Analysis of telomeric chromatin using pyrrole-imidazole polyamide probe

佐々木　飛鳥

博士（理学）

総合研究大学院大学
生命科学研究科
遺伝学専攻

平成３０（２０１８）年度

Analysis of telomeric chromatin using pyrrole-imidazole polyamide probe

Sasaki, Asuka

Doctor of Philosophy

Department of Genetics

School of Life Science

SOKENDAI

(Graduate University for Advanced Studies)

2018 (school year)

Contents

List of Figures

Abstract

Mammalian telomeres consist of a long array of repetitive sequence (TTAGGG) and cap chromosome ends to maintain chromosome stability. Without the telomerase activity that can elongate the telomeres, the telomere length shortens every cell division. Upon reaching a critical short length, the telomeres trigger cellular senescence, which can suppress abnormal cell growth relating to tumorigenesis. Telomerase reactivation circumvents the growth limitation by the telomere-dependent senescence and leads to immortalization. Induction of recombination between telomeres, termed alternative lengthening of telomeres (ALT), can also lead to a similar effect leading to immortalization and tumorigenesis. In this thesis, to investigate telomeres in cancer cells by imaging, proteomics, and RNA interactome, I performed two projects: telomere visualization in tissue sections and isolation of telomere-associated noncoding RNAs.

Initially, I report a development of a telomere-visualization method in mouse and human tissue sections using PI polyamide probes. PI polyamides are DNA binding compounds that can be designed to target predetermined DNA sequences. To visualize telomeres in cultured cells and tissue sections, fluorescent in situ hybridization (FISH) method has been used as a 'standard' method. However, this method needs a DNA denaturation step for probe hybridization by harsh heating and formamide treatments, and carries the risk of destroying cellular structures which make it difficult to co-stain telomeres with a protein using an antibody. Because PI polyamides bind to the target sequences in the minor groove of double-stranded DNA without denaturation step, this compound is

compatible with immunostaining and has an advantage compared to FISH method. As collaboration with HiPep Laboratories and Prof. Sugiyama's group at Kyoto University, a fluorescent PI polyamide probe (HPTH-59) that target the mammalian telomere sequence with high specificity was developed. I showed that HPTH-59 visualized telomeres not only in cultured cells but also in mouse and human frozen tissue sections. Double staining with HPTH-59 and antibodies were performed. By quantitatively measuring the telomere length in clinical tissues from an esophageal cancer patient, I found that a cell population positive for a proliferation marker, Ki-67 protein, had shorter telomeres than do the Ki-67-negative cells in non-tumor tissue sections. From these results, I propose that PI polyamides are promising alternative for telomere labeling in cell biology as well as clinical research.

Next, I investigated the chromatin composition of telomeres. Telomeric DNA is three-dimensionally organized as chromatin, where nucleosomes are associated with non-histone proteins and RNAs. It still remains unclear how the telomeric chromatin structure regulates telomere maintenance because of limited information of telomeric chromatin composition. Although various methods were previously developed to identify telomere-binding proteins, there is still no reliable technique for unbiased identification of both proteins and RNAs including non-coding RNAs (ncRNAs) associated with telomeric chromatin. Since HPTH-59 polyamides bind to telomere DNA with high affinity in a mild condition, I expected that HPTH-59 enable us to analyze telomeric chromatin not only by imaging but also by affinity purification for proteomics and RNAomics.

To dissect the telomeric chromatin composition, I show a novel approach of locus-specific chromatin purification using a telomere-targeting PI polyamide, and named PI

polyamide-based proteomics and RNA-omics of isolated chromatin segments (PI-PRICh). By applying PI-PRICh to mouse erythrocytes leukemia (MEL) cells, I was able to identify proteins known to associate with telomeres such as shelterin complex (TRF1, TRF2, POT1, and TIN2) using mass spectrometric analysis. At the same time, I also extracted RNA fraction associated with telomeric chromatin followed by next generation-sequencing (NGS) for comprehensive identification of telomeric chromatin-associated RNAs. PI-PRICh highly enriched (>500-folds) the well-known telomerase RNA component (TERC) and telomeric repeat-containing RNA (TERRA) bound to the telomeres.

To identify ncRNAs associated with the telomeres in human ALT cell in which telomeres are highly recombinogenic, I compared ncRNAs in ALT cells with those in telomerase positive cells. I found that several intronic ncRNAs specifically identified in ALT cells. A possible physiological role of the intronic ncRNAs will be discussed.

Finally, from my studies I emphasize that understanding the nature of telomeric chromatin in different types of cancers from various angles is important for basic and clinical aspects. In addition, I anticipate that PI polyamide will be a powerful tool for chromatin imaging and the chromatin composition analysis of other regulatory sequences.

Chapter 1

Telomere Visualization in Tissue Sections using Pyrrole–Imidazole Polyamide Probes

1.1. Introduction

Telomeres protect the end of chromosomes. In mammalian chromosomes, the telomeres consist of a repetitive sequence (TTAGGG)n to prevent chromosome instability and can be elongated by an enzyme "telomerase". In differentiated cells that do not have the telomerase activity, the telomere length gradually decreases with every cell division. Upon reaching to a critical length, the telomeres can induce cellular senescence, which suppresses further cell division. The senescence is thought to inhibit the abnormal growth associated with tumorigenesis (Blackburn 2010; Zakian 2012; Smogorzewska and de Lange 2004). Reactivation of the telomerase can induce cells to circumvent growth limitation by the telomere-dependent senescence and lead to immortalization (Shay and Bacchetti 1997; Kim et al. 1994). Additionally, induction of recombination between telomeres, termed alternative lengthening of telomeres (ALT), can also lead to a telomere elongation resulting to immortalization from the telomere-dependent senescence (Bryan et al. 1997; Bryan et al. 1995).

Telomere length can be used as a diagnostic marker to detect immortalized cells with short telomeres or ALT cells with much longer telomeres. Thus far, quantitative fluorescent in situ hybridization (Q-FISH) with a peptide nucleic acid (PNA) probe has

been widely used to visualize relative telomere length in individual cells (Lansdorp et al. 1996; Henderson et al. 1996; Levsky and Singer 2003). Many clinical studies of telomere length measurement by Q-FISH have demonstrated that there are types of cancer cells have shortened telomeres (Meeker, Gage, et al. 2002; Meeker, Hicks, et al. 2002; Ferlicot et al. 2003). However, the FISH method requires harsh treatment using heat and 50% formamide for probe hybridization, which carries the risk of destroying cellular structures. Indeed, only a few studies have performed telomere labeling along with immunostaining for a cell marker (e.g., tumor marker) that can provide results with physiological relevance in human tissue sections (Meeker et al. 2004). Additionally, simple and less time-consuming steps and labeling reproducibility, such as clinical studies with numerous samples, are preferable for high-throughput experiments. The 'gold standard' for labeling telomeres, FISH, still has issues to be resolved, including for instance, the time required and convenience.

N-methylpyrrole (P)−N-methylimidazole (I) (PI) polyamides bind to the minor groove of double-stranded DNA without denaturation and can recognize Watson–Crick base pairs (Figure 1) (Trauger, Baird, and Dervan 1996; White et al. 1998; Chenoweth and Dervan 2009; Dervan 2001; Dervan and Edelson 2003; Dervan, Doss, and Marques 2005; Bando and Sugiyama 2006; Blackledge and Melander 2013). PI polyamide is an alternative to the nucleic acid probes, which require denaturing genomic DNA, before probe hybridization. Maeshima *et al*. a tandem hairpin PI polyamide probe with a fluorophore (TH59) that target human telomere sequences $(TTAGGG)_n$ under mild conditions (Maeshima, Janssen, and Laemmli 2001). A hairpin structure is composed of two antiparallel linear compounds folded in a U-shape and is connected to another

5

hairpin moiety through a hinge segment (Figure 2). To increase the binding selectivity of TH59, my collaborators developed a method for the robust synthesis of TH59 and derivatives of TH59 with higher specificity for human telomeres in the cell (Kawamoto et al. 2013; Hirata et al. 2014; Kawamoto et al. 2015).

Here, I show that HPTH59-b, which has an optimal hinge connecting two hairpins, visualizes telomeres not only in cultured cells but also in mouse and human frozen tissue sections (Figure 3). The tissue sections were co-stained with HPTH59-b and antibodies for cell-specific markers to compare the telomere lengths in various cell populations in the tissue sections. By quantitatively measuring telomeres in clinical tissue samples from an esophageal cancer patient, I found that a cell population positive for a proliferation marker, Ki-67 has shorter telomeres than those in the Ki-67-negative cells in non-tumor tissue sections. Therefore, PI polyamide provides an advantageous alternative for the measurement of telomere length in clinical research.
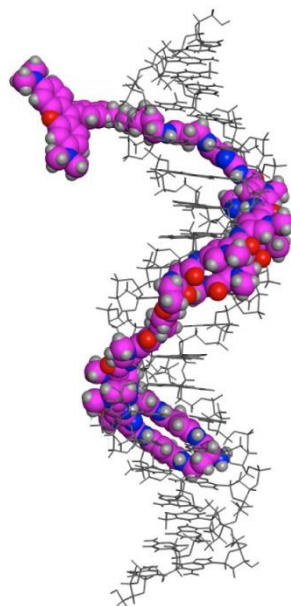
Figure 1. A structural model of HPTH59-b binding to DNA.
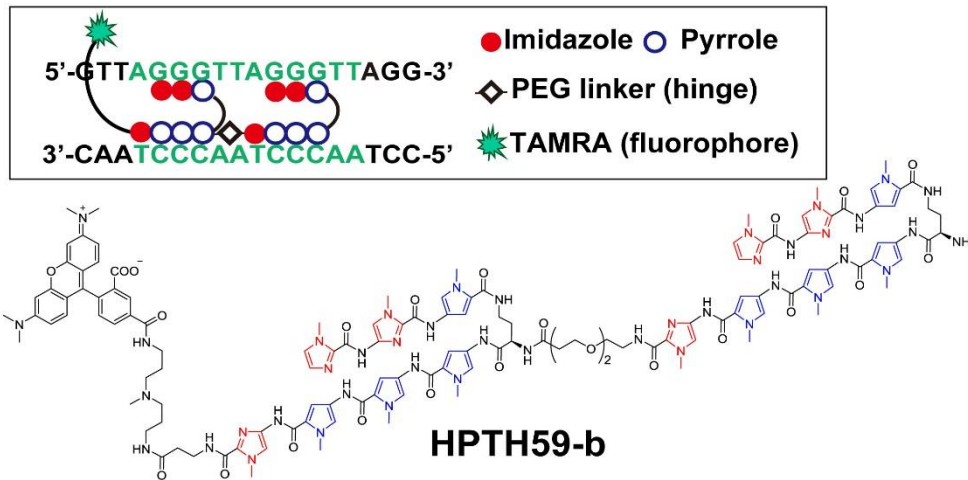
Figure 2. Chemical structure of HPTH59-b.

TAMRA is tetramethylrhodamine. In the box, the base recognition profile of HPTH59-b is shown.
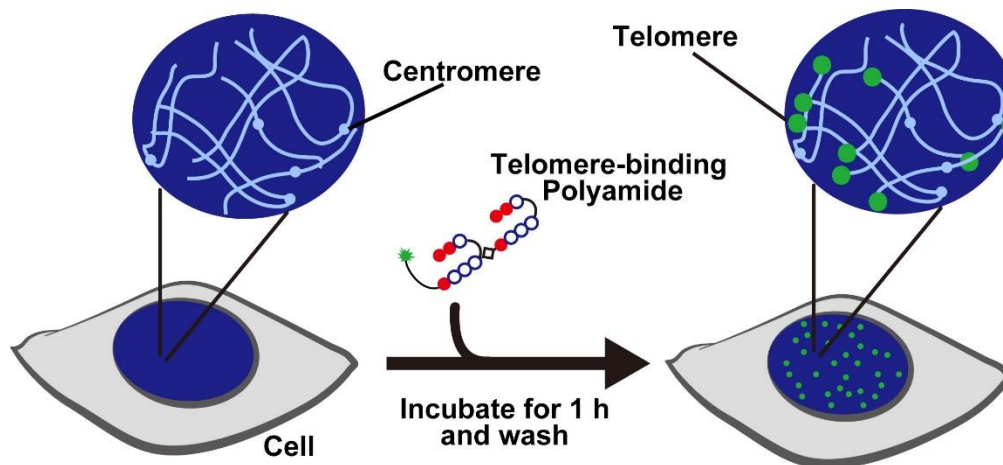
Figure 3. A simple scheme for telomere labeling using HPTH59-b.

Green dots indicate signals from telomere repeats.

1.2. Results

Telomere length measurement by quantitative PI polyamide labeling

Several studies have shown that telomeres in cancer tissues are often shorter than those in normal tissues (Hastie et al. 1990; Bryan et al. 1998; Sommerfeld et al. 1996), thus, quantitative telomere measurement should make an important contribution to investigations of tumorigenesis. To verify the quantification of telomere length by PI polyamide labeling, I stained the telomeres in cells having different lengths of telomere repeats with fluorescently labeled HPTH59-b (Hirata et al. 2014) and 4', 6-diamidino-2-phenylindole (DAPI) (Figure 4, 5). The telomere lengths of HeLaS3, HeLa1.3, and U2-OS ALT cells are 2–10 kb, ~23 kb, and <3 to >50 kb, respectively (Takai et al. 2010; Scheel et al. 2001).
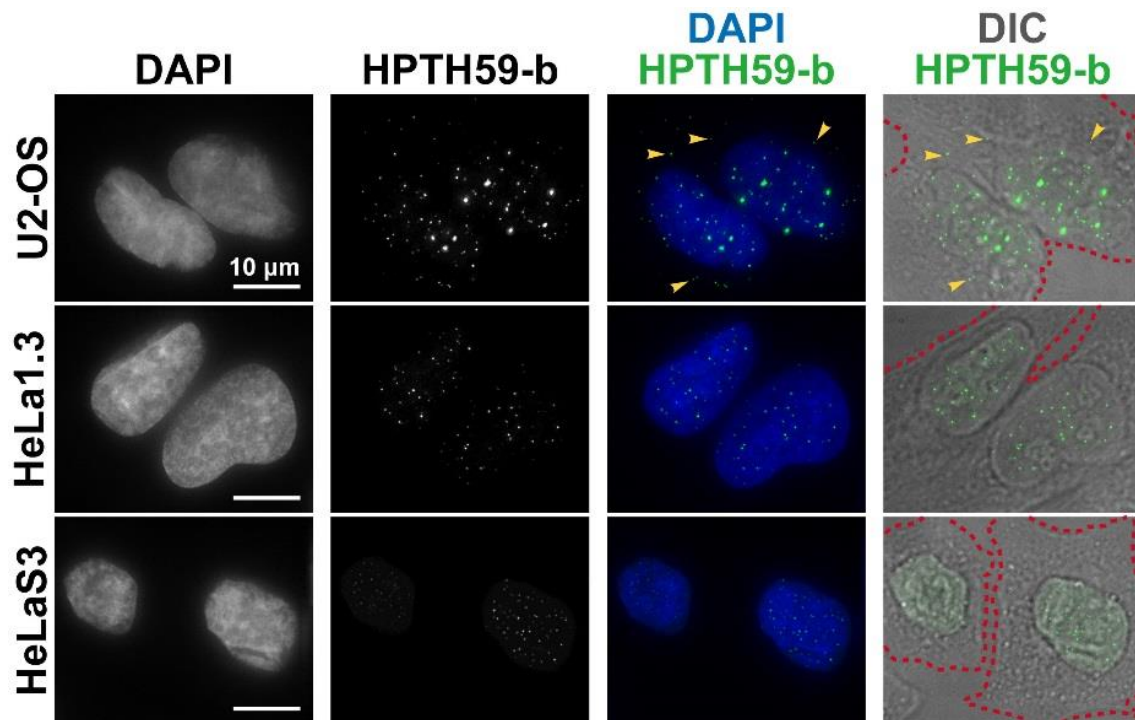
Figure 4. Telomere labeling with HPTH59-b in human cultured cells.

Telomerase-active cells (HeLa1.3 and HeLaS3) and alternative lengthening of telomere (ALT) cells (U2-OS) were stained with DAPI (first column) and HPTH59-b (second column). Merged images are in the third column. The fourth column shows merged images of HPTH59-b staining and differential interference contrast (DIC) images. Note that some U2-OS cells have extra-nuclear telomere signals (yellow arrowhead).
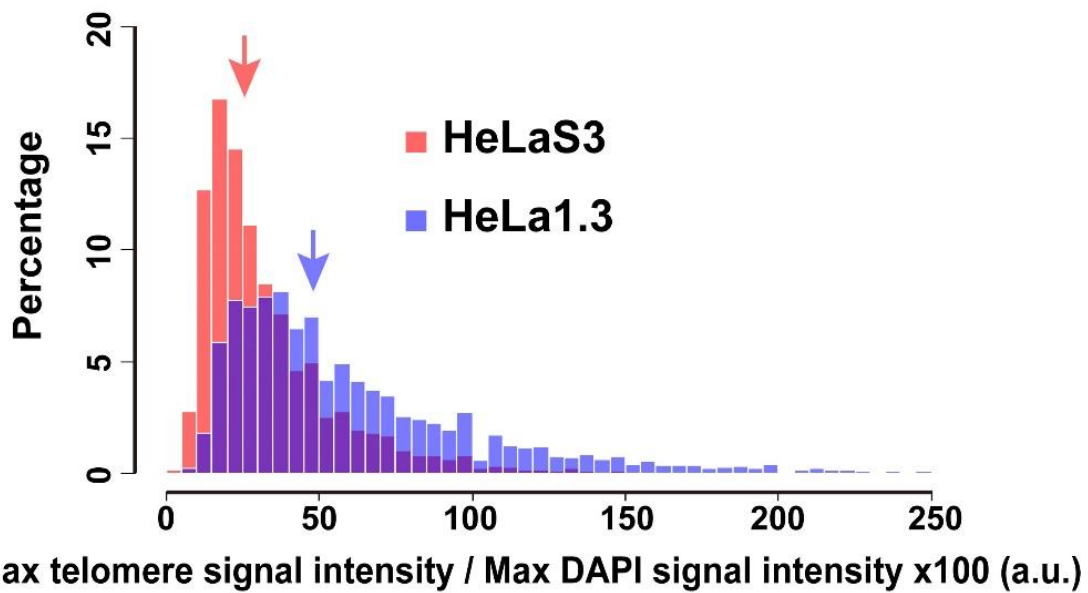
Figure 5. Distribution histograms of telomere signal intensities in HeLa1.3 and HeLaS3 cells. Signal intensities in HeLa1.3 (long telomere) and in HeLaS3 (short telomere) cells are shown in blue and red, respectively. The overlapping area of the two distributions is shown in purple. The numbers of telomere signals measured for HeLa1.3 and HeLaS3 cells are 13171 dots from 280 cells and 10805 dots from 277 cells, respectively. Median values of the intensity distributions for HeLa1.3 and HeLaS3 are 50.0 (blue arrow) and 26.7 (red arrow), respectively. To compare these median values, the Wilcoxon rank sum test was used (P<0.01).

As shown in Figure 4, HeLa1.3 cells exhibited more intense HPTH59-b signals than HeLaS3 cells. Quantitative measurement of their signal intensities normalized to the DAPI signal indicated that the telomere signals of HeLa1.3 cells were significantly higher than those of HeLaS3 (Figure 5). Moreover, quantitative measurement of telomere signals on mitotic chromosomal spreads also showed that the signals at every chromosomal end of HeLa1.3 cells were higher than those of HeLaS3 cells (Figure 6, 7), suggesting that the telomeres of HeLa1.3 are longer than those of HeLaS3 cells. In U2-OS cells, I observed the strongest telomere signals among the three cell lines (Figure 4). Additionally, U2-OS cells had several intense signals even outside the nucleus, which have been reported to be extra-chromosomal telomere DNA in the cytoplasm (Tokutake et al. 1998). These results suggested that the observed signal intensity of HPTH59-b can distinguish different telomere lengths among these cultured cells.
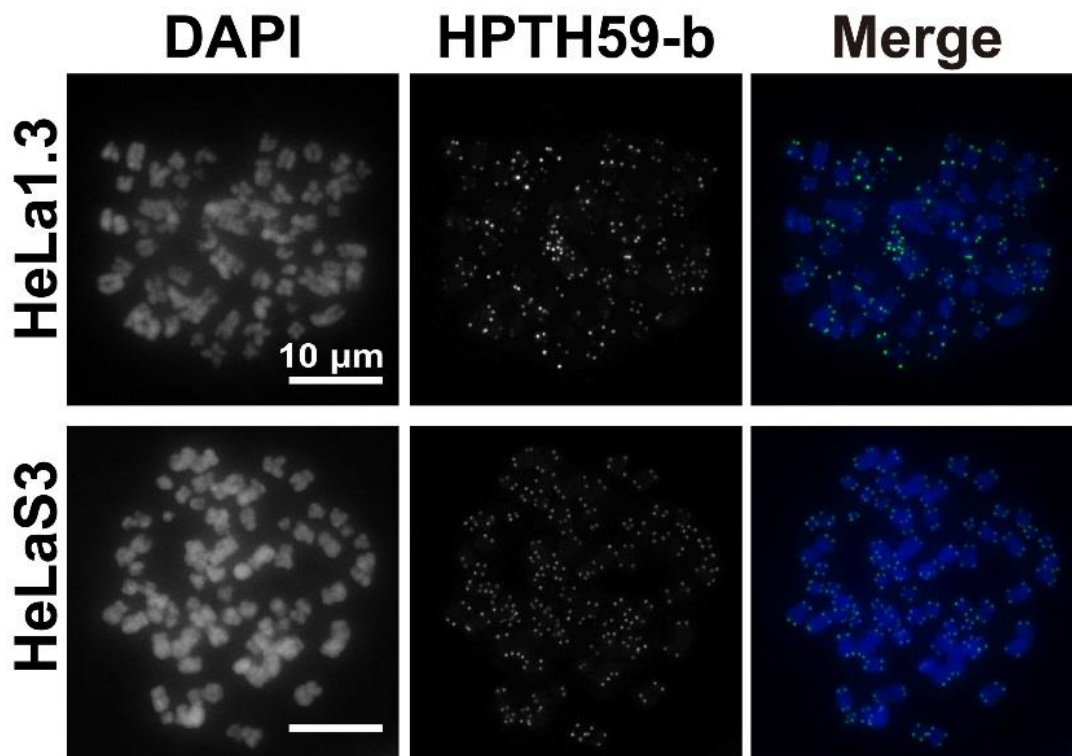
Figure 6. Telomere labeling in chromosomal spreads with HPTH59-b.

HeLa1.3 (1st row) and HeLaS3 (2nd row) cell spreads stained with DAPI (1st column) and HPTH59-b (2nd column). The 3rd column is the merged images.
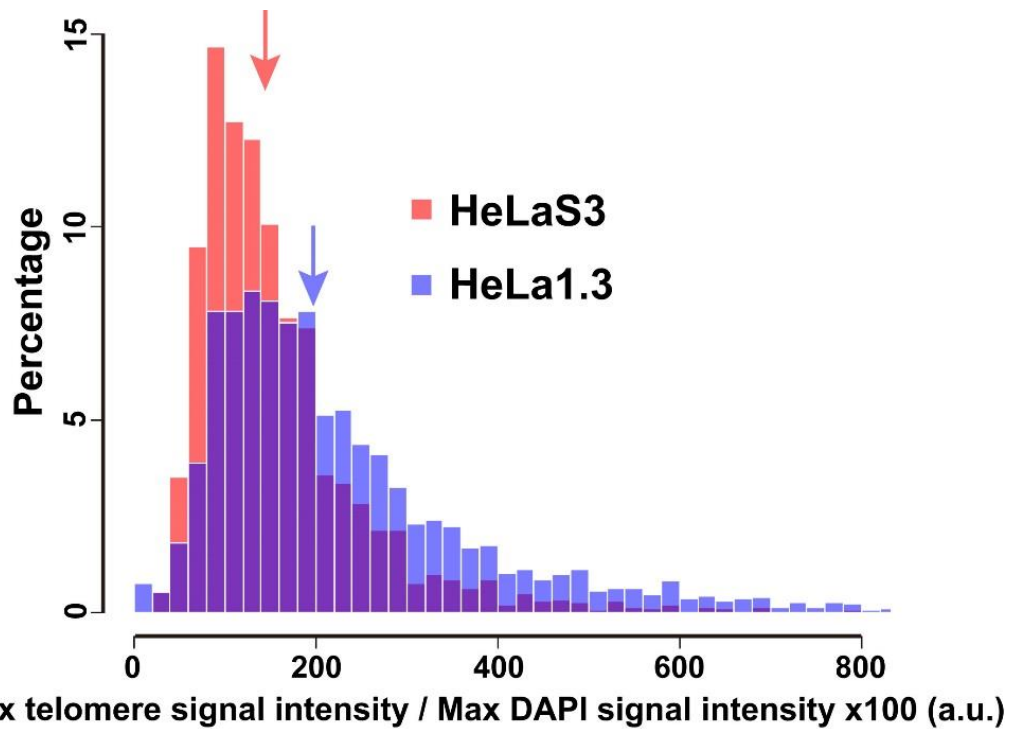
Figure 7. Distribution histograms of telomere signal intensities in HeLa1.3 and HeLaS3 cell spreads.

Telomere signal intensities in HeLa1.3 (6564 dots from 38 cells) and HeLaS3 (10895 dots from 43 cells) cell spreads are shown in blue and red, respectively.

Median values of signal intensities in non-tumor and tumor tissues are 194 (blue arrow) and 135 (red arrow), respectively. The overlapping area of the two distributions is shown in purple. To compare these median values, the Wilcoxon rank sum test was used (P<0.01).

Telomere labeling in tissue sections

Next, to examine whether the telomere labeling by HPTH59-b can be applied to cells in tissues, I prepared frozen tissue sections of a mouse (20-µm thickness) and stained them with fluorescent HPTH59-b (green). Images of lung and brain regions are presented in Figure 8. Intense HPTH59-b foci in the DAPI-staining DNA region were clearly detected in both tissues. To test the telomere-targeting efficiency of HPTH59-b, I stained mouse tissue sections with HPTH59-b and anti-TRF1 (telomere-binding protein) antibody simultaneously and compared the signals. I observed strong signals of HPTH59-b with low background noise, whereas TRF1 signals were weak with high background noise in the brain and lung (Figure 9A). Additionally, substantially more signals were detected by HPTH59-b than by TRF1 (Figure 9A, B). On the other hand, in cultured murine cells (MC12), the intensity and number of HPTH59-b signals were comparable to those of TRF1signals (Figure 9B). Thus, the HPTH59-b probe might label telomeres in tissue sections more efficiently than immunostaining.
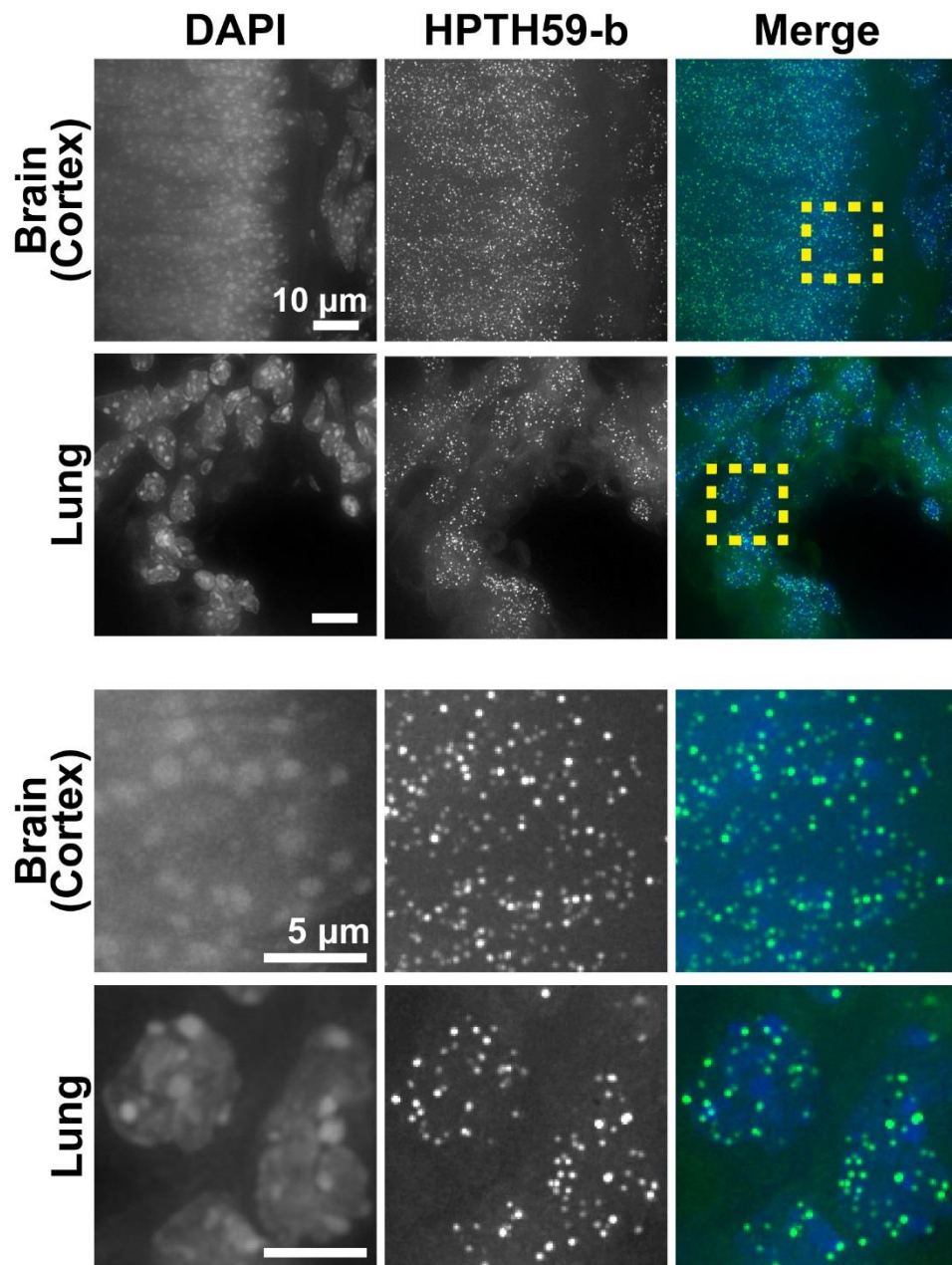
Figure 8. Co-staining with HPTH59-b and TRF-1 antibody in cultured cells and tissue sections. First column, DAPI signal; second column, HPTH59-b signal; third column, merged images of DAPI and HPTH59-b. Enlarged images of the boxed region in the upper part are shown in the lower part.
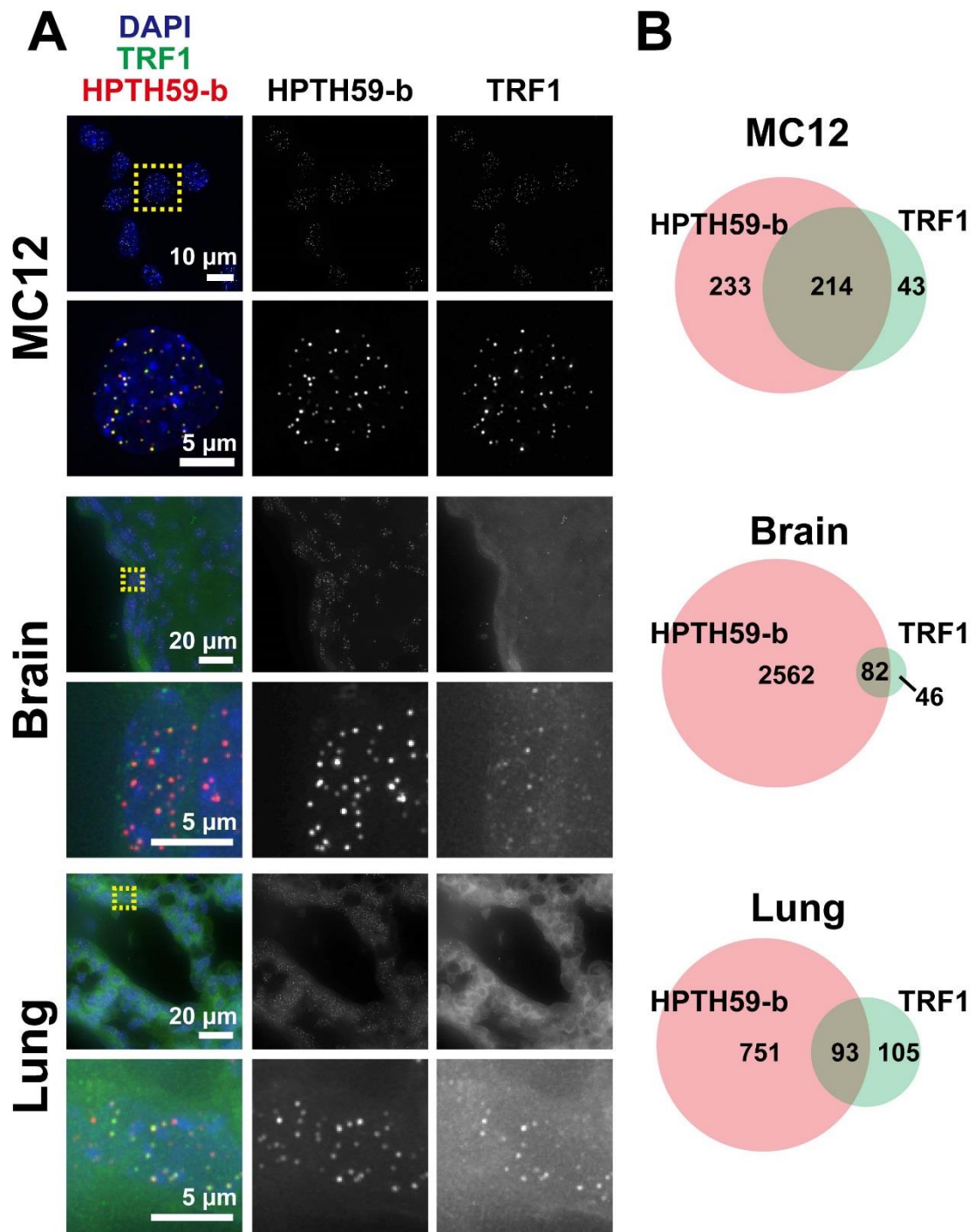
Figure 9. Co-staining with HPTH59-b and TRF-1 antibody in cultured cells and tissue sections. (A) Mouse embryonic carcinoma cells (MC12) and frozen mouse tissue sections stained with DAPI (blue), HPTH59-b (red) and TRF1 antibody (green). Enlarged images of the boxed region in the upper row are shown in the lower row. (B) Venn diagrams of telomere signals derived from HPTH59-b and TRF1. The number of dots was extracted from images in (A).

Knowing the telomere length of a specific cell population in a tissue, especially at the

single-cell level, is very useful. To explore telomere length in a specific cell type, I

focused on the telomere length in a germ cell line, which has high telomerase activity.

DEAD box proteins (DDX), putative RNA helicases, are specifically and highly

expressed in the germ cell lineage in both sexes, and they are widely used as a marker of

germ cell lineage (Fujiwara et al. 1994; Toyooka et al. 2000). Testis sections containing

gonadal tissues, where primordial germ cells (PGCs) are located, were treated with

HPTH59-b (red) and anti-DDX4/MVH antibody (green). Immunostaining with anti-

DDX4/MVH antibody specifically labeled the cytosol of PGCs, and HPTH59-b showed

clear foci in both the germ cells and somatic cells (Figure 10). The telomere signals in

PGCs seemed to be weaker than those in PGC marker-negative cells. The PGC marker-

positive cells showed a unique characteristic of nuclear organization: less DAPI staining

over the nucleus, which is consistent with a previous report (Yoshioka, McCarrey, and

Yamazaki 2009). The quantitative measurement of telomere signals normalized to the

DAPI signal suggested that PGCs had slightly longer telomeres than the PGC-negative

cells (Figure 11). These results demonstrated that HPTH59-b highlighted telomeres and

the antibody marked PGCs simultaneously, thereby demonstrating that double staining

with HPTH59-b and antibodies can be used to label telomeres in a specific cell lineage

in tissue.
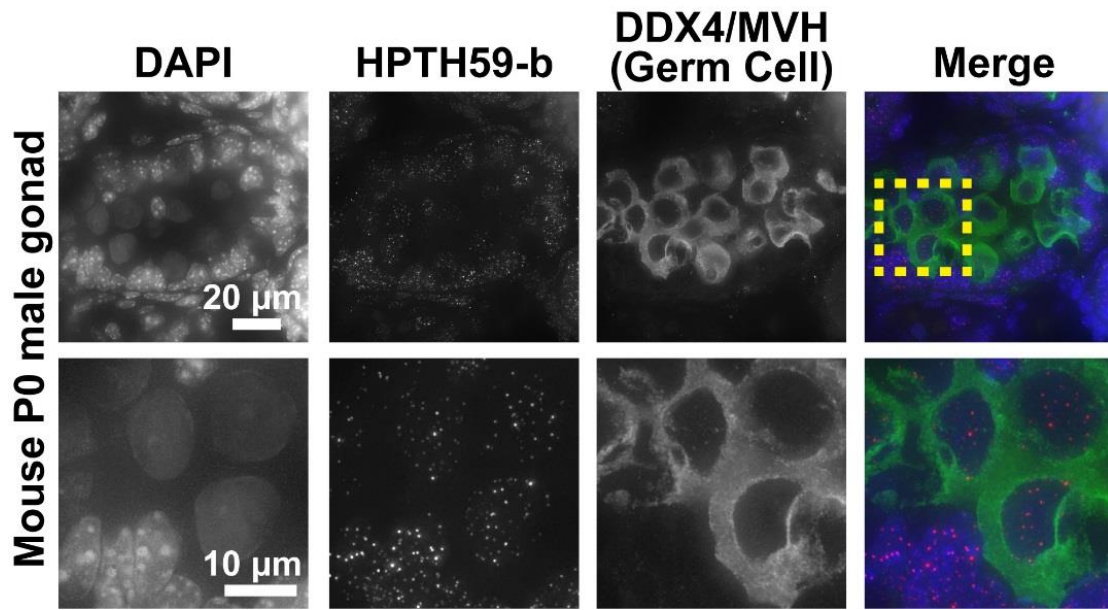
Figure 10. Telomere labelng in mouse germ cells and somatic cells.

Neonatal (P0) mouse gonads stained with DAPI (blue), anti DDX4/MVH antibody (germ cell marker; green), and HPTH59-b (red). Enlarged images of the boxed region in the first-row images are shown in the second row.
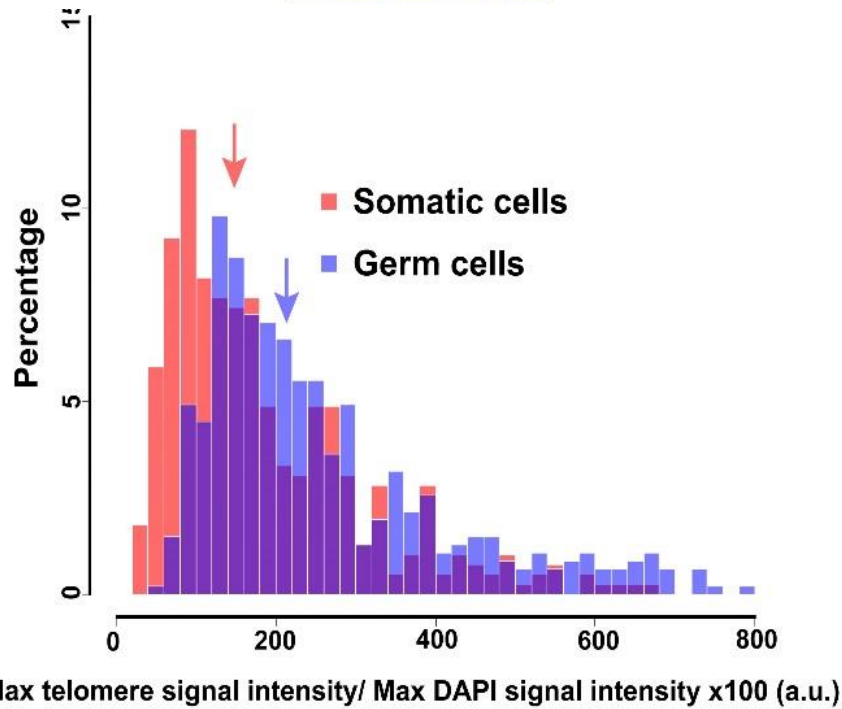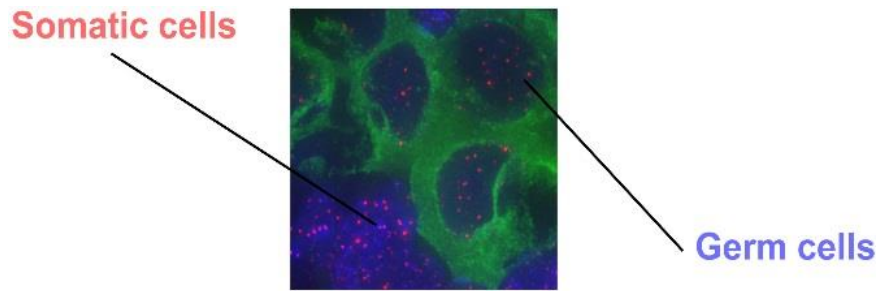
Figure 11. Distribution histograms of telomere signal intensities in germ cells and somatic cells. I analyzed 19 cells (469 dots) and 19 cells (390 dots) in germ (blue) and somatic (red) cells, respectively. The identical image containing PGCs in Figure 10 is shown at the top. Median values of signal intensities in germ and somatic cells are 220 (blue arrow) and 152 (red arrow), respectively. The overlapping area of the two distributions is shown in purple. To compare these median values, the Wilcoxon rank sum test was used (P<0.01).

Shorter telomeres in human tumor cells

Telomere length is strongly connected to cell immortalization and tumorigenesis. Indeed, telomere shortening has been observed in carcinoma derived from bladder, esophageal, gastric, head and neck, ovarian, and renal cells (Wentzensen et al. 2011; Zhang et al. 2015). To further investigate telomeres in human neoplastic cells, I performed simultaneous labeling in esophageal cancer tissue using HPTH59-b and an antibody for a proliferation marker Ki-67, which is highly expressed in cells that are actively dividing, but it is absent in cells under a quiescent state, such as cells in the G0 phase of the cell cycle (Schluter et al. 1993). I found that some cell fractions in the lesions were still Ki-67-positive (green) and the fluorescent intensity of telomere foci appeared lower than that of Ki-67-negative cells (Figure 12). Using digital image analysis, I quantitatively compared telomere signals (Figure 13). The distribution map of the telomere signals revealed two distinct peaks between tumor and non-tumor tissues. These results suggest that highly proliferating tumor cells have shorter telomeres than non-dividing cells, and HPTH59-b can detect differences in telomere length between tumor and non-tumor cells in the tissue.
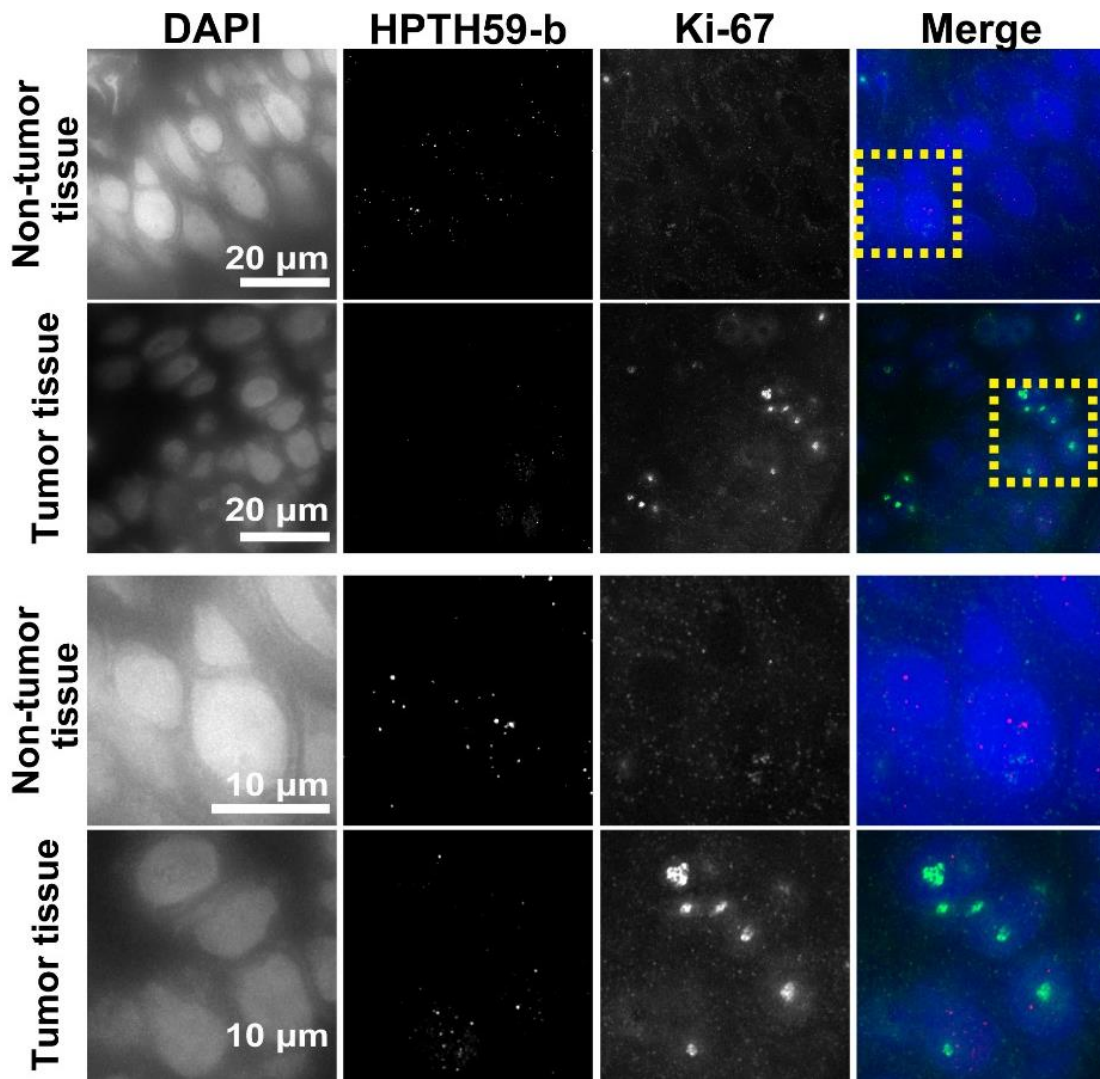
Figure 12. Different telomere lengths between human tumor and non-tumor tissue sections. Frozen sections of esophageal tumor/non-tumor tissue stained with DAPI (blue), anti-Ki-67 (growth marker; green) antibody, and HPTH59-b (red). Enlarged images of the boxed region in the upper images are shown in the lower part.
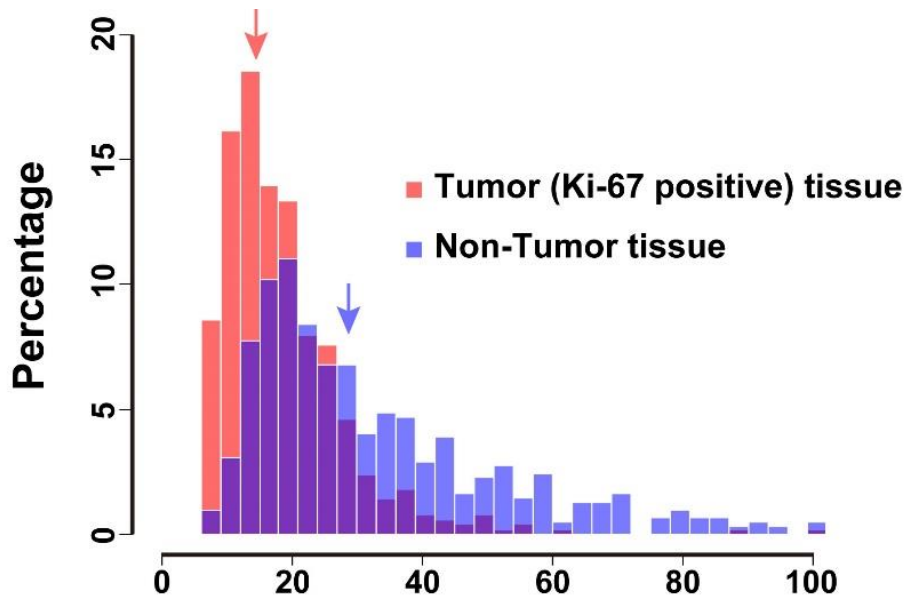
Figure 13. Distribution histograms of telomere signal intensities in tumor and non-tumor tissue sections.

Telomere signal intensities in tumor (511 dots from 54cells) and non-tumor (635 dots from 50 cells) are shown in red and blue, respectively. Median values of the signal intensities in non-tumor and tumor tissues are 27.8 (blue arrow) and 16.7 (red arrow), respectively. The overlapping area of the two distributions is shown in purple. To compare these median values, the Wilcoxon rank sum test was used (P<0.01).

1.3. Discussion

To demonstrate and verify the application of PI polyamide to clinical studies, I used PI polyamide HPTH59-b, along with immunostaining, to visualize telomeres in mouse and human frozen tissue sections. My quantitative analysis of the telomere signals also suggested that highly proliferating cells in tumor tissue had shorter telomeres than non-dividing cells in non-tumor tissue. Although the issue of whether telomere shortening is a consequence of cell proliferation during tumor expansion or a cause of tumorigenesis initiation is controversial, HPTH59-b provides a simple and quick detection method for telomere alteration at the single-cell level, thereby contributing to the genetic diagnosis of malignancy and drug response in patients.

Telomere visualization by HPTH59-b has several advantages over the Q-FISH method. First, the sample preparation process is much faster. Even with tissue sections, incubation with HPTH59-b for just 1 hour is sufficient for staining telomeres. Moreover, even when co-staining with HPTH59-b and an antibody for cell-specific markers, the whole procedure can be completed within 4.5 hour. The second advantage is high sensitivity. As shown in Figure 9A, HPTH59-b generated more telomere signals with lower background noise than did immunostaining with a TRF1 antibody, whereas with cultured cells, HPTH59-b showed a staining efficiency comparable to that of a TRF1 antibody. In the case of thicker tissue sections, the probe size might be critical to accessing telomeres, and HPTH59-b (2.67 kDa) is much smaller than antibodies (~150 kDa). The third advantage is that HPTH59-b labeling can be carried out under mild conditions. This probe can also detect telomere sequences in small DNA molecules in the cytoplasm of ALT cells, which have extra-chromosomal circular DNA of telomeric

repeats due to the high activity of recombination (Dunham et al. 2000; Nabetani and Ishikawa 2011). Only a few studies using the conventional FISH method have ever reported the extra-chromosomal DNA in cytoplasm (Tokutake et al. 1998). This is probably because, unlike large genomic DNA, such small circular plasmid-like DNAs may be easily washed away during the harsh hybridization process and/or easily re-annealed, preventing the nucleic acid probes from hybridization. In contrast, HPTH59-b binds to dsDNA TTAGGG repeats without denaturation, and it may have a similar labeling efficiency for the extra-chromosomal and chromosomal telomere repeats.

Two new telomere labeling methods that use genome editing systems have recently been reported. One involves telomere labeling using a transcription activator-like effector (TALE)-based strategy in both fixed and living mammalian cells (Ma, Reyes-Gutierrez, and Pederson 2013; Miyanari, Ziegler-Birling, and Torres-Padilla 2013). Although the TALE-based method is applicable to live-cell imaging, it is time consuming (i.e., for plasmid construction and establishment of stable cell lines expressing the fluorescent-TALE). Moreover, the TALE-based strategy cannot be applied to telomere labeling in tissue sections, especially to clinical samples from patients. On the other hand, HPTH59-b can label telomeres in both cultured cells and tissue sections without constructing plasmids and establishing stable cell lines.

Another recent approach uses the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated caspase 9 (Cas9) technique (Chen et al. 2013). Recently, Deng et al. reported the use of in vitro constituted nuclease-deficient CRISPR/Cas9 complexes as probes (Cas9-mediated fluorescence in situ hybridization,

CASFISH) (Deng et al. 2015). Telomere labeling by this method does not require DNA denaturation and can quickly (15 min) label telomeres in cultured cells and tissue sections. However, the production cost of large amounts of single-guide RNAs (sgRNAs) and dCas9 (nuclease-deficient) protein make it more expensive than the HPTH59-b method, which could be problematic for high-throughput applications, such as those involved in cancer diagnosis.

With the recent development of new technologies, our understanding of chromatin structure and dynamics is deepening (Maeshima et al. 2016). Because our sensitive telomere labeling method can be performed under mild conditions, another interesting application to telomere regions involves super-resolution imaging without harsh treatments. This technique could help to elucidate how telomere chromatin is organized in the cell nuclei. Therefore, telomere visualization using the PI polyamide-based approach discussed here would contribute to telomere biology and related medical science.

1.4. Material and Methods

Synthesis

HPTH59-b was synthesized as reported previously (Hirata et al. 2014).

Human tissues

The use of human tissues was approved by the committees of the National Center for Global Health and Medicine (#NCGM-G-001766-00) and was in accordance with the Declaration of Helsinki of the World Medical Association. Participants provided written informed consent. I analyzed one case of a patient who had a surgical operation at the National Center for Global Health and Medicine Hospital. Tissues were prepared from regions diagnosed as esophageal squamous cell carcinoma and adjacent normal tissue according to the manufacturer's protocol. Briefly, excised tissue was flash frozen in cold acetone with optimal cutting temperature (OCT) compound (Sakura Finetek Japan). Tissue sections (10 µm) were prepared by microtome and placed on the slide glass for telomere staining.

Telomere staining of HeLaS3, HeLa1.3, and U2-OS cells with HPTH59-b

HeLa cells were maintained at 37°C under 5% CO_2 atmosphere in DMEM containing 10% fetal bovine serum (FBS). For polyamide staining, cells were grown on coverslips coated with poly-lysine. The cell coverslips were washed in phosphate-buffered saline (PBS) twice and fixed with 1.85% formaldehyde in PBS. The fixed cells on coverslips were stained with HPTH59-b and then mounted as described previously (Kawamoto et al. 2013). Section images were recorded with a DeltaVision microscope and deconvolved to eliminate out-of-focus blur to obtain clearer pictures. The deconvolved

26

images were projected ('Quick Projection' tool) to obtain the maximum intensity of telomere signals. Non-deconvolved pictures were used for quantitative analysis of HPTH59-b signals.

Telomere staining of mouse tissue sections with HPTH59-b

Postnatal 0 (P0) mice were fixed in 4% paraformaldehyde in PBS overnight at 4°C and placed in 30% sucrose in PBS for 1 day at 4°C. Mouse whole bodies were then embedded in optimal cutting temperature (OCT) compound /30% sucrose (2:1) and incubated for 1 h at room temperature. Embedded samples were stored at -80°C until sectioning. Samples were sectioned (20 µm) using a CM3050S cryostat (Leica) and kept at -30°C until use. All experimental protocols were approved by the Animal Committee of the National Institute of Genetics and carried out according to the guidelines to minimize the pain and discomfort of the animals.

Preparation of human tumor/non-tumor tissue sections

Before staining, sections were incubated in HEN buffer (10 mM HEPES pH 7.5, 1 mM EDTA, 100 mM NaCl) overnight at 37°C. The sections were permeabilized with 0.1% Triton X-100 for 10 min and briefly washed twice with TEN (10 mM Tris-HCl pH 7.5, 1 mM EDTA, 100 mM NaCl). For blocking, the sections were treated with Normal Goat Serum (NGS) in TE buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA) for 30 min at room temperature. After a brief rinse with TE buffer, the sections were incubated with 10% NGS, 15 nM HPTH59-b, and 0.5 µg/mL DAPI for 2 h at 37°C. After washing five times for 3 min with TEN200 buffer (10 mM Tris-HCl pH7.5, 1 mM EDTA, and 200

27

mM NaCl), the sections were mounted and image acquisition was performed as described above.

Co-staining with polyamide and antibody

The procedure from sample preparation to blocking was the same as that described above. The sections were incubated with 10% NGS, 15 nM HPTH59-b, 0.5 µg/mL DAPI, and either anti-DDX4/MVH antibody (Abcam #ab13840, 1:500 dilution) or anti-Ki-67 antibody (Oncogene #NA59, 1:20 dilution) for 2 h at 37°C (Maeshima et al. 2006), following the process from washing to image acquisition as described above.

Quantification of telomere signals

Images of HeLa1.3 and HeLaS3 cells and human/mouse tissue sections were recorded with a DeltaVision microscope under identical conditions. Non-deconvolved images were projected and used as source images. Extraction of each telomere signal was performed as follows. The background fluorescent signals of HPTH59-b and DAPI in the source images were subtracted using the 'Rolling ball' tool (Fiji) (Schneider, Rasband, and Eliceiri 2012). By setting an arbitrary threshold value of HPTH59-b intensity, the telomere spots were contoured. The maximal signal intensities of HPTH59-b and DAPI were extracted from each telomere spot, and each HPTH59-b signal was then normalized with DAPI signals. The Wilcoxon rank sum test was used for statistical analysis, which was performed with R software (R core Team 2016).

Chapter 2

Capture of Locus-specific Chromatin to Identify Non-coding RNAs using a Pyrrole–Imidazole Polyamide Probe

2.1. Introduction

Telomeres protect the chromosome ends so that the ends are not recognized as DNA double-strand breaks (DSB) and also do not trigger DNA damage response (DDR). Telomere maintenance mechanisms involve three type of non-coding RNAs (ncRNAs). First, telomerase RNA component (TERC) is a classic trans-acting RNA that serves as the template and scaffold for the telomerase complex that elongates telomere (Blackburn and Collins 2011; Zappulla and Cech 2006). Second, telomeric repeat-containing RNA (TERRA) is transcribed from subtelomeric region to the telomere and coats the chromosome ends (Azzalin et al. 2007). Although some roles of TERRA in the maintenance of telomere length, heterochromatin stabilization, and homologous recombination were proposed (Arora et al. 2014; Arora and Azzalin 2015 ;Chu et al. 2017), its precise function remains enigmatic. Third, telomeric DNA damage response RNAs (tDDRNAs) are small double-stranded telomeric ncRNAs and are induced by telomere uncapping for DDR activation (Rossiello et al. 2017). While these ncRNAs are recognized as key regulators of the chromatin states for telomere maintenance, the comprehensive and unbiased identification of ncRNAs associated with telomeres have been technically challenging.

29

For comprehensive analysis of proteins associated with telomeres, a locus-specific chromatin purification called PICh was developed (Dejardin and Kingston 2009). In PICh, a locked nucleic acid (LNA) probe is used to capture telomeric chromatin through Watson–Crick base pairs to telomeric DNA sequence and the telomere-bound proteins are identified by mass spectrometry. As another method, quantitative telomeric chromatin isolation protocol (QTIP) using the antibodies against telomere binding proteins also revealed telomere proteins in various cell states (Grolimund et al. 2013). In addition to the LNA and the antibody probes, recent development of engineered DNA-binding molecules including TALE and CRISPR provided alternative methods for telomere chromatin purification. These systems, called enChIP (Fujita et al. 2015) and CAPTURE (Liu et al. 2017) identified telomere-specific chromatin-regulating protein complexes and ncRNAs, but overall enrichment rates of the telomeric proteins such as shelterin complex in the purified chromatin fractions were less than 10-fold. For accurate quantification and comprehensive identification of chromatin-associated RNAs, a methodological breakthrough is required.

Related to my research shown in Chapter 1, we and our collaborators have previously developed N-methylpyrrole (P)−N-methylimidazole (I) (PI) polyamides for the visualization of telomere and the assessment of telomere length (Maeshima, Janssen, and Laemmli 2001; Kawamoto et al. 2013; Hirata et al. 2014; Kawamoto et al. 2015; Kawamoto et al. 2016). PI polyamide binds to the minor groove of double-stranded DNA without denaturation and can recognize Watson–Crick base pairs (Trauger, Baird, and Dervan 1996; White et al. 1998; Chenoweth and Dervan 2009; Dervan 2001; Dervan and Edelson 2003; Dervan, Doss, and Marques 2005; Bando and Sugiyama

2006; Blackledge and Melander 2013). PI polyamide can be an alternative to the nucleic acid probes in any genomic analyses. I have demonstrated in chapter 1 that fluorescent labeled tandem hairpin PI polyamides that target human telomere sequences $(TTAGGG)_n$ visualized telomeres in cultured cells and tissue sections in a mild condition without denaturation (Maeshima, Janssen, and Laemmli 2001;Kawamoto et al. 2013).

Here, I show a novel approach of locus-specific chromatin purification with a telomeric-targeting PI polyamide, named PI polyamide-based proteomics and RNA-omics of isolated chromatin segments (PI-PRICh). By using PI-PRICh with mouse erythrocytes leukemia (MEL) cells, I was able to identify proteins bound to telomeres such as shelterin complex (TRF1, TRF2, POT1, and TIN2) using mass spectrometric analysis. At the same time, I also extracted RNAs associated with telomeric chromatin, which were subsequently analyzed by next generation-sequencing (NGS) for comprehensive identification of telomeric chromatin-associated RNAs. Using PI-PRICh, I obtained the following 3 points: First, TERC is the most enriched (>500-folds). Second, the telomeric sequence reads including TERRA occupied over 10% of the total reads. Finally, other ncRNAs transcribed from subtelomeric regions are >100-fold enriched. To identify telomeric ncRNAs in human ALT cells, whose telomeres are the highly recombinogenic, I compared identified ncRNAs in ALT cells with those in telomerase positive cells. I found that several intronic ncRNAs specifically appeared at ALT telomeres. A possible physiological role of the intronic ncRNAs will be discussed. PI polyamides are a promising alternative for identification of ncRNAs to investigate particular chromatin regions.

2.2. Results

Affinity purification of telomeric chromatin by a PI polyamide probe

I developed a method, PI-PRICh, which allows unbiased high-throughput identifications of telomeric chromatin-bound proteins and RNAs (Figure 14). Briefly, cultured cells were crosslinked with formaldehyde and their chromatin was extracted and homogenized. Telomere targeting PI polyamide, which has a biotin analogue for affinity purification (TH59-DB, Figure 15), was bound to telomeric chromatin. The TH59-DB-bound chromatin was isolated using magnetic streptavidin beads. The co-purified proteins and RNAs were eluted and then subjected to downstream assays for identification and quantitation. Using streptavidin covalently attached to a fluorescent dye, I confirmed that TH59-DB targeted the telomere regions in the crosslinked cells, which were immunostained by TRF2 antibody (Figure 16). Moreover, by TH59-DB, I succeeded in an efficient pull-down of a plasmid containing a 750-bp telomeric repeats fragment, whereas the empty plasmid was not retrieved (Figure17, 18). These results indicated that PI polyamide is competent for affinity purification of telomeric chromatin.
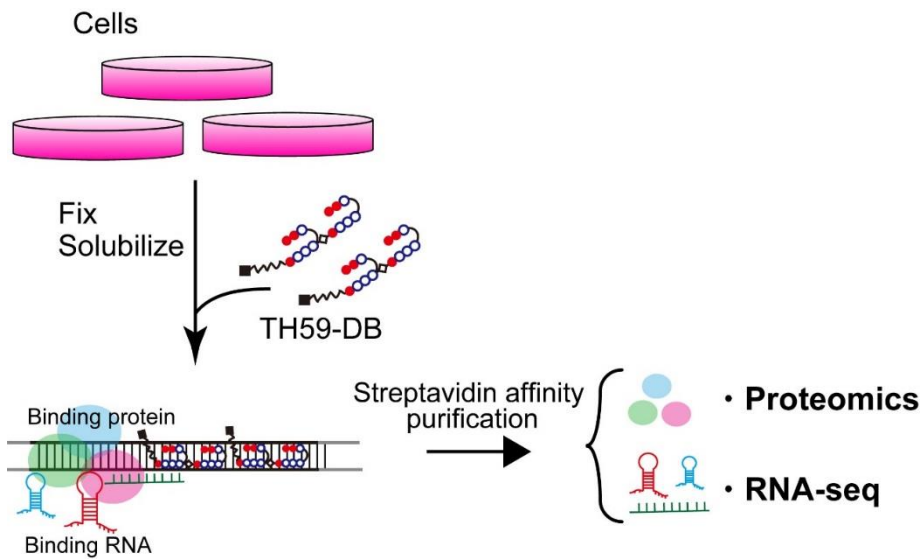
Figure 14. Scheme for telomeric chromatin isolation using telomere-targeting PI polyamide probe (TH59-DB).

Cells are crosslinked with formaldehyde. The chromatin is solubilized using a French press. The chromatin is mixed and incubated with TH59-DB, and the probe-chromatin complexes are isolated by streptavidin affinity purification. The isolated chromatin fractions are analyzed by mass spectrometry for the study of proteins and by next-generation sequencing for the study of ncRNAs, respectively.
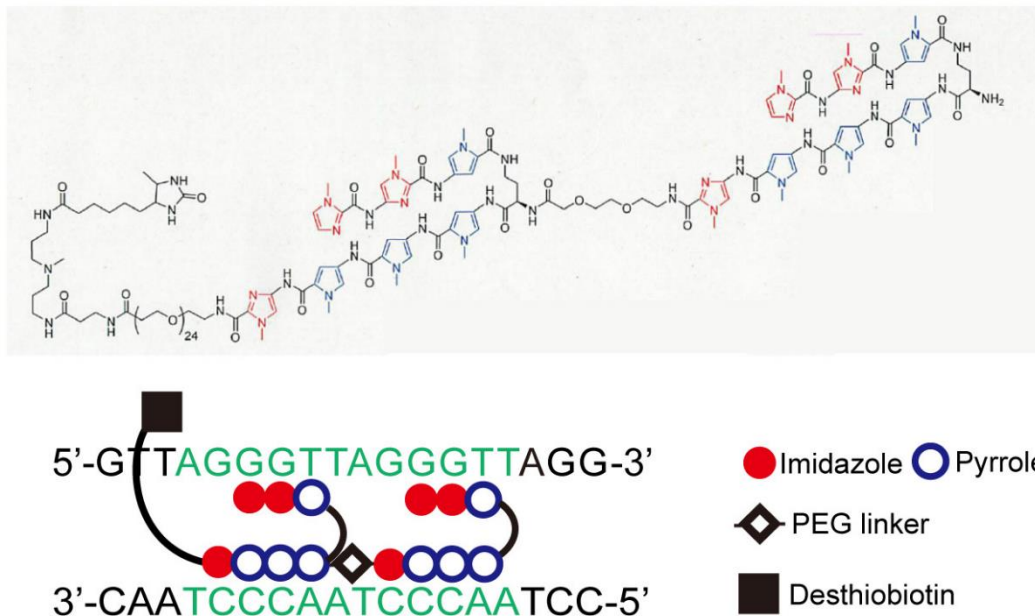
Figure 15. Chemical structure of TH59-DB.

Base recognition profile of TH59-DB is shown in lower part.

33
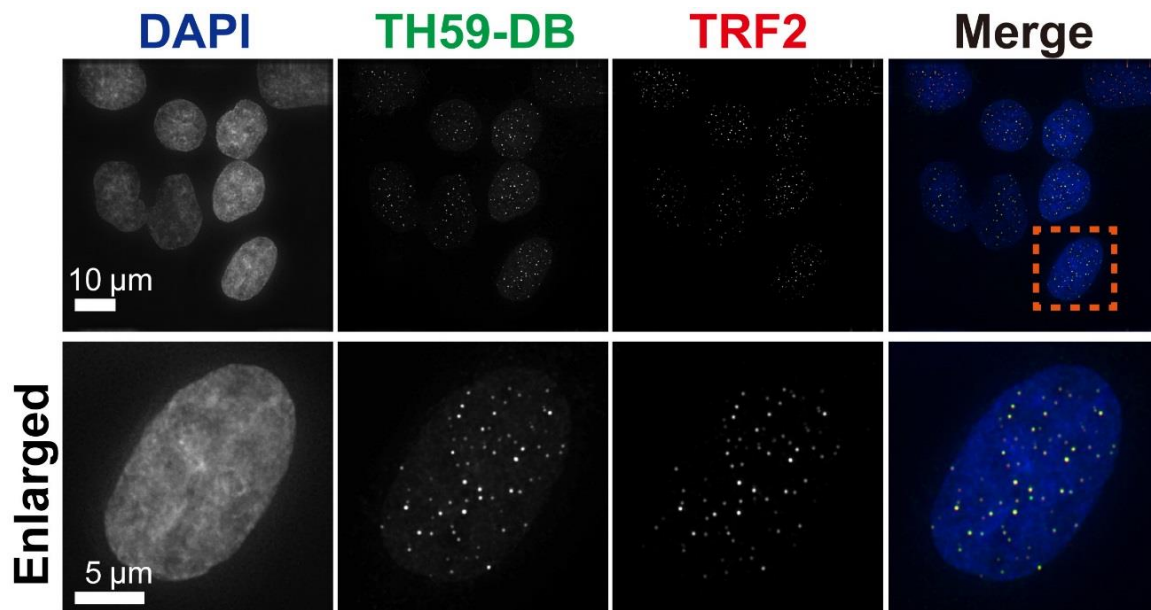
Figure 16. Telomere labeling with TH59-DB in HeLa1.3 cells.

Cells were stained with DAPI (first column), TH59-DB (second column) and anti-TRF2 antibody (third column). The merged images are in the fourth column.
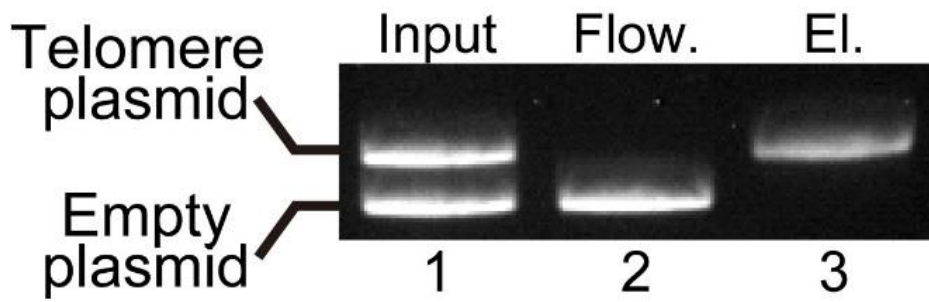
Figure 17. Purification of the telomeric repeat-containing plasmid with TH59-DB.
Each fraction (Input, Flow-through (Flow.), and Elution (El.)) was analyzed by agarose gel electrophoresis and EtBr staining. The positions of telomeric repeat-containing plasmid and the empty vector are indicated.
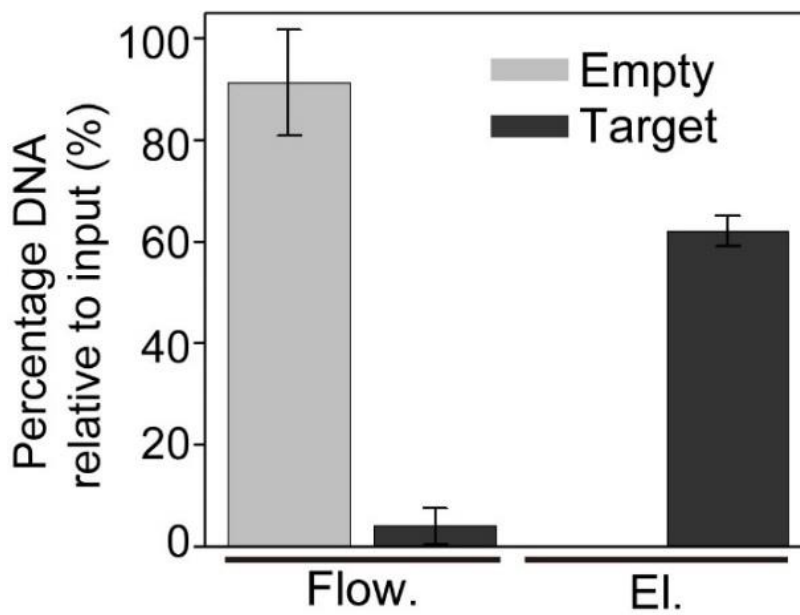
Figure 18. Bar graph quantifying telomeric DNA capture shown in Figure 17.
Error bars represent standard deviations.

I next performed telomeric chromatin isolation from mouse MEL cells using TH59-DB. As a negative control probe, I used the TH59-DB premixed with telomeric oligo DNA, prior to incubation with the chromatin extract (Figure 19). Actually, the masked TH59-DB failed to capture telomeric DNA in the plasmid pull-down assay (Figure 20, 21). After incubation of TH59-DB with solubilized MEL-derived chromatin and subsequent washing process, the pull-down fractions were electrophoresed and analyzed by silver staining for protein analysis. Several specific bands were indeed detected in the TH59-DB pull-down fraction (Figure 22). The band pattern was quite similar to that of telomeric proteins purified by PICh based on LNA (Figure 23, 24). To verify specific enrichment of telomeric proteins, I monitored the presence of the known telomere-associating protein TRF1 by immunoblotting (Figure 25). TRF1 was highly enriched in the TH59-DB pull-down fraction but not in the masked TH59-DB pull-down (negative control) and input fractions (Figure 25). Importantly, by mass spectrometry analysis, I identified the known telomeric proteins such as the components of the shelterin and chromosome passenger complexes (Figure 26). These data showed that the TH59-DB bound to telomeric sequence specifically and efficiently purified telomeric chromatin.
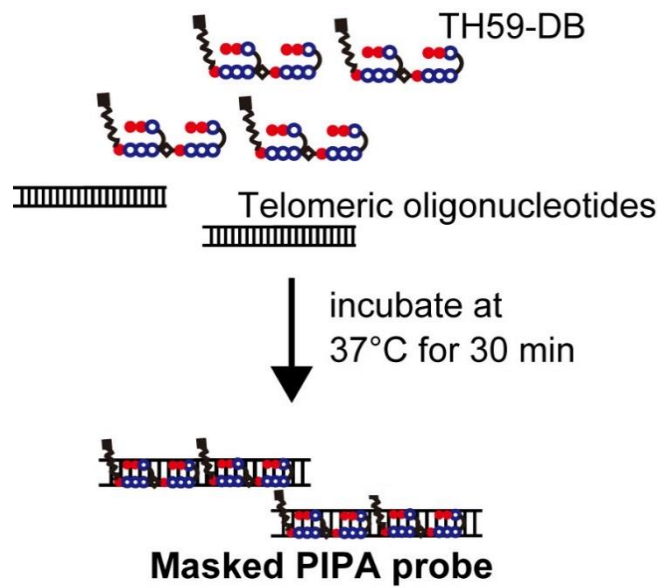
Figure 19. Preparation of a negative control probe for telomeric chromatin isolation.
TH59-DB is mixed and incubated with double-stranded oligonucleotide $(TTAGGG)_4$
$/(CCCTAA)_4$ (telomeric oligonucleotides) prior to incubation with telomeric repeat containing
plasmid. The oligonucleotide- premixed TH59-DB is named as "masked TH59-DB".
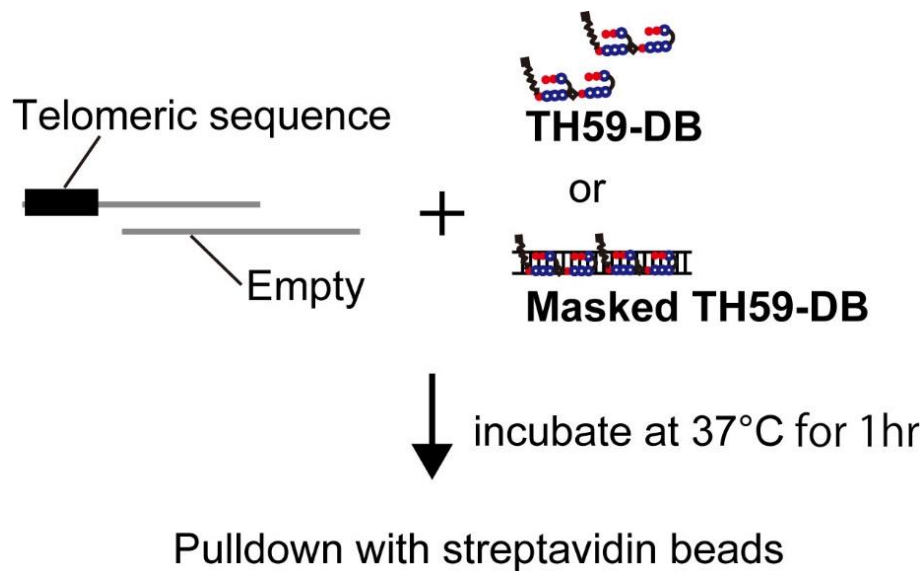
Figure 20. Outline of a plasmid pull-down assay.
Linearized plasmids (gray line) with or without the telomeric repeat (black thick line) are mixed
with TH59-DB or masked TH59-DB. The mixture is incubated at 37°C for binding of TH59-DB.
Plasmid-TH59-DB hybrids are captured using MyOne C1 streptavidin beads.
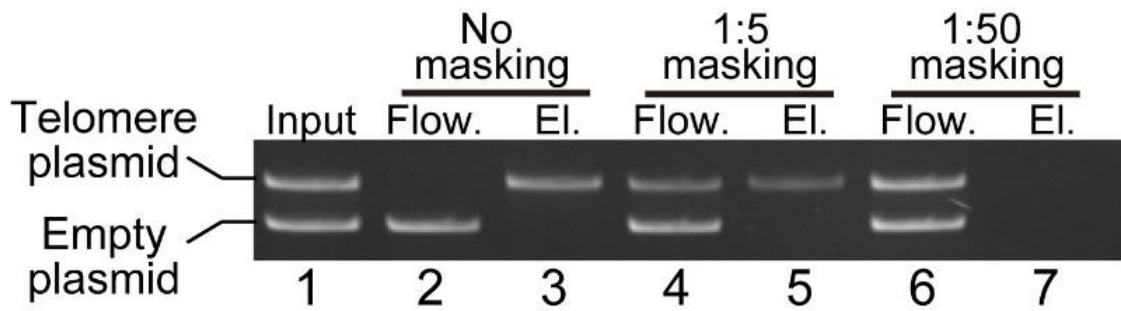
Figure 21. Purification of the telomeric repeat containing plasmid with TH59-DB or masked TH59-DB.

Telomeric repeat-containing plasmids were purified with TH59-DB (lane 2, 3) or masked TH59-DB with 5 times the amount of the telomeric oligonucleotide (lane 4, 5) or with 50 times the amount of the oligonucleotide (lane 6, 7). Each fraction (Input, Flow-through (Flow.), Eluate (El.)) was analyzed by agarose gel electrophoresis and EtBr staining. The positions of the telomeric repeat-containing plasmid and the empty vector are indicated.
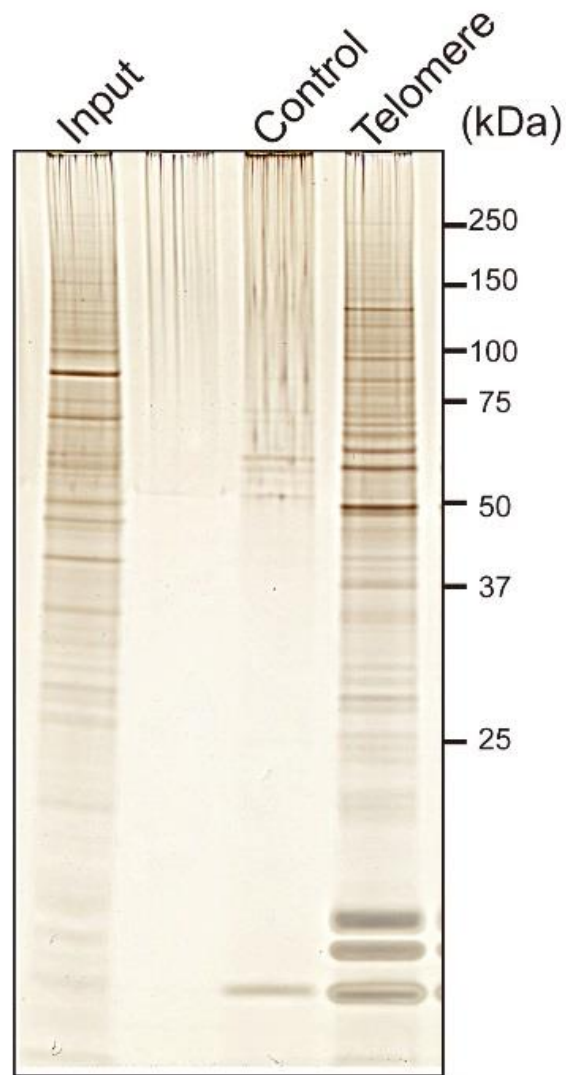
Figure 22. Silver staining of proteins obtained from telomeric chromatin purification with TH59-DB.

Chromatin isolation procedure were performed with TH59-DB and masked TH59-DB (See Figure 20, 21). Input represents 0.001% of the starting material (10^4 cells equivalent). 8 % of the materials of the TH59 pull-down fraction (Telomere) and masked TH59-DB pull-down fraction (Control) were analyzed.
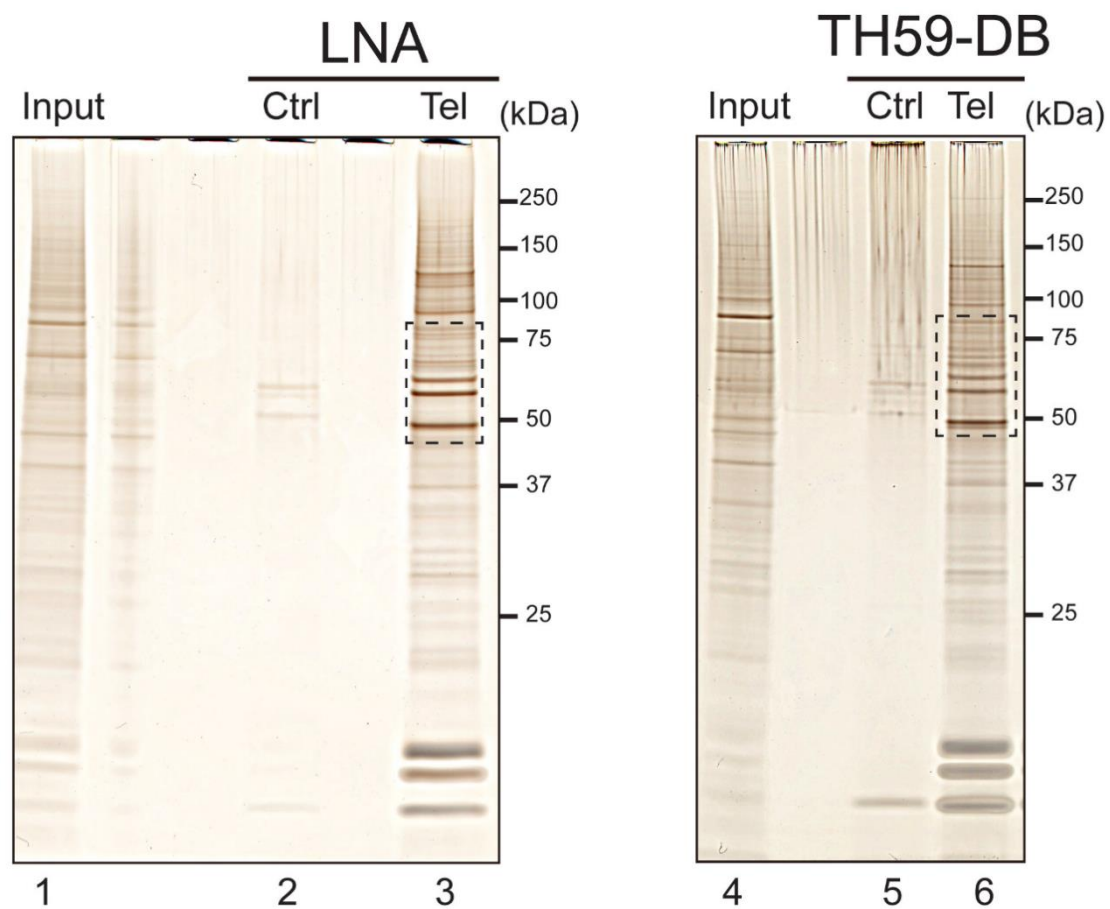
Figure 23. Comparison of silver staining between PICh and PI-PRICh.

Silver staining of proteins obtained from PICh with scrambled (lane 2, Ctrl) and telomere LNA probe (lane 3, Tel), and obtained from the telomeric chromatin isolation with masked TH59-DB (lane 5, Ctrl) and TH59-DB (lane 6, Tel) in mouse erythrocytes leukaemia (MEL) cells. Input represents 0.001% of the starting material (lane 1 & lane 4). The right panel is identical to Figure 22.
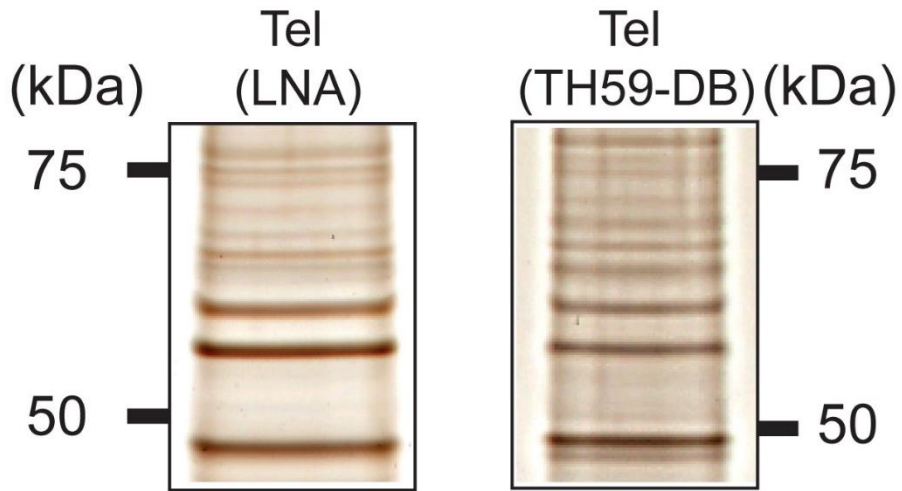
Figure 24. Enlarged images of the boxed region shown in Figure 23.

Enlarged images of the boxed region between 50 kDa and 70 kDa in the left and right panel are shown in Figure 23 to compare the band patterns.
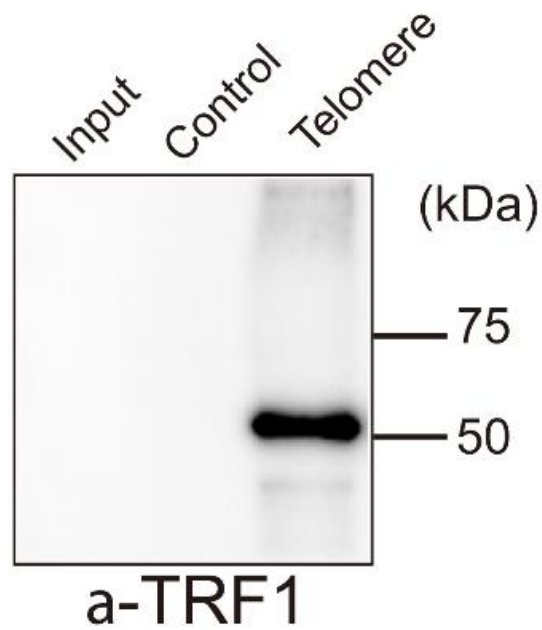
Figure 25. Western blot analysis for TRF1 in each fraction of PI-PRICh.

Input represents 0.0005% of the chromatin extracts. 4 % of the materials of masked TH59-DB pull-down fraction (Control) and TH59-DB pull-down fraction (Telomere) were loaded per lane.

| Name | Total number of peptides | |
	Control	Telomere-PIPA
TRF2	0	142
TRF1	0	23
RAP1	0	21
TIN2	0	14
POT1	0	18
ACACA1	0	16
PAXI1	0	12
INCENP	0	11
Borealin	0	10
AURKB	0	9

TRF2, TRF1, RAP1, TIN2, POT1 — Shelterin complex

INCENP, Borealin, AURKB — Chromosome passenger complex

Figure 26. List of proteins detected by mass spectrometry analysis of the material purified by TH59-DB from MEL cells.

The top 10 proteins are sorted by the total number of peptides.

Comprehensive identification of telomeric chromatin-associated RNAs

Next, RNA was extracted from the TH59-DB pull-down fraction containing telomeric chromatin and subjected to NGS analysis. Sequencing and mapping of these libraries yielded approximately 5.5 million and 2 million total and mappable read pairs in the TH59-DB pull-down fraction and 7.8 million and 3.5 million total and mappable read pairs in input. In the TH59-DB pull-down fraction, the substantial part of the RNAs in our dataset (742249 out of 5479729, 13.5%) had > 5 times-repeated telomeric sequences, which correspond to telomeric repeat-containing RNA such as TERRA & ARRET (Figure 27). The reads containing the $(TTAGGG)_5$ in the TH59-DB pull-down fraction was enriched 1000-fold over that of the input fraction (0.01%), suggesting much higher purity of the telomeric materials than that of the previous report (Fujita et al. 2015). To identify RNAs that are highly enriched in the telomeric chromatin, I plotted RNA level in the pull-down fraction (after purification) against that of input (before purification). Figure 28 showed that a cluster of RNAs were clearly enriched 100-fold over that of input. Importantly, the most enriched RNA was telomerase RNA component TERC (Figure 29). Almost all of the other RNAs included $(TTAGGG)_n$ repeat sequence (Supplementary Table S1), suggesting that the highly enriched RNAs derived from subtelomere regions or interstitial telomeric sequences (ITS). Actually, the sequence reads in the TH59-DB pull-down fraction were specifically mapped within last 30-kb region adjacent to the telomere sequence of each chromosome (Figure 30-32), which are considered to be subtelomeric region on q arm of each chromosome (Lopes et al. 2014). A part of the sequence reads were mapped to a validated TERRA transcript from the subtelomeric region of chromosome 18 in mice (Lopes et al. 2014, Figure 30). Altogether, I concluded that the telomeric chromatin purification using the PI polyamide probe is suitable to the

43

RNA-omics for telomeric chromatin-associated ncRNA, which has been never done by

PICh using LNA probes.
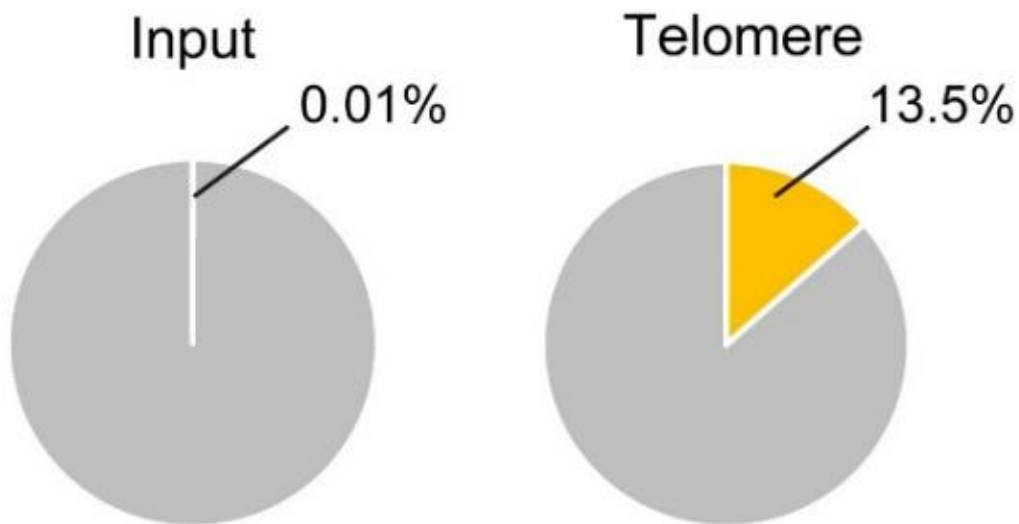
Figure 27. The percentage of telomeric repeat reads including TERRA transcripts in the input and TH59-DB pull-down fraction in MEL cells.

The number of the paired-end reads including $(TTAGGG)_5$ or $(CCCTAA)_5$ extracted from input and TH59-DB pull-down fraction (telomere) were divided by total number of reads. The percentage of telomeric repeat reads were 0.01% (telomere reads/total reads, 750/7843024) in input and 13.5% (742249/5479729) in telomere, respectively (shown in yellow color). Gray color shows the other reads.
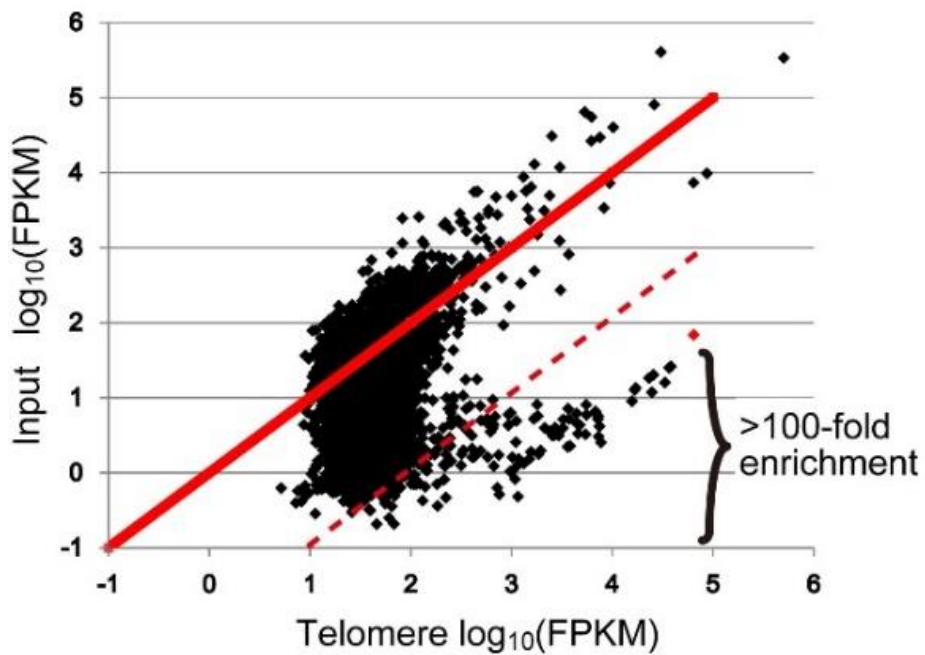
Figure 28. Scatter plot of fragments per kilobase million (FPKM) of TH59-DB pull-down fraction versus that of input sample for each RNA.

Telomerase RNA component, TERC, that is most enriched in TH59-DB pull-down fraction (telomere) is highlighted in red. RNAs that enriched more than100 fold in TH59-DB pull-down fraction are plotted below a red dotted line.
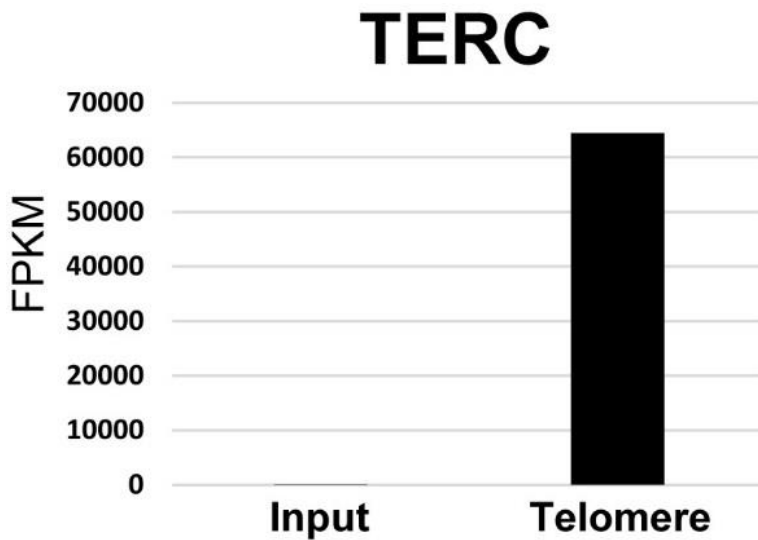
Figure 29. Bar graph of FPKM of telomerase RNA-component (TERC) in input and TH59-DB pull-down fraction (telomere).
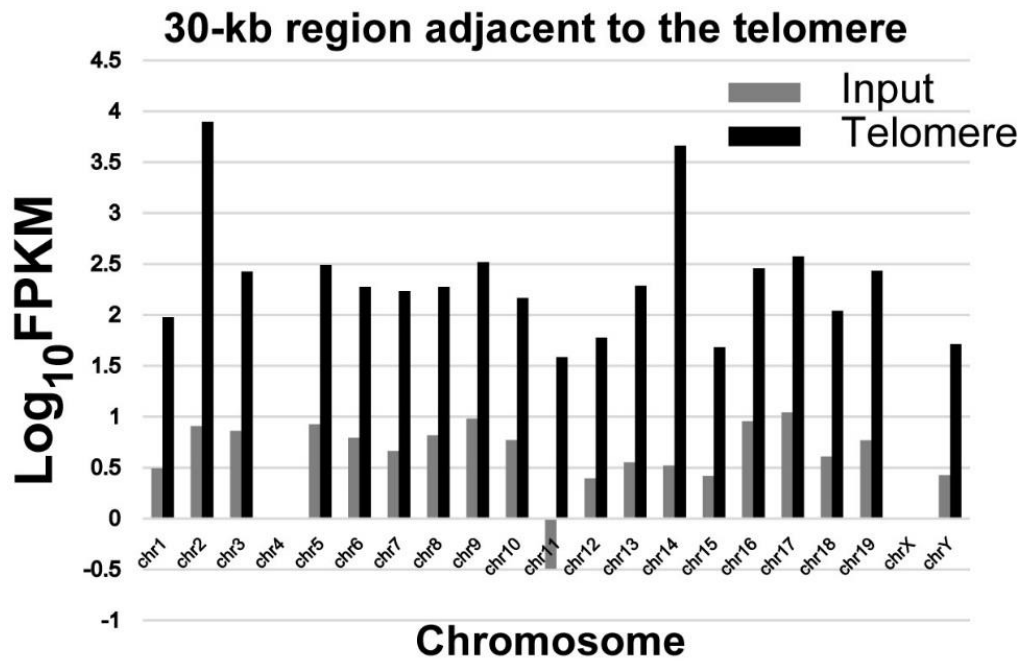
Figure 30. Bar graph of FPKM value of the transcripts at a 30-kb region adjacent to the telomere on q arm of each chromosome in input and TH59-DB pull-down fraction (telomere, Figure32).
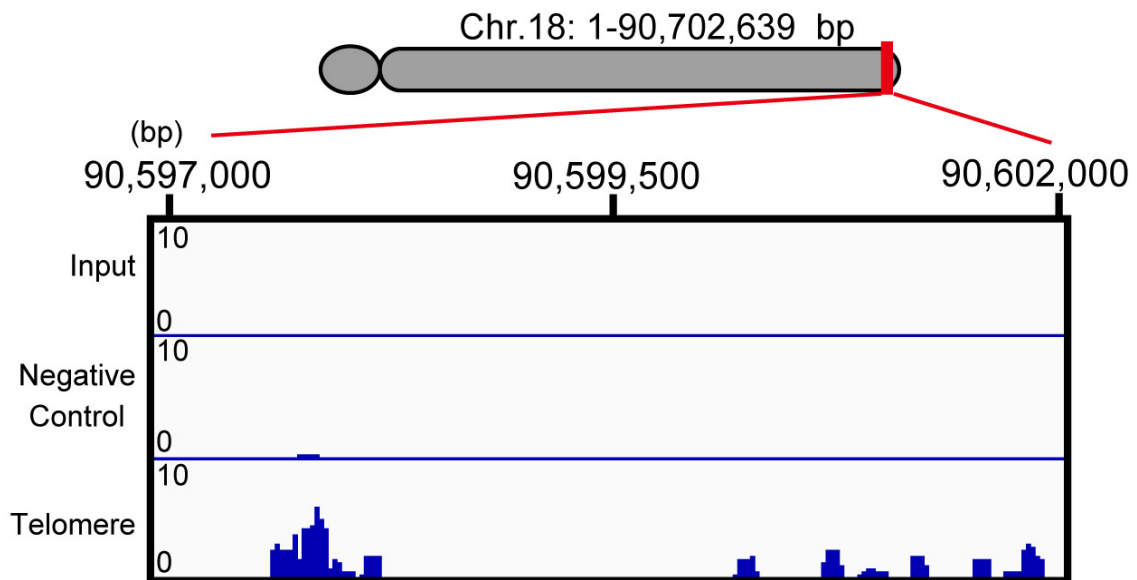
Figure 31. Mapped reads to a subtelomeric region in chromosome 18.
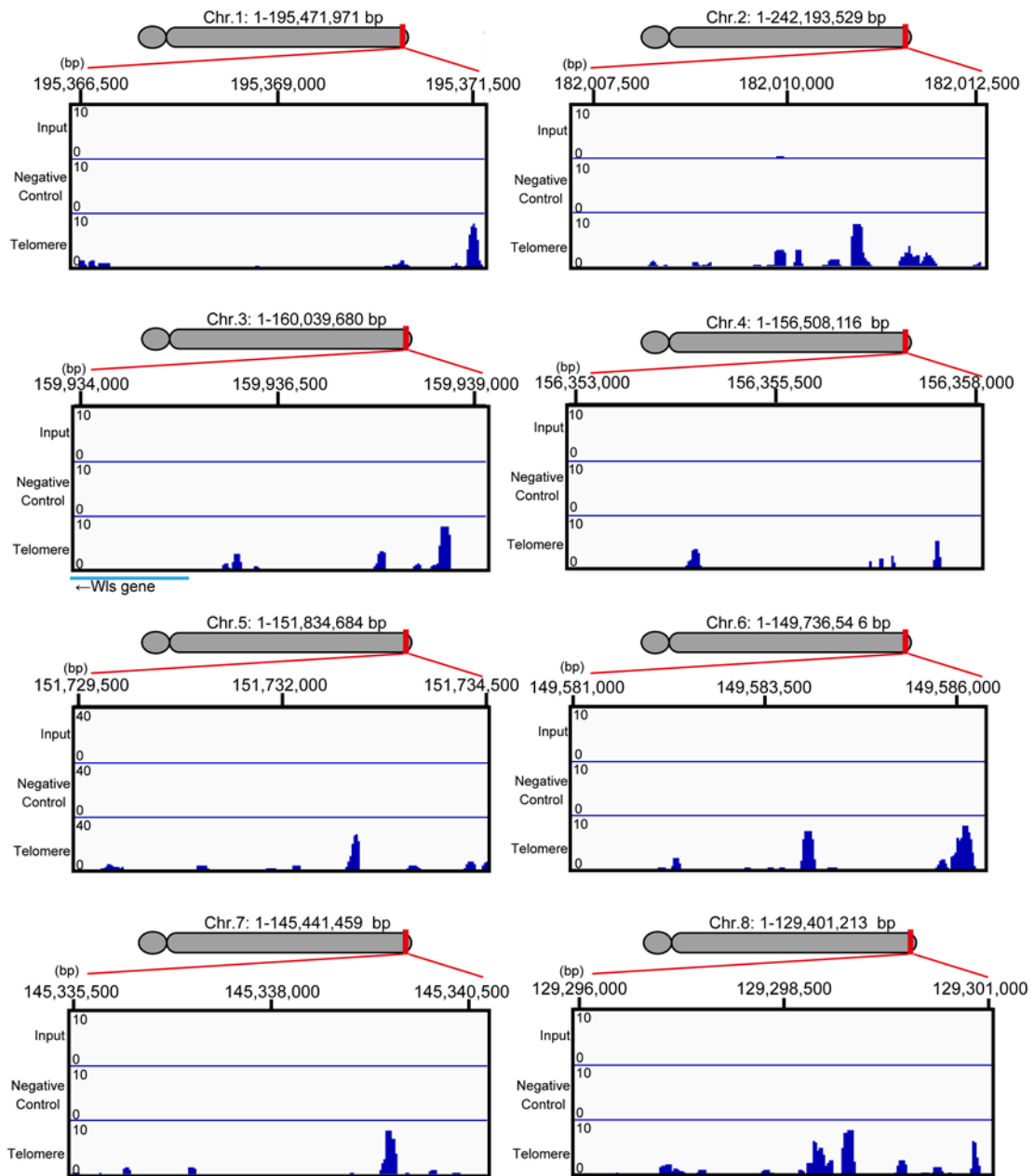
See legend on page 49.

See legend on page 49

48
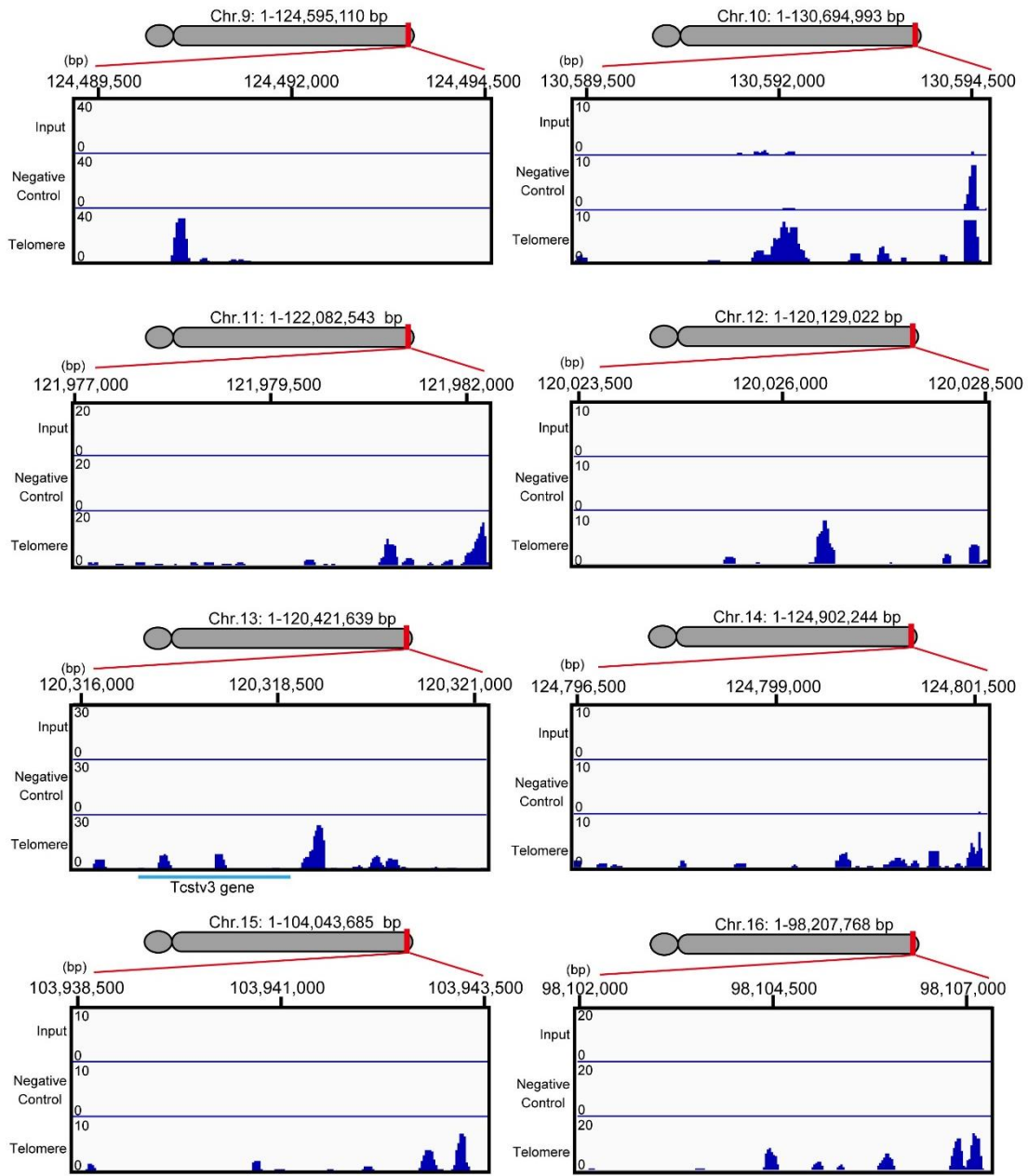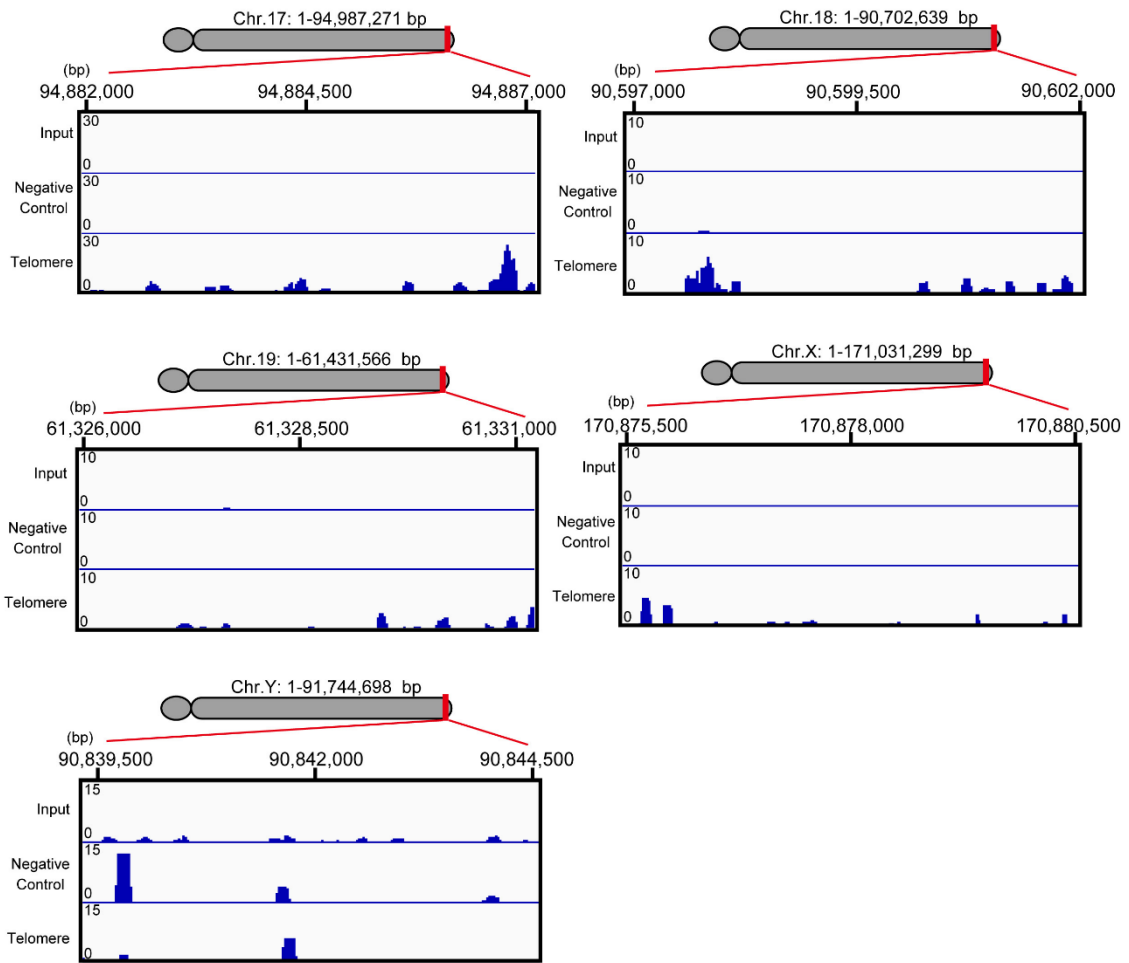
Figure 32. ncRNAs mapped to the region near the end of each chromosome.

Genomic view of the terminal regions of q arm of each chromosome based on mouse reference genome, GRCm38 (mm10). RNA were mapped to the terminal region of each chromosome.

ncRNAs on ALT telomeric chromatin

I then focused on ALT cells, which have highly recombinogenic telomeres. ALT telomeres have a different chromatin state from telomerase-driven telomere (Tardat and Dejardin 2018), and telomeric RNA, particularly TERRA, is highly expressed in ALT cells and might be a key player to promote the homologous recombination (Arora et al. 2014; Graf et al. 2017). Additionally, it is known that telomere elongation by ALT is associated with genomic alternation via the telomere repeats insertion into non-telomeric genome regions (Sakellariou et al. 2013; Marzec et al. 2015; Sieverling 2018). Based on these findings, I assumed that I might be able to pull-down telomeric chromatin as well as other genomic regions containing the telomeric repeat insertions distributed in the genome.

I applied PI-PRICh to two human cell lines with two distinct types of telomere: HeLa1.3 that is telomerase positive and has long telomeres, and the U2-OS ALT cell line. PI-PRICh data showed that TERC was detected in the telomerase-positive HeLa1.3 cells, but very few in the ALT cells (Figure 33). To compare the amounts of TERRA associated with the pull-down chromatin between the two cell lines, I mapped the reads on a reference sequence of a subtelomeric region called TelBam3.4 (Negishi et al. 2015; Nergadze et al. 2009; Figure 34). The TERRA amount calculated from the mapped reads in U2-OS cells was comparable to that of HeLa1.3 cells (Figure 34, 35). Furthermore, I compared the percentages of the reads including $>$(TTAGGG)$_5$ between the two cell lines, and found little difference (Figure 36).
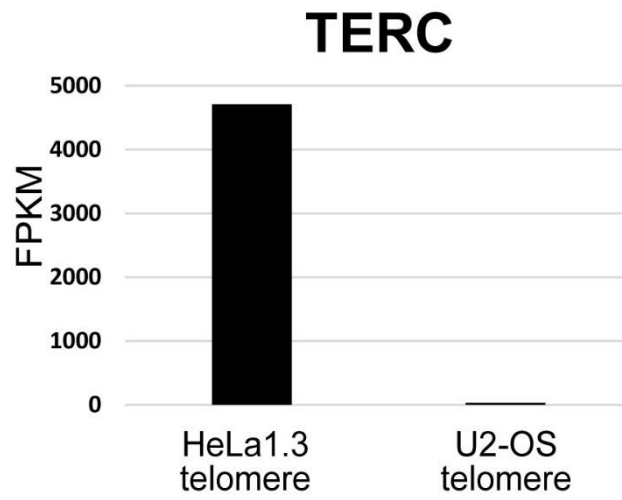
Figure 33. Bar graph of FPKM of telomerase RNA component (TERC) in telomerase-maintained telomere (HeLa1.3) and ALT telomere (U2-OS).
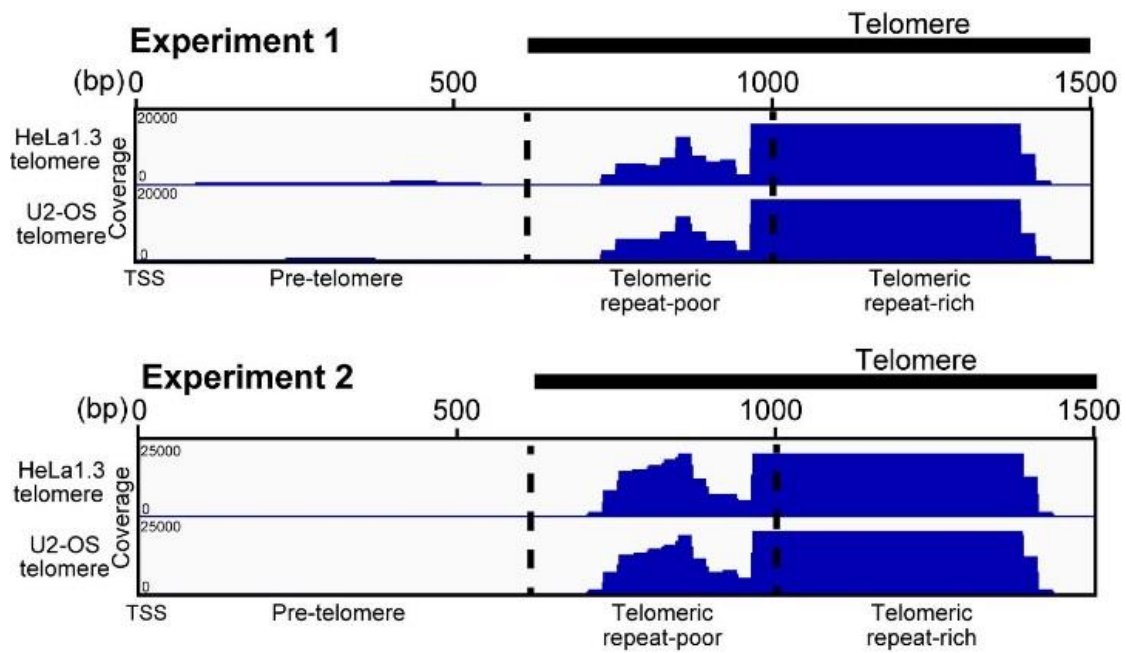
Figure 34. Comparison the telomeric chromatin-associated TERRA between telomerase positive and ALT cell lines.

A browser view of a chromatin-associated TERRA transcripts at a subtelomeric region downstream transcription start site (TSS) coding TelBam3.4 sequence using Integrative Genomic Viewer (IGV) in telomerase-maintained telomere (HeLa1.3) and ALT telomere (U2-OS). The numbers at left side in tracks indicate the track height with reads counts in RNA-seq. Telomere region in TelBam3.4 that was defined previously (Nergadze 2009, Usui 2015). Borderlines (vertical dotted black lines) are shown between pre-telomere, telomeric repeat-poor, and -rich regions based on TTAGGG repeat density.

52
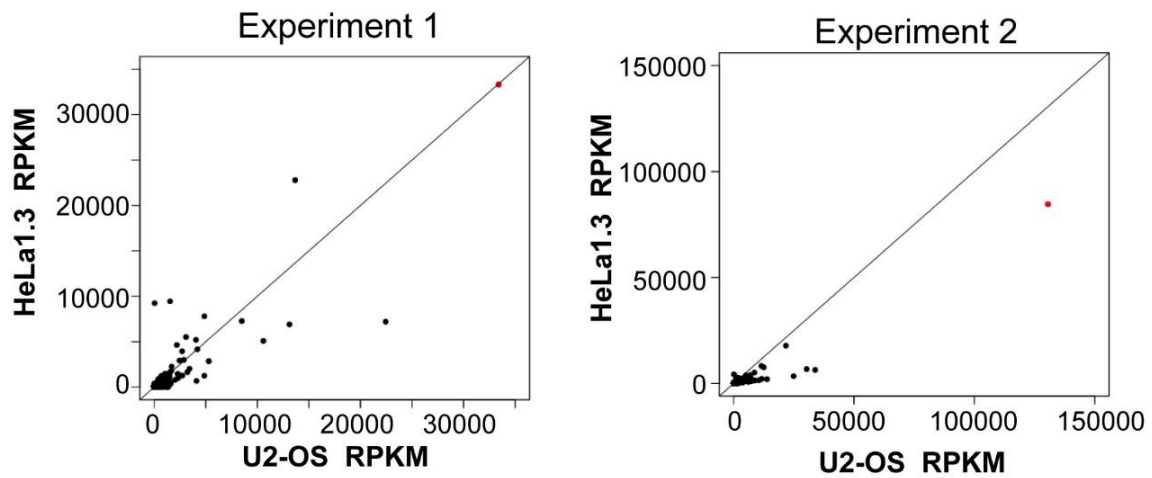
Figure 35. Scatter plots of read per kilobase with normalized count per million (RPKM) in ALT cells versus that of telomerase positive cell for each RNA from two independent PI-PRICh experiments.

RPKM of TH59-DB pull-down fraction in telomerase positive cells (HeLa1.3) and ALT cells (U2-OS) were plotted. Left part and right part show experiment 1 and experiment 2, respectively. Experiment 1 and 2 of PI-PRICh were performed independently. RNA that is transcribed from the subtelomeric region of TelBam3.4 are highlighted in red. The line y=x is shown as a black line.

(TTAGGG)×5 or (CCCTAA) ×5 in Telomeric fraction
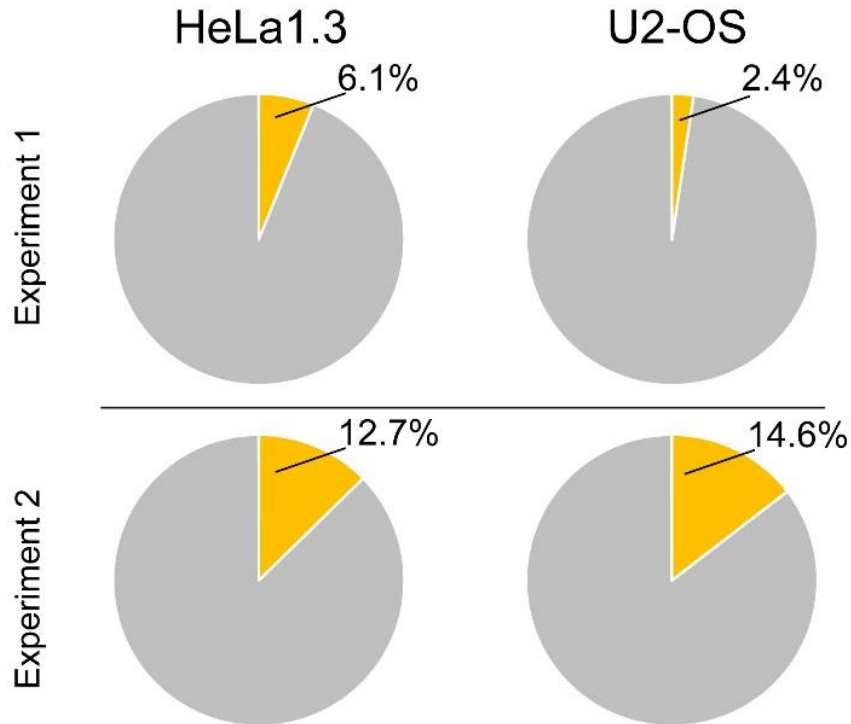
	HeLa1.3			U2-OS		
	Total read	(TTAGGG)x5 or (CCCTAA)x5	Ratito to total read	Total read	(TTAGGG)x5 or (CCCTAA)x5	Ratito to total read
Experiment1	2864776	175994	0.061434	3150959	75317	0.023903
Experiment2	2806838	355639	0.126704	3031148	441514	0.145659

Figure 36. Comparison of telomeric RNA between telomerase positive cells and ALT cells. Percentage of telomere sequence reads including TERRA/ARRET in the input and TH59-DB pull-down fraction of telomerase positive cells (HeLa1.3) and ALT cells (U2-OS). Upper part and lower part show experiment 1 and experiment 2, respectively. Experiment 1 and 2 of PI-PRICh were performed independently. Total read number, read number containing >(TTAGGG)$_5$/(CCCTAA)$_5$ and the ratio of the telomere sequence containing read to total reads are shown in a table. Pie chart shows TERRA/ARRET fraction in yellow and the other fractions in gray. Counted read numbers are shown in the table below the graph.

To search novel ncRNAs enriched in the pull-down chromatin from ALT cells, I used

three filtering criteria: (1) the fragments per kilobase of per million mapped reads

(FPKM) value. (2) the fold-enrichment over total RNA of input sample before the

telomeric purification. (3) the relative ratio of ncRNAs between U2-OS cells and

HeLa1.3 cells. Then I identified some intronic ncRNAs (SPOCK3 and USP16 gene),

which had more than 100 FPKM, were enriched more than 25-fold, and were detected

in U2-OS cells 50 times more than HeLa1.3 cells (Figure 37, 38, and Supplementary

Table S2-S6). On the other hand, no ncRNAs except TERC were enriched in HeLa1.3

cell by this criteria.

Cell	RNAs	length	FPKM in telomere fraction in experiment1 (HeLa1.3/U2-OS)	FPKM in telomere fraction in experiment2 (HeLa1.3/U2-OS)
HeLa1.3	Terc	462 bp	4709/53	4125/69
U2-OS	SPOCK3 intron	434 bp	0/827	1/179
U2-OS	USP16 intron	3299 bp	0/280	1/120

Figure 37. ncRNAs specifically enriched in TH59-DB pull-down fraction of ALT cells (U2-OS) or
telomerase positive cells (HeLa1.3) in two independent experiments.
Filtering criteria to search novel ncRNAs are shown in text.
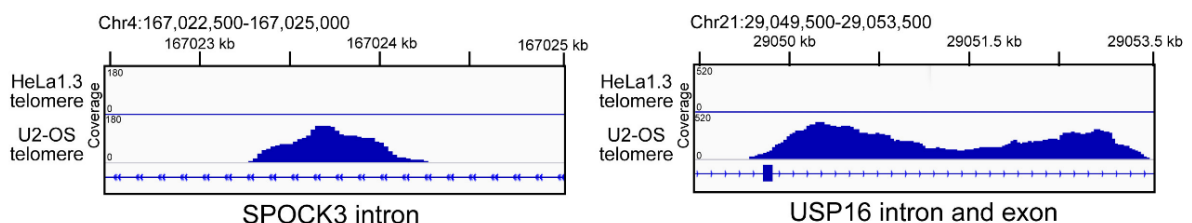
Figure 38. Telomeric chromatin-associated ncRNAs mapped to the intron of SPOCK3 and USP16 gene in
ALT cells.
These ncRNAs were highly enriched in TH59-DB fraction from ALT cells (U2-OS), not in that of
telomerase positive cells (HeLa1.3).

Telomere repeat insertion into the intron regions coding ncRNAs associated with ALT telomere

As described above, it is known that telomere elongation by ALT involves homologous recombination not only between the telomeres but also between a non-telomeric regions and telomeres. The latter case can induce the telomeric repeat insertions over the genome (Marzec et al. 2015; Sieverling et al. 2018). Therefore, I hypothesized that the identified intronic RNAs in ALT cells might be generated by transcription from around the inserted telomeric repeats in the intron regions. To test if the intron regions in ALT cells have flanking telomeric repeats insertions, I used genomic polymerase chain reaction (PCR) with a primer set: one primer annealing to the intron regions and the other to $(CCCTAA)_4$ or $(TTAGGG)_4$ (Primer sets (1) (2) in Figure 39). If the telomeric repeats locate close enough to the intron regions, PCR reaction can generate longer amplification products than the control ones by primer sets to amplify the target intron regions (Primer sets (3) in Figure 39). Genomic PCR in U2-OS cells provided the smeared and long PCR products with the primer set (1) and (2) annealing to the introns of SPOCK3 and the telomeric repeats (lane 5 & 6 in Figure 40) while such products were not amplified with genomic PCR in HeLa1.3 cells (lane 1 & 2 in Figure 40). In the case of USP16 intron, the smear and intense PCR products was amplified by genomic PCR only with primer set (2) in U2-OS cells, but not in HeLa1.3 cells (lane 9 & 13 in Figure 40). I also confirmed that the telomeric repeats were inserted near the intron regions by Sanger sequencing after cloning the smeared PCR products into a plasmid vector. Furthermore, I investigated five intron sequences of CKS1B, NRDC, PAM, RBBP8, and KIAA1671 genes, which correspond to the other intronic ncRNAs identified only in the experiment 1(Figure 41, Supplementary Table S2), to test if they

56

also have the flanking telomeric repeats (Figure 42). I obtained similar results to

SPOCK3 and USP16 introns (Figure 43).
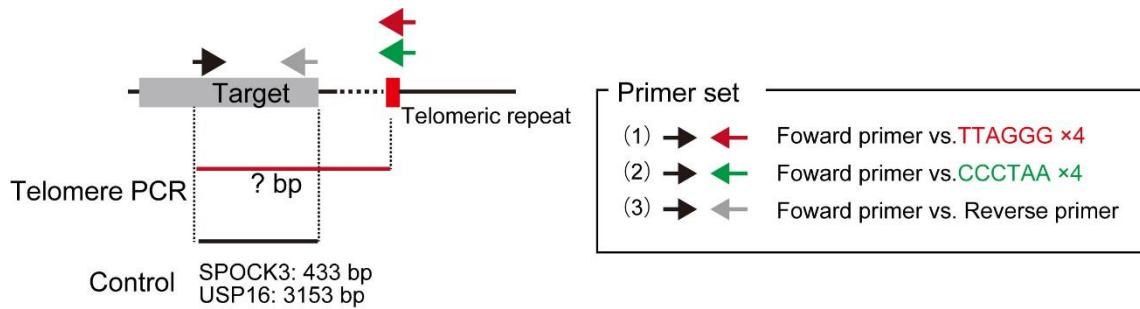
Figure 39. A scheme of genomic PCR to test telomeric repeat insertion into the flanking regions of
SPOCK3 and USP16 introns.

Primer sets and expected PCR products are shown. Primer set (1), an oligonucleotide annealing to
the target intron and $(TTAGGG)_4$; primer set (2), the same oligonucleotide annealing to the target
intron and $(CCCTAA)_4$; primer set (3), two oligonucleotides both of which anneals to the target
intron were used as primers for PCR.
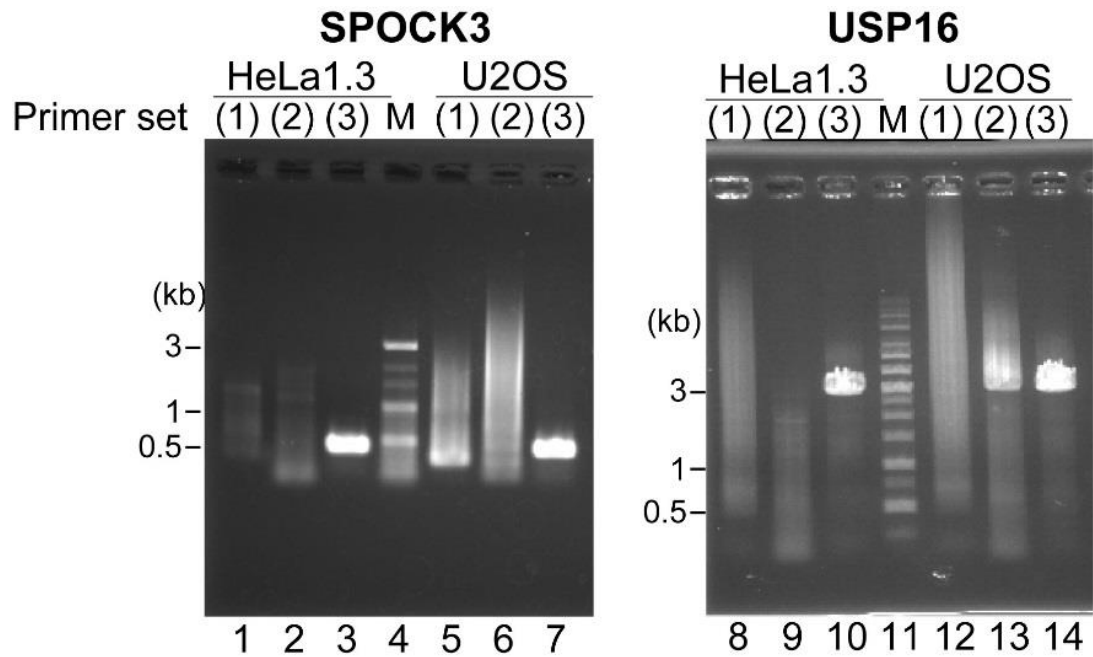
Figure 40. Genomic PCR to test telomere repeat insertion into SPOCK3 and USP16 intron. PCR products of SPOCK3 intron amplified by each primer set with telomerase positive (HeLa1.3) and ALT (U2-OS) genomic DNA are shown in lane 1-3 and lane 5-7 in the left panel, respectively. PCR products of USP16 intron amplified by each three primer set with HeLa1.3 and U2-OS genomic DNA are shown in lane 8-10 and lane 12-14 in the right panel, respectively. 100 bp and 1 kb ladder were loaded in lane 4 and lane 11 as size maker (M), respectively.
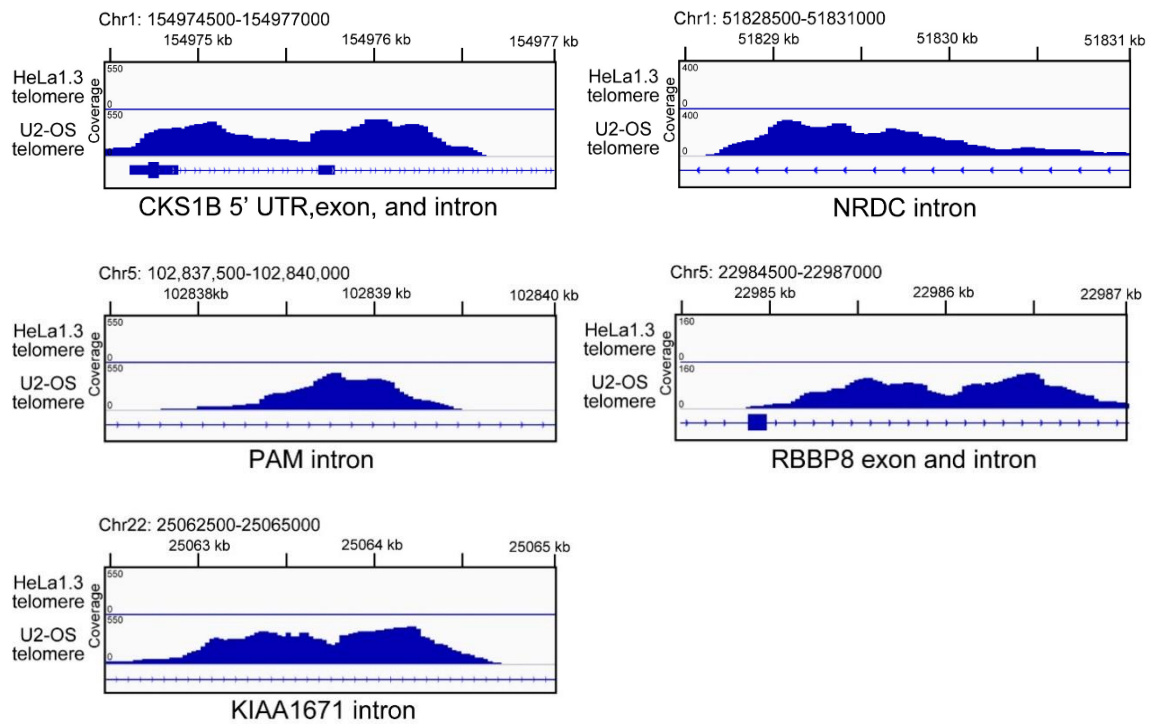
Figure 41. Telomeric chromatin-associated ncRNAs mapped to intron regions of CKS1B, NRDC, PAM, RBBP8, and KIAA1671 gene in ALT cells.

These ncRNAs were highly enriched in TH59-DB pull-down fraction of ALT cells (U2-OS), not in that of telomerase positive cells (HeLa1.3).
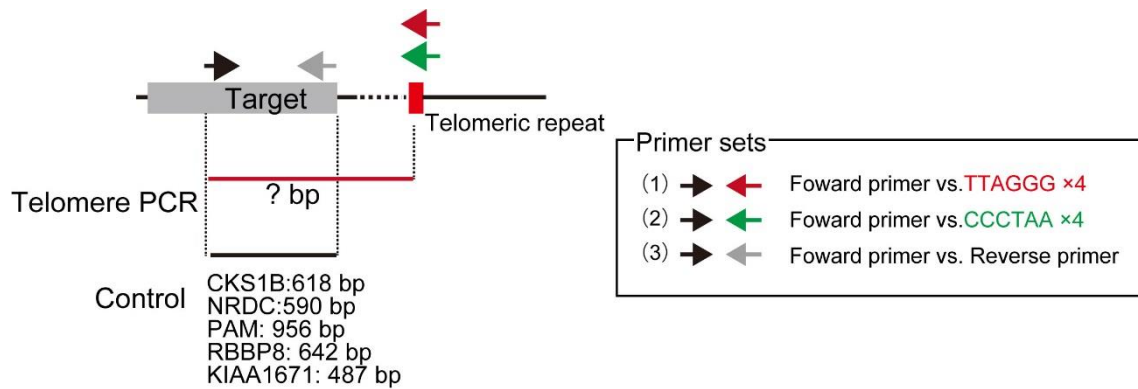
Figure 42. A scheme of genomic PCR to test telomeric repeat insertion into the flanking regions of CKS1B, NRDC, PAM, RBBP8, and KIAA1671 introns.

Primer sets and expected PCR products are shown. Primer set (1), an oligonucleotide annealing to the target intron and $(TTAGGG)_4$; primer set (2), the same oligonucleotide annealing to the target intron and $(CCCTAA)_4$; primer set (3), two oligonucleotides both of which anneals to the target intron were used as primers for PCR.
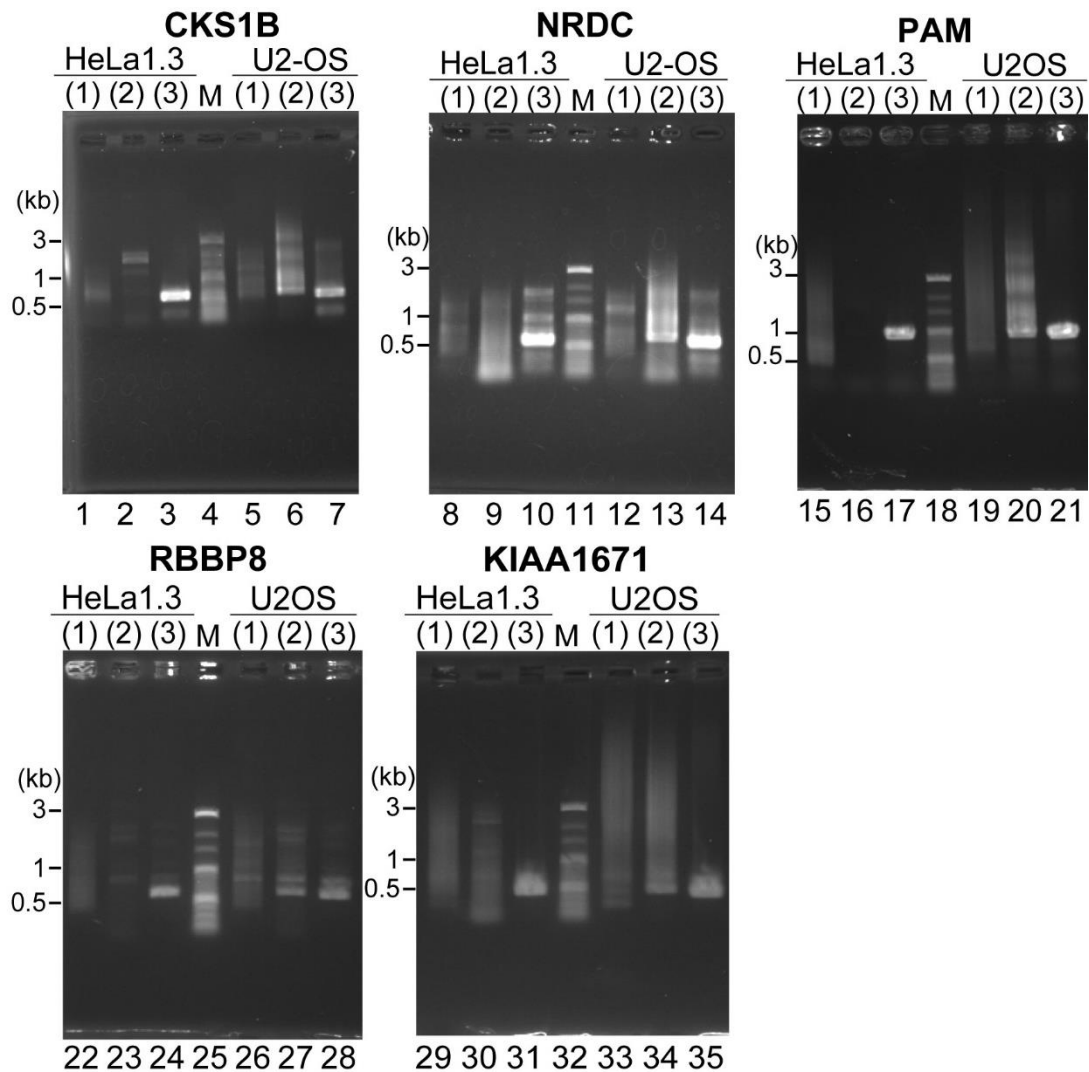
Figure 43. Genomic PCR to test telomeric repeat insertion into the flanking regions of the CKS1B, NRDC, PAM, RBBP8, and KIAA1671 introns.

PCR products of CKS1B intron amplified by each primer set with telomerase positive (HeLa1.3) and ALT (U2-OS) genomic DNA are shown in lane 1-3 and lane 5-7 in the left panel, respectively. PCR products of NRDC intron amplified by each three primer set with HeLa1.3 and U2-OS genomic DNA are shown in lane 8-10 and lane 12-14 in the right panel, respectively. PCR products of PAM intron amplified by each three primer set with HeLa1.3 and U2-OS genomic DNA are shown in lane 8-10 and lane 12-14 in the right panel, respectively. PCR products of RBBP8 intron amplified by each three primer set with HeLa1.3 and U2-OS genomic DNA are shown in lane 8-10 and lane 12-14 in the right panel, respectively. PCR products of KIAA1671 intron amplified by each three primer set with HeLa1.3 and U2-OS genomic DNA are shown in lane 8-10 and lane 12-14 in the right panel, respectively.

61

To confirm the existence of telomere repeat insertion into the intron regions by a different methodology, I analyzed the DNA regions captured by PI-PRICh, herein called DNA-Seq, because the TH59-DB can bind to the telomeric repeats in the intron regions and enrich those region from genomic DNA. DNA was extracted from TH59-DB pull-down fraction in HeLa1.3 and U2-OS and subject to NGS analysis. Figure 44 shows that intron regions were enriched in TH59-DB pull-down fraction in U2-OS compared to HeLa1.3. I found the chimeric sequence reads of the intron DNA sequence with $(TTAGGG)_n$ repeats, indicating that telomeric repeats were inserted into the intron regions (Figure 45).
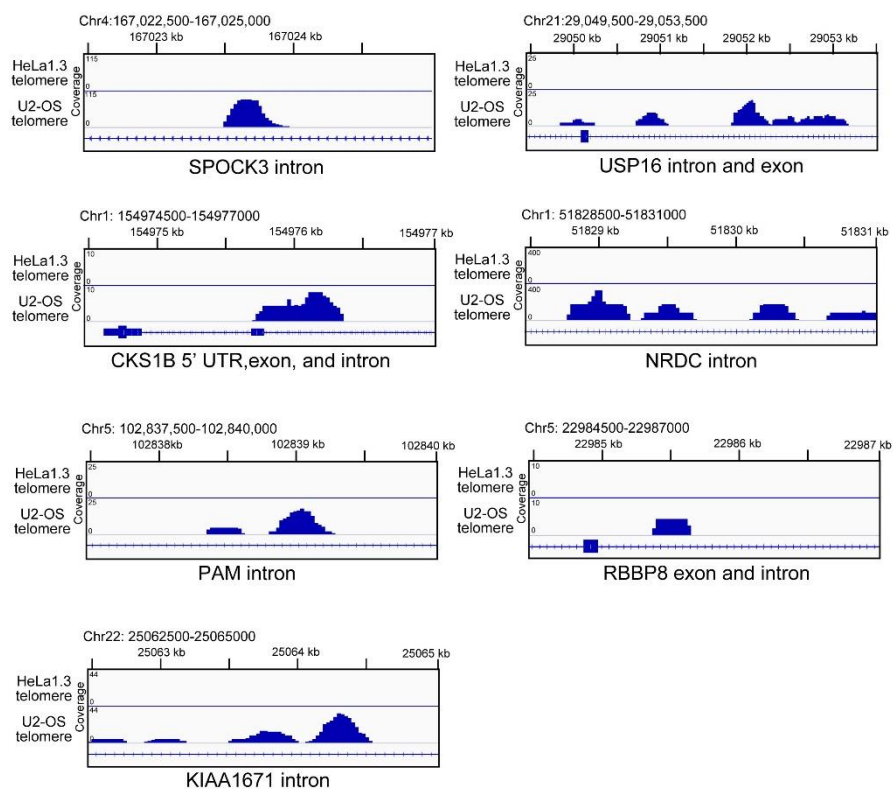
Figure 44. Genomic regions coding the introns of SPOCK3, USP16, CKS1B, NRDC, PAM, RBBP8, and KIAA1671 were specifically enriched in TH59-DB pull-down fraction of ALT cells (U2-OS).

These intron regions were enriched in TH59-DB fraction from ALT cells (U2-OS), not in that of telomerase positive cells (HeLa1.3).

SPOCK3   intron   Telomere repeat

>@M00453:54:000000000-C8TP6:1:1101:11178:15491 2:N:0:2
GAAAGTAAACCCTGCTGAATATCCCCCTAATGCAAAGCCAGACTGTTTTCTATCCTGCTC
TACGCTGTAAGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGGCTACAGTTAATT
ACTAACTTAGCTTTCATGACACCAGCCCAGATCGGAAGAGCGTCGTGTAGGAAAGAGTGT
GCCTCTATGTGTAGATCTCCGCGGTCGCCCTATCATTAAAAAAAAAAAAAAAAAAGAACCAC
CCCAACACAC

>@M00453:54:000000000-C8TP6:1:1101:11617:24152 2:N:0:2
TGGGCCTTTGTATGTTTCCCACGTGGAGAAATGGGCTAGAGAAAGGACAAGATTACACAT
GCTGGGGAGAAAAGAAAGTAAACCCTGCTGAATATCCCCCTAATGCAAAGCCAGACTGTT
ATCTATCCTGCTCTACGCTGTAAGTTAGGGTTAGGGTTAGGGTTAGGGTTAGCC

>@M00453:54:000000000-C8TP6:1:1101:18395:24383 2:N:0:2
TCCTGCTGAATATCCCCCTAATGCAAAGCCAGACTGTTTTCTATCCTGCTCTACGCTGTA
AGTTAGGGTTAGGGTTAGGGCTAGGGTTAGGGTTAGGGGCTACAGTTAATTACTAACTTA
GCTTTCATGACACCAGTGGTGACTTTAACAGCCGTAGTTCTTTGATGAGTTATTTTTCCCC
CTTTTTCTCTGCTTCTGTACTTAAAGAGCAATGGTGGCCG

>@M00453:54:000000000-C8TP6:1:1102:25172:16164 2:N:0:2
AGAAAGGACAAGATTACACATGCGGGGGAGAAAAGAAAGTAAACCCTGATGAATATCCC
CATAATGCAAAGCCAGACTGTTTTCTATCCTGCTCTACGCTGTAAGTTAGGGTTAGGGTT
AGGGTTAGGGTTAGGGTTAGGGGCTACAGTTAATTCCTAACTTAGCTTTCATGACACCAG
TGGGGACTTTAACACCCGTAGTCCCCCAATCGGAAGAACGACGTCAAGGAAAAAAGGGG
GCCTCTATGTGT

>@M00453:54:000000000-C8TP6:1:1104:11009:7524 2:N:0:2
CCTGCTGAATGTCCCCCTAATGCAAAGCCAGACTGTTTTCTATCCTGCTCTACGCTGTAA
GTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGGCTACAGGCC

See legend on page 65.

64

USP16 intron   telomere repeat

>@M00453:54:000000000-C8TP6:1:1101:6138:14030 2:N:0:2
GACGGAGAAGTACTAAGTTCTTCAGTGAATATGGAGGCACAGATTGGGAAGCTGAGTATA
GGAGACAGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTA
GGCC

>@M00453:54:000000000-C8TP6:1:1103:16887:22996 2:N:0:2
GGTGTGGTTGCCTAAAATGGAATGGCATTTAAAGTCCTGTGAGAATCATAGATGGACGGA
GAAGTACTAAGTTCTTCAGTGAATATGGAGGCACAGATTGGGAAGCTGAGTATAGGAGAC
AGGTTAGGGATAGGGTTAGGGTTAGGGTTAGGGTTAGGGGTAGGGGTAGGTGT
AGGGATAGGTGGAGGGTTGGGGGCACGGTCGCAGATACGGATACGGGTAGGGGGAAGGT
TAGGGGTACGATG

>@M00453:54:000000000-C8TP6:1:1105:18418:11869 2:N:0:2
GGACGGAGAAGTACTAAGTTCTTCAGTGAATATGGAGGCACAGATTGGGAAGCTGAGTAT
AGGAGACAGGTTAGGGTTAGGGTTAGGGTTAGGGTTCGGGATAGGGTTGTGGTTAGGGTT
AGGGTTAGGGTTAGTGTGGC

>@M00453:54:000000000-C8TP6:1:1107:13604:10641 2:N:0:2
CAATTCTGAGTGATGACAAGCCCAGACTCTAGTCCTGGGTGTGGTTGACTAAAATGGAAT
GGCATTTAAAGTCCTGTGAGAATCATAGATGGACGGAGAAGTACTAAGTTCTTCAGTGAA
TATGGAGGCACAGATTGGGAAGCTGAGTATAGGAGACAGGTTAGGGTTAGGGTTAGGGTT
AGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGC

>@M00453:54:000000000-C8TP6:1:1107:11764:22861 2:N:0:2
TCTGAGTGATGACAAGCCCAGACTCTAGTCCTGGGTGTGGTTGACTAAAATGGAATGGCA
TTTAAAGTCCTGTGAGAATCATAGATGGACGGAGAAGTACTAAGTTCTTCAGTGAATATG
GAGGCACAGATTGGGAAGCTGAGTATAGGAGACAGGTTAGGGTTAGGGTTAGGGTTAGG
ATTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTCGGGTTAGGGTTAGGGTTAGGGTTAGG
GTTAGGGTTCGG

Figure 45. Representative sequence reads containing both telomeric repeats and SPOCK3 or
USP16 intron region in TH59-DB pull-down fraction from U2-OS cells.

65

Localization of ncRNAs at telomere

To examine localizations of the four intronic ncRNAs (SPOCK3, USP16, CKS1B, NRDC genes), we performed their RNA-FISH analysis. I simultaneously stained the telomeres in U2-OS cell line using the fluorescent TH59 (Maeshima, Janssen, and Laemmli 2001; Kawamoto et al. 2013). As a positive control of RNA-FISH and telomere labeling, I confirmed the colocalization of TERRA signals with telomere signals (Figure 46). Using this method, foci of SPOCK3 and USP16 intronic RNA were detected. Some of them were co-localized with telomere signals in 50% of SPOCK3 positive cells and 25% of USP16 positive cells, respectively. (Figure 47-49). The RNA-FISH for CKS1B and NRDC intronic RNA also co-localized with telomeres in 10% of cells that RNA signals were detected. (Figure 49-51). The other RNA (KIAA1671, PAM, and RBBP8) signals were not detected probably because of their low amounts. These results indicated that the intronic RNAs often localize near telomeres in the cell nuclei.
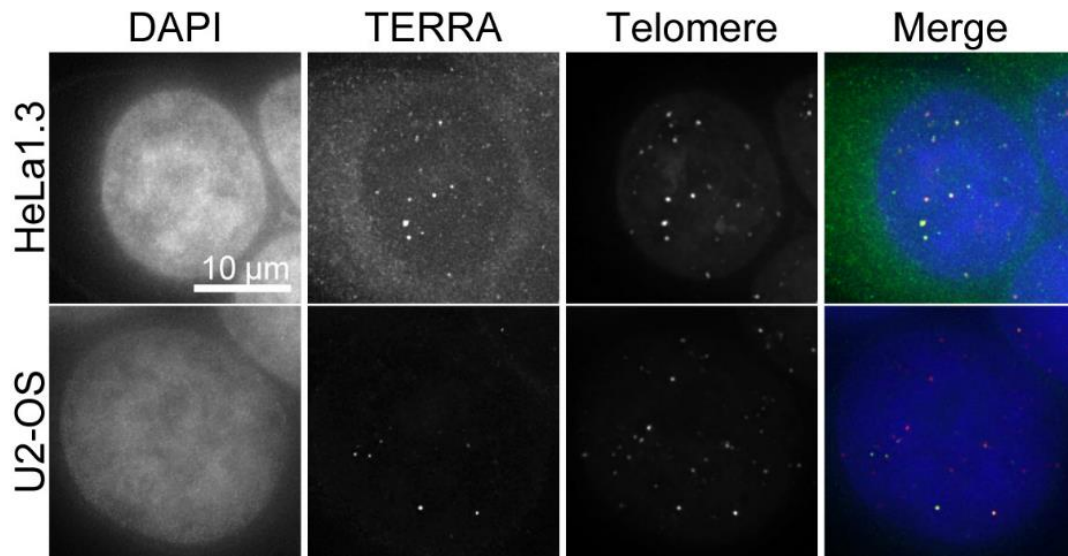
Figure 46. RNA-FISH for TERRA and the simultaneous labeling with fluorescent PI polyamide probe.

First column, DAPI signal; second column, ncRNA signal of TERRA; third column, fluorescent PI polyamide, fluorescent TH59 signal (telomere); fourth column, merged image of DAPI (blue), ncRNA (green) and telomere (red).
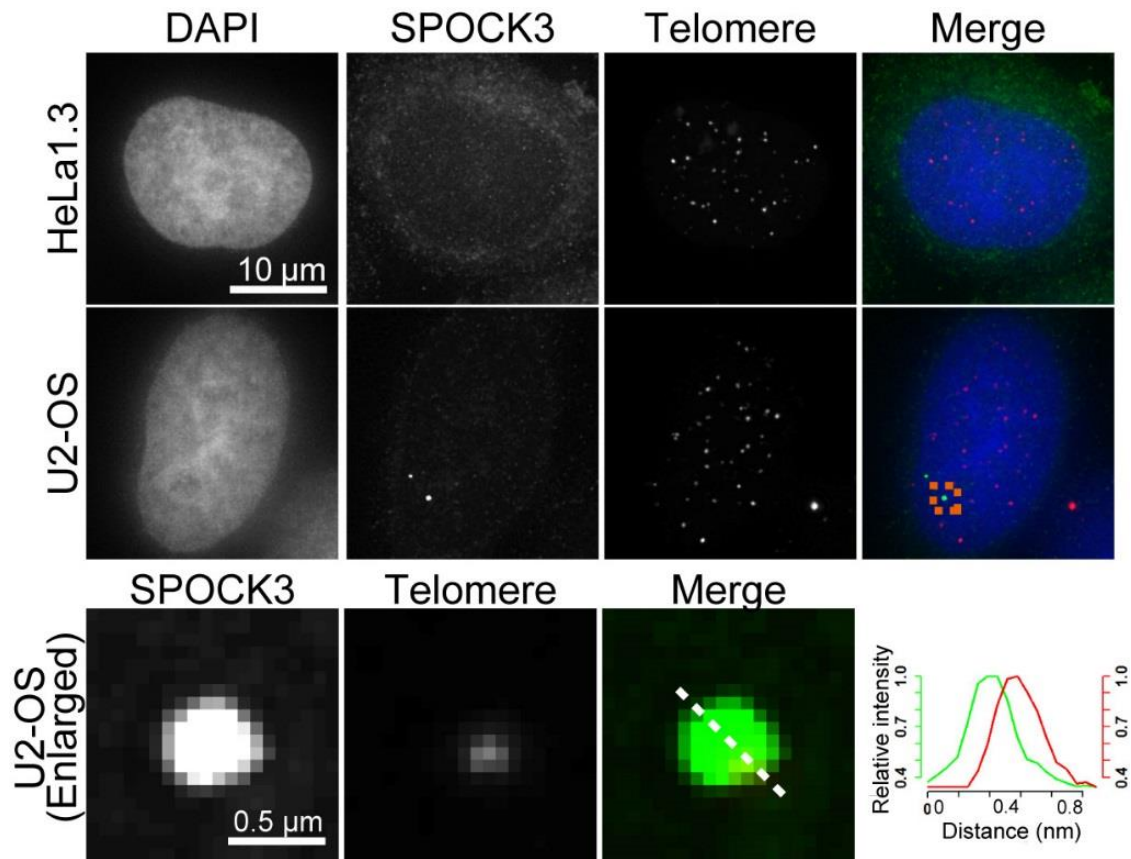
Figure 47. RNA-FISH for SPOCK3 intron and the simultaneous labeling with fluorescent PI polyamide probe.

First column, DAPI signal; second column, ncRNA signal of TERRA; third column, fluorescent PI polyamide, fluorescent TH59 signal (telomere); fourth column, merged image of DAPI (blue), ncRNA (green) and telomere (red). Enlarged images of the boxed region in the upper panel are shown in the lower panel. Line plot illustrates the signal of SPOCK3-intronic RNA and the telomeric signal on the white dotted line in the left merged image.
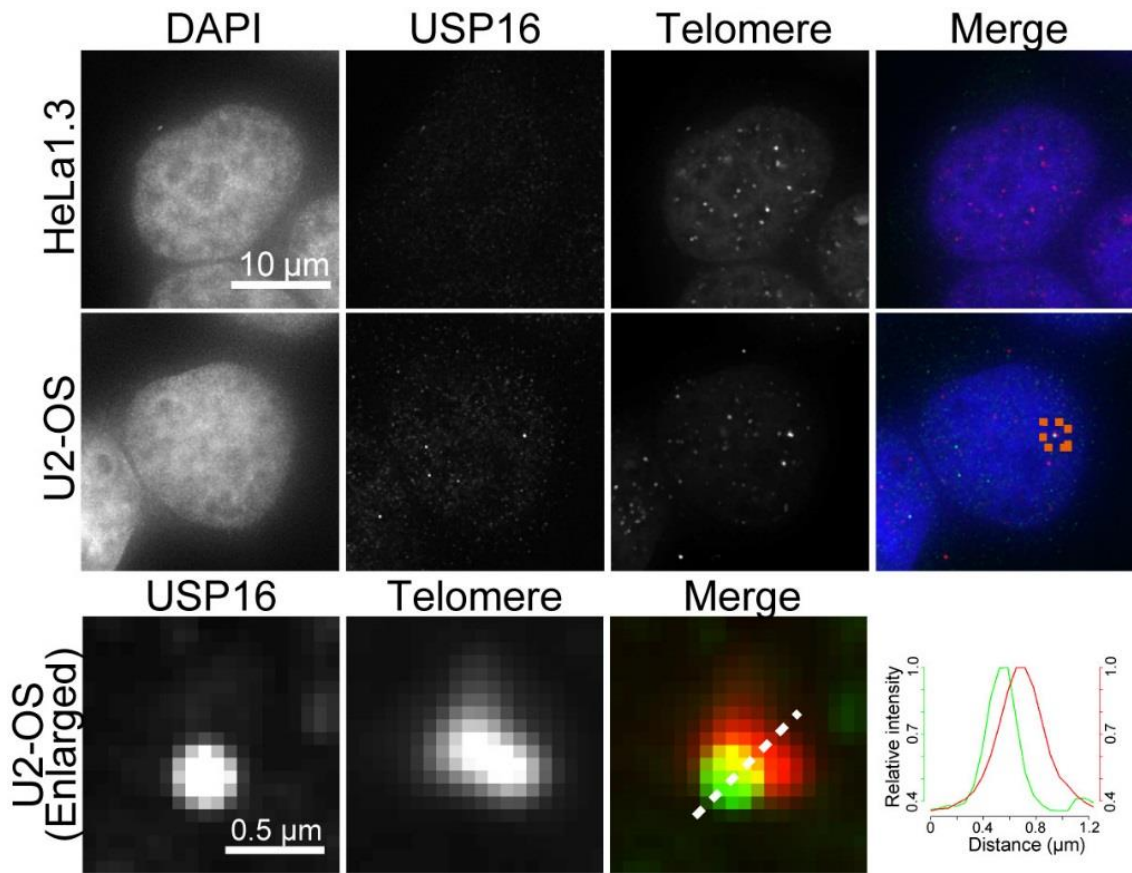
Figure 48. RNA-FISH for USP16 intron and the simultaneous labeling with fluorescent PI polyamide probe.

First column, DAPI signal; second column, ncRNA signal of TERRA; third column, fluorescent PI polyamide, fluorescent TH59 signal (telomere); fourth column, merged image of DAPI (blue), ncRNA (green) and telomere (red). Enlarged images of the boxed region in the upper panel are shown in the lower panel. Line plot illustrates the signal of USP16-intronic RNA and the telomeric signal on the white dotted line in the left merged image.
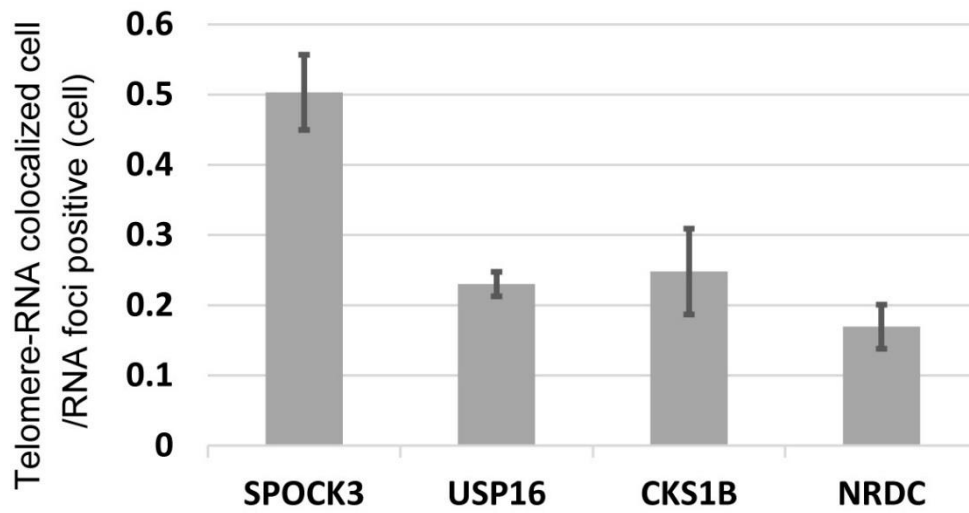
Figure 49. Bar graph of percentages of cells in which SPOCK3, USP16, CKS1B, and NRDC intronic RNA signals were colocalized with telomere signal.

Three independent experiments are performed and the numbers of ncRNA signals measured for U2-OS cells are from 100 cells in each experiment. Error bars show standard deviation.
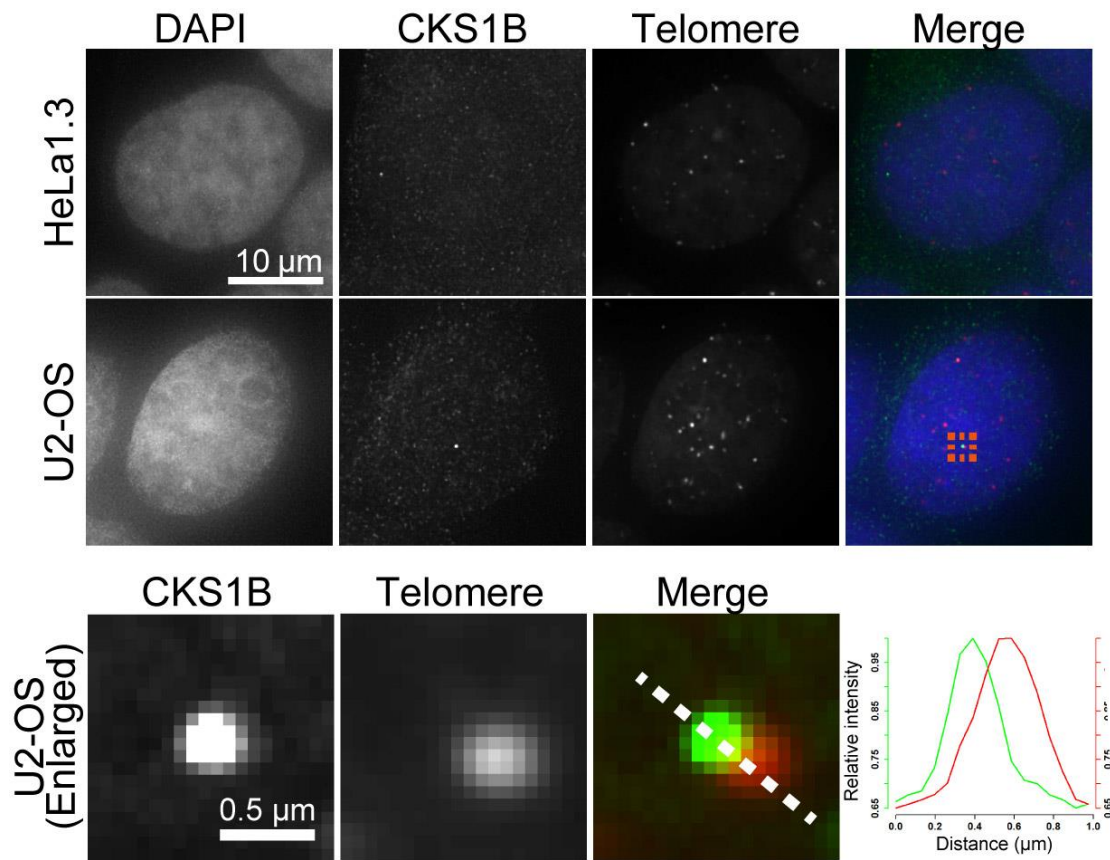
Figure 50. RNA-FISH for CKS1B intron and the simultaneous labeling with fluorescent PI polyamide probe.

First column, DAPI signal; second column, ncRNA signal of TERRA; third column, fluorescent PI polyamide, fluorescent TH59 signal (telomere); fourth column, merged image of DAPI (blue), ncRNA (green) and telomere (red). Enlarged images of the boxed region in the upper panel are shown in the lower panel. Line plot illustrates the signal of USP16-intronic RNA and the telomeric signal on the white dotted line in the left merged image.

71
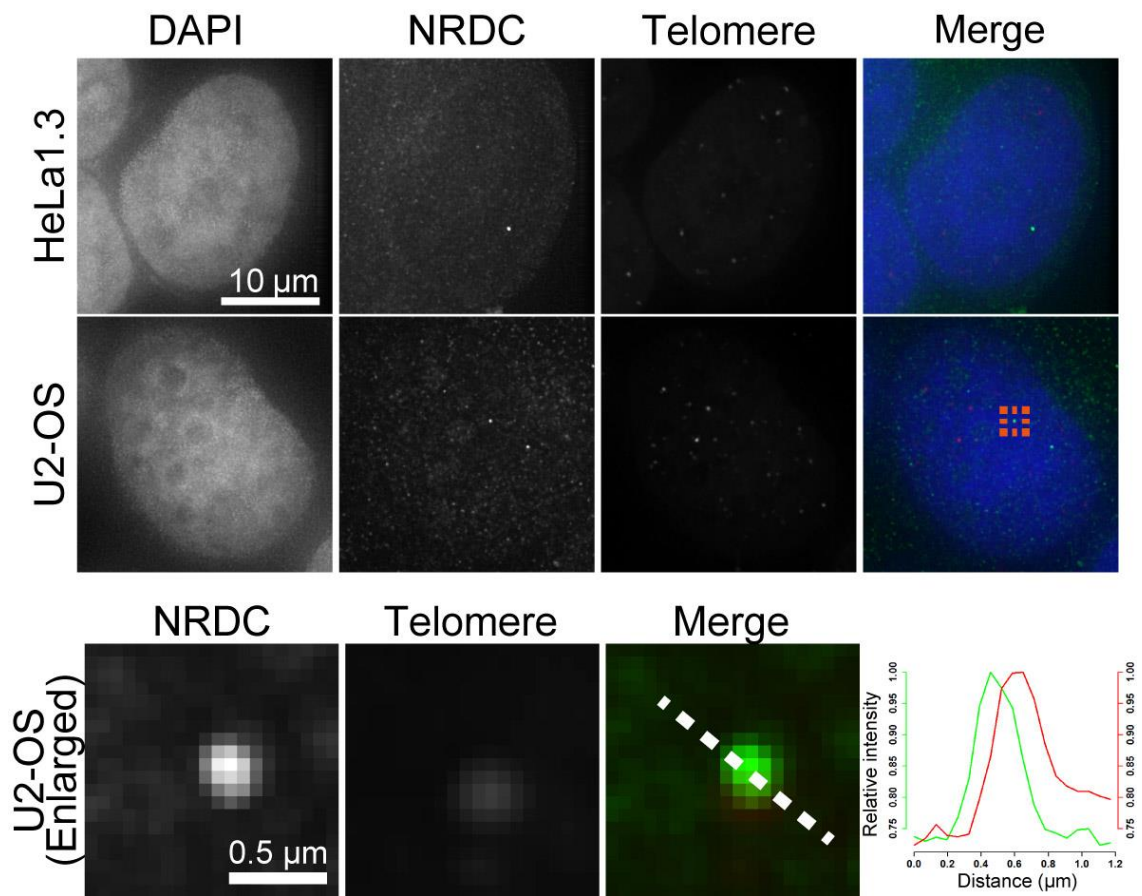
Figure 51. RNA-FISH for NRDC intron and the simultaneous labeling with fluorescent PI polyamide probe.

First column, DAPI signal; second column, ncRNA signal of TERRA; third column, fluorescent PI polyamide, fluorescent TH59 signal (telomere); fourth column, merged image of DAPI (blue), ncRNA (green) and telomere (red). Enlarged images of the boxed region in the upper panel are shown in the lower panel. Line plot illustrates the signal of NRDC-intronic RNA and the telomeric signal on the white dotted line in the left merged image.

Finally, I wondered if the ncRNAs associated with telomeres were common among other ALT cell lines, I performed RNA-FISH and simultaneous telomere labeling in Sa-OS, GM847, VA13 cell lines and U2-OS cell lines obtained from other laboratories. While I was able to observe the colocalization of SPOCK3 foci with telomere in U2-OS cell line from another laboratory (Figure 52), SPOCK3 intronic RNA foci in the other three ALT cell lines were not colocalized with telomere (Figure 53). These results showed that the intronic ncRNAs associated with telomeres are unique in the U2-OS cell line.
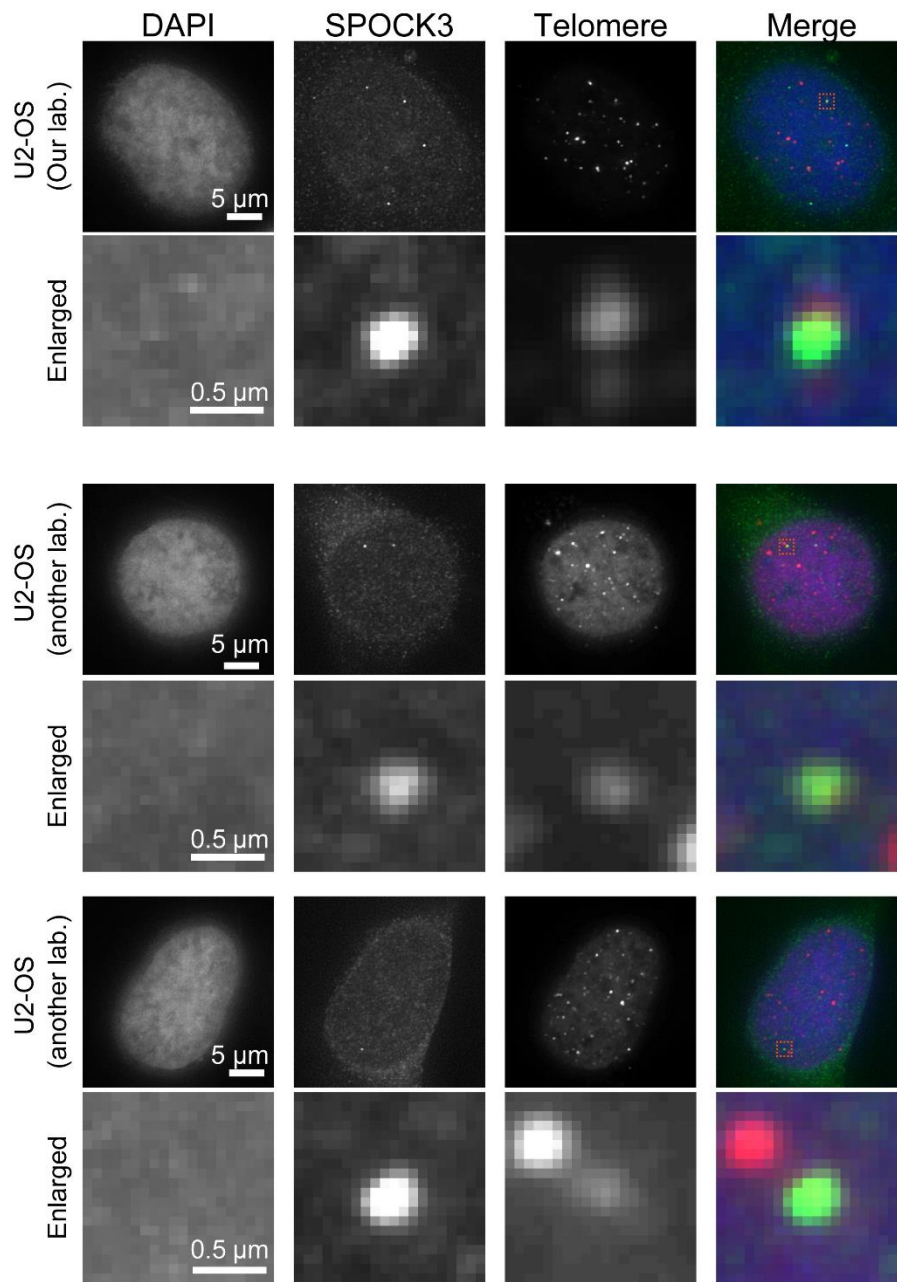
Figure 52. Localization of SPOCK3 intron RNA at telomeres in U2-OS cell lines.
RNA-FISH for ncRNAs and the simultaneous telomere labeling with fluorescent PI polyamide probe in two U2-OS cell lines from our laboratory and another laboratory, respectively. First column, DAPI signal; second column, ncRNA signal (SPOCK3 intron); third column, fluorescent PI polyamide, SiR-TDi59B signal (telomere); fourth column, merged image of DAPI (blue), ncRNA (green) and PI polyamide (red). Upper panel shows the result of our U2-OS cells. Middle and lower panels show the result of another U2-OS cell.
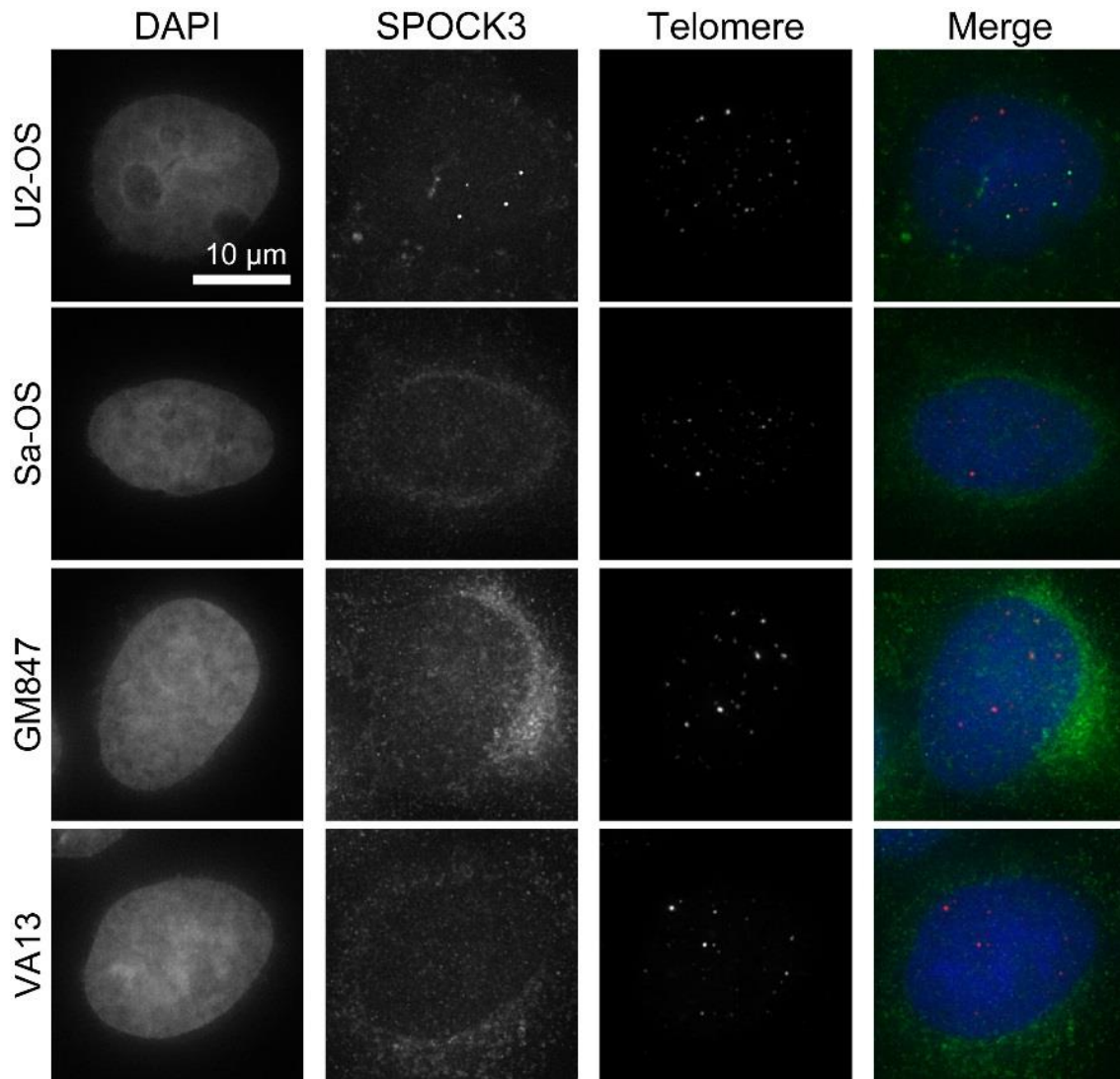
Figure 53. Localization of SPOCK3 intron RNA at telomeres in other ALT cell lines.
RNA-FISH for ncRNAs and the simultaneous telomere labeling with fluorescent TH59 in other
ALT cell lines (Sa-OS, GM847, and VA13). First column, DAPI signal; second column, ncRNA
signal (SPOK3 intron); third column, fluorescent PI polyamide, fluorescent TH59 signal (telomere);
fourth column, merged image of DAPI (blue), ncRNA (green) and telomere (red). Cells were
hybridized with the RNA probes (green), and then stained with telomere (red) and DAPI (blue).

2.3. Discussion

I demonstrated that PI-PRICh was able to be used for purification of both proteins and RNAs associating with the telomeric chromatin, whereas the previous methods could identify only protein composition (Dejardin and Kingston 2009; Grolimund et al. 2013; Liu et al. 2017). By using PI-PRICh, I succeeded in identification of the known three classes of ncRNAs associated with telomeric chromatin with >100 fold enrichment from mouse MEL cells: First class is a trans-acting ncRNA bound as a part of a ribonucleoprotein such as TERC of the telomerase complex. Second is a cis-acting ncRNA such as TERRA that acts as a major modulator of telomere maintenance. Third is ncRNAs associated as nascent transcripts from subtelomeric regions and interstitial telomeric sequences (ITS). This accomplishment of the comprehensive identification of ncRNAs with high enrichment relies on a unique binding mode of PI polyamide to the target sequences of double-stranded DNA through minor groove so that PI polyamide has no affinity to single-stranded RNA. This property reduces contamination of abundant messenger RNAs and ribosomal RNAs during telomeric chromatin isolation by PI-PRICh, whereas the chromatin isolation with nucleic acid probes can potentially have abundant contaminant RNAs, because the nucleic acid probes often hybridizes RNAs nonspecifically by mismatch pairing.

PI-PRICh is distinct from recently published methods for studying a genome-wide RNA-DNA contact, including MARGI, GRID-seq, and ChAR-seq (Sridhar et al. 2017; Li et al. 2017; Bell et al. 2018). With these techniques joint chromatin-associated RNAs are immobilized to DNA by proximity ligation after restriction digestion, forming RNA-DNA chimeric sequences for sequencing. The obtained global RNA-chromatin

interactomes, however, have to exclude the repetitive sequence regions due to lack of

the restriction recognition site. PI-PRICh is thus an excellent approach for the

comprehensive identification of the chromatin-associated ncRNAs on the repetitive

sequence occupying approximately 45% of the human genome (Lander et al. 2001). PI-

PRICh can also be applied to other repetitive sequences such as centromere for

identification of new chromatin-associated RNAs.

I used PI-PRICh to compare telomeric repeats associated ncRNAs between ALT cells

with highly recombinogenic telomeres and telomerase-active cells. PI-PRICh detected

TERRA on the telomere and telomeric repeats chromatin in both ALT and telomerase-

positive cells at a comparable amount. Although it was reported that TERRA is

expressed in ALT cells more than telomerase-positive cells and supposed to trigger

recombination through the hybridization with telomeric DNA (Arora and Azzalin 2015;

Graf et al. 2017), our results imply that not all of the TERRA might bind to telomeric

repeats in ALT cells.

I also identified novel ncRNAs derived from the introns of genic locus in the ALT cells.

The intron regions (SPOCK3, USP16, CKS1B, NRDC, KIAA1671, PAM, and RBBP8)

commonly included insertion of short telomeric repeats. As DNA-seq shows that the

intronic DNA regions were also enriched in the pull-down fraction from U2-OS cells,

TH59-DB may bind to the inserted telomeric repeats in the PI-PRICh experiment, and

pull-down those ncRNAs nascently transcribed by RNA polymerase II. Interestingly,

some of them were associated with telomeres in nucleus. One possible mechanism by

which the intronic RNAs localize at telomere is that the telomeric binding proteins such

as orphan nuclear receptors might tether the inserted telomeric repeats to telomeres (Marzec et al. 2015, Figure 54). This clustering of the intron regions and telomeres might be facilitated by ncRNA. The ncRNAs might make the unique chromatin state in ALT cells and involve their high recombinogenic property.
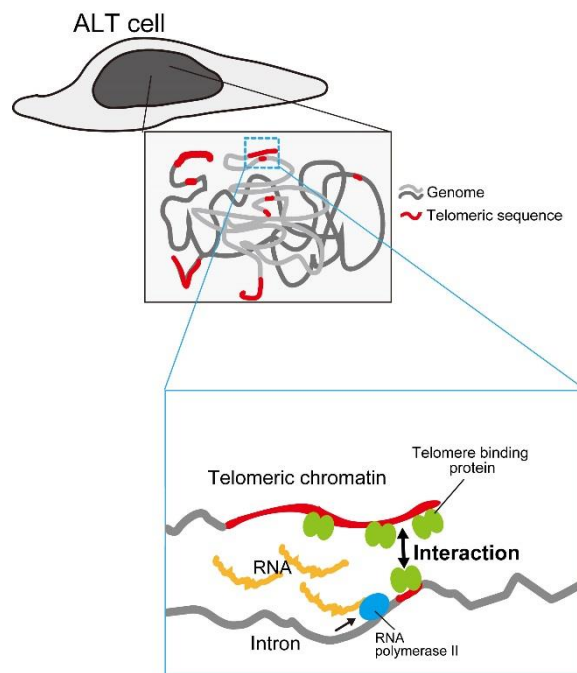
Figure 54. Model of how to associate de novo intronic RNAs with telomeric chromatin in ALT cells.

2.4. Material and Methods

Synthesis

Telomere PI polyamide probe (TH59-DB), which is tandem dimer TH59 labeled with a long spacer arm (24 PEG) and a biotin analogue (desthiobiotin), for chromatin affinity purification was synthesized as described in ref (Maeshima, Janssen, and Laemmli 2001; Kawamoto et al. 2013). Fluorescent TH59, silicon rhodamine (SiR)-TDi59B, (tandem dimer connected by a longer hinge) for telomere visualization was synthesized as described in ref (Hirata et al. 2014).

Telomere staining with TH59-DB

HeLa1.3 cells were maintained at 37°C (5% CO_2) in DMEM containing 10% FBS. For polyamide staining, cells were grown on coverslips coated with poly-lysine. The cell coverslips were washed in phosphate-buffered saline (PBS) twice and fixed with 1.85% formaldehyde in PBS for 15 min at room temperature. For blocking, the cells were treated with 10% normal goat serum (NGS) (Millipore) in TE buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA) for 30 min at room temperature. After a brief rinse with TE buffer, the cells were incubated with 10% NGS, 100 nM TH59-DB in DMF, mouse anti-TRF2 (abcam) and 0.5 μg/mL DAPI in TE buffer at 37°C for 1 hr. After washing with TEN200 buffer, cells were incubated with 10% NGS, streptavidin Alexa488 (Invitrogen) and anti-mouse Alexa594 (Invitrogen) in TE for 1 hr at room temperature. Slips were washed with TEN200 buffer (five times for 3 min), then mounted. Cell images were recorded with a DeltaVision microscope and deconvolved to eliminate out-of-focus blur to obtain clearer pictures. The deconvolved images were projected ('Quick Projection' tool) to obtain the maximum intensity of telomere signals.

Plasmid pull-down assay

This assay was based on ref (Ide and Dejardin 2015). For each assay, 200 ng linearized plasmids (telomeric repeats containing-plasmid and empty vector) were resuspended in LB3JD buffer (10 mM HEPES-NaOH, pH 7.7, 100 mM NaCl, 2 mM EDTA, 1 mM EGTA, 0.2% SDS, 0.1% SLS) containing 0.5 µM TH59-DB. The mixture was incubated at 37°C for 30 min. During incubation, 15 µl of MyOne C1 beads (Invitrogen) was added. After incubation at room temperature for 30 min with shaking, beads were immobilized on a magnetic stand. The supernatant was collected as flow-through. Beads were washed five times with 1 mL of LB3JD at room temperature, resuspended in LB3JD containing 12.5 mM D-biotin (Invitrogen, #B-20656), and incubated at 65°C for 15 min for elution. One-tenth volumes of the input, flow-through and elution fraction were analyzed by agarose gel electrophoresis, followed by ethidium bromide staining. Band intensities were quantified relative to input by ImageJ in three independent experiments. As a negative control, we incubated TH59-DB with a double strand DNA oligonucleotide (TTAGGG/CCCTAA)$_4$ for 30 min and 60 min at 37°C prior to the addition to linearized plasmids and chromatin solution, respectively. The telomeric oligo-premixed TH59-DB was used for telomeric chromatin isolation and was named "masked" TH59-DB.

Chromatin preparation of mouse (MEL) cells by TH59-DB

MEL cells were grown in DMEM with 10% FBS and 2 mM L-glutamine. 2 L of MEL cells (~2-4×10^6 cells/ mL) grown in a roller bottle were crosslinked with 3.7% formaldehyde in PBS for 30 min at room temperature. After washing four times with PBS, cells were transfered in sucrose buffer (0.3 M sucrose, 10 mM HEPES-NaOH, pH

7.7, 1% Triton X-100, 2 mM MgOAc) lysed with 20 strokes of a Dounce homogenizer with a tight pestle. Chromatin was pelleted by centrifugation at 3,200g for 10 min at 4°C. The pellet was resuspended in the same volume of glycerol buffer (25% glycerol, 10 mM HEPES-NaOH, pH 7.9, 0.1 mM EDTA, 0.1 mM EGTA, 5 mM MgOAc), and then frozen in liquid nitrogen and stored at −80°C or used immediately for telomeric-chromatin-isolation. The chromatin pellet was washed with PBS for five times and with LB3JD containing 1mM phenylmethylsulphonyl fluoride (LB3JD-PMSF). After centrifugation at 3,000g for 8 min, the pellet was resuspended in a 1.5 × volume of LB3JD-PMSF buffer and then passed three times through a French press at 25,000 p.s.i. at room temperature. The solubilized chromatin sample was collected at 20,000g for 15 min at 4°C and was heated at 58°C for 5 min. LB3JD-PMSF-pre-equilibrated streptavidin agarose beads (Thermo) were added and the sample was incubated at 4°C overnight. The mixture was applied to Sephacryl S-400-HR spin columns (Roche). Sample was centrifuged at 20,000g for 15 min, and SDS was added to the supernatant to a final concentration of 0.2 %. About 30 mg chromatin was incubated with 600 pmol of TH59-DB or masked TH59-DB for 2 hr at 37°C. The sample was then centrifuged at 20,000g for 15 min, and the supernatant was added to FG beads-streptavidin beads (Tamagawa-seiki) pre-equilibrated with LB3JD. The sample was incubated at room temperatue for 2 hr on nutator. Beads were washed with 10 mL of LB3JD for five times at room temperature. Beads were collected in 1.5 mL tube and additionally washed with shaking in LB3JD containing 30 mM NaCl at 42°C for 10 min and then in LB3JD containing 10 mM NaCl at 42°C for 10 min. Beads were resuspended in LB3JD containing 12.5 mM D-biotin, and incubated at 37°C for 1 hr for elution with shaking. Cleared supernatant was collected and kept at -80°C until use.

Chromatin sample preparation of adherent cells by TH59-DB

HeLa1.3 and U2-OS cells (~10^9 cells) were grown in DMEM with 10% FBS. The cells were washed with PBS and crosslinked with 3.7% formaldehyde in PBS for 30 min at room temperature. After washed with cold PBS twice, cells were scraped with scraping buffer (0.05% Tween20 in PBS). The following procedures are same as chromatin preparation of MEL cells described above.

RNA extraction and preparation of cDNA libraries for Next-Generation Sequencing

To extract total RNA, >2.5 mg isolated chromatins were incubated in de-crosslinking buffer (32 mM Tris-HCl pH 8.0, 320 μg/mL Protease K, 0.8% SDS) for 6 hr at 65°C. An equal volume of acidic phenol was added, then vortexed, and incubated for 5 min at room temperature. After adding an equal volume of chloroform:isoamylalcohol (24:1, v/v) and centrifuging at 16,000 g for 5 min at 4°C, upper aqueous phase was transferred to new tube. This step was repeated again. One-tenth volume of 3 M sodium acetate, glycogen, and double volume of ethanol were added and incubated overnight at -80°C. After centrifugation at 20,000 g for 30 min at 4°C, then pellet were retained, washed with 70% ethanol, dried up and dissolved in diethylpyrocarbonate (DEPC)-treated water and kept at -80°C until use. Total RNA content of each sample was measured by using Qubit (Thermo Fisher Scientific Inc.), and the quality of RNA samples was assessed by Agilent 2100 Bioanalyzer using Agilent RNA 6000 pico kit. cDNA libraries were synthesized by the SMARTer Stranded Total RNA-Seq Kit v2 (Takara). The size distribution of the libraries were checked by an Agilent 2100 Bioanalyzer using Agilent

High Sensitivity DNA kit. Pooled amplicon library was sequenced with paired-end $2 \times$ 150 bp reads on the Illumina MiSeq platform.

Detection of TERRA/ARRET in the RNA-Seq data

To estimate the number of putative TERRAs, reads containing $(TTAGGG)_5$ or $(CCCTAA)_5$ repeats were extracted from each fastq file using the grep command in the UNIX system.

Data processing and software for RNA-seq

First, adapter sequences were removed and the reads were trimmed for low quality with CUTADAPT and PRINSEQ, using a composite set of Illumina adapters, minimum quality score of 20, and a minimum length of 25 (Martin 2011; Schmieder and Edwards 2011). Filtered mouse and human sequence data were aligned to the mouse mm10 genome and human hg38 genome, respectively using HISAT2 (Kim, Langmead, and Salzberg 2015). To find telomere-enriched ncRNAs, the mapped fragments in TH59-DB pull-down fractions were assembled into RNA transcripts and annotated using Cufflinks software (Trapnell et al. 2010). The amounts of RNA transcripts were compared based on fragments per kilobase of per million mapped reads (FPKM) between samples using featureCounts and R (Liao, Smyth, and Shi 2014; R Core Team 2016). RNA-seq data was also visualized by igvtools (Robinson et al., 2011; Thorvaldsdottir, Robinson, and Mesirov 2013).

Mapping of telomere-enriched RNAs in MEL cells to a TERRA locus

For mapping to a mouse TERRA locus, we focused on a 30-kb region adjacent to the telomere repeat of each long (q) arm of chromosome (ref.Maria Blasco 2014 NatureCom). Centromere-adjacent telomeres on short (p) arm were not analyzed because they were not sequenced. Using featureCounts, FPKM values of subtelomeric region (supplementary Table S7) were compared between input and TH59-DB pull-down fraction.

Mapping of telomere-enriched RNAs in HeLa1.3 and U2-OS to a TERRA locus

For mapping to a human TERRA locus on X/Y chromosome, custom gene annotation that contains the sequence of TelBam3.4 (Nergadze et al. 2009; Negishi et al. 2015) was generated for quantification of telomere-associated TERRA. Individual reads of RNA-Seq data were mapped to the custom gene and human genome hg38 by Bowtie2 with default setting except an extraction of uniquely mapped reads, respectively (Langmead and Salzberg 2012). Mapped reads were assembled into transcripts and annotated by Cufflinks. The number of reads overlapped to region downstream of the transcription start site in TelBam3.4 (Negishi et al. 2015) were counted by featureCounts and normalized as reads per kilobase of exon model per million mapped reads (RPKM) on TelBam3.4 and human genome hg38.

DNA extraction and preparation of cDNA libraries for Next-Generation Sequencing

To extract DNA, the pull down chromatin fractions were incubated in de-crosslinking buffer (50 mM Tris-HCl pH 8.0, 200 μg/mL Protease K, 2% SDS) for overnight at

65°C. An equal volume of phenol:chloroform:isoamylalcohol (25:24:1,v/v) was added and voltexed, and centrifuge at 16,000 g for 5 min at room temperature. The upper aqueous phase was transferred to new tube. This step was repeated again. One-tenth volume of 3 M sodium acetate, glycogen, and the double volume of ethanol were added and incubated for 2 hr at -80°C. After centrifugation at 20,000 g for 30 min at 4ºC, then pellet were retained, washed with 70% ethanol, dried up and dissolved in 10 μg/mL RNaseA in DNA dilution buffer attached to DNA SMART ChIP-Seq Kit (Takara). After incubation for 1 hr at 37°C, DNA was purified by the phenorl-chloroform extraction and ethanol precipitation to remove RNaseA. DNA samples were kept at -30°C until use. DNA concentration of each sample was measured by Qubit (Thermo Fisher Scientific Inc.). cDNA libraries were synthesized by DNA SMART ChIP Seq Kit. The size distribution of the libraries were checked by an Agilent 2100 Bioanalyzer using Agilent High Sensitivity DNA kit. Pooled amplicon library was sequenced with paired-end 2×250 bp reads on the Illumina MiSeq platform.

Detection of chimeric sequences of ncRNA with telomeric repeats in the DNA-seq
To confirm telomere repeats were inserted in near the genomic region of ncRNA, sequence reads containing ncRNA sequences were extracted from each fastq file using the grep command in the UNIX system. The sequences for grep command were shown in Supplementary Table S8.

Data processing and software for DNA-seq
First, adapter sequences were removed and the reads were trimmed for low quality with

85

CUTADAPT and PRINSEQ, using a composite set of Illumina adapters, minimum quality score of 20, and a minimum length of 25. Filtered human sequence data was aligned to human hg38 genome using Bowtie2. DNA-seq data was also visualized by igvtools.

Probes for RNA-FISH

For construction of RNA-FISH probes, DNA fragments were amplified by PCR from U2-OS genomic DNA or HeLa1.3 genomic DNA and inserted into pGEM-Teasy plasmid (Promega) for *in vitro* transcription using In-Fusion HD Cloning Kit (Takara). The primers were shown in Supplementary Table S9. After linearization of each template plasmid with the appropriate restriction enzyme or amplification of DNA fragments containing the T7 promoter and intron region by PCR, digoxigenin (DIG)-labeled RNA probes were generated using DIG RNA Labeling Kit (Sigma). The quality of transcripts were checked by agarose gel electrophoresis. RNA probes were then purified with illustra MicroSpin G-25 Columns (GE Healthcare) to remove extra NTPs, and were then added into an equal volume of formamide (Sigma) and kept at -30°C until use. For RNA probe to visualize endogenous TERRA, the plasmid containing a 750-bp telomere fragment was used as a template for *in vitro* transcription using DIG RNA Labeling Kit.

RNA fluorescence in situ hybridization

Immnofluorescence combined with RNA-FISH was performed as described in ref (Kawaguchi et al. 2015). Briefly, cells were seeded on poly-lysine coated coverslips and

incubated overnight at 37°C in 5% CO_2. Coverslips were washed with PBS and fixed with 3.7% formaldehyde in PBS at room temperature for 10 min. After 5 min-washing with PBS, cells were permeabilized with 0.1% Triton X-100 in PBS for 10 min and then washed with PBS for 5 min. The coverslips were incubated with prehybridization solution [2×SSC, 1× Denhards solution, 50% (v/v) formamide, 10 mM EDTA pH 8.0, 100 μg/mL yeast tRNA, and 0.01% Tween-20] at 55°C for 1 hr. The prehybridized coverslips were then incubated with hybridization solution [2×SSC, 1× Denhards solution, 50% (v/v) formamide, 10 mM EDTA pH 8.0, 100 μg/mL yeast tRNA, and 0.01% Tween-20, 5% (w/v) dextran sulfate and DIG-labeled RNA probe] at 55°C for overnight. After hybridization, the coverslips were washed twice with wash buffer [2 × SSC, 50% (v/v) formamide and 0.01% Tween-20] at 55°C for 30 min twice. To remove non-specific binding of RNA probe, cell were incubated in RNaseA buffer [1 μg/mL RNaseA, 0.5 M NaCl, 10 mM Tris-HCl pH 8.0, 1 mM EDTA] for 1 hr at 37°C. After RNaseA treatment, cells were washed with wash buffer2 (2 × SSC, 0.01% Tween-20) and wash buffer3 (0.2 × SSC, 0.01% Tween-20) for 30 min at 55°C. For detection, coverslips were washed in TBST (Tris-buffered saline solution containing 0.01% Tween-20) for 5 min at room temperature, incubated with blocking solution [1× Blocking reagent (Sigma) in TBST] for 5 min at room temperature, and then incubated with anti-DIG Rhodamine (Sigma) and 15 nM SiR-TDi59B for 90 min at room temperature. After staining with 0.5 μg/mL DAPI in TBST for 5 min at room temperature, unbound antibodies were removed by washing three times with TBST for 5 min. Coverslips were then mounted and image acquisition was performed as described in Chapter 1. Signal intensities were measured by Plot Profile tools in Fiji and line plots were created using R software (Schindelin et al. 2012).

Genomic PCR to test telomeric sequence insertion

To extract genomic DNA, cells grown on φ10 cm dish were collected and incubated in

lysis buffer (10 mM Tris-HCl pH 8.0, 0.1M EDTA, 0.5% SDS, 20 µg/mL RNaseA) for

1hr at 37°C. After adding 100 µg/mL Proteinase K and the cell lysate was incubated for

3 hr at 50 °C. DNA was isolated using phenol-chloroform extraction and ethanol

precipitation. To test if telomere sequence locates in the flanking region of the introns,

the genomic DNA of HeLa1.3 and U2-OS were used for PCR reaction using primer sets

shown in Supplementary Table S10. PCR products were analyzed by agarose

electrophoresis, followed by ethidium bromide staining.

Acknowledgement

Chapter 2: Dr. Mahito Sadaie (Tokyo University of Science) for providing Sa-OS, GM847, and VA13 cell lines, Dr. Tom Misteli for providing another U2-OS cell line. Through my Ph.D. course, my committee members, Prof. Akatsuki Kimura, Prof. Hiroyuki Araki, Prof. Tatsumi Hirata, and Prof. Masato T. Kanemaki in National Institute of Genetics/SOKENDAI, always gave me helpful advices and supports. I would like to thank them very much.

I would also like to thank the members of Maeshima laboratory in National Institute of Genetics/SOKENDAI, Assistant Prof. Kayo Hibino, Dr. Tadasu Nozaki, Dr. Ryosuke Imai, Dr. Eloise Prieto, Dr. Michael Babokhov, Mr. Ryosuke Nagashima, Ms. Mai Tambo, Mr. Tomohiro Yamaguchi, Ms. Sachiko Tamura, Ms. Kaeko Nakaguchi, and Ms. Hiroko Ochi for helpful discussions, technical assistances, and kind encouragements.

I would also like to thank Atsuko Yamaguchi for her support for mental and physical health by making delicious food in National Institute of Genetics.

Finally, I would like to thank my family and relatives, especially my parents Yoshito and Kazue, my sister Reika, my grandmother Kuniko Hamada, and my aunt Fumie Hamada for their continuous support and encouragement through my life.

References

✧ Arora, R., & Azzalin, C. M. (2015). Telomere elongation chooses TERRA ALTernatives. RNA Biol, 12(9), 938-941.

✧ Arora, R., Lee, Y., Wischnewski, H., Brun, C. M., Schwarz, T., & Azzalin, C. M. (2014). RNaseH1 regulates TERRA-telomeric DNA hybrids and telomere maintenance in ALT tumour cells. Nat Commun, 5, 5220.

✧ Azzalin, C. M., Reichenbach, P., Khoriauli, L., Giulotto, E., & Lingner, J. (2007). Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. Science, 318(5851), 798-801.

✧ Balk, B., Maicher, A., Dees, M., Klermund, J., Luke-Glaser, S., Bender, K., & Luke, B. (2013). Telomeric RNA-DNA hybrids affect telomere-length dynamics and senescence. Nat Struct Mol Biol, 20(10), 1199-1205. doi:10.1038/nsmb.2662

✧ Bando, T., & Sugiyama, H. (2006). Synthesis and biological properties of sequence-specific DNA-alkylating pyrrole-imidazole polyamides. Accounts of chemical research, 39(12), 935-944.

✧ Bell, J. C., Jukam, D., Teran, N. A., Risca, V. I., Smith, O. K., Johnson, W. L., Straight, A. F. (2018). Chromatin-associated RNA sequencing (ChAR-seq) maps genome-wide RNA-to-DNA contacts. Elife, 7.

✧ Blackburn, E. H. (2010). Telomeres and telomerase: the means to the end (Nobel lecture). Angewandte Chemie, 49(41), 7405-7421.

✧ Blackburn, E. H., & Collins, K. (2011). Telomerase: an RNP enzyme synthesizes DNA. Cold Spring Harb Perspect Biol, 3(5).

✧ Blackledge, M. S., & Melander, C. (2013). Programmable DNA-binding small molecules. Bioorganic & medicinal chemistry, 21(20), 6101-6114.

✧ Bryan, T. M., Englezou, A., Dalla-Pozza, L., Dunham, M. A., & Reddel, R. R. (1997). Evidence for an alternative mechanism for maintaining telomere length in human tumors and tumor-derived cell lines. Nature medicine, 3(11), 1271-1274.

✧ Bryan, T. M., Englezou, A., Dunham, M. A., & Reddel, R. R. (1998). Telomere length dynamics in telomerase-positive immortal human cell populations. Experimental cell research, 239(2), 370-378.

✧ Bryan, T. M., Englezou, A., Gupta, J., Bacchetti, S., & Reddel, R. R. (1995). Telomere elongation in immortal human cells without detectable telomerase activity. The EMBO journal, 14(17), 4240-4248.

✧ Chen, B., Gilbert, L. A., Cimini, B. A., Schnitzbauer, J., Zhang, W., Li, G. W., Huang, B. (2013). Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. Cell, 155(7), 1479-1491.

✧ Chenoweth, D. M., & Dervan, P. B. (2009). Allosteric modulation of DNA by small molecules. Proceedings of the National Academy of Sciences of the United States of America, 106(32), 13175-13179.

✧ Chu, H. P., Cifuentes-Rojas, C., Kesner, B., Aeby, E., Lee, H. G., Wei, C., Lee, J. T. (2017). TERRA RNA Antagonizes ATRX and Protects Telomeres. Cell, 170(1), 86-101 e116.

✧ Dejardin, J., & Kingston, R. E. (2009). Purification of proteins associated with specific genomic Loci. Cell, 136(1), 175-186.

✧ Deng, W., Shi, X., Tjian, R., Lionnet, T., & Singer, R. H. (2015). CASFISH: CRISPR/Cas9-mediated in situ labeling of genomic loci in fixed cells. Proceedings of the National Academy of Sciences of the United States of America, 112(38), 11870-11875.

✧ Dervan, P. B. (2001). Molecular recognition of DNA by small molecules. Bioorganic & medicinal chemistry, 9(9), 2215-2235.

✧ Dervan, P. B., Doss, R. M., & Marques, M. A. (2005). Programmable DNA binding oligomers for control of transcription. Current medicinal chemistry. Anti-cancer agents, 5(4), 373-387.

✧ Dervan, P. B., & Edelson, B. S. (2003). Recognition of the DNA minor groove by pyrrole-imidazole polyamides. Current opinion in structural biology, 13(3), 284-299.

✧ Dunham, M. A., Neumann, A. A., Fasching, C. L., & Reddel, R. R. (2000). Telomere maintenance by recombination in human cells. Nature genetics, 26(4), 447-450.

✧ Ferlicot, S., Youssef, N., Feneux, D., Delhommeau, F., Paradis, V., & Bedossa, P. (2003). Measurement of telomere length on tissue sections using quantitative fluorescence in situ hybridization (Q-FISH). The Journal of pathology, 200(5), 661-666.

✧ Flynn, R. L., Centore, R. C., O'Sullivan, R. J., Rai, R., Tse, A., Songyang, Z., Zou, L. (2011). TERRA and hnRNPA1 orchestrate an RPA-to-POT1 switch on telomeric single-stranded DNA. Nature, 471(7339), 532-536.

✧ Fujita, T., Yuno, M., Okuzaki, D., Ohki, R., & Fujii, H. (2015). Identification of non-coding RNAs associated with telomeres using a combination of enChIP and RNA sequencing. PLoS One, 10(4), e0123387.

✧ Fujiwara, Y., Komiya, T., Kawabata, H., Sato, M., Fujimoto, H., Furusawa, M., & Noce, T. (1994). Isolation of a DEAD-family protein gene that encodes a murine

homolog of Drosophila vasa and its specific expression in germ cell lineage. Proceedings of the National Academy of Sciences of the United States of America, 91(25), 12258-12262.

✧ Graf, M., Bonetti, D., Lockhart, A., Serhal, K., Kellner, V., Maicher, A. Luke, B. (2017). Telomere Length Determines TERRA and R-Loop Regulation through the Cell Cycle. Cell, 170(1), 72-85 e14.

✧ Grolimund, L., Aeby, E., Hamelin, R., Armand, F., Chiappe, D., Moniatte, M., & Lingner, J. (2013). A quantitative telomeric chromatin isolation protocol identifies different telomeric states. Nat Commun, 4, 2848.

✧ Hastie, N. D., Dempster, M., Dunlop, M. G., Thompson, A. M., Green, D. K., & Allshire, R. C. (1990). Telomere reduction in human colorectal carcinoma and with ageing. Nature, 346(6287), 866-868.

✧ Henderson, S., Allsopp, R., Spector, D., Wang, S. S., & Harley, C. (1996). In situ analysis of changes in telomere size during replicative aging and cell transformation. The Journal of cell biology, 134(1), 1-12.

✧ Hirata, A., Nokihara, K., Kawamoto, Y., Bando, T., Sasaki, A., Ide, S., Sugiyama, H. (2014). Structural evaluation of tandem hairpin pyrrole-imidazole polyamides recognizing human telomeres. Journal of the American Chemical Society, 136(32), 11546-11554.

✧ Ide, S., & Dejardin, J. (2015). End-targeting proteomics of isolated chromatin segments of a mammalian ribosomal RNA gene promoter. Nat Commun, 6, 6674.

✧ Kawaguchi, T., Tanigawa, A., Naganuma, T., Ohkawa, Y., Souquere, S., Pierron, G., & Hirose, T. (2015). SWI/SNF chromatin-remodeling complexes function in noncoding RNA-dependent assembly of nuclear bodies. Proc Natl Acad Sci U S A, 112(14), 4304-4309.

✧ Kawamoto, Y., Bando, T., Kamada, F., Li, Y., Hashiya, K., Maeshima, K., & Sugiyama, H. (2013). Development of a new method for synthesis of tandem hairpin pyrrole-imidazole polyamide probes targeting human telomeres. Journal of the American Chemical Society, 135(44), 16468-16477.

✧ Kawamoto, Y., Sasaki, A., Chandran, A., Hashiya, K., Ide, S., Bando, T., Sugiyama, H. (2016). Targeting 24 bp within Telomere Repeat Sequences with Tandem Tetramer Pyrrole-Imidazole Polyamide Probes. J Am Chem Soc.

✧ Kawamoto, Y., Sasaki, A., Hashiya, K., Ide, S., Bando, T., Maeshima, K., & Sugiyama, H. (2015). Tandem trimer pyrrole–imidazole polyamide probes targeting 18 base pairs in human telomere sequences. Chemical Science, 6, 2307-2312.

✧ Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. Nat Methods, 12(4), 357-360.

✧ Kim, N. W., Piatyszek, M. A., Prowse, K. R., Harley, C. B., West, M. D., Ho, P. L., . . . Shay, J. W. (1994). Specific association of human telomerase activity with immortal cells and cancer. Science, 266(5193), 2011-2015.

✧ Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., International Human Genome Sequencing, Consortium. (2001). Initial sequencing and analysis of the human genome. Nature, 409(6822), 860-921.

✧ Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods, 9(4), 357-359.

✧ Lansdorp, P. M., Verwoerd, N. P., van de Rijke, F. M., Dragowska, V., Little, M. T., Dirks, R. W., . . . Tanke, H. J. (1996). Heterogeneity in telomere length of human chromosomes. Human molecular genetics, 5(5), 685-691.

✧ Levsky, J. M., & Singer, R. H. (2003). Fluorescence in situ hybridization: past, present and future. Journal of cell science, 116(Pt 14), 2833-2838.

✧ Li, X., Zhou, B., Chen, L., Gou, L. T., Li, H., & Fu, X. D. (2017). GRID-seq reveals the global RNA-chromatin interactome. Nat Biotechnol, 35(10), 940-950.

✧ Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics, 30(7), 923-930.

✧ Liu, X., Zhang, Y., Chen, Y., Li, M., Zhou, F., Li, K. Xu, J. (2017). In Situ Capture of Chromatin Interactions by Biotinylated dCas9. Cell, 170(5), 1028-1043 e1019.

✧ Lopez de Silanes, I., Grana, O., De Bonis, M. L., Dominguez, O., Pisano, D. G., & Blasco, M. A. (2014). Identification of TERRA locus unveils a telomere protection role through association to nearly all chromosomes. *Nat Commun, 5*, 4723.

✧ Ma, H., Reyes-Gutierrez, P., & Pederson, T. (2013). Visualization of repetitive DNA sequences in human chromosomes with transcription activator-like effectors. Proceedings of the National Academy of Sciences of the United States of America, 110(52), 21048-21053.

✧ Maeshima, K., Ide, S., Hibino, K., & Sasai, M. (2016). Liquid-like behavior of chromatin. Curr Opin Genet Dev, 37, 36-45.

✧ Maeshima, K., Janssen, S., & Laemmli, U. K. (2001). Specific targeting of insect and vertebrate telomeres with pyrrole and imidazole polyamides. The EMBO journal, 20(12), 3218-3228.

◇ Maeshima, K., Yahata, K., Sasaki, Y., Nakatomi, R., Tachibana, T., Hashikawa, T., Imamoto, N. (2006). Cell-cycle-dependent dynamics of nuclear pores: pore-free islands and lamins. J Cell Sci, 119(Pt 21), 4442-4451.

◇ Martin, M. (2011). Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads. EMBnet Journal, 17, 10-12.

◇ Marzec, P., Armenise, C., Perot, G., Roumelioti, F. M., Basyuk, E., Gagos, S. Dejardin, J. (2015). Nuclear-receptor-mediated telomere insertion leads to genome instability in ALT cancers. Cell, 160(5), 913-927.

◇ Meeker, A. K., Gage, W. R., Hicks, J. L., Simon, I., Coffman, J. R., Platz, E. A., De Marzo, A. M. (2002). Telomere length assessment in human archival tissues: combined telomere fluorescence in situ hybridization and immunostaining. The American journal of pathology, 160(4), 1259-1268.

◇ Meeker, A. K., Hicks, J. L., Gabrielson, E., Strauss, W. M., De Marzo, A. M., & Argani, P. (2004). Telomere shortening occurs in subsets of normal breast epithelium as well as in situ and invasive carcinoma. The American journal of pathology, 164(3), 925-935.

◇ Meeker, A. K., Hicks, J. L., Platz, E. A., March, G. E., Bennett, C. J., Delannoy, M. J., & De Marzo, A. M. (2002). Telomere shortening is an early somatic DNA alteration in human prostate tumorigenesis. Cancer research, 62(22), 6405-6409.

◇ Miyanari, Y., Ziegler-Birling, C., & Torres-Padilla, M. E. (2013). Live visualization of chromatin dynamics with fluorescent TALEs. Nature structural & molecular biology, 20(11), 1321-1324.

◇ Nabetani, A., & Ishikawa, F. (2011). Alternative lengthening of telomeres pathway: recombination-mediated telomere maintenance mechanism in human cells. Journal of biochemistry, 149(1), 5-14.

◇ Negishi, Y., Kawaji, H., Minoda, A., & Usui, K. (2015). Identification of chromatin marks at TERRA promoter and encoding region. Biochem Biophys Res Commun, 467(4), 1052-1057.

◇ Nergadze, S. G., Farnung, B. O., Wischnewski, H., Khoriauli, L., Vitelli, V., Chawla, R., Azzalin, C. M. (2009). CpG-island promoters drive transcription of human telomeres. RNA, 15(12), 2186-2194.

◇ R core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

❖ Rossiello, F., Aguado, J., Sepe, S., Iannelli, F., Nguyen, Q., Pitchiaya, S., d'Adda di Fagagna, F. (2017). DNA damage response inhibition at dysfunctional telomeres by modulation of telomeric DNA damage response RNAs. Nat Commun, 8, 13980.

❖ Sakellariou, D., Chiourea, M., Raftopoulou, C., & Gagos, S. (2013). Alternative lengthening of telomeres: recurrent cytogenetic aberrations and chromosome stability under extreme telomere dysfunction. Neoplasia, 15(11), 1301-1313.

❖ Sasaki, A., Ide, S., Kawamoto, Y., Bando, T., Murata, Y., Shimura, M., Maeshima, K. (2016). Telomere Visualization in Tissue Sections using Pyrrole-Imidazole Polyamide Probes. Sci Rep, 6, 29261.

❖ Scheel, C., Schaefer, K. L., Jauch, A., Keller, M., Wai, D., Brinkschmidt, C., Poremba, C. (2001). Alternative lengthening of telomeres is associated with chromosomal instability in osteosarcomas. Oncogene, 20(29), 3835-3844.

❖ Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. Nat Methods, 9(7), 676-682.

❖ Schluter, C., Duchrow, M., Wohlenberg, C., Becker, M. H., Key, G., Flad, H. D., & Gerdes, J. (1993). The cell proliferation-associated antigen of antibody Ki-67: a very large, ubiquitous nuclear protein with numerous repeated elements, representing a new kind of cell cycle-maintaining proteins. The Journal of cell biology, 123(3), 513-522.

❖ Schmieder, R., & Edwards, R. (2011). Quality control and preprocessing of metagenomic datasets. Bioinformatics, 27(6), 863-864.

❖ Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat Methods, 9(7), 671-675.

❖ Shay, J. W., & Bacchetti, S. (1997). A survey of telomerase activity in human cancer. European journal of cancer, 33(5), 787-791.

❖ Sieverling, L., Hong, C., Koser, S.D., Ginsbach, P., Kleinheinz, K., Hutter, B., Braun, D.M., Cortes-Ciriano, I., Xi, R., Kabbe, R., Park, P.J., Eils, R., Schlesner, M., Rippe, K., Jones, T.W.D., Brors, B., Feuerbach, L. (2018), Genomic footprints of activated telomere maintenance mechanisms in cancer. Nat Commun. 2018:in press, preprint; https://doi.org/10.1101/157560.

❖ Smogorzewska, A., & de Lange, T. (2004). Regulation of telomerase by telomeric proteins. Annual review of biochemistry, 73, 177-208.

❖ Sommerfeld, H. J., Meeker, A. K., Piatyszek, M. A., Bova, G. S., Shay, J. W., & Coffey, D. S. (1996). Telomerase activity: a prevalent marker of malignant human prostate tissue. Cancer research, 56(1), 218-222.

✧ Sridhar, B., Rivas-Astroza, M., Nguyen, T. C., Chen, W., Yan, Z., Cao, X., Zhong, S. (2017). Systematic Mapping of RNA-Chromatin Interactions In Vivo. Curr Biol, 27(4), 602-609.

✧ Takai, K. K., Hooper, S., Blackwood, S., Gandhi, R., & de Lange, T. (2010). In vivo stoichiometry of shelterin components. The Journal of biological chemistry, 285(2), 1457-1467.

✧ Tardat, M., & Dejardin, J. (2018). Telomere chromatin establishment and its maintenance during mammalian development. Chromosoma, 127(1), 3-18.

✧ Team, R Core. (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

✧ Thorvaldsdottir, H., Robinson, J. T., & Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform, 14(2), 178-192.

✧ Tokutake, Y., Matsumoto, T., Watanabe, T., Maeda, S., Tahara, H., Sakamoto, S., Furuichi, Y. (1998). Extra-chromosomal telomere repeat DNA in telomerase-negative immortalized cell lines. Biochemical and biophysical research communications, 247(3), 765-772.

✧ Toyooka, Y., Tsunekawa, N., Takahashi, Y., Matsui, Y., Satoh, M., & Noce, T. (2000). Expression and intracellular localization of mouse Vasa-homologue protein during germ cell development. Mechanisms of development, 93(1-2), 139-149.

✧ Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J. Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol, 28(5), 511-515.

✧ Trauger, J. W., Baird, E. E., & Dervan, P. B. (1996). Recognition of DNA by designed ligands at subnanomolar concentrations. Nature, 382(6591), 559-561.

✧ Wentzensen, I. M., Mirabello, L., Pfeiffer, R. M., & Savage, S. A. (2011). The association of telomere length and cancer: a meta-analysis. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology, 20(6), 1238-1250.

✧ Werner, M. S., & Ruthenburg, A. J. (2015). Nuclear Fractionation Reveals Thousands of Chromatin-Tethered Noncoding RNAs Adjacent to Active Genes. Cell Rep, 12(7), 1089-1098.

✧ White, S., Szewczyk, J. W., Turner, J. M., Baird, E. E., & Dervan, P. B. (1998). Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. Nature, 391(6666), 468-471.

✧ Yoshioka, H., McCarrey, J. R., & Yamazaki, Y. (2009). Dynamic nuclear organization of constitutive heterochromatin during fetal male germ cell development in mice. Biology of reproduction, 80(4), 804-812.

✧ Zakian, V. A. (2012). Telomeres: the beginnings and ends of eukaryotic chromosomes. Experimental cell research, 318(12), 1456-1460.

✧ Zappulla, D. C., & Cech, T. R. (2006). RNA as a flexible scaffold for proteins: yeast telomerase and beyond. Cold Spring Harb Symp Quant Biol, 71, 217-224.

✧ Zhang, C., Chen, X., Li, L., Zhou, Y., Wang, C., & Hou, S. (2015). The Association between Telomere Length and Cancer Prognosis: Evidence from a Meta-Analysis. PLoS One, 10(7), e0133174.

✧ Zhou, H. X., Rivas, G., & Minton, A. P. (2008). Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. Annu Rev Biophys, 37, 375-397.