

Viceconte et al

# **PAR-TERRA is the main contributor to telomeric repeat-containing RNA transcripts in normal and cancer mouse cells**

Nikenza Viceconte<sup>\*,1</sup>, Axelle Lorient<sup>\*,1</sup>, Patrícia Lona Abreu<sup>2</sup>, Marion Scheibe<sup>3</sup>, Albert Fradera Sola<sup>3</sup>, Falk Butter<sup>3</sup>, Charles De Smet<sup>1</sup>, Claus M. Azzalin<sup>2</sup>, Nausica Arnoult<sup>4\$</sup> and Anabelle Decottignies<sup>1\$</sup>

<sup>1</sup> Genetic and Epigenetic Alterations of Genomes, de Duve Institute, Université catholique de Louvain (UCLouvain), 1200 Brussels, Belgium

<sup>3</sup> Quantitative Proteomics, Institute of Molecular Biology (IMB), 55128 Mainz, Germany

<sup>2</sup> Instituto de Medicina Molecular João Lobo Antunes, Faculdade de Medicina da Universidade de Lisboa, 1649-028 Lisboa, Portugal

<sup>4</sup> MCB-D-University of Colorado Boulder, Boulder, CO 80309-0347, USA

\* equal contribution

\$ co-corresponding authors

Running head: PAR-TERRA as main contributor to mouse TERRA

Keywords: Telomeric RNA, PAR, TERRA interactome, Telomere

Viceconte et al

## **Abstract**

Telomeric repeat-containing RNA (TERRA) molecules play important roles at telomeres, from heterochromatin regulation to telomerase activity control. In human cells, TERRA is transcribed from subtelomeric promoters located on most chromosome ends and associates with telomeres. The origin of mouse TERRA molecules is however unclear, as transcription from the pseudoautosomal PAR locus was recently suggested to account for the vast majority of TERRA in embryonic stem cells (ESC). Here, we confirm the production of TERRA from both the chromosome 18q telomere and the PAR locus in mouse embryonic fibroblasts, ESC and various mouse cancer and immortalized cell lines, and we identify two novel sources of TERRA on mouse chromosome 2 and X. Using various approaches, we show that PAR-TERRA molecules account for the majority of TERRA transcripts, displaying an increase of 2 to 4 orders of magnitude compared to the telomeric 18q transcript. Finally, we present a SILAC-based pull-down screen revealing a large overlap between TERRA-interacting proteins in human and mouse cells, including PRC2 complex subunits, chromatin remodeling factors, DNA replication proteins, Aurora kinases, shelterin complex subunits, Bloom helicase, Coilin and paraspeckle proteins. Hence, despite originating from distinct genomic regions, mouse and human TERRA are likely to play similar functions in cells.

Viceconte et al

## Introduction

While telomeres have long been recognized as heterochromatic structures, they are active transcription sites. The first report of the existence of a transcriptional activity at telomeres dates back to 1989 with work from Rudenko and van der Ploeg in *Trypanosoma brucei* (Rudenko and van der Ploeg 1989). A few years later, transcription was also observed at telomeres of bird lampbrush chromosomes (Solovei et al. 1994).

The first evidences of telomeric transcription in human and mouse cells were provided nearly 15 years later (Azzalin et al. 2007; Schoeftner and Blasco 2008). Telomeric Repeat-containing RNA (TERRA) molecules were found to be transcribed from the C-rich telomeric DNA strand and to be detectable with (CCCTAA)<sub>n</sub> probes in RNA fluorescence *in situ* hybridization (FISH). From the beginning however, it emerged that mouse and human TERRA molecules may present some differences as their respective FISH patterns were quite distinct. In human interphase cells, TERRA-FISH signals were clearly detected at telomeres (Azzalin et al. 2007), while, in mouse cells, main TERRA-FISH signals appeared as only one or two large foci co-localizing with the q end of X/Y chromosomes (Schoeftner and Blasco 2008; Zhang et al. 2009; Schoeftner et al. 2009). On rare occasions (2-3 foci per nucleus), much smaller TERRA foci were also found to co-localize with mouse telomeres (López de Silanes et al. 2014).

In line with human TERRA detection at telomeres by RNA-FISH, many studies clearly demonstrated that human chromosome ends are transcribed from subtelomeric promoters that are located directly upstream of telomeric repeats (Nergadze et al. 2009; Deng et al. 2012; Porro et al. 2014a; Diman et al. 2016; Feretzaki and Lingner 2017; Sagie et al. 2017; Koskas et al. 2017; Feretzaki et al.

Viceconte et al

2019). A possible explanation to account for the distinct TERRA-FISH patterns between human and mouse cells was recently provided by RNA-seq experiments after TERRA capture from mouse ESC (Chu et al. 2017a). That study revealed the existence of a novel TERRA species, dubbed PAR-TERRA, originating from the pseudoautosomal locus of Xq/Yq subtelomeric regions, and accounting for more than 99% of all TERRA transcripts in mouse ESC. A few reads emanating from a small subset of telomeres were also identified in that study and included the previously reported Telo 18q TERRA (López de Silanes et al. 2014). The transcript emanating from Telo 18q was, however, about 4000-fold less abundant than PAR-TERRA in ESC (Chu et al. 2017a). Using the CHIRT technique, which combines ChIRP (Chromatin isolation by RNA purification) and CHART (Capture hybridization analysis of RNA targets), the same study also revealed that PAR-TERRA binds *in trans* to most chromosome ends and is the major  $(UUAGGG)_n$ -containing RNA species associating with mouse ESC telomeres. Altogether, these observations questioned the telomeric origin of mouse TERRA, as recently discussed (Diman and Decottignies 2018).

Here, we used several approaches, including the interrogation of publicly available RNA-seq datasets, to further investigate the genomic origin of mouse TERRA in mouse embryonic fibroblasts (MEF), ESC, normal tissues and cancer cells. Our results clearly indicated that, similarly to the previous observations in ESC, PAR-TERRA molecules account for the vast majority of TERRA transcripts in all mouse cell types. We also identified a new TERRA species transcribed from chromosome 2. However, both this newly identified Chr 2 TERRA and the Telo 18q TERRA appear to minimally contribute to the total cellular pool of  $(UUAGGG)_n$  repeats. To further probe the conservation of mouse and human TERRA, we

Viceconte et al

performed a TERRA interactome screen in mouse cells using a previously published SILAC-based pull-down approach (Scheibe et al. 2013). Cross-comparison of our results with an *in vivo* proteomic screen for TERRA-interacting proteins (Chu et al. 2017b) revealed that, despite distinct genomic origins and distinct nucleotide sequences, human and mouse TERRA are likely to interact with similar cellular proteins.

## Results

### **TERRA-FISH foci co-localize with telomeres and their intensity is proportional to telomere length in human cells**

In human primary and cancer cells, most TERRA-FISH foci co-localize with the telomeres of interphase nuclei (Azzalin et al. 2007; Arnoult et al 2012; Diman et al. 2016) (Fig. 1A). Consistent with previous northern blot experiments showing that the length of human TERRA molecules is proportional to telomere length (Yehezkel et al. 2008; Arnoult et al. 2012; Van Beneden et al. 2013), we confirmed, by TERRA-FISH, that fluorescence signal intensity mirrors telomeric DNA FISH intensity and, therefore, telomere length (Fig. 1B-D). The intensity of TERRA-FISH foci however does not reflect the total levels of *UUAGGG* repeats in human cells, as shown by slot-blot hybridization of total RNA isolated from HCA2 normal fibroblasts and LB37 lung cancer cells which, because of subtelomeric promoter hypomethylation (Nergadze et al. 2009; Diman et al. 2016; Sagie et al. 2017; Feretzaki et al. 2019; Le Berre et al. 2019), strongly up-regulate telomeric transcription (Fig. 1E-F). Taken together, these observations suggest that a substantial fraction of the human TERRA pool escapes detection by RNA-FISH and that the telomere-bound molecules may

Viceconte et al

likely be the only ones that are detected by RNA-FISH. The reasons for this are still unclear but may derive from the pre-extraction step of RNA-FISH protocol, which would wash out soluble TERRA and leave chromatin-bound TERRA as the only detectable species. It is also possible that soluble TERRA is not efficiently recognized by the RNA-FISH probe because folded into tight secondary structures such as G-quadruplexes (G4), or heavily bound to RNA-binding proteins.

**In mouse cells, most TERRA-FISH signals do not co-localize with telomeres and their intensity does not correlate with telomere length**

We next applied RNA-FISH to mouse cell lines with various telomere lengths in order to test whether signal intensity may similarly be proportional to telomere length. To do this, we first evaluated telomere length by TRF in the following cell lines: 3T3 (spontaneously immortalized MEF), L929 (spontaneously immortalized adult and adipose tissue-derived fibroblasts), RAW264.7 (Abelson murine leukemia virus transformed macrophages), J774A.1 (reticulum cell sarcoma), M1 (myeloblast cell line) and Neuro-2a (neuroblastoma) (Fig. 2A). We selected four mouse cell lines with either short (J774A.1, 4.5 kb), average (M1, 5.8 kb) or long (L929, 9.5 kb and Neuro-2a, 19.9 kb) telomeres for TERRA-FISH experiments and performed RNase A treatments as controls (Fig. 2B). Our data showed that the total TERRA-FISH signal intensity does not correlate with telomere length in mouse cells ( $p=0.53$ ) (Fig. 2C). As reported earlier, the RNA-FISH profiles are also very different from the ones obtained in human cells, with the presence of 2-3 prominent TERRA foci per nucleus (Fig. 2B and D). When TERRA-FISH signals were intense, like in J774A.1 and Neuro-2a cell lines, additional, but much smaller foci were detected using the same exposure conditions (Fig. 2B). To further investigate the co-localization of TERRA foci with

Viceconte et al

mouse telomeres, we combined RNA-FISH with immunofluorescence against the shelterin protein Terf1 in the Neuro-2a cell line. In sharp contrast with the consistent co-localization of TERRA-FISH signals with the other shelterin protein TRF2 in human cells (Fig. 1A), and in line with previous reports (López de Silanes et al. 2014; Chu et al. 2017b), only 2-4 small TERRA-FISH foci overlapped with Terf1 in mouse interphase nuclei (Fig. 2D). Conversely, in the same Neuro-2a cell line, Terf1 signals were co-localizing with telomeres in control DNA-FISH experiments using the same probe (Fig. 2E). The lack of correlation between telomeric DNA and TERRA-FISH signal intensity, together with the low frequency of co-localization events between TERRA and telomeres, suggest that telomere transcription may not be the main source of TERRA in mouse cells.

Similarly to human cells however, we found that *UUAGGG* levels, normalized to *ACTB* mRNA (Fig. 2F), do not correlate with *TTAGGG* levels in mouse cells (Fig. S1A-B and Fig. 2G,  $p=0.28$ ). This could be explained either by a non-telomeric origin of mouse TERRA or, like in human cells, by a distinct transcriptional activity at telomeres of various mouse cell lines.

Altogether, our experiments suggest very distinct regulations of TERRA in human and mouse cells (Fig. 2H), reinforcing the idea that the origin and molecular features of TERRA molecules may be very distinct in these two species.

### **TERRA molecules are produced from various mouse genomic loci**

Based on the above-described findings, we hypothesized that  $(UUAGGG)_n$ -containing sequences may result from the transcription of intrachromosomal *TTAGGG*-rich sequences. To identify these putative loci, we first scanned successive 2 kb-long regions of the mouse genome for the presence of at least 30 telomeric

Viceconte et al

repeats – not necessarily consecutive – including either pure (TTAGGG) or degenerate (TAAGGG, TGAGGG, TTGGGG, GTAGGG or TCAGGG) motifs (Fig. 3A). The arbitrary cut-off of 30 telomeric repeats was based on our preliminary analysis of the mouse genome, which suggested that the criteria was stringent enough to detect telomeric repeat-enriched regions above the background. A total of 105 intrachromosomal loci containing  $\geq 30$  telomeric motifs were identified. In a second step, six RNA-seq datasets with high sequencing coverage ( $>140$  million paired reads) from ES, forebrain, frontal lobe, B-cell lymphoma, MEF and MEL leukemia cells were selected for alignment with the identified intrachromosomal telomeric repeat-containing genomic loci. Using the same RNA-seq datasets, and whenever subtelomeric sequences were available, telomeric transcripts were also searched for by screening for subtelomeric reads directly adjacent to telomeres. As previously reported (López de Silanes et al. 2014), we detected few reads from the 18q subtelomere (Fig. 3B-C). A very limited number of reads appeared to similarly emanate from the 10q subtelomere, but not from any other chromosome ends with available sequence. Analysis of reads produced from the 105 identified telomeric repeat-containing intrachromosomal loci revealed the possible contribution of two loci to the production of  $(UUAGGG)_n$ -containing RNAs in all cell/tissue types: one on Chr 2 (3'UTR of *Polr3k* gene) and one on Chr X/Y (PAR locus-*Erdr1* gene) (Fig. 3B-C and Fig. S2). The *Tsix* locus on Chr X was identified as an additional intrachromosomal source of  $(UUAGGG)_n$ -containing transcripts in mouse ES cells exclusively (Fig. 3B-C and Fig. S2) and is known to be exclusively transcribed from the future active X chromosome in female cells. Although not detected in ES cells, a fourth intrachromosomal locus on Chr 8 may (Fig. 3B, red dotted line) also contribute to TERRA production in some cells, including brain cells. This region of Chr 8 was

Viceconte et al

however not considered as a strong contributor to overall TERRA production in mouse tissues as the total number of *TTAGGG* repeats was only 31 and the expression of the host gene, *Inpp4b*, is restricted to some tissue types only (not shown). The other candidate loci showed no or extremely low count number and were not selected as good candidates for contributing to TERRA production.

The pseudo-autosomal PAR locus was recently identified as the main source of TERRA in mouse ES cells and, through CHIRT analysis, PAR-TERRA was found to bind *in trans* to various genomic loci (Chu et al. 2017a). Careful analysis of the CHIRT data revealed that, while most PAR-TERRA and TERRA peaks indeed overlap in the mouse genome, PAR-TERRA is not detected at *Polr3k* 3'UTR locus, where TERRA peaks are readily detected (Chu et al. 2017a), suggesting that a TERRA molecule, unrelated to PAR-TERRA, is produced from- and binds to this locus. Together with our analysis, this shows that Chr 2 TERRA is a newly-identified source of mouse TERRA.

To compare the relative expression levels of the four major TERRA transcripts (Telo 18q, Chr 2, PAR and Tsix) in the selected RNA-seq datasets, we quantified the reads emanating from small unique regions located upstream of the  $(TTAGGG)_n$  repeats (Fig. 3C, light blue boxes). Strand specificity was taken into account to exclusively quantify the reads corresponding to  $(UUAGGG)_n$ -containing transcripts. The analysis revealed that PAR-TERRA was the most abundant species in ESC, MEF, B-cell lymphoma and MEL leukemia cell line, while similar levels of Chr 2 TERRA and PAR-TERRA were measured in brain tissues (Fig. 3D). Importantly, Telo 18q TERRA was barely detectable in all samples, suggesting that the contribution of this locus to mouse TERRA is extremely low (Fig. 3D).

### Candidate mouse TERRA molecules contain *UUAGGG* repeats

To validate the production of  $(UUAGGG)_n$ -containing sequences from the four genomic candidate loci, we designed primers located upstream of the *TTAGGG* repeats of each locus. For Telo 18q, primers are located within the possible 3'UTR of *LOC108168395* downstream of *Tmx3* gene, within the last 20 kb before 18q telomeric repeats (Fig. 4A). For Chr 2, primers are in the 3'UTR region of the *Polr3k* gene. Primers for PAR and *Tsix* loci are located within introns (Fig. 4A). Primer efficiency was tested (Fig. S3) and two distinct reverse transcriptions (RT) were performed: one with random hexamers and one with  $(CCCTAA)_5$  primers (Telo primers). If the candidate transcript contains long stretches of *UUAGGG* repeats, we expect a strong enrichment of the corresponding cDNA when RT is performed with Telo primers. If, on the other hand, the transcript contains fewer and/or interspaced *UUAGGG* repeats, the enrichment should be weaker. Finally, if the transcript is completely devoid of *UUAGGG* repeats, we expect the efficiency of the RT with Telo primers to be strongly reduced and the corresponding cDNA to be virtually absent.

Total RNA was isolated from J774A.1 and ES cells. Consistent with the respective abundance of consecutive *TTAGGG* repeats at the tested loci (Fig. S2), we observed either a strong (Telo 18q, Chr 2) or a moderate (PAR, *Tsix*) enrichment of cDNA species when Telo primers were used for the RT instead of random primers (Fig. 4B). Importantly, the levels of *Sod1*, used as negative control for  $(UUAGGG)_n$ -containing mRNA, were at background levels when Telo primers were used for RT (Fig. 4B). Altogether, these data confirm the presence of *UUAGGG* repeats in Telo 18q, Chr 2, PAR and *Tsix* TERRA species.

## **PAR-TERRA is the main source of *UUAGGG*-rich RNA molecules in ESC, MEF and immortalized mouse cell lines**

As the candidate TERRA loci likely produce RNA molecules with distinct *UUAGGG* repeat contents, we next sought to determine which of those TERRA loci is the main contributor to the pool of *UUAGGG* repeats in mouse cells. To do this, we performed RT-qPCR analyses on random primer-synthesized cDNA in MEF, ES cells and in the six mouse cell lines, and compared these measurements to *UUAGGG* levels measured by slot-blots. Our RT-qPCR analyses confirmed that PAR-TERRA levels are much higher than the other tested transcripts in all cell lines (Fig. 5A). The strong correlation ( $R=0.967$ ,  $p=0.0001$ ) that we obtained between PAR-TERRA levels and total *UUAGGG* levels, measured by slot-blot, further confirmed that PAR-TERRA is the main source of *UUAGGG* repeats in mouse cells (Fig. 5B-C). Conversely, the levels of either Telo 18q ( $R=-0.49$ ,  $p=0.223$ ) or Chr 2 ( $R=-0.25$ ,  $p=0.554$ ) did not correlate with total *UUAGGG* levels (Fig. 5C).

A short stretch of *TTAGGG* repeats located between the *LOC108168395* locus and the 18q telomere (Fig. S4A) was previously proposed as a possible source for TERRA (López de Silanes et al. 2014). However, our data failed to reveal any correlation between this subtelomeric 18q transcript and the relative *UUAGGG* levels in the tested cell lines (Fig. S4B).

## **Telomere deprotection through Terf2 depletion does not up-regulate telomere transcription**

Previous work in human cells revealed that TERRA transcription is up-regulated from TRF2-depleted telomeres (Caslini et al. 2009; Porro et al. 2014a and 2014b). To test the hypothesis that telomere deprotection may be similarly

Viceconte et al

associated with enhanced transcription of mouse telomeres, we extracted RNA from SV40-immortalized MEF conditionally knocked-out for *Terf2* (Fig. 6A). We also designed new pairs of primers for 5q and 11q chromosome ends located at 15 bp and 10 bp from the telomeric tract, respectively (Fig. 6B). Primers specifically amplified genomic DNA from J774A.1 cells (Fig. 6B and Fig. S3), however no amplification product was obtained with J774A.1 cDNA as template (Fig. 6B). Similarly, no product was obtained for 5q or 11q after 40 cycles of PCR on cDNA from *Terf2* F/+ or tamoxifen-treated *Terf2* F/F MEF, although Telo primers were used for the RT (Fig. 6C). In addition, 18q TERRA levels were not up-regulated upon *Terf2* depletion (Fig. 6C), suggesting, once again, a very distinct regulation of telomere transcription in human and mouse cells.

### Comparison of human and mouse TERRA interactomes

Despite the distinct chromosomal origin of mouse and human TERRA molecules, important cellular functions, including heterochromatin regulation (Bettin et al. 2019), appear to be exerted by TERRA in the two organisms. To better evaluate the extent of conservation of TERRA-associated functions in mouse and human, we performed a SILAC-based *in vitro* purification with a  $(UUAGGG)_8$  biotinylated TERRA-like oligonucleotide as a bait to identify mouse TERRA-interacting proteins (Scheibe et al. 2013). Even though PAR-TERRA appears to be the most abundant TERRA species in mouse cells, and does not consist of long uninterrupted  $(UUAGGG)_n$  sequences, we believe that the use of a  $(UUAGGG)_8$  biotinylated probe is appropriate, as a number of PAR-TERRA/protein interactions likely involve *UUAGGG* repeats (Chu et al. 2017b).

Viceconte et al

We incubated mouse R1/E ES cell extracts with a (UUAGGG)<sub>8</sub> biotinylated TERRA probe or with a biotinylated (GUGUGA)<sub>8</sub> probe as a control for specificity. A total of 307 candidate proteins were identified that showed at least a 4-fold enrichment over the control probe (Fig. S5A-B) and the number increased to 581 with a threshold of 2-fold (Table S1). To help identify functional classes of mouse TERRA-interacting proteins, we compared our dataset of 307 proteins with the one previously obtained through iDRiP (identification of Direct RNA interacting Proteins) in mouse ESC (134 proteins) (Chu et al. 2017b). Thirty candidates were common to both datasets (yellow rectangles in Fig. 7A) and additional 98 proteins, from either screen (candidates from the SILAC screen are shown in bold), belonged to similar protein families, leading to a total of 128 “common” candidate proteins (Fig. 7A). A STRING analysis was performed on the 128 proteins and revealed possible roles for mouse TERRA in chromatin remodeling, RNA metabolism, DNA replication, ribosome biogenesis or mitosis, with an enrichment in proteins related to centromeres, chromatid cohesion, mitotic spindle or cytokinesis (Fig. 7A). Consistent with the previously reported interaction between mouse or human TERRA and the PRC2 complex (Bettin et al. 2019), our pull-down approach recovered the mouse Eed, Suz12, Jarid2 and Mtf2 subunits, with Mtf2 being shared with the iDRiP screen (Fig. 7A). The Terf1 subunit of the shelterin complex was also recovered in both screens. Similarly, we confirmed the interaction of TERRA with many mouse proteins involved in RNA metabolism, including proteins from the hnRNP (heterogeneous nuclear ribonucleoprotein) family, as initially demonstrated by López de Silanes et al. (2010). Other interacting proteins included Orc1, Orc2 and Blm – with roles in DNA replication and DNA damage repair –, the Cajal body protein Coilin and a series of proteins with various functions in mitosis, like Aurora kinases b and c, the inner

Viceconte et al

centromere protein Incenp, Sgol2, involved in sister chromatid cohesion regulation, Tpx2, with a key role in spindle assembly during mitosis and Mki67, that associates with mitotic chromosomes (Fig. 7A). As the *in vivo* iDRiP screen mostly monitored the interactions between PAR-TERRA and mouse ES cell proteins, the consistency that we observed with the SILAC pull-down screen agrees with interactions happening through the *UUAGGG* repeats. To further validate our SILAC pull-down, we performed RNA immunoprecipitation (RIP) experiments using extracts from immortalized MEF and antibodies against two identified candidates, Suz12 and Blm. Because TERRA interacts with human TRF2 both *in vitro* and in cells (Deng et al. 2009; Lee et al. 2018), we also performed RIP using anti-Terf2 antibodies even though the protein was not recovered in our SILAC-pull down experiments. In line with previous RIP experiments in mouse iPS cells (Marión et al. 2019), PAR-TERRA co-immunoprecipitated with the anti-Suz12 antibody (Fig. 7B). Further validating our pull-downs, PAR-TERRA was also recovered using antibodies against Blm (Fig. 7B). Finally, PAR-TERRA was also detected in fractions isolated using the anti-Terf2 antibodies. In all pull-downs, Telo 18q RNA was barely detectable (Fig. 7C). Together, these results confirm the validity of our method and clearly establish that PAR-TERRA and human TERRA bind to similar sets of proteins. The differences in the amounts of PAR-TERRA measured in the three different RIP do not necessarily reflect the strength of PAR-TERRA interaction with Suz12, Blm and Terf2 due to different IP efficiencies. The fact that Terf2 was not detected in the SILAC screen might indicate that PAR-TERRA/Terf2 interactions are transient or not strong enough to be detected without cross-linking.

We next searched for common TERRA-interacting proteins in human and mouse cells by comparing our two SILAC-based pull-down screens, using a

Viceconte et al

threshold of  $\geq 2$ -fold enrichment for both screens (Table S1 and S2). Our analyses revealed that 188 out of the 454 (41%) human TERRA-interacting protein candidates are also present in the mouse SILAC screen while 188 out of the 581 mouse candidates (32%) are detected in the human screen (Fig. 8A and Fig. S6A-B). The common interacting proteins are involved in RNA metabolism, DNA replication, mitosis (Aurora kinases a/b/c, Mcm5, Rfc3/4, ...) or chromatin organization (including PRC2 complex) (Fig. 8B) and also include components of the CDC5L complex, the Cajal bodies or the paraspeckle subnuclear bodies.

Based on the large overlap of interacting proteins and associated pathways, it is conceivable that many TERRA-associated functions are shared between human and mouse cells.

## Discussion

From the first reports of TERRA in human and mouse cells, differences between the two species rapidly emerged as very distinct TERRA-FISH patterns were observed. While human TERRA was clearly detected at telomeres, co-localization of TERRA with mouse telomeres was rarely observed by RNA-FISH in interphase cells (López de Silanes et al. 2014). More recently, CHIRT experiments revealed that a TERRA-like transcript, produced from the pseudo-autosomal PAR locus of mouse ES cells, was able to bind, *in trans*, to most mouse telomeres (Chu et al. 2007a). Here, we confirmed the distinct origin of TERRA in human and mouse cells. Using the same probes and tools, and together with our previous studies (Azzalin et al. 2007; Arnoult et al. 2012; Van Beneden et al. 2013), we showed that TERRA results from the transcription of telomeres in human cells, while the vast

Viceconte et al

majority of the *UUAGGG* repeats comes from the PAR locus in all the mouse cell lines that we tested. PAR-TERRA corresponds to an intronic region of the *Erd1* gene on the X/Y chromosome, and thus reinforces the idea that spliced introns can be precursors of non-coding RNAs, including microRNAs or long non-coding RNAs (Hesselberth 2013). Our RNA-FISH experiments using a telomeric probe in mouse cells revealed a good correlation between TERRA-FISH intensity and the level of PAR-TERRA molecules measured by RT-qPCR. This contrasts with human cells in which TERRA-FISH intensity does not correlate with the total TERRA levels as FISH signals are only visible at telomeres. We propose that telomere-bound human TERRA molecules can be detected by RNA-FISH while the other TERRA molecules escape detection, possibly because of G4 structures forming along the long (*UUAGGG*)<sub>n</sub> tracts as previously shown (Biffi et al. 2012; Hirashima and Seimiya 2015) or because human TERRA, when not bound to telomeres, is lost during the pre-extraction step of the RNA-FISH protocol. PAR-TERRA, on the other hand, in light of its very distinct nucleotide sequence with interspaced telomeric repeats (Fig. S2), may not fold into G4 structures and/or may interact more strongly with chromatin. In this respect, CHIRT experiments in mouse ES cells clearly revealed the presence of PAR-TERRA binding sites throughout the genome and at the PAR locus (Chu et al. 2017a and 2017b).

The *in silico* search for intrachromosomal mouse candidate TERRA loci unveiled a new locus on chromosome 2, likely located within the 3'UTR region of the *Polr3k* gene. Long non-coding RNA functions have previously been proposed for 3'UTR regions of genes, whether as whole molecules or as cleaved fragments (Mayr 2017). Based on our analyses of RNA-seq datasets, this Chr 2 locus appears to contribute significantly to the pool of TERRA in some mouse tissues, including the forebrain and

Viceconte et al

the frontal lobe. Interestingly, a careful analysis of the CHIRT data from Chu et al. (2017a) revealed the presence of a TERRA peak, but not a PAR-TERRA peak, at the *Polr3k* locus, that likely corresponds to the binding of Chr 2 TERRA molecules. This, in turn, suggests that Chr 2 TERRA binds to the locus where it is produced. Future work is needed to address the possible functions of Chr 2 TERRA in mouse. We also identified an additional intrachromosomal source of  $(UUAGGG)_n$ -containing transcripts emanating from the *Tsix* locus – located within the X-inactivation center (Xic) – in mouse ES cells exclusively. Interactions between the Xic and the PAR loci were previously shown to be required for the initiation of X-chromosome inactivation in ES cells (Chu et al. 2017a). Interestingly, the reported CHIRT experiments suggested that the *Tsix* locus was bound not only by PAR-TERRA, but also by  $(UUAGGG)_n$ -containing transcripts not related to PAR-TERRA (Chu et al. 2017a). We therefore anticipate that, in ES cells, the  $(UUAGGG)_n$ -containing transcripts emanating from the *Tsix* locus may interact with PAR-TERRA to promote Xic:PAR pairing and X-chromosome inactivation.

The reason why mouse telomeres are mostly silent and Terf2 depletion is unable to de-repress telomere transcription is not clear. This may be related to the heterochromatic nature of mouse telomeres and subtelomeres that, unlike human telomeres, display high enrichment of H3K9me3 and H4K20me3 marks, together with abundant binding of Heterochromatin protein 1 (Blasco 2007; Rosenfeld et al. 2009). Alternatively, the lack of telomeric transcription may result from the absence, in the mouse genome, of dedicated subtelomeric promoters like in human cells. In line with this, production of the Telo 18q TERRA appears to result from the transcription of an upstream gene that initiates about 22 kb upstream of the first 18q

Viceconte et al

telomeric repeats, and not from a subtelomeric promoter directly driving the transcription of telomeric repeats.

Despite their distinct genomic origins, increasing evidence suggests that TERRA molecules may play similar roles in human and mouse cells (Bettin et al. 2019). This sketches a fascinating picture where  $(UUAGGG)_n$ -containing RNAs produced from different genomic loci may exert similar functions through their binding *in trans* to various genomic loci, including telomeres. Our comparative study of TERRA-interacting proteins, identified through SILAC-based pull-down screens in both species, reinforced the idea of common functions shared by human and mouse TERRA. Among the common interacting proteins, we found subunits of the PRC2 complex, Aurora kinases, proteins involved in DNA replication and repair, rRNA metabolism, pre-mRNA processing and Pol II-dependent transcription. Interestingly, interactions between Aurora kinase B (AURKB) and centromeric RNAs in mouse ES cells were previously reported to be involved in the recruitment of telomerase to chromosome ends in S phase (Mallm and Rippe 2015). AURKB was proposed to bind to centromeric RNAs before interacting with the telomerase complex to enhance its activity. It is therefore tempting to speculate that TERRA may similarly be implicated, through its interaction with AURKB, into the activation and/or recruitment of telomerase to telomeres, in line with previous observations in yeast (Cusanelli et al. 2013; Moravec et al. 2016). A more recent study reported on the localization of AurkB at mouse ESC telomeres where the kinase modulates Terf1 affinity for telomeres and participates in telomeric integrity (Chan et al. 2017). The interaction between AURKB and TERRA may participate in overall telomere protection through the modulation of shelterin affinity for telomeres. Common pulled-down proteins also included components of the paraspeckle subnuclear bodies and the Dazap1 RNA-

Viceconte et al

binding protein. Interaction between TERRA and the human NONO paraspeckle component was recently confirmed and shown to suppress RNA:DNA hybrid-induced telomere instability; the same study confirmed the interaction between TERRA and the DAZAP1 RNA-binding protein (Petti et al. 2019). HNRNP proteins and the BLM helicase, previously identified as human TERRA interactors (Deng et al. 2009; Flynn et al. 2011; Petti et al. 2019), were also recovered in the screens and we validated the interaction between PAR-TERRA and mouse Blm by RIP. Interestingly, two classes of proteins appeared to be differentially recovered in the screens. The first one comprises centromeric proteins that were abundantly recovered in the mouse pull-down screen, suggesting a function for mouse TERRA in centromere assembly/stability. The second group includes proteins from the exosome complex (EXOSC) that were enriched in the human pull-down screen, but not in the mouse screen, suggesting possible distinct degradation machineries in human and mouse.

As we found a good overlap between our pull-down screen using mouse cell extracts and the *in vivo* iDRiP screen in ES cells (Chu et al. 2017b), in which PAR-TERRA was likely the most abundant TERRA species authors were looking at, these results support the idea that, whether they originate from the transcription of telomeres or the PAR locus, these *UUAGGG*-rich non-coding RNA molecules interact with the same protein sets and are likely to serve similar functions in the cell. Atrx, previously reported to interact with mouse TERRA (Chu et al. 2017b) was, however, not recovered in our pull-down screens using the  $(UUAGGG)_8$  probe. This may be related to the observation that, although a 83-nt *in vitro* synthesized TERRA RNA probe was efficiently shifted by Atrx in electrophoretic mobility shift assays, a 30-nt TERRA was shifted to a much lesser degree (Chu et al. 2017b). This suggests that the oligonucleotide used in our studies might not be long enough to allow for

Viceconte et al

efficient interaction with Atrx. Our biotinylated (UUAGGG)<sub>8</sub> probe similarly failed to pull-down other mouse iDRiP candidates, including Rtel1, Rpa1/2, Ctc1, Stn1 or Pml (Chu et al. 2017b), with important functions in telomere biology. This may be explained either by the size of the probe, as explained above, or by the need for a chromatin context that is not recapitulated in the *in vitro* pull-down assay. Alternatively, interactions may involve non-(UUAGGG)<sub>n</sub> sequences.

Although the good overlap with the *in vivo* iDRiP screen in mouse ES cells indicates that our pull-down screen efficiently recovered TERRA-interacting proteins, future experiments will be required to properly address TERRA functions in human and mouse cells using tools to specifically deplete TERRA or PAR-TERRA transcripts.

## Materials and Methods

**Cell culture.** The following mouse cell lines were used in this study: 3T3 spontaneously immortalized embryonic fibroblasts (ATCC), L929 spontaneously immortalized adult and adipose tissue-derived fibroblasts (ECACC), M1 myeloblast cell line (ECACC), RAW264.7 Abelson murine leukemia virus transformed macrophages (ATCC), Neuro-2a neuroblastoma cell line (ATCC) and J774A.1 reticulum cell sarcoma cell line (monocytes/macrophages) (ATCC) and were kindly provided by Thomas Michiels (de Duve Institute, Brussels, Belgium). Mouse embryonic fibroblasts (MEF) were prepared from CD1 mice using standard protocols and were kindly provided by Frédéric Lemaigre (de Duve Institute, Brussels, Belgium). ES cells were kindly provided by Olivier De Backer (Université de Namur, Namur, Belgium). The human cells used in this study were previously described:

Viceconte et al

HCA2 (Arnoult et al. 2010) and HFF2 (Mattiussi et al. 2012) human foreskin fibroblasts, HeLa cervix cancer cell line (ATCC), LB37 non-small cell lung cancer cell line (Tilman et al. 2009) and SW39 SV40T-immortalized fetal lung fibroblasts (Tilman et al. 2009, kindly provided by W. Wright, UT Southwestern Medical Center, Dallas, USA). Cells were cultured in EMEM (L929, Neuro-2a, SW39, HFF2, HCA2, LB37), DMEM (3T3, J774A.1, RAW264.7, MEF, HeLa) or RPMI (M1). All media were from Gibco and enriched with 1x Penicillin-Streptomycin (Gibco) and 10% FBS (Gibco). Cells were grown at 37°C under 5% CO<sub>2</sub>. CreERT2 *Terf2* F/+ and *Terf2* F/F SV40T-immortalized MEF were a kind gift from Eros Lazzerini Denchi (The Scripps Research Institute, La Jolla, USA) and were previously described (Okamoto et al. 2013). For RNA-IP experiments, we used 2575i immortalized *mTert* +/+ MEF, kindly provided by Lea Harrington (Institute for Research in Immunology and Cancer, Université de Montréal, Canada). Cells were cultured in high glucose DMEM, GlutaMAX (Gibco) supplemented with 10% tetracycline-free fetal bovine serum (Pan BioTech) and 100 U/ml penicillin-streptomycin (Gibco). For *Terf2* knock-out, cells were treated with 0.6 µM 4-hydroxytamoxifen (OHT, Sigma-Aldrich) and collected 48 or 72 hours after treatment. *Terf2* deletion was confirmed by western blotting using a rabbit monoclonal anti-*Terf2* (Novus Biologicals, NB110-57130, 1:2000 dilution) and a mouse monoclonal anti-beta Actin (Abcam, ab8224, 1:5000 dilution) to control for loading. Secondary antibodies were HRP-conjugated goat anti-mouse and goat anti-rabbit IgGs (Bethyl Laboratories, A90-116P and A120-101P, 1:2000 dilution). Signal detection was performed using the ECL detection reagents (GE Healthcare) and a FluorChem HD2 imaging apparatus (Alpha Innotech).

**RNA extraction and RT-qPCR.** Total RNA was extracted from mouse cell lines

Viceconte et al

using TriPure (Sigma-Aldrich) as previously described (Arnoult et al. 2012). For CreERT2 *Terf2* F/+ and *Terf2* F/F SV40T-immortalized MEF, total RNA was extracted using the NucleoSpin RNA kit (Macherey-Nagel) by following the manufacturers' instructions. RNA concentrations were determined by Nanodrop. cDNA synthesis, and qPCR were performed as described previously (Arnoult et al. 2012) using Superscript III (Thermo Fisher Scientific), KAPA SYBR FAST (Sigma-Aldrich) and primers described in Table S3. For cDNA synthesis using (CCCTAA)<sub>5</sub> Telo primers, the mGAPDH-RT primer was added at the same concentration of 1  $\mu$ M.

**Genomic DNA extraction.** Genomic DNA was extracted from cell lines by overnight digestion at 45°C with 100  $\mu$ g/ml proteinase K and 50  $\mu$ M CaCl<sub>2</sub> in 600  $\mu$ l Lysis buffer (10 mM Tris-HCl, 10 mM EDTA, 5% SDS, pH 8.0, Sigma-Aldrich) followed by DNA purification with phenol-chloroform-isoamyl alcohol (25:24:1, Sigma-Aldrich) and ethanol precipitation in 0.3 M sodium acetate, pH 5.2 (Sigma-Aldrich).

**TRF.** After digestion of 5-10  $\mu$ g of genomic DNA with *Hinfl* and *RsaI* (Sigma-Aldrich), Telomere Restriction Fragment (TRF) analysis was performed by Southern blotting using a radioactive (CCCTAA)<sub>4</sub> probe 5' end-labeled with [ $\gamma$ -<sup>32</sup>P] ATP and polynucleotide kinase as described previously (Viceconte et al. 2017). Telomere length was then evaluated with the publicly available Telotool software (Göhring et al. 2014).

**RNA/DNA slot-blots.** Slot-blots were performed as described previously (Arnoult et al. 2012) using the radioactive (CCCTAA)<sub>4</sub> probe and the indicated amounts of RNA/DNA. To normalize TERRA measurements, a radioactive probe against mouse  $\beta$ -actin transcript was prepared as follows. A 0.8 kb fragment of mouse *ACTB* coding

Viceconte et al

sequence was amplified from MEF cDNA using 5'-ATATCGCTGCGCTGGTCGTC and 5'-ATGCCACAGGATTCCATACC primers. The probe was obtained by incubation for 1h at 37°C of 100 ng of boiled PCR product with 40  $\mu$ Ci [ $\alpha$ -<sup>32</sup>P] dCTP (Perkin Elmer), 1.5 mM dATP, dGTP and dTTP (Thermo Fisher Scientific), 5  $\mu$ M random hexamers (Thermo Fisher Scientific) and 5 U of large Klenow fragment (NEB). After purification of the probe with Microspin G-25 columns (GE Healthcare), hybridization was performed in UltraHyb solution (Thermo Fisher Scientific) and membranes were washed first in 2x SSC/0.1% SDS and then in 0.2x SSC/0.1% SDS. Signals were quantified using a Phosphorimager Typhoon Trio (GE Healthcare). In some experiments, TERRA slot-bot membranes were stripped by incubation for 30 min in 0.1 % SDS before hybridization with the  $\beta$ -actin probe.

**TERRA-FISH, Telo-FISH and immunofluorescence.** TERRA-FISH was performed as described previously (Arnoult et al. 2012) using the C-rich (FAM)CCCTAaCcCTaaCcCTAACCCCTaaCCCTaaCCCTaaCCCTaA(FAM) probe where small letters indicate LNA-modified bases (Exiqon). For control RNase treatments, cells were permeabilized in buffer lacking Ribonucleoside Vanadyl Complex (RVC) followed by fixation and incubation for 1 h at 37°C with 1 mg/ml RNase A (Sigma-Aldrich) in PBS. When combined with immunofluorescence (IF) against either Terf1 (mouse cells) or TRF2 (human cells), TERRA FISH was performed first, and, after removal of the probe, IF was performed as previously described (Arnoult et al. 2012) using the following antibodies: Terf1 (1:3000, #1448/1449, laboratory of Titia de Lange) and TRF2 (1:500, IMGENEX). To visualize telomeres by Telo-FISH, in combination with IF against Terf1, we followed the protocol described in Diman et al (2016). TERRA- and Telo-FISH images were captured with a confocal microscope (Cell Observer Spinning Disc, Zeiss) equipped

Viceconte et al

with 100x objective and AxioVision software. TERRA and telomeric foci were quantified using the ImageJ software (National Institutes of Health). After maximum intensity projection of the Z-stacks, each focus was measured using local maxima determination with a noise tolerance of 10. Total intensity of TERRA or telomeric foci was calculated as the sum of all foci for each nucleus. At least 25 nuclei from at least 7 different images were quantified for each condition after background removal.

**Bioinformatic analysis of publicly available RNA-seq experiments.** A Perl script was written to scan successive 2 kb-long sequences from the mouse genome (GRCm38) to identify regions containing at least 30 telomeric repeats including either pure (TTAGGG) or degenerate (TAAGGG, TGAGGG, TTGGGG, GTAGGG or TCAGGG) repeats. Publicly available mouse RNA-seq experiments obtained from NCBI's Sequence Read Archive (SRA) were downloaded. ESC, forebrain, frontal lobe, B-cell lymphoma cells, MEF and MEL leukemia cell line datasets correspond, respectively, to accessions SRR530639-SRR530640, SRR3192667-SRR3192668, SRR567478-SRR567479, SRR3192560-SRR3192561, SRR5870458-SRR5870452 and SRR5048192-SRR5048193. These are all strand-specific RNA-seq datasets with a high sequencing coverage (> 140 millions paired reads). Fastq files were all processed using the same pipeline. Read quality control was performed using FastQC v0.11.8 (Andrews 2010) and low-quality sequences were removed using Trimmomatic v0.38 (Bolger et al. 2014). Filtered reads were aligned on GRCm38 mouse genome using HISAT2 v2-2.1.0 (Kim et al. 2015). Gene expression levels were evaluated using featureCounts v2.0.0 (Liao et al. 2014) and *Mus\_musculus.GRCm38.94.gtf*. Regions corresponding to Chr 2, PAR, Telo 18q and Tsix PCR products were artificially introduced in the gtf file as fictive genes to allow their quantification. Strand specificity was taken into account and only unambiguously

Viceconte et al

mapped reads were considered for quantification. Bam files were further converted to tdf files using igvtools v2.3.98 (Robinson et al. 2011) before viewing alignment data in IGV browser.

**SILAC labeling and nuclear extract preparation.** Mouse R1/E embryonic stem cells (ATCC) were SILAC-labeled in DMEM (-Arg, -Lys) medium containing 42 mg/L  $^{13}\text{C}_6^{15}\text{N}_4$  L-arginine (Euriso-Top) and 73 mg/L  $^{13}\text{C}_6^{15}\text{N}_2$  L-lysine (Euriso-Top) or the corresponding concentration of unlabeled amino acids (Sigma-Aldrich). Medium was supplemented with 10% dialyzed fetal bovine serum (PAA), 1x non-essential amino acids (Gibco), 50  $\mu\text{M}$  2-mercaptoethanol (Gibco), 3  $\mu\text{M}$  CT-99021 (Biomol), 1  $\mu\text{M}$  PD-0325901 (Biomol), 1 mM sodium pyruvate, 100 U/ml LIF (Millipore), 100 U/ml penicillin and 100  $\mu\text{g/ml}$  streptomycin (Gibco). Nuclear extracts were prepared essentially as previously described (Scheibe et al. 2013) and were shock frozen in liquid nitrogen and stored at  $-80^\circ\text{C}$  until use.

**RNA pull-downs, mass spectrometry and data analysis.** RNA pull-downs with SILAC nuclear extracts were performed as previously described (Scheibe et al. 2013). Pull-downs were separated on a 4%–12% NOVEX gradient SDS gel (Life Technologies) for 50 min at 180 V in 1x MOPS buffer (Life Technologies) and processed into peptides as previously described (Scheibe et al. 2013). For mass spectrometry analyses, peptides were separated on a 20 cm self-packed column with 75- $\mu\text{m}$  diameter filled with ReproSil-Pur 120  $\text{C}_{18}$ -aq (Dr. Maisch GmbH) mounted to an EASY HPLC 1000 (Thermo Fisher) and sprayed online into an Q Exactive Plus mass spectrometer (Thermo Fisher). We used a 120-min gradient from 2% to 60% acetonitrile in formic acid at a flow of 225 nL/min. The mass spectrometer was operated with a top 10 MS/MS data-dependent acquisition scheme per MS full scan.

Viceconte et al

Mass spectrometry raw data were searched using the Andromeda search engine (Cox et al. 2011) integrated into MaxQuant suite 1.5.2.8 (Cox and Mann 2008) using the ENSEMBL Mus\_musculus.GRCm38 protein database (57,751 entries). The human data (Scheibe et al. 2013) was reanalyzed with MaxQuant suite 1.5.2.8 using the ENSEMBL Human GRCh38 protein database (102,915 entries). In both analyses, carbamidomethylation at cysteine was set as fixed modification while methionine oxidation and protein N-acetylation were considered as variable modifications. Match between run option was activated. Prior to bioinformatic analyses, reverse hits, proteins only identified by site, protein groups based on 1 unique peptide and known contaminants were removed. Further filtering and graphical representation was done in an R framework incorporating ggplot2 package (R core team 2017; Wickham 2016). To determine murine/human orthologues, we queried the Compara database provided within ENSEMBL (Zerbino et al. 2018) using the biomaRt software package (Durinck et al. 2005; Durinck et al. 2009). Further, we queried the gene ontology (GO) database (Ashburner et al. 2000; The Gene Ontology Consortium, 2017) for functional analysis. We tested our enriched proteins against all detected proteins using overrepresentation tests ( $FDR < 0.05$ ) as implemented in the clusterProfiler package (Yu et al. 2012). Finally, functional associations among enriched proteins were highlighted using STRING (version 10.5) (Szklarczyk et al. 2017).

**RNA immunoprecipitation (RIP).** Approximately  $10^7$  2575i MEF were cross-linked with 1% formaldehyde (Sigma-Aldrich) for 15 min at RT, before quenching with 125 mM glycine for 5 min at RT. After two washed with ice-cold PBS, cells were scraped from dishes and centrifuged at 1000  $xg$  for 5 min at RT. Pellets were resuspended in

Viceconte et al

2 ml PBS, 2 ml Nuclear isolation buffer (40 mM Tris-HCl, pH 7.5, 20 mM MgCl<sub>2</sub>, 4% Triton X-100, 1.28 M sucrose) and 6 ml water before incubation on ice for 20 min with frequent mixing. After centrifugation at 2500 *xg* for 15 min at 4°C, pellets were resuspended in 1 ml of RIP buffer (150 mM KCl, 25 mM Tris-HCl, pH 7.5, 5 mM EDTA, 0.5% Nonidet-P40, 1 mM DTT, 100 U/ml RNase Out (Thermo Fisher Scientific) and EDTA-free Protease Inhibitor Cocktail (Sigma-Aldrich)). Nuclear extracts were sonicated twice using a Bioruptor apparatus (Diagenode) before centrifugation at maximum speed for 10 min at 4°C. Sonicated extracts were pre-cleared by incubation with Dynabeads™ Protein A (Thermo Fisher Scientific), pre-blocked with BSA and *E. coli* tRNA, on a rotating wheel for 1h at 4°C. Pre-cleared lysates were then diluted in RIP buffer to a final concentration of 1 mg/ml and 1 ml was used for each IP with 2 µg of the respective antibodies (rabbit anti-Blm (Bethyl Laboratories A300-110A), rabbit anti-Suz12 (Abcam ab12073) and rabbit anti-Terf2 (Novus Biologicals NB110-57130)). After 3h incubation at 4°C on a rotating wheel, pre-blocked Dynabeads™ Protein A was added to the samples for an overnight incubation at 4°C. Beads were washed 3 times with RIP buffer and once with high salt RIP buffer (300 mM KCl) before elution with 25 mM Tris-HCl, pH 7.5, 5 mM EDTA, 0.5% SDS with 0.1 µg Proteinase K for 45 min at 45°C, and then overnight at 65°C to revert cross-link. Eluted RNA was purified using Nucleospin RNA kit (Macherey-Nagel) and cDNA was synthesized using Superscript IV (Thermo Fisher Scientific) and 10 µM Telo primers as described above.

## Acknowledgments

We are grateful to Titia de Lange (The Rockefeller University, New-York, USA) for Terf1 antibody, to Thomas Michiels (de Duve Institute, UCLouvain, Belgium) for

Viceconte et al

mouse cell lines, to Frédéric Lemaigre (de Duve Institute, Brussels, Belgium) for MEF, to Eros Lazzerini Denchi (The Scripps Research Institute, La Jolla, USA) for CreERT2 *Terf2* F/+ and F/F MEF and to Olivier De Backer (Université de Namur, Namur, Belgium) for the kind gift of mouse ES cells. We are most grateful to Amandine Van Beneden for help with RNA slot-blot experiments. N.V. and A.L. were supported by a grant from the D.G. Higher Education and Scientific Research of the French Community of Belgium (Actions de Recherche Concertées). A.D. is a recipient from the Fonds National de la Recherche Scientifique (FNRS). Work in the Azzalin laboratory was supported by the European Molecular Biology Organization (IG3576) and the Fundação para a Ciência e a Tecnologia (IF/01269/2015; PDTC/MED-ONC/28282/2017; PDTC/BIA-MOL/29352/2017). We are most grateful to the de Duve Institute for constant support.

Viceconte et al

## References

Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

Arnoult N, Schluth-Bolard C, Letessier A, Drascovic I, Bouarich-Bourimi R, Campisi J, Kim S-H, Boussouar A, Ottaviani A, Magdinier F et al. 2010. Replication timing of human telomeres is chromosome arm-specific, influenced by subtelomeric structures and connected to nuclear localization. *PLoS Genet* **6**: e1000920. doi: 10.1371/journal.pgen.1000920

Arnoult N, Van Beneden A, Decottignies A. 2012. Telomere length regulates TERRA levels through increased trimethylation of telomeric H3K9 and HP1 $\alpha$ . *Nat Struct Mol Biol* **19**: 948-956. doi: 10.1038/nsmb.2364

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT et al. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**: 25-29. doi: 10.1038/75556

Azzalin CM, Reichenbach P, Khoriantuli L, Giulotto E, Lingner E. 2007. Telomeric repeat-containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science* **318**: 798-801. doi: 10.1126/science.1147182

Bettin N, Oss Pegorar C, Cusanelli E. 2019. The emerging roles of TERRA in telomere maintenance and genome stability. *Cells* **8**: 246. doi: 10.3390/cells8030246

Viceconte et al

Biffi G, Tannahill D, Balasubramanian S. 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. doi: 10.1021/ja305734x

Blasco MA. 2007. The epigenetic regulation of mammalian telomeres. *Nat Rev Genet* **8**: 299-309. doi: 10.1038/nrg2047

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics* **30**: 2114-2120. doi: 10.1093/bioinformatics/btu170

Caslini C, Connelly JA, Serna A, Broccoli D, Hess JL. 2009. MLL associates with telomeres and regulates telomeric repeat-containing RNA transcription. *Mol Cell Biol* **29**: 4519-4526. doi: 10.1128/MCB.00195-09

Chan FL, Vinod B, Novy K, Schittenhelm RB, Huang C, Udugama M, Nunez-Iglesias J, Lin JI, Hii L, Chan J et al. 2017. Aurora kinase B, a novel regulator of TERF1 binding and telomeric integrity. *Nucleic Acids Res* **45**: 21. doi: 10.1093/nar/gkx904

Chu HP, Froberg JE, Kesner B, Oh HJ, Ji F, Sadreyev R, Pinter SF, Lee JT. 2017a. PAR-TERRA directs homologous sex chromosome pairing. *Nat Struct Mol Biol* **24**: 620-631. doi: 10.1038/nsmb.3432

Chu HP, Cifuentes-Rojas C, Kesner B, Aeby E, Lee HG, Wei C, Oh HJ, Boukhali M, Haas W, Lee JT. 2017b. TERRA RNA antagonizes ATRX and protects telomeres. *Cell* **170**: 86-101. doi: 10.1016/j.cell.2017.06.017

Viceconte et al

Cox J, Mann M. 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26**: 1367-1372. doi: 10.1038/nbt.1511

Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. 2011. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res* **10**: 1794-1805. doi: 10.1021/pr101065j

Cusanelli E, Romero CA, Chartrand P. 2013. Telomeric noncoding RNA TERRA is induced by telomere shortening to nucleate telomerase molecules at short telomeres. *Mol Cell* **51**: 780-791. doi: 10.1016/j.molcel.2013.08.029

Deng Z, Norseen J, Wiedmer A, Riethman H, Lieberman PM. 2009. TERRA RNA binding to TRF2 facilitates heterochromatin formation and ORC recruitment at telomeres. *Mol Cell* **35**:403-413. doi: 10.1016/j.molcel.2009.06.025

Deng Z, Wang Z, Xiang C, Molczan A, Baubet V, Conejo-Garcia J, Xu X, Lieberman PM, Dahmane N. 2012. Formation of telomeric repeat-containing RNA (TERRA) foci in highly proliferating mouse cerebellar neuronal progenitors and medulloblastoma. *J Cell Sci* **125**: 4383-4394. doi: 10.1242/jcs.108118

Diman A, Boros J, Poulain F, Rodriguez J, Purnelle M, Episkopou H, Bertrand L, Francaux M, Deldicque L, Decottignies A. 2016. Nuclear respiratory factor 1 and

Viceconte et al

endurance exercise promote human telomere transcription. *Sci Adv* **2**: e1600031.  
doi: 10.1126/sciadv.1600031

Diman A, Decottignies A. 2018. Genomic origin and nuclear localization of TERRA telomeric repeat-containing RNA: from Darkness to Dawn. *FEBS J* **285**: 1389-1398.  
doi: 10.1111/febs.14363

Durinck S, Moreau Y, Kasprzyk A, David S, De Moor B, Brazma A, Huber W. 2005. BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics* **21**: 3439-3440. doi: 10.1093/bioinformatics/bti525

Durinck S, Spellman PT, Birney E, Huber W. 2009. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc* **4**: 698-705. doi: 10.1038/nprot.2009.97

Feretzaki M, Lingner J. 2017. A practical qPCR approach to detect TERRA, the elusive telomeric repeat-containing RNA. *Methods* **114**: 39-45. doi: 10.1016/j.ymeth.2016.08.004

Feretzaki M, Renck Nunes P, Lingner J. 2019. Expression and differential regulation of human TERRA at several chromosome ends. *RNA* **25**: 1470-1480. doi: 10.1261/rna.072322.119

Viceconte et al

Flynn RL, Centore RC, O'Sullivan RJ, Rai R, Tse A, Songyang Z, Chang S, Karlseder J, Zou L. 2011. TERRA and hnRNPA1 orchestrate an RPA-to-POT1 switch on telomeric single-stranded DNA. *Nature* **471**: 532-536. doi: 10.1038/nature09772

Göhring J, Fulcher N, Jacak J, Riha K. 2014. TeloTool: a new tool for telomere length measurement from terminal restriction fragment analysis with improved probe sensitivity correction. *Nucleic Acids Res* **42**: e21. doi: 10.1093/nar/gkt1315

Hesselberth JR. 2013. Lives that introns lead after splicing. *WIREs RNA* **4**. doi: 10.1002/wrna.1187

Hirashima K, 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. doi: 10.1093/nar/gkv063

Kim D, Langmead B, Salzberg SL. 2015. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* **12**: 357-360. doi: 10.1038/nmeth.3317

Koskas S, Decottignies A, Dufour S, Pezet M, Verdel A, Vourc'h C, Faure V. 2017. Heat shock factor 1 promotes TERRA transcription and telomere protection upon heat stress. *Nucleic Acids Res* **45**: 6321-6333. doi: 10.1093/nar/gkx208

Le Berre G, Hossard V, Riou JF, Guieysse-Peugeot AL. 2019. Repression of TERRA expression by subtelomeric DNA methylation is dependent on NRF1 binding. *Int J Mol Sci* **20**. doi:10.3390/ijms20112791

Viceconte et al

Lee YW, Arora R, Wischnewski H, Azzalin CM. 2018. TRF1 participates in chromosome end protection by averting TRF2-dependent telomeric R loops. *Nat Struct Mol Biol* **25**: 147-153. Doi: 10.1038/s41594-017-0021-5

Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**: 923-930. doi: 10.1093/bioinformatics/btt656

López de Silanes I, Stagno d'Alcontres M, Blasco MA. 2010. TERRA transcripts are bound by a complex array of RNA-binding proteins. *Nat Commun* **1**: 33. doi: 10.1038/ncomms1032

López de Silanes I, Graña-Castro O, Luigia De Bonis M, Dominguez O, Pisano DG, Blasco MA. 2014. Identification of TERRA locus unveils a telomere protection role through association to nearly all chromosomes. *Nat Commun* **5**: 4723. doi: 10.1038/ncomms5723

Mallm JP, Rippe K. 2015. Aurora kinase B regulates telomerase activity via a centromeric RNA in stem cells. *Cell Rep* **11**: 1667-1678. doi: 10.1016/j.cellrep.2015.05.015

Marión RM, Montero JJ, López de Silanes I, Graña-Castro O, Martinez P, Schoeftner S, Palacios-Fábrega JA, Blasco MA. 2019. TERRA regulate the transcriptional

Viceconte et al

landscape of pluripotent cells through TRF1-dependent recruitment of PRC2. *eLife* **8**: e44656. doi: 10.7554/eLife.44656

Mattiussi M, Tilman G, Lenglez S, Decottignies A. 2012. Human telomerase represses ROS-dependent cellular responses to Tumor Necrosis Factor- $\alpha$  without affecting NF- $\kappa$ B activation. *Cell Signal* **24**: 708-717. doi: 10.1016/j.cellsig.2011.11.004

Mayr C. 2017. Regulation by 3'-untranslated regions. *Ann Rev Genet* **51**: 171-194. doi: 10.1146/annurev-genet-120116-024704

Moravec M, Wischnewski H, Bah A, Hu Y, Liu N, Lafranchi L, King MC, Azzalin CM. 2016. TERRA promotes telomerase-mediated telomere elongation in *Schizosaccharomyces pombe*. *EMBO Rep* **17**: 999-1012. doi: 10.15252/embr.201541708

Nergadze SG, Farnung BO, Wischnewski H, Khoriantuli L, Vitelli V, Chawla R, Giulotto E, Azzalin CM. 2009. CpG-island promoters drive transcription of human telomeres. *RNA* **15**: 2186-2194. doi: 10.1261/rna.1748309

Okamoto K, Bartocci C, Ouzounov I, Diedrich JK, Yates JR, Lazzerini Denchi E. 2013. A two-step mechanism for TRF2-mediated chromosome-end protection. *Nature* **494**: 502-505. doi: 10.1038/nature11873

Viceconte et al

Petti E, Buemi V, Zappone A, Schillaci O, Veneziano Broccia P, Dinami R, Matteoni S, Benetti R, Schoeftner S. 2019. SFPQ and NONO suppress RNA:DNA-hybrid-related telomere instability. *Nat Commun* **10**: 1001. doi: 10.1038/s41467-019-08863-1

Porro A, Feuerhahn S, Delafontaine J, Riethman H, Rougemont J, Lingner J. 2014a. Functional characterization of the TERRA transcriptome at damaged telomeres. *Nat Commun* **5**: 5379. doi: 10.1038/ncomms6379

Porro A, Feuerhahn S, Lingner J. 2014b. TERRA-reinforced association of LSD1 with MRE11 promotes processing of uncapped telomeres. *Cell Rep* **6**: 765-776. doi: 10.1016/j.celrep.2014.01.022

R core team. 2017. R: a language and environment for statistical computing. *Open J Statistics* **7**. <https://www.R-project.org/>

Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. *Nat Biotechnol* **29**: 24-26. doi: 10.1038/nbt.1754

Rosenfeld JA, Wang Z, Schones DE, Zhao K, DeSalle R, Zhang MQ. 2009. Determination of enriched histone modifications in non-genic portions of the human genome. *BMC Genomics* **10**: 143. doi: 10.1186/1471-2164-10-143

Viceconte et al

Rudenko G, Van der Ploeg LH. 1989. Transcription of telomere repeats in protozoa. *EMBO J* **8**: 2633-2638

Sagie S, Toubiana S, Hartono SR, Katzir H, Tzur-Gilat A, Havazelet S, Francastel C, Velasco G, Chédin F, Selig S. 2017. Telomeres in ICF syndrome cells are vulnerable to DNA damage due to elevated DNA:RNA hybrids. *Nat Commun* **8**: 14015. doi: 10.1038/ncomms14015

Scheibe M, Arnoult N, Kappei D, Buchholz F, Decottignies A, Butter F, Mann M. 2013. Quantitative interaction screen of telomeric repeat-containing RNA reveals novel TERRA regulators. *Genome Res* **23**: 2149-2157. doi: 10.1101/gr.151878.112

Schoeftner S, Blasco MA. 2008. Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. *Nat Cell Biol* **10**: 228-236. doi: 10.1038/ncb1685

Schoeftner S, Blanco R, Lopez de Silanes I, Muñoz P, Gómez-López G, Flores JM, Blasco MA. 2009. Telomere shortening relaxes X chromosome inactivation and forces global transcriptome alterations. *Proc Natl Acad Sci USA* **106**: 19393-19398. doi: 10.1073/pnas.0909265106

Solovei I, Gaginskaya ER, Macgregor HC. 1994. The arrangement and transcription of telomere DNA sequences at the ends of lampbrush chromosomes of birds. *Chromosome Res* **2**: 460-470. doi: 10.1007/bf01552869

Viceconte et al

Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P et al. 2017. The STRING database in 2017: Quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res* **45(D1)**: D362-D368. doi: 10.1093/nar/gkw937

The Gene Ontology Consortium. 2017. Expansion of the gene ontology knowledgebase and resources. *Nucleic Acids Res* **45(D1)**: D331-D338. doi: 10.1093/nar/gkw1108

Tilman G, Lorient A Van Beneden A, Arnoult N, Londoño-Vallejo JA, De Smet C, Decottignies A. 2009. Subtelomeric DNA hypomethylation is not required for telomeric sister chromatid exchanges in ALT cells. *Oncogene* **28**: 1682-1693. doi: 10.1038/onc.2009.23

Van Beneden A, Arnoult N and Decottignies A. 2013. Telomeric RNA expression: length matters. *Front Oncol* **3**: 178. doi: 10.3389/fonc.2013.00178

Viceconte N, Dheur MS, Majerova E, Pierreux CE, Baurain JF, van Baren N, Decottignies A. 2017. Highly aggressive metastatic melanoma cells unable to maintain telomere length. *Cell Rep* **19**: 2529-2543. doi: 10.1016/j.celrep.2107.05.046

Wickham H. 2016. Ggplot2: elegant graphics for data analysis. New York, NY: Springer-Verlag. <https://ggplot2.tidyverse.org/>

Viceconte et al

Yehezkel S, Segev Y, Viegas-Péquignot E, Skorecki K, Selig S. 2008. Hypomethylation of subtelomeric regions in ICF syndrome is associated with abnormally short telomeres and enhanced transcription from telomeric regions. *Hum Mol Genet* **17**: 2776-2789. doi: 10.1093/hmg/ddn177

Yu G, Wang LG, Han Y, He QY. 2012. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**: 284-287. doi: 10.1089/omi.2011.0118

Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J, Billis K, Cummins C, Gall A, Girón CG et al. 2018. Ensembl 2018. *Nucleic Acids Res* **46(D1)**: D754-D761. doi: 10.1093/nar/gkx1098

Zhang L-F, Ogawa Y, Ahn JY, Namekawa SH, Silva SS, Lee JT. 2009. Telomeric RNAs mark sex chromosomes in stem cells. *Genetics* **182**: 685-692. doi: 10.1534/genetics.109.103093

Viceconte et al

## Tables

### Figure legends

**Figure 1. TERRA-FISH signals correlate with telomere length but not with *UUAGGG* content in human cells.** **A.** Immunofluorescence against TRF2 (red) combined with RNA-FISH to detect TERRA with a  $(CCCTAA)_7$  LNA green probe in HT1080-ST human cell line. DNA is stained with DAPI (blue). Scale bar, 5  $\mu$ m. **B.** Comparison between TERRA-FISH (left) and telomeric FISH (right) in HCA2 (telomere length: 11.1 kb) and LB37 (telomere length: 3.8 kb) human cells using the  $(CCCTAA)_7$  LNA green probe and the same exposure time. Experimental conditions are indicated below. Scale bar, 5  $\mu$ m. **C.** Correlation between total nuclear TERRA-FISH intensity (expressed as % of intensity in HFF2 cells) and total Telo-FISH intensity (expressed as % of intensity in HCA2 cells). At least 50 nuclei were quantified for TERRA-FISH or Telo-FISH in each cell line. Cell line names are indicated.  $R=0.943$ ;  $p$  two-tailed=0.016. **D.** Correlation between TERRA-FISH intensity (expressed as % of intensity in HFF2 cells) and telomere length (kb) evaluated by TRF (Arnoult et al. 2012; Tilman et al. 2009).  $R=0.882$ ;  $p$  two-tailed=0.048. **E.** RNA slot-blots in LB37 and HCA2 human cells. Total RNA (2 or 4  $\mu$ g) isolated from LB37 or HCA2 cells was hybridized with either the  $(CCCTAA)_4$  probe or a probe against human *ACTB*. **F.** Quantification from B and E. Data are normalized to HCA2. Mean  $\pm$  S.D.

**Figure 2. TERRA-FISH signals correlate with *UUAGGG* content but not with telomere length in mouse cells.** **A.** TRF analysis on genomic DNA isolated from the indicated mouse cell lines. Position of the ladder is indicated on the left with the

Viceconte et al

corresponding size (kb). Telomere length, calculated with Telotool software, is indicated below. **B.** TERRA-FISH in J774A.1, Neuro-2a, L929 and M1 mouse cell lines using the  $(CCCTAA)_7$  LNA green probe with (right) or without (left) RNase A treatment. DNA is stained with DAPI (blue). Bar scale, 5  $\mu$ m. **C.** Correlation between measurements in A and B. At least 35 nuclei were quantified for TERRA-FISH in each cell line.  $R=-0.47$ ;  $p$  two-tailed=0.53. **D.** Immunofluorescence against Terf1 (red) combined with RNA-FISH to detect TERRA with a  $(CCCTAA)_7$  LNA green probe in Neuro-2a mouse cells. Arrows indicate co-localization events. DNA is stained with DAPI (blue). Bar scale, 5  $\mu$ m. **E.** Immunofluorescence against Terf1 (red) combined with FISH to detect telomeres with a  $(CCCTAA)_7$  LNA green probe in Neuro-2a mouse cells. DNA is stained with DAPI (blue). Bar scale, 5  $\mu$ m. **F.** RNA slot-blots in the indicated mouse cell lines. Total RNA (1, 0.5 or 0.25  $\mu$ g) was hybridized with the  $(CCCTAA)_4$  probe (upper panel) or with a probe against mouse *ACTB* (lower panel, 0.5  $\mu$ g). **G.** Correlation between measurements from F and *TTAGGG* content evaluated by slot-blot (Fig S1A-B).  $R=-0.37$ ;  $p$  two-tailed=0.28. **H.** Summary of the differences between human and mouse TERRA.

**Figure 3. *In silico* search for candidate mouse genomic loci producing UUAGGG-containing RNAs.** **A.** Workflow for the identification of mouse candidate genomic loci for TERRA production. **B.** Graphical representation of mouse chromosomes and putative loci for TERRA production in green (telomeric loci) or red (intrachromosomal loci). **C.** IGV screenshots showing coverage density over regions of interest for the indicated mouse cell lines or tissues. The red line on each chromosome indicates the region displayed below. Genes or portions of genes are represented, with dark blue boxes corresponding to exons and dark blue lines with

Viceconte et al

arrows corresponding to introns. *TTAGGG* or variant (as indicated in A) repeats are indicated in red. Regions corresponding to PCR products in the study appear in light blue with dotted lines on both sides. Displayed coverages are normalized (each value was multiplied by 1e6 and divided by the total number of reads) and correspond to reads unambiguously mapped and resulting from transcription in sense orientation exclusively. **D.** Quantification of reads aligned to the regions shown in C (“PCR products”). Counts are normalized (read counts were divided by the region length (kb) and multiplied by 1e6 / total number of reads).

**Figure 4. Candidate TERRA molecules contain *UUAGGG* repeats.** **A.** Graphical representation of the genomic context for the four mouse TERRA loci analyzed in this study. *TTAGGG*-rich regions, whether at telomeres or at intrachromosomal loci, are shown by red boxes and their length is indicated. PCR product positions are indicated with light blue boxes. **B.** Relative expression levels of Telo 18q, Chr 2, PAR or Tsix TERRA loci and *Sod1* in cDNA from J774A.1 and ES cells synthesized with either random or Telo and GAPDH primers. Values are normalized first to *GAPDH* and then to the ratio measured in J774A.1 cells when random primers are used for RT. Mean  $\pm$  S.D.

**Figure 5. TERRA transcripts mostly originate from the PAR locus in mouse cells.** **A.** Relative expression levels of PAR, Telo 18q, Chr 2 or Tsix TERRA loci in cDNA from the indicated cell lines synthesized with random primers. Values are normalized first to *GAPDH* and then to the PAR expression level in J774A.1 cells. Mean  $\pm$  S.D. **B.** RNA slot-blots in the indicated mouse cell lines. Total RNA (1  $\mu$ g) was hybridized with the (CCCTAA)<sub>4</sub> probe or with a probe against mouse *ACTB*. **C.**

Viceconte et al

Correlations between measurements from A and B as indicated (all measurements were normalized to J774A.1). R values and p two-tailed values are indicated for each graph.

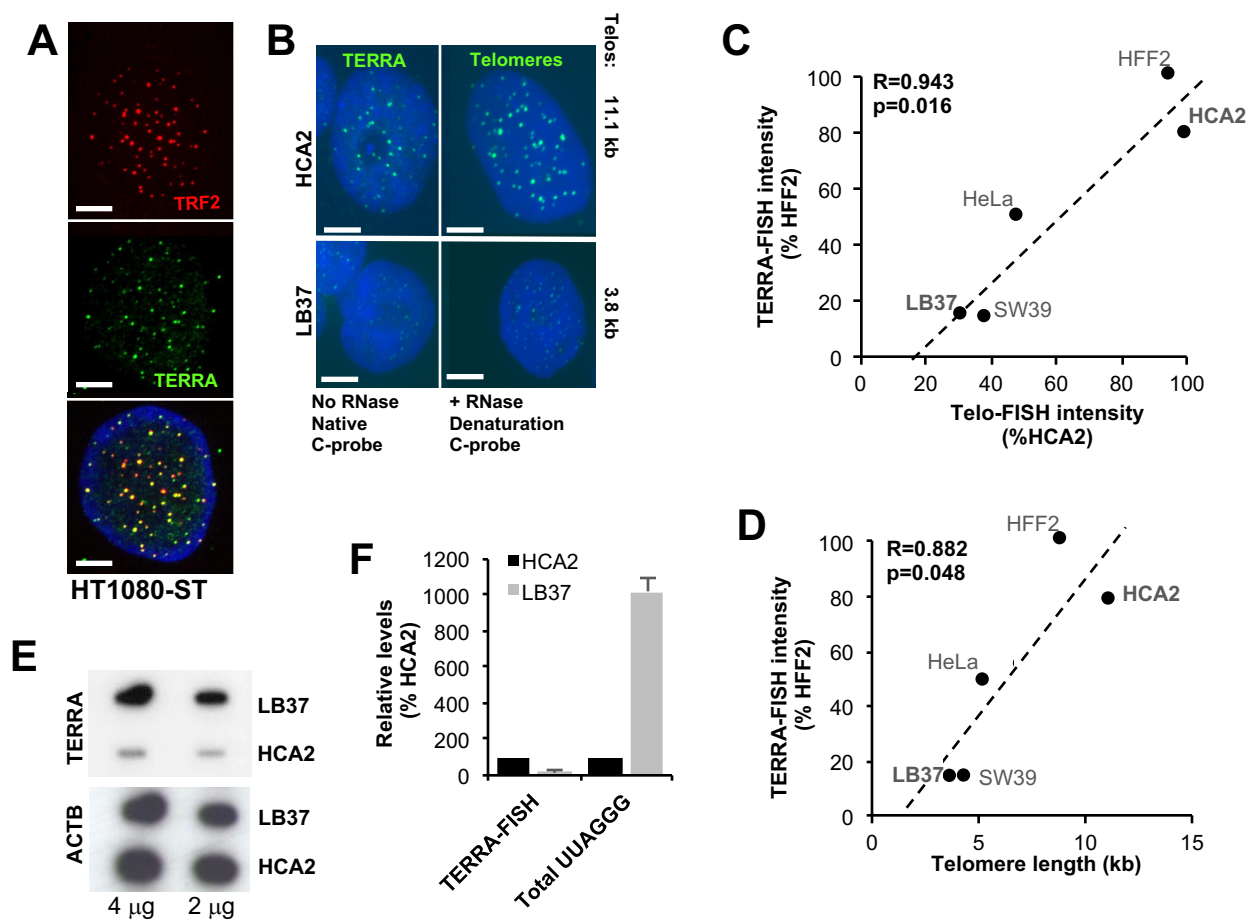
**Figure 6. Transcription of mouse telomeres is not increased upon Terf2 depletion.** **A.** Western blot analysis of Terf2 in total cell extracts from CreERT2 Terf2 F/+ and Terf2 F/F SV40T-immortalized MEF.  $\beta$ -actin is shown as loading control. **B.** Visualization of PCR products from Telo 18q, Telo 5q and Telo 11q obtained through amplification on either gDNA (upper panel) or cDNA (lower panel) prepared from J774A.1 cells. The distance between each PCR product and the respective telomere is indicated below (bp). **C.** Ct values from RT-qPCR experiments on cDNA synthesized with Telo and GAPDH primers from the F/+ or F/F fibroblast cell lines treated as indicated. Mean  $\pm$  S.D.

**Figure 7. SILAC screening of the quantitative mouse TERRA pull-down.** **A.** STRING network showing the 128 mouse TERRA-interacting candidates from either the pull-down screen (using  $\geq 4$ -fold enrichment as threshold) or the iDRiP screen (Chu et al. 2017b) that either showed perfect overlap between the two screens (highlighted in yellow) or were belonging to similar protein families. Candidates that were exclusively found in the SILAC pull-down screen are shown in bold. Line color of the edges indicates the type of interaction evidence between the nodes according to STRING default parameters. **B.** Validation of Suz12 and Blm as PAR-TERRA-interacting proteins, using the shelterin complex as positive control for binding. Three independent RNA-IP experiments were performed for each protein using nuclear extracts from 2575i MEF and Telo primers for RT. No antibody was added in the

Viceconte et al

control experiments (-Ab) and additional controls lacking reverse transcriptase (-RT) were performed to monitor gDNA contamination. PAR-TERRA transcript levels in the IP were normalized to input (%). Mean $\pm$ S.E.M, two-tailed unpaired *t*-tests. **C.** Average Ct values for PAR-TERRA and Telo 18q RT-qPCR in the input and the RIP samples. Mean $\pm$ S.D., n=3 independent experiments.

**Figure 8. Comparison of human and mouse TERRA interactomes. A.** Two-dimensional interaction plot for the TERRA pull-down under 250 mM sodium chloride washing conditions both for murine and human (Scheibe et al. 2013) samples. Dotted red line indicates the enrichment threshold. The top TERRA enriched proteins are annotated, highlighting in green those that are shared between human and mouse datasets. **B.** GO terms associated with the mouse 188 common interacting proteins were tested for over-representation (with Fisher's exact test) against GO terms associated to all detected proteins in the mouse dataset. Nodes represent GO categories and edges show overlapping gene sets. Color scale shows the significance (BH adjusted p-values) of the over-representation and node sizes are proportional to the number of genes per GO category.



**Figure 1**

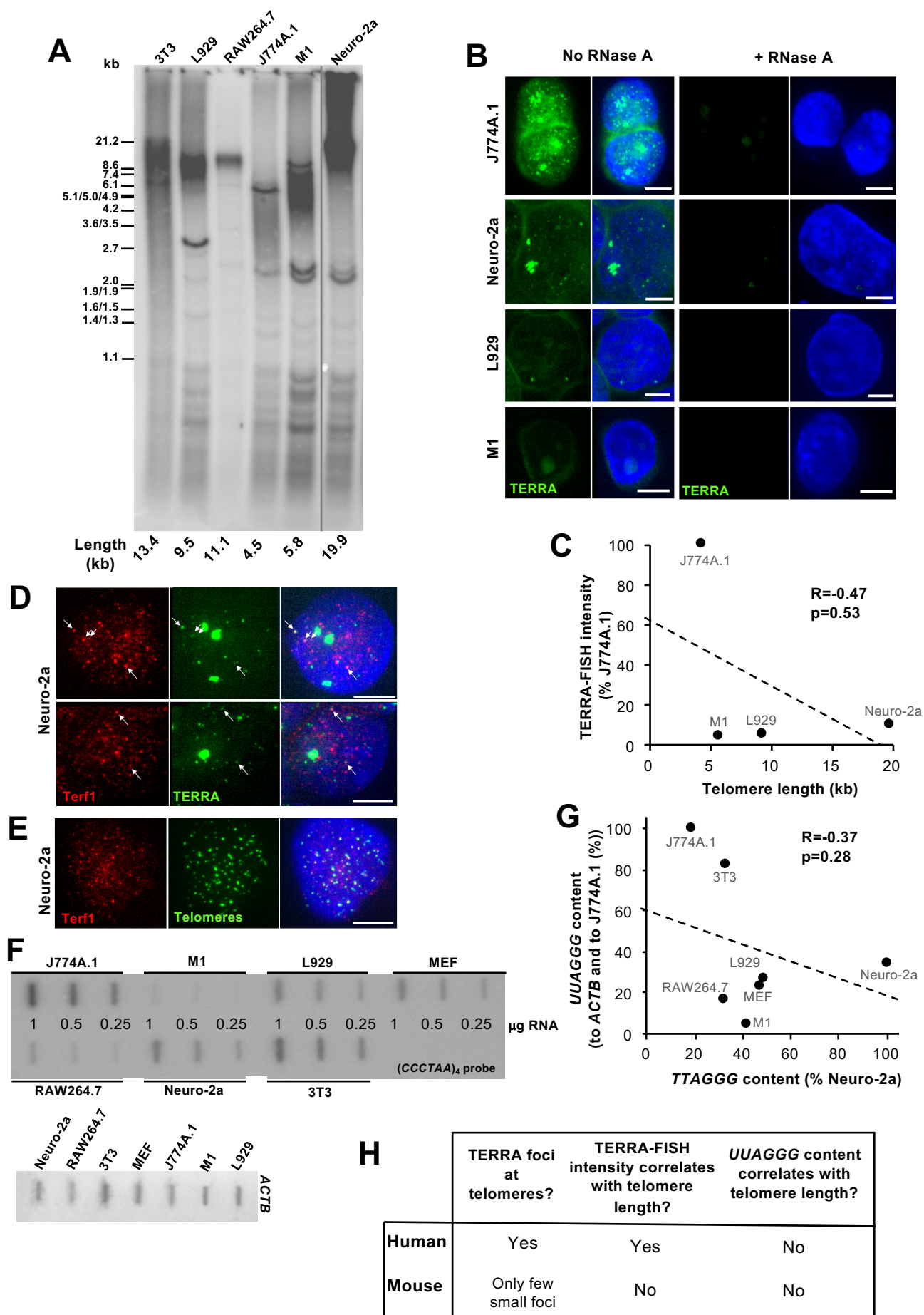
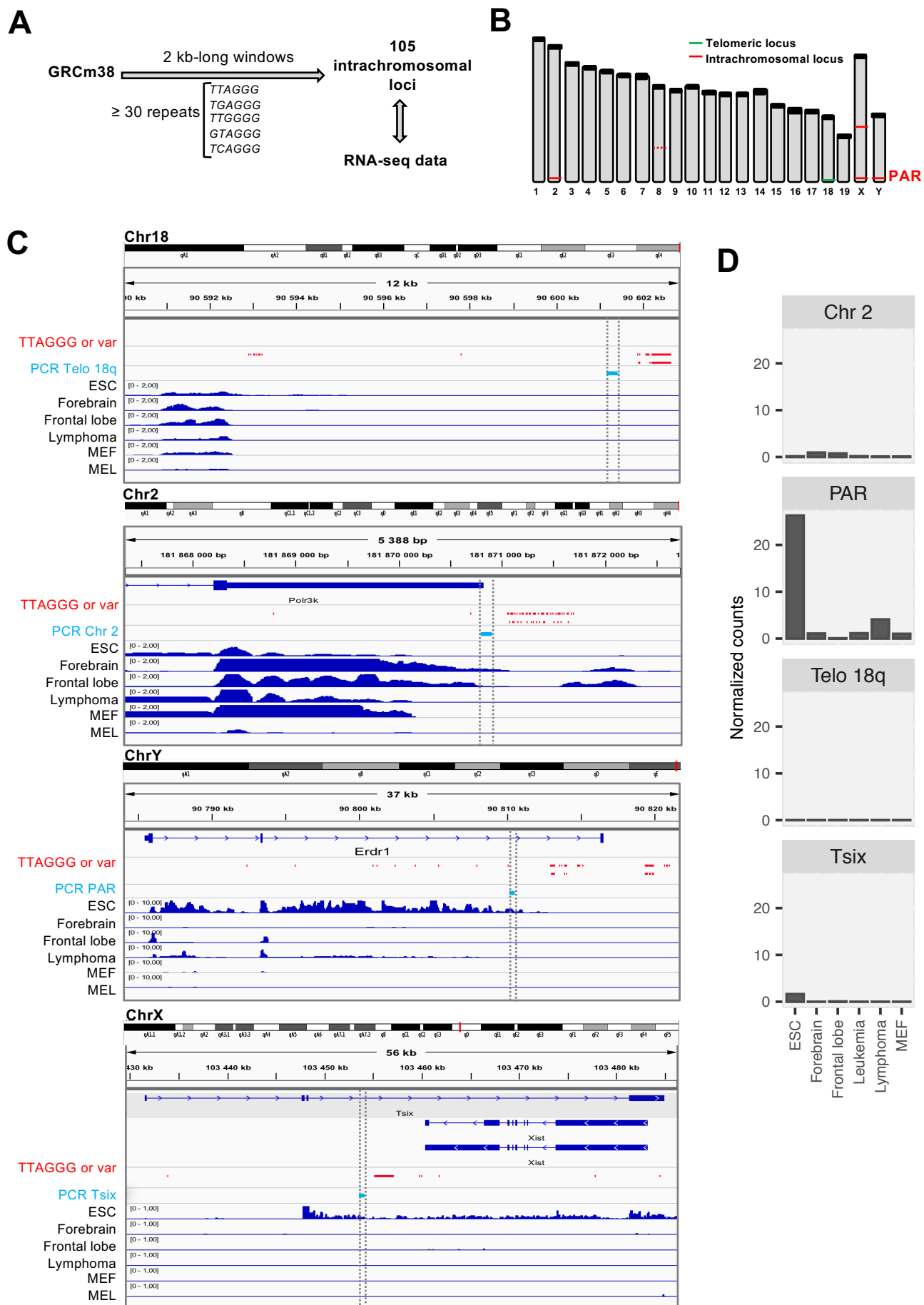


Figure 2



**Figure 3**

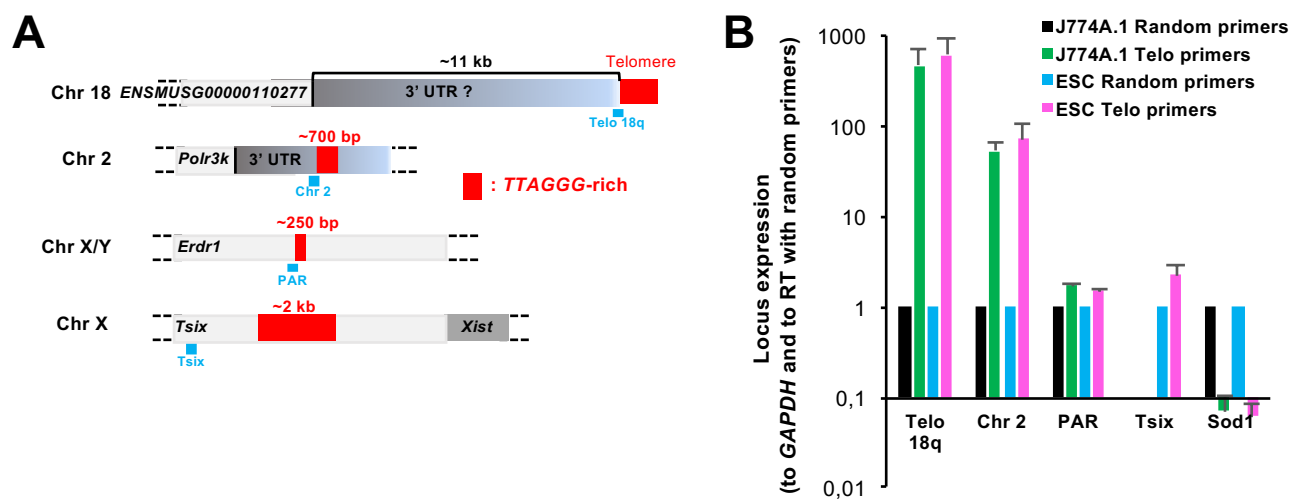
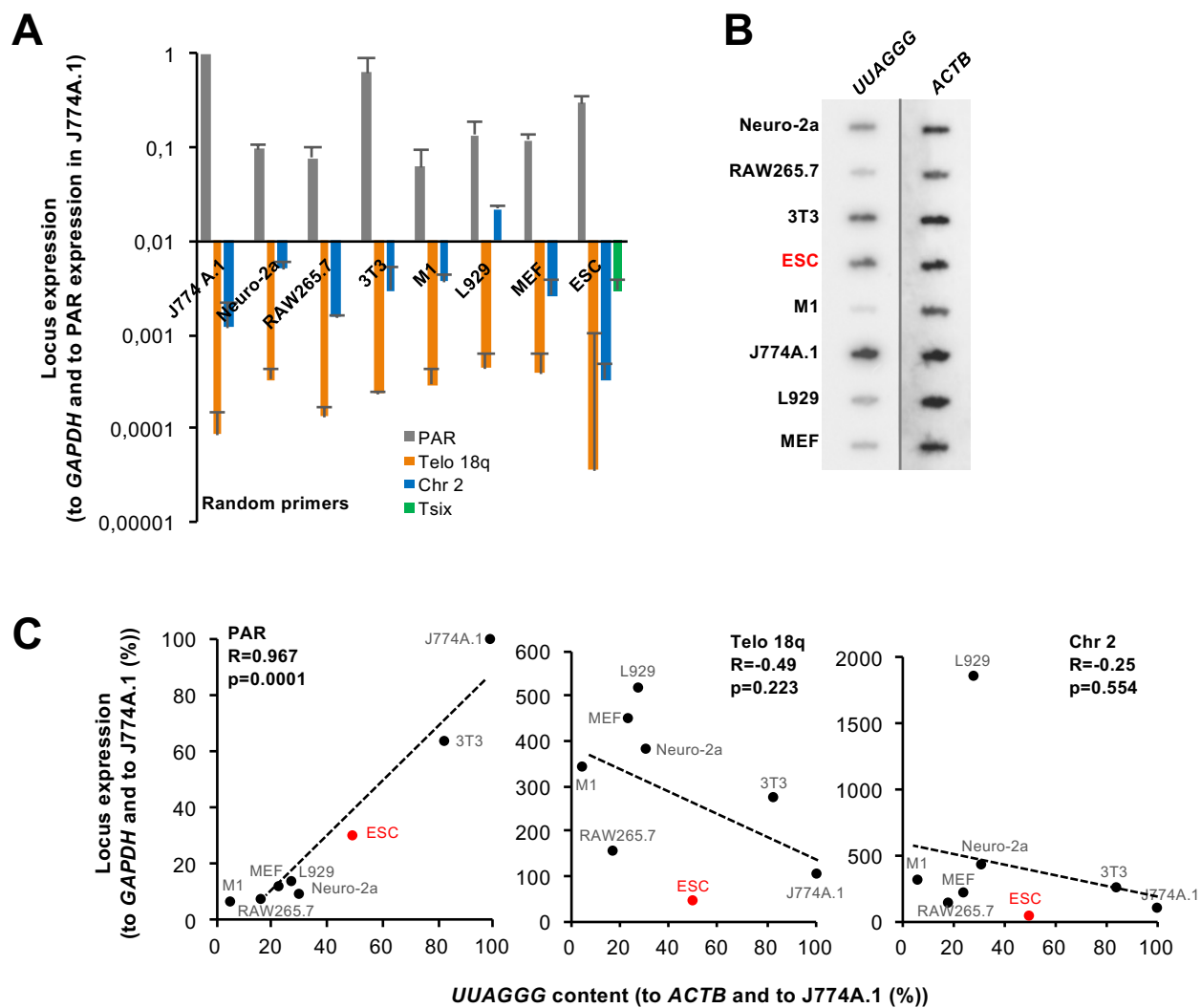


Figure 4



**Figure 5**

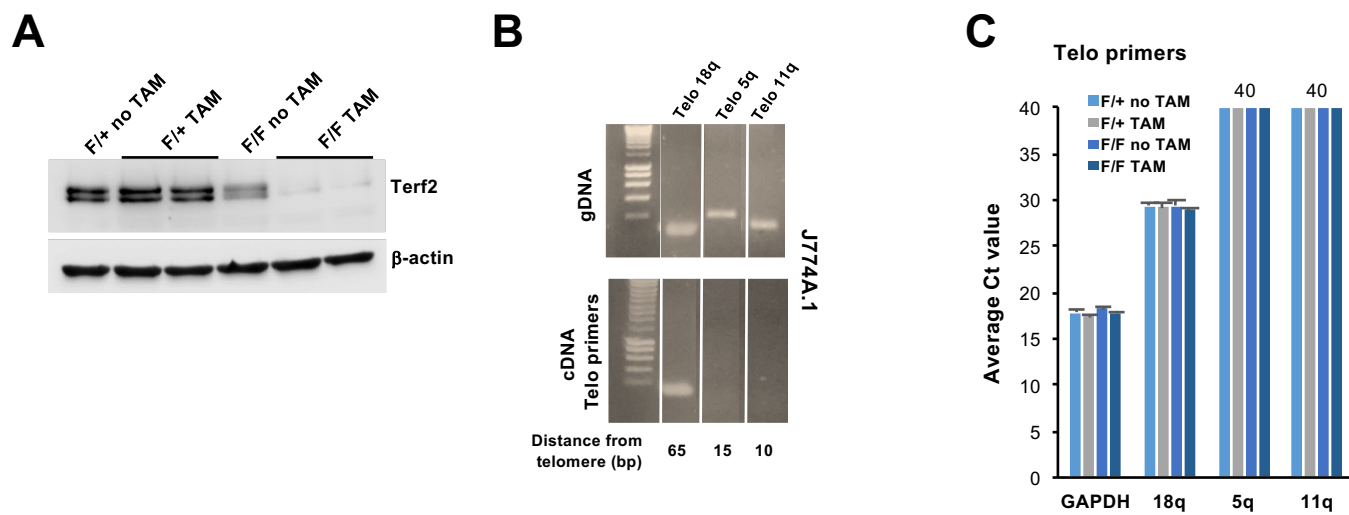


Figure 6

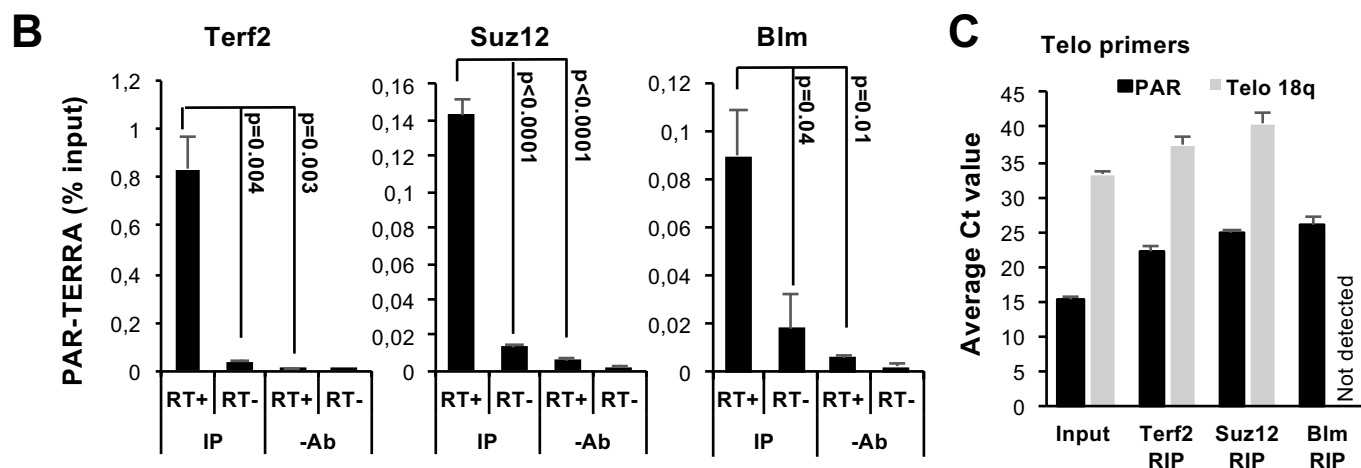
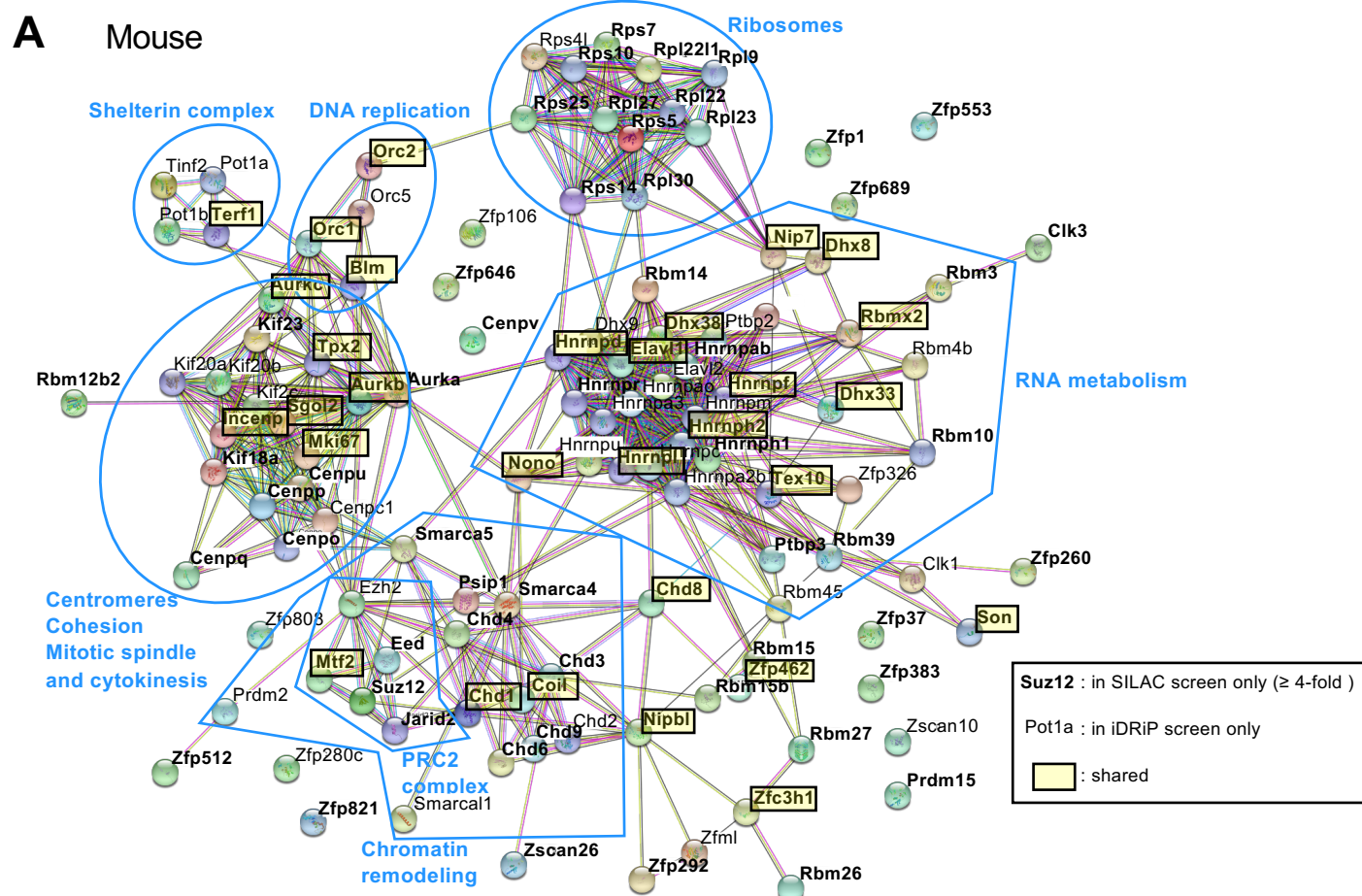
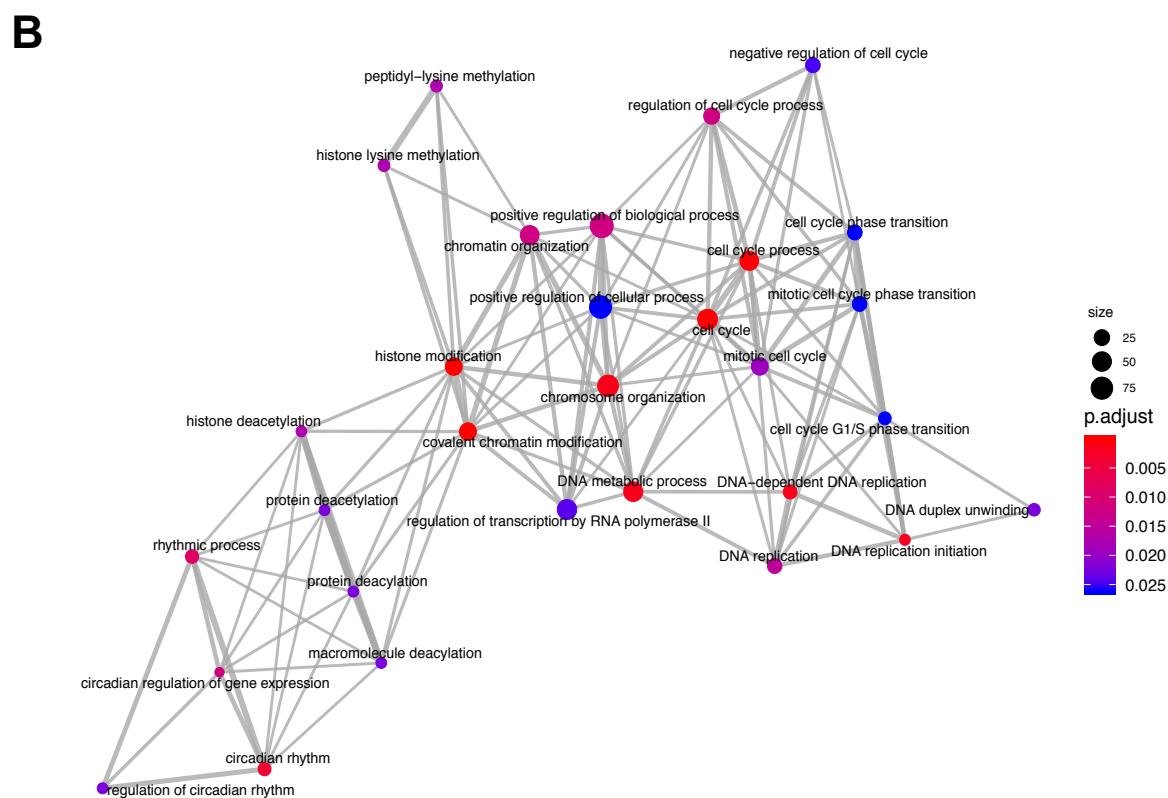
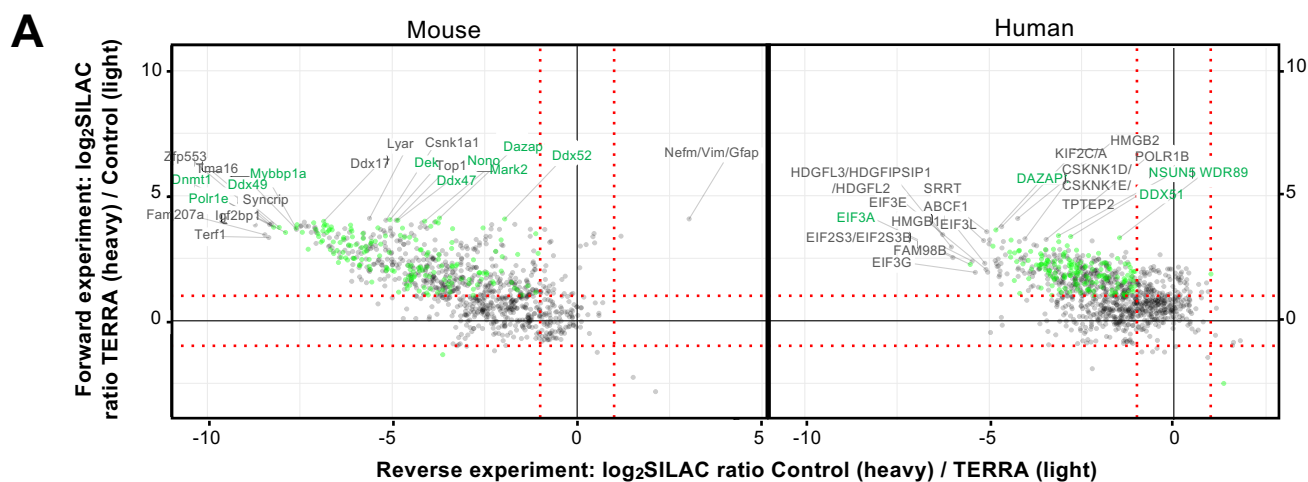


Figure 7



**Figure 8**



# RNA

A PUBLICATION OF THE RNA SOCIETY

## PAR-TERRA is the main contributor to telomeric repeat-containing RNA transcripts in normal and cancer mouse cells

Nikenza Viceconte, Axelle Lorient, Patricia Lona Abreu, et al.

RNA published online October 30, 2020

---

### Supplemental Material

<http://rnajournal.cshlp.org/content/suppl/2020/10/30/rna.076281.120.DC1>

### P<P

Published online October 30, 2020 in advance of the print journal.

### Accepted Manuscript

Peer-reviewed and accepted for publication but not copyedited or typeset; accepted manuscript is likely to differ from the final, published version.

### Creative Commons License

This article is distributed exclusively by the RNA Society for the first 12 months after the full-issue publication date (see <http://rnajournal.cshlp.org/site/misc/terms.xhtml>). After 12 months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

### Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

---

SMART<sup>®</sup> cDNA + library prep:  
now all from **one source**



---

To subscribe to RNA go to:  
<http://rnajournal.cshlp.org/subscriptions>

---