Integration of Telomere Length Dynamics into Systems Biology Framework: A Review

Libertas Academica
FREEDOM TO RESEARCH

Lilit Nersisyan

Group of Bioinformatics, Institute of Molecular Biology, National Academy of Sciences RA, Yerevan, Republic of Armenia.

ABSTRACT: Telomere length dynamics plays a crucial role in regulation of cellular processes and cell fate. In contrast to epidemiological studies revealing the association of telomere length with age, age-related diseases, and cancers, the role of telomeres in regulation of transcriptome and epigenome and the role of genomic variations in telomere lengthening are not extensively analyzed. This is explained by the fact that experimental assays for telomere length measurement are resource consuming, and there are very few studies where high-throughput genomics, transcriptomics, and/or epigenomics experiments have been coupled with telomere length measurements. Recent development of computational approaches for assessment of telomere length from whole genome sequencing data pave a new perspective on integration of telomeres into high-throughput systems biology analysis framework. Herein, we review existing methodologies for telomere length measurement and compare them to computational approaches, as well as discuss their applications in large-scale studies on telomere length dynamics.

KEYWORDS: telomere length, gene expression, age-related diseases, cancers

Introduction

Telomeres are nucleoprotein complexes located at the ends of eukaryotic chromosomes. The telomeric DNA sequence comprises short consecutive repeats, which in humans are in the form of "TTAGGG" and normally span 2–10 kb, depending on age, tissue type, and cell state.[1] They protect chromosome ends from being recognized as double strand breaks and from forming end-to-end fusions. Because of the "end replication problem", whereby the replicative machinery of the cell is not able to fully replicate the linear chromosomes, telomeres are gradually shortened after each cell division.[2–4] Replicative burden in most somatic cells leads to telomere shortening, and eventually to senescence and apoptosis,[5] whereas stem cells and cancer cells overcome the telomere shortening problem by developing mechanisms for their elongation.[6] Changes in telomere length dynamics have been implicated in physiological aging and in the development of age-related diseases and cancers.[7]

Based on the importance of telomere length dynamics in the development of certain diseases, they have been studied for a long time. A number of experiments have been performed to identify the genes affected by Telomere Position Effect (TPE),[8] long distance effects,[9] and genomic variations affecting telomere length.[10] These studies have been based on the experimental measurement of telomere length. Despite the importance of such studies, data on the association of

telomere length with gene regulatory mechanisms are scarce. This is caused by two main factors: (1) a part is based on low-throughput data, such as measuring telomeres and small number of genes in the same experiment, and (2) the other part is based on high-throughput data on genotypes, epigenetic modifications, and gene expression, for which experimental measurement of telomere lengths is also available. Unfortunately, the latter cases are extremely rare, thus limiting the data source for analysis of telomere length regulatory networks in the cell. Finally, the experimental assays of telomere length measurement usually require a large amount of source DNA material, or viable dividing cells, which lead to sample availability issues for the telomere measurement assays as well. For this reason, most of the studies on telomere length have been performed on blood leukocytes.[11]

Emergence of high-throughput sequencing technologies has provided grounds for the development of novel data analysis methodologies to capture the information on telomere length from the existing DNA sequencing data. This allows for integrating information on telomere length into linked data on gene variants, and possibly gene expression, chromatin immunoprecipitation sequencing (ChIP-seq), and DNA methylation. This new approach performed on high-throughput data should allow for fostering integrative analysis of telomere regulatory networks in health and disease.

The purpose of this review is to stress on the possibility of efficiently obtaining telomere-related information from the currently available high-throughput technologies and show the importance of such computations to foster telomere biology research. For this, we will sum up the state-of-the-art experimental methods for measurement of telomere length and their application in large-scale studies, and then will talk about current computational methods for telomere length measurement and existing data on integration of telomere length into high-throughput data analysis pipelines.

Telomere Length Dynamics in Aging and Disease

Regulation of telomere length dynamics is a complex interplay between telomere attrition and elongation processes. Studies indicate that replicative telomere attrition caused by the end replication problem accounts for only up to 50–100 bp shortening per each cell division.[12] Other factors, such as reactive oxygen species, radiation, and a number of additional stress-associated environmental factors, as well as genetic background, might contribute to faster attrition rates.[13] The main telomere elongation mechanism, which is employed by most cells with high proliferative capacity, is dependent on the expression of telomerase, an RNA-dependent reverse transcriptase, which elongates the telomeres.[6,14,15] A part of cancer cells elongate their telomeres via alternative elongation mechanisms, dependent on recombination events between sister chromatids.[16–19] These processes are regulated by various players in the telomere elongation and maintenance pathway.[20]

Changes in telomere length homeostasis are differently implicated in aging and disease.[7] While telomere lengthening and maintenance machinery are shown to be active in cancer cells, senescence-related diseases are mostly characterized by accelerated telomere shortening. The latter may lead to two main consequences: (1) extremely short telomeres lead to telomere dysfunction and chromosome instability[21,22] and (2) changes in telomere length up to critically short levels are associated with the changes in expression and epigenetic modifications of several genes.[8,9] This association has been explained by telomere position effect, reversible silencing of genes located near telomeres,[8] and by long distance interactions based on physical localization of chromosomes within the nucleus.[9] Alternative direct and indirect effects of telomere length dynamics on regulation of gene expression is still an open field for investigation.

Telomere shortening is considered as one of the factors leading to organismal aging, though it is not known whether it results from replicative arrest in somatic and stem cells, or activation of telomere shortening-induced signaling cascades.[23–25] The onset of many complex human diseases is associated with aging and partly attributed to the telomere shortening. A number of studies have implicated leukocyte telomere length (LTL) as a risk and/or prognostic marker, for cardiovascular diseases, ischemic strokes, type 2 diabetes, and neurodegenerative diseases; however, the direct link between telomere shortening and disease development is not always clear.[26–31] Additionally, telomere length is crucial for preserving replicative capacity and clonal exhaustion of actively dividing immune cells.[32] Age-dependent telomere shortening in B- and T-lymphocytes has been linked to defective immune responses and development of many age-related diseases. In some cases, telomere shortening in T-cells has been linked to development of autoimmune diseases.[33] Individuals with short telomeres or those suffering from premature aging syndromes have susceptibility of developing infectious diseases, because of telomere shortening-associated decline in immune functions.[32] Some diseases are directly linked to telomere dysfunction, such as dyskeratosis congenita, where mutations occur in genes responsible for telomere maintenance.[34]

The role of telomeres in development of cancers is more complex. In normal cells, telomere shortening up to critical levels leads to cell growth arrest and apoptosis. Overcoming the cell crisis checkpoint at this stage leads to chromosomal instability in many cancer cells.[22,35] While chromosomal instability is considered an important factor for cancer onset, maintenance of telomeres is important for cell survival in cancer progression.[22] Most of the cancer cells express telomerase to elongate the shortened telomeres,[14] while others develop alternative lengthening mechanisms.[17] Depending on the cancer type, both short and long telomeres have been described as markers of risk.[36] In some cases, telomerase expression is considered as a marker of aggressiveness and poor prognosis.[37] Therapies inhibiting the action of telomerase have been shown to be effective in limiting cell growth in some cancers.[38]

Telomere Length Measurement

The existing methodologies for telomere length measurement can be roughly divided into experimental and computational approaches. Experimental methods assume direct measurement of telomere lengths, while computational methods assume meta-analysis of the existing sequencing data for length estimation (Table 1).

Experimental methods. Experimental methods for telomere length measurement include, but are not limited to, terminal restriction fragment (TRF) analysis, quantitative polymerase chain reaction (qPCR), quantitative fluorescence in situ hybridization (qFISH),[39] etc., which measure the mean and/or chromosome-specific telomere lengths.

The first method developed for measuring the mean telomere length (MTL) was TRF. It first performs a test to select intact DNA, and then uses restriction enzymes to cut the whole DNA sequence except for the telomeric regions. The remaining long telomeric fragments are then separated using agarose gel electrophoresis. The isolated fragments are analyzed with Southern blot using telomere probe ligation.[40] Since the fragments are of nonuniform length, they appear as dispersive smears. A crucial factor in the accuracy of the average telomere length calculation is accounting for the smear length and the intensity of probe binding. TRF results

Table 1. Comparison of experimental and computational methods for telomere length measurement.

METHOD	TRF[39,40]	qPCR[39]	STELA[44]	qFISH[47]	TELSEq[50]	COMPUTEL[52]
Type	Experimental	Experimental	Experimental	Experimental	Computational	Computational
Required source	$1–3 \times 10^6$ cells (0.5–10 µg DNA)	Blood or DNA sample (20 ng DNA/ reaction)	$1–1 \times 10^5$ cells	Actively dividing cells (15–20 metaphases)	Illumina WGS data	Illumina WGS data
Estimate of telomere length	Mean length for total cell population	Ratio of telomere amplification to that of a single copy gene	Length at a subset of individual chromosomes	Cell average length	Mean length per single chromosome set	Mean length per single chromosome set
Advantage	Gold standard; numerous studies for comparison	Less resource consuming than TRF; standardization to a single copy gene	Detects critically short telomeres	Detects telomere length distribution at individual chromosome ends; Detects end-to-end fusions and telomere-free ends	Extracts information from WGS data	Extracts information from WGS data; does not account for interstitial telomeric repeats; is not dependent on read length; is robust against sequencing errors
Disadvantage	Requires large amount of DNA; Resource consuming; polymorphisms in subtelomeres may impact the results; subtelomeres may add up to 4 kb bias; different experiments should be compared with caution	Pipetting errors; problems with absolute telomere length measurement, since appropriate standards are lacking; prone to bias from calibration steps	Is not suited for all the chromosomes; does not measure very long telomeres	Labor intensive; requires viable cells; length is expressed in relative fluorescence units;	Arbitrary repeat threshold; doesn't work with short/long reads; biased by interstitial telomeric repeats; sensitive to sequencing errors; doesn't report end-specific telomere lengths	Doesn't report end-specific telomere lengths

obtained from different experiments are dependent on various factors, such as source DNA quality and quantity, choice of restriction enzymes, gel density, signal intensity calculations, and length adjustment.[41] Moreover, the distance of the furthest subtelomeric restriction site to the telomeres is estimated to be 2.5–4 kb, depending on the chromosomes and the restriction enzyme used. Thus, the results of different studies should be compared with caution and account for the difference of restriction enzymes used. Finally, TRF requires large amounts of source DNA and is also not capable of capturing short telomeres.[39] Since the TRF has been the first technique for MTL measurement, it has served as a reference for further emerging methods and is thus considered the "gold standard".

The quantitative PCR approach is less elaborate compared to TRF. It is based on the amplification of telomeric regions via telomere-specific primers. The basic assumption is that the longer the telomeric sequence of the source DNA, the more there are places for the primers to attach, and more amplicons will be generated. Comparison of their relative quantity and that of the amplicons generated from single-copy gene PCR gives a telomeric over single-copy gene signals ratio, which is correlated with the overall telomeric content of the cell. One major disadvantage of this method is that it is strictly dependent on the initial calibration steps, and the results derived from different experiments are difficult to compare.[39]

Compared to TRF, qPCR is thought to be more prone to measurement errors, which are being addressed by further modifications and amendments.[42] For example, in traditional qPCR, the reactions for the single-copy gene and telomeres are performed in different test tubes, which is a source of potential pipetting error. Thus, a modified version of this approach, such as monochrome multiplex qPCR, has been suggested, where the single-gene and telomere amplifications are performed in the same tube.[43] Besides measurement errors, telomeric over single-copy gene signals ratios obtained with these methods should be treated with caution, accounting for possible copy number and chromosome number variations.[39]

Other methods are used less frequently and serve more specific purposes. Single telomere length analysis (STELA) utilizes a 3′ overhang specific linker and a subtelomeric primer to amplify telomeres at specific chromosome ends with PCR.[44] STELA uses subtelomeric primers of known lengths, and their distance to the true telomeric region is known, which makes it more accurate than the TRF. Additionally, it estimates telomere lengths at specific chromosomes and is able to measure short telomeres, which is important in telomere shortening–induced senescence studies and for accounting for the telomere length variability across chromosomes. However, sequence variability at most of the subtelomeric regions allows only for capturing telomeres at chromosome arms Xp, Yp, 2p, 11q, 12q, and 17p.[45] Another limitation of STELA is that it is

not able to capture long telomeres (lengths more than 8 kb are usually not captured).[39,44]

qFISH is used for measuring telomere lengths at metaphase chromosomes via ligation of telomere-specific fluorescent probes. The signal intensity is then compared to a standard of known telomere length.[46] This method is accurate and is advantageous of providing arm-specific telomere lengths. The main drawback is that cells should be able to divide, and thus, it is not applicable to cell cycle arrested cells. Also, because the method is hybridization based, it makes it difficult to quantify extremely short telomeres.

There is a wide variety of other techniques, each aimed at overcoming limitations of the existing ones.[39] However, there have been concerns about the comparability of the results obtained in different experiments, raising the need for proper calibration of the techniques based on the gold standard reference.[47,48]

Another major limiting factor for understanding the complex picture of telomere regulatory network is the lack of high-throughput data coupled with experimental results on telomere length. Only a limited number of studies exist, where telomere length measurement experiments have been coupled with high-throughput genome, transcriptome, or epigenome data. Importantly, in order to put the data obtained from these two experiments in the same context, these should be performed on the same samples at the same time. Existing data do not allow for making appropriate analysis, thus posing the necessity of computational techniques that could obtain telomere length information from genome sequencing data.

Computational methods. The advent of next-generation sequencing (NGS) techniques has allowed for obtaining high-throughput whole genome, transcriptome, and epigenome data. Compared to low-throughput experiments, these techniques allow for performing systems biology analysis with integration of data from different sources into a single outlook. Currently, large amounts of coupled whole genome sequencing (WGS), RNA sequencing, ChIP-seq, and bisulfite NGS data for individual organisms are available.[49] These data are valuable for making systems-level association studies with telomere length dynamics.

Telomere length information can be obtained from WGS data, which is presented in the form of short read fragments (20–150 bp). During the last five years, a few methodologies have been developed for estimating telomere length from these data, all of which are based on capturing "telomeric" reads, which are presumably derived from telomeric regions of the genome. The first attempts were based on simple counts of reads that contain a certain number of intact telomeric TTAGGG repeats.[50,51] Read count based software TelSeq scans for reads containing more than a threshold amount of TTAGGG repeats and compares it to the number of genomic reads with the same GC content. This relative count is then multiplied by a GC-normalized genome length constant, which gives an estimate of absolute MTL.[50] This method provides correlation

of telomere length estimates with experimental results in certain settings (for 100 bp long Illumina reads); however, it has several limitations. The output of TelSeq is dependent on the threshold chosen for telomeric repeat counts, and there is no established way of calibrating this threshold in different sequencing settings, thus making the results of TelSeq-based counts noncomparable across studies. Accounting for this repeat count limitation with optimum performance with a threshold of 7 repeats in 100 bp reads, TelSeq performs poorly on read lengths less than 50 bp or more than 100 bp. The results are also nonrobust to sequencing errors, which easily obscure the telomeric repeat pattern and change the length estimate. Finally, TelSeq is hard-coded for GC-normalized lengths and is not suitable for calculations based on nonhuman chromosomes or nondiploid chromosome sets.

Another software package Computel has been recently developed for MTL estimation from WGS data.[52] It estimates MTL by aligning short reads to a special telomeric reference. The reference is designed in a way to uniquely capture telomeric reads coming from pure telomeric regions of the genome, or from the junctions of telomeres and subtelomeres. It also minimizes the possibility of capturing interstitial telomeric repeat rich reads, which do not originate from telomeres, but from other regions of the genome.[53] In their study, the authors have conducted detailed analysis of the performance of Computel on synthetic and experimental reads and performed comparative analysis with TelSeq.[52] Importantly, Computel is more robust to sequencing errors and is flexible for variations in sequencing platforms, accounting for read length and depth of coverage.[52] Finally, the results of Computel obtained from various settings are reliably comparable, which makes it suitable for integrative analysis in telomere length association studies performed on high-throughput sequencing data, across experiments. One major limitation of Computel, as well as TelSeq, is that these software do not count the telomere length at individual chromosomes, since the latter is an important marker of genome stability.[54] However, despite this limitation, MTL is proven to be an informative surrogate marker of individual telomere lengths in a large amount of association studies.[55–57]

Except for WGS data, other types of *-omics* data are not suited for direct measurement of telomere length. Whole exome sequencing libraries mostly do not cover telomeric sequences. Although the discussed computational approaches can be used to calculate "telomere abundance" in these data, obtained results can be poorly correlated with the WGS-measured telomere length[50] and do not report the absolute length values. Telomeric and subtelomeric sequences, such as the telomeric repeat containing RNA (TERRA) repeats, also get transcribed,[58] and NGS RNA sequencing data can be used to compute relative abundance of these transcripts. Although their role is still to be defined,[59] these measurements may also be useful. Finally, ChIP-seq data contain the information on the DNA regions enriched with bound proteins, such as chromatin-binding proteins and transcription factors.

Computational measurement of telomeric sequences in this type of data may be useful for the identification of chromatin states at telomeres.[60]

Application of Telomere Length Measurement Approaches in Large-Scale Studies

Epidemiological studies on telomere association with age and age-related diseases. Epidemiological studies on telomere length and its association with aging and diseases in humans have mostly been performed using TRF and qPCR techniques discussed above. These assays require a large amount of source DNA, which is easily available from blood leukocytes. For this reason, LTL has been considered as a biomarker of biological aging in a considerable amount of epidemiological studies. Interestingly, LTL is assumed to be correlated with telomere lengths of other body cells.[5] Most of these studies point on reverse association of LTL with chronological age, but there are some exceptions to this rule, which are at times hypothesized to be population or ethnicity specific.[61–63] Some epidemiological studies indicate that telomere length is also predictive of mortality, but a meta-analysis of these studies indicates that this predictive association diminishes with age.[64] Two studies performed on the very old population have measured telomere length in blood cells using qPCR and TRF assays and have shown that there is no association between telomere length and mortality in these cohorts.[65,66] Some studies pinpoint on the fact that interindividual differences between LTL are largely determined at birth or within the first few years after birth, and then diminish at constant rates during life.[1,67] As indicated by a recent review, existing data on the association of telomere length with age does not imply causality, and the role of telomeres in the biology of aging still needs further investigation.[68]

Based on its association with age in the general population, telomere length has been studied in the context of its association with diseases of the aging population, such as cardiovascular diseases, cerebrovascular diseases, and type 2 diabetes. Moreover, since chronological age is not the perfect marker of organismal aging, telomeres have been considered as better indicators of biological aging.[11] The results obtained from various studies clearly indicate that atherosclerosis is associated with shortened telomeres in the cells of the vascular endothelium, as well as in leukocytes. The latter is explained by the correlation of telomere length in different tissues in the body, as well as by clonal exhaustion of immune cells and hematopoietic somatic cells in the presence of chronic inflammation.[69] Myocardial infarction and ischemic stroke, as well as other cardiovascular and cerebrovascular diseases, have also been shown to be mostly associated with shorter LTL.[11] However, there is a lot of discrepancy in the associations found, which raise doubts about comparability, reproducibility, and accuracy of the applied measurement techniques, especially stressing the difference in technologies and intra-study measurement errors.[70] A meta-analysis of 15 cohort and 12 case–control studies performed before 2013 by D'Mello et al has shown that myocardial infarction, ischemic stroke, and type 2 diabetes, but not coronary artery disease, are significantly associated with shorter LTL.[11] In five of these studies, LTL was measured with TRF essay, while the rest were measured with qPCR. Interestingly, the studies using TRF essay showed a greater effect size for stroke, which can be explained by measurement bias in qPCR,[11] which reduces the effect of possible association.

An interesting study has been conducted recently to measure the association of telomere length in the largest cohort to date in a single experimental setting via robotics-based automation of qPCR experimental procedures.[71] The study has confirmed that telomere length negatively correlates with age up to 75 years, while having positive correlation with age in older individuals. It has also confirmed that after the age of 50 years, females have longer telomeres than males. Finally, the study has found that telomere length variance between individuals increases with age.[71] Due to automation, the authors have eliminated the inter-study variation of telomere length measurements and have increased the statistical power of the tests. This partially overcomes the limitations of experimental approaches for telomere length measurement, but stresses the resources required to perform accurate telomere length association studies.

Genome-wide association studies on MTL. Evidence points that telomere length at birth and attrition rates may be largely conditioned by genetic background. Thus, genome-wide association studies (GWASs) for MTL are valuable to find the possible causality and reveal the role of these genes in the regulation of telomere length and telomere biology. To date, several large GWASs for MTL have been performed on various populations mainly via qPCR- or TRF-based measurement of LTL and genotyping arrays. The first GWAS has been performed for mean TRF length of leukocytes and found *DDX11* gene as an associated locus.[72] Several GWAS studies have been conducted until then, with a trend of finding new loci in each experiment.[10,73–76] A meta-analysis on 37,684 individuals has identified seven loci associated with telomere length, including genes that have an established role in telomere biology, such as *TERC*, *OBFC1*, *NAF1*, and *RTEL1*, and two other loci, namely, *ACYP2* and *ZNF208*.[10] All these genes, except for *ZNF208*, were further validated within the largest GWAS for telomere length, in the Collaborative Oncology Gene-environment Study project.[55] This project also implicated new telomere-associated genes, such as *PXK*, *ZNF311*, and *BCL2L1*.[55] However, the association of these genes is still dependent on the population or the study design, and there is no gene to be associated with telomere length variations in all or most of the GWASs.[77] Only a few gene variant associations have been replicated in different studies, such as *TERC*, *TERT*, *MRE11*, and *OBFC1*.[10,75,76,78] The causal relationship of these SNPs with telomere length dynamics has been addressed to a lesser extent.[10,75]

In contrast to GWAS, few attempts have been made to link MTL to high-throughput gene expression studies. An interesting study on integration of Hi-C and 3D-FISH along with gene expression microarray analysis has revealed a mechanism for telomere length-regulated expression of three genes, namely, *ISG15*, *DSP*, and *C1S*.[9] *ISG15*, along with a number of other genes, has been identified to be differentially expressed in telomere length-dependent manner in telomerase overexpression settings via gene expression microarray and TRF analysis in cancer cells.[79] A limited number of studies have addressed high-throughput gene expression in the context of replicative senescence, but the link between global gene expression and telomere length dynamics has not been established yet.[80] There is also little or no extensive study on association of high-throughput epigenome data and telomere length dynamics, although a number of studies have analyzed NGS data on epigenomic determinants affecting telomere length.[81,82]

Use of computational approaches for integration of telomere length into high-throughput data analysis framework. Based on the wide availability of transcriptome and epigenome NGS data coupled with WGS, computational approaches for telomere length measurement will soon fill the gap of knowledge in the biology of telomere length dynamics. One of the pioneering works in this direction has been the study by Parker et al on association of telomere length dynamics in pediatric cancers, where gain and loss of telomeric DNA has been estimated from WGS data by counting the number of short reads containing at least four consecutive telomeric repeats.[51] Although the method they have used for estimation of telomere length was not calibrated to return accurate results, they have shown that it is a promising approach. The number of similar studies started to increase with the introduction of computational pipelines for estimation of telomere length from WGS data (TelSeq[50] and Computel[52]). These software packages have already been used in a number of studies.[62,83–87]

Using TelSeq, it has been shown that adversity in major depression is associated with shorter telomeres and increased copies of mitochondrial DNA.[83] Analysis of WGS data in diffuse glioma has revealed that mutations in *ATRX* gene associated with chromatin remodeling, but not TERT are linked to increased telomere length, where the latter was measured with TelSeq.[84]

Another set of studies have used Computel for measurement of telomere length from the available NGS data. A quantitative trait association study performed on whole genomes of 168 South Asians has revealed *ADARB2* gene, with SNPs significantly associated with long telomeres, as well as revealed that the subpopulation of Sikhs has significantly longer telomeres.[62] Previously, *ADARB2* gene has been linked to extreme longevity, but its role in telomere biology has to be investigated.[85] Computel has also been used in integrative systems analysis of telomere length association in lung adenocarcinoma cell lines, for which whole genome, transcriptome, and epigenome data are available. Combined analysis of genome and transcriptome data from these cell lines revealed a partial correlation network with five genes directly correlated with telomere length, namely, *PIR*, *GPR19*, *FECH*, *ASUN*, and *NDUFA1*.[86] Two of these genes, namely, *PIR* and *GPR19*, have been previously implicated in overcoming the senescence barrier and acceleration of mesenchymal transition through cell cycle in lung cancers.[86,87] The role of the other three (*TASUN*, *FECH*, and *NDUFA1*) in telomere dynamics and lung cancer still needs further investigation, though they have known roles in the control of cell cycle and mitochondrial processes.

In summary, the applications of computational methods to high-throughput data show the potential of integrative view on telomere length dynamics and its association with molecular events. Most of the controversies and knowledge gaps on telomere biology stem from the fact that those are not really viewed at systems level. Telomeres are peculiar structures that are tightly regulated by complex protein machinery and themselves regulate the expression of certain genes. These relationships should form a complex network including positive and negative feedback loops. Development of appropriate instruments to model this network and analyze its behavior is a key to get deeper insight into the role that telomeres play in the cell development.

Conclusion

Biological role of telomeres in the establishment of cellular processes and cell fate still remains a mystery. There is a vast amount of studies aimed at finding an association and a causal relationship between telomere length dynamics and development of cellular senescence and organismal aging, as well as development and progression of complex human diseases and cancers. However, these studies, although mostly linking telomere length with aging, age-related diseases, and some cancers, fail to establish the causality of these associations and the biological role of telomere length dynamics. The various low-throughput assays for measurement of telomere length that have been used to date, such as TRF, qPCR, and qFISH, have drawbacks and limitations and are resource consuming. While high-throughput sequencing and technologies have opened a new era in systems biology research allowing for integration of data on genome, transcriptome, and epigenome for thousands of genes simultaneously, very limited number of studies have also performed coupled assays for telomere length measurement. Novel computational approaches for measuring telomere length from the already available WGS data allow for integrating telomeres into the high-throughput analysis frameworks. These approaches have short history, but have already been used in several studies, and will likely contribute to a large amount of findings in telomere biology in the future.

Abbreviations

TRF: terminal restriction fragment
qPCR: quantitative polymerase chain reaction

qFISH: quantitative fluorescence in situ hybridization
MTL: mean telomere length
T/S: telomeric over single-copy gene signals
STELA: single telomere length analysis
NGS: next-generation sequencing
WGS: whole genome sequencing
RNA-seq: RNA sequencing
ChIP-seq: chromatin immunoprecipitation sequencing
LTL: leukocyte telomere length
GWAS: genome-wide association studies
TPE: telomere position effect
TERRA: telomeric repeat containing RNA

Author Contributions

Wrote the manuscript: LN. Agreed with manuscript results and conclusions: LN. Made critical revisions and approved the final version: LN. The author reviewed and approved the final manuscript.

REFERENCES

1. Monaghan P. Telomeres and life histories: the long and the short of it. *Ann N Y Acad Sci*. 2010;1206:130–42.
2. Blackburn EH. Structure and function of telomeres. *Nature*. 1991;350(6319): 569–73.
3. Zakian VA. Telomeres: the beginnings and ends of eukaryotic chromosomes. *Exp Cell Res*. 2012;318(12):1456–60.
4. Sfeir A, de Lange T. Removal of shelterin reveals the telomere end-protection problem. *Science*. 2012;336(6081):593–7.
5. Daniali L, Benetos A, Susser E, et al. Telomeres shorten at equivalent rates in somatic tissues of adults. *Nat Commun*. 2013;4:1597.
6. Hug N, Lingner J. Telomere length homeostasis. *Chromosoma*. 2006;115(6): 413–25.
7. Xi H, Li C, Ren F, Zhang H, Zhang L. Telomere, aging and age-related diseases. *Aging Clin Exp Res*. 2013;25(2):139–46.
8. Baur JA, Zou Y, Shay JW, Wright WE. Telomere position effect in human cells. *Science*. 2001;292(5524):2075–7.
9. Robin JD, Ludlow AT, Batten K, et al. Telomere position effect: regulation of gene expression with progressive telomere shortening over long distances. *Genes Dev*. 2014;28(22):2464–76.
10. Codd V, Nelson CP, Albrecht E, et al. Identification of seven loci affecting mean telomere length and their association with disease. *Nat Genet*. 2013;45(4):e1–2.
11. D'Mello MJJ, Ross SA, Briel M, Anand SS, Gerstein H, Paré G. Association between shortened leukocyte telomere length and cardiometabolic outcomes: systematic review and meta-analysis. *Circ Cardiovasc Genet*. 2015;8(1): 82–90.
12. Huffman KE, Levene SD, Tesmer VM, Shay JW, Wright WE. Telomere shortening is proportional to the size of the G-rich telomeric 3′-overhang. *J Biol Chem*. 2000;275(26):19719–22.
13. Epel ES, Blackburn EH, Lin J, et al. Accelerated telomere shortening in response to life stress. *Proc Natl Acad Sci U S A*. 2004;101(49):17312–5.
14. Shay JW, Wright WE. Role of telomeres and telomerase in cancer. *Semin Cancer Biol*. 2011;21(6):349–53.
15. Palm W, de Lange T. How shelterin protects mammalian telomeres. *Annu Rev Genet*. 2008;42:301–34.
16. Neumann AA, Watson CM, Noble JR, Pickett HA, Tam PPL, Reddel RR. Alternative lengthening of telomeres in normal mammalian somatic cells. *Genes Dev*. 2013;27(1):18–23.
17. Reddel RR. Alternative lengthening of telomeres, telomerase, and cancer. *Cancer Lett*. 2003;194(2):155–62.
18. Cesare AJ, Reddel RR. Telomere uncapping and alternative lengthening of telomeres. *Mech Ageing Dev*. 2008;129(1–2):99–108.
19. Nittis T, Guittat L, Stewart SA. Alternative lengthening of telomeres (ALT) and chromatin: is there a connection? *Biochimie*. 2008;90(1):5–12.
20. Zhao Z, Pan X, Liu L, Liu N. Telomere length maintenance, shortening, and lengthening. *J Cell Physiol*. 2014;229(10):1323–9.
21. Takai H, Smogorzewska A, De Lange T. DNA damage foci at dysfunctional telomeres. *Curr Biol*. 2003;13(17):1549–56.
22. Greenberg RA. Telomeres, crisis and cancer. *Curr Mol Med*. 2005;5(2):213–8.
23. Blasco MA. Telomere length, stem cells and aging. *Nat Chem Biol*. 2007; 3(10):640–9.
24. Kappei D, Londoño-Vallejo JA. Telomere length inheritance and aging. *Mech Ageing Dev*. 2008;129(1–2):17–26.
25. Mather KA, Jorm AF, Parslow RA, Christensen H. Is telomere length a biomarker of aging? A review. *J Gerontol A Biol Sci Med Sci*. 2011;66 A(2):202–13.
26. Kota LN, Bharath S, Purushottam M, et al. Reduced telomere length in neurodegenerative disorders may suggest shared biology. *J Neuropsychiatry Clin Neurosci*. 2015;27(2):e92–6.
27. Farzaneh-Far R, Cawthon RM, Na B, Browner WS, Schiller NB, Whooley MA. Prognostic value of leukocyte telomere length in patients with stable coronary artery disease: data from the heart and soul study. *Arterioscler Thromb Vasc Biol*. 2008;28(7):1379–84.
28. Martin-Ruiz C, Dickinson HO, Keys B, Rowan E, Kenny RA, Von Zglinicki T. Telomere length predicts poststroke mortality, dementia, and cognitive decline. *Ann Neurol*. 2006;60(2):174–80.
29. Salpea KD, Talmud PJ, Cooper JA, et al. Association of telomere length with type 2 diabetes, oxidative stress and UCP2 gene variation. *Atherosclerosis*. 2010;209(1):42–50.
30. Huzen J, van Veldhuisen DJ, van Gilst WH, van der Harst P. Telomeres and biological ageing in cardiovascular disease. *Ned Tijdschr Geneeskd*. 2008; 152(22):1265–70.
31. Cawthon RM, Smith KR, O'Brien E, Sivatchenko A, Kerber RA. Association between telomere length in blood and mortality in people aged 60 years or older. *Lancet*. 2003;361(9355):393–5.
32. Georgin-Lavialle S, Aouba A, Mouthon L, et al. The telomere/telomerase system in autoimmune and systemic immune-mediated diseases. *Autoimmun Rev*. 2010;9(10):646–51.
33. Andrews NP, Fujii H, Goronzy JJ, Weyand CM. Telomeres and immunological diseases of aging. *Gerontology*. 2010;56(4):390–403.
34. Nelson ND, Bertuch AA. Dyskeratosis congenita as a disorder of telomere maintenance. *Mutat Res*. 2012;730(1–2):43–51.
35. Campbell PJ. Telomeres and cancer: from crisis to stability to crisis to stability. *Cell*. 2012;148(4):633–5.
36. Fernndez-Marcelo T, Morn A, De Juan C, et al. Differential expression of senescence and cell death factors in non-small cell lung and colorectal tumors showing telomere attrition. *Oncology*. 2012;82(3):153–64.
37. Bertorelle R, Rampazzo E, Pucciarelli S, Nitti D, De Rossi A. Telomeres, telomerase and colorectal cancer. *World J Gastroenterol*. 2014;20(8):1940–50.
38. Ruden M, Puri N. Novel anticancer therapeutics targeting telomerase. *Cancer Treat Rev*. 2013;39(5):444–56.
39. Aubert G, Hills M, Lansdorp PM. Telomere length measurement-Caveats and a critical assessment of the available technologies and tools. *Mutat Res*. 2012;730(1–2):59–67.
40. Kimura M, Stone RC, Hunt SC, et al. Measurement of telomere length by the Southern blot analysis of terminal restriction fragment lengths. *Nat Protoc*. 2010;5(9):1596–607.
41. Göhring J, Fulcher N, Jacak J, Riha K. TeloTool: a new tool for telomere length measurement from terminal restriction fragment analysis with improved probe intensity correction. *Nucleic Acids Res*. 2014;42(3):e21.
42. Eisenberg DTA, Kuzawa CW, Hayes MG. Improving qPCR telomere length assays: controlling for well position effects increases statistical power. *Am J Human Biol*. 2015;27(4):570–5.
43. Cawthon RM. Telomere length measurement by a novel monochrome multiplex quantitative PCR method. *Nucleic Acids Res*. 2009;37(3):e21.
44. Baird DM, Rowson J, Wynford-Thomas D, Kipling D. Extensive allelic variation and ultrashort telomeres in senescent human cells. *Nat Genet*. 2003;33(2):203–7.
45. Britt-Compton B, Rowson J, Locke M, Mackenzie I, Kipling D, Baird DM. Structural stability and chromosome-specific telomere length is governed by cis-acting determinants in humans. *Hum Mol Genet*. 2006;15(5):725–33.
46. O'Sullivan JN, Finley JC, Risques R-A, Shen W-T, Gollahon KA, Rabinovitch PS. Quantitative fluorescence in situ hybridization (QFISH) of telomere lengths in tissue and cells. *Curr Protoc Cytom*. 2005;Chapter 12:Unit12.6.
47. Verhulst S, Susser E, Factor-Litvak PR, et al. Commentary: the reliability of telomere length measurements. *Int J Epidemiol*. 2015;44(5):1683–6.
48. Aviv A. Commentary: raising the bar on telomere epidemiology. *Int J Epidemiol*. 2009;38(6):1735–6.
49. Leinonen R, Sugawara H, Shumway M. The sequence read archive. *Nucleic Acids Res*. 2011;39(suppl 1):D19–21.
50. Ding Z, Mangino M, Aviv A, Spector T, Durbin R, UK10 K Consortium. Estimating telomere length from whole genome sequence data. *Nucleic Acids Res*. 2014;42(9):e75.
51. Parker M, Chen X, Bahrami A, et al. Assessing telomeric DNA content in pediatric cancers using whole-genome sequencing data. *Genome Biol*. 2012; 13(12):R113.
52. Nersisyan L, Arakelyan A. Computel: computation of mean telomere length from whole-genome next-generation sequencing data. *PLoS One*. 2015;10(4): e0125201.

53. Azzalin CM, Nergadze SG, Giulotto E. Human intrachromosomal telomeric-like repeats: sequence organization and mechanisms of origin. *Chromosoma*. 2001;110(2):75–82.

54. Hemann MT, Strong MA, Hao LY, Greider CW. The shortest telomere, not average telomere length, is critical for cell viability and chromosome stability. *Cell*. 2001;107(1):67–77.

55. Pooley KA, Bojesen SE, Weischer M, et al. A genome-wide association scan (GWAS) for mean telomere length within the COGS project: identified loci show little association with hormone-related cancer risk. *Hum Mol Genet*. 2013;22(24):5056–64.

56. Zee RYL, Michaud SE, Germer S, Ridker PM. Association of shorter mean telomere length with risk of incident myocardial infarction: a prospective, nested case-control approach. *Clin Chim Acta*. 2009;403(1–2):139–41.

57. Zee RYL, Castonguay AJ, Barton NS, Buring JE. Mean telomere length and risk of incident colorectal carcinoma: a prospective, nested case-control approach. *Cancer Epidemiol Biomarkers Prev*. 2009;18(8):2280–2.

58. Luke B, Lingner J. TERRA: telomeric repeat-containing RNA. *EMBO J*. 2009;28(17):2503–10.

59. Maicher A, Lockhart A, Luke B. Breaking new ground: digging into TERRA function. *Biochim Biophys Acta*. 2014;1839(5):387–94.

60. Vega-Palas M, Vaquero-Sedas M. Integrating telomeres into genome-wide ChIP-seq analyses; 2013. doi: http://dx.doi.org/10.1038/protex.2013.039.

61. Bhupatiraju C, Saini D, Patkar S, Deepak P, Das B, Padma T. Association of shorter telomere length with essential hypertension in Indian population. *Am J Hum Biol*. 2012;24(4):573–8.

62. Hakobyan A, Nersisyan L, Arakelyan A. Quantitative trait association study for mean telomere length in the South Asian genomes. *Bioinformatics*. 2016; 32(11):1697–700. doi: 10.1093/bioinformatics/btw027.

63. Jodczyk S, Fergusson DM, Horwood LJ, Pearson JF, Kennedy MA. No association between mean telomere length and life stress observed in a 30 year birth cohort. *PLoS One*. 2014;9(5):e97102.

64. Boonekamp JJ, Simons MJP, Hemerik L, Verhulst S. Telomere length behaves as biomarker of somatic redundancy rather than biological age. *Aging Cell*. 2013;12(2):330–2.

65. Martin-Ruiz CM, Gussekloo J, van Heemst D, von Zglinicki T, Westendorp RGJ. Telomere length in white blood cells is not associated with morbidity or mortality in the oldest old: a population-based study. *Aging Cell*. 2005;4(6):287–90.

66. Bischoff C, Petersen HC, Graakjaer J, et al. No association between telomere length and survival among the elderly and oldest old. *Epidemiology*. 2006; 17(2):190–4.

67. Okuda K, Bardeguez A, Gardner JP, et al. Telomere length in the newborn. *Pediatr Res*. 2002;52(3):377–81.

68. Simons MJP. Questioning causal involvement of telomeres in aging. *Ageing Res Rev*. 2015;24(pt B):191–6.

69. Aviv A, Levy D. Telomeres, atherosclerosis, and the hemothelium: the longer view. *Annu Rev Med*. 2012;63:293–301.

70. Sanders JL, Newman AB. Telomere length in epidemiology: a biomarker of aging, age-related disease, both, or neither? *Epidemiol Rev*. 2013;35(1):112–31.

71. Lapham K, Kvale MN, Lin J, et al. Automated assay of telomere length measurement and informatics for 100,000 subjects in the genetic epidemiology research on adult health and aging (GERA) cohort. *Genetics*. 2015;200(4): 1061–72.

72. Vasa-Nicotera M, Brouilette S, Mangino M, et al. Mapping of a major locus that determines telomere length in humans. *Am J Hum Genet*. 2005;76(1):147–51.

73. Walsh KM, Codd V, Smirnov IV, et al. Variants near TERT and TERC influencing telomere length are associated with high-grade glioma risk. *Nat Genet*. 2014;46(7):731–5.

74. Codd V, Mangino M, van der Harst P, et al. Common variants near TERC are associated with mean telomere length. *Nat Genet*. 2010;42(3):197–9.

75. Mirabello L, Yu K, Kraft P, et al. The association of telomere length and genetic variation in telomere biology genes. *Hum Mutat*. 2010;31(9):1050–8.

76. Shen Q, Zhang Z, Yu L, et al. Common variants near TERC are associated with leukocyte telomere length in the Chinese Han population. *Eur J Hum Genet*. 2011;19(6):721–3.

77. Pooley KA, Tyrer J, Shah M, et al. No association between TERT-CLPTM1L single nucleotide polymorphism rs401681 and mean telomere length or cancer risk. *Cancer Epidemiol Biomarkers Prev*. 2010;19(7):1862–5.

78. Shen J, Gammon MD, Wu H-C, et al. Multiple genetic variants in telomere pathway genes and breast cancer risk. *Cancer Epidemiol Biomarkers Prev*. 2010; 19(1):219–28.

79. Hirashima K, Migita T, Sato S, Muramatsu Y, Ishikawa Y, Seimiya H. Telomere length influences cancer cell differentiation *in vivo*. *Mol Cell Biol*. 2013;33(15):2988–95.

80. Shelton DN, Chang E, Whittier PS, Choi D, Funk WD. Microarray analysis of replicative senescence. *Curr Biol*. 1999;9(17):939–45.

81. Blasco MA. The epigenetic regulation of mammalian telomeres. *Nat Rev Genet*. 2007;8(4):299–309.

82. García-Cao M, O'Sullivan R, Peters AH, Jenuwein T, Blasco MA. Epigenetic regulation of telomere length in mammalian cells by the Suv39h1 and Suv39h2 histone methyltransferases. *Nat Genet*. 2004;36(1):94–9.

83. Cai N, Chang S, Li Y, et al. Molecular signatures of major depression. *Curr Biol*. 2015;25(9):1146–56.

84. Ceccarelli M, Barthel FP, Malta TM, et al. Molecular profiling reveals biologically discrete subsets and pathways of progression in diffuse glioma. *Cell*. 2016;164(3):550–63.

85. Sebastiani P, Montano M, Puca A, et al. RNA editing genes associated with extreme old age in humans and with lifespan in *C. elegans*. *PLoS One*. 2009;4(12):e8210.

86. Nersisyan L, Hakobyan A, Arakelyan A. Telomere-associated gene network in lung adenocarcinoma. *Eur Respir J*. 2015;46(suppl 59). Available at: http://erj.ersjournals.com/content/46/suppl_59/OA3493.abstract.

87. Komai K, Niwa Y, Sasazawa Y, Simizu S. Pirin regulates epithelial to mesenchymal transition independently of Bcl3-Slug signaling. *FEBS Lett*. 2015; 589(6):738–43.