cells, extended for several kilobases, covering size ranges comparable to those covered by the inducible telomere itself [15]. Nevertheless, measurement of the UUAGGG tract length by reverse transcription in the absence of dGTP indicated that a large fraction of human TERRA molecules does not contain cytosine-lacking stretches that exceed 400 bases, even though telomeres in the same cells extend for several kilobases [12]. This suggests that the pure UUAGGG-tract length is considerably shorter than its C-rich telomeric DNA template, or that the DNA template contains interdispersed mutant guanosines as recently documented in several human cells [16]. Nevertheless, we suspect that telomeric transcription represents a challenge for polymerases and that only a fraction of telomeres are fully transcribed.

Human TERRA promoters comprising CpG dinucleotide-rich tandem repeats of 39 and 37 base pairs have been identified in approximately half of human subtelomeres (Figure 1) [11]. They lie at variable distances from the telomeric tract immediately upstream of TERRA transcription start-sites. RNAPII binds to TERRA promoters *in vivo*, and transcriptional activity is repressed by CpG methylation cooperatively established by the DNA methyltransferase enzymes DNMT1 and DNMT3b [11]. TERRA CpG island promoters are often preceded by a third repetitive element comprising 61 bp tandem repeats, which are bound by the transcription regulator CTCF (CCCTC-binding factor) and the cohesin Rad21 (radiation-sensitive 21). Depletion of CTCF diminished TERRA levels as well as RNAPII and cohesin binding to subtelomeres [17]. This