AN IMAGING APPROACH TO EXAMINE TELOMERE DYNAMICS AND REGULATION OF GENE EXPRESSION WITH AGING

APPROVED BY SUPERVISORY COMMITTEE

Gaudenz Danuser, Ph.D.

Jerry W. Shay, Ph.D.

Yang Xie, Ph.D.

Khuloud Jaqaman, Ph.D.

Daniel J. Siegwart, Ph.D.

DEDICATION

To Mom and Dad.

AN IMAGING APPROACH TO EXAMINE TELOMERE DYNAMICS AND REGULATION OF GENE EXPRESSION WITH AGING

by

NING ZHANG

DISSERTATION / THESIS

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

August, 2020

ACKNOWLEDGMENTS

First of all, I would like to thank my mentors, Dr. Gaudenz Danuser and Dr. Jerry W. Shay for their tremendous support on my research projects. The scientific training I received from the graduate school at UT Southwestern has laid the foundation of critical thinking and continuous learning ability in my mind. Second, I appreciate the encouragement and help from my family, colleagues, and friends in the last a few year, which made me feel warm and energetic on the way of my mountain climb. Third, I really enjoy the intellectual freedom and heuristic discussions that happened in the labs, which helped me to grow from a toddler in science to a young scientist that has developed something useful to this community.

I would like to thank my thesis committee members, Dr. Yang Xie, Dr. Khuloud Jaqaman, Dr. Daniel Siegwart, for their support and valuable feedback from each meeting.

I would like to thank all present and previous members of Danuser lab. It is grateful to work with geniuses in the fields of computational biology, experimental biology, and microscopy. Thanks to the guidance and support from Dr. Sangyoon Han, Dr. Xiao Ma, Dr. Maria Bagonis, Dr. Kevin Dean, Dr. Assaf Zaritzky, Dr. Zhengwen Zhang, Dr. Zhuo Gan, Dr. Claudia Schaefer, Dr. Liya Ding, Dr. Jin Suk Park, Dr. Ashwathi (Abbee) Mohan, Dr. Orlando Argüello-Miranda, Dr. Christoph Burckhardt, Dr. Meghan Driscoll, Dr. Tadamoto Isogai, Dr. Marcel Mettlen, Dr. Benjamin Nanes, Dr. Jungsik Noh, Dr. Erik Welf, Dr. Dana "Kim" Reed, Dr. Qiongjing "Jenny" Zou, Dr. Andrew Jamieson, Dr. Reto Fiolka, Dr. Murat Can Cobanoglu, Liqiang Wang, Dr. Kushal Bhatt, Joseph Chi, Dr. Jaewon Huh, Dr. Hanieh Mazloom-Farsibaf, Dr. Gabriel Muhire Gihana, Dr. Vasanth S. Murali, Dr. Philippe Roudot, Dr. Dagan Segal, Dr.

Qiuyan Shao, Dr. XinXin Wang, Dr. Andrew Weems, Xuexia Jiang, Andres Nevarez, David Saucier, and Byron Weiss.

I would like to thank all Shay lab members and give special thank you to Dr. Woodring E. Wright, who dedicated his life to scientific research. I truly appreciate the working experience together with Dr. Jerome Robin, Dr. Wanil, Kim, Dr. Ejun "Elijah" Huang, Dr. Enzo Tedone, Ryan O'Hara, Dr. Andrew Ludow, Dr. Tsung-Po Lai, Crystal Cornelius, Summer Barron, Kimberly Batten, Dr. Brody Holohan, Dr. Yanhui Li, Dr. Sang Bum Kim, Dr. Lu Zhang, Dr. Mariam El-Ashmawy, Dr. Ronald Bozeman, Dr. Ilgen Mender, Dr. Laura Yuan, Melissa Coquelin, Sei Sho, Dr. Mohammed Sayed, Dr. Mazvita Maziveyi, Dr. Busola Alabi, and Dr. Silvia Siteni.

I would like to further express my thanks to the BioHPC, Bioinformatics Core Facility (BICF), Live Cell Imaging Core Facility (LICF), and the funding support from Cancer Prevention and Research Institute of Texas (CPRIT).

I would like to give thanks to faculty members and administrators in the Cancer Biology program and Computational & Systems Biology track. Thanks for the support from Dr. Rolf Brekken, Dr. Asaithamby Aroumougame, Priyarama Sen, Mardreana Reed, Dr. Khouloud Jaqaman, Dr. Kimberly Reynolds, Erica Garza, and Rebekah Craig.

Finally, to my parents, Yinfeng Wang and Liang Zhang. Thank you for the unconditional love and support throughout my life.

AN IMAGING APPROACH TO EXAMINE TELOMERE DYNAMICS AND REGULATION OF GENE EXPRESSION WITH AGING

Ning Zhang, Ph.D.

The University of Texas Southwestern Medical Center at Dallas, 2020

Supervising Professor: Gaudenz Danuser, Ph.D. and Jerry W. Shay, Ph.D.

Abstract

Telomeres are repetitive non-coding nucleotide sequences (TTAGGG)n capping the ends of chromosomes. Improved methods to measure the shortest (not just average) telomere lengths (TLs) are needed. Progressive telomere shortening with increasing age has been associated with shifts in gene expression through models such as the telomere position effect (TPE), which suggests reduced interference of the telomere with transcriptional activity of increasingly more distant genes. A modification of the TPE model, referred to as Telomere Position Effects over Long Distance (TPE-OLD), explains why some genes 1-10 MB from a

telomere are still affected by TPE, but genes closer to the telomere are not. Therefore, demonstrating the regulatory roles of telomere length shortening on genes with accurate TL measurement will improve our understanding to the 3D genomic DNA landscape including telomeres.

In this doctoral dissertation, I developed a user-friendly software for automatic electrophoresis gel quantification and contributed to developing the Telomere Shortest Length Assay (TeSLA), a technique that detects telomeres from all chromosome ends from <1 kb to 18 kb using small amounts of input DNA. Using cells with more TL information provided by TeSLA, I conducted an imaging approach to systematically examine the occurrence of TPE-OLD at the single cell level. Compared to existing methods, the pipeline allows rapid analysis of hundreds to thousands of cells, which is necessary to establish TPE-OLD as an acceptable mechanism of gene expression regulation.

I examined two human genes, for which TPE-OLD has been described before, ISG15 (Interferon Stimulated Gene 15) and TERT (TElomerase Reverse Transcriptase). For both genes I found less interaction with the telomere on the same chromosome in old cells compared to young cells. Experimentally elongated telomeres in old cells rescued the level of telomere interaction for both genes. However, the dependency of the interactions on the age progression from young to old cells varied. One model for the differences between ISG15 and TERT may relate to the markedly distinct interstitial telomeric sequence arrangement in the two genes. Overall, this provides a strong rationale for the role of telomere length shortening in the regulation of gene expression.

TABLE OF CONTENTS

PRIOR PUBLICATIONS

Zhang N, Li Y, Lai TP, Danuser G, Shay JW. Imaging assay to probe the role of telomere length shortening on telomere-gene interactions in single cells. Under review.

Li Y, Zhou G, Bruno IG, Zhang N, Sho S, Tedone E, Lai TP, Cooke JP, Shay JW. Transient introduction of human telomerase mRNA improves hallmarks of progeria cells. Aging Cell. 2019 Aug;18(4):e12979.

Lai TP, Zhang N, Noh J, Mender I, Tedone E, Huang E, Wright WE, Danuser G, Shay JW. A method for measuring the distribution of the shortest telomeres in cells and tissues. Nat Commun. 2017 Nov 7;8(1):1356.

Kim W, Ludlow AT, Min J, Robin JD, Stadler G, Mender I, Lai TP, Zhang N, Wright WE, Shay JW. Regulation of the Human Telomerase Gene TERT by Telomere Position Effect-Over Long Distances (TPE-OLD): Implications for Aging and Cancer. PLoS Biol. 2016 Dec;14(12):e2000016.

Holohan B, Kim W, Lai TP, Hoshiyama H, Zhang N, Alazami AM, Wright WE, Meyn MS, Alkuraya FS, Shay JW. Impaired telomere maintenance in Alazami syndrome patients with LARP7 deficiency. BMC Genomics. 2016 Oct 17;17(Suppl 9):749.

Liu X, Han D, Somel M, Jiang X, Hu H, Guijarro P, Zhang N, Mitchell A, Halene T, Ely JJ, Sherwood CC, Hof PR, Qiu Z, Pääbo S, Akbarian S, Khaitovich P. Disruption of an Evolutionarily Novel Synaptic Expression Pattern in Autism. PLoS Biol. 2016 Sep;14(9):e1002558.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

TL – Telomere Length

DNA – Deoxyribonucleic Acid

RAP1 – Repressor/Activator Protein 1

TRF1 – Telomere Repeat binding Factor 1

TRF2 – Telomere Repeat binding Factor 2

TIN2 – TRF1-Interacting Nuclear protein 2

TPP1 – TIN2-interacting Protein 1

POT1 – Protection of Telomeres 1

TPE – Telomere Position Effect

TPE-OLD – Telomere Position Effect Over Long Distance

qPCR – Quantitative Polymerase Chain Reaction

TRF – Terminal Restriction Fragment

FISH – Fluorescent in situ Hybridization

HT Q-FISH – High Throughput Quantitative Fluorescent in situ Hybridization

STELA – Single Telomere Length Analysis

TeSLA – Telomere Shortest Length Assay

ALT – Alternative Lengthening of Telomeres

ITS – Interstitial Telomeric Sequences

TRAP – Telomere Repeat Amplification Protocol

TERT – Telomerase Reverse Transcriptase

ISG15 – Interferon Stimulating Gene 15

CLIA – Clinical Laboratory Improvement Amendments

FACS – Fluorescence Activated Cell Sorting

PD – Population Doubling

ddPCR – digital droplet Polymerase Chain Reaction

ddTeSLA – digital droplet Telomere Shortest Length Assay

RL – Reinforcement Learning

CRISPR – Clustered Regularly Interspaced Short Palindromic Repeats

CHAPTER ONE

INTRODUCTION

1.1 Telomere and Telomerase

1.1.1 The discovery of telomeres

The concept of telomeres was first raised by Hermann Muller in 1938, given the observation that only chromosome pieces that retained their natural ends survived after X-ray induced chromosomal breaks (Muller 1938). Its name was derived from the Greek words “telos (τέλος)” for "end" and “meros (μέρος)” for "part". Later in 1941, Barbara McClintock charactered the “natural ends” of chromosomes from maize (corn) cells and pointed that those parts behaved differently than other broken chromosome ends induced by X-irradiation (X-ray) (McClintock 1941).

Before the 1960s, it was widely believed that normal cells can grow indefinitely under appropriate conditions in tissue culture (Shay and Wright 2000). The Nobel laureate Alexis Carrel claimed in 1921 that chick heart fibroblasts were continuously cultured for 34 years and the observations of finite cell replication elsewhere were due to the ignorance of cell culture conditions (Carrel and Ebeling 1921). However, in 1961, that central dogma was challenged and overturned by Leonard Hayflick and Paul Moorhead, who mixed young female fibroblasts with old male fibroblasts and then observed the extinction of old cells under the same culture conditions (Hayflick and Moorhead 1961). It was speculated that Carrel’s original observation

was due to the some experimental errors that fresh cells were added into the culture when chicken embryo tissue extract was fed (Hayflick 1965; Hayflick and Moorhead 1961). Later in 1974, the Australian Nobel laureate, Sir Macfarlane Burnett, used the term "Hayflick limit" to indicate the finite replication capacity of normal cells in his book "Intrinsic Mutagenesis" for the first time (Burnet 1974). Although there was a debate arguing that premature senescence of normal cells is largely due to the harsh cell culture conditions in most laboratories (Rubin 1998), the Hayflick limit concept has been generally accepted for most cell types.

With the discovery of the DNA molecular structure by James Watson and Francis Crick in 1953 (Watson and Crick 1953), researchers noticed that there was a problem of DNA replication at the end of the complementary double-stranded helix. Because DNA polymerase always synthesizes the new strand from 5' to 3' (Bessman et al. 1958), the leading strand can be continuously replicated, while the lagging strand is synthesized as a set of ligated short Okazaki fragments (Okazaki et al. 1968; Sugimoto et al. 1968), each requiring a primer binding to the template strand to initiate the replication. Therefore, the lagging strand cannot be synthesized all the way through the end of a linear chromosomal template, because there is no template sequence beyond the end point for a primer to bind and synthesize the very last Okazaki fragment. That end replication problem was first pointed out by James Watson in 1972 (Watson 1972), and further speculated to be linked to chromosome shortening along with cell division by Alexy Olovnikov (Olovnikov 1973). Those findings suggest that the leading strand is always possessing a 3' overhang at the end of a chromosome.

A few years later, in 1978, Elizabeth Blackburn characterized the first telomere sequence from Tetrahymena thermophila (Blackburn and Gall 1978). Later those telomeric sequences were proved to be protective on artificial minichromosomes introduced into yeast (Szostak and Blackburn 1982). In the following years, telomeric sequences from a large variety of organisms were reported. The repetitive sequence 5'-TTAGGG-3' was first identified in humans in 1988 (Moyzis et al. 1988), and then found to be conserved in all mammals (Gomes et al. 2011; Meyne et al. 1989). Together with the discovery of telomerase in 1985 (Greider and Blackburn 1985; Greider and Blackburn 1987), the 2006 Lasker Award and 2009 Nobel Prize were awarded to Elizabeth H. Blackburn, Carol W. Greider, and Jack W. Szostak, for their discovery of how chromosomes are protected by telomeres and for the discovery of the enzyme telomerase (Zakian 2009).

1.1.2 The molecular biology of telomeres

Telomeres cap the end of chromosomes and protect the chromosomes from fusion, degradation, and being recognized as double strand breaks (O'Sullivan and Karlseder 2010a). The conservative telomeric functions across species are implemented by a set of proteins associated with telomeric DNA sequences (Lewis and Wuttke 2012). All mammals including human share the same tandem repeats of TTAGGG telomeric sequence and associated proteins termed the shelterin protein complex (de Lange 2005). The shelterin complex consists of six subunits (Figure 1.1): Repressor/activator protein 1 (Rap1); Telomere repeat binding factor 1 (TRF1); Telomere repeat binding factor 2 (TRF2); TRF1-interacting nuclear protein 2 (TIN2);

TIN2-interacting protein 1 (TPP1) and Protection of telomeres 1 (POT1). TRF1 and TRF2 both bind to double stranded TTAGGG telomeric repeats on chromosome ends and can also bind to interstitial telomeric sequences (ITS) (Simonet et al. 2011). TIN2 interacts with TRF1 and TRF2 (Kim et al. 1999), while RAP1 interacts with TRF2 on binding to telomeric repeats. POT1 binds to the single stranded 3'-overhang and interacts with TIN2 through the binding with TPP1 (O'Connor et al. 2006). All those components and interactions protect the integrity and stability of telomeres (Sfeir and de Lange 2012). With the help of shelterin protein complex, the single-stranded 3'-overhang invades into the double-stranded sequence to form a t-loop (Griffith et al. 1999), which protects the chromosome ends from exposure to DNA damage repair responses.

1.1.3 Introduction of telomerase

Telomerase was discovered from T. thermophila (Greider and Blackburn 1985) by Carol Greider, a grad student in Elizabeth Blackburn's lab in 1985. It had enzymatic activity that could elongate telomeres and was first named "activity terminal transferase", later known as telomerase (Greider 1991; Greider and Blackburn 1987). Telomerase was then identified as a ribonucleoprotein complex, and its RNA component in T. thermophila was confirmed in 1989 (Greider and Blackburn 1989). In the next a few years, telomerase activity was reported from a variety of species, including humans (Morin 1989; Shippenlentz and Blackburn 1989; Yu et al. 1990; Zahler and Prescott 1988).

In 1994, the telomere repeat amplification protocol (TRAP), a PCR based assay was developed to measure telomerase activity (Kim et al. 1994). It confirmed telomerase activity in 85-90% human cancers (Shay and Bacchetti 1997), and raised the idea of inhibiting telomerase for cancer therapeutics (Herbert et al. 1999; Mender et al. 2015; Mender et al. 2018; Shay and Wright 2005; Shay and Wright 2006). In the following years, the human telomerase protein and RNA components (known as TERT and TERC) were cloned (Feng et al. 1995; Meyerson et al. 1997; Nakamura et al. 1997). Although TERC is ubiquitously expressed in all somatic cells, there is no TERT full length protein detected in most normal human cells (Feng et al. 1995; Nakamura et al. 1997), which is consistent with the TRAP assay results from cancer cells.

1.1.4 New functions of telomerase

Telomerase is well known for its function of extending telomeres since it was discovered in 1985 (Greider and Blackburn 1985; Shay and Wright 2019). Beyond that scope, after 2000, researchers have identified additional telomerase functions that are unrelated to telomere maintenance, but relevant to cell proliferation (Cong and Shay 2008; Hanahan and Weinberg 2011). Most of those noncanonical functions are associated with its protein subunit, known as telomerase reverse transcriptase (TERT). For example, there were studies showing telomerase's roles in cell apoptosis resistance, DNA damage repair, and RNA-dependent RNA polymerase function (Kang et al. 2004; Maida et al. 2009; Masutomi et al. 2005). TERT was also able to amplify the Wnt pathway signaling, by serving as a cofactor of the β-catenin/LEF

transcription factor complex (Park et al. 2009). Therefore, telomerase activity inhibition in cancer therapeutics needs to consider those new functions associated with telomerase and avoid side effects on normal cell proliferation. Additional telomerase functions and its regulation remain to be fully elucidated.

1.1.5 Telomere length shortening and effects

Telomerase activity only exists in human germline cells and stem cells (Shay 2016). Therefore, all dividing somatic cells show telomere length shortening after each cell division (Figure 1.2). The progressive telomere shortening has many effects related to aging and cancer. It can cause in vitro replicative senescence and may lead to chromosome instability and cancer development by forming bridge-fusion-breakage cycles (Fumagalli et al. 2012; Hemann et al. 2001; Herbig et al. 2004; von Zglinicki et al. 2005; Zou et al. 2004). 85-90% of cancer cells have reactivated their telomerase activity to maintain the stable short telomeres (Shay and Bacchetti 1997). Other cancer types developed mechanisms like alternative lengthening of telomeres (ALT) to make the cells survive (Bryan et al. 1997). With telomere length shortening, the interaction between telomeres and other genes on chromosomes may change, and therefore lead to the gene expression level variation (Kim et al. 2016; Kim and Shay 2018; Robin et al. 2014). To better understand the dynamics and regulatory roles of telomere shortening, accurate measurement of short telomeres and visualization of telomere-gene interaction are necessary for current research.

1.2 Telomere length (TL) measurement assays

Due to the end replication problem (Watson 1972), telomere lengths (TLs) become shorter along with cell division in all somatic cells, and even in proliferating cells with transient telomerase activity. It therefore links the TL measurement to aging and numerous human diseases associated with progressive telomere shortening. Considering the TL heterogeneity of different chromosome ends (Lansdorp et al. 1996), short telomeres are more likely to abandon their protective roles and lead to chromosome fusion and DNA damage response. Studies have shown that in both mice and humans, it is the shortest telomere, not the average TL, that plays an important role in cell viability and chromosome stability (Hemann et al. 2001; Zou et al. 2004). Different assays have been developed so far to quantitatively measure the TL distribution (Lai et al. 2018).

1.2.1 Quantitative Polymerase Chain Reaction (qPCR)

The qPCR assay (Cawthon 2002) is easy to establish for most labs with only ~50 ng starting DNA . This method measures the telomere signal (T) and a single-copy reference gene signal (S), then calculates the T/S ratio (Figure 1.3) to determine relative TLs (Cawthon 2002; Cawthon 2009). This assay has been widely used in many studies because it's easy to conduct and can be applied to high-throughput experiments. However, there are several limitations so that qPCR cannot be recognized as the gold standard of TL measurement.

First, qPCR measures the T/S ratio, which is probational to average TL, without giving the kilobase (kb) pairs information. The exact TL can only be acquired by comparing to a reference

cell line with known TL measured by other methods. Second, there is 10% or even higher variability of this assay (Aubert et al. 2012), which may lead to different results between independent laboratories. Third, qPCR only provides average TL information and is unable to accurately measure the shortest telomeres. Finally, it fails to measure cancer samples with aneuploidy (Holland and Cleveland 2009) where zero or multiple copies of the reference gene may exist. Therefore, the application of qPCR is limited to average TL measurement on normal diploid samples.

1.2.2 Terminal Restriction Fragment (TRF) analysis

The current "gold standard" of TL measurement is Terminal Restriction Fragment (TRF) analysis, which is a Southern blot based assay (Figure 1.4) to measure the average TL using a TTAGGGn-labelled probe (Kimura et al. 2010; Lansdorp et al. 1996; Mender and Shay 2015b). TRF digests genomic DNA with a combination of restriction enzymes but keeps the telomeric repeats intact because the TTAGGGn do not contain any specific sites that can be recognized by the enzymes. Gel electrophoresis followed by Southern blot shows a smear of telomeric repeats and indicates the range of TL due to the heterogeneity of telomere attrition (Bourgeron et al. 2015). The average TL can be determined by quantification of the smear size distribution.

Although TRF analysis has been widely accepted for TL measurement, it requires relatively larger amount of starting DNA (~3 ug) and longer time to perform this assay. Therefore, it's high-throughput application is limited. In addition, the choices of restriction enzymes lead to

different cutting sites in the subtelomeric region and may result in the smear size variation even for the same sample. Finally, TRF is less sensitive for short telomeres with the length of 2 kb or less (Lai et al. 2018).

1.2.3 Quantitative Fluorescent in situ Hybridization (Q-FISH)

There are several Q-FISH methods publicly or commercially available for TL measurement. Interphase Q-FISH uses peptide nucleic acid (PNA) probes CCCTAA3 to hybridize the telomeric repeats and quantify the fluorescence intensity to estimate TL (Canela et al. 2007). Metaphase Q-FISH applies the same probe, but requires proliferating cells to quantify TL for each individual chromosome end (Lansdorp et al. 1996). (SEEMS OUT OF PLACE) Flow FISH is CLIA (Clinical Laboratory Improvement Amendments) certified and commercially available for clinical diagnose to measure telomere fluorescence intensity in single interphase cells using fluorescence activated cell sorting analysis (FACS) (Rufer et al. 1998).

Due to the limit of fluorescence hybridization probes, Q-FISH cannot detect short telomeres with size below the PNA probe hybridization threshold. Secondly, those PNA probes may accidentally bind to interstitial telomeric sequences (ITSs), which are telomeric repeats located in the genomic DNA in vertebrates (Meyne et al. 1990). Finally, clusters of several telomeres may magnify the fluorescence Q-FISH signal intensity and lead to inaccurate TL quantification. One major contribution of the fluorescence Q-FISH technique is to demonstrate that in mice, it is not the average TL, but the shortest telomeres, that are critical for genome stability and cell viability (Blasco et al. 1997; Erdmann et al. 2004; Hemann et al. 2001).

1.2.4 Single Telomere Length Analysis (STELA)

STELA was developed to measure TL on individual chromosomes (Baird et al. 2003). It combines ligation, PCR, and Southern blot and emphasizes on shortest telomeres quantification. With allele specific primers, STELA can accurately measure TL of several chromosomes and reveal telomere allelic variation. However, not all chromosome ends have specific primers for STELA. Therefore, STELA is limited to only a subset of chromosomes ends.

To solve this problem, Universal STELA (U-STELA) was developed in 2010 to measure TL for all chromosome ends with universally designed primes (Bendix et al. 2010). But it is less sensitive to detect TLs over 8 kb, as previously described (Aubert et al. 2012).

1.2.5 Telomere Shortest Length Assay (TeSLA)

TeSLA is an improved version of U-STELA, which requires less than 1 ug of starting DNA and measures TL distribution from 18 kb to less than 1 kb with user-friendly Southern blot image quantification software (Lai et al. 2017b). Although limited by its low-throughput, TeSLA provides comprehensive TL distribution information for human studies and can accurately measure the dynamics of TL shortening along with cell divisions. Compared to Q-FISH, this assay will not be interfered by ITS regions inside chromosomes. It also provides much more details compared to TRF analysis (Lai et al. 2018). Table 1 compares all advantages and limitations of major existing TL measurement methods.

Table 1.1. Advantages and limitations of different telomere length measurement methods adapted from (Lai et al. 2018).

Method	Advantages	Limitations
qPCR	Easy to conduct; small amount of starting DNA required; many population-based studies for comparisons	Large variations among different laboratories, but reproducibility is better in commercial setting; not useful in cancer studies due to aneuploidy; only average TL is provided as a relative ratio
TRF	Common method for research studies; highly reproducible in some labs; many published studies for comparative research	Larger amounts of starting DNA required; provides most information on average TL; need to standardize restriction enzymes used to compare studies between laboratories (subtelomeric polymorphisms can alter data obtained); labor intensive
Interphase Q-FISH	Can be conducted on fixed tissues and cells	Labor intensive; TLs expressed as relative fluorescence units (not

Method	Advantages	Limitations
		actually TLs) but using standards measured by TRF actual TLs can be inferred
HT Q-FISH	Same as interphase Q-FISH; very reliable and reproducible results; CLIA certified	Does not distinguish telomere clustering in interphase cells; does not recognize telomere-free ends
Flow FISH	Same as interphase Q-FISH; can provide cell type specific information on mostly average telomere lengths; reproducible; CLIA certified	Requires an expensive FACS instrument; almost universally uses peripheral blood mononuclear cells (PBMCs)
Metaphase Q-FISH	Can potentially detect telomeres on all chromosomes	Does not detect the telomeres that are very short that do not hybridize with probes (appear as telomere-free ends); requires highly skilled cytogeneticist for chromosome-specific analyses

Method	Advantages	Limitations
STELA	Can detect the shortest telomeres on specific chromosomes	Works on only a small subset of individual human chromosomes; low throughput; labor intensive
Universal STELA	Measures mainly the shortest telomeres	Does not detect larger telomeres and can detect ITSs; manual quantitation; low throughput; labor intensive
TeSLA	Measures all the telomeres less than 1 kb and up to 18 kb on all chromosome ends; works on many animal types; automatic quantitation of telomere sizes using user-friendly software	Low throughput; labor intensive

1.3 Telomere position effect and genes

In organisms from yeast to humans, previous studies show that telomere length shortening can regulate gene expression level. Originally described in Drosophila and extended to yeast, it was named telomere position effect (TPE) (Gottschling et al. 1990; Sandell and Zakian 1992; Stavenhagen and Zakian 1998). A luciferase reporter inserted into Hela cells also showed

inhibited expression when inserted at telomeres but not when inserted into the genome at long distances away from the telomere (Baur et al. 2001). The first endogenous human gene (ISG15) regulated by telomere length was reported (Lou et al. 2009) and later, extended in human fibroblasts and myoblasts, including genes identified as being regulated by telomere length that were 5-10 Mb from the telomere but not genes closer to the telomere (Kim et al. 2016; Robin et al. 2014). This modification of TPE was termed telomere position effect over long distance (TPE-OLD) or telomere looping (Kim et al. 2016; Mukherjee et al. 2018; Wood et al. 2015; Wood et al. 2014).

Classic TPE is dependent on telomere length and is characterized by an "all or nothing" effect that is heritable and generally stable (Baur et al. 2001; Gottschling et al. 1990; Wright and Shay 1992a). Thus, when cells have long telomeres, genes near telomeres may be silenced due to chromatin effects near telomeres. As cells age (and telomeres progressively shorten), there may be de-repression of genes near telomeres eventually leading to reactivation of previously silenced genes. This could occur on all or only a subset of chromosome ends. The dependence of TPE on telomere length provides a mechanism for the modification of gene expression throughout the replicative lifespan of human cells and potentially in organisms.

As is illustrated in Figure 1.5, TPE-OLD or telomere looping is proposed to explain why some genes relatively far away from a telomere can be regulated by TPE but genes closer to the telomere are not. It thus becomes important to not only identify additional genes that may be regulated by TPE-OLD but also to develop higher throughput methods.

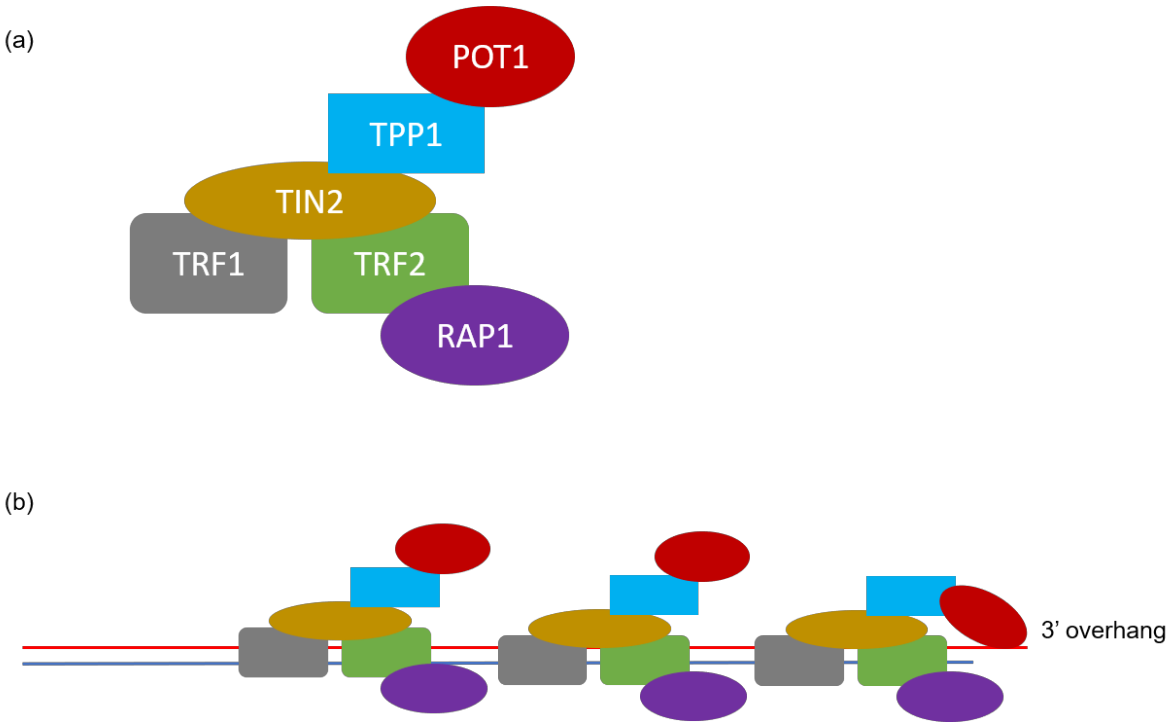


Figure 1.1 Shelterin Protein Complex. (a) The six subunits of shelterin and protein interactions. Repressor/activator protein 1 (Rap1); Telomere repeat binding factor 1 (TRF1); Telomere repeat binding factor 2 (TRF2); TRF1-interacting nuclear protein 2 (TIN2); TIN2-interacting protein 1 (TPP1) and Protection of telomeres 1 (POT1). (b) Schematic review of shelterin binding on telomeric DNA. Only POT1 can bind to the single-stranded 3' end. TRF1 and TRF2 bind directly to the double-stranded telomeric DNA. Other subunits interact indirectly with the telomeric DNA.

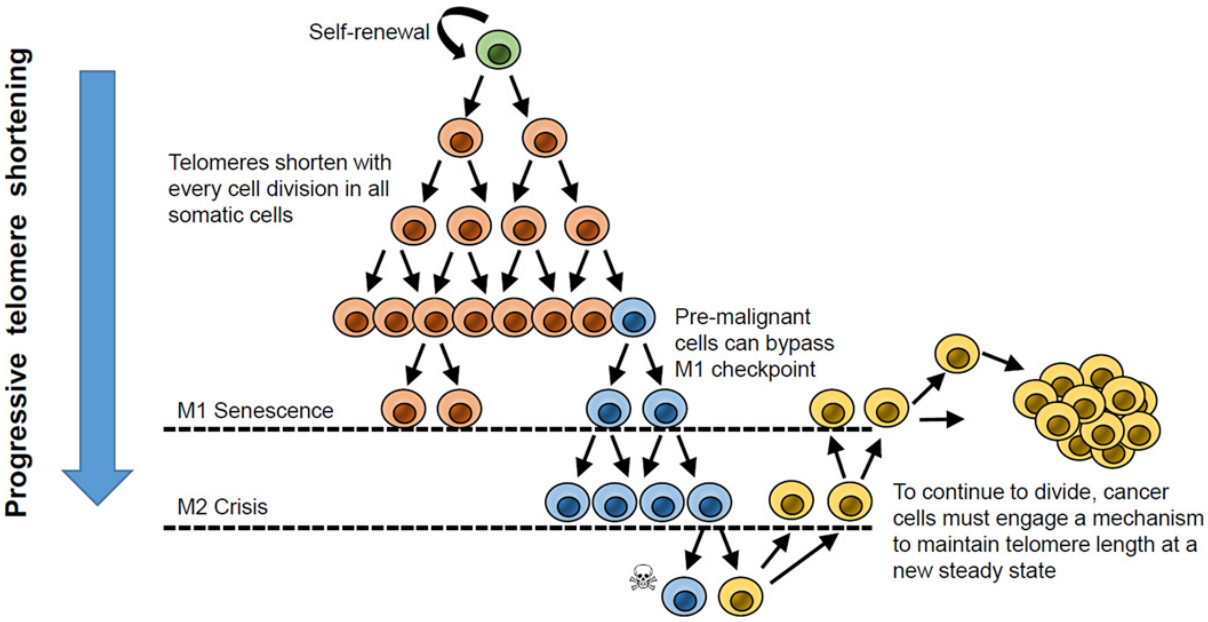


Figure 1.2 With increasing cell divisions, telomeres progressively shorten. Even stem cells that self-renew, there is a gradual shortening of telomeres. After a finite number of cell doublings, eventually the cells have sufficient short telomeres that they undergo a growth arrest called senescence or the Mortality Stage I (M1). This has also been termed the Hayflick limit. Premalignant cells that have obtained a number of oncogenic changes can bypass M1 and enter into an extended lifespan period. This has been termed the extended lifespan period but eventually these cells also slowdown in proliferation and enter a period called crisis. In crisis there is a balance between cell growth and apoptosis and the vast majority of the cell population dies. A rare cell can upregulate telomerase or the much rarer ALT pathway and continue to growth. The hallmark of cells escaping crisis is almost universally, stable but short telomere lengths and telomerase activity (Shay 2016).

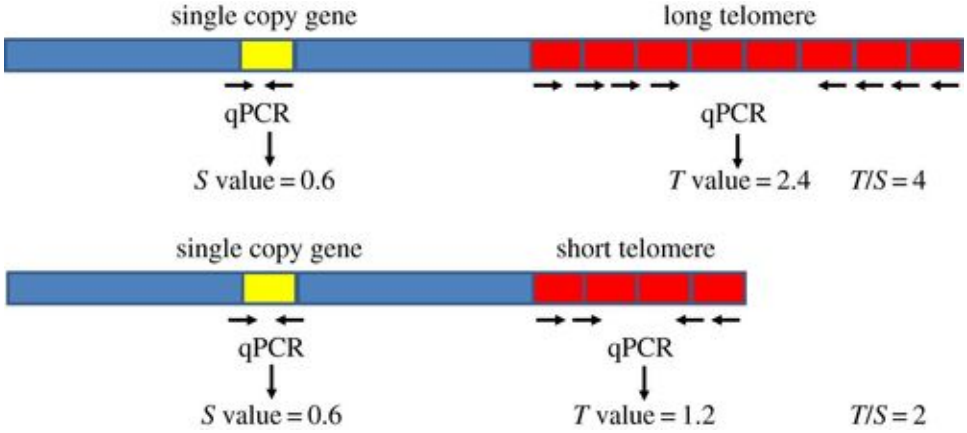


Figure 1.3 Q-PCR method is commonly used because it is a relatively easy procedure, does not require a lot of input DNA, and can be conducted with high-throughput PCR methods. The method provides a relative telomere length (*T*) compared to a single (*S*) copy gene and results are expressed as a *T*/*S* ratio as depicted (Lai et al. 2018).

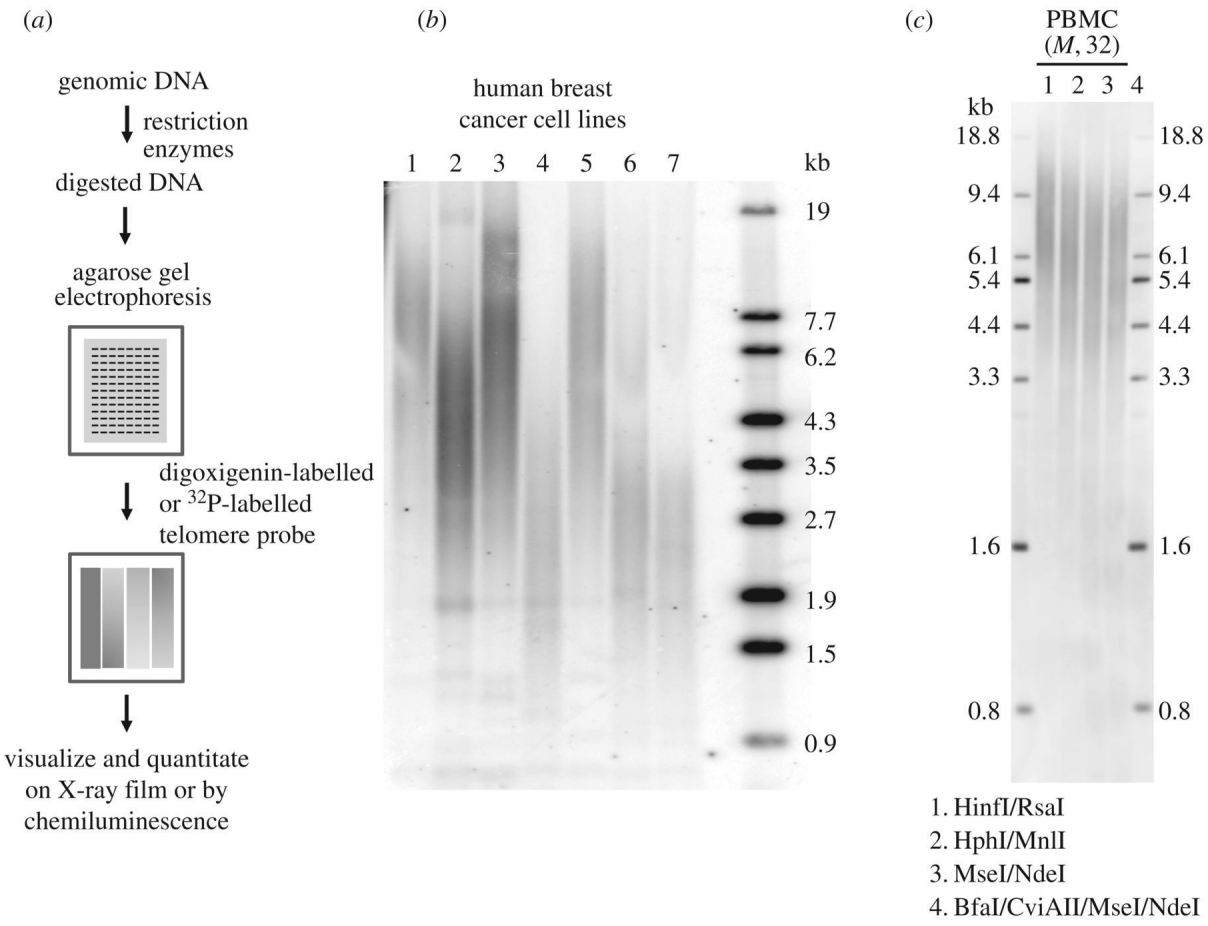


Figure 1.4 TRF is the 'gold standard' for TL measurements. (a) The methodology and (b) examples of a panel of human breast cancer cell lines of varying TLs. (c) Depending on the panel of restriction enzymes used to digest genomic DNA, the same DNA can vary greatly in TL. It is difficult with this Southern blot method to detect and quantitate the shortest telomeres (Lai et al. 2018).

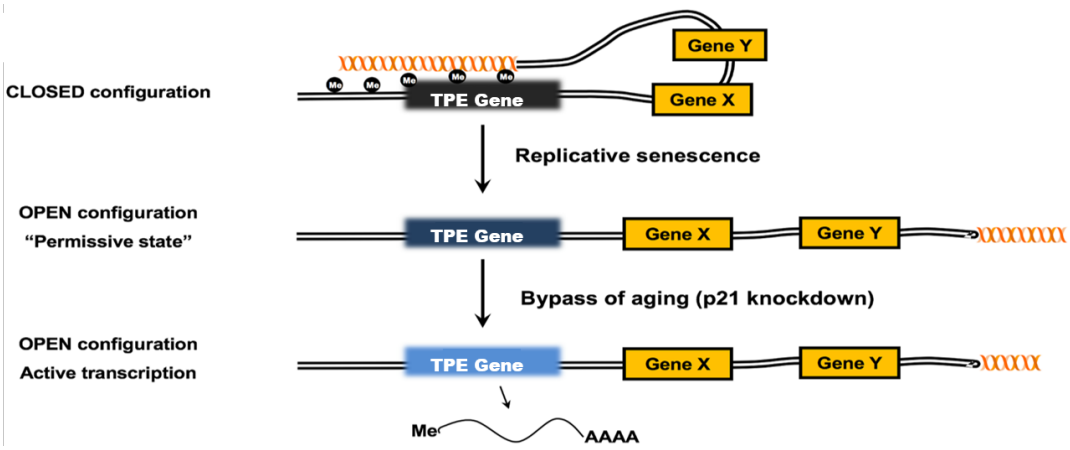


Figure 1.5 Simplified model of how TPE-OLD regulates hTERT expression in human cells during aging and cancer progression. Adapted from (Kim et al. 2016)

CHAPTER TWO

A METHOD FOR MEASURING THE DISTRIBUTION OF THE SHORTEST TELOMERES IN CELLS AND TISSUES

2.1 Introduction

In vertebrates, telomeres consist of conserved sequence repeats (TTAGGGn) with a single-strand 3′ G-rich overhang. Telomeres reside at the ends of chromosomes and combined with shelterin proteins, help maintain genomic stability (Palm and de Lange 2008). Telomerase is composed of two core components, telomerase RNA (TR or TERC) and telomerase reverse transcriptase (TERT) that add repetitive DNA to telomeres, but is not detected in most human adult somatic cells. Therefore, telomeres progressively shorten with each cell division due to the "end replication problem" (Olovnikov 1973; Watson 1972) In humans, telomere shortening has been implicated as a risk factor in numerous diseases such as cancer, diabetes mellitus, liver cirrhosis, and cardiovascular disease (Fitzpatrick et al. 2007; Sampson et al. 2006; Shay 2016; Wiemann et al. 2002). In addition, genetic diseases have been identified that have defects in the telomere maintenance machinery termed telomere spectrum disorders (or telomeropathies) (Holohan et al. 2014). Patients with these syndromes display accelerated telomere attrition and much shorter telomeres compared with age-matched healthy controls (Armanios and Blackburn 2012; Opresko and Shay 2017).

It is well established that it is the shortest telomeres, not average telomere length (TL), that are able to activate DNA damage responses and subsequently trigger a cell-cycle progression arrest, termed cellular senescence (Fumagalli et al. 2012; Hemann et al. 2001; Herbig et al. 2004; Zou et al. 2004). Senescence correlates with a variety of age-associated diseases and serves as a tumor suppressor mechanism in large long-lived species to protect genome integrity and prevent accumulation of oncogenic changes by limiting the maximum number of cell divisions (Campisi 2013). Furthermore, an increase in the percent of the shortest telomeres has been proposed to be a predictor of lifespan in mammals (Vera et al. 2012). Thus, quantitative information about the shortest telomeres may serve as a biomarker for telomere-associated aging disorders including cancer.

Various approaches have been developed for quantifying TL, and generally provide information on average TL (Montpetit et al. 2014; Nussey et al. 2014; Vera and Blasco 2012).Terminal restriction fragment (TRF) analysis estimates the intensity and size distribution of the "telomeric smear" by Southern blot analysis (Kimura et al. 2010). The TRF technique requires a large amount of starting genomic DNA, and due to the lower hybridization signal of the shortest telomeres, underestimates the abundance of the shortest telomeres. The quantitative PCR (qPCR) TL measurement (Cawthon 2002) has been widely used for high-throughput (HT) testing to overcome the amount of the DNA requirement for TRF analysis and measures ratios of telomere signals (T) to a single copy gene signal (S). However, qPCR only provides relative TL quantitation that is proportional to the average TL from a reference

sample (T/S ratio). The qPCR method is not suitable to quantify TL for cancer studies since most cancer cells are aneuploidy (Holland and Cleveland 2009).

TL can also be measured by quantitative fluorescence in situ hybridization (Q-FISH) methods. Although metaphase Q-FISH (Lansdorp et al. 1996) can detect TL from each chromosome, it requires dividing cells as well as a skilled cytogeneticist (Montpetit et al. 2014; Vera and Blasco 2012). Flow-FISH is adapted to estimate mean TL of interphase cells (typically human lymphocytes). While this approach is an improvement over qPCR, it requires expensive equipment and the probe not only binds to telomeric repeats but also interacts with non-specific components in the cytoplasm (Aubert et al. 2012; Wieser et al. 2006). HT Q-FISH is able to quantify each individual telomere signal in each nucleus, however, telomere clustering has been reported in both lower eukaryotes (Gasser et al. 2004) and mammalian cells (Ramirez and Surralles 2008). In addition, Q-FISH methods depend on probe hybridization kinetics and do not permit quantitation of the shortest telomeres.

Single telomere length analysis (STELA) (Baird et al. 2003) was designed to generate high-resolution TL measurements including the shortest telomeres on individual chromosomes. Using ligation and PCR-based methods combined with Southern blot analysis, STELA is able to provide information about the abundance of the shortest telomeres. The major limitation of STELA is that TLs can only be detected on a subset of individual chromosome ends. The Universal STELA (U-STELA) (Bendix et al. 2010) method was reported to solve this problem by detecting telomeres from each chromosome. However, U-STELA is not efficient in detecting TL over 8 kb (Bendix et al. 2010) affecting the detection and accuracy of TL

distribution. Computational HT analyses have been developed to evaluate TLs using whole-genome sequencing data (Castle et al. 2010; Ding et al. 2014; Nersisyan and Arakelyan 2015). The sequencing-based assays only provide average TL and do not correlate well with TRF analysis (Lee et al. 2017).

Here, we developed a method called Telomere Shortest Length Assay (TeSLA). This method allows more sensitive, efficient, and specific TL detection when directly compared to other methods for TL measurement. We used TeSLA in combination with image-processing software that automatically measures TL after Southern blot analysis. We are thus able to detect telomere dynamics in a range from <1 to ~18 kb in normal aging processes, cancer progression, and telomere-related disorders in humans. Also, we applied the TeSLA method to different mammals such as mice and bowhead whales to demonstrate the utility of this improved technology.

2.2 Methods

2.2.1 Cell culture

BJ, Calu 6, HeLa, HeLa LT, HT1080, and NIH 3T3 cells were maintained in Media-X with 10% cosmic calf serum (Hyclone) at 37 °C in 5% CO2. C106, CEM, Jurkat, and RAJI cells were cultured in RPMI with 10% of fetal bovine serum at 37 °C in 5% CO2. H2087 cells were cultured at 37 °C in 5% CO2 in Media-X containing 10% cosmic calf serum with or without 1 μM imetelstat. HBECs were cultured in bronchial epithelial growth medium (BEGM) (Lonza, Allendale, NJ) with an antibiotic solution (penicillin G-streptomycin–amphotericin B)

and incubated in low oxygen (2–3%) at 37 °C. Bowhead whale lung fibroblasts were maintained in F12/DMEM (1/1) with 15% of fetal bovine serum and incubated in low oxygen (2–3%) on 0.1% gelatin-coated flasks at 32 °C. All cells were tested regularly for mycoplasma contamination.

2.2.2 TRF analysis

Isolated DNA (3 μg of each sample) was digested with different combinations of REs (as described in the figures) and then separated on a 0.7% agarose gel at 2 V/cm for 16 h. The telomere signals were detected by Southern blot analysis as previously indicated (Lai et al. 2016). In brief, after DNA was transferred from gel to a positive-charged nylon membrane, the transferred DNA was fixed by UV crosslinking. The cross-linked membrane was then hybridized with the hypersensitive DIG-labeled telomere probe overnight at 42 °C. After hybridization, the membrane was washed with buffer 1 (2× SSC, 0.1% SDS) at room temperature for 15 min and then washed twice with buffer 2 (0.5× SSC, 0.1% SDS) at 55 °C for 15 min. Next, the membrane was washed with DIG wash buffer (1× maleic acid buffer with 0.3% Tween-20) for 5 min. Then the membrane was incubated with 1× DIG blocking solution for 30 min at room temperature. After blocking, the membrane was incubated with anti-DIG antibody (Roche) in 1× blocking solution (1 to 10,000 dilution) for 30 min at room temperature. The membrane was then washed with DIG buffer two times at room temperature for 15 min. After washing, telomere signals were detected by incubating with CDP-star for 5 min. Image quantification was performed using Image Quant software to measure the intensity value of

each telomere smear. The intensity value of each sample was then adjusted with the background graphed from a lane with no DNA sample (Mender and Shay 2015b).

2.2.3 Quantitative florescence in situ hybridization

Cells in interphase were fixed with freshly prepared methanol/acetic acid (3/1 vol/vol) and then were stored at −20 °C. Fixed cells were dropped onto methanol/acetic acid pre-wet slides and subsequently air-dried overnight. We used the Q-FISH protocol as previously described (Min et al. 2017b). In brief, slides with fixed cells were washed three times with PBS for 5 min and then with PBS containing 0.5% Triton X-100 for 10 min. Fixed cells were washed three times with PBS for 5 min and then dehydrated with 70, 90, and 100% ethanol for 5 min each. Cells were hybridized with FAM-TelC probe (panagene) in hybridization buffer (70% foramide, 0.6× SSC, 5% MgCl2, and 0.0025% blocking reagent (Roche)) at 80 °C for 7 min and then incubated at room temperature for 12 h. After incubation, slides were washed with wash buffer (70% foramide and 0.6× SSC) for 1 h and 2× SSC for 30 min. Cells were washed with PBS for 5 min and then dehydrated with 70, 90, and 100% ethanol for 5 min each. Slides were mounted with VECTASHIELD Antifade Mounting Medium with DAPI (Vector Labs). Imaging was acquired using a Personal DeltaVision wide-field fluorescent microscope (Mender and Shay 2015a). Image quantification was performed using TFL-Telo V2 image analysis software (Poon and Lansdorp 2001).

2.2.4 Genomic DNA extraction

Genomic DNA was extracted using the Gentra Puregen DNA Extraction Kit (Qiagen) according to the manufacturer's instructions. Each DNA sample was evaluated on a Nanodrop (Thermo Scientific) for concentration and purity, and integrity of DNA was determined by resolving 20 ng of DNA on a 1% agarose gel (Kimura et al. 2010).

2.2.5 XpYp STELA and Universal STELA

We used the XpYp STELA protocol as previously indicated (Sfeir et al. 2005) with some modifications. In brief, 10 ng of EcoRI digested DNA was ligated with a mixture of C-telorettes 1–6 (10−3 μM in 10 μl reaction at 35 °C for 12 h. About 250–500 pg of ligated DNA was used to perform multiple PCRs with 0.1 μM of XpYpE2 and Teltail primers (26 cycles of 95 °C for 15 s, 58 °C for 20 s, and 72 °C for 15 min). The PCR products were resolved on a 0.85 % agarose gel (1.5 V/cm for 19 h). The Southern blot analysis was performed as described for TRF analysis to detect amplified telomeres.

For the Universal STELA, we used the protocol as previously described (Bendix et al. 2010). In brief, 10 ng of MseI and NdeI digested DNA with 50 μM 42-mer and 11 + 2-mer oligonucleotides were incubated at 65 °C and then gradually cooled down to 16 °C for 1 h. The mixture was then incubated at 16 °C with T4 DNA ligase for 12 h in 10 μl reaction. The mixture was incubated with 10−3 μM of C-telorettes (as indicated for XpYp STELA) in 25 μl ligation reaction at 35 °C for 12 h. About 5–40 pg of ligated DNA was used to perform multiple PCRs with 0.1 μM of Adapter and Teltail primers (1 cycle at 68 °C for 5 min, 26 cycles at 95 °C for

15 s, 58 °C for 20 s, and 72 °C for 15 min). The PCR products were separated by a 0.85% agarose gel. Telomere signals were detected by Southern blot analysis as described for TRF analysis.

Oligonucleotides for XpYp STELA and Universal STELA:

C telorette1, 5′-TGCTCCGTGCATCTGGCATCCCCTAAC-3′;

C telorette 2, 5′-TGCTCCGTGCATCTGGCATCTAACCCT-3′;

C telorette 3, 5′-TGCTCCGTGCATCTGGCATCCCTAACC-3′;

C telorette 4, 5′-TGCTCCGTGCATCTGGCATCCTAACCC-3′;

C telorette 5, 5′-TGCTCCGTGCATCTGGCATCAACCCTA-3′;

C telorette 6, 5′-TGCTCCGTGCATCTGGCATCACCCTAA-3′;

XpYpE2, 5′-TTGTCTCAGGGTCCTAGTG-3′;

Teltail, 5′-TGCTCCGTGCATCTGGCATC-3′;

Adapter, 5′-TGTAGCGTGAAGACGACAGAA-3′;

11 + 2-mer, 5′-TACCCGCGTCCGC-3′;

42-mer, 5′-TGTAGCGTGAAGACGACAGAAAGGGCGTGGTGCGGACGCGGG-3′.

2.2.6 TeSLA for TL measurement

Before starting the TeSLA procedure, we make stocks of short double-stranded 5′ AT and TA overhang adapters for ligation at genomic and subtelomeric regions. To make 40 μM AT and TA adapters, 40 μl of 100 μM TeSLA adapter short oligonucleotide (ONT) was mixed with 40 μl of 100 μM of TeSLA adapter TA and TeSLA adapter AT ONTs individually to make the

final volume of 100 µl in 1× TSE buffer (10 mM Tris pH 8.0, 50 mM NaCl, 1 mM EDTA). The mixtures were incubated at 95 °C for 5 min and subsequently allowed to gradually cool down to room temperature.

To ligate TeSLA-T (TeSLA-T 1–6) to each telomere overhang, 1000 units of T4 DNA ligase (New England Biolabs), 1 mM ATP, and 10−3 µM of TeSLA-Ts were added to a final volume of 20 µl in 1× CutSmart buffer (New England Biolabs) with 50 ng of isolated genomic DNA (without RE digestion). The mixture was incubated at 35 °C for 12–16 h and then heat inactivated at 65 °C for 10 min. The inactivated mixture including two units of CviAII in 10 µl 1× CutSmart buffer was incubated at 25 °C for 2 h to generate genomic DNA fragments with 5′ AT overhangs. After CviAII digestion, two units of BfaI, NdeI, and MseI (New England Biolabs) in 10 µl 1× CutSmart buffer was added to the CviAII digested mixture to further digest genomic DNA as well as to generate DNA fragments with a 5′ TA overhang by incubating at 37 °C for 2 h. One unit of Shrimp Alkaline Phosphatase (rSAP; New England Biolabs) in 10 µl 1× CutSmart buffer was subsequently added to the digested mixture to remove 5′ phosphate from each DNA fragment to improve the specificity of ligation between overhang adapters and genomic DNA fragments (37 °C for 30–60 min) and subsequently heat inactivated at 80 °C for 20 min. After heat inactivation, the mixture was allowed to gradually cool down to 25 °C.

For adaptor ligation, 10 µl of the inactivated mixture was combined with 1000 units of T4 DNA ligase to a final volume of 20 µl in 1× CutSmart buffer with 1 mM ATP, 1 µM of AT adapter, and 1 µM of TA adapter. The mixture was incubated at 16 °C for 12–16 h. After adapter ligation, the mixture was heat inactivated at 65 °C for 10 min.

After adapter ligation, multiple PCR reactions were performed (initial melt at 94 °C for 2 min followed by 26 cycles at 94 °C for 15 s, 60 °C for 30 s, and 72 °C for 15 min) using 2.5 units of FailSafe Enzyme Mix (Epicenter) with 1× FailSafe buffer H in 25 µl reaction containing 0.25 µM primers (AP and TeSLA-TP) and 20–40 pg of ligated DNA. PCR products were resolved on a 0.85 % agarose gel (1.5 V/cm for 19 h). After gel electrophoresis, the Southern blot analysis was performed as described for TRF analysis to detect amplified telomeres.

2.2.7 Image quantification and statistical analysis

TeSLA is able to provide information for shortest TL <3 kb, which is often missed in other TL quantification methods. For this assay, we developed user friendly software that can automatically detect bands from sample images, annotate band size, calculate average TL, the percent of the shortest telomeres (1.6 kb and shorter), and obtain other relevant statistics. We have documented the validity of this assay in a spectrum of samples. The executable file, Matlab code, and detailed step-by-step manual of this software are attached in the Supplementary Item.

2.2.8 Bootstrapping method to assess estimation accuracy

We applied a bootstrapping method to 32 TeSLA PCRs of HBEC cells to evaluate the estimation accuracy that would be achieved by using n reactions ($1 \leq n \leq 32$) so that we could determine the number of reactions that give reasonably stable TL measurements. We computed the variation of mean TLs as well as the coverage rates to determine whether the detected TLs

are sufficiently dense within 0–10 kb. In the bootstrapping method, we randomly selected n reactions with replacement, computed the mean TL, and repeated this procedure 500 times for each number of reactions. We then computed SDs of the bootstrapped mean TLs, which showed the estimation accuracy as a function of the number of reactions (Figure 2.4c). To compute coverage rates, we considered bins with a size of 0.5 kb ranging from 0 to 1. Coverage rates at n reactions were defined to be the percentage of bins containing at least one reaction when randomly chosen n reactions are used. The coverage rate curve together with confidence bounds estimated from variation of bootstrapping repeats is shown in Figure 2.4d.

2.2.9 Permutation test for comparison of paired EDCs

We developed a new statistical test based on permutation for pairwise comparison of two sets of TL distribution curves to perform a systematic analysis of TL changes between two experimental conditions along the entire length spectrum or in a particular, yet a priori unknown sub-spectrum. This test was applied, e.g., to compare TeSLA measurements for PBMCs from 15 subjects measured at baseline and 1 year after (Figure 2.6c). Without loss of generality, we explain the testing procedure for this data set. We first represented TL measurements at baseline and 1 year after with EDCs for each subject (Supplementary Figure 2.5c, d), and computed 1-year differences in cumulative frequencies at every value of TLs (Supplementary Figure 2.5e, f). The difference curves were averaged over subjects. The resulting averaged difference curve (Figure 2.6c, d) visualizes 1-year changes in TL distributions.

To test how much the observed averaged difference curve deviates from zero, i.e., the null hypothesis that TL distributions do not change over 1 year, we randomly permuted for each subject independently the temporal order of the experimental condition, i.e., baseline and 1 year after. Accordingly, 50% of subjects, on average, retained the original order, whereas 50% of subjects had the order reversed. This procedure was repeated 105 times to simulate permuted averaged difference curve under the reference situation without difference between the conditions. Next, we derived 95% confidence envelope curves about the reference situation (Figure 2.6c, d). If the experimentally observed averaged difference curve goes outside the envelope curve at any TL value, then the null hypothesis is rejected with a p-value < 5%, i.e., the distributions did likely change between the two compared conditions.

To derive the one-sided 95% confidence envelope curve, we considered the EDCs measured at baseline and in 1-year period $(F_{0i}(x), F_{1i}(x)),$ for all subjects i = 1, 2, …, n, with x > 0 denoting a TL. One-year changes at the distribution level can be visualized by the observed averaged difference curve $\bar{D}(x) = \frac{1}{n}\sum_{i=1}^{n}\{F_{1i}(x) - F_{0i}(x)\}$.

Let $\bar{D}^{(k)}(x)$ denote the averaged difference curve from k-th permuted data set. These permuted difference curves displayed heterogeneous variance at different values of TLs, which is intrinsic to the data structure. Before quantifying deviation of the observed difference curve from the reference difference curve, the permuted difference curves are standardized to account for the heterogeneous variance as follows:

$$Z^{(k)}(x) = \frac{\bar{D}^{(k)}(x) - m_{\text{perm}}(x)}{\sigma_{\text{perm}}(x)},\ k = 1, 2, \ldots, K,$$

where

$$m_{\text{perm}}(x) = \frac{1}{K}\sum_{k=1}^{K}\bar{D}^{(k)}(x),\ \sigma_{\text{perm}}(x)$$

$$= \left\{\frac{1}{K-1}\sum_{k=1}^{K}\left(\bar{D}^{(k)}(x) - m_{\text{perm}}(x)\right)^2\right\}^{1/2},$$

and K is the total number of permutations. We then obtained permutation distances of standardized difference curves from the reference function in positive direction by

$$T_{\text{perm}}^{(k)} = \max_{x>0} Z^{(k)}(x).$$

The distance between the observed and the reference difference curves is computed by

$$T_{\text{obs}} = \max_{x>0}\left\{\left(\bar{D}(x) - m_{\text{perm}}(x)\right)/\sigma_{\text{perm}}(x)\right\}$$

using the same standardization. The significance of the observed distance can now be assessed by comparing it with the permutation distances and computing the p-value,

$$P_{\text{perm}} = \#\left\{T_{\text{perm}}^{(k)} > T_{\text{obs}}\right\}/K.$$

The one-sided 95% confidence envelope curve is obtained by $Q(0.95)\times\sigma_{\text{perm}}(x) + m_{\text{perm}}(x)$, where Q(0.95) is the 0.95th quantile of the permutation distances.

2.2.10 Quantitative PCR

The relative amount of DNA from mTERT +/− mouse liver tissue and three individual DNA preps from the same mTERT −/− mouse liver tissue was measured by qPCR (initial melt at 95 °C for 3 min followed by 35 cycles at 95 °C for 10 s, 60 °C for 1 min) using SsoFast EvaGreen Supermix (Bio-Rad) in a LightCycler 480 II (Roche Molecular Biochemicals). Relative DNA levels of each mouse DNA were calculated and normalized to the mean of amplification level of mouse B1 repeats from 10 ng of genomic DNA from NIH 3T3 cells. Primers for mouse B1 repeats are listed in Supplementary Table 2.1.

2.3 Results

2.3.1 The principle of TeSLA

A schematic presentation of the TeSLA method is shown in Figure 2.1. TeSLA is a ligation and PCR-based approach to detect amplified TRFs from all chromosomes. TeSLA significantly improves the specificity and the sensitivity for TL measurements when compared to other methods.

First, we compared TeSLA to U-STELA that uses restriction enzyme (RE)-digested genomic DNA to ligate the terminal adapters (telorettes) to the 5′ end of telomeric C-rich strands. However, the U-STELA strategy may result in ligation between subtelomeric sequences and digested genomic DNA fragments. To avoid this possibility and increase the specificity of terminal adapters to anneal and ligate to the 5′ end at each telomeric C-rich strand, the TeSLA method uses extracted DNA (no RE digestion) with a mixture of newly designed terminal

adapters (TeSLA-T 1-6; Supplementary Table 1). Each TeSLA-T contains seven nucleotides of telomeric C-rich repeats at the 3′ end, which is complementary to the G-rich overhang followed by a unique sequence derived from bacteriophage MS2 for PCR.

Second, to minimize the measurement of subtelomeric regions and to create specific ends for ligation of adapters at the proximal ends of telomeric repeats, TeSLA uses a combination of four REs (BfaI/CviAII/MseI/NdeI) to digest TeSLA-T-ligated genomic DNA. BfaI and MseI digest DNA at the telomere variant region that is adjacent to the canonical telomere repeats in subtelomeric regions. CviAII and NdeI increase the frequency of generating 5′ AT and 5′ TA overhangs at genomic and subtelomeric regions. We performed TRF analysis using genomic DNA from human BJ fibroblasts and different cancer cells (C106, CEM, HeLa, and RAJI) with REs for TeSLA or two additional different RE mixtures commonly used (AluI/HaeIII/HhaI/HinfI/MspI/RsaI, and HphI/MnlI) that significantly reduce the detection of subtelomeric regions (Kimura et al. 2010; Steinert et al. 2004). We observed that the RE mixture for TeSLA further reduces detection of subtelomeric regions when compared to the other combinations of REs (Supplementary Figure 2.1a, b). After RE digestion, we performed 5′ dephosphorylation to prevent non-specific ligation between the telomeric DNA fragments and the digested genomic DNA fragments, which could potentially add extra sequences to subtelomeric regions during the next step of TeSLA for adapter ligation.

Third, to increase the ligation efficiency and the specificity of PCR for telomeric DNA amplification, two double-stranded adapters (5′ AT and 5′ TA overhangs) to tag genomic and subtelomeric sequences were generated (Methods; Figure 2.1a; Supplementary Table 1).

Adapters contain phosphorylated 5′ AT or TA overhang and C3 spacers at each 3′ end to facilitate ligation between the 5′ end of adapters and the 3′ end of genomic/telomeric C-rich DNA fragments. Adapters also contain a unique 3′ overhang complementary to the AP primer for the subsequent PCRs.

Fourth, we used a long-ranged PCR enzyme (FailSafe PCR enzyme mix) that is reliable for telomeric DNA amplification (Hockemeyer et al. 2005; Sfeir et al. 2005) to perform multiple TeSLA PCRs. Because the C-rich telomeric DNA fragments are tagged with TeSLA-Ts and 5′ TA or AT overhang adapters on both ends, multiple copies of tagged telomeres can be amplified using the AP primer together with the TeSLA-TP primer that is identical to the 5′ tail of TeSLA-Ts at each cycle of PCR. Since genomic DNA fragments are ligated only at the 5′ end, the genomic DNA fragments only amplify at one copy per PCR cycle. Amplified DNA fragments are then separated using 0.85% agarose gel and subsequently transferred to positive-charged nylon membrane to perform Southern blot analysis using a hypersensitive telomere-specific probe (Lai et al. 2016).

To validate the specificity of TeSLA for TL measurement, we used DNA from BJ fibroblasts to perform TeSLA (Supplementary Figure 2.2a). In the absence of either RE digestion, adapters, or T4 DNA ligase, no telomere-specific products are detected. We determined the effect of DNA degradation on TeSLA for TL detection using human Jurkat leukemic cells. To obtain Jurkat cells with variations in the number of viable cells, we reduced nutrients and allowed the culture medium to become more acidic. We then evaluated the integrity of DNA in cells with different percent (95, 75, 45, and 35%) viable cells and then performed TeSLA.

We observed that cells with lower viability have more degraded DNA and a higher frequency of short telomeres (Supplementary Figure 2.2b, c). Thus, DNA integrity is essential to obtain reliable results using TeSLA for TL measurement.

2.3.2 TeSLA compared to STELA and U-STELA

STELA, U-STELA, and TeSLA are designed to analyze telomere dynamics especially the distribution of the shortest telomeres. Therefore, we compared the sensitivity and specificity of TeSLA to U-STELA and XpYp STELA. In addition to the canonical telomeric ends of chromosomes, short telomere repeats (between 2 and 25 repeats) called interstitial telomeric sequences (ITSs) are present in numerous intra-chromosomal locations (Ruiz-Herrera et al. 2008; Wood et al. 2014). To preferentially amplify tagged telomeres, the U-STELA uses a "panhandle" design of the proximal adapter to ligate both ends of genomic DNA fragments to suppress subsequent PCR amplification (Bendix et al. 2010). However, this suppression PCR strategy is designed for low-molecular weight (MW) products (Lavrentieva et al. 1999). Therefore, the U-STELA method may not completely suppress the amplification of ITSs.

We next compared the specificity of PCR amplification between U-STELA and TeSLA. In the absence of one or both primers, TeSLA does not have any detectable telomere products. However, there are several non-specific PCR products that can be observed after using only one of the primers to amplify the tagged DNA from U-STELA (Figure 2.1b). This indicates that the "panhandle" structure may not suppress non-telomeric DNA amplification such as

ITSs. In contrast, TeSLA is able to specifically detect only chromosomal telomeric repeats without including ITSs (Figure 2.1b).

We also performed PCRs on serial dilutions of ligated DNA by TeSLA and U-STELA to compare the sensitivity for telomere detection. Using the same amount of input DNA, we detected more telomere signals using TeSLA compared to U-STELA (Figure 2.1c). Thus, TeSLA is more sensitive and efficient compared to U-STELA for telomere detection.

In a normal human cell, each chromosome end has different TLs (Lansdorp et al. 1996). Thus, we directly compared the distribution of the telomere amplification products between TeSLA and XpYp STELA using the same BJ fibroblast DNA. The results show TeSLA is able to identify a wider distribution range of TL in comparison to XpYp STELA using considerably less input DNA (Figure 2.1d). When compared to XpYp STELA, TeSLA provides more precise information not only about the average TL but also about all the shortest telomeres, not just those of a specific chromosome end.

2.3.3 TeSLA is more sensitive compared to TRF and Q-FISH

To examine the utility of TeSLA in studying telomere dynamics, we evaluated TL in a telomerase positive non-small-cell lung cancer (NSCLC) cell line (H2087). These cells were subjected to long-term telomerase inhibition treatment with imetelstat and then released to observe the dynamics of telomere re-elongation. The results were compared using TeSLA, TRF, and telomere Q-FISH. Imetelstat is a lipid modified thio-phosphoramidate oligonucleotide that binds to the active site of telomerase RNA to inhibit telomerase activity

(Herbert et al. 2005). A previous study (Frink et al. 2016) demonstrated long-term imetelstat treatment shortened average TL in multiple NSCLC cell lines. In the present study, H2087 cells were continuously treated with 1 μM imetelstat (three times per week) for 18 weeks and then released from imetelstat treatment for 5 weeks. Cells were collected at 10 and 18 weeks with imetelstat treatment and 1–5 weeks after drug removal for TL measurement. Although we observed average TL shortening at 10 and 18 weeks treatment with TRF analyses, the relative intensity of shorter TL measurements, as expected, was significantly reduced (Figure 2.2a). We also examined TL using Q-FISH with 0 and 18 weeks of treatment, and 5 weeks in the absence of imetelstat (Figure 2.2b). We observed that only a few telomere signals with relatively low intensity were detected in H2087 cells with 18 weeks treatment compared to cells with 0 weeks treatment and 5 weeks after drug removal. This demonstrates that interphase Q-FISH is not sufficiently sensitive to measure TL for cells with extremely short telomeres. In contrast, TeSLA detects TL in both imetelstat-treated cells and cells released from imetelstat in a quantitative manner (Figure 2.2c). Using TeSLA, we observed not only the average TL changes but also the changes in the distribution of the shortest telomeres from all chromosomes.

2.3.4 Software for TeSLA quantification

To quantify the TeSLA images efficiently and accurately, we developed user friendly software based on MATLAB programming for automatic detection and size annotation of telomere bands (See Chapter 4 for software manual). The quantitation workflow is shown in Figure 2.3a.

With the pre-processed image (Figure 2.3b), the lane profile is generated by summing up the normalized pixel intensity values along each vertical line from left to right (Figure 2.3c). The software then detects the center of each lane with watershed segmentation of the lane profile to determine significant peaks. Next, the software estimates the average lane width based on peak-to-peak intervals, and crops the regions of each individual lane for band detection (Figure 2.3c). With each individual lane, the software generates band profiles by summing up the pixel intensity values horizontally from top to bottom, followed by segmentation of the significant peaks marking the centers of telomere bands (Figure 2.3d). Band intensity is recorded by averaging the pixel intensity values among the local region of bands. The software displays all detected bands and gives users the ability to manually adjust results.

The software then fits the user-defined ladder size standards at each region of the image. The band size is annotated and recorded by comparing each band to its pixel position and intensity value. Telomeres with similar sizes that cannot be separated by electrophoresis results in more intense bands. The software is able to identify overlapping bands by comparing the intensity of each band to the intensity of neighboring bands. Significantly brighter bands are attributed with a double or triple count and indicated with different colors in the final output (Figure 2.3e). The TL distribution is also plotted for each sample (Figure 2.3f). With the annotated band sizes, the software rapidly calculates average TL, the percent of the shortest telomeres (with user-defined threshold), and other relevant statistics.

2.3.5 TeSLA is not biased for detecting short TLs

Since TeSLA measures TL using PCR to amplify tagged telomeres, the resolution of TL is limited by the efficiency of PCR amplification for GC-rich telomeric DNA. To examine the upper size limit of TeSLA for TL measurement, we compared TeSLA and TRF on HeLa LT cells, which are telomerase positive and have long telomeres (Min et al. 2017a). Using the same panel of REs for TeSLA to perform TRF analysis, we observed that the majority of telomeres distribute at higher MW range (>18.8 kb) with a relatively small amount of telomeres at the lower MW range (<18.8 kb) (Supplementary Figure 2.3a). After using ligated HeLa LT DNA to perform 16 TeSLA PCRs, we observed that TeSLA detects TL reliably up to 18 kb (Supplementary Figure 2.3b, c). Although we did not use other PCR enzymes to perform TeSLA, we determined that TeSLA's upper limit of telomere detection is ~18 kb, which covers TL detection for the vast majority of human normal and cancer cells.

We next tested whether TeSLA PCR amplification is biased toward amplifying the shortest telomeres. We used DNA from normal human bronchial epithelial cells (HBECs) at early passage (age 24, female), which have relatively long telomeres and a NSCLC cell line, Calu 6, which has very short telomeres (Frink et al. 2016) to perform TeSLA (Figure 2.4a). After ligations, the same amount of ligated DNA from HBEC and Calu 6 was mixed (HBEC:Calu 6 = 1:1) and used for PCR. We observed that the mean TL from the mixed DNA TeSLA (3.56 kb) was very close to the average of HBEC (4.97 kb) in combination with Calu 6 (2.04 kb) results (Figure 2.4a). We further visualized the distribution of TL from HBEC, Calu 6, and the mixed DNA using a kernel density estimation (Figure 2.4b). The results showed that

the distribution of TL from mixed DNA is similar to the distribution of TL from the reference mixture (HBEC:Calu 6 = 1:1) indicating that TeSLA PCR does not have a bias for over-amplifying the shortest telomeres.

2.3.6 TeSLA for TL measurements

To achieve representation of all the telomeres in a population of cells, we analyzed the MW of each detected telomere from 32 TeSLA PCRs of HBECs at the same time. Since each reaction detects around 10–15 different MWs (Supplementary Figure 2.4a), we determined the number of PCRs that are necessary to obtain a reliable estimate of the entire distribution of TLs in HBECs. We used bootstrapping to draw from the 32 PCRs, many repeats of simulated TL distributions with $1 \leq n \leq 32$ PCRs. We then computed the distribution of mean TLs and corresponding standard deviations (SDs) of the mean TLs to estimate the expected accuracy achievable when "n" PCRs were employed to measure the TL (Figure 2.4c). When using eight PCRs, the SD was 0.26 kb, about one-third of the SD when using one reaction. Since the gain in estimated accuracy is small with 24 additional reactions, using eight PCRs is reasonable to estimate the mean TL. We also computed coverage rates as a function of the number of reactions used. We observed that eight reactions are sufficient to cover 87% of the bins with a size of 0.5 kb ranging from 0 to 10 kb (Figure 2.4d). Based on these analyses, we used eight PCRs as a unit to analyze the distributions of TL. Next, we divided the 32 PCRs into four data sets and visualized TL distributions from each data set by empirical distribution curves (EDCs)

(Figure 2.4e) to calculate an intra-assay coefficient of variation (%). For the mean TL, this value is 4.3%.

To determine the inter-variation of TeSLA, we used extracted DNA from human peripheral blood mononuclear cells (PBMCs) provided by two healthy male donor volunteers (age 32 and 72 years old) to perform three independent TeSLAs on different days. We found that PBMCs from the 32-year-old donor have considerably fewer critically short telomeres and longer average TL compared to the 72-year-old donor's PBMCs (Figure 2.4f). Then, we visualized the distribution of TL from each triplicate TeSLA result using EDCs (Figure 2.4g). The inter-variation between the triplicates was small and the inter-assay coefficient of variations (%) of mean TL was 1.6% for DNA from the 32-year-old donor and 3.9% for DNA from the 72-year-old donor.

To further evaluate whether eight PCRs are sufficient to measure telomere shortening in vitro, we used extracted DNA from two different BJ population doublings (PDs) to perform TeSLA and then quantify TL from both samples (Supplementary Figure 2.4b). The results showed that the telomere shortening rate is ~70 bp for each cell division, consistent with previous studies using chromosome-specific STELA and TRF analysis (Baird et al. 2003; Harley et al. 1990).

2.3.7 TeSLA in cancer and telomere-related disease progression

Short telomeres correlate with genetic alterations in cancer initiation (DePinho and Polyak 2004). To investigate the relationship between cancer progression and telomere dynamics, we performed TeSLA to measure TLs of normal colon epithelium, adenomas (villous and tubular

polyps), and colorectal cancer tissues from a patient (Druliner et al. 2016). We observed using TeSLA that DNA extracted from adenomas and cancer tissues have shorter mean TLs and more of the shortest telomeres below 1.6 kb compared to DNA isolated from normal colon tissue (Figure 2.5a).

TRF, Q-FISH, and qPCR methods are widely used to assess TL in telomere spectrum disorders such as idiopathic pulmonary fibrosis (IPF) (Armanios et al. 2007; Cronkhite et al. 2008; de Leon et al. 2010; Tsakiri et al. 2007). We used TeSLA to examine leukocyte TLs from eight individuals from a kindred with familial pulmonary fibrosis (four were affected and four were unaffected). We found that the affected family members have more short telomeres and shorter mean TL compared to unaffected family members (Figure 2.5b–e).

While the TeSLA results of colon cancer progression and IPF development are consistent with previous studies measuring TL by other methods (Cronkhite et al. 2008; de Leon et al. 2010; Druliner et al. 2016; Stuart et al. 2015), TeSLA provides unprecedented detail of the distribution of telomeres and supports the notion that critically short telomeres may provide information about the stage when short telomeres contribute to cancer initiation and onset of telomere spectrum disorders.

2.3.8 TeSLA monitors telomere dynamics in normal human aging

Previous reports indicate that life stresses, infectious diseases, and inflammatory diseases can cause acute telomere shortening (Epel et al. 2004; Pawelec et al. 2005; Vallejo et al. 2004). We used DNA of PBMCs from healthy subjects who served as placebo-treated volunteers in a

clinical trial over a 1-year period (Salvador et al. 2016) to measure TLs by TeSLA and TRF analysis (Supplementary Figure 2.5a, b). After comparing changes of TL from each subject over a 1-year period (collected at baseline and 1 year), telomere dynamic changes (15 subjects; 8 females and 7 males, age 51–69) were detected by TeSLA (Figure 2.6a) but not by TRF analysis (Figure 2.6b).

To further understand the differences between TeSLA and TRF for TL measurements, we used EDCs to represent TL distributions measured by TeSLA and TRF analysis (Supplementary Figure 2.5c, d) and then statistically tested changes of TL distributions of each subject in 1 year by comparing differences of cumulative frequency of EDCs from each pair at baseline and after 1 year (Supplementary Figure 2.5e, f). Figure 2.6c, d show the differences of cumulative frequencies averaged (15 pairs) from TeSLA and TRF analysis. These analyses demonstrate that the effect of TL shortening is at relatively short telomeres (<6 kb) in both TeSLA and TRF analysis. Using a permutation-based estimate of the 95% confidence envelope curve, we found that TeSLA not only showed TLs decreased over a 1-year period (one-sided p-value 0.0275) but also that the effect of telomere attrition was the most significant in the shortest telomeres (~1 kb) (Figure 2.6c). In the TRF analysis (Figure 2.6d), although the differences in cumulative frequencies still indicate the effect of TL shortening at short telomeres, the changes in TL distributions during 1 year was not significant (one-sided p-value 0.364), consistent with the results of mean, median, and 20th percentile TLs (Figure 2.6b).

Next, we directly compared the TL distributions measured by TeSLA and TRF analysis with all 30 DNA samples (15 pairs) at baseline and 1 year. The mean TLs by TeSLA (average

3.97 kb) were consistently 10% shorter than the ones by TRF (average 4.40 kb) (Figure 2.6e, paired t-test p-value < 0.0001). However, the median TLs by TeSLA and TRF analysis were not statistically different (Figure 2.6f, p-value 0.634). To compare and contrast differences of TeSLA and TRF measurements at the distribution level, we visualized the EDCs from TeSLA and TRF analysis by integrating all 30 TL measurements. The results showed large discrepancies at the shortest telomeres, while the differences of the two distribution curves were moderate at other TLs (Figure 2.6g). The averaged difference curve in cumulative frequencies (TeSLA-TRF) together with 95% confidence envelope curves revealed marked differences between TL distributions by TeSLA and TRF analysis in the 0.6–2.8 kb range (Figure 2.6h). However, the difference curve stayed within 95% confidence limits in the range of 3.0–7.8 kb indicating that TeSLA and TRF lead to consistent cumulative frequencies in the middle region and further explains why the median TLs by TeSLA and TRF were not significantly different (Figure 2.6f). Thus, TeSLA is able to measure changes of human PBMC TL over 1 year and uncovers distributions of the shortest telomeres that have not been fully addressed by TRF analysis.

2.3.9 TeSLA for TL measurements in other organisms

The links between human diseases and TL are studied using tissue samples and cultured cells from humans as well as laboratory mice. Previous studies using Q-FISH analysis demonstrated that progressive telomere shortening causes loss of tissue function in mTERT-deficient mice (Erdmann et al. 2004; Meznikova et al. 2009). Most laboratory mice have many telomerase

positive tissues with average TLs up to 40 kb (Lejnine et al. 1995), which is above TeSLA's upper size limit (~18 kb). However, using TeSLA we tested if we could detect the proportion of the shortest telomeres in mice with TERT deficiency. We extracted DNA from liver tissues of the mTERT +/− and the 4th generation mTERT −/− mice in a C57BL/6 background. We could detect increasing amounts of telomeres under 18 kb in the 4th generation mTERT −/− mouse compared to mTERT +/− mouse using the same amount of input DNA (Supplementary Figure 2.6a, b) to perform TeSLA (Figure 2.7a). In the heterozygous mouse, most of the telomeres are not detected by TeSLA. However, in the mTERT −/− mouse (three individual genomic DNA preparations from the same mouse to perform three individual TeSLAs) more short telomeres were detected (Figure 2.7a).

DNA sequencing of the longest-living mammal, the bowhead whale, revealed that duplication and loss of genes related to DNA damage responses and repair may be involved in longevity and cancer resistance (Keane et al. 2015). The mean TL of cultured lung fibroblast cells from the bowhead whale is <10 kb by TRF analysis (Gomes et al. 2011). To examine the telomere dynamics of the bowhead whale in vitro by TeSLA, we used DNA isolated from cultured lung fibroblast cells at different PDs. Even though the DNA from these cells was not degraded, we found that subsets of very short telomeres (<1.6 kb), which have not been reported by TRF or Q-FISH analysis are detected in both early (PD 23) and late passage (PD 81) whale cells. We also observed telomere shortening and increasing amounts of the shortest telomeres at late passage (Figure 2.7b).

2.4 Discussion

We developed TeSLA as a method to determine TLs from all chromosomes and to monitor the changes and distribution of the shortest telomeres, the average TL, as well as TLs up to 18 kb. By directly comparing TeSLA to U-STELA and XpYp STELA for TL measurements (Figure 2.1), we demonstrated that TeSLA is more sensitive and specific for TL detection and generates more information of the spectrum of telomere distributions. We also compared TeSLA to both TRF and Q-FISH methods for detecting extremely short telomeres after treating a NSCLC cell line with a telomerase inhibitor (imetelstat) and then after imetelstat was removed (Figure 2.2). With TeSLA, but not TRF analysis or Q-FISH, we were able to quantitate the length of the shortest telomeres. In addition, we developed software to automatically quantify TeSLA results (Figure 2.3).

We demonstrated that TeSLA is able to measure telomeres up to 18 kb reliably (Supplementary Figure 2.3). Although TeSLA is sufficient to detect TL in almost all human normal somatic and cancer cells, TeSLA might not be suitable for cells, which have long and heterogeneous TLs such as alternative lengthening of telomeres (ALT) cells (Henson et al. 2002). In addition, we demonstrated that TeSLA is not biased for amplifying the shortest telomeres and is a highly reproducible method for TL measurements (Figure 2.4). Together, these analyses document that TeSLA measures the TL distribution in cells with a higher degree of confidence compared to existing TL measurement approaches.

To illustrate the changes of TL in human diseases, we examined telomere dynamics in a colon cancer progression series and a familial kindred with IPF (Figure 2.5). With TeSLA, we

observed not only short mean TL but also increasing amounts of very short telomeres that correlate with cancer and IPF progression. Going forward, TeSLA may serve as a tool to detect telomere shortening of clinical disease onset.

With TeSLA, but not TRF analysis, we detected changes in TL of PBMCs from healthy subjects over a 1-year period (Figure 2.6). Others have demonstrated the complexity of TLs measured in PBMCs by identifying variations in TLs and the rates of TL changes that are cell type-specific in vivo (Lin et al. 2016; Lin et al. 2015). TeSLA is very sensitive for detecting the shortest telomeres in a heterogeneous telomere background. Thus, TeSLA is capable of measuring sub-populations of cells in PBMCs, such as CD28- T cells, that have a lower capacity for cell division, shorter telomeres, and higher TL shortening rate compared to other sub-types of cells in PBMCs (Weng et al. 2009). Thus, TeSLA may be able to identify critically short telomeres in specific subsets of immune cells that are important in human aging. Using newly designed statistical analysis tools to compare TL distribution over a 1-year period, we were able to determine the most dramatic effect on telomere shortening is on some of the shortest telomeres (around 1 kb in length) from a group of healthy human volunteers. This suggests that TeSLA could be used to detect pathological thresholds of disease at an earlier stage than previously possible. Early diagnoses may result in the implementation of more effective interventions.

Besides studying changes of TLs in humans, TeSLA can be applied to evaluate the shortest telomeres in other animals (Figure 2.7). We demonstrated that TeSLA detects the distribution of the shortest telomeres in mTERT knockout mice. A recent study reported that telomere

shortening is a critical factor for age-dependent cardiac disease in the NOTCH1 haploinsufficiency mouse model (Theodoris et al. 2017). Thus, while TeSLA may not be able to detect the longest telomeres in mice, it can serve as a powerful tool to study the relationship between changes of the shortest telomeres and age-dependent diseases in mouse models with deficiencies in telomere maintenance.

TeSLA provides resolutions of all the telomeres including the shortest and requires very small amounts of starting DNA without limitation of specific cell types or tissue samples. TeSLA, therefore, will be useful for studying the changes of the shortest telomeres in disease development (pathological thresholds). Finally, TeSLA can provide information about telomere dynamics in human cells as well as other animals.

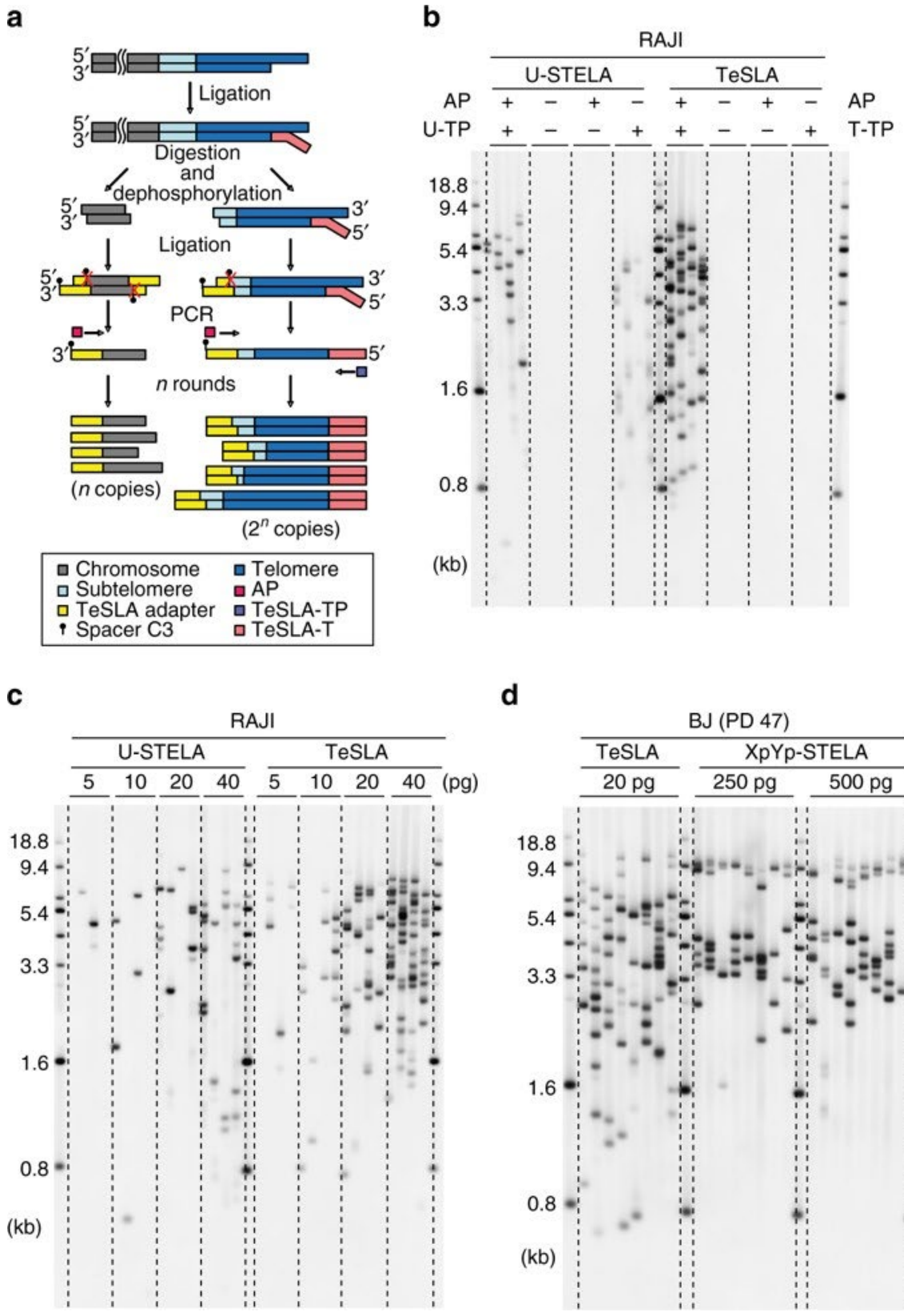

a
Ligation
Digestion and dephosphorylation
Ligation
PCR
n rounds
(n copies)
(2^n copies)
Chromosome
Subtelomere
TeSLA adapter
Spacer C3
Telomere
AP
TeSLA-TP
TeSLA-T
b
RAJI
U-STELA
TeSLA
AP
U-TP
T-TP
18.8
9.4
5.4
3.3
1.6
0.8
(kb)
c
RAJI
U-STELA
TeSLA
5 10 20 40 5 10 20 40 (pg)
18.8
9.4
5.4
3.3
1.6
0.8
(kb)
d
BJ (PD 47)
TeSLA
XpYp-STELA
20 pg
250 pg
500 pg
18.8
9.4
5.4
3.3
1.6
0.8
(kb)

Figure 2.1 Overview of Telomere Shortest Length Assay (TeSLA) and comparison to Universal STELA (U-STELA) and XpYp STELA. **(a)** Schematic of overall TeSLA methods. Extracted genomic DNA is ligated with TeSLA-Ts (each TeSLA-T contains seven nucleotides of telomeric C-rich repeats at the 3′ end) at the overhangs of telomeres and then digested with a restriction enzyme panel. Digested DNA is subsequently ligated with doubled-stranded TeSLA adapters at the proximal end of telomeres and genomic DNA fragments. After adapter ligation, PCR is performed to amplify ligated telomeric DNA. **(b)** About 40 pg of DNA from RAJI cells was used in each U-STELA and TeSLA reaction to test specificity of primers for telomere amplification and was tested as indicated (AP, adapter primer; U-TP, U-STELA teltail primer; T-TP, TeSLA-TP). **(c)** The sensitivity of U-STELA and TeSLA was compared by serial dilution of DNA from RAJI cells from 5 to 40 pg. **(d)** Using TeSLA (20 pg DNA for each reaction) and XpYp STELA (250 and 500 pg of DNA for each reaction) to detect TL in BJ cells.

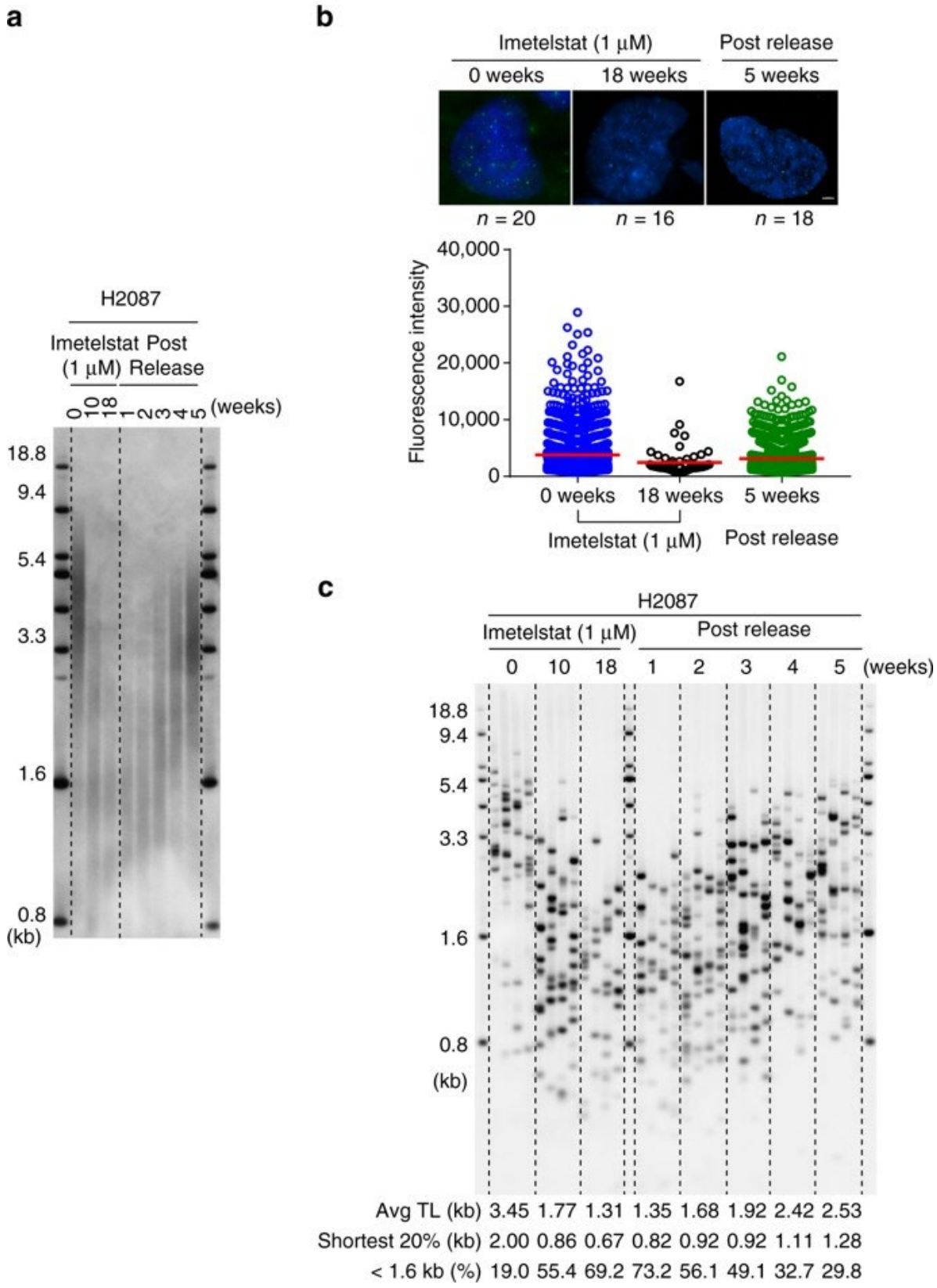

a
H2087
Imetelstat (1 μM)
Post Release
0 10 18 1 2 3 4 5 (weeks)
18.8
9.4
5.4
3.3
1.6
0.8
(kb)
b
Imetelstat (1 μM)
Post release
0 weeks
18 weeks
5 weeks
n = 20
n = 16
n = 18
Fluorescence intensity
40,000
30,000
20,000
10,000
0
0 weeks
18 weeks
5 weeks
Imetelstat (1 μM)
Post release
c
H2087
Imetelstat (1 μM)
Post release
0 10 18 1 2 3 4 5 (weeks)
18.8
9.4
5.4
3.3
1.6
0.8
(kb)
Avg TL (kb) 3.45 1.77 1.31 1.35 1.68 1.92 2.42 2.53
Shortest 20% (kb) 2.00 0.86 0.67 0.82 0.92 0.92 1.11 1.28
< 1.6 kb (%) 19.0 55.4 69.2 73.2 56.1 49.1 32.7 29.8

Figure 2.2 Measuring TL in long-term telomerase inhibition with imetelstat (1 μM) treatment and after the removal of drug in H2087 cells. **(a)** Isolated DNA from H2087 cells with 1 μM imetelstat treatment for 0, 10, and 18 weeks, and post released from 18 weeks for 1, 2, 3, 4, and 5 weeks was digested with the same REs for TeSLA (*Bfa*I/*Cvi*AII/*Mse*I/*Nde*I) and then separated on 0.7% agarose gel for TRF analysis. **(b)** Interphase Q-FISH; cells from H2087 with 0 and 18 weeks 1 μM imetelstat treatment, and 5 weeks after drug removal were used to measure TL by Q-FISH. The results were quantified using TFL-Telo software (*n*: numbers of nuclei were quantified for each time point as indicated above). Scale bar, 3 μm. **(c)** Results of TeSLA using DNA as indicated. Four TeSLA PCRs (30 pg of each reaction) were performed for each DNA sample.

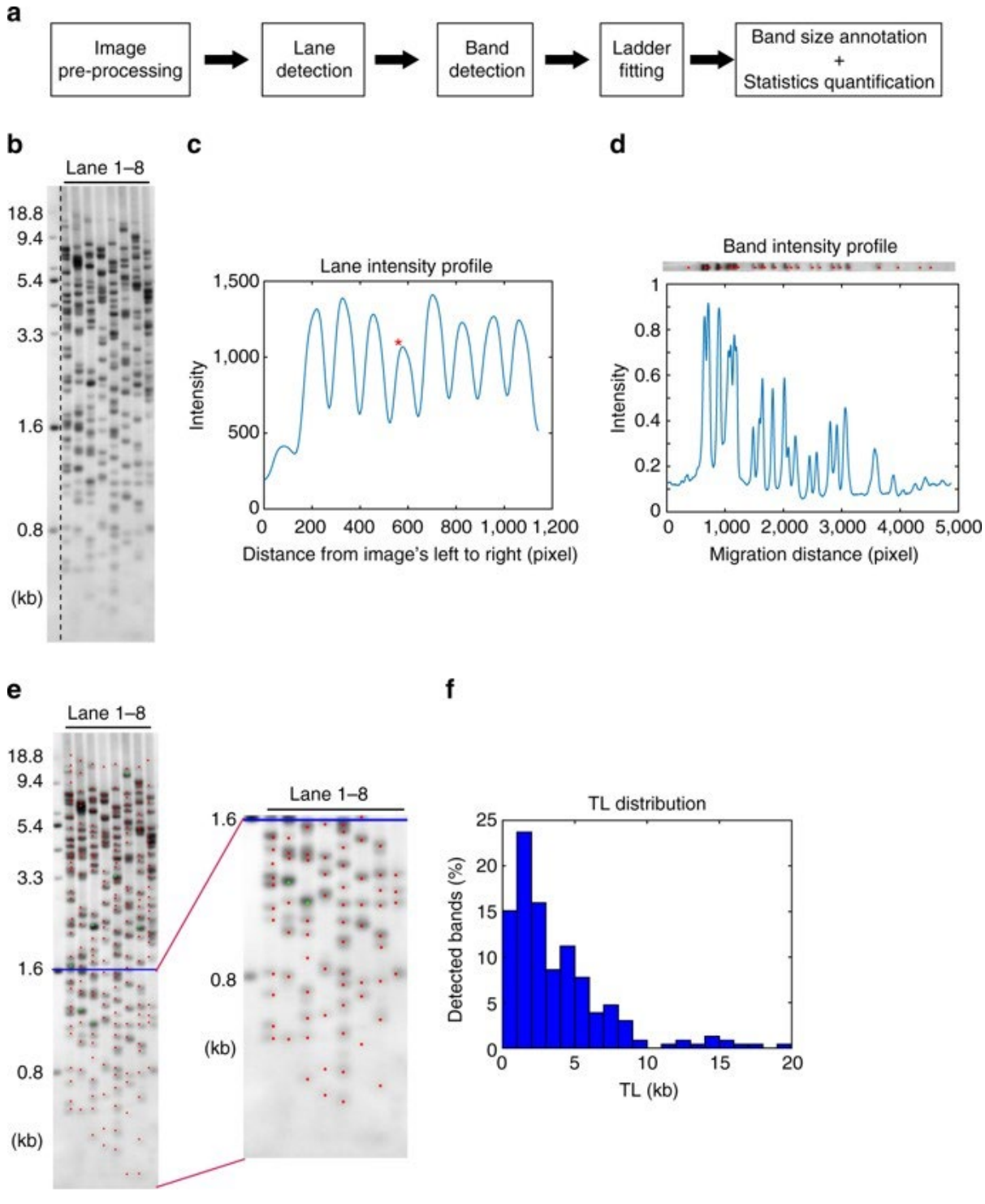

a
Image pre-processing
Lane detection
Band detection
Ladder fitting
Band size annotation + Statistics quantification
b
Lane 1–8
18.8
9.4
5.4
3.3
1.6
0.8
(kb)
c
Lane intensity profile
Intensity
1,500
1,000
500
0
0 200 400 600 800 1,000 1,200
Distance from image's left to right (pixel)
d
Band intensity profile
Intensity
1
0.8
0.6
0.4
0.2
0
0 1,000 2,000 3,000 4,000 5,000
Migration distance (pixel)
e
Lane 1–8
18.8
9.4
5.4
3.3
1.6
0.8
(kb)
Lane 1–8
1.6
0.8
(kb)
f
TL distribution
Detected bands (%)
25
20
15
10
5
0
0 5 10 15 20
TL (kb)

Figure 2.3 Overview of TeSLA image quantification software. **(a)** The computational analysis pipeline automatically detects each telomere band location, then annotates the band sizes, and calculates the relevant statistics (e.g., average TL, the ratio of shortest TL, TL at 20th percentile, and telomeres below 1.6 kb). **(b)** Example of input TeSLA image, tiff format recommended. On the left side of the image is the ladder lane with a standard size of 0.8–18.8 kb. **(c)** Lane profile, generated by summarizing the pixel intensity values vertically from left to right. Each peak indicates one lane detected by the software. **(d)** Band profile of lane 4, marked with the asterisk in **c**. The band profile is generated for each lane by horizontally summarizing pixel intensity values. Each significant peak refers to an individual band. **(e)** Example of the final output with the zoom-in of shortest telomere bands. Red dots indicate individual bands. Green dots mark the overlapping bands that are counted twice or three times. The blue line crosses the marker of 1.6 kb, which is the default threshold of shortest TL that other methods cannot reach. The software can calculate the ratio of TL below any given threshold. **(f)** Histogram of TL distribution, which covers the range of 0–20 kb.

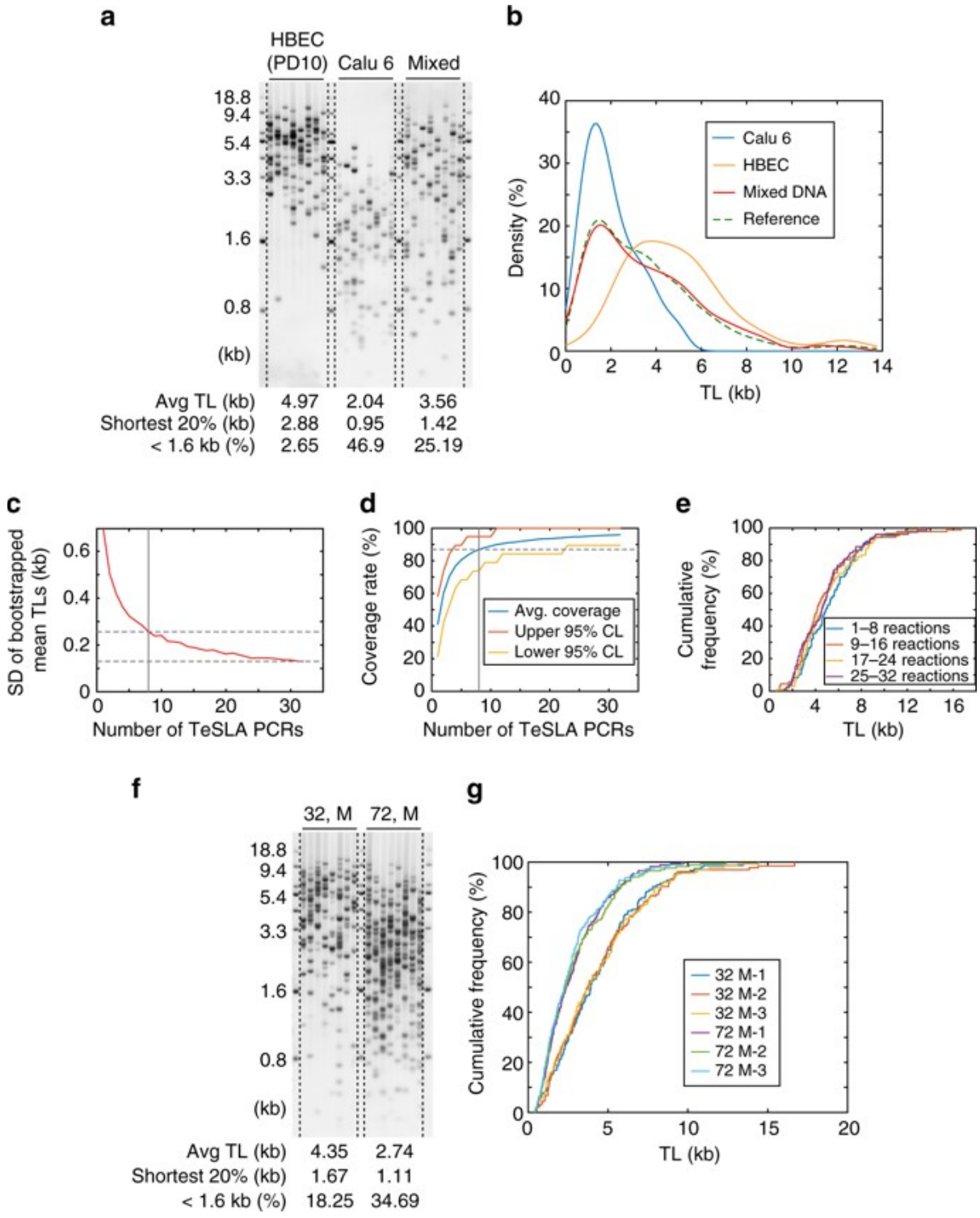

a
HBEC (PD10)
Calu 6
Mixed
18.8
9.4
5.4
3.3
1.6
0.8
(kb)
Avg TL (kb) 4.97 2.04 3.56
Shortest 20% (kb) 2.88 0.95 1.42
< 1.6 kb (%) 2.65 46.9 25.19
b
Density (%)
TL (kb)
Calu 6
HBEC
Mixed DNA
Reference
c
SD of bootstrapped mean TLs (kb)
Number of TeSLA PCRs
d
Coverage rate (%)
Number of TeSLA PCRs
Avg. coverage
Upper 95% CL
Lower 95% CL
e
Cumulative frequency (%)
TL (kb)
1–8 reactions
9–16 reactions
17–24 reactions
25–32 reactions
f
32, M
72, M
18.8
9.4
5.4
3.3
1.6
0.8
(kb)
Avg TL (kb) 4.35 2.74
Shortest 20% (kb) 1.67 1.11
< 1.6 kb (%) 18.25 34.69
g
Cumulative frequency (%)
TL (kb)
32 M-1
32 M-2
32 M-3
72 M-1
72 M-2
72 M-3

Figure 2.4 Assessment of different variations of TeSLA. **(a)** TeSLA of human bronchial epithelial cells (HBECs), Calu 6 (lung cancer cell line), and mixed DNA (HBEC: Calu 6 = 1:1) using 30 pg of DNA for each TeSLA PCR. **(b)** The kernel density estimation of TL from TeSLA results of Calu 6 (blue), HBEC (orange), and mixed DNA (red). The reference line (green dashed) represents a theoretical density function of TeSLA results when HBEC: Calu 6 = 1:1. **(c)** Standard deviations of bootstrapped distributions of mean TLs computed from 32 TeSLA PCRs of HBECs (Supplementary Figure 2.4a) show the estimation accuracy achieved by using n ($1 \leq n \leq 32$) lanes to estimate the mean TL. **(d)** Coverage rates estimated from 32 TeSLA PCRs of HBECs (Supplementary Figure 2.4a) based on bootstrapping. When eight PCRs were randomly selected from the 32 PCRs, 87% of all telomeres were detected with bin sizes 0.5 kb ranging from 0 to 10 kb. Red (yellow) line indicates upper (lower) 95% confidence bounds of the coverage rates. **(e)** Empirical distribution curves of quadruplets (eight TeSLA PCRs of each) from TeSLA results for HBECs show no significant changes in each eight TeSLA PCRs. **(f)** Representative TeSLA results of PBMCs from a young (age 32) and an old (age 72) individual. **(g)** Empirical distribution curves of TLs from the TeSLA of the triplicate results (blue, red, and yellow lines) for the 32-year-old male and triplicate results for the 72-year-old male (purple, green, and sky blue lines).

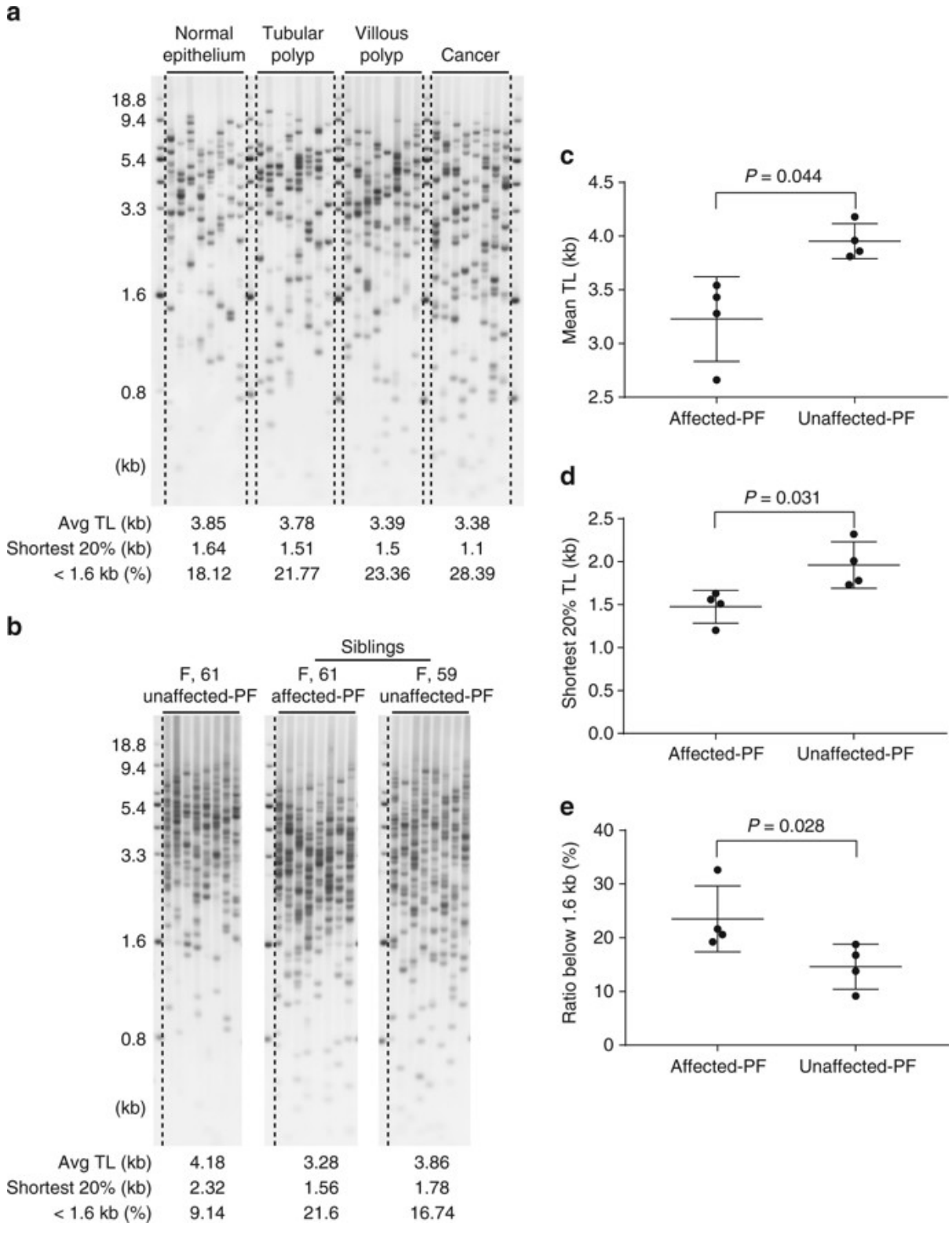

a
Normal epithelium
Tubular polyp
Villous polyp
Cancer
18.8
9.4
5.4
3.3
1.6
0.8
(kb)
Avg TL (kb) 3.85 3.78 3.39 3.38
Shortest 20% (kb) 1.64 1.51 1.5 1.1
< 1.6 kb (%) 18.12 21.77 23.36 28.39
b
Siblings
F, 61 unaffected-PF
F, 61 affected-PF
F, 59 unaffected-PF
18.8
9.4
5.4
3.3
1.6
0.8
(kb)
Avg TL (kb) 4.18 3.28 3.86
Shortest 20% (kb) 2.32 1.56 1.78
< 1.6 kb (%) 9.14 21.6 16.74
c
P = 0.044
Mean TL (kb)
4.5
4.0
3.5
3.0
2.5
Affected-PF
Unaffected-PF
d
P = 0.031
Shortest 20% TL (kb)
2.5
2.0
1.5
1.0
0.5
0.0
Affected-PF
Unaffected-PF
e
P = 0.028
Ratio below 1.6 kb (%)
40
30
20
10
0
Affected-PF
Unaffected-PF

Figure 2.5 Using TeSLA to determine TL and distribution of telomeres in colon cancer progression and idiopathic pulmonary fibrosis (IPF) siblings compared to age-matched normal controls. **(a)** TeSLA results of normal colorectal epithelium, adenomas (tubular polyp and villous polyp), and colon cancer tissues from one colon cancer patient show shorter mean TL and increasing amount of the shortest telomeres in adenomas and cancer tissues compared to normal colorectal epithelium. **(b)** Using TeSLA to determine TLs of DNA isolated from circulated leukocytes of the unrelated normal control, siblings with and without IPF. The age and gender are indicated above each TeSLA results. **(c**–**e)** Scatter plots of mean TL of TeSLA (**c**), the shortest 20% of telomeres (**d**), and percent of the shortest TL (<1.6 kb) (e) are shown for family members that have no IPF (three unrelated controls and a family member without IPF) and four family members with IPF. (**c**–**e** mean and s.e.m., $n = 4$)

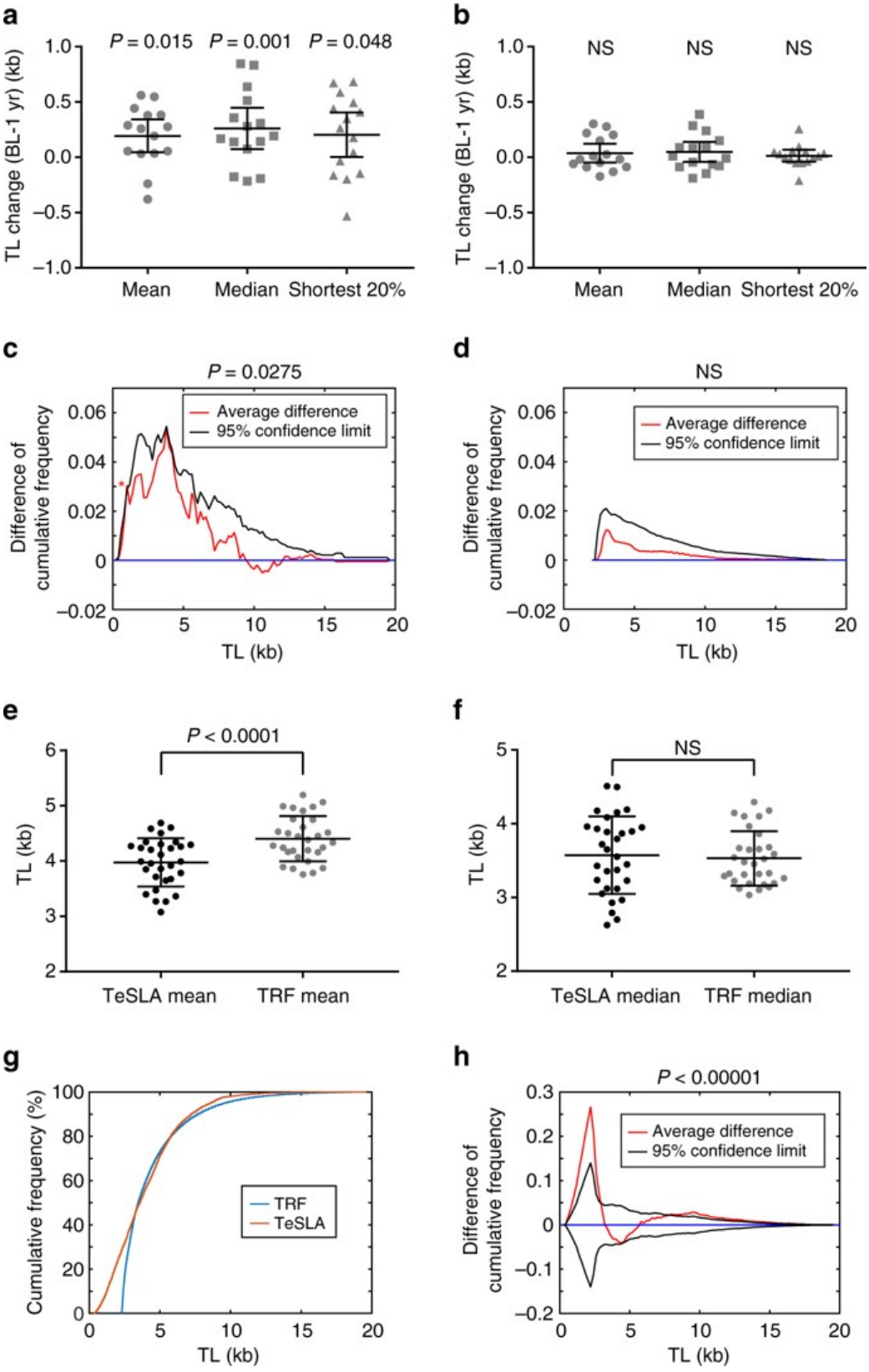

a
P = 0.015
P = 0.001
P = 0.048
TL change (BL-1 yr) (kb)
1.0
0.5
0.0
−0.5
−1.0
Mean
Median
Shortest 20%
b
NS
NS
NS
TL change (BL-1 yr) (kb)
Mean
Median
Shortest 20%
c
P = 0.0275
Average difference
95% confidence limit
Difference of cumulative frequency
0.06
0.04
0.02
0
−0.02
0
5
10
15
20
TL (kb)
d
NS
Average difference
95% confidence limit
Difference of cumulative frequency
TL (kb)
e
P < 0.0001
TL (kb)
6
5
4
3
2
TeSLA mean
TRF mean
f
NS
TL (kb)
5
4
3
2
TeSLA median
TRF median
g
Cumulative frequency (%)
100
80
60
40
20
0
TRF
TeSLA
TL (kb)
h
P < 0.00001
Average difference
95% confidence limit
Difference of cumulative frequency
0.3
0.2
0.1
0
−0.1
−0.2
TL (kb)

Figure 2.6 TeSLA is sensitive enough to detect changes of TLs in a 1-year period of normal human aging. **(a), (b)** Scatter dot blots comparing TLs in PBMCs measured at baseline and in 1-year period by TeSLA (**a**) and TRF analysis (**b**). The mean, the median, and the shortest 20% TLs of 15 normal healthy subjects (age from 51 to 69) were averaged. *P*-values from paired *t*-tests are shown as indicated above. BL, baseline; 1 yr, one year after; NS, not significant. **(c)**, (**d)** The average changes of TL distributions in PBMCs in a 1-year period of 15 subjects measured by TeSLA (**c**) and TRF analysis (**d**). One-year differences in cumulative frequencies from each subject were computed (see Supplementary Figure 2.5 e, f as examples). The average of 1-year changes in TL distributions of 15 subjects is shown in red, and one-sided 95% confidence limit (black) is derived from permutation. The asterisk represents the value (~1 kb of TL) that lies outside the 95% confidence limit, which indicates the most significant effect on telomere shortening. **(e)**, (**f)** Scatter plots comparing TeSLA and TRF analysis for mean (**e**) and median (**f**) TL measurements ($n = 30$) in PBMCs. **(g)** Comparison of TeSLA and TRF analysis of empirical distribution curves of pooled TLs from all 30 DNA samples. **(h)** The averaged differences (red) in cumulative frequencies (TeSLA-TRF) by the same method used in **(c)** and **(d)** show large difference between TeSLA and TRF in the short TL analyses (0.6–2.8 kb). Black lines are 95% confidence limits obtained from permutation. (**(a)**, **(b)**; mean and s.e.m., $n = 15$) (**(e)**, **(f)**; mean and s.e.m., $n = 30$)

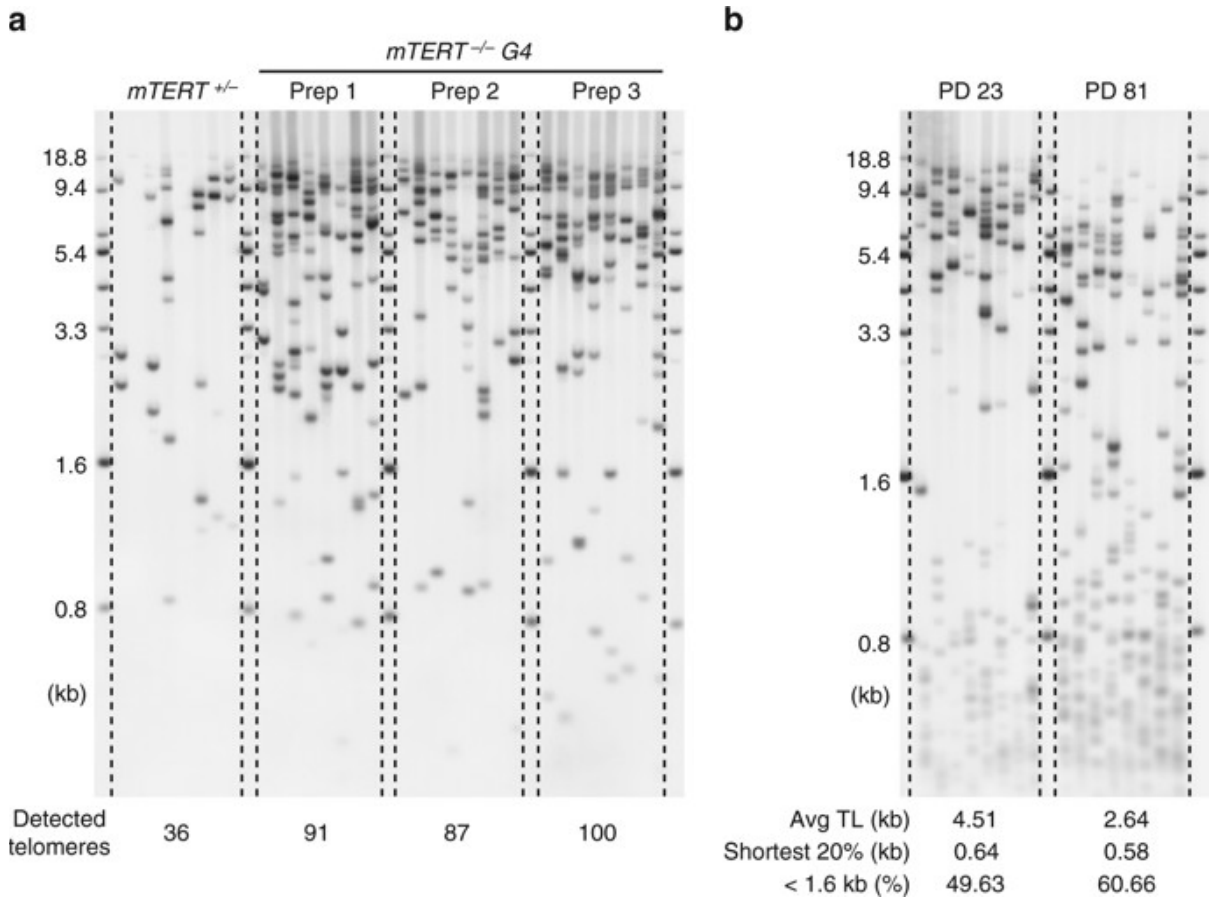
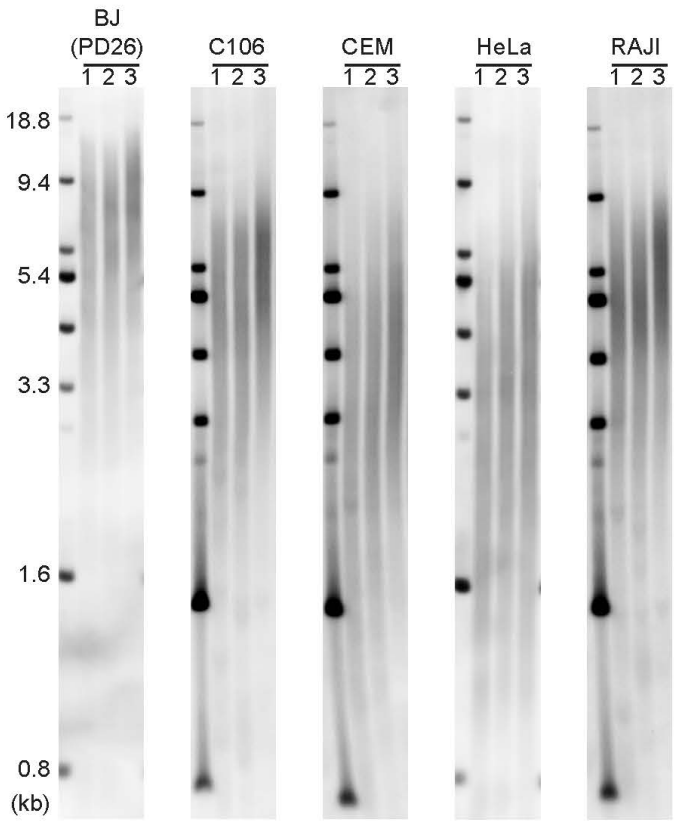
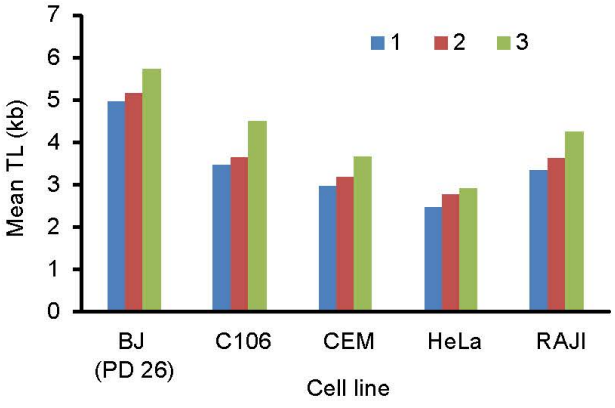


Figure 2.7 TeSLA for telomere detections in *mTERT* knockout mice and lung fibroblasts from a bowhead whale. **(a)** DNA extracted from *mTERT* $^{+/-}$ and *mTERT* $^{-/-}$ (4th generation, G4) mouse liver tissues were used to perform TeSLA (30 pg for each TeSLA reaction). Detected telomeres of three individual genomic DNA preps from the same *mTERT* $^{-/-}$ G4 mouse (91, 87, and 100 bands) are considerably more than telomeres that were detected from *mTERT* $^{+/-}$ liver tissue (36 bands). **(b)** TeSLA results of high-quality DNA extracted from early (PD 23) and late (PD 81) passages of cultured bowhead whale lung fibroblasts. Both early and late passage cells contain a subset of the shortest telomeres that have not been identified by TRF analysis

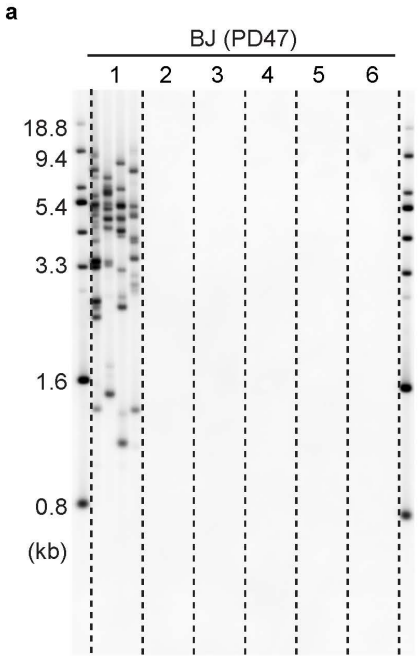
a

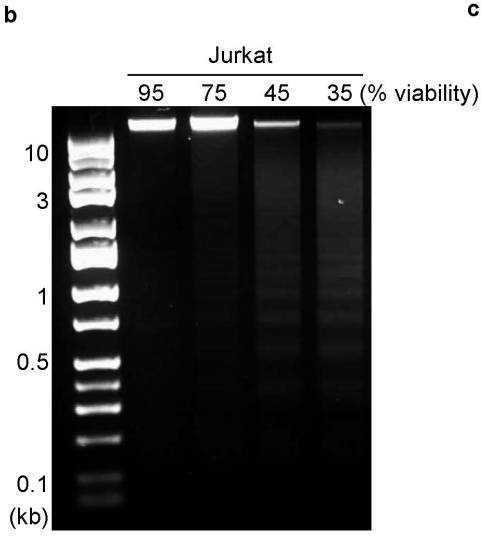

BJ
(PD26)
C106
CEM
HeLa
RAJI
1 2 3
18.8
9.4
5.4
3.3
1.6
0.8
(kb)

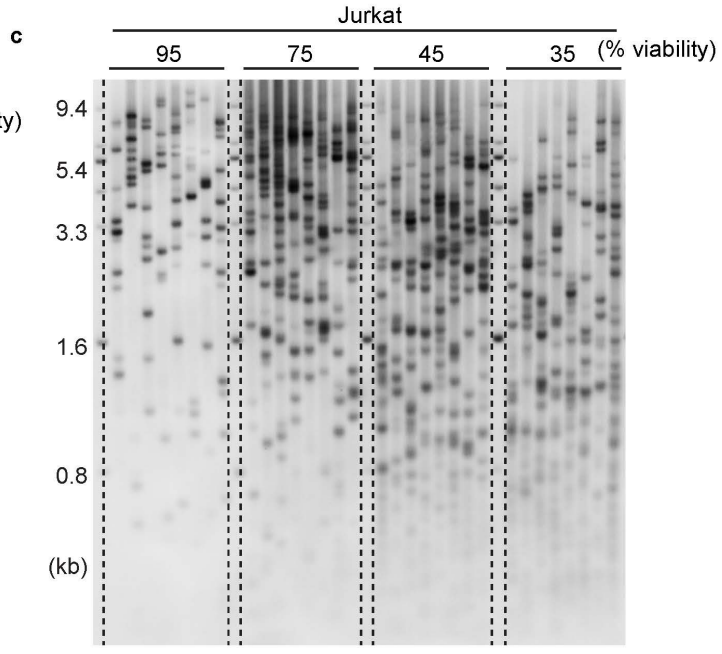
b


1
2
3
Mean TL (kb)
0
1
2
3
4
5
6
7
BJ
(PD 26)
C106
CEM
HeLa
RAJI
Cell line

Supplementary Figure 2.1 TRF analysis using different combinations of restriction enzymes. (**a**) DNA of BJ, C106, CEM, HeLA, and RAJI cells were digested with 1, BfaI/CviAII/MseI/NdeI; 2, HphI/MnlI; 3, AluI/HaeIII/HhaI/HinfI/MspI/RsaI. (**b**) Quantification results of each cell line's TRF analysis in (**a**).

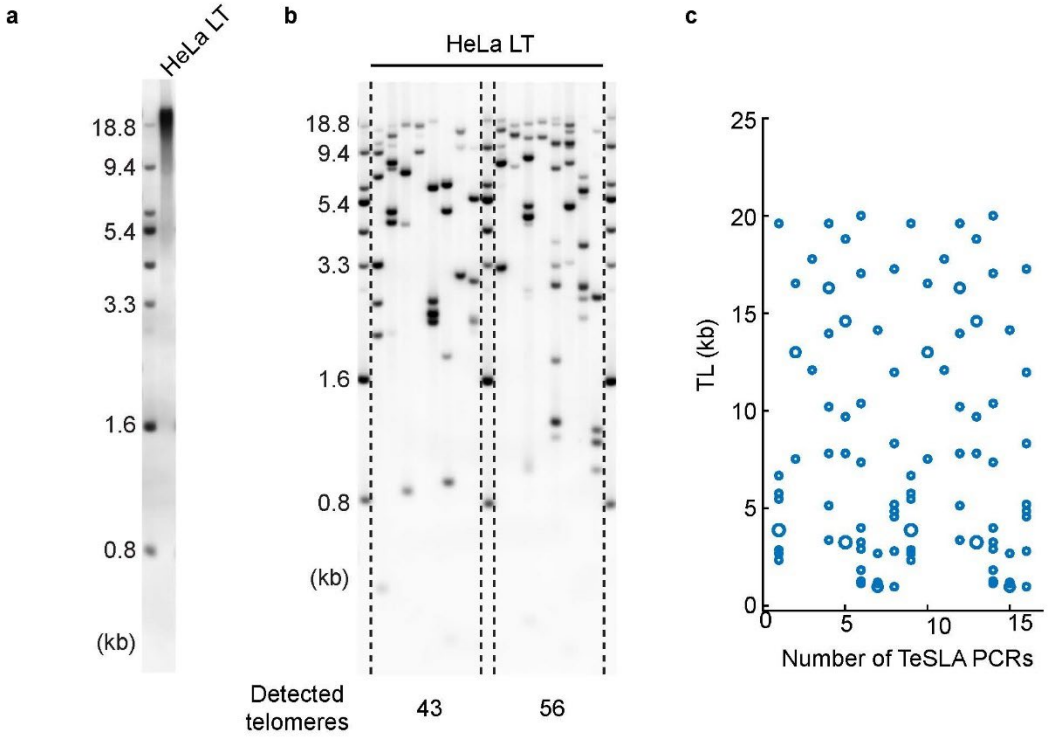

a
BJ (PD47)
1
2
3
4
5
6
18.8
9.4
5.4
3.3
1.6
0.8
(kb)


b
Jurkat
95
75
45
35 (% viability)
10
3
1
0.5
0.1
(kb)

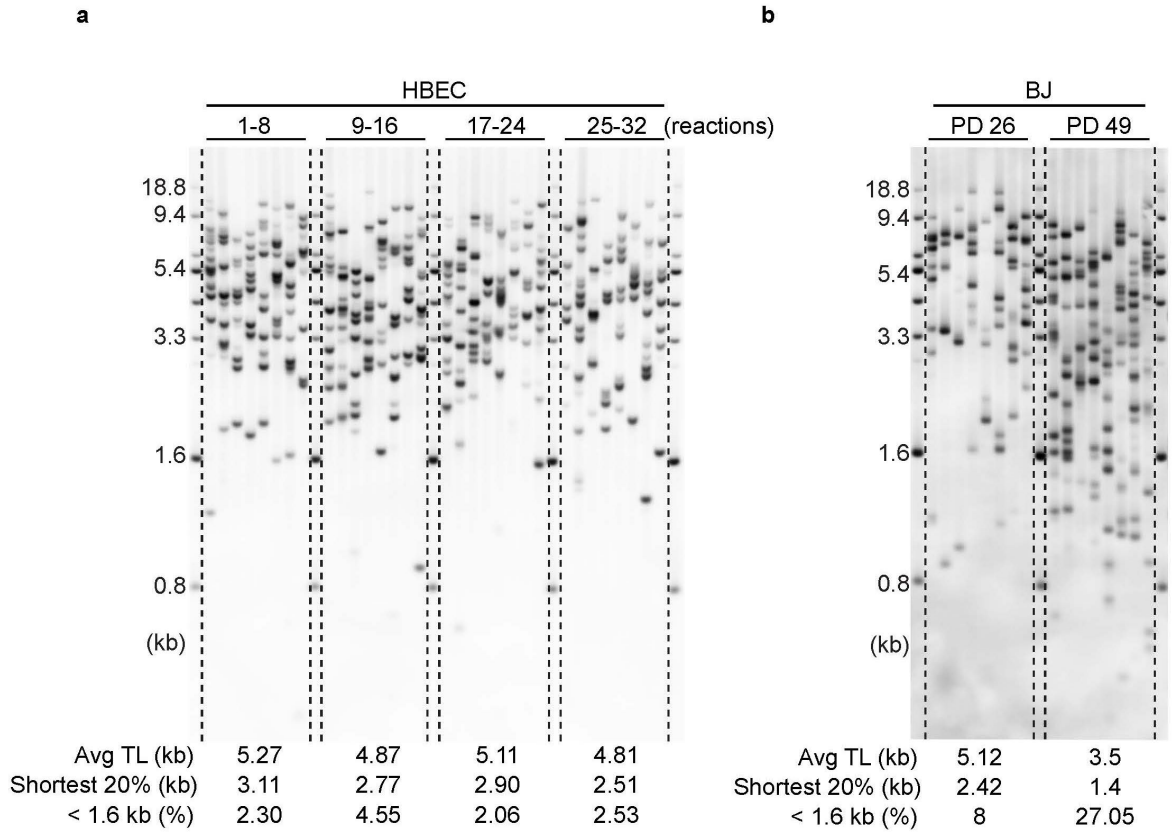

c
Jurkat
95
75
45
35 (% viability)
9.4
5.4
3.3
1.6
0.8
(kb)

Supplementary Figure 2.2 Validation of TeSLA.

(a) DNA from BJ cells was used to perform TeSLA; 1, positive control; 2, negative control without primers for PCR; 3, no TeSLA-Ts for ligations at telomere overhangs; 4, without digestion with REs; 5, no TeSLA adapters for ligations at genomic and subtelomeric DNA; 6, no ligase for any ligation reactions. (**b**) Same amount of genomic DNA (20 ng of each) from Jurkat cells with different viabilities was separated on 1% agarose gel to evaluate DNA integrity. (**c**) TeSLA of Jurkat cells with different percent viable cells.

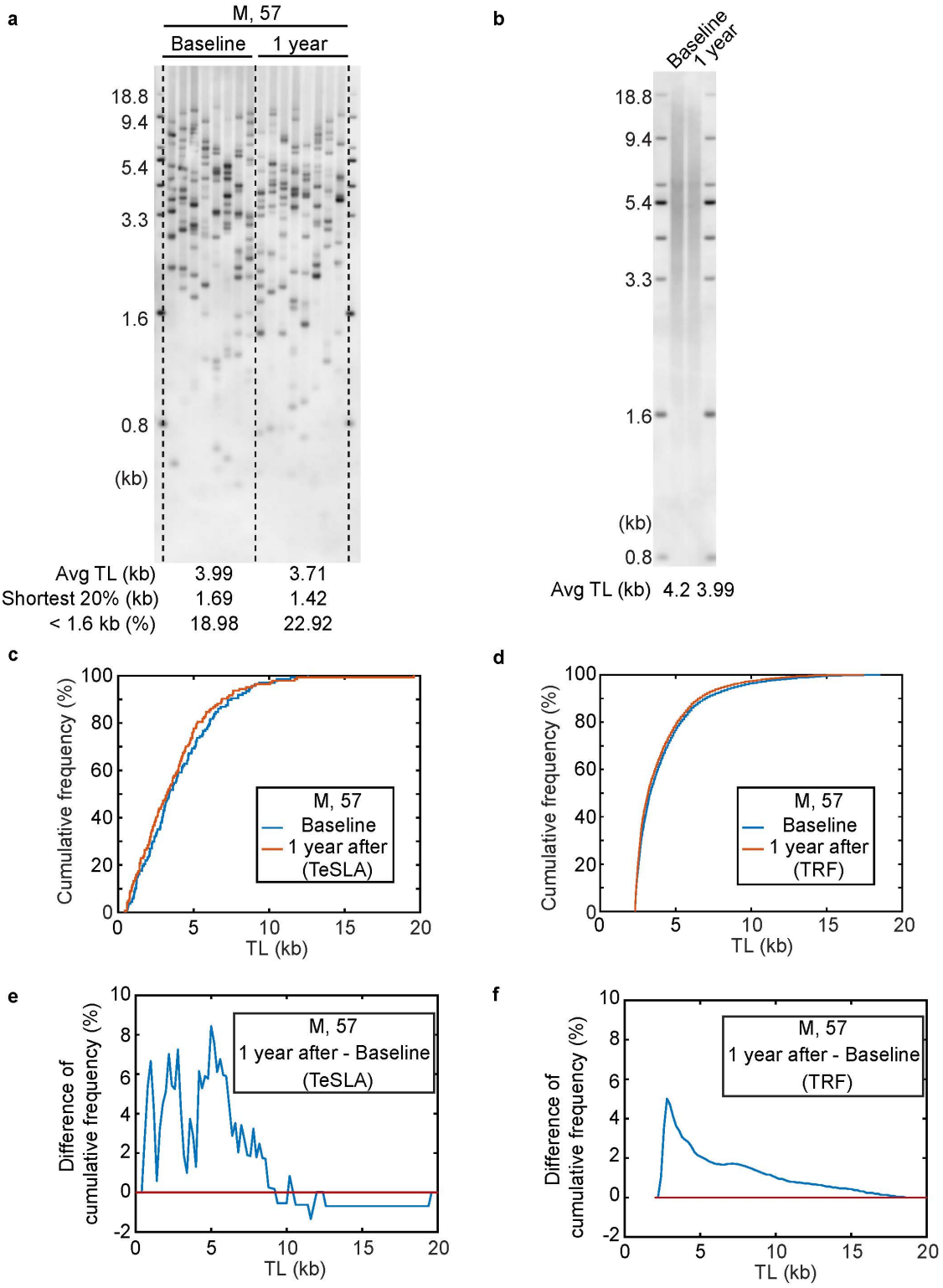


Supplementary Figure 2.3 TLs of HeLa LT cells determined by TRF analysis and TeSLA. **(a)** TLs of HeLa LT cells determined by TRF analysis using the same restriction enzyme combination as TeSLA. **(b)** TeSLA results of HeLa LT (30 pg for each TeSLA PCR reaction, 16 reactions) demonstrated that the upper size limit of TeSLA for TL detection is around 18 kb. **(c)** The scatter plot represents distributions of TLs from TeSLA results (16 reactions) of HeLa LT cells. Each circle represents a particular TL that was detected by TeSLA. The circle size indicates single (small circle) or multiple (large circle) counts for a particular TL.



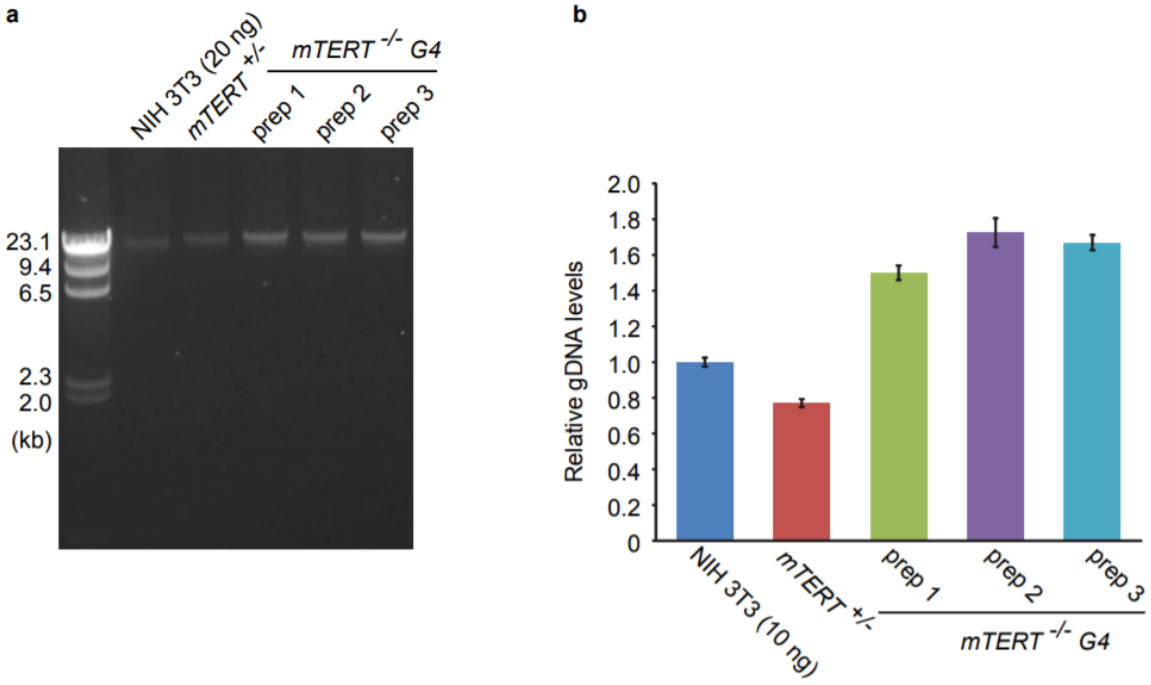
Supplementary Figure 2.4 Intra-variation of TeSLA and TLs determined by TeSLA of BJ cells with different PDs.

(a) TeSLA results of HBEC (30 pg for each TeSLA reaction, 32 reactions) to determine the intra-variation of TeSLA. (**b**) DNA extracted from different PDs (PD 26 and PD 49) of BJ cells were used to determine TL by TeSLA.


a
M, 57
Baseline
1 year
18.8
9.4
5.4
3.3
1.6
0.8
(kb)
Avg TL (kb) 3.99 3.71
Shortest 20% (kb) 1.69 1.42
< 1.6 kb (%) 18.98 22.92
b
Baseline
1 year
18.8
9.4
5.4
3.3
1.6
0.8
(kb)
Avg TL (kb) 4.2 3.99
c
Cumulative frequency (%)
M, 57
Baseline
1 year after
(TeSLA)
TL (kb)
d
Cumulative frequency (%)
M, 57
Baseline
1 year after
(TRF)
TL (kb)
e
Difference of cumulative frequency (%)
M, 57
1 year after - Baseline
(TeSLA)
TL (kb)
f
Difference of cumulative frequency (%)
M, 57
1 year after - Baseline
(TRF)
TL (kb)

Supplementary Figure 2.5 Using TeSLA and TRF analysis to determine changes of TLs for human longitudinally over a one year period.

(**a**, **b**) TLs of DNA isolated from PBMCs from a healthy male at age 57 (baseline) and 1 year later were measured by TeSLA (**a**) and TRF analysis (**b**). (**c**, **d**) Empirical distribution curves based on TeSLA (**c**) and TRF analysis (**d**) results to represent TL distributions. The blue (red) lines are TL distributions at baseline (in 1 year after). The increase of cumulative frequency at a TL indicates the effect of TL shortening at the TL. (**e**, **f**) One-year differences in cumulative frequencies as a function of TLs describe one- year change in TL distribution by TeSLA (**e**) and TRF analysis (**f**).



Supplementary Figure 2.6 Quality control and q-PCR quantification of extracted DNA from *mTERT*$^{+/-}$ and *mTERT* $^{-/-}$, G4 mice.

(a) Extracted genomic DNA from mTERT +/- mouse liver tissue, and 3 individual DNA preps from the same *mTERT* $^{-/-}$ mouse liver tissue was separated on 1% agarose gel to evaluate DNA integrity. 20 ng of DNA from NIH 3T3 cells was served as a standard and positive control. (b) Relative DNA levels of *mTERT* $^{+/-}$ and *mTERT* $^{-/-}$ DNA compared with NIH 3T3. Relative DNA levels of each mouse DNA were calculated and normalized to 10 ng of genomic DNA from NIH3T3 cells by q-PCR to amplify mouse B1 repeats. After normalizing DNA concentrations of each sample, the same amount of DNA (50 ng of each) was used to perform TeSLA in **Figure 2.7a**. (**b**; mean and s.e.m., $n = 3$)

Oligos for TeSLA	sequence
TeSLA-T1	5'-ACT GGC CAC GTG TTT TGA TCG ACC CTA AC-3'
TeSLA-T2	5'-ACT GGC CAC GTG TTT TGA TCG ATA ACC CT-3'
TeSLA-T3	5'-ACT GGC CAC GTG TTT TGA TCG ACC TAA CC-3'
TeSLA-T4	5'-ACT GGC CAC GTG TTT TGA TCG ACT AAC CC-3'
TeSLA-T5	5'-ACT GGC CAC GTG TTT TGA TCG AAA CCC TA-3'
TeSLA-T6	5'-ACT GGC CAC GTG TTT TGA TCG AAC CCT AA-3'
TeSLA adapter short	5'-GGT TAC TTT GTA AGC CTG TC[SpcC3]-3'
TeSLA adapter TA	5'-[Phos] TAG ACA GGC TTA CAA AGT AAC CAT GGT GGA GAA TTC TGT CGT CTT CAC GCT ACA TT [SpcC3]-3'
TeSLA adapter AT	5'-[Phos] ATG ACA GGC TTA CAA AGT AAC CAT GGT GGA GAA TTC TGT CGT CTT CAC GCT ACA TT [SpcC3]-3'
AP	5'-TGT AGC GTG AAG ACG ACA GAA-3'
TeSLA-TP	5'-TGG CCA CGT GTT TTG ATC GA-3'
Oligos for mouse genomic DNA	sequence
mB1F	5'-CAG AGG CAG GCG GAT TT-3'
mB1R	5'-GAC AGG GTT TCT CTG TAG CC-3'

[Phos] represents 5' phosphorylation; [SpcC3] represents C3 spacer

Supplementary Table 2.1 Oligonucleotides used for TeSLA and mouse genomic DNA quantification.

CHAPTER THREE

IMAGING ASSAY TO PROBE THE ROLE OF TELOMERE LENGTH SHORTENING ON TELOMERE-GENE INTERACTIONS IN SINGLE CELLS

3.1 Introduction

Telomeres are repetitive nucleotide sequences (TTAGGGn) capping the end of chromosomes. The length of telomeres becomes progressively shorter after each cell division (Harley et al. 1990), which appears to serve as a clock or replicometer of human cellular lifespan. Telomeres prevent chromosome ends from fusion, degradation, and being recognized as double strand DNA breaks (O'Sullivan and Karlseder 2010b; Webb et al. 2013). Cells undergo replicative senescence when a single or perhaps a few telomeres become very short and unprotected, which results in DNA damage at telomeres (Zou et al. 2004). Importantly, previous work has shown that entry into cellular senescence and chromosome instability are dictated by the shortest telomere length, not the average telomere length (Hemann et al. 2001; Zou et al. 2004). Germline cells and some highly proliferative stem-like cells can transiently express the ribonucleoprotein enzyme complex, telomerase, that can partially maintain telomere lengths (Wright et al. 1996). In contrast, the vast majority of adult somatic cells do not express telomerase. Thus, cells lose a small amount of telomeric sequences after each cell division due to incomplete DNA lagging strand synthesis also known as the end replication problem

(Watson 1972). In addition, other factors such as oxidative stress may contribute to more rapid telomere erosion (Reichert and Stier 2017; von Zglinicki 2002). With progressive telomere shortening, cells will gradually reach a senescence checkpoint (Wright and Shay 1992b). Premalignant cells can bypass this checkpoint by acquiring p53 or pRB/p16 mutations to keep dividing (extended lifespan) until reaching a crisis checkpoint (Wright et al. 1989), at which step almost all cells will die. Only very few rare cells are capable of acquiring a method to maintain their telomere length in order to continue to divide (Shay and Wright 1989). While telomerase is almost universally activated at this stage, other mechanisms such as the Alternative Lengthening of Telomeres (ALT) DNA recombination pathway have also been identified (Cesare and Reddel 2010). Once a telomere maintenance mechanism is achieved, cells reach a new steady state with unlimited potential to divide and this enables additional genetic and epigenetic changes leading to cancer development (Mathon and Lloyd 2001; Shay 2014).

Studies have shown that the gradual decrease in telomere lengths can also regulate gene expression. Originally discovered in yeast, it was named Telomere Position Effect (TPE) (Gottschling et al. 1990; Sandell and Zakian 1992; Stavenhagen and Zakian 1998; Wright and Shay 1992a), indicating that expression of genes adjacent to telomeres can be repressed. In mammalian cells, insertion of a luciferase reporter into the genome at short vs long distances from telomeres showed significant expression variation (Baur et al. 2001) with reduced luciferase expression when adjacent to a telomere but not when inserted far distances from a telomere. Interferon Stimulating Gene 15 (ISG15) was the first endogenous mammalian gene

reported as regulated by telomere length (Lou et al. 2009). It has low expression when telomeres are long, and gradually higher expression as telomeres get progressively shorter. The coupling between telomere length and gene expression for ISG15 was later reproduced in human fibroblasts and myoblasts (Robin et al. 2014; Stadler et al. 2013). In these studies, re-elongation of telomeres in older cells with short telomeres by expression of telomerase, reversed the expression of ISG15. Curiously, telomere length dependent expression was not observed for genes located between the telomere and specific target genes (e.g. ISG15 which is approximately 1MB from the telomere). This modified form of TPE was termed Telomere Position Effect over Long Distance (TPE-OLD) or telomere looping (Kim et al. 2016; Mukherjee et al. 2018; Wood et al. 2015; Wood et al. 2014). The hypothetical model for TPE-OLD is that telomeres can loop back and interact with target genes with the help of shelterin protein complex (de Lange 2005; Kim et al. 2016). Short telomeres in old cells are unable to maintain such interactions and dissociate more frequently from the target genes, which leads to changes in gene expression (Figure 3.1a). Therefore, TPE-OLD may explain why some genes relatively far away (up to 10 MB) from a telomere are regulated by telomere length shortening, yet other genes closer to the telomere are not.

To further investigate TPE-OLD genes and understand the mechanisms involved, we need assays to probe gene expression variation in aging cell populations with a systematic and unbiased identification of interactions between target gene candidates and associated telomeres. It thus becomes important to develop higher throughput methods that can confirm the regulatory role of telomere length shortening on gene expression. Sequencing-based assays

like chromosome conformation capture (3C) and Hi-C (de Wit and de Laat 2012; Denker and de Laat 2016; Lieberman-Aiden et al. 2009) have been widely applied on research of genome organization with large cell numbers. One previous study showed that Hi-C could capture potential long-range interactions on certain chromosomes (Robin et al. 2014). However, these assays require millions of cells and lack the distinction of gene-telomere interaction in individual cells.

In the present study, we applied a 3D-FISH imaging approach to quantify distances between the loci of a TPE-OLD gene and a unique subtelomeric region on the same chromosome. We observed that for ISG15 and TERT (TElomerase Reverse Transcriptase), the loci of fluorescent probe pairs have increased mutual distances (separation) when telomeres become short, indicating that telomere length shortening may regulate gene expression via mutual interactions in 3D. We also found that these distances were reversed in old cells whose telomeres were experimentally elongated by ectopic introduction of hTERT (catalytic rate limiting component of telomerase). We observed that at the population level distance increases between TPE-OLD genes and telomeres are fairly rare events. Therefore, we implemented an automated pipeline, relying on 3D epifluorescence imaging, to acquire robust statistics of distance distributions. This allowed us to scrutinize the regulatory roles of progressive telomere shortening on gene expression levels. We observed that the loci of the ISG15 and TERT genes both have more interactions with their corresponding telomere on the same chromosome in young cells compared to old cells. With telomere length shortening, the telomere associated

interactions are gradually decreasing (like a rheostat) with protein levels increasing for ISG15. In contrast, the TERT progression appears to be more similar to an off/on switch.

The new results of this study and the technologies developed provide an easier platform for future work to systematically probe the significance of TPE-OLD as a mechanism regulating gene expression in normal human aging. The image analysis pipeline also provides automated and time saving solutions on quantifying genome interaction in 3D.

3.2 Methods

3.2.1 Droplet digital PCR (ddPCR)

ddPCR was performed on samples using previously described protocols (Kim et al. 2016; O'Hara et al. 2019). mRNA was extracted from cell pellets using RNeasy plus mini kit (Qiagen) and reverse-transcribed using cDNA synthesis kit (Bio-Rad). Each 20 μL ddPCR reaction contained a final concentration of 1× EvaGreen ddPCR Supermix (Bio-Rad), 100 nM primers, and 10 ng cDNA sample. After PCR, fluorescent intensity of each oil droplet was measured using QX100 droplet reader (Bio-Rad). The threshold for positive droplets could be automatically or manually determined based on the baseline fluorescence level. The final software output generated quantitative results of template molecule concentration in 10ng RNA.

3.2.2 Telomere shortest length assay (TeSLA)

Genomic DNA was extracted from cell pellets using Gentra Puregene DNA Extraction Kit (Qiagen). TeSLA was performed as previously described (Lai et al. 2017b). DNA (50 ng) was

first ligated with TeSLA-T 1-6, followed by digestion using four restriction enzymes (*Bfa*I/*Cvi*AII/*Mse*I/*Nde*I, New England Biolabs). After dephosphorylation with rSAP (New England *Biolabs*), digested DNA was ligated with double-stranded adapters to ensure the amplification of telomeric DNA by PCR (94 °C for 2 min followed by 26 cycles at 94 °C for 15 s, 60 °C for 30 s, and 72 °C for 15 min).

PCR products were separated on a 0.85% agarose gel at 2 V/cm for 16 h. The DNA bands were first transferred to a Hybond-N+ membrane (GE) and then fixed by UV crosslinking. The membrane was then hybridized with DIG-labeled telomere probe at 42 °C overnight, followed by washing with buffer 1 (2× saline-sodium citrate (SSC), 0.1% sodium dodecyl sulfate (SDS)) at RT for 15 min, washing with buffer 2 (0.5× SSC, and 0.1% SDS) at 60 °C for 15 min twice, and washing with buffer 3 (1× maleic acid buffer with 0.3% Tween-20) for 5 min. The membrane was then incubated with 1× DIG blocking solution at RT for 30 min, followed by incubation of anti-DIG antibody (Roche) at RT for 30 min in 1× blocking solution (1 to 10,000 dilution). After washing with DIG buffer twice, telomere signals on membrane were detected by incubating in dark with CDP-star (Roche) for 5 min. Telomere bands image was taken by G:box. The average telomere length and percentage of short telomeres (below 1.6 kb) were quantified using TeSLA Quant software.

3.2.3 Cell culture and fixation

Primary human fibroblasts (BJ cells) were cultured in Medium X (DMEM : Media199 = 4:1, Hyclone) containing 10% cosmic calf serum (Hyclone) at 37°C. Cells with different population

doublings (PD) were passaged using 0.05% Trypsin-EDTA (1X, gibco) and 20-50K cells were seeded in each chamber of the 4-chamber glass slide (LAB-TEK).

Slide chambers were manually removed after overnight incubation. Cells on slides were fixed with 4% paraformaldehyde (PFA), followed by cytoplasm washout with 2% PFA/0.5% Triton X-100. Then cells were permeabilized with 0.5% Triton X-100 in PBS and incubated in 20% glycerol/PBS for at least 20 min. 3X liquid nitrogen freezing-thawing cycles were then performed on slides for further permeabilization. Slides were incubated again in 20% glycerol/PBS at RT for 1 min after each cycle, followed by 5 min wash in 1X PBS and 30 min incubation in 0.1M HCl. Finally, slides were preserved at 4°C in 50% formamide (pH 7) /2X SSC in a Coplin jar shielded from light.

3.2.4 Probes hybridization

Slides with cells were stained with hybridization probes ISG15-FITC, TERT-FITC (Empire Genomics); 1pter and 5pter subtelomere specific probes, Texas Red (Cytocell) following the vendor's protocol. The hybridization lasted overnight at 37˚C in a light-tight humidified chamber. The next day, slides were sequentially washed with 0.2X SSC, 2X SSC, and 1X PBS, stained in 1ug/ml DAPI solution and mounted with Fluoromount-G (ThermoFisher).

3.2.5 Image acquisition

Images were acquired in 3D using a Nikon Ti-Eclipse widefield microscope equipped with 60X/NA=1.4 oil lens, a CMOS camera, and filters for DAPI, FITC and TRITC. The alignment

of microscope channel was qualitatively checked using TetraSpeck Microspheres (Invitrogen). For the fine correction of residual chromatic aberration, we implemented a compensation schema at the level of the distance calculation (see Image Analysis, below). For each slide, multiple, non-overlapping fields of view for acquisition of 3D stacks were defined on the positions of a square grid with an interval of 0.5 mm. After the stage was moved to the stack position, stacks were acquired over a range of +/- 10um above and below the set focal plane with a voxel size of 108 nm x 108 nm x 293 nm.

3.2.6 Image analysis

Image preprocessing

Multiple 3D stack images were sequentially read and analyzed using a custom-written image processing pipeline. After reading a stack the voxel intensities in each channel were normalized to the min/max range of [0, 1].

Nuclei mask segmentation in 3D

Nuclei masks were segmented from the DAPI channel. The pre-processed DAPI channel was compressed to 2D by maximum intensity projection and then convolved with a Gaussian filter with $\sigma_{X/Y}$ = 3.75 um, matching approximately the size of a fibroblast nucleus. Local intensity maxima in the filtered image were marked as the centers of 45 x 45 um^2 squares. Bounding boxes of 45 x 45 um^2 x stack height were defined about these local maxima. Only boxes

containing one maximum were further processed. The original cell seeding density was calibrated to minimize the overlap bounding boxes.

The actual nucleus volume in every bounding box was segmented slice by slice in 2D and then assembled in 3D. The 2D segmentation integrated the information from multi-scale filtering and thresholding, followed by majority voting to determine whether pixels fell inside or outside the nucleus volume. Given the fact that top/bottom slices are blurred and should contain less nuclei voxel, global thresholding was performed to refine the 3D mask so that any voxels with intensities less than 0.6 x global threshold were marked as background. Nuclei with extreme volume sizes or touching the image boundary were removed automatically. On average a single field of view contained ~10 valid nuclei for further processing.

Spot detection

Nuclei masks extracted from the DAPI channel were then applied to the fluorescent channels to detect in each nucleus independently FISH probes as diffraction-limited spots using a previously described pipeline (Aguet et al. 2013; Roudot et al. 2017). These algorithms apply statistical testing for the selection spot signals deemed as significantly brighter than the background. P-values ranged from 0.05 to 0.1, intentionally set to suppress false negatives at the risk of a higher false positive rate. The break-down of spot detection into a nucleus-by-nucleus protocol was essential to account for the vast difference in fluorescent background between nuclei.

Spot pairing

Corresponding pairs of FITC and TRITC spots were identified by solving linear assignment problem in bipartite graph. In brief, the graph was computed by considering all possible spot pairs i, j in the FITC and TRITC channels, respectively, with a 3D distance less than 5 um. The spot candidate intensities (I_i and I_j) were also recorded to calculate a pairing score $S_{ij} = d_{ij}/\sqrt{I_i I_j}$., i.e. spot pairs with high brightness and short distance had the lowest scores. We then applied a modified Hungarian algorithm (Kuhn 1955) to identify among all possible pairs in the graph the two mutually exclusive pairs among the pair assignments with overall smallest score. The scores of the two selected pairs tested against the scores of all other pair assignments. We required that their scores be 50% smaller than the following spot pair score (Supplementary Figure 3.1). Nuclei for which this requirement was not fulfilled were eliminated from the data set.

Correction of chromatic shifts

Before compiling the selected spot pairs into distance distributions reflecting the telomere-target gene interaction under a particular experimental condition we eliminated distance bias due to FITC-TRITC channel misalignment (chromatic aberration). To accomplish this for one particular experiment we computed the mean 3D displacement vector from the FITC to the TRITC spot and subtracted it from the individual displacements, i.e. the corrected displacement vector distribution has a mean value of [0, 0, 0]. Distance distributions for statistical analysis were then computed based on these corrected vectors. The underlying assumption of this

correction protocol is that the vast majority of displacement vectors represent short, random distances between interacting subtelomeric and target gene sequences with no preferred spatial directionality.

3.2.7 Analysis of interstitial telomeric sequences

ITS analysis was performed using IGV software (Thorvaldsdottir et al. 2013). Specific motifs of TTAGGG and the complementary sequence CCCTAA were searched using reference genome hg19. We defined the search region on chromosomes 1 and 5, respectively, to cover the entire gene with flanking sequences of 19 kb for ISG15 and 56 kb for TERT. The searching region was visualized in the software output panel with indicated positions of RefSeq Genes and target motifs. We further plot red arrows to highlight the ITS positions.

3.3 Results

3.3.1 Gene expression level changes with telomere length shortening

The relation between telomere length shortening and gene expression regulation under the TPE-OLD mechanism is thought to rely on the interaction between telomere and gene locus, which is mediated by the shelterin protein (Kim et al. 2016; Kim and Shay 2018; Robin et al. 2015; Robin et al. 2014; Stadler et al. 2013) (Figure 3.1a). To measure how the interaction is affected by telomere length shortening in individual cells, we first prepared human fibroblasts (BJ cells) with different ranges of population doubling (PD). Young cells (PD 14 - 32) were expected to have the longest telomeres, and mid-age cells (PD 33 - 50) and old cells (PD > 50)

were expected to have gradually shorter telomeres. To validate this expectation, we performed a Telomere Shortest Length Assay (TeSLA) on all three cell populations. TeSLA allowed us to measure the telomere lengths in a mixed population of cells with a sensitivity for short telomeres below 1.0 kb (Lai et al. 2017a). As expected, the average telomere length became shorter when cells grew older, and the percentage of the shortest telomeres (< 1.6 kb) increased (Figure 3.1b). We then tested the expression level of a previously identified TPE-OLD gene, ISG15, using droplet digital PCR (ddPCR). Indeed, in young cells, the RNA copy number was 10-fold lower than in old cells (Figure 3.1c) and correlates with increase in ISG15 protein levels [22].

3.3.2 Determining telomere-gene interactions using single cell imaging

We next developed an automated image acquisition and analysis pipeline to determine the level of interaction between genes and subtelomeric regions on the same chromosome by single cell screening in 3D. To visualize the interaction between telomeres and TPE-OLD genes, we labeled the locus of a chromosome-specific, subtelomeric sequence and the locus of a target gene of interest using FITC- and TRITC-tagged FISH probes, respectively. We also labeled the overall nuclear volume using a DAPI stain. After fixation and labeling, 3D image stacks were acquired to measure the distances of telomere – gene pairs. Figure 3.2a shows a typical field of view in maximum intensity projection (MIP), with the DAPI channel outlining the nuclei and the zoom-in windows presenting the FITC and TRITC channels for two neighboring nuclei. Both FISH-probe channels indicate four clearly discernible bright spots in the same

location, suggesting co-localization of the subtelomeric probe and the gene of interest, and thus presumptive interaction, for both chromosomes in each nucleus.

We expected that in a population of cells the presence or absence of telomere-gene pairing would be heterogeneous, both because of variability in the DNA sequence interaction per se and in variability in telomere length at the individual cell level (Figure 3.1b). Although previous studies have used manual analyses of small cell populations to demonstrate the shift in telomere-gene interactions between cells with long and short telomeres (Robin et al. 2014), it was unclear for a general case how many cells would be necessary to detect shifts in TPE-OLD regulation between two conditions. To investigate this question using an unbiased data set, we developed a protocol based on a motorized microscope stage to acquire 3D stacks of hundreds to thousands of cells on a single slide. For each slide we typically sampled 50-80 randomly placed fields of views, each containing ~15 cells, on average. We employed widefield epi-fluorescence imaging because of the higher detection sensitivity and the ten-fold faster image acquisition rate compared to confocal microscopy.

We then implemented a fully automated image processing pipeline, delineated in Figure 3.2b, to systematically measure the relation between the FISH markers of the subtelomeric region and TPE-OLD gene of interest in every nucleus. After image pre-processing steps, such as pixel intensity normalization and background removal, we segmented each nucleus from the DAPI channel (Figure 3.2c, see methods for details) and then applied a sub-pipeline for FISH marker detection for each nucleus separately. This eliminated contaminating signals in the space between nuclei (Figure 3.2c, yellow arrows) and allowed an adaptive tuning of image

filters and thresholds for FISH probe identification between nuclei with very different fluorescent backgrounds.

Specifically, in each nuclear volume we first detected spots in 3D in both fluorescent channels, adapting a 3D version of the previously published algorithm by Aguet et al. (Aguet et al. 2013). Statistical spot selection was performed at a p-value of 0.1, which tends to err on the side of false positive candidates. This ensured that the initial spot sets included the signals of all FISH markers with high confidence.

Next, we sought to identify spot pairs between the FITC and TRITC channels that would represent with high likelihood the interaction between a telomere and TPE-OLD gene marker. We made the assumption that both markers produce a relatively bright spot and that the proximity in 3D of corresponding markers, on average, is much greater than the proximity of randomly paired spots, even in the case where the DNA sequences of telomere and TPE-OLD gene do not interact. To capture this model, we computed a pairing score matrix between FITC and TRITC channel spots. Scores were low for bright and proximal spots, whereas scores were high for dim and distant spots. Based on this score matrix we assigned spot pairs by solving the linear assignment problem (LAP) (Jaqaman et al. 2008; Jonker and Volgenant 1987), which identified among all pairing configurations the one with the overall smallest sum of scores. Due to the detection of an unequal number of spots in both channels our LAP implementation relaxed the conditions that every spot must be paired. Finally, our algorithm checked that the pairing scores of the two lowest score assignments were significantly less than the scores of any other pairing (Supplementary Figure 3.1). Only nuclei fulfilling this condition were

accepted as containing valid states of telomere and TPE-OLD gene interactions. Figure 3.2d provides an example of spot detection and pairing. Note that in this particular case the FITC channel contains several nearly identically bright spots; however, the selection of the two relevant FISH markers is unambiguous when considering the detections in the TRITC channel. Supplementary Figure 3.1 displays additional examples of high- and unacceptably low confidence pairings.

3.3.3 The increased separation between the gene of interest and subtelomere along with cell replicative aging

Equipped with this imaging pipeline, we first investigated the gradual separation of a well-established TPE-OLD gene, ISG15 (Lou et al. 2009), from the corresponding telomere on chromosome 1p as cells get older. Regardless of cell age, represented by the population doubling (PD), the vast majority of parings had a 3D distance of less than 500 nm, i.e. the spots in FITC- and TRITC-channel fell within the same point spread function and thus appear visually co-localized (Figure 3.3a). With increasing age, an increasing sub-population of nuclei with distances of between 500 nm and 3 um, are detected suggesting that a larger number of telomeres dissociated from the ISG15 locus. Importantly, at all ages these longer distance pairs describe the exception to the rule. This implies that the expression shifts of TPE-OLD genes (Figure 3.1c) are driven by only a small sub-population of cells, and bulk measurements of DNA-DNA interactions, like 3C and Hi-C, are relatively insensitive in detecting TPE-OLD. Even with a single-cell assay as described here, TPE-OLD can only be confirmed based on a

statistical sample large enough to capture a representative outlier population. To visualize the shift in the outlier population we present the cumulative distributions (Figure 3.3b). In this representation it becomes obvious that interactions between ISG15 and the subtelomere is decreasing as the PD increases. The significance of these shifts is quantified by the Kolmogorov-Smirnov test statistics (Massey 1951) (Figure 3.3a).

3.3.4 Large cell number quantification is required for statistical robustness

According to our analysis thus far, detecting of TPE-OLD therefore relies on the assessment of the abundance of data outliers. This is a notoriously difficult task, strongly depending on a representative and geometrically unbiased sampling of the subtelomere-gene distances for a particular experimental condition.

Therefore, we next investigated how many cells should be imaged to confidently identify a TPE-OLD gene. Specifically, we determined by random sampling of increasingly larger cell subpopulations the number of cells required to reveal the distance distributions difference between age groups. For each subpopulation size we bootstrapped 500 samples from the full cell populations of two age groups (younger and older cells) and then computed the percentage of samples yielding a significant difference between younger versus older cells. We set the condition for a high-confidence distinction between any two age groups as 98% or more of sample pairs be assessed as different. As illustrated in Supplementary Figure 3.2 at least 150 cells are needed to recapitulate the difference between distance distributions of young and old

BJ cells for ISG15 associated telomere interactions. More cells are required to capture the distribution differences if the experimental conditions have a closer age gap.

3.3.5 Identification of an age-controlled switch in the TPE-OLD-based repression of the TERT gene

TERT is the reverse transcriptase protein core component of the telomerase complex, which promotes elongation and partial maintenance of telomere length in stem cells and stable maintenance of telomere length in more than 90% of cancer cells (Shay 2016). A previous study suggested that TERT is another TPE-OLD gene in large long-lived mammals (Kim et al. 2016). We sought to confirm this in our 'aging'-model of human cells. In contrast to the ISG15 gene, we observed that the distance distributions between TERT and associated telomere remained constant up to PD 54, with most mutual distances less than 1 um. (Figure 3.3c). Beyond PD54 there was an abrupt change in the distribution, where a significant fraction (49% of PD 60 fibroblasts) fell outside the 99%-quantile (2.13 um) of the distance distributions of PD54 and younger. To better visualize the switch-like shift we present the cumulative distributions (Figure 3.3d). Since TERT has no activity in most somatic cells (Wright et al. 1996), in previous work we have proposed that TPE-OLD is among the primary mechanisms of repressed telomerase activation during human fetal development [26]. Our present data supports this notion with a direct experiment in a cell-level 'aging' model and identifies a switch-like release of the repression beyond a critical replication count.

3.3.6 TERT immortalized cells observe more interaction between TPE-OLD genes and the subtelomeres in late passage normal cells

Our observation of an increased dissociation of subtelomeres from both the ISG15 and TERT gene loci in old cells (PD60) led to the obvious hypothesis that the loss of interaction directly relates to the shortening of telomeres in an aging cell population, i.e. shorter telomeres have lower probability of interacting with a TPE-OLD gene. Testing this directly would require a concurrent assay of subtelomere-gene interaction and telomere length at the single cell level. Such an assay is currently not feasible. To nonetheless probe the effect of telomere length on the TPE-OLD genes we tested BJ cells with hTERT reintroduced in late passage (Bodnar et al. 1998) and observed the distance distributions for the ISG15 and the TERT genes (Figure 3.4). Confirmed by TeSLA, the telomere length of immortalized BJ cells is significantly re-elongated (Supplementary Figure 3.3). Intriguingly, for both genes hTERT expression restored telomere length in old cells back to the distance distributions of PD18 cells, directly supporting the concept that the TPE-OLD mechanism is controlled by telomere length.

3.4 Discussion

Telomere length shortens along with cell division and aging in all normal human somatic cells. It has been postulated that this shortening can affect the expression level of genes at distances up to 10 MB from the chromosome end, a mechanism referred to a TPE-OLD or telomere looping (Kim et al. 2016; Kim and Shay 2018; Lou et al. 2009; Robin et al. 2015; Robin et al. 2014; Stadler et al. 2013; Wood et al. 2015). To provide more insights of gene-telomere

interactions, we developed an automated pipeline for measuring in hundreds of cells individually the proximity of a unique subtelomere sequence of a chromosome and a potential TPE-OLD regulated gene locus on the same chromosome. This amount of data is required to capture with sufficient confidence the consequences of intrinsically stochastic and convoluted processes that determine the proximity of subtelomere and target gene marker. The pipeline encompasses image preprocessing, nucleus segmentation, and fluorescent spot detection and pairing, from which distance distributions are derived as proxies of telomere-gene interactions. Our analyses of two TPE-OLD genes show that age-related separation of a subtelomere unique region and a gene sequence far from the telomere is likely to be a relatively rare event. Even in the oldest cell population we were technically able to analyze, the vast majority of measured gene-telomere distances were short, in the likely range of interaction. Technically, this implies that any further dissection of the TPE-OLD mechanism requires high-throughput measurements in order to capture in an unbiased fashion the subtle interaction shifts between experimental conditions. Biologically, this implies that TPE-OLD defines a mechanism for stepwise activation of transcriptional activity and protein expression of a small set of select genes in individual cells.

Intriguingly, because of the unbiased acquisition of single cell measurement, our data revealed two genes with differential long-term kinetics of gene-transcription changes with telomere shortening. While the ISG15 gene becomes gradually released from telomere interactions with progressive telomere shortening, the TERT gene displayed no significant interaction level change between young and middle-aged cells. The release of the TERT gene from telomeric

interactions occurred only in the oldest population we could analyze. Given the fact that about 90% of cancer cells with short telomeres have telomerase activity (Jafri et al. 2016; Kim et al. 1994; Shay and Bacchetti 1997), there might be a mechanism that enhances telomerase activation specifically in old cells. A previous study (Kim et al. 2016) also found the enrichment of telomere associated shelterin protein component TRF2 near the hTERT promoter via ChIP. Young cells tend to have more TRF2 interactions with TERT, which is consistent with the observation from our large-scale image data analysis.

An emerging area of interest is the role of telomere 3D looping with interstitial telomeric sequences (ITS) (Simonet et al. 2011), i.e. TTAGGG, through interactions with telomere shelterin proteins (Wood et al. 2014). Telomeres form loops that interact with ITS at distal genomic regions. In the case of ISG15 the interactions likely occur via single ITS within 3 kb from the start codon (Figure 3.5a), while the TERT gene includes many more ITS that provide potential interaction sites (Figure 3.5b). Hence, in the ISG15 case the interaction probability is directly proportional to the telomere length, and thus gradual shortening in a heterogeneous population of cells leads to gradual loss of TPE-OLD. In contrast, the multiple ITS in and surrounding the TERT gene may result in a binding cooperativity that causes the interaction to be insensitive to telomere shortening up to the point at which the telomere is too short to bridge multiple ITS (Figure 3.5c). This potentially explains the switch-like loss of telomere-gene interactions for TERT.

In summary, our assay sets the foundation for a systematic validation and mechanistic analysis of candidate genes whose expression regulation may be co-regulated by TPE-OLD. In a

previous study we found that DNA methylation and histone modifications in the hTERT promoter region showed significant changes as cells developed shorter telomeres and that TRF2, a TTAGGG shelterin protein, may have important roles in these age-dependent genomic changes. These observations offer a model and a partial explanation for how age-dependent changes in the genome structure potentially affect the regulation of genes without initiating a DNA damage response from a critically shortened telomere. In conclusion, changes in telomere looping with increased age (and progressive telomere shortening) may be one mechanism of how cells time changes in physiology over decades. With the improved higher throughput single cell imaging approach described here, it will now be possible to acquire more knowledge of TPE-OLD genes.

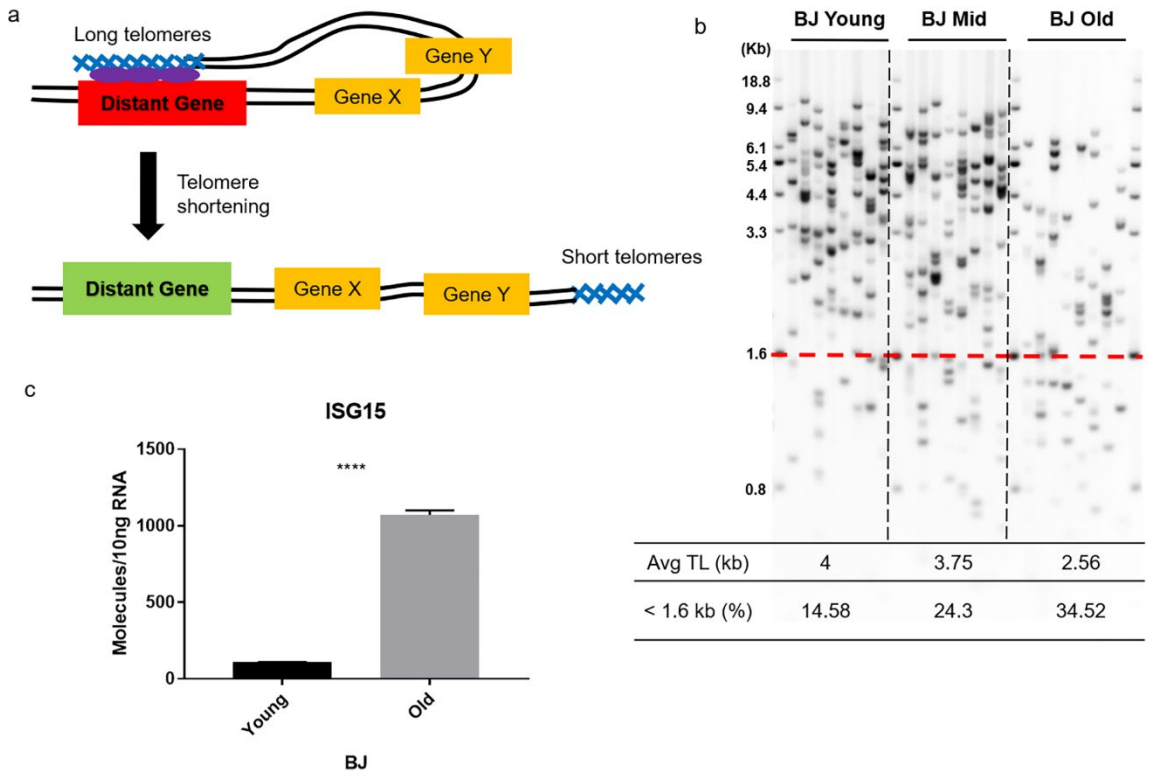


Figure 3.1 TPE-OLD model and gene expression level change in young/old cells. (a) Current model of telomere position effect over long distance (TPE-OLD). Long telomeres can form a loop and therefore interact with target genes over long distance. Red color indicates a repressed state, while green color indicates an activated state. Other genes (yellow color) sitting in between are not affected. The interaction might be mediated by the shelterin protein complex (purple color) and becomes weak as telomeres get short. (b) Telomere length distribution revealed by Telomere Shortest Length Assay (TeSLA) for human fibroblasts (BJ cells) with different population doublings (PD). Old cells tend to have smaller average telomere length and more short telomeres. Quantification of Southern blot bands was accomplished using the software described in (Lai et al. 2017b). (c) ISG15 is a TPE-OLD candidate gene. The RNA

expression level in young BJ cells is repressed (~10 fold lower than in old cells; p-value < 0.0001; n = 3 experiments; error bars indicate standard deviation).

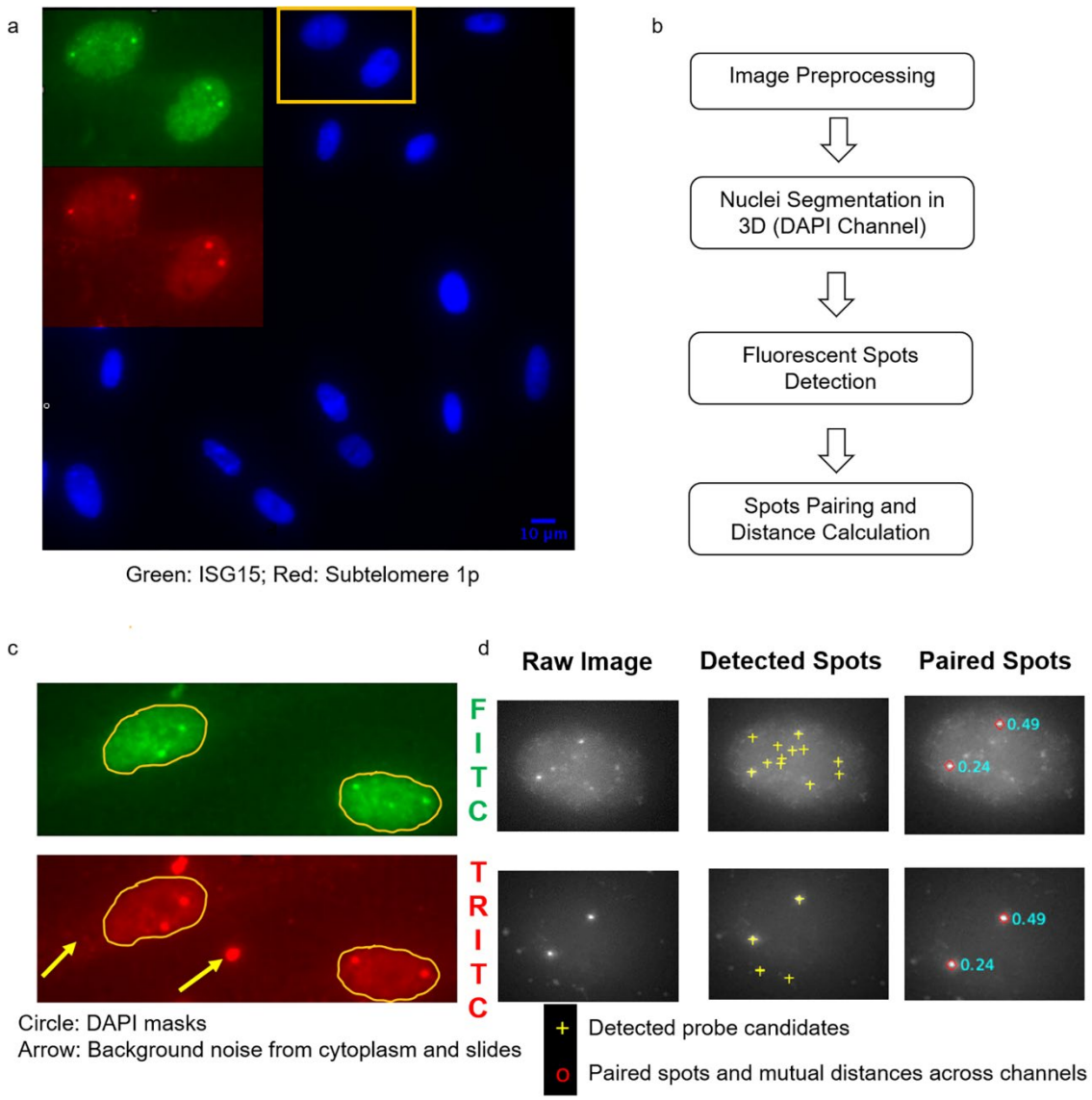


Figure 3.2 Visualization and analysis of the relative position of chromosome-specific subtelomere sequence and target gene locus by FISH. (**a**) Maximum intensity projection (MIP) of a 3D sample image showing nuclei in DAPI channel. Inset: Zoomed version of MIP of FITC (green; ISG15 loci) and TRITC (red; specific subtelomeric region near chromosome 1p) channels associated with the nuclei indicated in the overview. The selected nuclei show a

prototypical scenario with two clearly discernible spots per channel, which co-localize between channels. (**b**) Overview of image analysis pipeline. (**c**) MIP of FITC and TRITC channels indicating representative scenarios of FISH probe clutter (arrows) outside the nuclear perimeter as well as ambiguous spot signals inside the nucleus. **(d)** MIPs of FITC and TRITC channels indicating representative scenarios of FISH probe signals (left). Both channels contain more than two discernible spots, all of which are correctly detected as probe candidates based on the statistical significance of foreground to background contrast (middle, yellow crosses). Based on a pairing algorithm and selection of the overall two brightest pairs FISH probes tagging bona fide subtelomeric and target gene sequence are identified (right, red circles). Cyan numbers in the right column indicate the distance between corresponding probes in um.

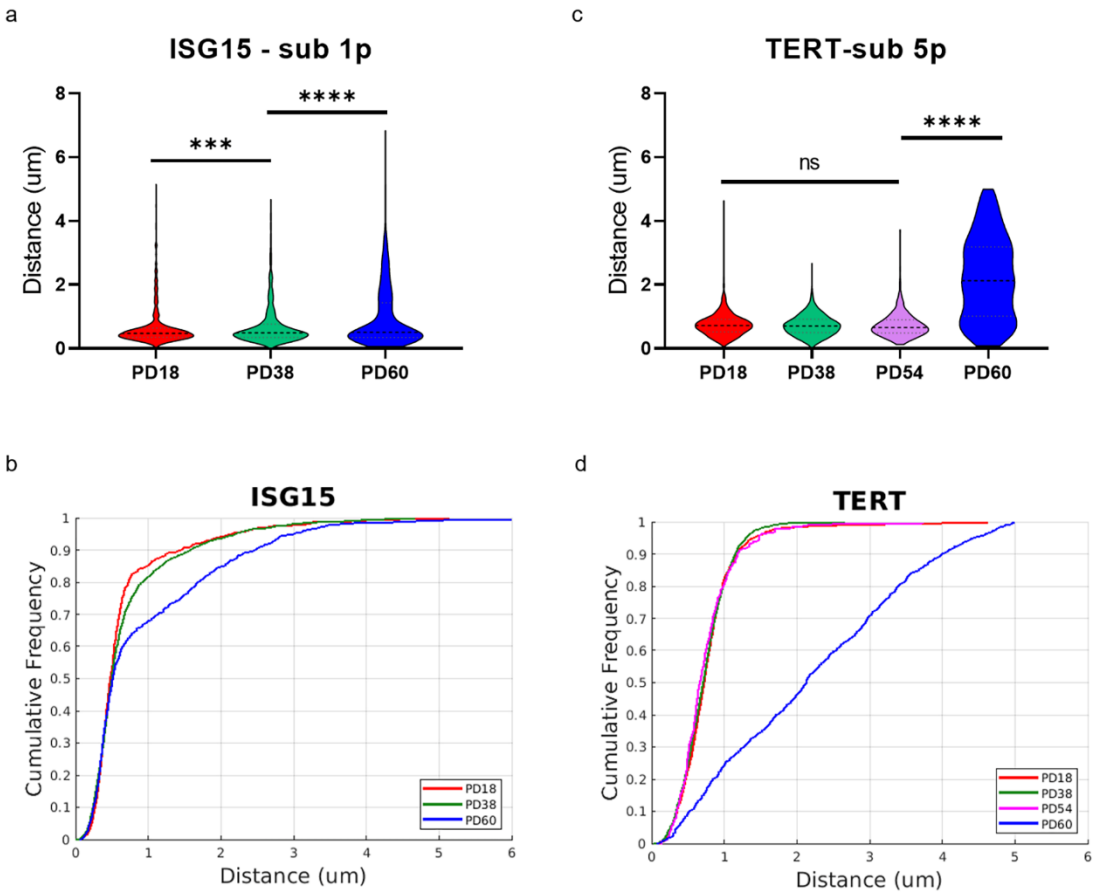


Figure 3.3 TPE-OLD genes interact more with long telomeres. **(a)** Violin plots of distances between ISG15 and subtelomere 1p for BJ fibroblasts with PD18 (n = 414 cells), PD38 (n = 1079 cells), and PD60 (n = 416 cells). 25%, 50%, and 75% quantile lines are plotted. Kolmogorov–Smirnov (KS) test was applied to determine the significance of differences between the distributions under different conditions. The significant p values were observed between any two distance distributions ($p < 0.001$). **(b)** Cumulative histograms of distances between ISG15 and subtelomere 1p for BJ fibroblasts with PD18, PD38, and PD60. Young cells have more short-distance pairs, suggesting closer interaction between the telomere and ISG15. **(c)** Violin plots of mutual distances between TERT and subtelomere 5p for BJ

fibroblasts with PD18 (n = 704 cells), PD38 (n = 835 cells), PD54 (n = 162 cells), and PD60 (n = 536 cells). KS test was applied to check if two distributions are significantly different from each other. The significant p value was only observed between PD60 cells and other groups ($p < 0.0001$). **(d)** CDF plots of paired distances between TERT and subtelomere 5p for BJ fibroblasts with PD18, PD38, PD54, and PD60. Young cells have more adjacent pairs, which indicates closer interaction between the telomere and TERT.

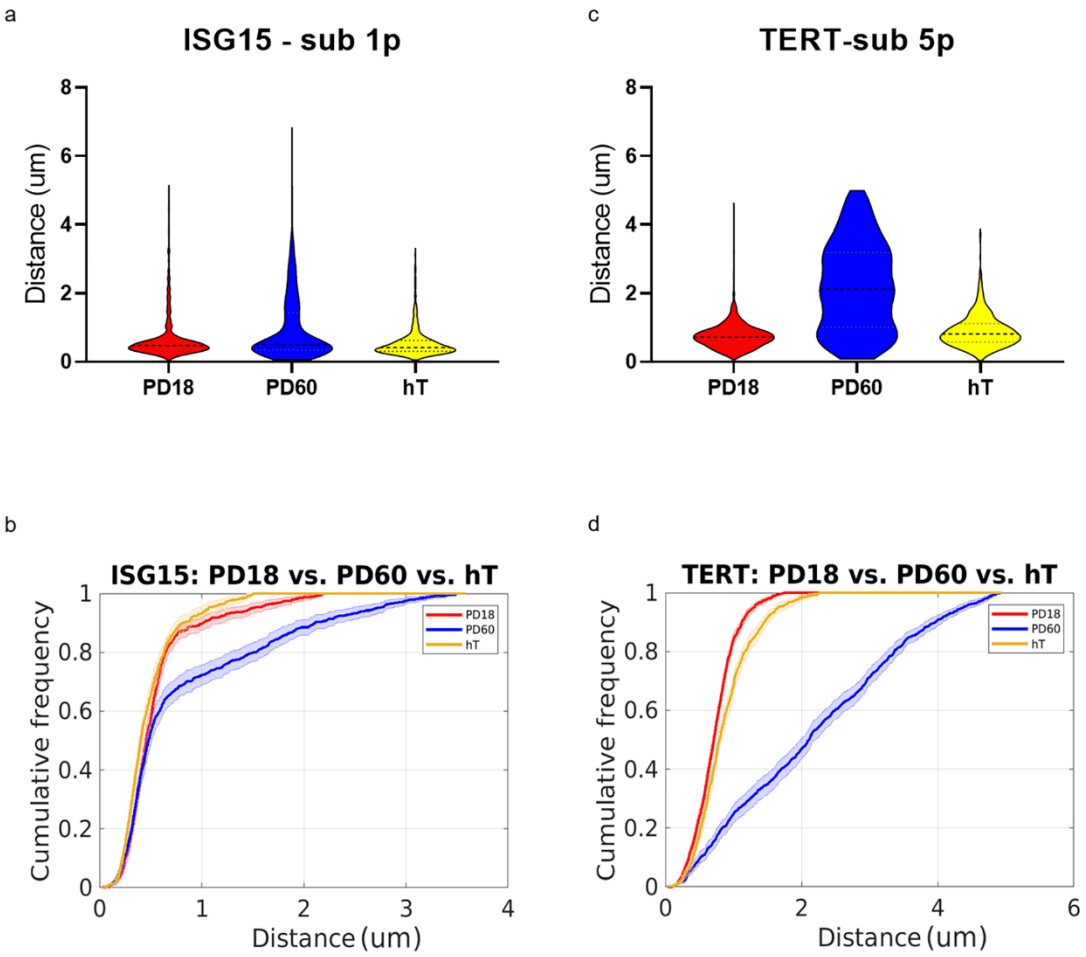


Figure 3.4 TERT immortalized cells reverse the telomere-gene interaction pattern. **(a)** Violin plots of distances between ISG15 and subtelomere 1p for BJ fibroblasts with PD18 (n = 414 cells), PD60 (n = 416 cells), and TERT immortalized (hT; n = 345 cells) cells. 25%, 50%, and 75% quantile lines are plotted. **(b)** Cumulative histograms of the same distances. **(c)** Violin plots of distances between TERT and subtelomere 5p for BJ fibroblasts with PD18 (n = 704 cells), PD60 (n = 536 cells), and TERT (n = 437 cells) immortalized cells. 25%, 50%, and 75% quantile lines are plotted. **(d)** Cumulative histograms of the same distances.

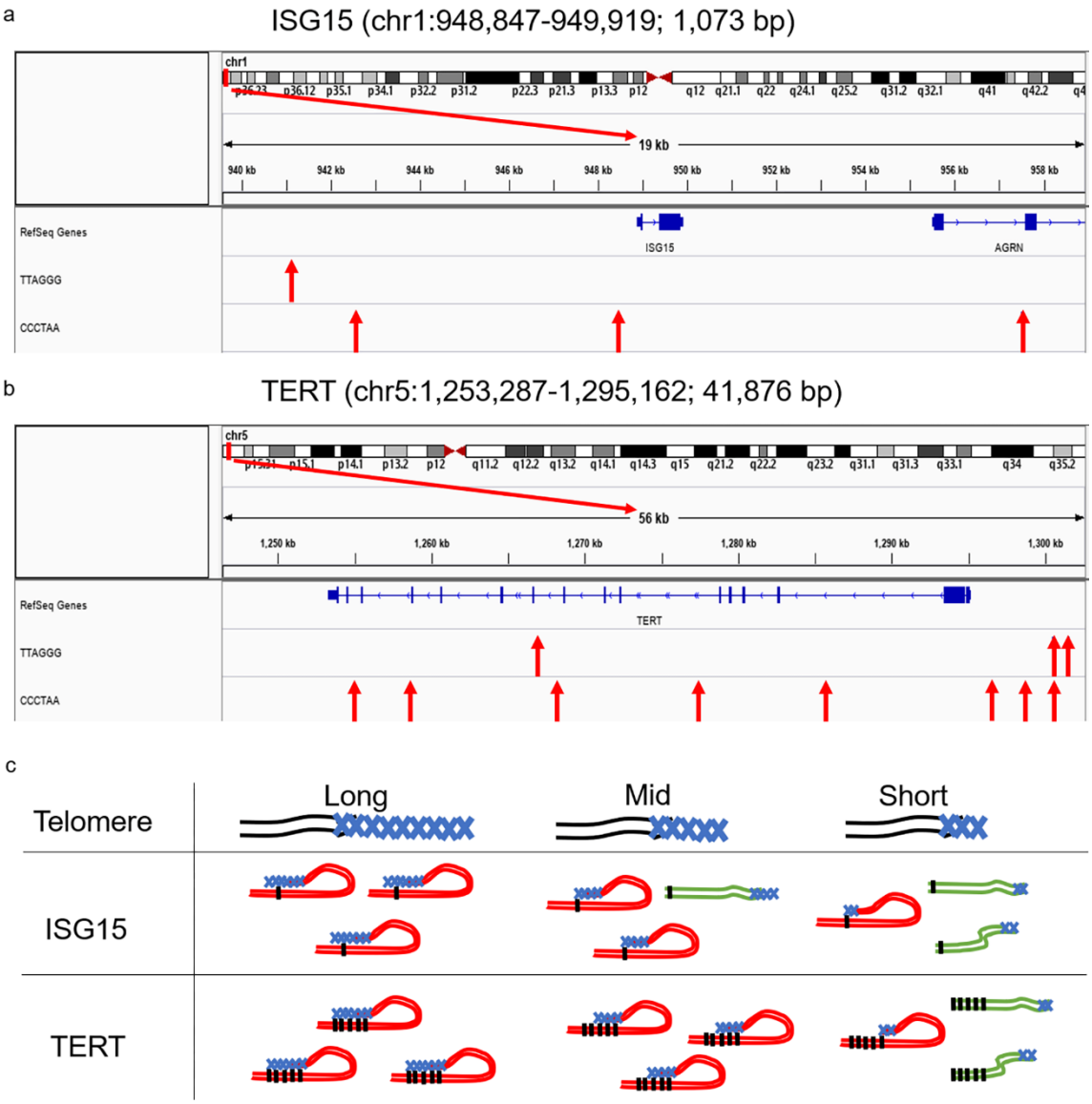
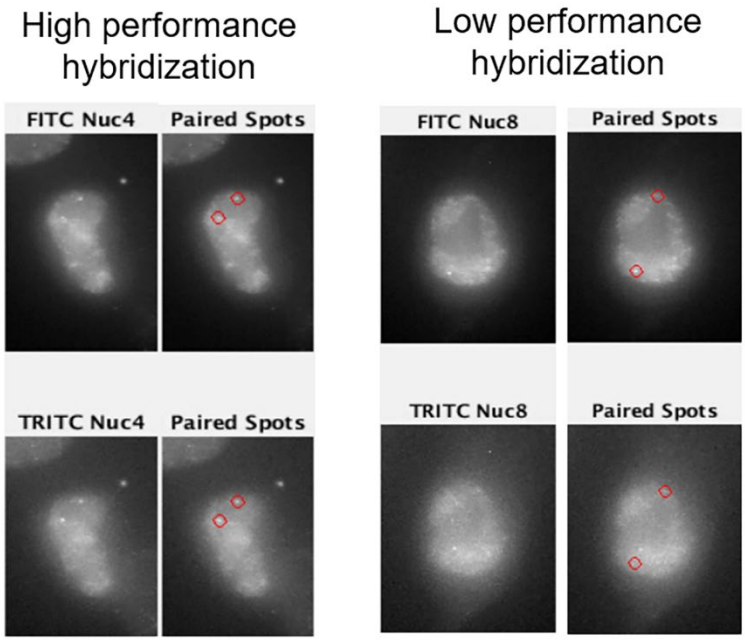


Figure 3.5 Model of telomere-gene interactions modulated by interstitial telomeric sequences. (**a**) Positions of interstitial telomeric sequence (ITS, TTAGGG) and complementary sequence (CCCTAA) indicated by red arrows within and flanking the ISG15 gene. ISG15 gene length and position are indicated. Total lengths of flanking regions, 19 kb. (**b**) Positions of (TTAGGG) and complementary sequence (CCCTAA) indicated by red arrows within and

flanking the TERT gene. TERT gene length and position are indicated. Total lengths of flanking regions, 56 kb. (**c**) Illustrations of gene-telomere interaction change along with telomere length shortening for ISG15 and TERT. For ISG15, the interaction probability is proportional to the telomere length due to the limited number of ITS. Thus, gradual shortening leads to gradual loss of interaction. In contrast, the multiple ITSs within and flanking TERT may result in a binding cooperativity that causes the interaction probability to be insensitive to telomere shorting up to the point at which the telomere is too short to bridge multiple ITSs.

a

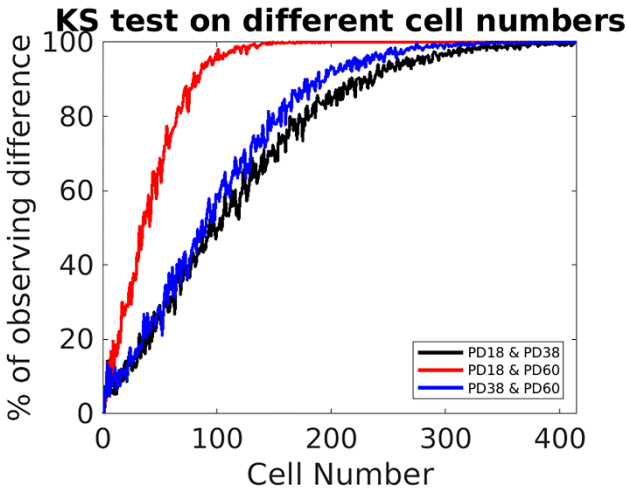


b

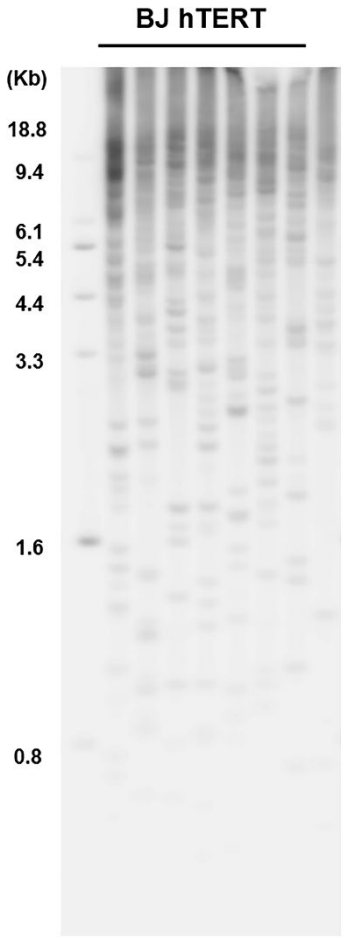
Ranked pairing score table of representative nuclei

Nucleus	Pair 1	Pair 2	Pair 3	Pair 4	Pair 5	Pair 6
High Performance	2.07	3.97	13.99	18.91	28.05	31.09
Low Performance	8.53	45.11	47.93	48.65	60.02	NaN

Supplementary Figure 3.1 Quality control of spot pairing. (**a**) Examples of high-performance hybridization and low performance hybridization in spot pairing. Red circles indicate the positions of top 2 pairs in each nucleus. (**b**) Ranked pairing scores of selected nuclei. With high performance hybridization, the top two candidates have significantly lower score compared to the other scores. With low performance hybridization the scores of the two candidates are not clearly different from the other scores.



Supplementary Figure 3.2 Large cell number quantification is required for statistical robustness. Large cell number is required for quantification. Designated number of cells were randomly selected from the PD18, PD38 and PD60 cell pools using bootstrap. Only with more than 150 cells can the difference between PD18 and PD60be reproducibly observed More cells are required if the two samples have closer PDs.



	Avg TL (kb)	< 1.6 kb (%)
BJ hTERT	4.09	18.75

Supplementary Figure 3.3 TeSLA of TERT immortalized BJ cells shows the re-elongated telomere length. Telomere length distribution revealed by Telomere Shortest Length Assay (TeSLA) for hTERT immortalized old human fibroblasts (BJ cells). Its telomere length was re-elongated due to the function of telomerase after cultivation.

CHAPTER FOUR

SOFTWARE MANUAL

4.1 TeSLA-Quant User's Manual

4.1.1 General Introduction

TeSLA-Quant is designed to automatically quantify TeSLA Southern blot images and generate statistical outcomes. It can automatically mark the band positions, detect band intensity, and calculate relevant statistics, such as average size, percentage of short bands, etc. It supports multiple image formats as input. We recommend using tiff as the default input format.

This software gives users the freedom to use different sets of ladders, manually optimize the detection results, and adjust the detection threshold level. It is also able to detect bands on other gel images, such as Western blots and DNA fingerprints.

4.1.2 System Requirements

TeSLA-Quant is programmed and compiled with Matlab 2016b (Mathworks) and its image processing toolboxes. The required compiler MCR 2016b (9.1) has been incorporated in the installation file 'TeSLAQuant.exe'. Users can simply double click the executable file to install the software in a local computer. The current version is designed for 64-bit Windows 8 or 10

operation system. It may take a few minutes to start the software on computers with low memory or early version systems.

4.1.3 Software Input and Output

Sample Input

After the software installation, the user can click ‘Load Image’ to take the input. It requires the 2D single image for each analysis. We provide a sample image for user’s reference (Figure 4.1). The expectation is to have low background noise, vertically aligned lanes, and one separated lane for ladders.

Users can click ‘Save Figure’ after the detection to save the output panel with .jpg format. The software will automatically generate a result folder on the image’s path, which has the same name with the input image, plus a number indicating the analysis time. That folder contains the following output files:

.txt file: Quantification results

.mat file: Matlab variables containing analysis results

.mat file with ‘imageInput’: Input image stored in a 2D matrix

.jpg file: Output software panel

.xlsx file: Excel form containing each band’s intensity and count number

4.1.4 Step by Step Guide

1. Load image: Click on the ‘Load Image’ button once the software panel has been launched. The default file format is tiff. The software can remember the user defined path and start and will use the same folder when the next image is to be uploaded (Figure 4.2).

2. Crop image: The user is asked to crop the region of interest. The marker lane needs to be included. User can change the cropped region by dragging the frame and finally confirming the choice by double clicking within the frame (Figure 4.3).

3. Run: The user will get preliminary detection results after clicking the ‘Run’ button. Each red dot marks an individual detected band, including the marker lane. The user will have the freedom to manually adjust band detection results by adding or removing bands. After clicking on the ‘Add Bands’, the user can click on any suspicious band center to manually add one band. One or multiple bands can also be removed using the ‘Delete Bands’ function. The user needs to click on the image and drag a frame which contains the unwanted bands. Those bands will be removed after user double clicks within the frame. If the image noise level is different from the sample input, the user can adjust the lane/band threshold factor to achieve fine or coarse detection results. The higher number indicates coarser detection power (Figure 4.4).

4. Set Ladders: The user needs to click on the image to define the left and right boundary of the ladder lane. The software will report an error message if the user clicks outside the image region. Our default ladder set consists of eight markers from 0.8kb to 18.8 kb. The users are always free to switch to another ladder set as long as they enter ladder number and size accordingly. The software will then fit the ladders with a piecewise log transform regression

model and annotate band size by comparing the pixel position of detected bands to ladders (Figure 4.5).

5. Output: The software will give output of average telomere length, ratio of shortest telomeres, and the telomere length of the shortest 20%. Users need to click on the 'Calculate' button and then define the short telomere threshold. The default is 1.6kb. The software will also display the telomere length distribution histogram. To store the final output, a folder with the same name as the input image will be created in the same path. The individual band information will be saved in a mat file. The statistics (e.g. mean, ratio) will be saved in a txt file. The final output interface will be saved in a jpg file if the user clicks 'Save Figure' button. The preprocessed input image will be saved in a mat file as well (Figure 4.6).

4.2 3D image analysis pipeline manual

4.2.1. General Introduction

This pipeline is designed to automate 3D FISH image analysis for sparse nuclei. There are three major steps in this pipeline: 3D nuclei segmentation, fluorescent spots detection, and spots pairing. After image preprocessing, multi-scale automatic analysis is applied to DAPI channel to segment nuclei. With the mask of segmented nuclei, it applies sensitive spot detection algorithms (Aguet et al. 2013; Roudot et al. 2017) to each fluorescent channel. There are several strategies for the pipeline to autonomously validate the quality of nuclei segmentation and spots detection in each channel. Spots pairing across channels in the

following step calculates spatial distances between paired gene-subtelomeric region loci labeled with different colors.

4.2.2. System requirement

The 3D image analysis pipeline is programmed and compiled with Matlab 2019b (Mathworks) and its image processing toolboxes. Users can type commands in Matlab to implement each function block designed in the pipeline workflow. Large memory and parallel computing is required for image data processing.

4.2.3. Input image

This pipeline takes 3D image series acquired from Nikon Ti-Eclipse widefield microscope. Each file contains 40-80 3D stacks and each stack includes 10-20 fibroblast nuclei. Cells are sparsely distributed and cell density was well controlled in cell culture and fixation experiments.

4.2.4. Output

Results are saved to the output directory predefined by user, including mutual distance of each detected spot pair, nuclei segmentation mask, image metadata, relative positions of each nucleus in original 3D stack, and spot candidates used for pairing. The detailed step-by-step introduction and intermediate outputs are given in Chapter 3.2.6.

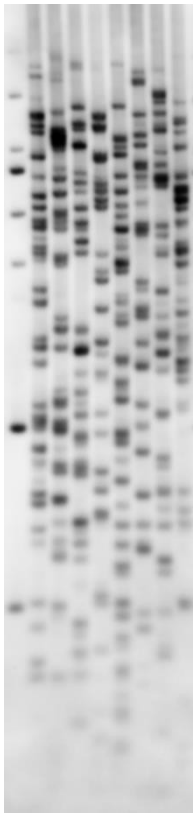


Figure 4.1 Software sample input image.

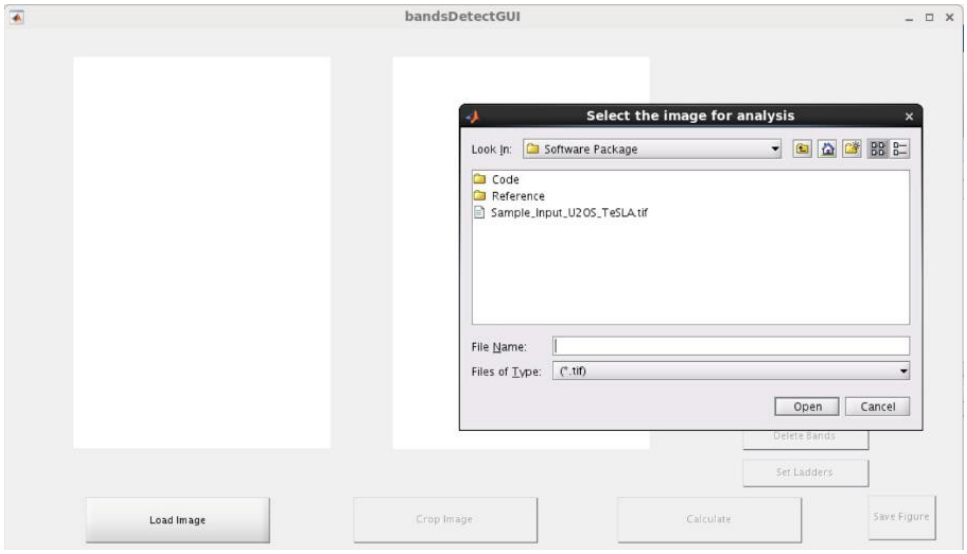


Figure 4.2 Software load image panel.

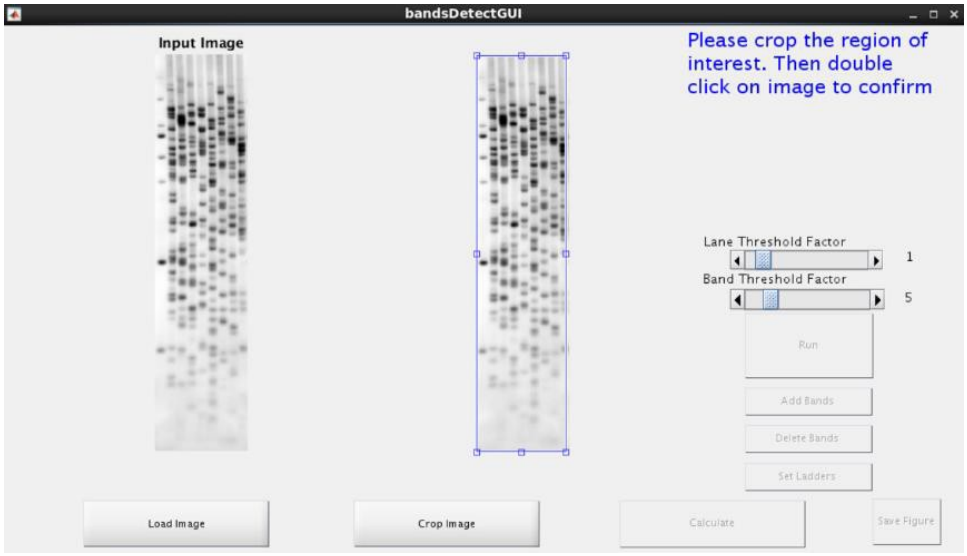
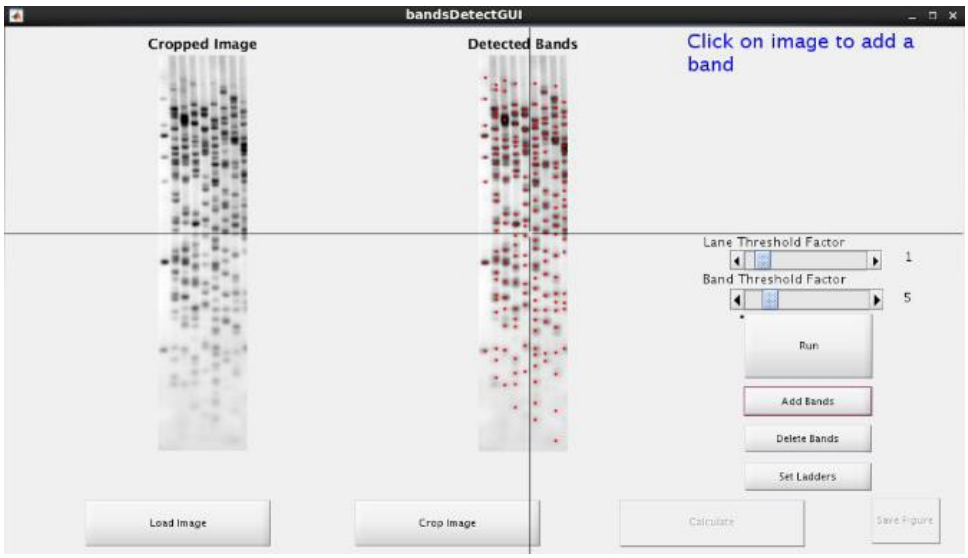


Figure 4.3 Software crop image panel.

a



b

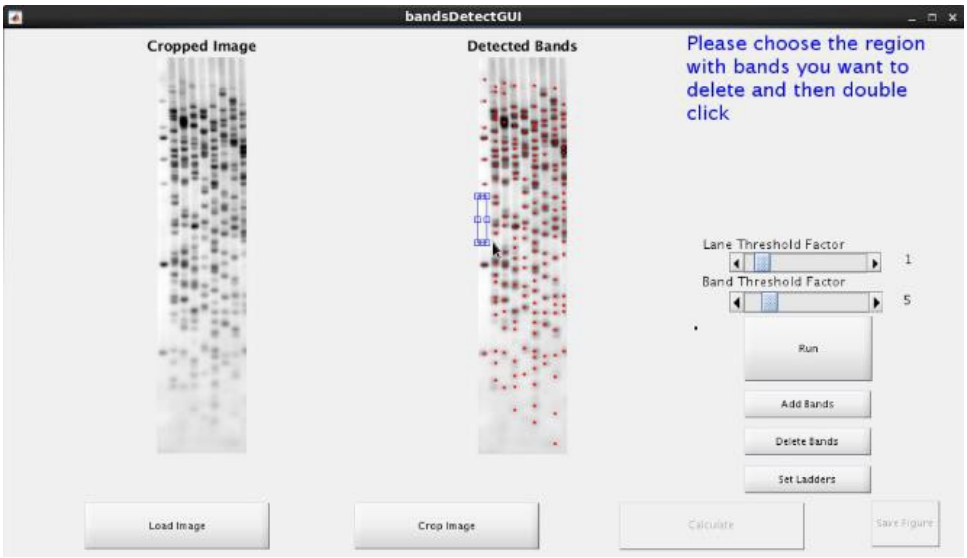


Figure 4.4 Bands detection result is manually adjustable by users. **(a)** Add a band **(b)** Delete bands within a region.

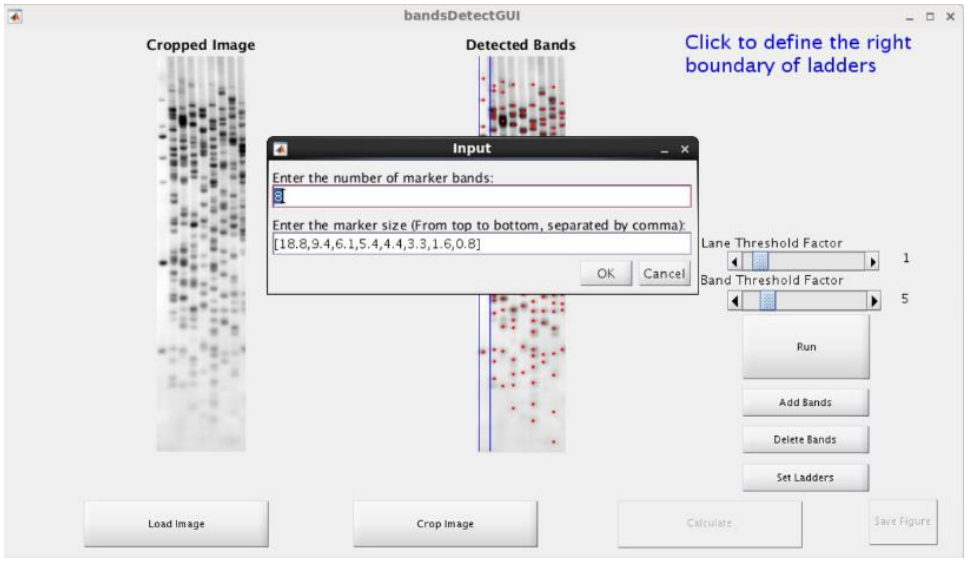


Figure 4.5 Software ladder setting panel.

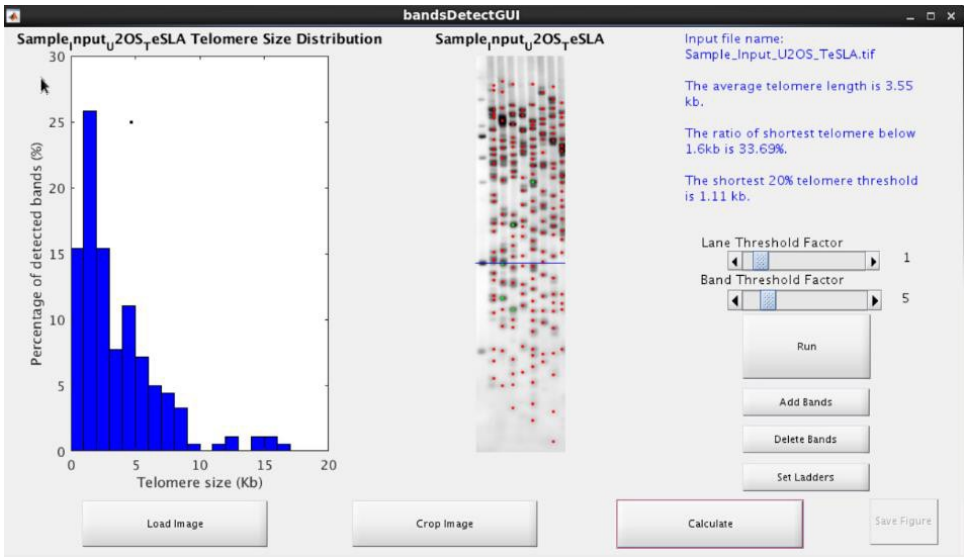


Figure 4.6 Software final output panel.

CHAPTER FIVE

CONCLUSIONS AND FUTURE PERSPECTIVES

5.1 Conclusions

Telomeres and telomerase have been intensively studied during the past many decades (Shay and Wright 2019). However, there are still some aspects that remain elusive. Among them, telomere length shortening associated effects and telomerase regulation with aging have been the biological targets of my doctoral research. Instead of intensively using molecular biology tools, I took an imaging approach to examine telomere dynamics and regulation of gene expression along with cell aging.

First, I developed an electrophoresis gel analysis software and reduced at least 50% of time on accurate telomere length quantification using TeSLA Southern blot images. With detailed telomere length information, I then applied 3D-FISH on cells with different population doublings (PDs) and developed another automated 3D image analysis pipeline to quantify telomere-gene interactions. That pipeline has achieved statistical analytics power of hundreds to thousands of cells. With that, I visualized how progressive telomere shortening alters the interaction patterns with certain genes and inferred a mechanism based on the potential binding sites of telomere associated shelterin proteins such as TRF2. Taken together, my work has improved the understanding of telomere regulatory roles on gene

expression and contributed new user-friendly tools to help researchers do a better, more accurate, and more time efficient analyses of their experimental telomere results.

5.2 Future perspectives

Digital droplet Telomere Shortest Length Assay (ddTeSLA)

TeSLA is a ligation and PCR based assay that can measure the telomere length distribution in a broader range (Lai et al. 2017b). It is the first method to measure the shortest telomere length distribution of all the telomeres in a population of cells but is limited by the low throughput of Southern blots (Lai et al. 2018). With the help of mature fluorescent labelling technology on single nucleotides, there could be an improved version of TeSLA that incorporates fluorescence labelled nucleotides into the PCR products. After ligation, instead of running PCR with regular dNTP (nucleoside triphosphate with deoxyribose), fluorophore labelled dUTP can be added to the system. Therefore, the fluorescent intensity of each amplified telomere product would be proportional to its length. By separating molecules to single oil droplets and measuring the intensity distribution of those droplets, telomere length can be quantified on molecular lever in a high-throughput manner. Artificial TTAGGG repetitive sequence with known length can be used as the control reference to accurately quantify telomere length.

Software enhancement using machine learning

The current software developed for Southern blot bands detection and analysis is based on a pre-written algorithm which quantifies local maxima using image pixel intensities. Although it offers users the freedom to manually add and drop bands, those higher-level decisions made by human experts do not contribute to a good improvement of the software future performance. Thus, this software can be optimized and becomes smarter by machine learning techniques. Reinforcement learning (RL) is a subclass of machine learning paradigms, alongside supervised learning and unsupervised learning (van Otterlo and Wiering 2012). It evaluates the environment feedback to software agent actions and gives either a reward or penalty to better train the software to improve its future performance. AlphaGo, the first computer program that defeated a professional human Go player, is a landmark achievement of the RL algorithm (Silver et al. 2016; Silver et al. 2017). With RL features added to this software, it may have a better performance based on users' feedbacks provided to prior band detection results.

Live cell imaging on telomere activity visualization

Clustered regularly interspaced short palindromic repeats (CRISPR) mediated live cell imaging on genomic loci has been widely studied in the last a few years (Chen et al. 2013; Qin et al. 2017; Wang et al. 2019; Wu et al. 2019). Researchers also revealed the dynamics of telomerase recruitment to telomeres using live cell imaging approach (Schmidt et al. 2016). It therefore provides the possibility to visualize telomere-gene interaction in live

cells along with progressive telomere shortening. In order to better classify telomere dynamics, probes need to be designed to distinguish telomere ends on different chromosomes. Also, the type of cell line and CRISPR off-target effect need to be evaluated in the experiment design.

The visualization of cell dynamics using imaging approaches has achieved long promising progress. In combination with the telomere study, it may reveal the dynamics of cellular aging relevant regulations and could lead to new approaches that target telomere or telomerase in diseases such as cancer.

BIBLIOGRAPHY

Aguet F, Antonescu CN, Mettlen M, Schmid SL, Danuser G (2013) Advances in analysis of low signal-to-noise images link dynamin and AP2 to the functions of an endocytic checkpoint. Developmental Cell 26:279-291 doi:10.1016/j.devcel.2013.06.019

Armanios M, Blackburn EH (2012) The telomere syndromes. Nature Reviews Genetics 13:693-704 doi:10.1038/nrg3246

Armanios MY et al. (2007) Telomerase mutations in families with idiopathic pulmonary fibrosis. New England Journal of Medicine 356:1317-1326 doi:10.1056/NEJMoa066157

Aubert G, Hills M, Lansdorp PM (2012) Telomere length measurement-Caveats and a critical assessment of the available technologies and tools. Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis 730:59-67 doi:10.1016/j.mrfmmm.2011.04.003

Baird DM, Rowson J, Wynford-Thomas D, Kipling D (2003) Extensive allelic variation and ultrashort telomeres in senescent human cells. Nature Genetics 33:203-207 doi:10.1038/ng1084

Baur JA, Zou Y, Shay JW, Wright WE (2001) Telomere position effect in human cells. Science 292:2075-2077 doi:10.1126/science.1062329

Bendix L, Horn PB, Jensen UB, Rubelj I, Kolvraa S (2010) The load of short telomeres, estimated by a new method, Universal STELA, correlates with number of senescent cells. Aging Cell 9:383-397 doi:10.1111/j.1474-9726.2010.00568.x

Bessman MJ, Lehman IR, Adler J, Zimmerman SB, Simms ES, Kornberg A (1958) Enzymatic Synthesis of Deoxyribonucleic Acid. Iii. The Incorporation of Pyrimidine and Purine Analogues into Deoxyribonucleic Acid. Proceedings of the National Academy of Sciences of the United States of America 44:633-640 doi:10.1073/pnas.44.7.633

Blackburn EH, Gall JG (1978) A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in Tetrahymena. Journal of Molecular Biology 120:33-53 doi:10.1016/0022-2836(78)90294-2

Blasco MA, Lee HW, Hande MP, Samper E, Lansdorp PM, DePinho RA, Greider CW (1997) Telomere shortening and tumor formation by mouse cells lacking telomerase RNA. Cell 91:25-34 doi:10.1016/s0092-8674(01)80006-4

Bodnar AG et al. (1998) Extension of life-span by introduction of telomerase into normal human cells. Science 279:349-352 doi:10.1126/science.279.5349.349

Bourgeron T, Xu Z, Doumic M, Teixeira MT (2015) The asymmetry of telomere replication contributes to replicative senescence heterogeneity. Scientific Reports 5:15326 doi:10.1038/srep15326

Bryan TM, Englezou A, DallaPozza L, Dunham MA, Reddel RR (1997) Evidence for an alternative mechanism for maintaining telomere length in human tumors and tumor-derived cell lines. Nature Medicine 3:1271-1274 doi:10.1038/nm1197-1271

Burnet FMS (1974) Intrinsic mutagenesis, a genetic approach to ageing. New York J. Wiley,

Campisi J (2013) Aging, Cellular Senescence, and Cancer. In: Julius D (ed) Annual Review of Physiology, vol 75. Annual Review of Physiology. pp 685-705. doi:10.1146/annurev-physiol-030212-183653

Canela A, Vera E, Klatt P, Blasco MA (2007) High-throughput telomere length quantification by FISH and its application to human population studies. Proceedings of the National Academy of Sciences of the United States of America 104:5300-5305 doi:10.1073/pnas.0609367104

Carrel A, Ebeling AH (1921) Age and Multiplication of Fibroblasts. Journal of Experimental Medicine 34:599-623 doi:10.1084/jem.34.6.599

Castle JC et al. (2010) DNA copy number, including telomeres and mitochondria, assayed using next-generation sequencing. BMC Genomics 11 doi:10.1186/1471-2164-11-244

Cawthon RM (2002) Telomere measurement by quantitative PCR. Nucleic Acids Research 30 doi:10.1093/nar/30.10.e47

Cawthon RM (2009) Telomere length measurement by a novel monochrome multiplex quantitative PCR method. Nucleic Acids Research 37:e21 doi:10.1093/nar/gkn1027

Cesare AJ, Reddel RR (2010) Alternative lengthening of telomeres: models, mechanisms and implications. Nature Reviews: Genetics 11:319-330 doi:10.1038/nrg2763

Chen B et al. (2013) Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. Cell 155:1479-1491 doi:10.1016/j.cell.2013.12.001

Cong Y, Shay JW (2008) Actions of human telomerase beyond telomeres. Cell Research 18:725-732 doi:10.1038/cr.2008.74

Cronkhite JT, Xing C, Raghu G, Chin KM, Torres F, Rosenblatt RL, Garcia CK (2008) Telomere shortening in familial and sporadic pulmonary fibrosis. American Journal of Respiratory and Critical Care Medicine 178:729-737 doi:10.1164/rccm.200804-550OC

de Lange T (2005) Shelterin: the protein complex that shapes and safeguards human telomeres. Genes & Development 19:2100-2110 doi:10.1101/gad.1346005

de Leon AD et al. (2010) Telomere Lengths, Pulmonary Fibrosis and Telomerase (TERT) Mutations. PloS One 5 doi:10.1371/journal.pone.0010680

de Wit E, de Laat W (2012) A decade of 3C technologies: insights into nuclear organization. Genes and Development 26:11-24 doi:10.1101/gad.179804.111

Denker A, de Laat W (2016) The second decade of 3C technologies: detailed insights into nuclear organization. Genes and Development 30:1357-1382 doi:10.1101/gad.281964.116

DePinho RA, Polyak K (2004) Cancer chromosomes in crisis. Nature Genetics 36:932-934 doi:10.1038/ng0904-932

Ding ZH, Mangino M, Aviv A, Spector T, Durbin R, Consortium UK (2014) Estimating telomere length from whole genome sequence data. Nucleic Acids Research 42 doi:10.1093/nar/gku181

Druliner BR et al. (2016) Time Lapse to Colorectal Cancer: Telomere Dynamics Define the Malignant Potential of Polyps. Clinical and Translational Gastroenterology 7 doi:10.1038/ctg.2016.48

Epel ES, Blackburn EH, Lin J, Dhabhar FS, Adler NE, Morrow JD, Cawthon RM (2004) Accelerated telomere shortening in response to life stress. Proceedings of the National Academy of Sciences of the United States of America 101:17312-17315 doi:10.1073/pnas.0407162101

Erdmann N, Liu Y, Harrington L (2004) Distinct dosage requirements for the maintenance of long and short telomeres in mTert heterozygous mice. Proceedings of the National Academy of Sciences of the United States of America 101:6080-6085 doi:10.1073/pnas.0401580101

Feng JL et al. (1995) The RNA Component of Human Telomerase. Science 269:1236-1241 doi:10.1126/science.7544491

Fitzpatrick AL et al. (2007) Leukocyte telomere length and cardiovascular disease in the cardiovascular health study. American Journal of Epidemiology 165:14-21 doi:10.1093/aje/kw346

Frink RE, Peyton M, Schiller JH, Gazdar AF, Shay JW, Minna JD (2016) Telomerase inhibitor imetelstat has preclinical activity across the spectrum of non-small cell lung cancer oncogenotypes in a telomere length dependent manner. Oncotarget 7:31639-31651 doi:10.18632/oncotarget.9335

Fumagalli M et al. (2012) Telomeric DNA damage is irreparable and causes persistent DNA-damage-response activation. Nature Cell Biology 14:355-+ doi:10.1038/ncb2466

Gasser SM, Hediger F, Taddei A, Neumann FR, Gartenberg MR (2004) The function of telomere clustering in yeast: The circe effect. Cold Spring Harbor Symposia on Quantitative Biology 69:327-337 doi:10.1101/sqb.2004.69.327

Gomes NM et al. (2011) Comparative biology of mammalian telomeres: hypotheses on ancestral states and the roles of telomeres in longevity determination. Aging Cell 10:761-768 doi:10.1111/j.1474-9726.2011.00718.x

Gottschling DE, Aparicio OM, Billington BL, Zakian VA (1990) Position effect at S. cerevisiae telomeres: reversible repression of Pol II transcription. Cell 63:751-762 doi:10.1016/0092-8674(90)90141-z

Greider CW (1991) Telomeres. Current Opinion in Cell Biology 3:444-451 doi:10.1016/0955-0674(91)90072-7

Greider CW, Blackburn EH (1985) Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. Cell 43:405-413 doi:10.1016/0092-8674(85)90170-9

Greider CW, Blackburn EH (1987) The telomere terminal transferase of Tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity. Cell 51:887-898 doi:10.1016/0092-8674(87)90576-9

Greider CW, Blackburn EH (1989) A telomeric sequence in the RNA of Tetrahymena telomerase required for telomere repeat synthesis. Nature 337:331-337 doi:10.1038/337331a0

Griffith JD, Comeau L, Rosenfield S, Stansel RM, Bianchi A, Moss H, de Lange T (1999) Mammalian telomeres end in a large duplex loop. Cell 97:503-514 doi:10.1016/s0092-8674(00)80760-6

Hanahan D, Weinberg RA (2011) Hallmarks of cancer: the next generation. Cell 144:646-674 doi:10.1016/j.cell.2011.02.013

Harley CB, Futcher AB, Greider CW (1990) Telomeres shorten during ageing of human fibroblasts. Nature 345:458-460 doi:10.1038/345458a0

Hayflick L (1965) The Limited in Vitro Lifetime of Human Diploid Cell Strains. Experimental Cell Research 37:614-636 doi:10.1016/0014-4827(65)90211-9

Hayflick L, Moorhead PS (1961) The serial cultivation of human diploid cell strains. Experimental Cell Research 25:585-621 doi:10.1016/0014-4827(61)90192-6

Hemann MT, Strong MA, Hao LY, Greider CW (2001) The shortest telomere, not average telomere length, is critical for cell viability and chromosome stability. Cell 107:67-77 doi:10.1016/s0092-8674(01)00504-9

Henson JD, Neumann AA, Yeager TR, Reddel RR (2002) Alternative lengthening of telomeres in mammalian cells. Oncogene 21:598-610 doi:10.1038/sj.onc.1205058

Herbert BS et al. (2005) Lipid modification of GRN163, an N3 ' -> P5 ' thio-phosphoramidate oligonucleotide, enhances the potency of telomerase inhibition. Oncogene 24:5262-5268 doi:10.1038/sj.onc.1208760

Herbert BS, Pitts AE, Baker SI, Hamilton SE, Wright WE, Shay JW, Corey DR (1999) Inhibition of human telomerase in immortal human cells leads to progressive telomere shortening and cell death. Proceedings of the National Academy of Sciences of the United States of America 96:14276-14281 doi:10.1073/pnas.96.25.14276

Herbig U, Jobling WA, Chen BPC, Chen DJ, Sedivy JM (2004) Telomere shortening triggers senescence of human cells through a pathway involving ATM, p53, and p21(CIP1), but not p16(INK4a). Molecular Cell 14:501-513 doi:10.1016/s1097-2765(04)00256-4

Hockemeyer D, Sfeir AJ, Shay JW, Wright WE, de Lange T (2005) POT1 protects telomeres from a transient DNA damage response and determines how human chromosomes end. EMBO Journal 24:2667-2678 doi:10.1038/sj.emboj.7600733

Holland AJ, Cleveland DW (2009) Boveri revisited: chromosomal instability, aneuploidy and tumorigenesis. Nature Reviews Molecular Cell Biology 10:478-487 doi:10.1038/nrm2718

Holohan B, Wright WE, Shay JW (2014) Telomeropathies: An emerging spectrum disorder. Journal of Cell Biology 205:289-299 doi:10.1083/jcb.201401012

Jafri MA, Ansari SA, Alqahtani MH, Shay JW (2016) Roles of telomeres and telomerase in cancer, and advances in telomerase-targeted therapies. Genome Medicine 8:69 doi:10.1186/s13073-016-0324-x

Jaqaman K, Loerke D, Mettlen M, Kuwata H, Grinstein S, Schmid SL, Danuser G (2008) Robust single-particle tracking in live-cell time-lapse sequences. Nature Methods 5:695-702 doi:10.1038/nmeth.1237

Jonker R, Volgenant A (1987) A Shortest Augmenting Path Algorithm for Dense and Sparse Linear Assignment Problems. Computing 38:325-340 doi:Doi 10.1007/Bf02278710

Kang HJ et al. (2004) Ectopic expression of the catalytic subunit of telomerase protects against brain injury resulting from ischemia and NMDA-induced neurotoxicity. Journal of Neuroscience 24:1280-1287 doi:10.1523/JNEUROSCI.4082-03.2004

Keane M et al. (2015) Insights into the Evolution of Longevity from the Bowhead Whale Genome. Cell Reports 10:112-122 doi:10.1016/j.celrep.2014.12.008

Kim NW et al. (1994) Specific association of human telomerase activity with immortal cells and cancer. Science 266:2011-2015 doi:10.1126/science.7605428

Kim SH, Kaminker P, Campisi J (1999) TIN2, a new regulator of telomere length in human cells. Nature Genetics 23:405-412

Kim W et al. (2016) Regulation of the Human Telomerase Gene TERT by Telomere Position Effect-Over Long Distances (TPE-OLD): Implications for Aging and Cancer. PLoS Biology 14:e2000016 doi:10.1371/journal.pbio.2000016

Kim W, Shay JW (2018) Long-range telomere regulation of gene expression: Telomere looping and telomere position effect over long distances (TPE-OLD). Differentiation 99:1-9 doi:10.1016/j.diff.2017.11.005

Kimura M et al. (2010) Measurement of telomere length by the Southern blot analysis of terminal restriction fragment lengths. Nature Protocols 5:1596-1607 doi:10.1038/nprot.2010.124

Kuhn HW (1955) The Hungarian Method for the Assignment Problem. Naval Research Logistics Quarterly 2:83--97 doi:10.1002/nav.3800020109

Lai TP, Wright WE, Shay JW (2016) Generation of digoxigenin-incorporated probes to enhance DNA detection sensitivity. Biotechniques 60:306-309 doi:10.2144/000114427

Lai TP, Wright WE, Shay JW (2018) Comparison of telomere length measurement methods. Philosophical Transactions of the Royal Society B-Biological Sciences 373 doi:10.1098/rstb.2016.0451

Lai TP et al. (2017a) A method for measuring the distribution of the shortest telomeres in cells and tissues. Nature Communications 8:1356 doi:10.1038/s41467-017-01291-z

Lai TP et al. (2017b) A method for measuring the distribution of the shortest telomeres in cells and tissues. Nature Communications 8:1356 doi:10.1038/s41467-017-01291-z

Lansdorp PM et al. (1996) Heterogeneity in telomere length of human chromosomes. Human Molecular Genetics 5:685-691 doi:10.1093/hmg/5.5.685

Lavrentieva I, Broude NE, Lebedev Y, Gottesman, II, Lukyanov SA, Smith CL, Sverdlov ED (1999) High polymorphism level of genomic sequences flanking insertion sites of human endogenous retroviral long terminal repeats. FEBS Letters 443:341-347 doi:10.1016/s0014-5793(99)00004-6

Lee M, Napier CE, Yang SF, Arthur JW, Reddel RR, Pickett HA (2017) Comparative analysis of whole genome sequencing-based telomere length measurement techniques. Methods 114:4-15 doi:10.1016/j.ymeth.2016.08.008

Lejnine S, Makarov VL, Langmore JP (1995) Conserved Nucleoprotein Structure at the Ends of Vertebrate and Invertebrate Chromosomes. Proceedings of the National Academy of Sciences of the United States of America 92:2393-2397 doi:10.1073/pnas.92.6.2393

Lewis KA, Wuttke DS (2012) Telomerase and telomere-associated proteins: structural insights into mechanism and evolution. Structure 20:28-39 doi:10.1016/j.str.2011.10.017

Lieberman-Aiden E et al. (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326:289-293 doi:10.1126/science.1181369

Lin J et al. (2016) Systematic and Cell Type-Specific Telomere Length Changes in Subsets of Lymphocytes. Journal of Immunology Research doi:10.1155/2016/5371050

Lin Y et al. (2015) Age-associated telomere attrition of lymphocytes in vivo is co-ordinated with changes in telomerase activity, composition of lymphocyte subsets and health conditions. Clinical Science 128:367-377 doi:10.1042/cs20140481

Lou Z, Wei J, Riethman H, Baur JA, Voglauer R, Shay JW, Wright WE (2009) Telomere length regulates ISG15 expression in human cells. Aging 1:608-621

Maida Y et al. (2009) An RNA-dependent RNA polymerase formed by TERT and the RMRP RNA. Nature 461:230-235 doi:10.1038/nature08283

Massey FJ (1951) The Kolmogorov-Smirnov Test for Goodness of Fit. Journal of the American Statistical Association 46:68-78 doi:10.1080/01621459.1951.10500769

Masutomi K et al. (2005) The telomerase reverse transcriptase regulates chromatin state and DNA damage responses. Proceedings of the National Academy of Sciences of the United States of America 102:8222-8227 doi:10.1073/pnas.0503095102

Mathon NF, Lloyd AC (2001) Cell senescence and cancer. Nature Reviews: Cancer 1:203-213 doi:10.1038/35106045

McClintock B (1941) The Stability of Broken Ends of Chromosomes in Zea Mays. Genetics 26:234-282

Mender I, Gryaznov S, Dikmen ZG, Wright WE, Shay JW (2015) Induction of Telomere Dysfunction Mediated by the Telomerase Substrate Precursor 6-Thio-2'-Deoxyguanosine. Cancer Discovery 5:82-95 doi:10.1158/2159-8290.Cd-14-0609

Mender I et al. (2018) Telomerase-Mediated Strategy for Overcoming Non-Small Cell Lung Cancer Targeted Therapy and Chemotherapy Resistance. Neoplasia 20:826-837 doi:10.1016/j.neo.2018.06.002

Mender I, Shay JW (2015a) Telomere Dysfunction Induced Foci (TIF) Analysis. Bio Protocol 5 doi:10.21769/bioprotoc.1656

Mender I, Shay JW (2015b) Telomere Restriction Fragment (TRF) Analysis. Bio Protocol 5 doi:10.21769/bioprotoc.1658

Meyerson M et al. (1997) hEST2, the putative human telomerase catalytic subunit gene, is up-regulated in tumor cells and during immortalization. Cell 90:785-795 doi:10.1016/s0092-8674(00)80538-3

Meyne J et al. (1990) Distribution of non-telomeric sites of the (TTAGGG)n telomeric sequence in vertebrate chromosomes. Chromosoma 99:3-10 doi:10.1007/bf01737283

Meyne J, Ratliff RL, Moyzis RK (1989) Conservation of the human telomere sequence (TTAGGG)n among vertebrates. Proceedings of the National Academy of Sciences of the United States of America 86:7049-7053 doi:10.1073/pnas.86.18.7049

Meznikova M, Erdmann N, Allsopp R, Harrington LA (2009) Telomerase reverse transcriptase-dependent telomere equilibration mitigates tissue dysfunction in mTert heterozygotes. Disease Models & Mechanisms 2:620-626 doi:10.1242/dmm.004069

Min J, Wright WE, Shay JW (2017a) Alternative lengthening of telomeres can be maintained by preferential elongation of lagging strands. Nucleic Acids Research 45:2615-2628 doi:10.1093/nar/gkw1295

Min J, Wright WE, Shay JW (2017b) Alternative Lengthening of Telomeres Mediated by Mitotic DNA Synthesis Engages Break-Induced Replication Processes. Molecular and Cellular Biology 37 doi:10.1128/mcb.00226-17

Montpetit AJ et al. (2014) Telomere Length A Review of Methods for Measurement. Nursing Research 63:289-299 doi:10.1097/nnr.0000000000000037

Morin GB (1989) The human telomere terminal transferase enzyme is a ribonucleoprotein that synthesizes TTAGGG repeats. Cell 59:521-529 doi:10.1016/0092-8674(89)90035-4

Moyzis RK et al. (1988) A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes. Proceedings of the National

Academy of Sciences of the United States of America 85:6622-6626 doi:10.1073/pnas.85.18.6622

Mukherjee AK et al. (2018) Telomere length-dependent transcription and epigenetic modifications in promoters remote from telomere ends. Plos Genetics 14:e1007782 doi:10.1371/journal.pgen.1007782

Muller HJ (1938) The remaking of chromosomes. Collecting Net 8:182–195

Nakamura TM et al. (1997) Telomerase catalytic subunit homologs from fission yeast and human. Science 277:955-959 doi:10.1126/science.277.5328.955

Nersisyan L, Arakelyan A (2015) Computel: Computation of Mean Telomere Length from Whole-Genome Next-Generation Sequencing Data. PloS One 10 doi:10.1371/journal.pone.0125201

Nussey DH et al. (2014) Measuring telomere length and telomere dynamics in evolutionary biology and ecology. Methods in Ecology and Evolution 5:299-310 doi:10.1111/2041-210x.12161

O'Connor MS, Safari A, Xin HW, Liu D, Songyang Z (2006) A critical role for TPP1 and TIN2 interaction in high-order telomeric complex assembly. Proceedings of the National Academy of Sciences of the United States of America 103:11874-11879 doi:10.1073/pnas.0605303103

O'Hara R, Tedone E, Ludlow A, Huang E, Arosio B, Mari D, Shay JW (2019) Quantitative mitochondrial DNA copy number determination using droplet digital PCR with single-cell resolution. Genome Research 29:1878-1888 doi:10.1101/gr.250480.119

O'Sullivan RJ, Karlseder J (2010a) Telomeres: protecting chromosomes against genome instability. Nature Reviews Molecular Cell Biology 11:171-181 doi:10.1038/nrm2848

O'Sullivan RJ, Karlseder J (2010b) Telomeres: protecting chromosomes against genome instability. Nature Reviews: Molecular Cell Biology 11:171-181 doi:10.1038/nrm2848

Okazaki R, Okazaki T, Sakabe K, Sugimoto K, Sugino A (1968) Mechanism of DNA chain growth. I. Possible discontinuity and unusual secondary structure of newly synthesized chains. Proceedings of the National Academy of Sciences of the United States of America 59:598-605 doi:10.1073/pnas.59.2.598

Olovnikov AM (1973) A theory of marginotomy. The incomplete copying of template margin in enzymic synthesis of polynucleotides and biological significance of the phenomenon. Journal of Theoretical Biology 41:181-190 doi:10.1016/0022-5193(73)90198-7

Opresko PL, Shay JW (2017) Telomere-associated aging disorders. Ageing Research Reviews 33:52-66 doi:10.1016/j.arr.2016.05.009

Palm W, de Lange T (2008) How Shelterin Protects Mammalian Telomeres. Annual Review of Genetics 42:301-334 doi:10.1146/annurev.genet.41.110306.130350

Park JI et al. (2009) Telomerase modulates Wnt signalling by association with target gene chromatin. Nature 460:66-72 doi:10.1038/nature08137

Pawelec G, Akbar A, Caruso C, Solana R, Grubeck-Loebenstein B, Wikby A (2005) Human immunosenescence: is it infectious? Immunological Reviews 205:257-268 doi:10.1111/j.0105-2896.2005.00271.x

Poon SSS, Lansdorp PM (2001) Quantitative fluorescence in situ hybridization (Q-FISH). Current Protocol in Cell Biology Chapter 18:18 14 11-18 14 21 doi:10.1002/0471143030.cb1804s12

Qin P et al. (2017) Live cell imaging of low- and non-repetitive chromosome loci using CRISPR-Cas9. Nature Communications 8:14725 doi:10.1038/ncomms14725

Ramirez MJ, Surralles J (2008) Laser confocal microscopy analysis of human interphase nuclei by three-dimensional FISH reveals dynamic perinucleolar clustering of telomeres. Cytogenetic and Genome Research 122:237-242 doi:10.1159/000167809

Reichert S, Stier A (2017) Does oxidative stress shorten telomeres in vivo? A review. Biology Letters 13 doi:10.1098/rsbl.2017.0463

Robin JD et al. (2015) SORBS2 transcription is activated by telomere position effect-over long distance upon telomere shortening in muscle cells from patients with facioscapulohumeral dystrophy. Genome Research 25:1781-1790 doi:10.1101/gr.190660.115

Robin JD et al. (2014) Telomere position effect: regulation of gene expression with progressive telomere shortening over long distances. Genes and Development 28:2464-2476 doi:10.1101/gad.251041.114

Roudot P, Liya D, Jaqaman K, Kervrann C, Danuser G (2017) Piecewise-Stationary Motion Modeling and Iterative Smoothing to Track Heterogeneous Particle Motions in Dense Environments. IEEE Transactions on Image Processing 26:5395-5410 doi:10.1109/TIP.2017.2707803

Rubin H (1998) Telomerase and cellular lifespan: ending the debate? Nature Biotechnology 16:396-397 doi:10.1038/nbt0598-396

Rufer N, Dragowska W, Thornbury G, Roosnek E, Lansdorp PM (1998) Telomere length dynamics in human lymphocyte subpopulations measured by flow cytometry. Nature Biotechnology 16:743-747 doi:10.1038/nbt0898-743

Ruiz-Herrera A, Nergadze SG, Santagostino M, Giulotto E (2008) Telomeric repeats far from the ends: mechanisms of origin and role in evolution. Cytogenetic and Genome Research 122:219-228 doi:10.1159/000167807

Salvador L, Singaravelu G, Harley CB, Flom P, Suram A, Raffaele JM (2016) A Natural Product Telomerase Activator Lengthens Telomeres in Humans: A Randomized, Double Blind, and Placebo Controlled Study. Rejuvenation Research 19:478-+ doi:10.1089/rej.2015.1793

Sampson MJ, Winterbone MS, Hughes JC, Dozio N, Hughes DA (2006) Monocyte telomere shortening and oxidative DNA damage in type 2 diabetes. Diabetes Care 29:283-289 doi:10.2337/diacare.29.02.06.dc05-1715

Sandell LL, Zakian VA (1992) Telomeric position effect in yeast. Trends in Cell Biology 2:10-14 doi:10.1016/0962-8924(92)90138-d

Schmidt JC, Zaug AJ, Cech TR (2016) Live Cell Imaging Reveals the Dynamics of Telomerase Recruitment to Telomeres. Cell 166:1188-1197 e1189 doi:10.1016/j.cell.2016.07.033

Sfeir A, de Lange T (2012) Removal of shelterin reveals the telomere end-protection problem. Science 336:593-597 doi:10.1126/science.1218498

Sfeir AJ, Chai WH, Shay JW, Wright WE (2005) Telomere-end processing: The terminal nucleotides of human chromosomes. Molecular Cell 18:131-138 doi:10.1016/j.molcel.2005.02.035

Shay JW (2014) Are Short Telomeres Hallmarks of Cancer Recurrence? Clinical Cancer Research 20:779-781 doi:Doi 10.1158/1078-0432.Ccr-13-3198

Shay JW (2016) Role of Telomeres and Telomerase in Aging and Cancer. Cancer Discovery 6:584-593 doi:10.1158/2159-8290.Cd-16-0062

Shay JW, Bacchetti S (1997) A survey of telomerase activity in human cancer. European Journal of Cancer 33:787-791 doi:10.1016/s0959-8049(97)00062-2

Shay JW, Wright WE (1989) Quantitation of the frequency of immortalization of normal human diploid fibroblasts by SV40 large T-antigen. Experimental Cell Research 184:109-118 doi:10.1016/0014-4827(89)90369-8

Shay JW, Wright WE (2000) Hayflick, his limit, and cellular ageing. Nature Reviews Molecular Cell Biology 1:72-76 doi:10.1038/35036093

Shay JW, Wright WE (2005) Mechanism-based combination telomerase inhibition therapy. Cancer Cell 7:1-2 doi:10.1016/j.ccr.2004.12.012

Shay JW, Wright WE (2006) Telomerase therapeutics for cancer: challenges and new directions. Nature Reviews Drug Discovery 5:577-584 doi:10.1038/nrd2081

Shay JW, Wright WE (2019) Telomeres and telomerase: three decades of progress. Nature Reviews Genetics 20:299-309 doi:10.1038/s41576-019-0099-1

Shippenlentz D, Blackburn EH (1989) Telomere Terminal Transferase Activity From Euplotes Crassus Adds Large Numbers of TTTTGGGG Repeats Onto Telomeric Primers. Molecular and Cellular Biology 9:2761-2764 doi:10.1128/mcb.9.6.2761

Silver D et al. (2016) Mastering the game of Go with deep neural networks and tree search. Nature 529:484-489 doi:10.1038/nature16961

Silver D et al. (2017) Mastering the game of Go without human knowledge. Nature 550:354-359 doi:10.1038/nature24270

Simonet T et al. (2011) The human TTAGGG repeat factors 1 and 2 bind to a subset of interstitial telomeric sequences and satellite repeats. Cell Research 21:1028-1038 doi:10.1038/cr.2011.40

Stadler G et al. (2013) Telomere position effect regulates DUX4 in human facioscapulohumeral muscular dystrophy. Nature Structural & Molecular Biology 20:671-678 doi:10.1038/nsmb.2571

Stavenhagen JB, Zakian VA (1998) Yeast telomeres exert a position effect on recombination between internal tracts of yeast telomeric DNA. Genes and Development 12:3044-3058 doi:10.1101/gad.12.19.3044

Steinert S, Shay JW, Wright WE (2004) Modification of subtelomeric DNA. Molecular and Cellular Biology 24:4571-4580 doi:10.1128/mcb.24.10.4571-4580.2004

Stuart BD et al. (2015) Exome sequencing links mutations in PARN and RTEL1 with familial pulmonary fibrosis and telomere shortening. Nature Genetics 47:512-U114 doi:10.1038/ng.3278

Sugimoto K, Okazaki T, Okazaki R (1968) Mechanism of DNA chain growth, II. Accumulation of newly synthesized short chains in E. coli infected with ligase-defective T4 phages. Proceedings of the National Academy of Sciences of the United States of America 60:1356-1362 doi:10.1073/pnas.60.4.1356

Szostak JW, Blackburn EH (1982) Cloning yeast telomeres on linear plasmid vectors. Cell 29:245-255 doi:10.1016/0092-8674(82)90109-x

Theodoris CV, Mourkioti F, Huang Y, Ranade SS, Liu L, Blau HM, Srivastava D (2017) Long telomeres protect against age-dependent cardiac disease caused by NOTCH1 haploinsufficiency. Journal of Clinical Investigation 127:1683-1688 doi:10.1172/jci90338

Thorvaldsdottir H, Robinson JT, Mesirov JP (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Briefings in Bioinformatics 14:178-192 doi:10.1093/bib/bbs017

Tsakiri KD et al. (2007) Adult-onset pulmonary fibrosis caused by mutations in telomerase. Proceedings of the National Academy of Sciences of the United States of America 104:7552-7557 doi:10.1073/pnas.0701009104

Vallejo AN, Weyand CM, Goronzy JJ (2004) T-cell senescence: a culprit of immune abnormalities in chronic inflammation and persistent infection. Trends in Molecular Medicine 10:119-124 doi:10.1016/j.molmed.2004.01.002

van Otterlo M, Wiering M (2012) Reinforcement Learning and Markov Decision Processes. In: Wiering M, van Otterlo M (eds) Reinforcement Learning: State-of-the-Art. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 3-42. doi:10.1007/978-3-642-27645-3_1

Vera E, Blasco MA (2012) Beyond average: potential for measurement of short telomeres. Aging-Us 4:379-392 doi:10.18632/aging.100462

Vera E, de Jesus BB, Foronda M, Flores JM, Blasco MA (2012) The Rate of Increase of Short Telomeres Predicts Longevity in Mammals. Cell Reports 2:732-737 doi:10.1016/j.celrep.2012.08.023

von Zglinicki T (2002) Oxidative stress shortens telomeres. Trends in Biochemical Sciences 27:339-344 doi:10.1016/s0968-0004(02)02110-2

von Zglinicki T, Saretzki G, Ladhoff J, di Fagagna FD, Jackson SP (2005) Human cell senescence as a DNA damage response. Mechanisms of Ageing and Development 126:111-117 doi:10.1016/j.mad.2004.09.034

Wang H et al. (2019) CRISPR-mediated live imaging of genome editing and transcription. Science 365:1301-1305 doi:10.1126/science.aax7852

Watson JD (1972) Origin of concatemeric T7 DNA. Nature New biology 239:197-201 doi:10.1038/newbio239197a0

Watson JD, Crick FH (1953) The structure of DNA. Cold Spring Harbor Symposia on Quantitative Biology 18:123-131 doi:10.1101/sqb.1953.018.01.020

Webb CJ, Wu Y, Zakian VA (2013) DNA repair at telomeres: keeping the ends intact. Cold Spring Harbor Perspectives in Biology 5 doi:10.1101/cshperspect.a012666

Weng NP, Akbar AN, Goronzy J (2009) CD28(-) T cells: their role in the age-associated decline of immune function. Trends in Immunology 30:306-312 doi:10.1016/j.it.2009.03.013

Wiemann SU et al. (2002) Hepatocyte telomere shortening and senescence are general markers of human liver cirrhosis. FASEB Journal 16:935-942 doi:10.1096/fj.01-0977com

Wieser M, Stadler G, Bohm E, Borth N, Katinger H, Grillari J, Voglauer R (2006) Nuclear Flow FISH: Isolation of cell nuclei improves the determination of telomere lengths. Experimental Gerontology 41:230-235 doi:10.1016/j.exger.2005.09.013

Wood AM, Laster K, Rice EL, Kosak ST (2015) A beginning of the end: new insights into the functional organization of telomeres. Nucleus 6:172-178 doi:10.1080/19491034.2015.1048407

Wood AM et al. (2014) TRF2 and lamin A/C interact to facilitate the functional organization of chromosome ends. Nature Communications 5:5467 doi:10.1038/ncomms6467

Wright WE, Pereira-Smith OM, Shay JW (1989) Reversible cellular senescence: implications for immortalization of normal human diploid fibroblasts. Molecular and Cellular Biology 9:3088-3092 doi:10.1128/mcb.9.7.3088

Wright WE, Piatyszek MA, Rainey WE, Byrd W, Shay JW (1996) Telomerase activity in human germline and embryonic tissues and cells. Developmental Genetics 18:173-179 doi:10.1002/(SICI)1520-6408(1996)18:2<173::AID-DVG10>3.0.CO;2-3

Wright WE, Shay JW (1992a) Telomere positional effects and the regulation of cellular senescence. Trends in Genetics 8:193-197 doi:10.1016/0168-9525(92)90232-s

Wright WE, Shay JW (1992b) The two-stage mechanism controlling cellular senescence and immortalization. Experimental Gerontology 27:383-389 doi:10.1016/0531-5565(92)90069-c

Wu X, Mao S, Ying Y, Krueger CJ, Chen AK (2019) Progress and Challenges for Live-cell Imaging of Genomic Loci Using CRISPR-based Platforms. Genomics, Proteomics & Bioinformatics 17:119-128 doi:10.1016/j.gpb.2018.10.001

Yu GL, Bradley JD, Attardi LD, Blackburn EH (1990) In vivo alteration of telomere sequences and senescence caused by mutated Tetrahymena telomerase RNAs. Nature 344:126-132 doi:10.1038/344126a0

Zahler AM, Prescott DM (1988) Telomere terminal transferase activity in the hypotrichous ciliate Oxytricha nova and a model for replication of the ends of linear DNA molecules. Nucleic Acids Research 16:6953-6972 doi:10.1093/nar/16.14.6953

Zakian VA (2009) The ends have arrived. Cell 139:1038-1040 doi:10.1016/j.cell.2009.11.033

Zou Y, Sfeir A, Gryaznov SM, Shay JW, Wright WE (2004) Does a sentinel or a subset of short telomeres determine replicative senescence? Molecular Biology of the Cell 15:3709-3718 doi:10.1091/mbc.e04-03-0207