Studies on TERRA : telomeric repeat containing RNA

Reelina Basu

2015

Reelina Basu. (2015). Studies on TERRA : telomeric repeat containing RNA. Doctoral thesis, Nanyang Technological University, Singapore.

https://hdl.handle.net/10356/62529

https://doi.org/10.32657/10356/62529

STUDIES ON TERRA: Telomeric Repeat Containing RNA

REELINA BASU

SCHOOL OF BIOLOGICAL SCIENCES

A thesis submitted to Nanyang Technological University in partial fulfillment of the requirement for the degree of Doctor of Philosophy

2015

ACKNOWLEDGMENT

One of the joys of completion is to look over the journey past and remember everyone who has helped and supported me along this long but fulfilling road. First and foremost, I would like to express my sincere gratitude to my supervisor, Assistant Professor Dr Zhang Li-Feng, who has been my mentor throughout the years of my graduate research life at Nanyang Technological University. He has inspired and encouraged me to treat my Ph.D with utmost dedication. He has also provided me with scientific advice, knowledge and insightful discussions about my research. A person with an amicable and positive disposition, Prof. Zhang always made himself available to clarify my doubts despite his busy schedules and I consider it a great opportunity to have done my doctoral program under his guidance and to have learnt from his research expertise. His constant support motivated me in following the right track to complete my mission.

I would also like to thank the members of my thesis advisory committee, Prof. Zbynek Bozdech and Prof. Koh Cheng Gee, who have provided precious technical guidance to me throughout my Ph.D. Their encouraging and constructive feedback has helped me immensely and I truly value their role in shaping my work. I extend my sincerest gratitude to Dr. Shao Fangwei and Zhenyu Meng from SPMS, NTU, for sharing their technical expertise in the amino labeled probe design project.

I gratefully acknowledge Chua Song Lin, Chelliah Rosi, Lai Lan-Tian, Hong Ru, Norbert, Ang Zhi Wei, Li Wei and all my other lab mates. I gained a lot of inspiration from them through our personal and scholarly interactions and their suggestions at various points of my research program. I am particularly indebted to Song Lin for his extensive help in completing the 'Telomere length measurement' project, as it would not have been possible without him. I would like to convey my special thanks to Lan-Tian for her assistance in the amino-labeled probe project. I am extremely grateful to Rosi for always showing me the brighter side of things and showering her unconditional love on me. I would also like to extend my gratitude to Luah Yen Hoon and Devendra from Prof. Zbynek Bozdech's lab for their help with malarial cell culture and slides.

A good support system is important for surviving and staying sane in grad school and I have been very lucky to find a wonderful group of friends who have always been beside me. Naren, Divya and Margaret, you guys have been my source of eternal sunshine through my thesis writing times. Words wouldn't be able to suffice my gratitude towards you, Naren. Thank you for being my rock, always, especially when I have been at my worst. You have added a special flavour to my journey here. Divya and Margaret, you both have been like my sisters. I wouldn't have been able to get through my darkest times without your help and unstinted support. Your friendship has made me a stronger person. My sincerest gratitude and love to both of you. I have been extremely lucky to have friends like Archana, Alolika, Abhishek, Margaret, Ravi, Siddharth, Ishani, Manisha and Manish who took time out of their busy schedules for helping me with proofreading and preparation for my Ph.D. Oral Examination. I would specially like to thank Hari, Aunty (his mom) and Archana for their support, inspiration and love. I would also convey my gratitude to Vinay, Siddharth, Mrinal, Nandhini, Manikantan, Durga, Laxmi and all my other friends for their company and encouragement. Thank you all for providing me with the support and friendship that I needed. Apart from my friends I have been extremely fortunate to have had the help of Salim and Shaheen, my lovely house owners, in this journey. You have treated me as your family and I would never be able to repay your love and kindness.

I gratefully acknowledge the funding sources that made my Ph.D. work possible. Nanyang Technological University has provided me the financial assistance through the years of my graduate life which has allowed me to realize my long cherished dream of obtaining a doctorate in life sciences. I am sincerely thankful to Dr. Zhang for extending financial support to me in my final year of Ph.D.

To end this acknowledgement I would like to express my heartfelt gratitude to my family. My hard-working parents have sacrificed their lives for my brother and myself and provided us with unconditional love and care. I thank my brother for all his advice and support. I would not have made it this far without the love and understanding of my family. Thank you for having faith in me and my dreams and for giving me the wings to fly to achieve them. I dedicate this thesis to my devoted parents, my brother and my loving grandparents especially to the memory of my grandfather, Dr. Ashok Shankar Mittra, whose role in my life was, and remains, immense.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

µg	microgram
µl	microliter
APS	Ammonium persulfate
ATP	Adenosine-5'-triphosphate
ATM-kinase	Ataxia telangiectasia mutated - kinase
BSA	Bovine serum albumin
cAMP	Cyclic adenosine monophosphte
Ci	Curie units
cm	centimetre
C-rich	Cytosine rich
CpG	Cytosine-phosphate-Guanine
CRISPR	Clustered regularly interspaced short palindromic repeats
CTCF	CCCTC-binding factor (protein)
CTCF	CCCTC-binding factor (RNA)
DEPC	Diethyl pyrocarbonate
DIC	Differential interference contrast
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl sulfoxide
DNA	Deoxy ribonucleic acid
DNase	Deoxyribonuclease
DNMT1	DNA (cytosine-5)-methyltransferase 1
DNMT3b	DNA (cytosine-5-)-methyltransferase 3 beta
ssDNA	Single stranded deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
dsRNA	Double-stranded ribonucleic acid

dUTP	Deoxyuridine triphosphate
EB	Ethidium bromide
EF	Embryonic fibroblasts
eq	Equivalence
EScell	Embryonic stem cell
EST1A/SMG6	Ever shorter telomeres 1A
Evf2	Embryonic ventral forebrain-1 RNA
G-rich	Guanine-rich
FBS	Fetal bovine serum
FISH	Fluorescence *in situ* hybridization
FITC	Fluorescein Isothiocyanate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
hr	Hour
k/d	knockdown
let -7	lethal-7 gene
lin-14	abnormal cell LINeage gene
LIF	Leukemia inhibitory factor
LINE	Long INterspersed Elements
lincRNA	Large intergenic non-coding RNA
lncRNA	long non-coding RNA
MEF	Mouse embryronic fibroblasts
mES	murine Embyonic Stem cells
min	Minutes
miRNA	micro RNA
mM	milli Molar
MOI	Multiplicity of infection
mRNA	messenger Ribonucleic acid

miRNA	micro RNA
ncRNA	non-coding RNA
NEAT1	Nuclear Enriched Abundant Transcript 1
NFAT	Nuclear factor of activated T cells
ng	nanogram
NMD	Nonsense mediated decay
NRON	ncRNA repressor of the nuclear factor of activated T cells
nt	nucleotide
PAGE	Polyacrylamide Gel Electrophoresis
PBS	Phosphate-buffered saline
PcG	Polycomb group proteins
Pf	*Plasmodium falciparum*
PNA	Peptide nucleic acid
PNK	Polynucleotide Kinase
POT1	Protection of telomeres protein 1
PTGS	Post transcriptional gene silencing
PRC	Polycomb Repressive Complex
Rad21	RAD21 homolog (*S. pombe*) (Protein)
Rad21	RAD21 homolog (*S. pombe*) (RNA)
RAP1	Ras-related protein 1
RBC	Red Blood Cell
Rb1	retinoblastoma protein 1
Rbl1	retinoblastoma like 1
Rbl2	retinoblastoma like 2
RISC	RNA induced silencing complex
RNA PolII	RNA PolymeraseII
RNA	Ribonucleic acid

RNase	Ribonuclease
RNasin	RNase inhibitor
RNP	Ribonucleoprotein
RT-PCR	Real time - polymerase chain reaction
rpm	Revolutions per minute
shRNA	small hairpin RNA
siRNA	small interfering RNA
SINE	Short INterspersed Elements
SMG1	Serine/threonine-protein kinase
snRNA	small nucleolar RNA
sncRNA	short noncoding RNA
snoRNA	small nucleolar RNA
SSC	Saline-sodium citrate
ssRNA	single stranded RNA
Suv39h1	Suppressor of variegation 3-9 homolog 1
Suv39h2	Suppressor of variegation 3-9 homolog 2
Suv4-20h1	Suppressor of variegation 4-20 homolog 1
TANK1	TRAF family member-associated NF-kappa-B activator 1
TANK2	TRAF family member-associated NF-kappa-B activator 2
TARE	Telomere associated repetitive element
TBE	Tris/Borate/EDTA
TEMED	Tetramethylethylenediamine
TERRA	Telomeric end repeat containing RNA (Human)
Terra	Telomeric end repeat containing RNA (Mouse)
Temp	Temperature
Tet	Tetracycline
TIN2	TRF1-interacting protein 2

TPP1	Tripeptidyl-peptidase 1
TRF1	Telomeric repeat-binding factor 1
TRF2	Telomeric repeat-binding factor 2
tRNA	transfer RNA
Xa	active X chromosome
Xi	inactive X chromosome
XCI	X chromosome inactivation
XIC	X inactivation centre
XIST	X-inactive specific transcript (human)
Xist	X-inactive specific transcript (mouse)
y*TERRA*	Yeast *TERRA*

ABSTRACT

To study *Terra* (telomeric repeat-containing RNA), an lncRNA transcribed from the telomeres, we first developed a robust RNA-DNA FISH method, with which we achieved direct observation on the co-localization pattern of *Terra* and telomeric DNA in the nuclear context. We show that not all the telomeres are associated with *Terra* and not all the *Terra* signals are overlapped with telomeres. Furthermore, the length of *Terra*-associated telomeres is not significantly longer than the length of those non-associated. These results do not lend support to the hypothesis that *Terra* is transcribed from the longer telomeres in a given cell and functions as a telomerase inhibitor to direct telomerase activity to the shorter telomeres.

We monitored the X chromosome telomere length change over a long-term cell culture of mouse ES cells. Taking advantage of the fact that one of the two ends of X chromosome is specifically associated with a bright and large *Terra* signal, we directly compared the telomere length of the two ends of X chromosome. Our observation, in compliance with the RNA-DNA results, does not support *Terra*'s role as a telomerase inhibitor *in vivo*. Interestingly, we found that X chromosome is under high pressure to maintain its telomere length. In a mutant ES cell line, in which telomerase activity is abolished, the X chromosome is one of the earliest chromosomes to show chromosome fusion, a sign of a chromosome suffering severe telomere length shortening. These observations are in fact consistent with a new hypothesis put forth by a recent live cell imaging study that *Terra* may recruit telomerase onto a telomere to maintain its length. With this hypothesis, our observations can be reasonably explained. X chromosome is under high pressure to maintain its telomere length; therefore *Terra* is transcribed at high levels on its telomere to recruit telomerase.

We used shRNA mediated knockdown to study roles of several transcription factors in regulating *Terra* expression in normal cells and in cells under heat shock. It is known that *Terra* expression is sharply up regulated in response to heat shock. We found that CTCF and Rad21 are involved in regulating *Terra* expression in normal cells. Interestingly, telomeres from different chromosome ends responded differently to the shRNA mediated knockdown of CTCF and Rad21. In addition, CTCF, Rad21

and NFκB, a master transcription factor involved in stress response, are not involved in regulating *Terra*'s expression in response to heat shock.

In order to study the functional roles of *Terra in viv*o, it is of utmost importance to establish a robust experimental system to artificially manipulate *Terra* expression *in vivo*. We attempted to knockdown *Terra* using shRNA and overexpress *Terra* using CRISPR/Cas9 system. However, we are not able to achieve satisfying results. We discuss our current results and suggest new directions for future research.

The subtelomere regions of the malaria parasite harbour an important class of *var* genes which provide them with antigenic variation, giving the study of presence of *Terra* in malaria parasite a special importance. In this study, we confirmed the presence of *Terra* in malaria parasite by Northern blot, RNA FISH and RNase protection assay.

CHAPTER 1:

INTRODUCTION

1. INTRODUCTION

The RNA world hypothesis suggests that life, as we know it today, had its beginnings in a self-replicating RNA molecule which could store genetic information and catalyze chemical reactions in primitive cells. During the course of evolution, under various selection pressures and with rising complexity at organismal levels, different functions of the RNA were delegated to other cellular molecules. The storage of genetic information was entrusted with more stable DNA, while the catalytic functions were transferred to more versatile proteins. Consequently, despite the fact that RNA played a critical role in a number of cellular processes as splicing and translation, its role was considered to be an intermediary between DNA and protein. This idea was summed up in the central dogma of molecular biology as 'DNA makes RNA makes protein' (Crick, 1970). Our understanding was that most of the information encoded in DNA specifying biological form and phenotype is expressed as proteins, which not only fulfill diverse catalytic and structural functions, but also regulate the activity of the biological systems in various ways. Hence, when the Human Genome Project was carried out to identify human genes almost 100,000-150,000 genes were expected to be present in the human genome (Collins et al., 1998). However, at the end of the project only ~20,000 genes could be identified. This showed that protein-coding sequences barely represented 2% of our whole genome. Based on the central dogma, the rest 98% of the genome was considered as junk and devoid of any important cellular function. Meanwhile RNA was classified into two categories (a) protein coding 'messenger RNA' (mRNA) and (b) non coding RNAs comprised of 'ribosomal RNA' (rRNA) and 'transfer RNA' (tRNA). However, recent high-throughput transcriptomic analyses revealed that eukaryotic genomes transcribe up to 90% of the genomic DNA (ENCODE Project Consortium, 2004), of which only 1-2% are protein coding transcripts while the vast majority are non-coding RNAs. The percentage of non-coding transcripts has grown considerably with the evolution of simpler life-forms to complex organism while the protein coding fraction remained almost constant (Mattick, 2004). These findings lead to speculations regarding a

significant contribution of the non-coding transcripts in the development of complex organisms and prompted more research in this direction (Amaral and Mattick, 2008).

1.1. Non-coding RNA

The term non-coding RNA (ncRNA) is commonly employed for functional RNA molecules that are not translated into proteins. Initially, most of the non-coding RNAs were thought to be involved in fulfilling relatively generic functions in cells, such as the rRNAs (ribosomal RNA) and tRNAs (transfer RNA) in protein translation, small nucleolar RNAs (snoRNAs) in modification of rRNAs. These constitutively expressed ncRNAs are now categorized as **infrastructural ncRNA**. Another category of ncRNAs includes the more recently discovered microRNAs (miRNA), piwi-interacting RNAs (piRNAs), small interfering RNAs (siRNAs) and long non-coding RNAs (lncRNA) (Ponting et al., 2009). They are grouped together as **regulatory ncRNAs**. These ncRNAs are transcribed from DNA sequences referred to as RNA gene. Bioinformatic and transcriptomic studies have revealed the existence of thousands of ncRNAs, although information about RNA genes is still sparse (Van Bakel et al., 2010; Wilusz et al., 2012). There is a lot of ongoing work concentrated on elucidating certain functional elements in the primary sequence of non-coding genes that define their roles (Wapinski and Chang, 2011).

Recent studies on ncRNAs have helped us towards gaining a progressive understanding of their biogenesis and functions. An important outcome from the work being done is the knowledge that the expression of many ncRNAs is regulated by diverse epigenetic, transcriptional and post-transcriptional mechanisms that affect biogenesis of primary transcript, post-transcriptional processing and maturation steps, as well as, changes in the conformation and trafficking of ncRNAs within and between the cell and its milieu. As a result, the non-coding transcriptome is now bestowed with an extraordinary diversity in function and response to environmental changes (Qureshi and Mehler, 2012). Examples of epigenetic and transcriptional regulators of the ncRNA expression include tumor protein 53 (TP53), repressor element 1-silencing transcription factor (REST), zinc finger protein 143 (ZNF143) and cAMP response

element-binding protein (CREB) (Anno et al., 2010; Huarte, 2010; Hung, 2011; Wu, 2006). ncRNAs are transcribed by specific RNA Polymerase enzymes (White, 2011). RNA Polymerase II (RNAP II) is responsible for synthesizing many types of non-coding RNAs (ncRNAs) including snRNAs and miRNAs. However, a range of essential ncRNAs for e.g. tRNAs, 5S rRNAs are also generated by Pol III (Dieci et al., 2007).

Various classes of ncRNAs can also be subjected to extensive post-transcriptional regulation (Faghihi, 2010; Hansen, 2011). Furthermore, a variety of post-transcriptional processes (such as non-templated modifications, alternative splicing, RNA editing, polyadenylation and 5′ capping) have the ability to modify ncRNAs (Jopling et al., 2005; Kapranov, 2007; Nishikura, 2010). Emerging evidence suggests that these and other post-transcriptional modifications may result in important functional consequences. For example, Adenosine-to-Inosine RNA editing can modify miRNA sequence and cause alterations in the interactions between a miRNA and its target mRNA. This modification also occurs in ncRNAs derived from retrotransposons (LINE-1 (L1) and *Alu* sequences), which form the largest fraction of substrates for primate RNA editing (Nishikura, 2010). Moreover, ncRNAs can undergo nuclear-cytoplasmic and nuclear-mitochondrial trafficking facilitated by ribonucleoprotein complexes. This process can be a critical factor for ensuring correct spatio-temporal distribution and ultimately function of mRNAs, ncRNAs and RNA-binding proteins (Mansfield, 2009). These observations imply that understanding of ncRNA life cycles is necessary to understand the regulation and development of different cellular systems.

1.1.1. Short non-coding RNAs

Non-coding RNAs that are shorter than 200 nt can be collectively referred to as short noncoding RNAs (sncRNAs) (Nagano and Fraser, 2011). These include small nuclear (sn) RNAs, 5S and 5.8S ribosomal RNAs (rRNAs), small nucleolar (sno) RNAs and transfer RNAs (tRNAs) (Matera Ag, 2007). More recently, a plethora of additional sncRNAs have been identified in eukaryotes including miRNAs, siRNAs,

piRNAs, etc (Ghildiyal and Zamore, 2009; Kim et al., 2009; Lee et al., 1993a; Wightman et al., 1993). Despite their small size, they have emerged as key players in regulating diverse cellular activities. Some of the known functions of sncRNAs include mRNA degradation, gene silencing, translational repression, post transcriptional modifications, genome stability and modulation of protein functions (Goodrich, 2006; Hüttenhofer et al., 2005; Maciej Szymañski, 2003; Matera Ag, 2007; Storz et al., 2005). Based on their mechanism of biogenesis and protein factors that they associate with, the sncRNAs are classified into:

- piwi-interacting RNAs (piRNAs)
- small interfering RNAs (siRNAs)
- microRNAs (miRNAs)

Another method of classification involves categorizing the sncRNAs based on their Dicer (RNaseA III family endoribonuclease specific for dsRNA) dependence during biogenesis into *Dicer-dependent* (miRNAs and siRNAs) and *Dicer-independent group* (piRNAs).

- ***Piwi interacting RNAs (piRNAs)***

piRNAs are Dicer-independent 26-30 nt long ssRNAs (single stranded RNAs), first discovered in the germline of fruit fly *D. melanogaster* using a genetic approach (Aravin et al., 2001; Aravin et al., 2003). To date, piRNAs have been additionally identified in vertebrates, nematodes, cnidarians, and poriferans (Aravin et al., 2001; Batista et al., 2008; Grimson, 2008; Grivna et al., 2006). These ncRNAs are produced from intergenic repetitive elements (e.g., retrotransposons) through several incompletely understood biogenesis pathways including an amplification mechanism called the *'ping-pong'* cycle (Kim, 2006). piRNAs have been shown to physically associate with the PIWI subfamily of Argonaute proteins, hence their name (Ghildiyal and Zamore, 2009; Saito et al., 2006; Thomson and Lin, 2009; Vagin et al., 2006). They possess phosphorylated 5' ends and a 2'-O-methyl (2'-O-me) modifications at their 3' ends (Aravin et al., 2008; Horwich et al., 2007; Saito et al., 2007; Vagin et al., 2006). piRNAs that are transcribed from transposon elements are involved in

suppressing transposon activity (function essential for germline integrity), whereas those derived from piRNA clusters are implicated in modulating gene expression (Brennecke et al., 2008; Girard and Hannon, 2008; Malone and Hannon, 2009; Xu et al., 2008a). Recent studies have also established the expression of piRNAs in somatic cells including neurons (Yan et al., 2011).

- ***Short interfering RNAs (siRNAs)***

siRNAs are a class of sncRNAs ~21 nt in length, found in a variety of organisms such as fungi (e.g, *Schizosaccharomyces pombe*), protozoa (e.g., *Trypanosoma*), plants (e.g., *Arabidopsis*), invertebrates (e.g., *Drosophila* and *Caenorhabditis elegans*) and vertebrates (e.g., human and mouse) (Ambros, 2004; Aravin et al., 2003; Reinhart et al., 2000). They are processed from longer double stranded RNA (dsRNA) precursors though repeated fragmentation steps catalyzed by endoribonuclease Dicer releasing double stranded (ds) siRNA intermediates from one of the dsRNA ends. These siRNA intermediates are composed of two complementary ~21 nt long strands that form a ~19 base pair duplex featuring 2-nt single-stranded overhang at each of its two 3' ends (Ghildiyal and Zamore, 2009). In *Drosophila*, dsRNA fragmentation is catalyzed by one of the two fly Dicer enzymes called Dicer2, whereas in mammals, only the Dicer protein is responsible for this reaction. One of the two strands of the double stranded siRNA precursor subsequently associates with an appropriate Argonaute (AGO) protein. This siRNA strand is known as the guide strand. The siRNA guide strand and the Argonaute protein form the core of RNA-induced silencing complex (RISC). RISC targets mRNA for silencing, commonly by degradation (Czech et al., 2008; Kawamura et al., 2008).

- ***MicroRNAs (miRNAs)***

miRNAs are 19-25 nt long single-stranded RNAs generated by RNase III-type enzymes from local hairpin structures embedded within endogenous transcripts (Ambros et al., 2003a; Ambros et al., 2003b; Chen, 2007). They are encoded in most

eukaryotic organisms and involved in regulation of a variety of essential and regulatory pathways, which include control of organ development, cell differentiation, proliferation and apoptosis (Dong et al., 2010; Hicks et al., 2009; Kim, 2005; Ucar, 2010). Human and mouse genomes encode hundreds of distinct miRNAs, each potentially targeting an extensive set of mRNA targets (Ambros, 2004; Bartel, 2004; Plasterk, 2006; Zamore and Haley, 2005).

MicroRNAs were discovered in 1993 by Victor Ambros, Rhonda Feinbaum, Gary Ruvkun and Rosalind Lee during their study on regulation of a transcription factor gene *lin-14* during larval development of the nematode *C. elegans* (Lee et al., 1993b; Wightman et al., 1993). They found that *lin-14* expression was regulated by the *lin-4* gene product known to control the timing of *C. elegans* larval development. However, in this study by the Ambros and Ruvkun teams, the fact that *lin-4* gives rise to short RNA products, instead of encoding for a protein, was strikingly surprising. It was also noticed that a 61-nt precursor RNA transcribed from the *lin-4* gene was processed to a 22-nt RNA that contained sequences, which, on base pairing in an antisense manner, were partially complementary to 7 cognate sites in the 3' UTR of the *lin-14* mRNA. This complementarity was found to be both necessary and sufficient to inhibit *lin-14* mRNA translation to the LIN-14 protein (Reinhart et al., 2000). Although, originally considered a distinctive feature of nematodes, in hindsight, the *lin-4* small RNA was the first known example of an extensive class of microRNAs expressed in most eukaryotic organisms (Lagos-Quintana et al., 2001; Lau et al., 2001; Lee and Ambros, 2001).

miRNAs function as guide molecules in post-transcriptional gene silencing (PTGS) by either repressing translation or/and destabilizing target mRNAs (Fvorstemann et al., 2005; Kim, 2005; Valencia-Sanchez et al., 2006). In plants, direct cleavage and degradation of the target mRNA occurs by complete or nearly complete complementary base pairing of miRNAs with their mRNA targets through a mechanism involving the '*Slicing activity'* of AGO protein (RNAi) machinery (Tang et al., 2003; Zhang, 2005). In contrast, most animal miRNAs interact in a partially complementary manner with the 3' untranslated regions (3' UTRs) of their targets. This normally leads to inhibition of protein synthesis or/and mRNA destabilization via slicer-independent mechanisms that preserves the stability of the mRNA target. Other

types of regulation, such as translational activation (Fazi and Nervi, 2008; Filipowicz, 2008; Inui, 2010; Majid, 2010) and heterochromatin formation (Gonzalez et al., 2008; Kim et al., 2008; Michael et al., 2003; Osada and Takahashi, 2007; Wu, 2010) have also been documented for miRNAs.

1.1.2. Long non-coding RNAs

The majority of ncRNAs belong to the group of long non-coding RNAs (lncRNA) that provisionally combines a wide range of non-protein coding transcripts with lengths of at least 200 nt (Ponting et al., 2009). lncRNAs have recently emerged as regulators of several important cellular processes (Huarte and Rinn, 2010; Wapinski and Chang, 2011). Although, clearly a heterogenous group, most lncRNAs share some common traits including nuclear localization, low expression levels, low level of interspecies sequence conservation and their occurrence in both poly A+ and poly A- forms (Consortium et al., 2005; Ewan Birney, 2007; Jopling et al., 2005; Ponting et al., 2009). They were initially described during the full-length sequencing of cDNA libraries in mouse (Okazaki, 2002). lncRNAs can be transcribed from intergenic regions (large intergenic ncRNAs, or lincRNAs) (Guttman, 2009; Khalil et al., 2009) in sense, antisense, bidirectional and overlapping orientations relative to the adjacent protein-coding genes. Other lncRNAs can be produced from genomic sequences encoding regulatory regions of protein-coding genes e.g. introns (Mercer et al., 2011), UTRs (Mercer et al., 2011), promoters (Tiffany Hung, 2011) and enhancers (Orom et al., 2010) as well as from specialized chromosomal regions e.g., telomeres; (Azzalin et al., 2007). Several lncRNAs have been shown to be derived from the mitochondrial genome (Rackham et al., 2011).

lncRNAs can be subject to post-transcriptional processing, including capping, polyadenylation, alternative splicing, RNA editing, as well as regulated intracellular trafficking (Carninci et al., 2005; Kapranov et al., 2007). Some lncRNAs can be involved in the regulation of the genomic loci from which they originate or of the neighboring ones. For example, a subset of lincRNAs have been shown to have enhancer-like activity (Orom et al., 2010). While other lncRNAs, such as HOX

transcript antisense RNA (*HOTAIR*) (Rinn et al., 2007) and X inactive-specific transcript (*XIST*) (Wutz, 2011), stimulate the formation of repressive chromatin environments across large regions in the genome and even entire chromosomes, respectively. Many lncRNAs show spatio-temporal patterns of expression, indicating that lncRNA expression is often tightly regulated (Dinger et al., 2008; Mercer et al., 2011). This might be beneficial given the broad range of lncRNA functions. Besides previously mentioned ones, it includes lncRNA roles in transcriptional and epigenetic control via recruitment of transcription factors and chromatin-modifying complexes to specific sites in the nucleus and genome. lncRNAs have also been observed to play roles in alternative splicing and other post-transcriptional RNA modifications, in translational control and nuclear-cytoplasmic shuttling (Wang and Chang, 2011). lncRNAs can also act as precursors for small ncRNAs, such as small nucleolar RNAs (snoRNAs) and miRNAs. Depending on their molecular functions lncRNAs are classified into 4 archetypes (Fig. 1.1):

Archetype I: as signals, lncRNA expression acts as a faithful indicator of the combinatorial actions of signaling pathways or transcription factors (colored ovals) reflecting gene regulation in time and space for e.g. *Xist*.

Archetype II: as decoys, titration of transcription factors and other proteins from chromatin or titration of protein factors into nuclear subdomains can be carried out by lncRNAs for e.g. *TERRA*, *MALAT1* etc.

Archetype III: as guides, recruitment of chromatin-modifying enzymes to target genes that are located either near the site of lncRNA production (i.e., in *cis*), or to distant target genes (i.e, in *trans*) is also done by lncRNAs like. *Xist* (acts in *cis*), HOTAIR (acts in *trans*) etc.

Archetype IV: as scaffolds, lncRNAs can form ribonucleoprotein complexes by bringing together multiple proteins. Histone modifications can be influenced by the lncRNA-RNP via their action on chromatin. In other instances, the structural properties of lncRNA scaffold leads to stabilization of nuclear structures or signaling complexes. An e.g. of archetype IV lncRNA is *TERC* acting as a scaffold in telomerase assembly.

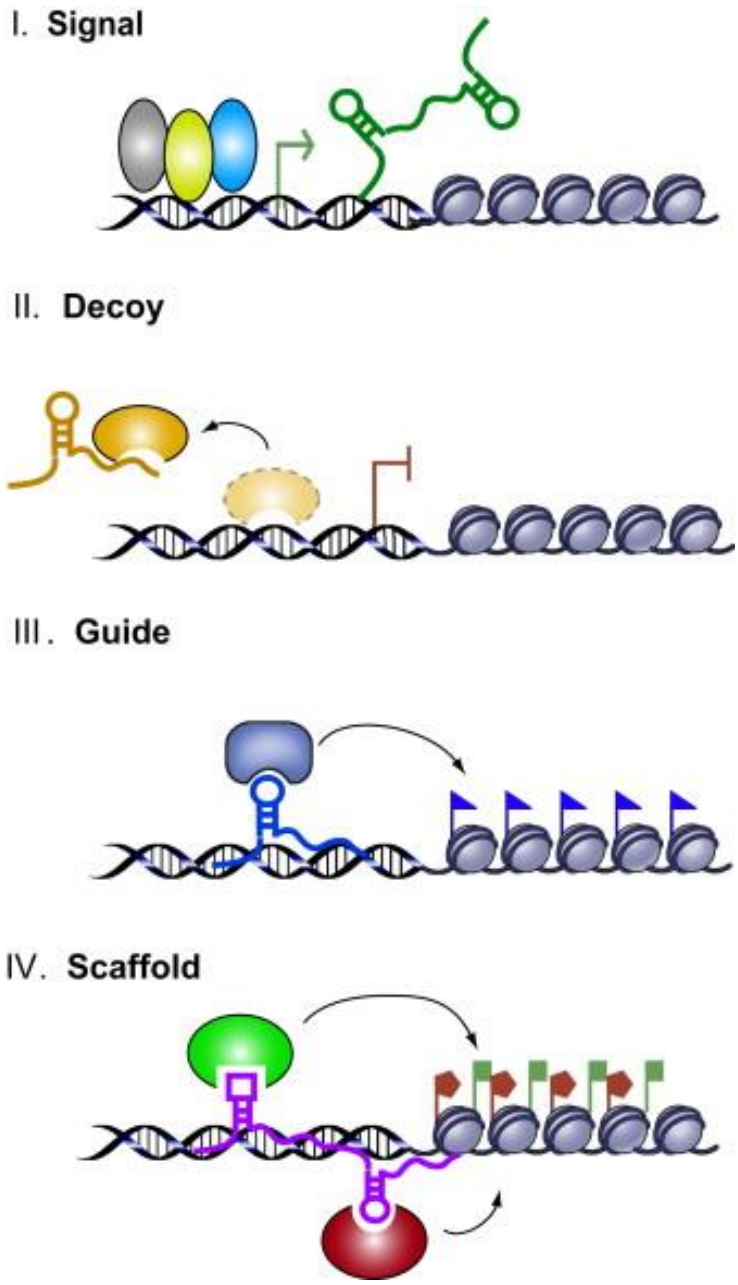


Figure 1.1. Diagram summarizing the four mechanistic archetypes of lncRNAs (Wang and Chang, 2011).

The broad functional repertoire of lncRNAs includes roles in higher-order chromosomal dynamics, transcriptional regulation and subcellular structural organization. These roles are discussed in brief in the following section.

- ***Transcriptional Regulation***

Many lncRNAs have been demonstrated as negative regulators of transcription. One such example is the regulation of dihydrofolate reductase (*DHFR*). The encoding gene has a major and a minor promoter. The lncRNAs are transcribed from the minor promoter and bind with the major promoter and transcription factor IIB, leading to dissociation of preinitiation complex (Martianov et al., 2007). The protein p53 is also known to activate expressions of many lncRNAs for e.g. lncRNA p21 which further regulate the expression of multiple downstream genes in conjunction with p53 (Huarte et al., 2010). While some other lncRNAs act as ligands for transcription factors and perform roles of coactivators like the Evf2 ncRNA (Feng et al., 2006).

- ***Subcellular Structural Organization***

In eukaryotes the nucleus is highly compartmentalized. It contains a number of membraneless subnuclear bodies. The information regarding assembly and association of these bodies is limited but they do seem to be associated with distinct lncRNAs. A prominent example is that of paraspeckles. These are cell cycle regulated nuclear foci, which depend on RNA for their structural integrity. Their functions include nuclear retention of a subset of mRNAs that have undergone Adenosine (A)-to-Inosine (I) editing (Chen and Carmichael, 2009; Chen et al., 2008; Prasanth et al., 2005). The lncRNA, NEAT1, plays a crucial role in paraspeckle structural integrity (Chen and Carmichael, 2009). Studies show that depletion of NEAT1 (Chen and Carmichael, 2009; Clemson et al., 2009) disrupted paraspeckle formation, while overexpression (Clemson et al., 2009) led to an increase in paraspeckle number, suggesting NEAT1 is an essential component of paraspeckle formation.

- ***Chromatin Re-modeling***

The lncRNAs have also been proposed to regulate transcription by recruiting chromatin-remodeling factors, which lead to epigenetic changes (Ponting et al., 2009). Epigenetics is referred to as the study of inheritable changes in phenotype and gene

expression caused by mechanisms other than the changes in DNA sequences. One of the most well described repressive complex which is involved in transcriptional and epigenetic regulation is the Polycomb-group (PcG) complex. PcG is comprised of two multiprotein complexes- polycomb repressive complex 1 (PRC1) and 2 (PRC2). PRC2 trimethylates histone 3 lysine 27 (H3K27me3) causing silencing of the chromatin region. PRC1 proteins, then bind to the H3K27me3 and ubiquitinate lysine 119 on histone 2A (H2AK119ub). The two complexes PRC1 and PRC2 also have RNA binding ability (Rinn et al., 2007). It is speculated that locus-specific silencing mediated by PcG might be guided by bound lncRNA. A classic example of this mechanism is X-chromosome inactivation. It has been shown that PcG complex binds to lncRNA *XIST* (Kohlmaier et al., 2004; Plath et al., 2003; Schoeftner et al., 2006). The coating of *Xist* RNA along the chromosome territory, in consequence, establishes multiple layers of epigenetic modifications on the silenced chromosome, such as H3K27me3 and H2AK119ub. Another example of a lnc RNA involved in chromatin remodeling is *TERRA* or telomeric end repeat containing RNA (Schoeftner and Blasco, 2008) . It originates from the subtelomeric region and contains variable numbers of telomeric repeats (UUAGGG). It is been speculated that a regulatory loop connecting telomere length, *TERRA* abundance and telomeric heterchromatinization may exist (Schoeftner and Blasco, 2008).

1.2. Telomeric Repeat containing RNA (*Terra*)

About 20 years before the discovery of DNA double helix structure by Watson and Crick, Hermann J. Muller showed that the ends of the linear eukaryotic chromosomes behaved differently from the rest of it. He exposed flies to ionizing conditions and obtained mutants with various chromosomal aberrations like deletions, inversions and translocations which encompassed different regions of the genome except the terminal ends. He summed up his findings by stating that even though chromosomes appear as homogeneous cytological entities, their extreme ends, which he termed telomeres ('telos' meaning 'end' and 'meros' for part), perform the unique function of sealing the chromosome (Muller, 1932).

Around 1940 Barbara McClintock provided further evidence on protective role of telomeres. She observed that when chromosomes were forcibly broken, it led to chromosome fusion, whereas chromosomes with intact telomeric ends never fused (Mcclintock, 1941). Nearing 1979, Elizabeth H. Blackburn isolated and defined the telomere sequences of *Tetrahymena* (Yao et al., 1981). In her work with Jack S. Szostak, she experimentally confirmed the previously hypothesized protective role of telomeres by McClintock and Muller, by showing that the telomeric repeats of *Tetrahymena* had a stabilizing effect on telomeres (Shampay et al., 1984; Szostak and Blackburn, 1982). In another couple of years, Blackburn and Carol W. Geider identified telomerase, a specialized enzyme for telomere elongation (Greider and Blackburn, 1987).

In the following years, research on telomere structure and function has come a long way, but to this date there are still many facts about telomeres that intrigue scientists. An astonishing revelation occurred in this telomere tale when the long standing belief that telomeres are transcriptionally silent was challenged by the discovery of non-coding RNA being transcribed from the ends of the chromosomes. These RNA molecules were termed *TE*lomere *R*epeat-containing *RNA* (*TERRA*). These RNA molecules were shown to be transcribed from the telomeres and associate with them (Azzalin et al., 2007; Schoeftner and Blasco, 2008; Zhang et al., 2009). In this section, we provide a brief discussion on telomere structure and function and

extensively review the current knowledge on *TERRA* biogenesis, regulation and potential functions.

1.2.1. Telomeres

Telomeres are specialized ribonucleoprotein structures that protect the natural ends of linear chromosomes. In the absence of telomeres, these ends would resemble DNA breaks and their repair would lead to deleterious chromosome fusions. Structurally, telomeres are composed of both DNA and protein components. The DNA component is composed of long tracts of double-stranded G-rich repeats, mostly ending in a protruding 3' single-stranded overhang from the G-rich strand called the G-overhang. The protein component, on the other hand, comprises of several factors that are involved in protection of the ssDNA and prevention of DNA ends being recognized as double-strand breaks (DSBs).

The telomere repeat sequences and protein components differ between organisms. In budding yeast, *Saccharomyces cerevisiae,* the DNA component consists of tandem arrays of short 5' to 3' G-rich repeats with the consensus sequence TG_{1-3} or $TG_{2-3}(TG)_{1-6}$, with a size ranging from 250-400 bp (Vega et al., 2003; Wang and Zakian, 1990). The double stranded repeats terminate in a single stranded (ss) G-rich 3' overhang, which ranges from 12-14 nt. The G-overhang is protected by binding of Cell division control protein 13 (Cdc13p) (Garvik et al., 1995; Lin and Zakian, 1996) In addition Cdc13p forms a multi-protein complex with two other ssDNA binding proteins (Stn1p and Ten1p), and is also involved in the maintenance of telomere integrity and length by regulating telomerase (Gao et al., 2007). In *Drosophila*, the telomere sequences consist of a mixed array of variably 5' truncated retrotransposons. The identified telomeric retrotansposons are *He*T-A, *T*ART, *T*AHRE (HTT). Multiple copies of HTT are present at the terminal ends of telomeres in *Drosophila* and the termini are capped by a multiprotein complex which can correctly recognize the ends even in the absence of retrotransposons (Rong, 2008). Telomere elongation in *Drosophila* takes place either by gene conversion or by retrotransposition as telomerase is absent. Telomeric DNA consists of TTTAGGG tandem repeats in most

plants. In *Arabidopsis thaliana*, telomere length ranges between 2 and 9 kilobases (kb), with a variation in length observed among the telomeres of the same cell as well as between different cell-types. (Richards and Ausubel, 1988). The G-overhang varies in length from 20-30 nt in *Arabidopsis* and is found only on half of the telomeres in seedlings and 35% of telomeres in leaves (Riha et al., 2001). Telomere replication in plants is taken care of by telomerase.

Telomeres in mammalian cells are made of TTAGGG repeats in tandem with a length ranging from around 5 kb to 50 kb in different chromosomes, cell types and organisms. The 3' length of the G-overhang can vary from 50 to 500 nt. A multiprotein complex called 'shelterin', caps the ends of the chromosomes in mammals and is involved in telomere length regulation. The shelterin complex is comprised of six proteins, consisting of the double stranded telomeric repeat binding factors TRF1 and TRF2, the TRF2 interacting factor RAP1, the bridging molecules TIN2 and TPP1 and the telomeric protection factor POT1, together covering the double and single stranded repeats (Palm and De Lange, 2008) (Fig. 1.2). POT1 binds specifically to the single stranded G-overhang and interacts with TPP1 for controlling the access of telomerase and protecting the chromosome ends from stimulating an ATR kinase signaling dependent DNA damage response (Denchi and De Lange, 2007; Hockemeyer et al., 2006; Hockemeyer et al., 2007; Palm et al., 2009; Wang et al., 2007; Wu et al., 2006). TRF1 and TRF2 bind to double stranded telomeric repeats and interact with POT1 via TPP1 (Baumann and Cech, 2001; Palm and De Lange, 2008; Ye et al., 2004). Both TRF1 and TRF2 are negative regulators of telomere length. TRF1 regulates telomere length by binding to TANK1 and TANK2 poly (ADP)-ribosylases (tankyrase1 and 2), which are positive telomere length regulators (Celli and De Lange, 2005; Karlseder et al., 2004; Munoz et al., 2005). TRF2, in association with RAP1, protects telomeres by inhibiting ATM-kinase pathway activation. TRF1 interacting nuclear factor 2 (TIN2) negatively regulates telomere length by preventing inhibition of TRF1 by tankyrase 1 (Karlseder et al., 2004; Ye and De Lange, 2004).

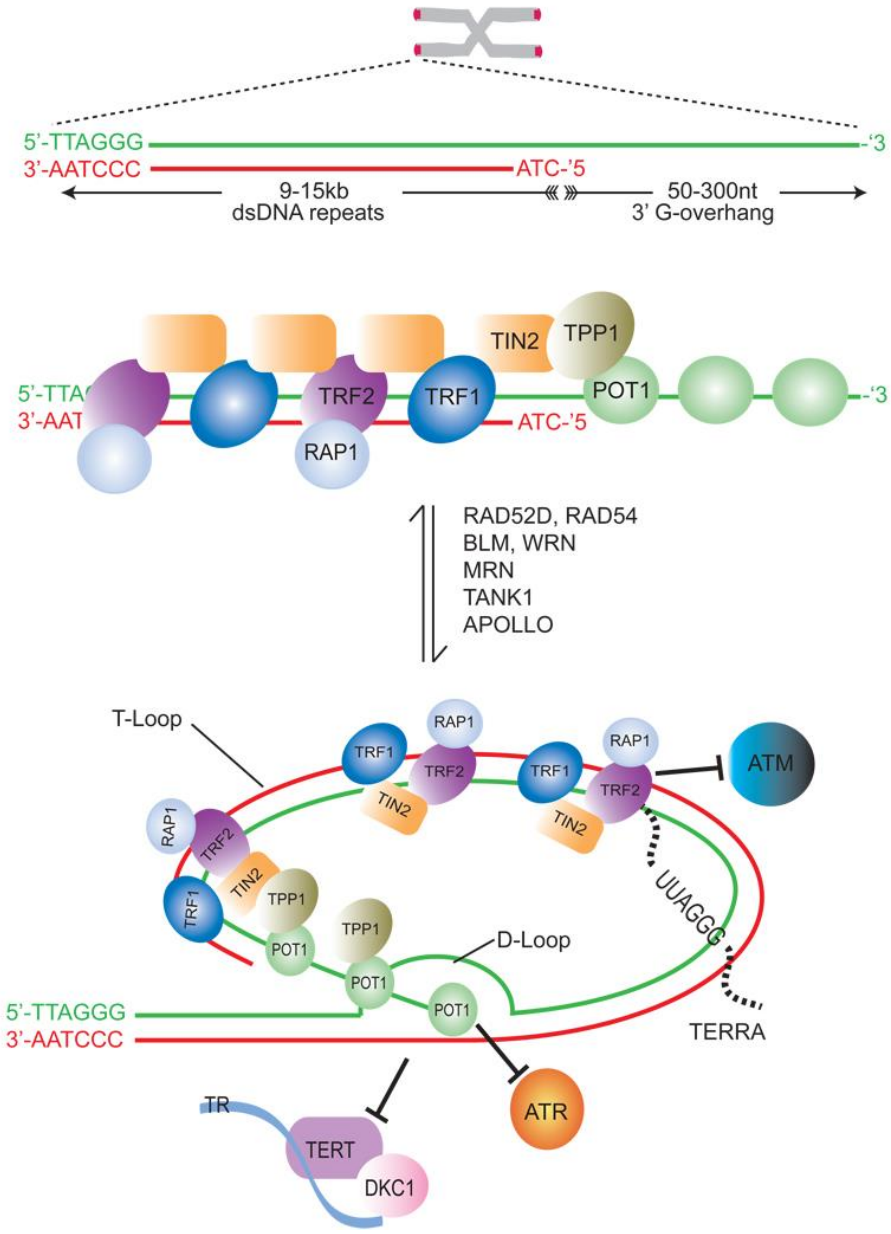


Figure 1.2. The structure of human telomeres (O'sullivan and Karlseder, 2010)

Telomere replication is more complicated than the rest of the chromosome and involves the intervention of a reverse transcriptase known as telomerase. Telomeres of eukaryotes shorten in length during each round of cell-division due to the *end replication problem*. (Fig. 1.3) DNA polymerases replicate DNA by adding nucleotides in a 5' to 3' direction to the 3'-hydroxyl end of a short RNA primer deposited by 'primase', a specialised RNA polymerase, at specific genomic loci. DNA replication does not start from the chromosomal ends, instead it is initiated at multiple

points along the chromosome and progresses smoothly on the leading strand where the DNA polymerase can move in the 5' to 3' direction. While on the lagging strand synthesis of DNA takes place in small stretches, each initiated from the RNA primer. This primer is degraded at the end of each replication cycle and replaced by DNA using the complementary strand as template. However, near the 3' terminus of the lagging strand exists a problem; removal of the RNA primer leaves a gap behind in the newly synthesized DNA due to the inability of DNA polymerase to replicate it. The 5' end containing strand is smaller than the 3' overhang containing strand, hence it cannot provide a template for it. This leads to the loss of sequences from the ends in each round of replication and is known as the *end replication problem*. Hence, an expected loss of sequence is observed in both the lagging and the leading strand during the replication cycle leading to loss of genetic material and an ultimate shortening of the telomeres (Bianchi and Shore, 2008; Watson, 1972). The cell has devised ways to tackle this loss of genetic material. One of them is through the enzyme *telomerase*. Telomerase consists of a catalytic protein subunit with reverse transcriptase function termed as telomerase reverse transcriptase (TERT) and a RNA moiety called telomerase RNA component (*TERC*). *TER*C comprises of a short sequence complimentary to the G-rich repeats of telomere. Telomerase is a RNA dependent polymerase involved in lengthening of the telomere ends using its RNA component as template for reverse transcription (Artandi and Depinho, 2010; Bianchi and Shore, 2008).

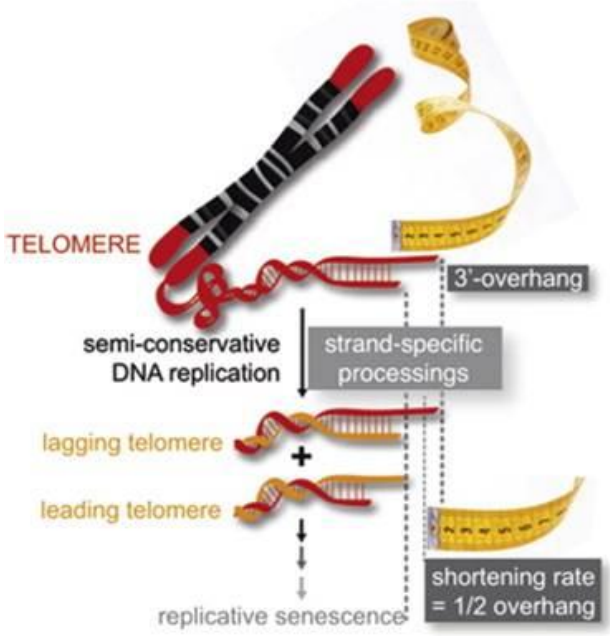


Figure 1.3. DNA end-replication problem in *Saccharomyces cerevisiae* (Teixeira et al., 2004)

Another common feature among the telomeres of mammals, yeast and flies is their ability to silence reporter genes inserted in the subtelomeric region by telomere position effect (TPE) (Baur et al., 2001; Koering et al., 2002; Mason et al., 2008; Rusche et al., 2002). The chromatin at mammalian telomeres is hypoacetylated and exhibits hallmark features of other heterochromatinised elements such as the pericentric repeats (Benetti et al., 2007; Michishita et al., 2008). Suv39h and Suv4-20h are the two main histone methyltransferases (HMTases) involved in the regulation of the chromatin structure at telomeres and subtelomeres. Suv39h1 and Suv39h2 cause enrichment of the histone 3 lysine 9 tri-methylation (H3K9me3) (Peters et al., 2001; Schotta et al., 2008). The H3K9me3 mark attracts binding of the heterochromatin protein 1 (HP1) to telomeres (Lachner et al., 2001; Nakayama et al., 2001). HP1 interacts with-and recruits Suv4-20h1 and Suv4-20h2 and they establish histone 4 lysine 20 trimethylation (H4K20me3). Another HMTase Dot1L mediates telomeric histone 3 lysine 79 trimethylation (H3K79me3), which enhances H4K20me3. Rb1, Rbl1 and Rbl2 belonging to the retinoblastoma (Rb) family of tumor repressive proteins interact with Suv4-20h HMTases to direct H4K20me3 to telomeric and centromeric repeats (Garcia-Cao et al., 2002; Gonzalo and Blasco, 2005; Lipinski and

Jacks, 1999) (Fig. 1.4). Chromatin compaction occurs by the interplay of all these modifications. These constitutive heterochromatin markers were seen to extend into the subtelomeres as well, giving rise to the general idea that the chromosome ends were transcriptionally silent.

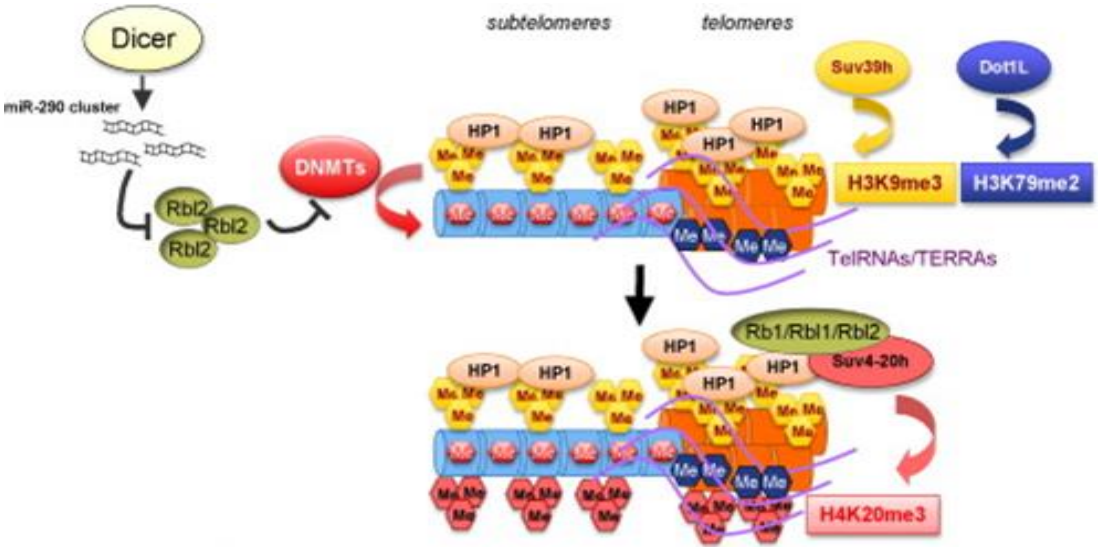


Figure 1.4. Assembly of mammalian telomeric and subtelomeric heterochromatin (Schoeftner and Blasco, 2010).

However recently, heterochromatic repetitive elements, such as mouse major satellite or human heterochromatic satellite III repeats were reported to give rise to non-coding transcripts (Jolly et al., 2004; Rizzi et al., 2004). In line with this, it was not surprising when telomeric C rich strand was found to be frequently transcribed by RNA Polymerase II in *M. musculus*, *H. sapiens*, *S. cerevisiae* and *D. rerio*, giving rise to UUAGGG-repeat containing non-coding RNAs (*TERRA* or *TelRNA*) that associate with telomeres (Azzalin et al., 2007; Schoeftner and Blasco, 2008; Zhang et al., 2009). In the next section we describe *TERRA* in detail.

1.2.2. Origin of *Terra*

The ends of the chromosome were believed to be transcriptionally silent for a very long time due to its heterochromatin nature and low gene density, until recent

discovery of *TERRA*. *TERRA* is a nuclear RNA which is actively transcribed from the telomeric sequences. Presence of *TERRA* has been demonstrated in humans, budding yeast, birds, rodents and zebra fish showing that this lncRNA is evolutionarily conserved (Azzalin et al., 2007; Luke et al., 2008; Schoeftner and Blasco, 2008; Solovei et al., 1994). *TERRA* is a nuclear RNA with length ranging from 100 to more than 9000 nt. Transcription of *Terra* takes place from the C-rich strand generating a RNA species composed of G-rich repeats (UUAGGG in mammals). Transcriptions from the G-rich strands have not been detected by standard techniques. This suggests that the transcription of *Terra* is unidirectional (Azzalin et al., 2007; Schoeftner and Blasco, 2008). Demonstration of RNA Polymerase II (RNAPII) in association with the telomeric and subtelomeric region and considerable cessation of *TERRA* levels in the presence of RNAPII blockers has led to the understanding that *TERRA* is transcribed mainly by RNAPII (Azzalin and Lingner, 2008; Schoeftner and Blasco, 2008). The role of other RNAPs in *TERRA* transcription is not ruled out, as *TERRA* molecules are detectable even on prolonged treatment with α-amanitin, an RNAPII inhibitor. In mammals, *in vivo* interaction of RNAPII with telomeres and TRF1 has been demonstrated (Schoeftner and Blasco, 2008). Another feature of *TERRA* that links it to RNAPII dependent transcription is the fact that about 7% of *TERRA* is 3' polyadenylated and polyadenylation is a hallmark of RNAPII transcription (Azzalin and Lingner, 2008; Schoeftner and Blasco, 2008). A canonical polyadenylation signal is absent in the UUAGGG repeat sequence of *TERRA* making us wonder which factors stimulate polyadenylation of *TERRA*.

Studies on *S. cerevisiae* have shown that the loss of polyadenylation has led to disappearance of *TERRA*, prompting speculations regarding a role played by the poly-A tail in *TERRA* stabilization (Luke et al., 2008). Analysis of *TERRA* sequences have demonstrated that *TERRA* comprises of both subtelomeric and telomeric sequences. We discuss more on *Terra* transcription in the next section. *TERRA* is capable of adopting a four-stranded structure referred to as parallel G-quadruplex, which derives from non-canonical Hoogsteen G-G base pairs (Xu et al., 2010).

1.2.3. Transcription of *Terra*

Transcription of *TERRA* has been hypothesized to originate in the subtelomeric region as a number of RT-PCR and Northern blot experiments (Azzalin et al., 2007; Zhang et al., 2009) have demonstrated the presence of both subtelomere and telomere sequences in individual *TERRA* molecules in humans. Confirmation of this hypothesis was provided by the discovery of promoter regions in the subtelomere of a number of human chromosomes which were committed to *TERRA* transcription.

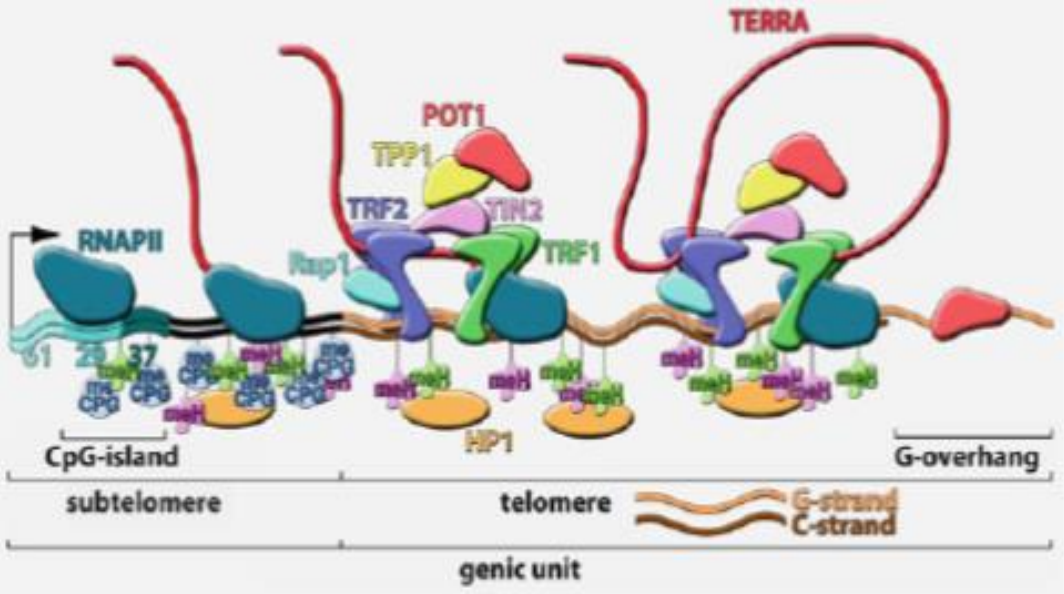


Figure 1.5. *Terra* transcription from human telomeres (Ugarkovic, 2011).

These subtelomeric promoters were found to be located about 250 bp away from the subtelomere to telomere transition. They were comprised of three repetitive tracts of DNA. Among them, the one towards the centromere is composed of 61 bp tandemly repeated units, the central tract constituted of 29 bp repeats and the third telomere proximal tract had tandemly repeated 37 bp units. They were located immediately upstream of the transcription start sites of several *TERRA* molecules and were cumulatively referred to as ’61-29-37’ repeats (Nergadze et al., 2009). Out of the three repeats, the 37 bp and 29 bp repeats exhibited high CpG dinucleotide content, a feature common to a number of RNAPII associated promoter regions. Further studies by Nergadze et al. confirmed the association of RNAPII with *TERRA* promoter *in vivo*

(Nergadze et al., 2009) (Fig. 1.5). DNA FISH and nucleotide sequence BLAST have helped us localize the 61-29-37 repeats at 1p, 2p, 3q, 4p, 5p, 6p, 8p, 9p, 9q, 10q, 11p, 12p, 15q, 16p, 17p, 19p, 20p, 21q, Xq, and Yq subtelomeres in human chromosomes. Studies have shown that 11q, Xp/Yp subtelomeres also transcribe *TERRA* molecules (Azzalin and Lingner, 2008) even though they lack the repeats, implying the involvement of other promoters or regulatory mechanisms in *TERRA* biogenesis.

Another interesting feature of human *TERRA* promoters is their association with CTCF and cohesion (Deng et al., 2012b). Colocalization studies have shown the presence of RNAPII, cohesion and CTCF at the '61-29-37' promoters of *TERRA* and downregulation of CTCF and cohesion molecule Rad21 led to a decrease in *TERRA* expression. Figure 1.6 demonstrates the enrichment profile peaks of CTCF, Rad21 and RNAPII at the 61-29-37 promoter region. The CTCF and Rad21 peak is positioned over the 61-mer repeat, centromeric to the CpG island. RNAPII peak is centered over the 29-mer repeat and the CpG island. These observations suggest that CTCF and Rad21 are activators of *TERRA* transcription originating in human chromosomes from the known promoter site (Deng et al., 2012b). Fibroblasts derived from Immunodeficiency, Centromeric region instability, Facial anomalies (ICF) syndrome patients, which suffer from a loss of DNMT3b activity, display highly increased *TERRA* levels, which correlate with a loss of subtelomeric CpG methylation (Yehezkel et al., 2008). This indicates DNMT3b controls *TERRA* transcription by exerting its repressive effect at the promoter. Furthermore, Cockayne Syndrome group B protein (CSB) has been implicated in *TERRA* transcriptional regulation at a subset of human telomeres (Batenburg et al., 2012).

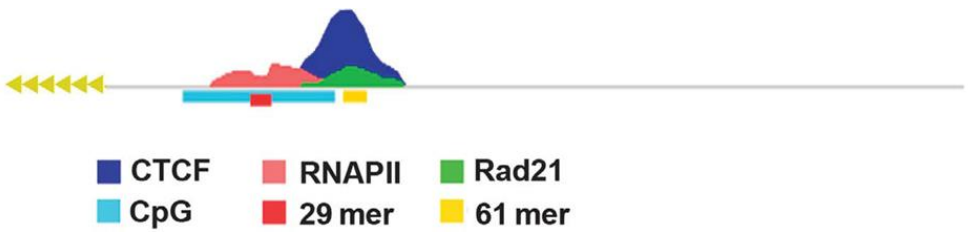


Figure.1.6. Enrichment profile model for CTCF, Rad21 and RNAPII at *TERRA* promoters (Deng et al., 2012b).

1.2.4. *Terra* localization and regulation

TERRA has been found to be exclusively present in the nuclear compartment implying its functions are restricted as a lncRNA in the nucleus. A recent study by Zhang et al. demonstrated that *Terra* localized near both sex chromosomes in mouse embryonic stem (ES) cells while, in differentiated fibroblasts cells it was associated with the heterochromatic sex chromosome [inactive X (Xi) in females and Y chromosomes in male] (Zhang et al., 2009). ES cells had two major *Terra* pinpoint signals in RNA FISH associated with the both the sex chromosomes in male and female cells, while in fibroblasts there was one major pinpoint signal associated with the sex chromosome that was transcriptionally less active (Xi in female and Y in male). This expression pattern suggests that *Terra* may have a functional relationship with the sex chromosomes in mouse cells, for e.g. a role in dosage compensation by X inactivation in female sex chromosomes. Another observation made in this study was the association of *Terra* with the distal telomeric end of mouse chromosomes (Zhang et al., 2009). Mouse chromosomes are telocentric i.e. they have centromere at one end of the chromosomes, with a proximal or centromere associated telomere and a distal telomere present at the other end of the chromosome.

In human cells, Porro et al. reported that *TERRA* can be divided into two subfractions: a 3' polydenylated fraction and a non-polyadenylated fraction (Porro et al., 2010). The differently polyadenylated fractions were localized in distinct areas in the nucleus. The non-polyadenylated fraction was seen to associate with chromatin whereas poly A+ fraction was mostly nucleoplasmic. This observation was explained by suggesting that, perhaps the non polyadenylated fraction was unable to detach from the telomeres or the polyadenylated fraction associates transiently with telomeres (Porro et al., 2010). Subsequent microscopic colocalization experiments using 3D reconstruction strategies showed that only about half of the *TERRA* foci colocalized with telomeres (chromatin associated *TERRA*) in human cells whereas the other half was not telomere associated (free *TERRA*) (Le et al., 2013). Consistent with previous reports (Azzalin et al., 2007), *TERRA* foci were detected at about 25% of telomeres (Le et al., 2013).

Association of *TERRA* with the telomeres in human cells is by the virtue of its binding with various protein factors. Pull-down experiments have identified shelterin components TRF1 and TRF2, telomerase associated Dyskerin, DNA PKc (a kinase that is involved in capping of leading strand of telomeres) and a number of heterochromatic proteins as *TERRA* interacting partners in humans (Deng et al., 2009; Scheibe et al., 2013). The interaction with TRF2 might require the formation of an intramolecular G-quadruplex by *TERRA* (Biffi et al., 2012; Xu et al., 2010). In addition, intermolecular G-quadruplexes formed by *TERRA* and the telomeric G-rich strand might contribute to *TERRA*'s localization to telomeres (Xu et al., 2008b).

Expression of *TERRA* has been found to be regulated over the cell-cycle. High levels of *TERRA* have been reported in G1/early S phase (Porro et al., 2010). As the cells progress from S phase into G2 phase, a decrease in *TERRA* levels occur, which then gradually increase through the mitotic phase into the G1 phase. The regulation of *TERRA* levels through cell cycle has been correlated with its role at the telomere.

TERRA expression regulation takes place directly at the level of promoters by DNA methyltransferases. The 61-29-37 promoter is heavily methylated by DNMT1 and DNMT3b, indicating a repressive effect on *TERRA* transcription (Nergadze et al., 2009). The nonsense mediated RNA decay (NMD) effectors viz, EST1A/SMG6, SMG1 and UPF1 also repress *TERRA* levels in human cells. In the absence of these NMD proteins, *TERRA* was reported to dramatically accumulate at chromosome ends although total *TERRA* output and turnover rate was not affected (Azzalin et al., 2007). In budding yeast, *yTERRA* is rapidly degraded by 5'to 3' nuclear exonuclease, Rat1 (human Xrn2) (Luke et al., 2008). Histone deacetylase Sir2 represses *yTERRA* levels. Other *yTERRA* repressive factors include telomeric binding proteins Rif1, Rap1 and Rif2 (Iglesias et al., 2011).

Multiple levels of negative regulation of *TERRA* expression is observed by factors impacting its localization, transcription and turnover. This control mechanism could be due to the role played by *TERRA* in cells. We discuss more on *TERRA* function in Section 1.2.6.

1.2.5. *Terra* and cellular stress

TERRA expression was reported to respond to cellular stress and genotoxicity. When *Xist*'s antisense *Tsix* was deleted in mouse cells, it lost its ability to count the number of X chromosomes and this led to the appearance of cells with abnormal Xi (X inactive) (Lee, 2002, 2005). In almost 64% of these cells *Terra* speckling was observed. *Terra* upregulation was also observed in response to gamma radiation in mouse fibroblasts (Zhang et al., 2009). Even Dicer deficiency leads to *Terra* upregulation in the cells (Zhang et al., 2009). ICF (Immunodeficiency, centromeric region instability, facial anomalies) syndrome, an autosomal recessive disease, caused due to a mutation in DNMT3b leads to short telomeres and increased *TERRA* transcription (Yehezkel et al., 2008). *TERRA* expression is altered in cancer cells with most cell-lines showing some degree of speckling (Zhang et al., 2009). When TERRA expression was studied by RNA-FISH in a primary human lung fibroblast cell-line and in HeLa cells, approximately 30% of the cells had 3 to 7 foci, whereas murine renal cancer cells and human osteosarcoma cells exhibited 20 to 40 foci in almost 80 to 100% of the cells (Azzalin et al., 2007).

1.2.6. *Terra* functions

A fraction of *TERRA* is telomere associated and appears to function in maintaining telomere structure and facilitating telomere replication (Deng et al., 2009). Long before naturally occurring telomeric transcripts were first identified, placement of a strong promoter upstream of a yeast telomere suggested that telomere transcription inhibits their elongation (Sandell et al., 1994). Most of the known regulators of *TERRA* have been found to negatively affect its expression by repressing it at transcription level, being involved in its degradation or its displacement. A logical conclusion from these studies was that, high level of *TERRA* transcription could have adverse effects on the telomere. Biochemical experiments later confirmed that telomeric RNA oligonucleotides bind to and potently inhibit human telomerase (Redon et al., 2010). Further support for *TERRA* inhibiting telomerase came from the observation that *yTERRA* levels are low in wild-type budding yeast but are

dramatically increased in cells defective in the Rat1 exonuclease (Luke et al., 2008). These cells have short telomeres, a phenotype that can be rescued by exogenous expression of RNaseH, indicating that *TERRA* could also inhibit telomerase via a DNA/RNA hybrid formation. In contrast, *TERRA* upregulation in human cancer cells did not impair telomere elongation (Farnung et al., 2012). Instead, telomere elongation repressed *TERRA* expression through increased H3K9 trimethylation suggesting that telomere shortening is associated with increased transcription (Arnoult et al., 2012). This idea of *TERRA* inhibiting telomerase has recently been challenged by *in vivo* assays using live-cell imaging and transcriptionally induced telomeres (Cusanelli et al., 2013; Farnung et al., 2012; Maicher et al., 2012). By utilizing transcriptionally inducible telomere in human cells having telomerase, a negative effect of heightened *TERRA* transcription on telomere length of the induced chromosome or the neighboring ones was absent. There was no adverse effect on telomerase activity or telomerase binding ability in these cells (Arora et al., 2012). Recent live-cell imaging studies in yeast also challenged the idea of *yTERRA* having a negative effect on telomerase binding. They concluded from their studies that *yTERRA* could help in recruiting telomerase at a short chromosome by interacting with telomerase and guiding it to the telomere of origin (Cusanelli et al., 2013). Thus, functional roles of *TERRA* in cells have not been defined with clarity and more studies on *TERRA* function need to be conducted.

1.3. Malaria

The word 'Malaria' originated from the Latin words, *Mala*-Bad and *aria*-air. It is one of the most prevalent parasitic infections in the world. The disease is caused by the protozoan pathogen ***Plasmodium***, which belongs to the phylum *Apicomplexa,* and is transmitted by blood-feeding female Anopheles mosquitoes. The genus *Plasmodium* has more than 200 species that are parasitic to reptiles, birds and mammals.

Human malaria is one of the world's most lethal infections, affecting 300 million people and killing 1-2 million people every year. It is mainly caused by five species of *Plasmodium, P. flaciparum, P. vivax, P. ovale, P. malariae* and *P. knowlesi.* Among these, the most lethal and virulent form of human malaria is caused by *P. falciparum*. The life cycle of the parasite alternates between the human host and the insect vector. Extensive research has been done to study the mechanism of malarial infection. Sir Ronald Ross was among the pioneers in malarial research field. He demonstrated that mosquito transmits malaria in 1897. Sir Charles Louis Alphonse Laveran demonstrated the presence of pigmented crescent shaped *Plasmodium* parasite in human blood. He received a Nobel Prize for his work in 1907. The infection of human erythrocytes is ultimately responsible for all the clinical pathologies associated with the disease.

1.3.1. Life-cycle of *Plasmodium*

P. falciparum sporozoites are transferred directly into the human bloodstream by female *Anopheles* mosquitoes during a blood meal. Sporozoites circulate briefly (i.e., minutes) in human blood before invading hepatocytes, where an asexual life cycle starts. After 10–12 days, infected hepatocytes rupture, releasing thousands of daughter merozoites back into the blood. They invade circulating erythrocytes and begin intra-erythrocytic cycle of asexual replication during the following 48 hours. *P. falciparum* differs from other human malarial species as the parasitized erythrocytes do not remain in the circulating blood for their entire life cycle. After 24–32 hours, when young parasites mature from the ring to the trophozoite stage, parasitized erythrocytes adhere to endothelial cells in the microcirculation of various organs (termed

sequestration). Trophozoites mature into schizonts, which eventually rupture and release 16–32 daughter merozoites that invade fresh erythrocytes to perpetuate the asexual life cycle. Some parasites inside red blood cells differentiate (in response to stress or other cues) into male and female gametocytes. Upon ingestion by a feeding female mosquito, they fuse, undergo complex sexual development and eventually form infective sporozoites that can be introduced into the human host at the next blood meal, thereby ensuring the continuation of the parasite life cycle (Fig. 1.7).

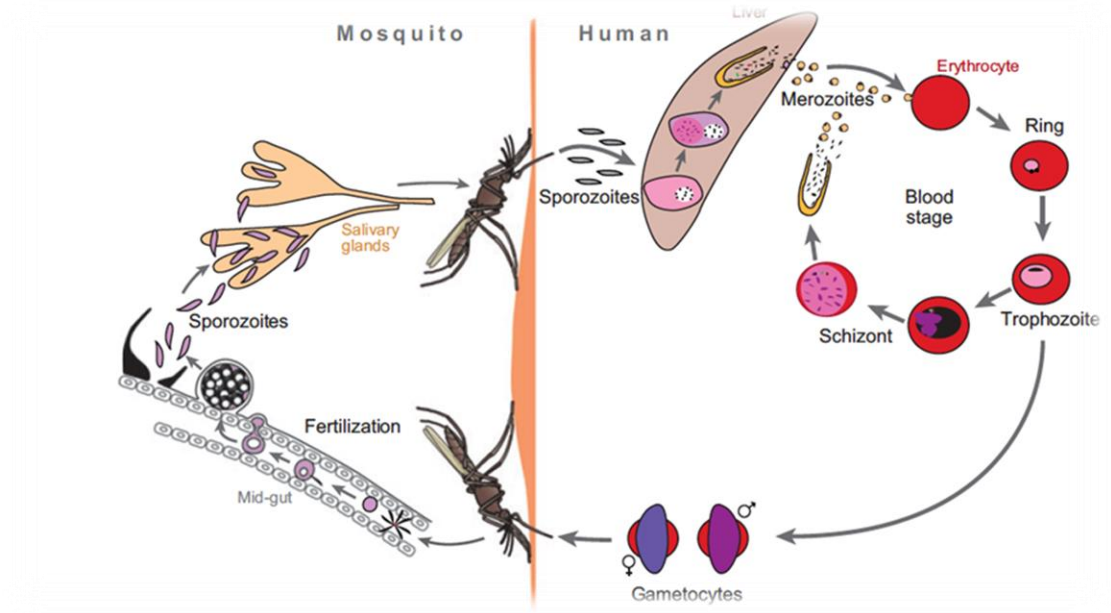


Figure 1.7. The life cycle of human malarial parasite *Plasmodium falciparum* (Scherf et al., 2008a).

1.3.2. Chromosomal organisation of *Plasmodium*

The haploid nuclear genome of *P. falciparum* is ~30Mb in size and is extremely AT rich (~82% AT). It is organized into 14 linear chromosomes varying from 0.7 to 3.4Mb. The chromosomal ends are guarded by telomeres. Telomeric DNA has been cloned from several malarial species and been found to consist of degenerate G-rich tandem repeats, with **GGGTT(T/C)A** as the most frequently repeated sequence. The mean length of the telomeric array shows significant interspecies variation. It is about 1.2kb in *P. falciparum*, while in *P. vivax* it is about 6.7kb (Figueiredo et al., 2000)

1.3.3. Evasion of host immune system

Vaccines against malarial infections are yet to be discovered because of the parasites' ability to evade host responses. Infections caused by *Plasmodium falciparum* are persistent and recurrent as the parasite is able to undergo antigenic variation and evade attack by the host immune system. Antigenic variation is the coordinated expression of variant surface molecules, which diminishes the immune clearance of the parasite by the host and allows the establishment of prolonged chronic infections with successive waves of parasitemia. *P. falciparum* parasites generate high levels of variability in gene families that comprise virulence determinants, such as the *var* genes (Scherf et al., 2008a). These genes encode the major variable parasite protein (*P. falciparum* erythrocyte membrane protein-1), and are expressed in a mutually exclusive manner at the surface of the erythrocyte infected by *P. falciparum.* There are 40–50 *var* genes per haploid genome, which are predominantly situated in the subtelomeric regions of all 14 chromosomes. FISH analysis of chromosomal ends have shown that telomeres physically associate to form clusters and are visualised as four to seven spots near the nuclear-periphery of the sexual and asexual blood stage parasite. These clusters include regions 30 to 40 kb beyond the telomeric repeats, including the subtelomeric region that encodes the variant-surface-antigen families. DNA sequences in this region, for e.g. the *var* genes, undergo ectopic (non-allelic) recombination at much higher rate than expected for homologous recombination (Freitas-Junior et al., 2000). A possible reason for this occurrence could be that the physical alignment of the heterologous chromosome ends brings homologous sequences from different chromosomes together, allowing efficient DNA recombination (Su et al., 1999). *var* gene sequences undergo recombination at frequencies much higher than those expected from homologous crossover events alone, leading to gene conversion and promotes the diversity of antigenic phenotypes. This decreases the ability of the host immune system to attack the parasite (Su et al., 1999).

1.3.4. lncRNAs in *Plasmodium*

The lack of sequence-specific *cis* regulating agents, transcription factors as well as the absence of DNA methylation machinery and RNA interference in the parasite has confounded scientists for a long time (Baum et al., 2009; Cui and Miao, 2010). The recent discovery of an expanded family of 27 apicomplexan apetela 2 (ApiAP2) transcription factors partially explained the regulation of parasite genome (Balaji et al., 2005; Campbell et al., 2010). Additionally, the important role of chromatin remodeling and epigenetic memory in blood stage-specific expression and antigenic variation of virulence genes is becoming highly evident (Chaal et al., 2010; Merrick and Duraisingh, 2010; Scherf et al., 2008b). Interestingly, while the parasite lacks many of the conventional regulatory mechanisms of other organisms, it has a repertoire of conserved histone modifying enzymes, and a higher than average number of RNA-binding proteins (Coulson et al., 2004; Cui and Miao, 2010).

The role of ncRNAs in epigenetic regulations has been established in eukaryotes ranging from yeasts to humans. The long- and short non-coding RNAs interact with different RNA binding proteins and chromatin remodeling complexes to modulate their target genomic loci, for instance interactions of *TERRA* with various telomere binding proteins which are important for telomere length regulation and maintenance. A DNA tiling array designed to investigate putative regulatory roles of lncRNAs in *Plasmodium* identified and characterized 60 lncRNA candidates based on their evolutionary conservation, G+C content, correlation with neighbouring genes and expression profile (Broadbent et al., 2011). Further investigations led to the discovery of a multigene family of lncRNA TAREs. Their loci have been mapped on 22 out of 28 *P. falciparum* chromosome ends. Another interesting observation on these lines was, the presence of an upstream sequence type B (upsB-type) *var* virulence gene adjacent to each of the lncRNA TARE genes and an enrichment of transcription factor binding sites near lncRNA-TARE genes which suggests the possibility of transcription from telomere in the parasite. The telomere associated RNA could interface with chromatin reading and writing factors and modulate the precise epigenetic regulation of nearby subtelomeric *var* loci.

CHAPTER 2: MATERIALS AND METHODS

2. MATERIALS AND METHODS

2.1. Cell lines

Mouse embryonic stem cells (ES) were cultured in embryonic stem cell media. Two types of male undifferentiated mouse embryonic stem cells were used for telomere length assay: J1, the wild type, male cell-line and F19, an early generation male *Terc* knockout (*Terc*-/-) cell line. ES cells were grown in the presence of a layer of irradiated male mouse embryonic fibroblasts (IMF) feeder cells in T25 flasks.

2.2. Generation of irradiated mouse embryonic fibroblast (IMF) feeder cells

Mouse embryonic fibroblasts (MEF) were extracted from day 13.5 DR4 mouse embryos. Fibroblasts were cultured and expanded in embryonic fibroblast (EF) medium (Dulbecco's Modified Eagle's medium (DMEM) (HyClone, #12100-061), 7.5% $NaHCO_3$ (Sigma, #S6014), HEPES (GIBCO, #15630-056), 100X non-essential amino acids (GIBCO, #11140-050), 100 mM glutamine (GIBCO, #35050061), penicillin-streptomycin (GIBCO, #15140-122), 2-mercaptoethanol (Sigma, #M6250), research grade fetal bovine serum (FBS) (HyClone, #SH3007103). MEF cells were harvested by trypsinization, irradiated at 30 Gray, and slow-frozen to -80°C after the addition of an equal volume of freezing medium, at $4x10^6$ cells per cryovial.

2.3. ES culture media

5X Dulbecco Modified Eagle's Medium (5X DMEM), 7.5% $NaHCO_3$, 1 M HEPES, 100X non-essential amino acids, 100 mM glutamine, 100X penicillin-streptomycin, 2-mercaptoethanol, Characterized FBS and Leukemia inhibitory factor (LIF) at final concentration of 500 U/ml (Millipore #ESG1107).

2.4. ES cell culture

Mouse embryonic stem (ES) cells were cultured over a layer of irradiated male fibroblasts (IMF) (Section 2.2). IMFs were seeded at least one day prior to seeding ES cells. Two hours before seeding ES cells, fresh media was added to the flasks containing IMF cells for conditioning the media. After media conditioning, ES cells were seeded. The cells were grown to ~80% confluence before proceeding to experiments.

2.5. Cytospin (Fixation of interphase cells on slides)

ES cells were cultured in T25 flasks till about 80% confluence. The cells were trypsinized and washed with 1X PBS. To prepare cytospin slides, ES cells at the concentration of 0.7 x 10^6 cells per ml were used and 100 µl of cell suspension containing approximately 700,000 cells was plated on each slide. Slides were made using Cytocentrifuge at the speed of 1,500 rpm for 10 minutes followed by washing in ice-cold 1X PBS for 5 minutes at 60 rpm on a shaker (OS-20 Orbital Shaker). The slides were then incubated in 4% Paraformaldehye (PFA), (Sigma Aldrich, #P6148) for 10 minutes and stored in 70% ethanol at 4°C for later use.

2.6. Metaphase chromosome spread

The ES cell culture was grown till confluence. ES cells were treated with 0.3 µg/ml Colcemid (GIBCO™, Kary MAX®, #15212-012) for 2 hours. This was followed by trypsinization of ES cells for 5 minutes and harvesting into a 15 ml falcon tube. The cells were then resuspended in 10 ml of pre-warmed 10 mM potassium chloride (KCl) solution at 37° C for 15 minutes. Meanwhile, ice-cold methanol-acetic acid fixative was made freshly in the ratio of 3:1 respectively. 5 drops of the fixative were then added to the cells and mixed well. The cells were then centrifuged at 1000 rpm for 5 minutes and resuspended in 6 ml of methanol-acetic acid fixative followed by two rounds of centrifugation at 1000 rpm for 5 minutes. The cells were finally

resuspended in 500 µl of fixative and transferred into a 1.5 ml Eppendorf tube. Using a pipette 50 µl of cell suspension was removed and dropped forcefully from 10 cm height onto a microscopic slide. The slide was left to air-dry for 5 minutes. It was then fixed in 4% PFA for 10 minutes. The remaining cell suspension was stored at -20° C until further use.

2.7. Nick translation

Probes for RNA- and DNA-FISH were prepared by labeling 2.5 µg of the respective large DNA constructs with Cy3-dUTP (GE Healthcare, #PA53022) using the Roche Nick Translation Kit. Reactions were performed at 15°C for 2 hours before heat-inactivation at 65°C for 10 minutes. Probes were ethanol-precipitated together with 25 µg mouse Cot-1 DNA and 10% 3 M NaOAc (Sodium acetate). Following ethanol precipitation, they were dissolved in hybridization buffer (50% formamide, 2x saline-sodium citrate (SSC) (pH 7.4), 2 mg/ml BSA, and 10% dextran sulphate.

2.8. RNA FISH (Fluorescence *in situ* hybridization)

Slides were dehydrated in 80%, 90% and 100% ethanol, 2 minutes each, dried and warmed to 42°C. The corresponding Cy3 or FITC-labeled DNA probes were denatured at 80°C for 10 minutes, pre-hybridized at 42°C for 10 minutes simultaneously, and applied to the slides. Coverslips (Fisher Brand microscope cover-glass, #12541A) were placed on the slides over the specific region where probe was added and was incubated at 42°C for 3 hours in dark and humid conditions. Slides were washed at 45°C under shaking, thrice with 50% formamide in 2x SSC, and thrice with 2x SSC. Slides were then washed in 1X PBS + 0.2% Tween 20, counterstained with Vectashield anti-fade medium (Vector Laboratories, #H-1200) containing 0.2 µg/ml of DAPI, and sealed with 50 mm x 50 mm coverslips. Slides were examined by fluorescent microscopy.

2.9. Fluorescent microscopy

Fluorescent images were obtained by a Nikon Eclipse Ti-E inverted microscope with Nikon software (NIS-Elements AR 3.2). The images were processed on Adobe Photoshop and Adobe Illustrator. DIC images at 20X magnification were also taken by the Nikon inverted microscope Eclipse Ti- E.

2.10. Calculations and tabulation of telomere length data

Cell generations and speed of growth of cells

Cell Generation = $\log_2 \left(\frac{\text{Final number of cells}}{\text{Initial number of cells}}\right)$

Cumulative cell generation = $\sum_{i=0}^{n} n$, where n is the number of cell divisions

Speed of Growth of Cells = $\frac{\text{Cell Generations}}{\text{Time in days}}$

Fluorescence Intensities Measurements

Average Distal Telomere Fluorescence Intensity

$= \frac{\text{Sum of fluorescence intensities of distal telomeres from 2 sister chromatids}}{2}$

Average Centromeric Telomere Fluorescence Intensity

$= \frac{\text{Sum of fluorescence intensities of centromeric telomeres from 2 sister chromatids}}{2}$

Average Sx9 Fluorescence Intensity

$= \frac{\text{Sum of fluorescence intensities of Sx9 from 2 sister chromatids}}{2}$

Relative distal telomere length = $\frac{\text{Average Distal Telomere Fluorescence Intensity}}{\text{Average Centromere Fluorescence Intensity}}$

Relative centromeric telomere length =

$\frac{\text{Average Centromeric Telomere Fluorescence Intensity}}{\text{Average Centromere Fluorescence Intensity}}$

2.11. Simultaneous RNA-DNA FISH

Fluorescence *in situ* hybridization (FISH) probes

Telomere DNA and *Terra* were detected with PNA probes (Panagene, # F1002, F1009 and F1010). The amino-labeled oligonucleotide probe for detecting *Terra* and the probes for single molecule RNA FISH were synthesized in-house.

RNA FISH

RNA FISH was performed using the same protocol as mentioned in Section 2.8 but instead of adding mounting media to the slides after washing here we proceed to the post-fixation step.

Post-fixation

Prior to DNA FISH, the slides were treated in 4% PFA for 10 minutes at room temperature and rinsed twice in PBS for post fixation of RNA FISH signals.

DNA FISH

Slides were denatured in 70% formamide, 2x SSC at 80°C for 10 minutes. Next, the denatured slides were serially dehydrated in ice-cold 80%, 90%, 100% ethanol for 2 minutes each. Simultaneously, the probes for DNA FISH were denatured at 75°C for 10 minutes and pre-annealed at 42°C for 10 minutes. The denatured probe was applied onto the slides for an overnight hybridization at 42°C in dark and humid environment. After hybridization the slides were washed thrice in 50% formamide and 2x SSC at 45°C for 5 minutes each and another three times in 2x SSC at 45°C for 5 minutes each. They were briefly immersed in PBS-Tween-20 and mounted with DAPI (0.2 μg/ml) containing mounting media (Vectashield). Coverslips were sealed with nail polish. Slides were examined under a fluorescent microscope.

2.12. Synthesis of amino-labeled oligonucleotide probes

Amino-labeled oligonucleotides were synthesized on an automated DNA synthesizer (Mermade 4, Bioautomation, USA) by using amino-dT phosphoramidite (Glen Research, #10-1039). The standard coupling method was applied to incorporate amino-dT and all of the natural nucleosides. FITC labels were conjugated to DNA probes via solid phase phosphoramidite chemistry. FITC and amino dual labeled probes were deprotected by AMA treatment at room temperature for 2 hours and were purified by reverse phase HPLC (Shimadzu, C18 column, 10 mm x 250 mm, 0% - 15% ACN over 50 mM TEAA buffer, pH 7.0 over 30 minutes). To prepare Cy5-labeled oligonucleotide probes, DNA strands with or without amino-dT were first extended with a C6 alkyl amino linker at 5'-end on solid phase. Cy5 NHS ester (Lumiprobe, #23020) was then coupled to amino terminal in the presence of 10 eq of DIPEA in anhydrous DMF for 2 hours at room temperature. Cy5 labeled DNA strands were cleaved from resin and deprotected in saturated ammonium solution simultaneously for 12 hours at 60˚C. HPLC purification was then carried out for the crudes and peaks with both DNA at 260 nm and Cy5 at 640 nm were collected. All of the oligonucleotide probes were confirmed by ESI-MS (Sangon, Shanghai, China)

2.13. Electrophoric mobility shift assay

20% native PAGE and 15% denaturing PAGE were applied to EMSA on duplex of probe and telomeric sequence (3 μL of 200 nM) in the presence or absence of Bovine Serum Albumin (66.4 mg/mL, Amersco, #9048-46-8) before and after formaldehyde fixation. Native and denaturing PAGE gels were run under 7 V for 7 hours and 10 V for 4 hours, respectively. All of the gels were then visualized by FITC fluorescence on a phosphorimager (Typhoon Trio, GE).

2.14.Lipofectamine transfection of EF cells

For transfection, transformed mouse embryonic fibroblasts (MEF) cells were seeded on coverslips and allowed to become 50% confluent. Before transfection, the cells were incubated in transfection medium (EF medium without serum and antibiotics) for 2 hours. Transient transfection of pEAK12GFP, an EGFP-expression plasmid, was carried out using Lipofectamine 2000 (Invitrogen, #11668-019). 2 µg of plasmid was added in every round of transfection. After 48 hours, the cells adhering to coverslips were fixed in 4% PFA for 10 minutes and preserved in 70% ethanol.

2.15.Single molecule RNA FISH

For single molecule RNA FISH, an oligonucleotide probe set was designed by Stellaris Probe Designer (https://www.biosearchtech.com/) to target eGFP mRNA. The oligonucleotide probe set was labeled with one Cy5 dye at 5' end. The corresponding amino-labeled probe set was synthesized with two or three amino labels in each oligonucleotide. The detailed experimental protocol has been described in literature (Raj et al., 2008). In brief, alcohol was removed from the previously fixed samples and they were rehydrated for 5 minutes in a solution of 2x SSC and 10% formamide. Following this hybridization reactions were performed in 10 µl of hybridization buffer containing 10% dextran sulfate (Sigma, #42867), 2 mM vanadyl-ribonucleoside complex (Sigma, #94742), 0.02% RNAse-free BSA (Invitrogen, #AM2614), 50 µg Yeast RNA, 2x SSC, 10% formamide and an empirically determined amount of probe cocktail (typically around 10-50 nM) overnight at 30˚C in dark and humid environment. The probe was added to a fresh slide and the coverslip with the cells fixed onto it was slowly placed on the probe such that the probe spreads over the entire surface of the coverslip. After hybridization the cells were washed twice for 30 minutes at 30˚C using a wash buffer (10%, formamide and 2x SSC). Nuclear staining was performed by adding DAPI (Sigma, #D9542) to the wash solution during the second wash. For Cy5 imaging we freshly prepared 100 µl mounting media by adding 1 µl of 3.7 mg/mL glucose oxidase (Sigma, #G2133) and 1

µl of catalase suspension (Sigma, #C3515) to 10 mM Tris HCl pH 8.0, 2xSSC, and 0.4% glucose. The coverslips were mounted on glass slides and sealed with nail polish.

For performing smRNA FISH coupled with DNA FISH, after overnight hybridization with the RNA probes, the coverslips were washed in wash buffer for 30 minutes as mentioned above. This was followed by post fixation of RNA FISH signal in 4% PFA for 15 minutes. The slides were rinsed twice in PBS. They were then treated with PBS + 0.5% Triton X and washed twice with PBS + 0.2% Tween20 for 5 minutes each. Denaturation was carried out for 5 minutes in 70% formamide, 2x SSC at 75°C. Simultaneously, probes for DNA FISH were denatured and pre-annealed. Probes were added to the cells and overnight incubation was performed at 42°C. The washing steps have been previously described in DNA FISH protocol. The coverslips were then mounted on slides with oxygen depleted mounting media and sealed using nail polish. The slides were viewed under a fluorescent microscope.

2.16.NFκB inhibitor treatment of cells

BAY-11-7082 (from Dr. Peter Cheung's lab, Sigma Aldrich #B5556) was used to inhibit NFκB activity in the cells. Mouse embryonic stem cells were grown in a T25 flask. 40 µl of 10 mM of the inhibitor was added to each flask containing 5 ml of media. Cells were incubated for an hour in the presence of the inhibitor. The media containing the inhibitor was pipetted out after an hour and preserved in a 15 ml tube. The cells were trypsinized followed by neutralization of the enzyme. They were then resuspended in media containing the BAY-11-7082 inhibitor which was preserved earlier. The cells were subjected to heat shock treatment.

2.17.Heat shock

ES cells were grown under standard conditions till 80% confluent. For heat shock, cells were trypsinized and resuspended in culture medium. Each sample was divided equally into two parts. The set undergoing heat shock was incubated at 45°C

for 45 minutes in a water bath. For the control set, cells were incubated at 37°C for 45 minutes in a water bath. After 45 minutes the samples were pelleted down and washed with PBS. Following this, either RNA was isolated from the cells or they were fixed on slides using the Cytocentrifuge (Thermo Scientific Cytospin-4).

2.18.RNA Isolation from cell lines

Cells were cultured till 80% confluent and lysed by adding 1 ml of TRIzol (Ambion, #15596018) reagent to a T25 flask. The cell lysate was pipetted up and down several times for homogenization and incubated for 5 minutes at room temperature. The lysate was mixed with 200 µl of chloroform in a 1.5 ml Eppendorf, vortexed vigorously, incubated at room temperature for 3 minutes, and centrifuged at 12,000g for 15 minutes at 4°C. The colourless upper aqueous phase containing RNA was collected into a fresh tube, mixed with 500 µl of isopropyl alcohol and incubated at room temperature for 10 minutes followed by centrifugation at 12000 g for 10 minutes at 4°C. Supernatant was removed and the pellet was washed with 1 ml of 75% ethanol for 5 minutes at 7500g at 4°C. The pellet was air dried and dissolved in appropriate amount of RNase-free water. DNase (Promega, #M6101) digestion using 6 units of enzyme was performed at 37°C for 1 hour followed by Phenol/choloform/isoamyl alcohol extraction. RNA was ethanol precipitated overnight. It was pelleted down by centrifugation at 12000g at 4°C for 15 minutes followed by a 75% ehtanol wash for 5 minutes. The pellet was re-dissolved in appropriate amount of RNase free (DEPC Treated) water.

2.19.Reverse trascription (RT)

1 µg of DNase-treated RNA was used with an iScript Reverse Transcription Supermix (Biorad, #170-8840) for RT reaction. The preblended 5x supermix contains iScript MMLV-RT ($RNaseH^+$), RNase inhibitor, dNTPs, oligo (dT), random primers, buffer, $MgCl_2$ and stabilisers. 4 µl of the supermix was mixed with the RNA and nuclease treated water was added to makeup a final volume of 20 µl. The reaction was

incubated at 25°C for 5 minutes, followed by 42°C for another 30 minutes and at 85°C for 5 minutes. RT-negative controls were prepared in a similar manner, by using the no-RT control supermix. Both RT-positive and negative reaction products were placed at -80°C for prolonged storage.

2.20. Quantitative PCR

Quantitative analysis of gene expression by real-time PCR was carried out using a CFX Connect™ Real-Time PCR Detection System (Biorad). Each Real-Time PCR reaction (20 µl) was made up of 2 µl of cDNA, 10 µl of 2x SsoAdvanced™ SYBR Green supermix (Biorad, #172-5262), 250 nM each of forward and reverse primers. All reactions were performed using the following thermal cycler conditions: 95°C for 30 seconds, followed by 35 cycles of a three-step reaction: denaturation at 95°C for 5 seconds, annealing and extension at 60°C for 30 seconds. The reaction was followed by a melting curve analysis to ensure amplification specificity. Transcript levels were normalized to Beta Actin or GAPDH mRNA levels. Relative fold change in expression was calculated using the "ΔΔCt" method as recommended by Biorad.

2.21. Plasmids for shRNA knockdown of *CTCF* and *Rad21*

To generate constructs encoding shRNA targeting *CTCF* and *Rad21* expression pSUPER RNAi System (Oligogene) was utilized. For each candidate gene, 64 nt DNA oligos were commercially synthesized by Sigma Aldrich which contained the 21 nt shRNA sequences targeting the candidate genes.

For *CTCF* the oligonucleotide set designed was:
CTCF_39021Forward:
5'gatcccc**CCCATTAACATAGGAGAGCTT**ttcaagagaAAGCTCTCCTATGTTAATGGGttttta-3'

CTCF_39021Reverse:
5'agcttaaaaa**CCCATTAACATAGGAGAGCTT**tctcttgaaAAGCTCTCCTATGTTAATGGGgggg-3'.

Similarly the oligonucleotides designed for targeting *Rad21* were:
Rad21_174832 Forward:
5'gatcccc**GCTTAGCGATTATTCTGATA**TttcaagagaATATCAGAATAATCGCTAAGCttttta-3'
Rad21_174832 Reverse:
5'agcttaaaaa**GCTTAGCGATTATTCTGATAT**tctcttgaaATATCAGAATAATCGCTAAGCggg-3'

The sequences written in bold are shRNA sequences targeting the gene of interest. The set of sequences in capital but not bold, are the reverse complement of the shRNA target sequence. The oligos were dissolved in ddH_2O to a concentration of 3 µg/µl. For each set, 1 µl of each oligo (F and R oligo) was combined with 48 µl of annealing buffer (100 mM NaCl and 50 mM HEPES pH 7.4) and the solution was heated up to 90°C for 4 minutes, followed by incubation at 70°C for 10 minutes, then slowly cooled down to 20°C within 20 minutes and then to room temperature. Meanwhile the template, pGH77 was digested with the restriction enzymes Bgl II and Hind III, purified using PCR purification kit (QIAquick PCR Purification Kit, #28106) and dissolved in ddH_2O at a concentration of 300 ng/µl. A ligation reaction was set-up with 2 µl of reconstituted oligos and enzyme digested pGH77 backbone (1µl) in the presence of T4 DNA Ligase. This was followed by Bgl II digestion of the ligated product for 30 minutes at 37°C. The product was transformed into TOP 10 (One Shot® TOP10 Chemically Competent *E. coli*, #C404003) cells and Ampicillin selected colonies were grown in liquid culture. The plasmid constructs isolated from these colonies were sent for sequencing. For each oligo set, one transformed colony was identified containing the correct sequence of the insert. These colonies were cultured and Midi Prep Plasmid isolation (QIAGEN Plasmid Midi Kit, #12143) was performed. The extracted plasmid was linearized using NotI restriction enzyme. Following this male mouse embryonic stem cells were stably transfected with the linearized plasmid DNA using Lipofectamine 2000 (Invitrogen Catalog, #11668-019) according to the manufacturer's instructions.

2.22.Lentiviral transduction and colony selection

Irradiated male fibroblasts (IMF) were seeded in 2 ml ES+LIF in a 6-well tissue culture dish and incubated at 37°C, 5% CO_2 for 24 hours. After 24 hours the growth medium was aspirated from the IMFs and conditioned media for ES cells was prepared by incubating the IMFs in ES+LIF for about 2 hours. Around 4000 mouse ES cells were seeded per well of the 6-well plate and incubated for 24 hours at 37°C, 5% CO_2. MISSION RNAi lentiviral particles were designed and ordered from Sigma Aldrich. Three set of inserts were designed having 3 repeat units of the telomeric repeat (-TTAGGG-) in all 3 reading frames. The lentiviral transduction media was prepared by mixing appropriate volume of lentiviral particles required for infection at multiplicity of infection (MOI) 10 with 2.5 µl of Polybrene (Sigma, #AL118) (2 mg/ml) in 1 ml of media. The ES cells were incubated in the virus-containing medium for 24 hours. The viral particles were removed; and the cells were cultured in ES cell culture medium without antibiotic selection for additional 24 hours. The cells were then subjected to Puromycin (Hyclone, #SV3007501) selection.

Calculations for lentiviral transduction experiment

Total Transducing Units (TU) needed = (total number of cells per well) x (Desired MOI)

Total volume of lentiviral particles to add to each well = $\frac{(\text{total TU needed})}{(\text{TU/ml reported})}$

Table 2.1. p24 Elisa Titre of shRNA Z1, Z2 and Z3

Clone ID	p24 Elisa Titre
pLKO.1Puro-Zhang1 (Z1)	3.6 x 10^7 TU/ml
pLKO.1Puro-Zhang2 (Z2)	1.9 x 10^7 TU/ml
pLKO.1Puro-Zhang3 (Z3)	5.9 x 10^6 TU/ml

Volume of each clone to be used to achieve MOI of 10 is-

Zhang1 = $\frac{4000\times10}{3.6\times10*7}$ = 1.1ul

Zhang2= $\frac{4000\times10}{1.9\times10*7}$ = 2ul

Zhang3= $\frac{4000\times10}{5.9\times10*6}$ = 6.7ul

Puromycin selection and picking colonies of ES cells

24 hours after removal of the virus containing media from the cells, 2 µg/ml of Puromycin was added. After 5-6 days of selection, small round colonies were visible on the plate. Colonies were grown till they were big enough to be picked. The colonies were picked from the dish with a 20 µl tip. Each colony was placed into one well of a 96 well plate containing Trypsin/EDTA for trypsinization. This was followed by mechanical dissociation of the colonies by pipetting and transferring to a 96 well plate containing feeder cells (IMF). They were cultured further into 24 well, 12 well and 6 well plates before being frozen down. The established cell lines were maintained in ES culture medium containing 1 µg/ml of Puromycin.

2.23.Polyacrylamide gel electrophoresis of RNA (Northern blotting)

Prepared a 15% PAGE from Sequagel® UreaGel System (National Diagnostics, #EC833) using 24 ml of concentrate, 12 ml of diluent and 4 ml of buffer. Added 200 µl of 10% APS (Sigma, #A3678) and 20 µl TEMED (Sigma, #T9281) to this mixture before pouring it in gel casting apparatus. For the gel apparatus, washed the glass plates with RNaseZap (Ambion, #AM9780) and dried them using 100% ethanol. The inner side of one glass plate was silanized using Sigmacote (Sigma, #SL-2), and dried in a chemical fume hood for 20 minutes, followed by 100% ethanol wash. After assembling the gel running system, the gel was poured into the system, set for an hour and pre-run for 30 minutes.

20-30 µg of total RNA was ethanol precipitated, and the pellet was re-suspended in 10 µl DEPC H_2O and 12 µl 8 M Urea RNA Loading Dye (48% Urea, 0.25% Bromophenol Blue, 0.25% Xylene Cyanol and 20 mM EDTA). Samples were heated at 80°C for 10 minutes and loaded on to the gel. The gel was run at 150V with 0.5X TBE (Tris-Borate EDTA) for 1 hour to allow samples to enter the gel gently. Then voltage was increased to 300 V for about 3 hours, until the loading dye runs off the gel. The silianized glass plate was removed and the gel was stained with 4 µg/ml of

Ethidium Bromide (EB) in 0.5X TBE for 5 minutes and visualized in Gel Scanner (Genesnap, Syngene). RNA was transferred to Hybond N+ membrane (GE Healthcare, #RPN119B) using semi-dry transfer apparatus (AA Hoefer, #TE70XP) for ~60 minutes at 3.3 mA/cm^2 in 0.5x TBE. RNA transferred to membrane was cross-linked with 254 nm UV at 150,000 μJ/cm^2 using UV cross-linker (Hybri Linker, UVP Laboratories, # HL-2000). Blocking of the membrane was carried out in pre-warmed commercially available ULTRAhyb® Ultrasensitive Hybridization buffer (Ambion, #AM8670) for 2 hours. Probes were labeled with γP^{32}-ATP in a T4 Polynucleotide Kinase Reaction and hybridization was carried overnight at 50°C to detect RNA. Blot was washed twice for 30 minutes each with 2x SSC, 0.5% SDS, at 50°C. Membrane was exposed to phosphorimage screen (GE Healthcare, #28-9564-74) for 3 hours to overnight and scanned using a Typhoon Trio Variable Mode Imager. The same protocol was followed for all the probes. The membrane was stripped of signal by boiling in 0.1% SDS for 10 minutes.

2.24.5' end labeling by T4 Polynucleotide Kinase reaction

For the reaction, 2 µl of 10x of T4 PNK buffer (NEB), 1 µM of DNA or RNA oligonucleotide probe, 6 µl of γP^{32}-ATP (Perkin Elmer, 6000Ci/mmol), 9 µl of Millipore water, and 1 µl of T4 PNK enzyme (10,000 units/ml, NEB #M0201) were mixed together to a final volume of 20 µl. Reaction sample was incubated at 37°C for 60 minutes followed by dilution to 100 µl in miliQ water (Millipore miliQ). Labeled oligonucleotides were purified through G-50 spin columns (GE Healthcare, #27-5330-01) according to the manufacturer's instructions.

2.25.Plasmids for CRISPR dependent upregulation of *Terra*

The plasmids encoding the Cas9 mutant (Cas9m4, #47316) and Cas9 mutant with VP64 transcription activator (Cas9m4-VP64, #47319) were ordered from Addgene. The Cas9 mutants (Cas9m4) lacked measurable nuclease activity due to mutations in four amino acids from HNH and RuvC-like catalytic domains (primarily

involved in nuclease activity of Cas9 protein). Cas9m4-VP64 was generated by fusing the nuclease deficient Cas9 (Cas9m4) with the VP64 activation domain at C-terminus (Mali et al., 2013b). The plasmids contain an NLS domain for directing the protein to the nucleus. The gRNA (guide RNA) expression construct for the target was assembled into a gRNA cloning vector (Addgene, #41824). gRNA construct cloning and design is explained in the next section.

Construction of gRNA expression plasmid

For constructing gRNA expression plasmid, first a 23bp genomic site in the form 5'-N_{20}NGG-3' was chosen. The highlighted 19bp of the selected target site 5'-N**NNNNNNNNNNNNNNNNNNN**NGG-3' was incorporated into two 60 mer oligonucleotides in direct and reverse complement orientation. For targeting *Terra* expression the two designed oligos were:

Insert_*Terra*_F
TTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCG**AGGGTTAGGGTTAGGGTTA**

Insert_*Terra*_R
GACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAAC**TAACCCTAACCCTAACCCTC**

The two oligos were annealed and extended to obtain a 100 bp double stranded DNA fragment using Phusion Polymerase (NEB, #M0530) according to manufacturer's protocol. The product was electrophoresed and the 100 bp fragment was cut from the get and purified. Simultaneously the gRNA cloning vector (Addgene, #41824) was linearized using AflII and purified.

The 100 bp DNA fragment and the linearized gRNA cloning vector were ligated using the Gibson assembly master mix (NEB, #2611). For Gibson assembly reaction, 50-100 ng of vector was used in the presence of 5 fold excess of insert. The reaction mixture contained the vector backbone (~50 ng), the DNA insert (~150 ng), 10 µl of 2x Gibson assembly master mix and deionized water was added finally to make up the final volume to 20 µl. The mixture was incubated at 50°C for 30 minutes. The assembled product from above was then transformed into One Shot TOP 10 competent cells (Life Technologies, #C4040). For transformation, the assembled product was

diluted 4-fold with water prior to transformation. 2 µl of the diluted product was incubated with 50 µl of competent cells on ice for 30 minutes. Heat shock was provided at 42°C for 30 sec. Post heat-shock, the cells were transferred to ice for 2 minutes and then 950 µl of room temperature SOC media was added. They were allowed to recover at 37°C for 60 minutes with vigorous shaking. 100 µl of the transformed product was plated on LB plates containing Kanamycin (50 µg/ml). Colonies were picked and the isolated plasmids were sequenced to confirm insertion.

2.26. Transfection of ES cells

For transfection, 2 x 10^5 mouse ES cells were seeded. Next day, transient plasmid transfection was carried out using Lipofectamine 2000 (Invitrogen, #11668-019) as per manufacturer's protocol. 2 µg of each plasmid was added in every round of transfection. After 48-72 hours, the cells were harvested and fixed onto slides using cytospin.

2.27. Malaria parasite culture

P. falciparum DD2 strain parasites were cultured in human red blood cells in RPMI 1640 media supplemented with 0.25% Albumax (Invitrogen, E003000PJ), 2 g/l sodium bicarbonate (Sigma, #G5761), 50 µl/l gentamycin (Invitrogen, #15710072), and 0.1 mM hypoxanthine (Sigma, #H9377). The culture was maintained at 2% hematocrit. Human erythrocytes were purified from whole blood obtained from anonymous donors who signed a written consent. The blood donation scheme for these studies has been approved by the internal ethics committee of the Nanyang Technological University. Cultures were grown in Prof. Zbyneck Bozdech's lab and provided by Luah Yen Hoon.

2.28.RNA FISH on malaria parasites

Cultures were washed twice in PBS and smears of the cultures were prepared on microscopic slides. RNA FISH was performed according to the same protocol mentioned above in Section 2.8 with minor modifications. Briefly, the blood smears were fixed in 4% Paraformaldehyde for 10 minutes followed by a PBS wash for 5 minutes. Slides were then treated with PBS + 0.5% Triton X-100 (Sigma, #T8787) for better probe penetration during RNA FISH. The slides were then washed with PBS thrice for 5 minutes each and dehydrated sequentially in 80%, 90% and 100% ethanol for 2 minutes each. A peptide nucleic acid (PNA) probe was designed and commercially synthesized (Panagene). The probe contains the telomeric repeats of the parasite.

Pf *Terra* PNA = 5'-FITC-OO-TGAACCCTGAACCCT-3'

The FITC-labeled PNA probe was denatured at 80°C for 10 minutes, pre-hybridized at 42°C for 10 minutes simultaneously, and applied to slides. Coverslips were placed on slides and slides were incubated at 42°C for 3 hours in a dark and humid environment for probe hybridization. Slides were then washed at 45°C under shaking: thrice with 50% formamide in 2x SSC and thrice with 2x SSC. Slides were then briefly dipped in PBS + 0.2% Tween 20 (Sigma, #P1379), counterstained with Vectashield anti-fade medium (Vector Laboratories) containing 0.2 μg/ml of DAPI, and sealed with 50 mm x 50 mm cover-slips. Slides were examined by fluorescence microscopy.

Mouse Cot1 DNA (Invitrogen, #18440016) was added 20 times in excess of malaria *Terra* probe to test the specificity of the probe in detecting *Terra*.

2.29.RNA isolation from malaria parasite

Prior to RNA isolation, the culture was collected in falcon tubes and stored at -80°C. To isolate RNA from malaria parasites, the cells were first thawed in a 65°C water bath for 2 minutes. Once thawed, 10 times volume of TRIzol (Ambion,

#15596018) was added to the cells. This was followed by adding chloroform twice the volume of blood and the sample was mixed gently and left on ice for 5 minutes and centrifuged at 4000 rpm for 15 minutes at 4°C. The top layer was pipetted out in a fresh 15 ml tube and an equal volume of isopropanol was added and mixed by inverting. The mixture was stored over night at -20°C. It was then centrifuged at 9000 rpm for 1 hour at 4°C. The supernatant was carefully decanted and 10 ml of 70% ethanol was added for washing. After centrifuging the tube for 10 minutes supernatant was discarded and the pellet was dried for 1 to 2 hours at room temperature. The dried pellet was dissolved in 100 µl of nuclease-free water. The RNA sample was then quantified using a nanodrop and stored at -80°C.

2.30. RNase protection Assay

Probe preparation for RNase protection assay by *in-vitro* transcription

For synthesis of the radioactively labeled RNA, we first amplified the 200 nt oligo template, 'Telo-C', with the primers Telo-C_Forward and Telo-C_Reverse. Mouse_Telo-C was designed as template to detect mouse *Terra* and Pf_Telo-C for *Plasmodium falciparum* (Pf) *Terra*. Mouse_Telo-C contains $(TAACCC)_{25}$ repeat sequence along with adaptors for amplification on both ends. –TAACCC- is the antisense of the mouse *Terra* repeat -GGGTTA-. Similarly Pf_Telo-C contains $(TAAACCC)_{22}$ repeats. The forward and the reverse primer had a HindIII and an EcoRI restriction site inserted in their 5' end respectively. After PCR amplifying the template DNA, the product was digested with EcoRI and HindIII and ligated into the vector backbone pGEM3Zf at its EcoRI and HindIII site. Upon successful ligation, the newly generated plasmid was named '**pGEM_Mouse_Telo' and "pGEM_Pf_Telo" for detecting *Terra* from mouse and *Plasmodium*** respectively. This plasmid was used for *in vitro* transcription reaction to generate RNA probes used in RPA experiment.

Table 2.2. Sequence of Telo C probe and the primers used for its amplification.

Probe	Sequence
Mouse_Telo-C	5' **GGT AAC AAT TTT CCC GCC ATG TG**T AAC CCT AAC **GGT ACT GCG TTC CGT TTA TTT CC** 3'
Mouse_Telo-C _Forward	GCGAAGCTTCCCGCCATGTGTAACCCTAACCC
Mouse_Telo-C_Reverse	CGCGAATTCAACGCAGTACCGTTAGGGTTAGG
Pf_Telo-C	5'GGTAACAATTTTCCCGCCATGTGAA CCCTAAACCCTAAACCCTAAACCCTAAA CCCTAACCCTAAACCCTAAACCCTAAAC CCTAAACCCTAAACCCTAAACCCTAAAC CCTAAACCCTAAACCCTAAACCCTAAAC CCTAAACCCTAAACCCTAAACCCTAAAC CCTAAACCCTAAACCCTAAACCCTAAAC CCTAAACCCTAGGTACTGCGTTCCGTT TATTTCC 3'
Pf_Telo-C_Forward	GCGAAGCTTCCCGCCATGTGAACCCTAAACC
Pf_Telo-C_Reverse	CGCGAATTCAACGCAGTACCTAGGGTTTAGG

An RNA probe uniformly labelled with [α-^{32}P]rCTP was synthesized by bacteriophage RNA polymerases (SP6 and T7) which initiate transcription of the anti-sense or sense transcripts from the genomic fragment, containing the targeted sequence, cloned downstream of the specific phage promoters in the common plasmid vector pGEM-3Zf(+) (Promega, #P2271) by *in vitro* transcription. The following components were added at room temperature in the listed order: 5 µl of 5X Transcription buffer, 2 µl of 100 mM DTT, 40 units of Recombinant RNasin Ribonuclease inhibitor, 4 µl of rATP, rGTP and rUTP mix having 2.5 mM of each, 1 µl of linearized template DNA (0.5 µg/ml) and 5 µl of [α-^{32}P]rCTP (50 µCi at 10 µCi/ml, Perkin Elmer) and 20 units of SP6 or T7 RNA polymerase (as required). The final reaction volume was 20 µl. The reaction was incubated at 37°C for 1 hour. This was followed by DNase digestion by RQ1 RNase-free DNase (1 unit per µg of DNA template) to remove the DNA template for 15 minutes at 37°C. After DNase digestion,

phenol-chloroform extraction and ethanol precipitation was performed. The RNA pellet was dried and resuspended in nuclease free water.

RNase protection assay

The synthesized RNA probe was dissolved in nuclease free water such that the specific activity is ~10^9 cpm/µg. 10-20 µg of sample RNA and 1-5 x 10^5 cpm of probe were co-precipitated and re-dissolved in 30 µl of hybridization buffer (1 mM EDTA, 80% formamide, 400 mM NaCl, 40 mM PIPES). The reaction was incubated at 95°C for 5 minutes to denature the RNA and immediately transferred to 42°C for hybridization of the sample RNA with the labeled probe. The hybridization reaction was carried out overnight. This was followed by ribonuclease digestion. 350 µl of the RNase digest buffer [10 mM Tris (pH 7.4), 300 mM NaCl, 5mM EDTA] containing RNase A/T1 cocktail (Invitrogen, #AM2286) at 1:100 dilution was added to above reaction mixture. The reaction was incubated at 37°C for 60 minutes. To stop the RNase digestion reaction, 10 µl of 20% SDS and 50 µg of Proteinase K were added in the reaction and the reaction was incubated at 37°C for 15 minutes. The reaction was extracted once with 400 µl of phenol/chloroform/ isoamyl alcohol. The RNA sample from the reaction was ethanol precipated in the presence of 1 µg Yeast RNA (Life Technologies, #AM7120) at -80°C. The pellet was dried after ethanol precipitation and then dissolved in 7 µl of formamide loading buffer (Life Technologies, #AM8546). The sample was then incubated at 95°C for 3-4 minutes and run on a denaturing 5% urea gel.

Denaturing Urea PAGE gel

Denaturing PAGE was prepared for a 5% Urea gel using the following recipe: 5ml of 40% 19:1 acrylamide: bisacrylamide (Biorad, #161-0144), 19.2g of urea, 4ml of 10X TBE and made up to a final volume of 40 ml with RNase-free water. The solution was briefly warmed in a 37°C water bath to help the urea to dissolve. This was followed by addition of 20 µl of TEMED and 200 µl of freshly prepared 10% ammonium persulphate (APS). The gel casting system was set up and the solution was poured. It was allowed to set for an hour. The samples prepared in the formamide

loading buffer were loaded onto the gel and electrophoresed until bromophenol blue reached the bottom of the gel. The gel was wrapped in cling film and exposed to phosphorimager screen (GE Healthcare, #28-9564-74) for 3 hours to overnight. The phosphoimager screen was scanned using a Typhoon Trio Variable Mode Imager.

CHAPTER 3: OBJECTIVES AND MOTIVATION

3. OBJECTIVES AND MOTIVATION

- *Terra* is a non-coding RNA transcribed from chromosome ends. Our first specific aim was to develop a robust system for simultaneous detection of RNAs and their corresponding DNA molecules. This method can be utilized to investigate the functions of various ncRNAs. In the present study, we gain new insights into *Terra* and telomeric DNA interactions using this technique.

- Our second specific aim was to study the effect of *Terra* on telomere length regulation. *Terra*, telomerase and telomeric DNA interaction and regulation have been studied *in vitro*. Our objective was to detect the effect of *Terra* on telomere elongation *in vivo*.

- Our third aim was to study the effect of cellular stress on *Terra* expression. *Terra* levels are known to be susceptible to cellular conditions and genotoxicity. We investigate the effect of heat shock on *Terra* expression pattern and transcriptional regulators of *Terra* in the nucleus.

- For studying biological role of *Terra* it is extremely helpful to modify *Terra* expression and observe its impact on the cell. With this objective we attempted to downregulate and upregulate *Terra* expression in cells as the fourth aim of this study.

- Apart from studying *Terra* expression and function in mammalian cells, we investigate the presence of *Terra* in the malaria parasite *Plasmodium.* Malaria is a dreaded disease around the world due to the parasites' ability to evade antigenic response by activating specific gene families in the telomere associated regions. As the fifth objective of this study, we decided to test *Terra* transcription from *Plasmodium* telomeres.

CHAPTER 4:

RESULTS

4. RESULTS

4.1. Observations on *Terra*

The telomere repeat containing RNA or *Terra* is transcribed from telomeres in a strand-specific manner and its transcription initiation is thought to start from the subtelomeric region and progress towards the chromosomal end, terminating within the telomeric repeat tract (Azzalin et al., 2007; Schoeftner and Blasco, 2008; Zhang et al., 2009). Although several functions have been proposed for *Terra*, its biological role remains largely elusive. To obtain a better understanding of its roles in the biological system, our first target was to study the subcellular localization of *Terra* and observe its interaction with the telomere DNA. For attaining this objective we developed a technique which allowed us to perform simultaneous *in situ* visualisation of RNA and DNA.

Also in this section, apart from RNA-DNA FISH method, we detected *Terra* in mouse cells by RNase protection assay (RPA). RNA can be detected by Northern hybridization, RNA FISH, RPA etc. but, among them RPA is a more sensitive technique for studying RNA. We also attempted to detect *Terra* antisense using RPA. *Terra* antisense has been shown to be present in *Arabidopsis* (Vrbsky et al., 2010). We endeavoured to study if an antisense of *Terra* was transcribed in mouse.

4.1.1. Developing tools for RNA-DNA FISH

Recent research has shown that >90% of the human genome is likely to be transcribed into non-coding RNAs. The list of lncRNAs being discovered in our biological system has been steadily growing (Wilusz et al., 2009). Many of them play important roles in nuclear architecture and epigenetic regulation (Guttman and Rinn, 2012; Lee, 2012). Therefore, it is important to study where the lncRNAs are localized

and how they interact with their DNA targets in the nucleus. Recent improvements towards this objective have been made on three front lines.

The first line, live-cell imaging, can be applied using fluorescent protein tags but has not been particularly used to target RNA and DNA simultaneously. Second, is single molecule detection (not measurement), which had been previously achieved only on RNA molecules in fixed cells (Raj et al., 2008).The third line is to perform simultaneous detection of the subcellular localization of RNA and DNA molecules by RNA-DNA FISH (Fluorescence *in situ* hybridization). Experimental approaches for RNA-DNA FISH involve (a) sequential hybridization of RNA and DNA FISH probes, wherein the difficulty lies in preserving the RNA through the harsh DNA FISH steps; (b) superimposition of RNA and DNA FISH signals after performing the experiments separately, a tedious and time-consuming approach; and (c) by the post fixation approach wherein RNA FISH signal is fixed by means of a chemical fixative before the slides are subjected to DNA FISH. This is incredibly useful, yet technically challenging, as the harsh conditions involved in denaturing the DNA template in DNA FISH easily destroys the fragile RNA signals. In our previous study (Lai et al., 2013), we solved the problem by introducing protein components into the RNA FISH signal detection steps. This, followed by a fixation step of formaldehyde, can efficiently protect the RNA FISH signals from being damaged by the subsequent DNA FISH. The fixation step is used to crosslink the RNA FISH signal with the cellular proteins in its close proximity, such that the RNA targets may still be damaged in the subsequent DNA FISH, but the RNA FISH signals survive. For the fixation step to work, it is critical to introduce protein components (immunostains) into the signal detection steps of RNA FISH, because the chemistry of formaldehyde-mediated fixation via the formation of covalent chemical bonds among amino groups, determines that formaldehyde-mediated crosslink (Fig. 4.1) can only be efficiently applied to proteins, which contains multiple lysine residues to offer the amino group as crosslinking sites.

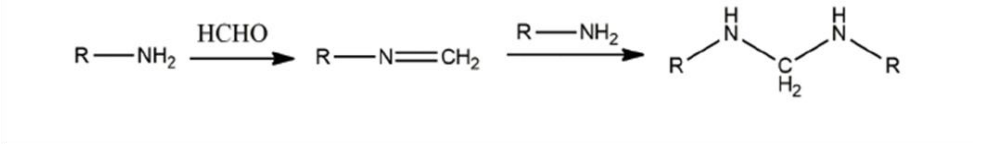


Figure 4.1. Chemical principle of formaldehyde fixation. Formamide fixation involves the formation of amide bonds between amino groups lying in close vicinity.

Although this method worked robustly, it had some shortcomings. Firstly, including multiple immunostains into the signal detection steps of RNA FISH made the procedure tedious. Secondly, addition of immunostains amplified the signal and made it quantitatively irrelevant. Thirdly, the slight background noise generated by immunostains is high enough to defeat the purpose of single RNA molecule detection.

To overcome these disadvantages, we designed and synthesized a novel probe where amino groups were added directly onto the nucleotide probes for RNA FISH. This new probe is described in detail in the next section.

4.1.2. RNA and DNA FISH with amino labeled probe

To circumvent the shortfalls of the previous RNA-DNA FISH method, we designed a new probe where amino labels were introduced directly on the nucleotide probe for RNA FISH. We first synthesized an amino-labeled 36 nt DNA oligonucleotide probe with the sequence $(TAACCC)_6$ to detect *Terra*. Alkyl amino groups are introduced to the oligonucleotide probe by adding amino-modified thymine (amino-dT) during the solid phase synthesis with a density of at least one amino group every six nucleotides. Amino-dT can survive the conjugation conditions for incorporating fluorescent dyes via either direct phosphoramidite synthesis or post-synthetic coupling. The diagrammatic representation of the amino labeled *Terra* probe and its crosslinking mechanism to cellular proteins in presence of formaldehyde is provided in Figure 4.2.

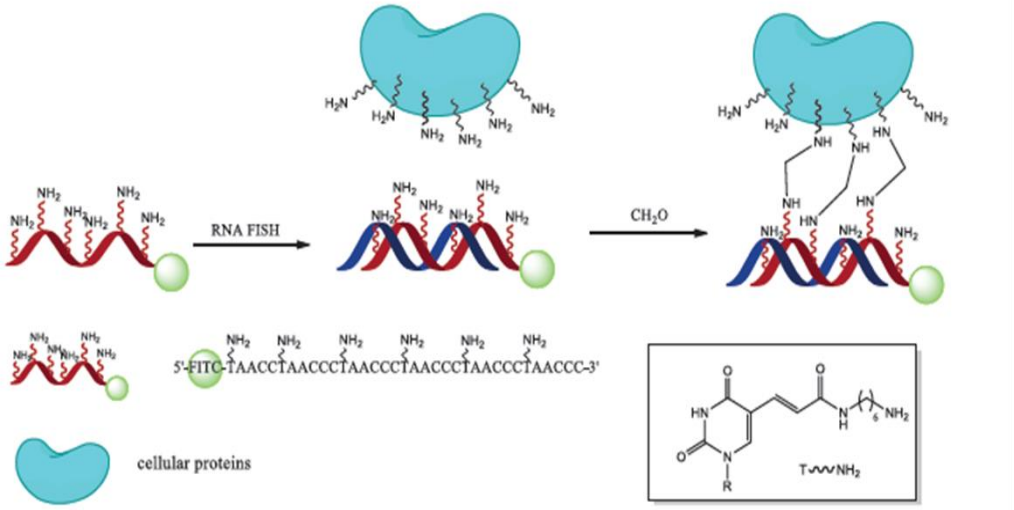


Figure 4.2. Scheme of formaldehyde mediated crosslink between amino-labeled probe and cellular proteins in vicinity. The probe has amino-dT introduced once every six nucleotides and is tagged with FITC. The oligonucleotide probes first bind to their target RNAs. Post fixation with paraformaldehyde involved formation of amide linkages between the amino labels on the probe and its surrounding proteins.

The oligonucleotide probes had dual labels of alkylamine and fluorophore. They were prepared in high yields after HPLC purification and exhibited UV-vis absorption features of both DNA and the fluorophore (Fig. 4.3A).

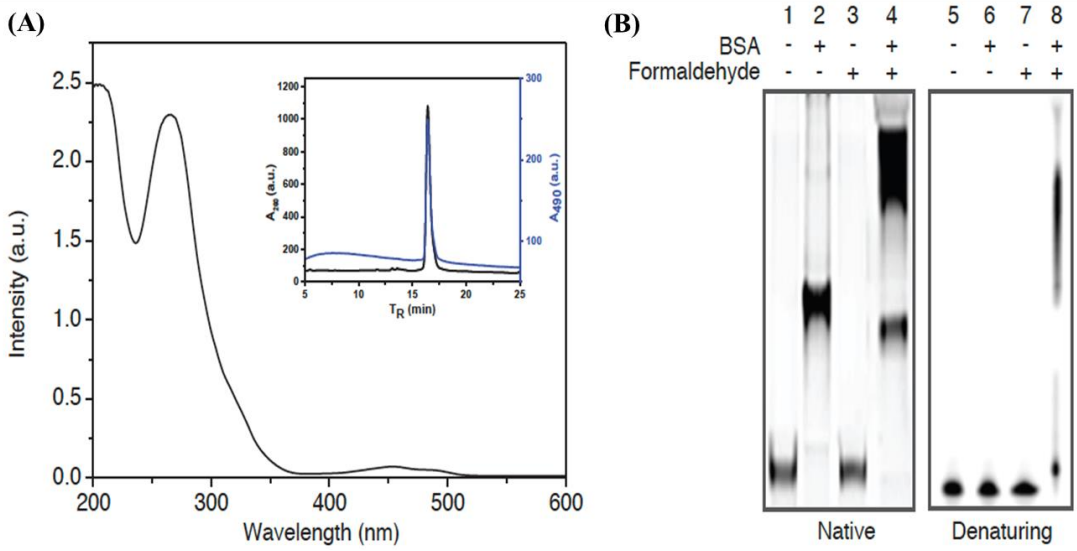


Figure 4.3. Properties of the synthesized amino-labeled oligonucleotide probe for formaldehyde fixation. (A) UV-vis absorption spectra of the oligonucleotide probe. Insert: HPLC purification trace of amino-labeled probes. (B) Electrophoretic mobility shift assay (EMSA) of formaldehyde mediated crosslink between the oligonucleotide probes and bovine serum albumin (BSA) on native (Lane 1 to 4) and denaturing PAGE (Lane 5 to 8).

Formaldehyde cross-linking reactions of the amino-labeled oligonucleotide probes were first studied *in vitro* (Fig. 4.3B). By introducing modifications on C5 of Thymine (amino group with long alkyl chain) the probe can be accommodated well in the major grooves of hybrid duplex DNA/RNA (Telser et al., 1989). Unlike the crosslinking agents such as iodo-uridine, with reaction site immediately next to nucleobases (Willis et al., 1993), the long alkyl linkage of amino-dT can send the crosslinking agent to a much wider reaction radius (> 1nm) around the probe and RNA targets, increasing the chance of the amino-labeled probes to be cross-linked with the surrounding proteins. Intra-strand crosslinking is absent in our amino labeled probe as no new bands and no alterations on the mobility of original probe band have been observed in either native or denaturing PAGE (Fig. 4.3B, Lane 3 and 7). Crosslinking reactions between the amino-labeled probes and bovine serum albumin (BSA) were observed as smear bands with much slower mobility than the probes (Fig. 4.3B, Lane 4 and 8). The fact that formaldehyde fixation further retards the mobility of the protein bound probes (Fig. 4.3B) indicates that crosslinking reaction may be extended to occur between the probes and non-bound protein which happens to be in the vicinal area. Similarly, during RNA FISH, the enlarged reaction radius by extending amino group with long linkage, would enable the amino-labeled probes to be cross-linked with the surrounding proteins and hence to survive the subsequent DNA FISH.

4.1.2.1. Simultaneous detection of *Terra* and telomere DNA

Next we carried out *Terra* RNA FISH using the amino-labeled oligonucleotide probe and a peptide nucleic acid (PNA) probe. The PNA probe is chosen as the control, because the probe exhibits extraordinary thermal stability and is not susceptible to hydrolytic (enzymatic) cleavage (Ray and Norden, 2000). The RNA signals of *Terra* were directly fixed by formaldehyde in our experiments. All the immunostains were skipped. As shown in Figure 4.4, *Terra* signals detected by the PNA oligonucleotide probe were lost after DNA FISH, despite the extraordinary thermal stability of the probe. In contrast, *Terra* signals detected by the amino-labeled probe survived DNA FISH. The surviving *Terra* signals were bright and clear (Fig. 4.4).

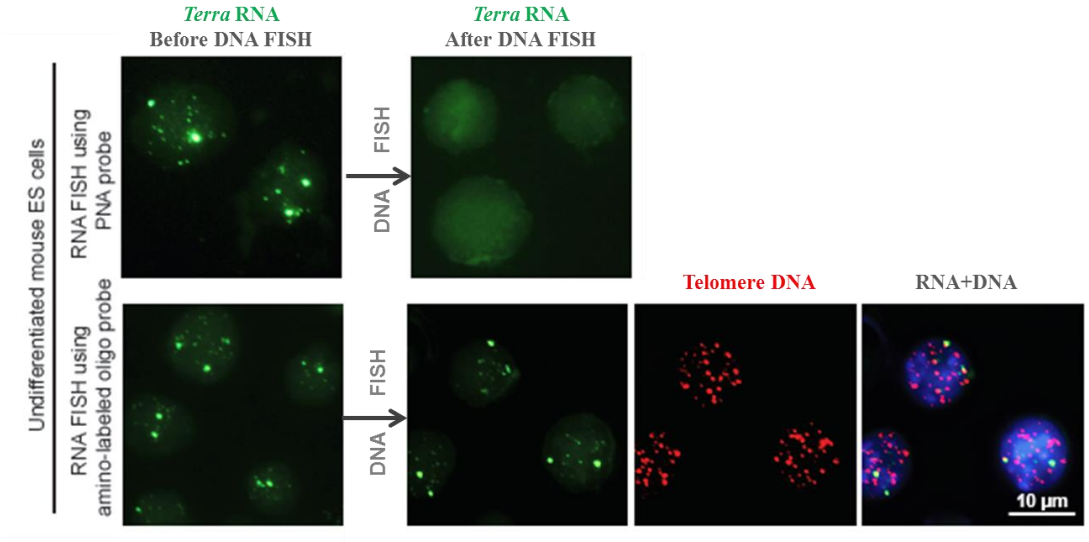


Figure 4.4. Simultaneous detection of *Terra* and telomere DNA. *Terra* was detected by FITC-labeled oligonucleotide probes (green) with the sequence of $(TAACCC)_6$. Telomere DNA was detected by Cy3-labeled oligonucleotide probes (red) with the sequence of $(GGGTTA)_6$. PNA probe and amino-labeled probe were used for RNA FISH in this study. Nuclei were stained by DAPI (blue). Scale bar, 10 μm.

Furthermore, the major biological characteristics of *Terra* were also preserved with our amino-labeled probe. Two major *Terra* RNA signals, which are significantly brighter and bigger than other *Terra* signals, were detected in the DAPI-stained nucleus in most of the mouse embryonic stem cells. Not all telomeric DNA ends in the nucleus are associated with *Terra* RNA. *Terra* transcripts were located in close proximity to the telomere DNA, but some of the *Terra* signals did not overlap with telomere DNA. All these characteristics of *Terra* were consistent with previous observations (Lai et al., 2013; Zhang et al., 2009). Thus, the RNA-DNA FISH experiment worked well with the amino-labeled oligonucleotide probe. We also compared *Terra's* signal size and signal number per nucleus before and after DNA FISH (Fig. 4.5). The results show that RNA signals were well protected from the damage of DNA FISH without significant loss in the RNA signal's quantity and quality.

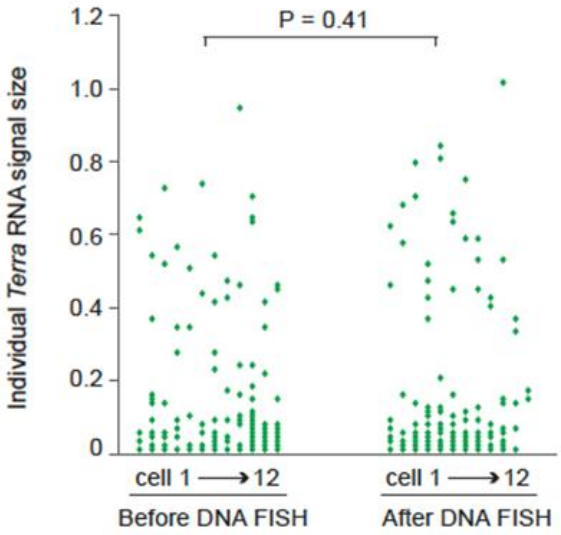


Figure 4.5. *Terra* RNA FISH signal with amino labeled probe quantified before and after DNA FISH. Signal number per nucleus and the signal sizes before and after DNA FISH were compared. A z-series image of 3 μm thickness was collected. Images were deconvolved and merged. Individual *Terra* signals in each cell were computationally recognized by setting a fluorescence intensity threshold. The size of each *Terra* signal (μm^2) was computed. *Terra* signals measured from one cell are plotted as discrete spots vertically on the X-axis. The P-value was calculated using Student t-test.

The biological functions of *Terra* are mostly unknown. Based on biochemistry evidence that *Terra* inhibits telomerase activity *in vitro*, it has been hypothesized that *Terra* associates with the long telomeres in a cell so that the telomerase activity is directed to concentrate on the short telomeres (Schoeftner and Blasco, 2008). This hypothesis predicts that the telomeres associated with *Terra* are longer than those not associated with *Terra*. To confirm this prediction, we measured the DNA FISH signal intensity of individual telomeres to quantitate the telomere length. Our results show that the average length of telomeres associated with *Terra* is slightly longer than non-associated telomeres (Fig. 4.6). However, the difference is not statistically significant. Our established RNA-DNA FISH method provided the first direct evidence of individual telomere length of telomeres associated and non-associated with *Terra*. We discuss this result in the discussion section.

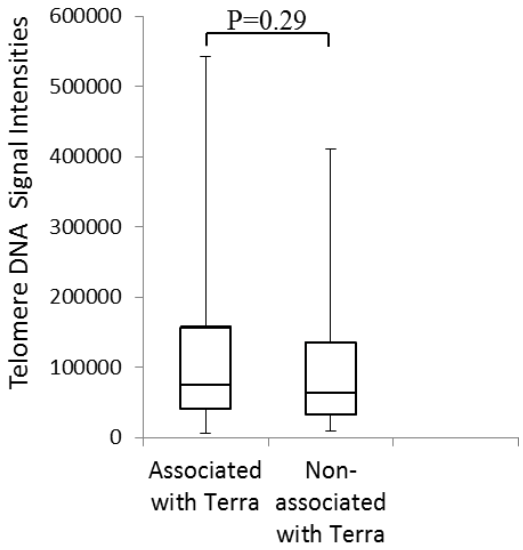


Figure 4.6. Comparison of length of telomeres associated and non-associated with *Terra*. DNA FISH signal intensity of telomeres was measured to quantitate individual telomere length. The telomeres associated and non-associated with *Terra* were separated into two groups (n=100 for each group). The data was presented in box-and-whisker plots. P-value was calculated using Student t-test.

4.1.2.2. Simultaneous detection of single mRNA molecule with DNA

The newly designed amino labeled probe could bind directly to the cellular proteins upon formaldehyde fixation. This allowed us to skip the immunostaining steps of the old method (Lai et al., 2013). The result was a simplified method of RNA-DNA FISH, which was less time-consuming and avoided unnecessary signal amplification along with reduction in the background noise. To take advantage of this, we further applied the amino-labeled oligo probes to detect single RNA molecules.

In the single molecule RNA FISH technique (Raj et al., 2008), a set of oligo probes were designed to detect a target mRNA (Fig.4.7). Previous studies have demonstrated that each pinpoint signal corresponds to a single RNA molecule (Vargas et al., 2005) and the image of single molecule RNA signals can be further processed by Laplacian of Gaussian (LoG) filter to obtain a sharper view (Tyagi, 2010).

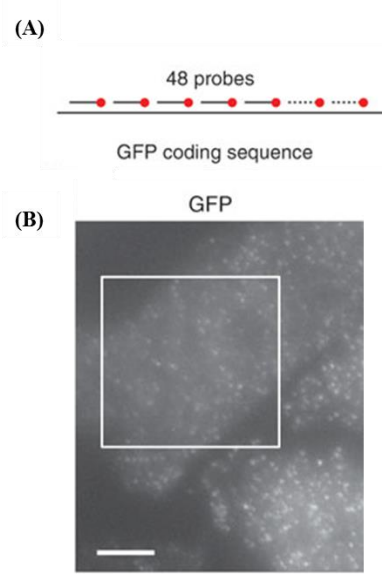


Figure 4.7. Single molecule RNA detection for enhanced green fluorescent protein (GFP) mRNA. (A) 48 probes labeled with Alexa 594 were used to detect the GFP coding sequence. (B) Fluorescent images of single molecules of EGFP mRNAs in CHO cells taken in the Alexa 594 channel made by maximum intensity merges of a pair of z-series. The white box marks a region containing GFP single molecule mRNA signal. Scale bar, 5 µm (Raj et al., 2008).

In their study Raj et al. tested the signal intensity of single molecule RNA FISH of GFP mRNA using 12, 24 or 36 probes from their set of 48 probes and found that detection worked with fewer probes for this mRNA although with lowered intensity (Raj et al., 2008). Taking this knowledge into consideration, in our study we designed and synthesized a set of 20 oligos to target the mRNA of the enhanced green fluorescent protein (EGFP) (Table 4.1). Two sets of probes were generated: a regular set without amino modification and an amino labeled set. We were able to successfully detect the single molecules of EGFP mRNA in the cytoplasm of a mouse fibroblast cell population transiently transfected by an EGFP-expressing plasmid (Fig. 4.8A) with the regular oligo probe set.

Table 4.1. Sequences of probe set for single molecule RNA FISH. In the amino labeled set the thymines (T) highlighted in red are amino-labeled thymines.

No.	Sequence	No.	Sequence
1	GTCCAGCTCGACCAGGATGG	11	CGATGCCCTTCAGCTCGATG
2	CTGAACTTGTGGCCGTTTAC	12	ATGTTGCCGTCCTCCTTGAA
3	AGCTTGCCGTAGGTGGCATC	13	GTAGTTGTACTCCAGCTTGT
4	GTGGTGCAGATGAACTTCAG	14	CATGATATAGACGTTGTGGC
5	CACTGCACGCCGTAGGTCAG	15	ATGCCGTTCTTCTGCTTGTC
6	GACTTGAAGAAGTCGTGCTG	16	GCGGATCTTGAAGTTCACCT
7	TGGACGTAGCCTTCGGGCAT	17	CGCTGCCGTCCTCGATGTTG
8	CTTGAAGAAGATGGTGCGCT	18	TGTGATCGCGCTTCTCGTTG
9	CGGGTCTTGTAGTTGCCGTC	19	GTCACGAACTCCAGCAGGAC
10	GTTCACCAGGGTGTCGCCCT	20	GTCCATGCCGAGAGTGATCC

The amino-labeled probe set had an alkylamino group conjugated to the thymine base as shown in Table 4.1. The regular and amino-labeled oligo probes were conjugated to a Cy5 fluorophore via solid phase NHS ester coupling with decent yields and simple HPLC purification. The results in Figure 4.8 show that the amino-labeled probe set worked well in single molecule RNA FISH (Fig. 4.8A), demonstrating that addition of amino group did not alter the probe's ability to detect single mRNA molecules.

To demonstrate that the amino-labeled probe set can be used to detect single mRNA molecule simultaneously with DNA, we carried out simultaneous single molecule EGFP mRNA FISH and telomere DNA FISH. As shown in Figure 4.8B, the single molecule RNA signals detected by the regular probe set suffered severe damage, while the signals detected by the amino-labeled probe set survived the DNA FISH. The surviving signal quality is satisfying. Although the pinpoint signals are not as sharp as the signals detected by RNA FISH alone, the surviving RNA signals remained as discrete pinpoint signals, present only in the cytoplasm of the EGFP-expressing cells.

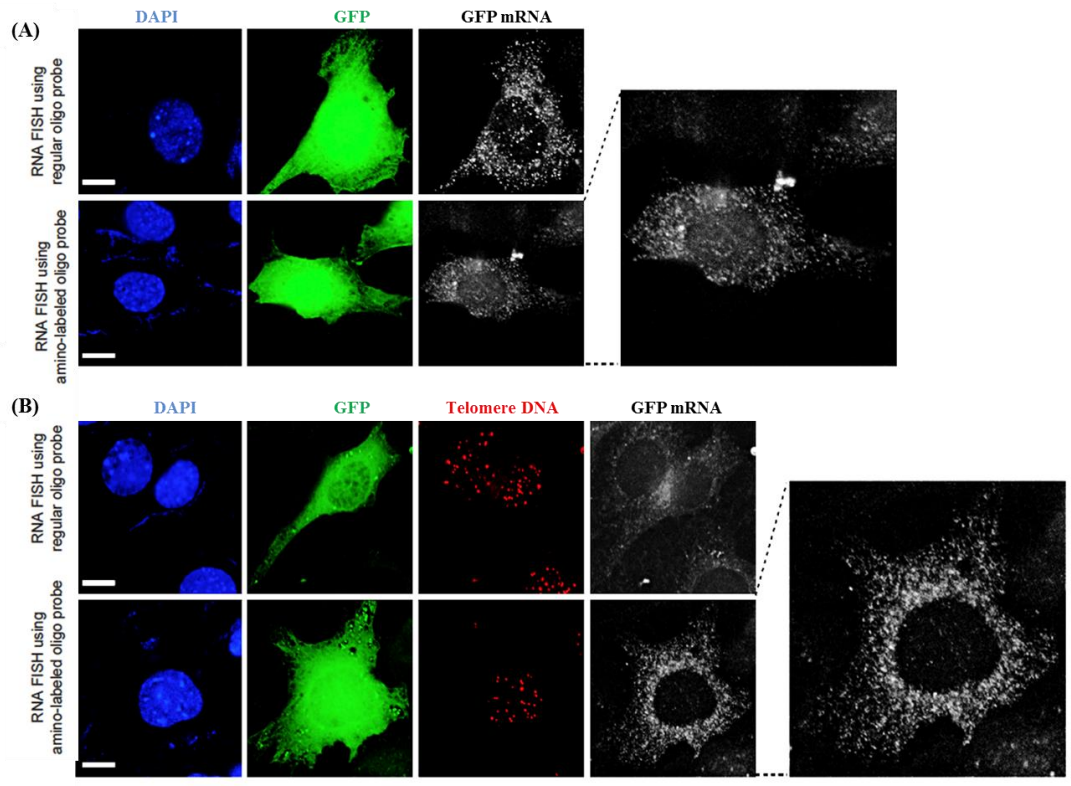


Figure 4.8. RNA-DNA FISH of the single mRNA molecules of EGFP and the telomeres (DNA). Telomere DNA was detected by Cy3-labeled oligonucleotide probes (shown in red). Nuclei were stained by DAPI (blue). The cells used were mouse fibroblast cells transiently transfected by an EGFP-expressing plasmid. The single mRNA molecules of EGFP were detected by a set of 20 oligonucleotides labeled by Cy5 (shown in grey). Two enlarged images of the single molecule RNA signals are shown on the right side. The images of the single molecule RNA signals are deconvolved single-layer z-sections. A scale bar of 10 μm is shown in each DAPI image. (A) RNA FISH only. (B) RNA-DNA FISH.

To quantitatively measure the surviving RNA signal quality, the total fluorescent signal intensity of the individual EGFP mRNA signals per cell was measured. The EGFP protein fluorescent signal was used to mark the cell boundary. The total RNA signal intensity was divided by the total EGFP protein signal intensity to normalize the experimental variation from slide to slide and the EGFP expression level variation from cell to cell. Our results show that the surviving RNA signals detected by the amino-labeled oligo probes are significantly brighter than the surviving RNA signals detected by the regular oligo probes (Fig. 4.9).

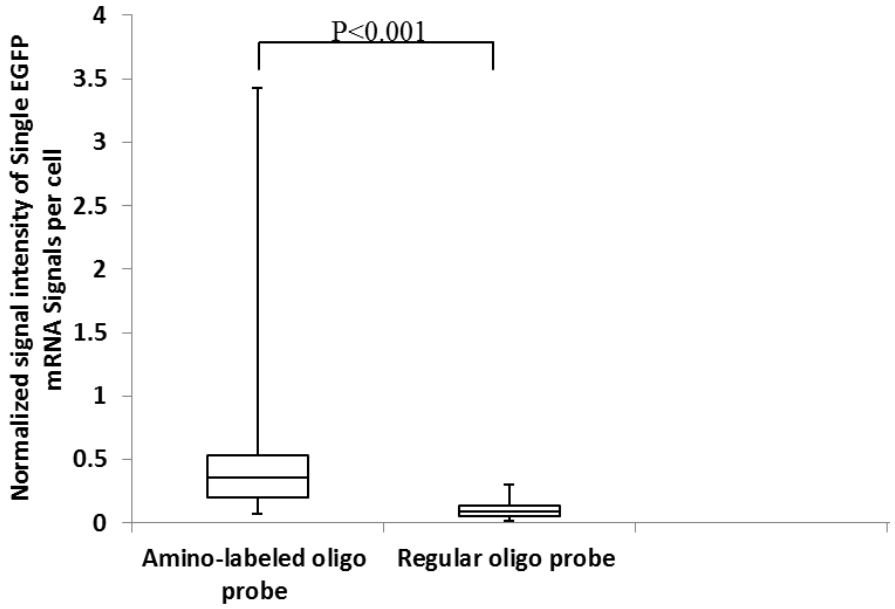


Figure 4.9. Comparative analysis of fluorescence intensity of amino labeled probe and regular probe upon DNA FISH. The total fluorescent signal intensity of the EGFP single mRNA molecule signals per cell (n=50) measured from the Cy5 channel was normalized by the total EGFP fluorescent intensity per cell measured from the "green" channel. The data was presented in box-and-whisker plots. The P-value was calculated using Student t-test.

4.1.3. RNase Protection Assay to study *Terra*

While in the previous section we observed *Terra* expression and its association with telomeric ends *in situ* using modified probes for RNA-DNA FISH, here we use RNase protection assay (RPA) as a method to detect *Terra* in RNA isolated from cells. RPA is a very sensitive technique for quantification and detection of RNA. *Terra* transcription was considered to be unidirectional (Azzalin et al., 2007) but, recent experiments have shown the presence of *Terra* antisense in some organisms for e.g. *Arabidopsis* (Vrbsky et al., 2010). We intend to detect if *Terra* antisense is transcribed in mouse using RPA due to its high sensitivity.

For RPA reactions, we synthesized a radioactively labeled RNA probe complementary to part of the target RNA to be analysed. As shown in Figure 4.10, transcription with T7 RNA polymerase resulted in a RNA product containing a sequence identical to *Terra*. When SP6 RNA polymerase was used, the RNA product contains a sequence complementary to *Terra.* The labeled probes and sample RNA were incubated under conditions that favour hybridization of complementary

sequences. After hybridization, the mixture was treated with ribonuclease to degrade unhybridized probe. Labeled probe that is hybridized to complementary RNA from the sample would be protected from ribonuclease digestion, and was separated on a polyacrylamide gel and visualized by autoradiography.

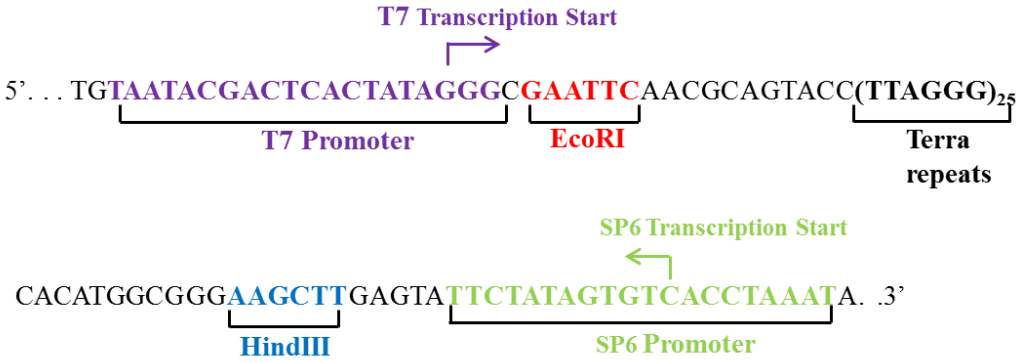


Figure 4.10. Schematic of 'pGEM_Mouse_Telo' plasmid showing the region between T7 and SP6 polymerase. The product from PCR amplication of Telo-C probe is inserted in between the T7 and SP6 polymerase site pf pGEM3Zf plasmid. The insert contains 25 repeats of the mouse *Terra* sequence.

For RPA, we wanted to avoid the presence of longer probe transcripts which contain vector sequences due to transcription run-off. For this purpose, the plasmid was digested with EcoRI before SP6 polymerase reactions and likewise with HindIII before T7 polymerase reactions. The *in vitro* transcription reaction produced a 188 nt radiolabeled RNA probe which protected a 153 nt sequence of mouse *Terra*.

Mouse *β-actin* was used as a positive control for our RPA experiment. The *β-actin* RNA probe was generated from the template plasmid 'pTRI-Actin-Mouse' obtained from Promega. Upon *in vitro* transcription using T7 RNA polymerase 304 nt RNA transcript was produced, which carried a 245 nt sequence complementary to β-actin. A successful RNA protection assay for *β-actin* should result in a 245 nt RNA signal. As shown in Figure 4.11, *in vitro* transcribed *β-actin* RNA probe gave a sharp band on PAGE gel (Fig. 4.11, lane 1). During RNase treatment, the probe was protected by mouse RNA sample and generated a smaller RNA signal on the PAGE gel (Fig. 4.11, lane 5). These results show that the RPA experiment is working.

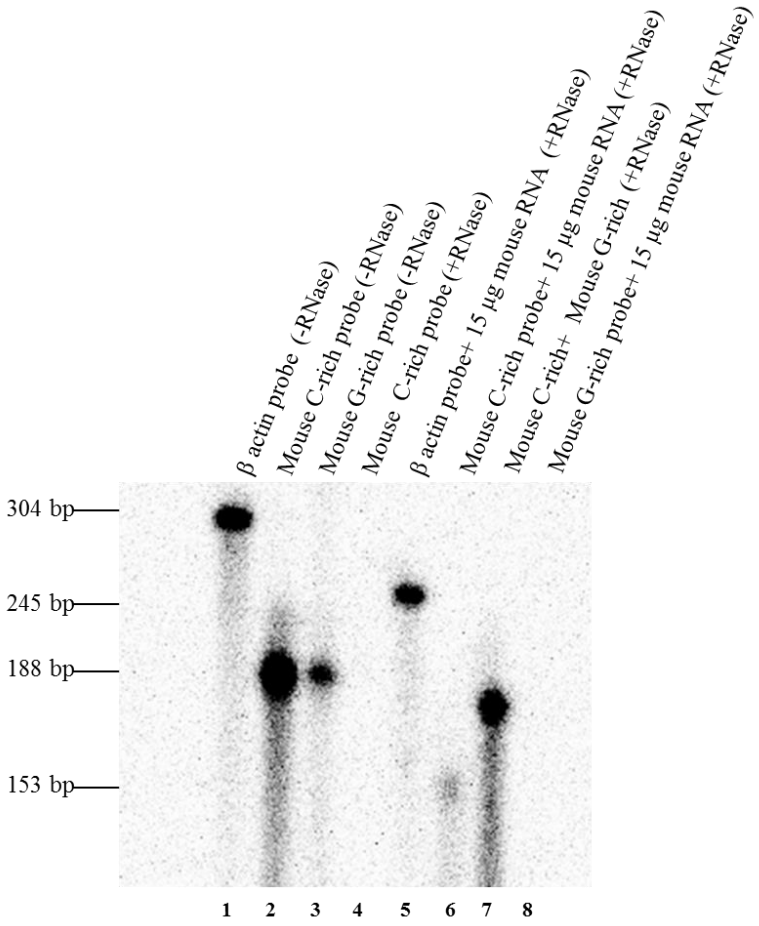


Figure 4.11. RNase protection assay to detect *Terra* in mouse total RNA sample. Mouse *β-actin* probe was used as a control in RPA. Sample RNA was isolated from wild type mouse ES cells.

Lanes 2 and 3 in Figure 4.11 contain unprotected C-rich and G-rich probes without RNase treatment. The C-rich probe is anti-sense to *Terra* and is used to detect mouse *Terra*, while the G-rich has the same sequence as mouse *Terra* transcript and is used as the control probe and to detect *Terra* antisense in the experiment. Both probes showed up as sharp signals with the correct sizes (188 nt) on the PAGE gel. The unprotected C-rich probe was completely destroyed by the RNase treatment (Fig. 4.11, lane 4). On protection of 15 µg of total RNA isolated from the mouse ES cells with C-rich probe, a smaller RNA signal (~153 nt) was detected (Fig. 4.11, lane 6). The protected RNA signal is weak. When the C-rich probe and the G-rich probe were combined before RNase treatment, the two probes protected each other and a resultant RNA signal of ~188 nt was observed. These results show that a *Terra* RNA signal was

detected by RPA from the mouse total RNA sample. The signal could be weak due to the low level of *Terra* presented in the mouse RNA sample probably because of the short half-life. We tried to detect antisense *Terra* but could not detect any signal when we probed for the antisense (Fig. 4.11, lane 8). This may be as a result of even lower quantities of antisense present in our sample. From our study we were not able to confirm or negate the presence of antisense *Terra* in mouse cells.

In summary, we detected *Terra* expression using two sensitive techniques. First we established a simple and robust method for simultaneous RNA-DNA FISH, which can be used to image RNA and DNA targets at the single-cell and single-molecule level. By conjugating alkylamino functionality to nucleotide probes either by oligonucleotide synthesis or nick translation, the current protocol can be applied to detect any potential RNA target and it provides a highly useful tool to study the large pool of nuclear IncRNA *in vivo*. Secondly, by setting up an RNase protection assay system for successfully detecting *Terra* we were able to use it for detection of *Terra* expression in later sections. In the next section we perform further studies to examine the effect of *Terra*-telomere interactions *in vivo*.

4.2. *Terra* and telomere length

Terra is known to be associated with the telomeric heterochromatin and a role of *Terra* in telomere length regulation was first suggested when *in vitro* experiments demonstrated that *TERRA* mimicking oligonucleotides were able to inhibit telomerase by duplexing with its RNA moiety (Schoeftner and Blasco, 2008). Another conceivable scenario advocating telomere length regulation by *TERRA* is, by binding to the telomeric DNA repeats it blocks access of telomerase to the chromosome ends. This scenario was supported by findings showing upregulation of Ribonuclease H (RNaseH) in budding yeast led to telomere lengthening, suggesting that RNA-DNA hybrid might inhibit telomerase action (Luke et al., 2008). As mentioned earlier, *Terra* signals have been seen to associate with the distal telomere, but not with the centromeric end in sex chromosomes of mouse cells (Zhang et al., 2009). With this background information, in the present section we study the impact of *Terra* on telomere *in vivo* by comparing telomere lengths between a wild type mouse cell-line and a telomerase knockout cell-line (*Terc* -/-). Taking into consideration that *Terra* associates specifically on the distal end of X chromosomes in male ES cells, we monitored the telomere length of both ends of the X chromosome in male ES cells over many cell divisions in a long-term cell culture.

4.2.1. Design of telomere length measurement assay

To measure telomere length we decided to use male mouse ES cells for two main reasons. Firstly, they have only one X chromosome, allowing us to specifically monitor its length. Secondly, as ES cells are immortal, we could measure the telomere length over a longer duration. Two male ES cell lines used for long term culture (~100 cell divisions) in this study were J1 (wild type) and F19 (early generation of *Terc* knockout).

Fluorescence intensity of telomere signal in DNA FISH was quantitatively measured for determining the telomere length. We chose DNA FISH signal intensity for our assay, as it was the only available method to provide chromosome specific

information. As shown in Figure 4.12A, ES cells were harvested at selected time points over the long term culture followed by preparation of metaphase chromosome spreads and DNA FISH.

In order to distinguish the X chromosome from the rest, Sx9 probe was used (Fig. 4.12A). Sx9 is a P1 DNA construct carrying a genomic fragment of X chromosome. It could be used as a template for generation of a probe by nick translation. This probe was used as the X chromosome marker and it specifically detected a genomic fragment with a constant length during the experiments. This feature allowed the fluorescence signal intensity of this probe to be used as an internal control to standardize the data from different experiments.

Mouse cells are telocentric in nature. As shown in Figure 4.12B, their centromere is located at one of the terminal ends of each chromosome. The telomeres of the X chromosome could be distinguished as the centromeric end telomeres and the distal end telomeres. In DNA FISH, the telomere signals at the centomeric end could be recognized by the overlapping centromere signals from the DAPI stain **(**Fig. 4.12A**)**. DAPI is a blue fluorescence dye which is readily incorporated into DNA double helix, thus staining all chromosomes. Centromere, being a piece of highly compact heterochromatin, could be easily recognized as a bright signal on one end of each chromosome. An example of DNA FISH is shown in Figure 4.12A.

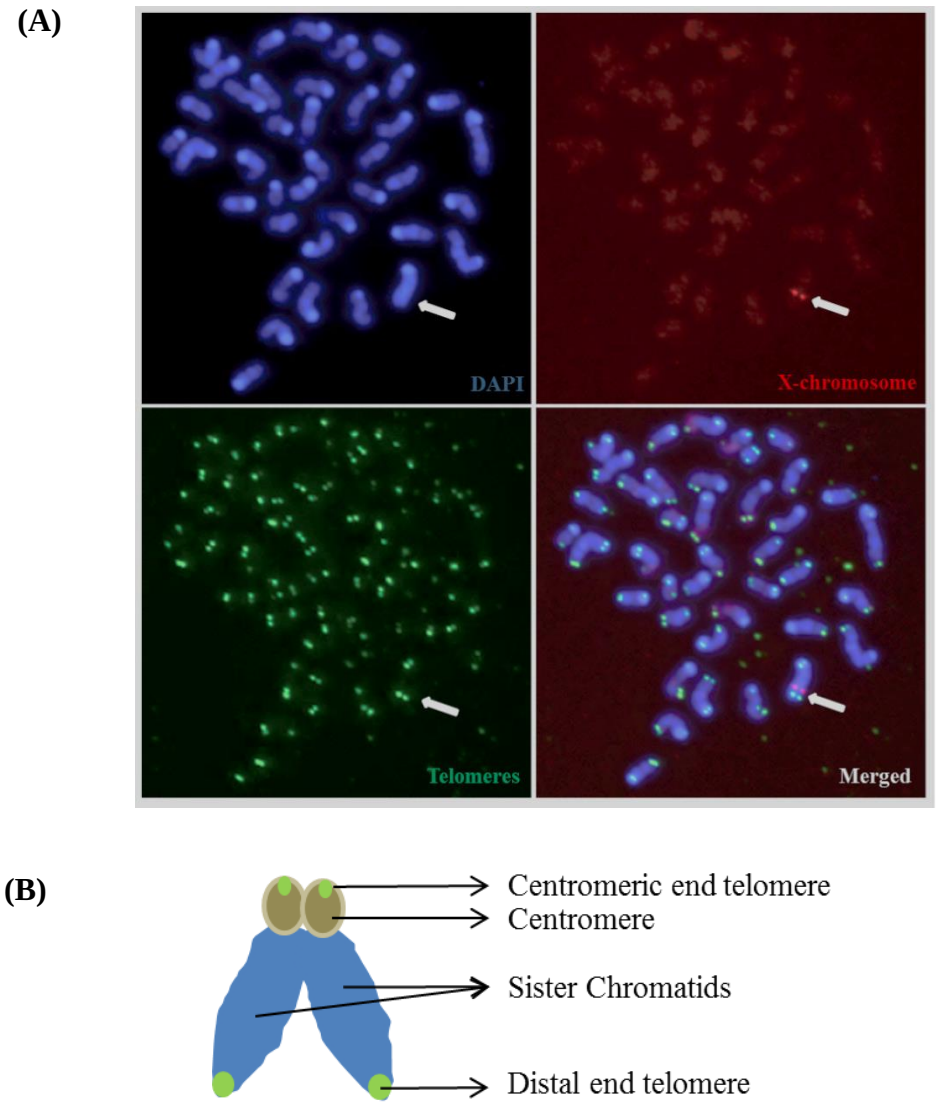


Figure 4.12. DNA FISH on mouse metaphase chromosome spread (A) Image of a cross-section with X chromosome marked by white arrow. Sx9 probe (the X chromosome marker) was labeled by Cy3 (red). Telomere probe, $(TAACCC)_7$, was labeled by FITC (green). Chromosomes are stained by DAPI (blue). (B) Diagram of a typical telocentric mouse metaphase chromosome.

4.2.2. Optimization of fluorescence intensity measurement

We optimized fluorescence intensity measurement with our microscope system Nikon Eclipse Ti-E (Nikon software NIS-Elements AR 3.2). Telomere signals from two sister chromatids of one metaphase chromosome were used as "standard" for system optimization, as the telomere length of both sister chromatids were identical. In

addition, the two Sx9 signals from the two sister chromatids of X chromosome were also used as standards in system optimization.

For fluorescent intensity measurement, each signal was manually defined as "ROI" (region of interest). The fluorescence intensity measurements were available in three formats: mean fluorescence intensity, sum of intensity and maximum fluorescence intensity. We measured telomere signals and Sx9 signals from sister chromatids. For each data pair from one metaphase chromosome, we generated a ratio of their intensity data. As the signals from two sister chromatids should be almost identical (Bekaert et al., 2002; Lansdorp et al., 1996), a perfect ratio in this analysis should be equivalent to "1". We compared the data from mean intensity, sum intensity and maximum intensity measurements and concluded that mean fluorescence intensity measurement provided the highest accuracy during measurement (Fig. 4.13A), as its mean ratio was closest to 1 and it showed the least standard deviation among the three.

An internal control was necessary in this study, which allowed the accurate comparison of telomere lengths across different samples and experiments. To this effect the telomere length measurement was normalized as *relative telomere length*, where the relative telomere length was the ratio of telomere fluorescence intensity and internal control fluorescence intensity. Since the DNA length of Sx9 and centromere was constant across cell divisions, the florescence intensity from both these signals were suitable internal controls. We compared the fluorescent intensity measurements of centromeres and Sx9 signals and selected centromere signal as the internal control in our study (Fig. 4.13B & C).

To validate the established protocol, we measured and compared the X chromosome relative telomere length of J1 (WT) and F19 (*Terc*-/- mutant). As shown in Figure 4.13D, our protocol was able to detect the telomere length difference between the wild type and telomerase knockout cell-line.

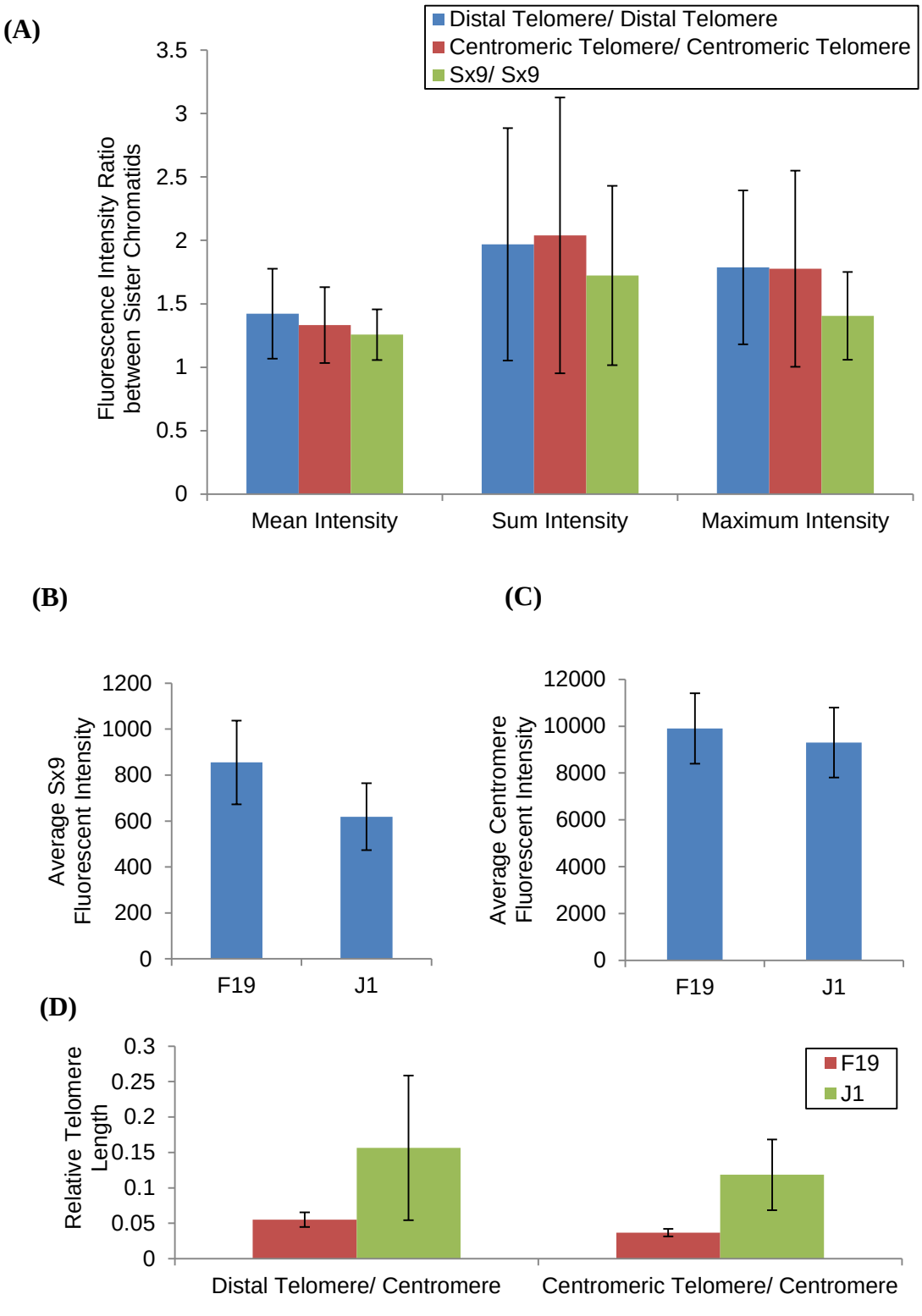


Figure 4.13. Optimization of fluorescence intensity measurements. (A) On J1 cells, telomere and Sx9 fluorescence intensity data were collected from two sister chromatids in three formats: mean, sum and maximum fluorescence intensities (N=50). (B) Fluorescence intensity of Sx9 (N=50). (C) Fluorescence intensity of centromere (N=50). (D) Relative telomere length measurements on two ends of X chromosome in J1 and F19 cells (N=50).

4.2.3. Long term culture of J1 and F19 cells

J1 and F19 ES cells were continuously grown in culture over a period of three months. Cells were sub-cultured every 3 days. Cell numbers was counted at each sub-culture and cell growth was tracked for the entire cell culture period. Figure 4.14 shows the cumulative cell divisions over passages. J1 had 150 cell divisions while F19 had 99 cell divisions within the same frame of time (Table 4.2). J1 consistently showed higher proliferation rate as compared to F19. This phenotype of F19 cells was due to the absence of a functional telomerase leading to the loss of telomere repeats during each round of replication thereby reducing the cellular proliferation rate of this cell-line and leading to growth arrest.

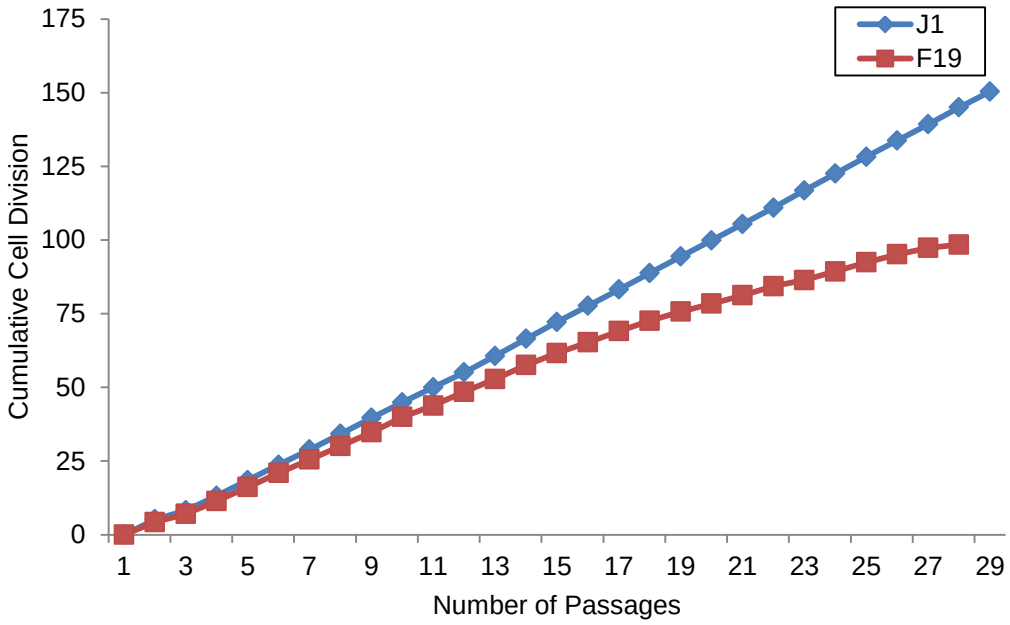


Figure 4.14. Growth curves of J1 and F19 ES cells over 100+ cell-divisions.

During the long-term cell culture, we harvested cells in metaphase spread format at 5 selected time points (Table 4.2). DNA FISH and telomere length measurements were carried out on the harvested samples.

Table 4.2. Cells harvested at the five mentioned time points. The corresponding cell passages and cell divisions have been given.

J1 (WT)		**F19 (*Terc*-/-)**	
Passage number (P)	Corresponding cell Division	Passage number (P)	Corresponding Cell Division
1	0	1	0
8	34	8	30
15	72	15	62
22	111	22	84
29	150	28	99

4.2.4. Telomere length measurements

With the established protocol, we analyzed the samples harvested from the long term cell culture. We successfully detected the gradual telomere shortening in F19 cells during the long term cell culture (Fig. 4.15). These data showed that our method was sensitive enough to monitor the telomere shortening of F19 cells across 100 cell divisions.

We observed in the F19 cells that the distal end of the X chromosome was shortened at a similar rate as the centromeric end of the X chromosome (Fig. 4.15A & B). In samples from all the selected time points, the telomere at the centromeric end of the X chromosome was significantly shorter than the telomere at the distal end. We analyzed a random sampling of autosomal ends. The data showed that the significantly shortened centromeric end was a feature specific to the X chromosome.

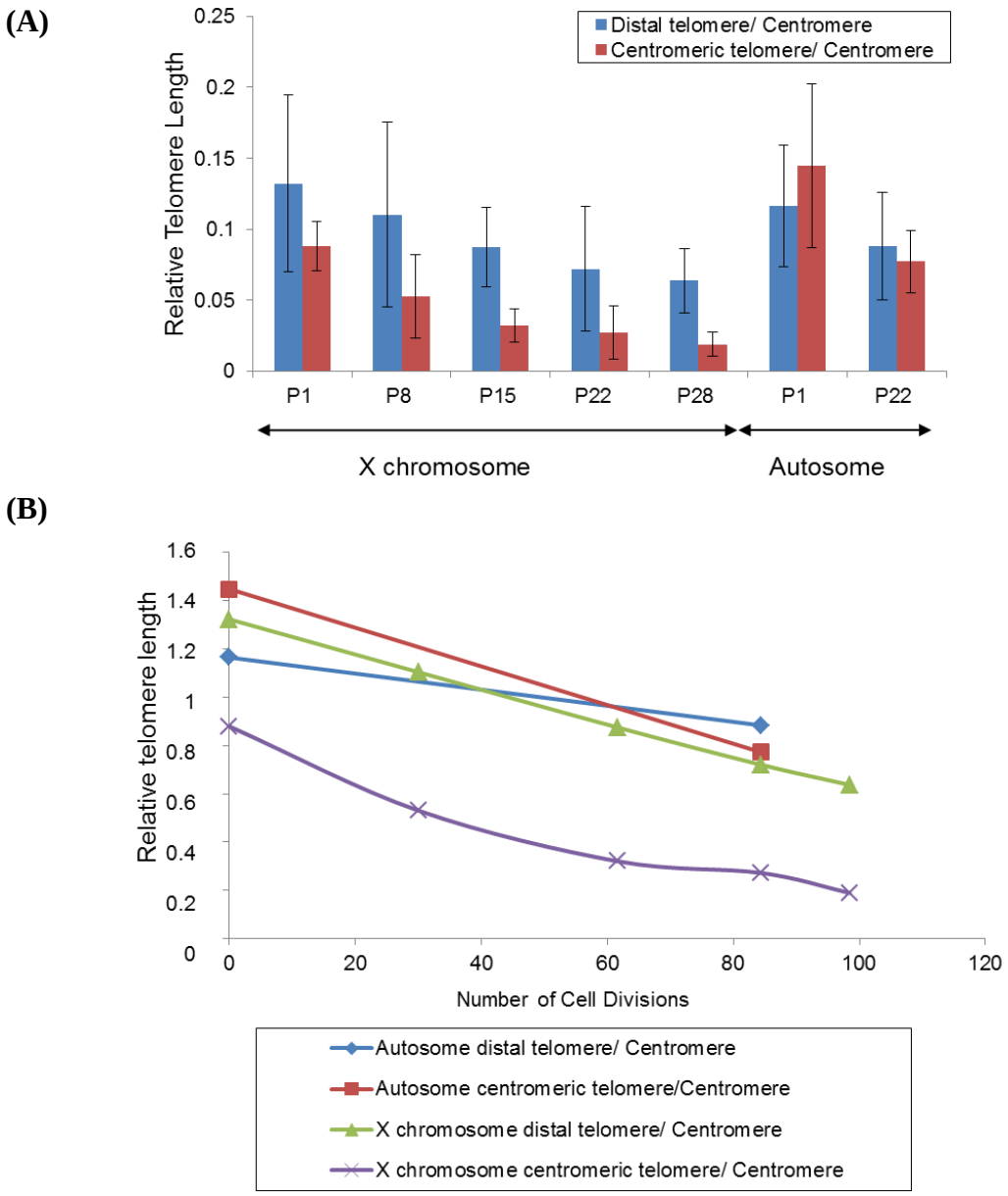


Figure 4.15. To monitor telomere length changes in telomerase knockout F19 cells during long term cell culture. (A) Bar graph displaying relative telomere length variation between centromeric and distal end over ~28 passages (P=passage number). (B) Line graph displaying relative telomere length variation across 100 cell divisions (N=20). Autosome telomere length was measured as mean telomere length from a random sampling of autosomes

Next, we analyzed the samples from J1 cells (wild type). The telomere length in J1 cells showed a fluctuated pattern across the long term cell culture (Fig. 4.16A & B). Comparing the distal end of X chromosome, which is specifically covered by *Terra*,

with the centromeric end, we did not recognize any significant difference between the two ends in their telomere length patterns.

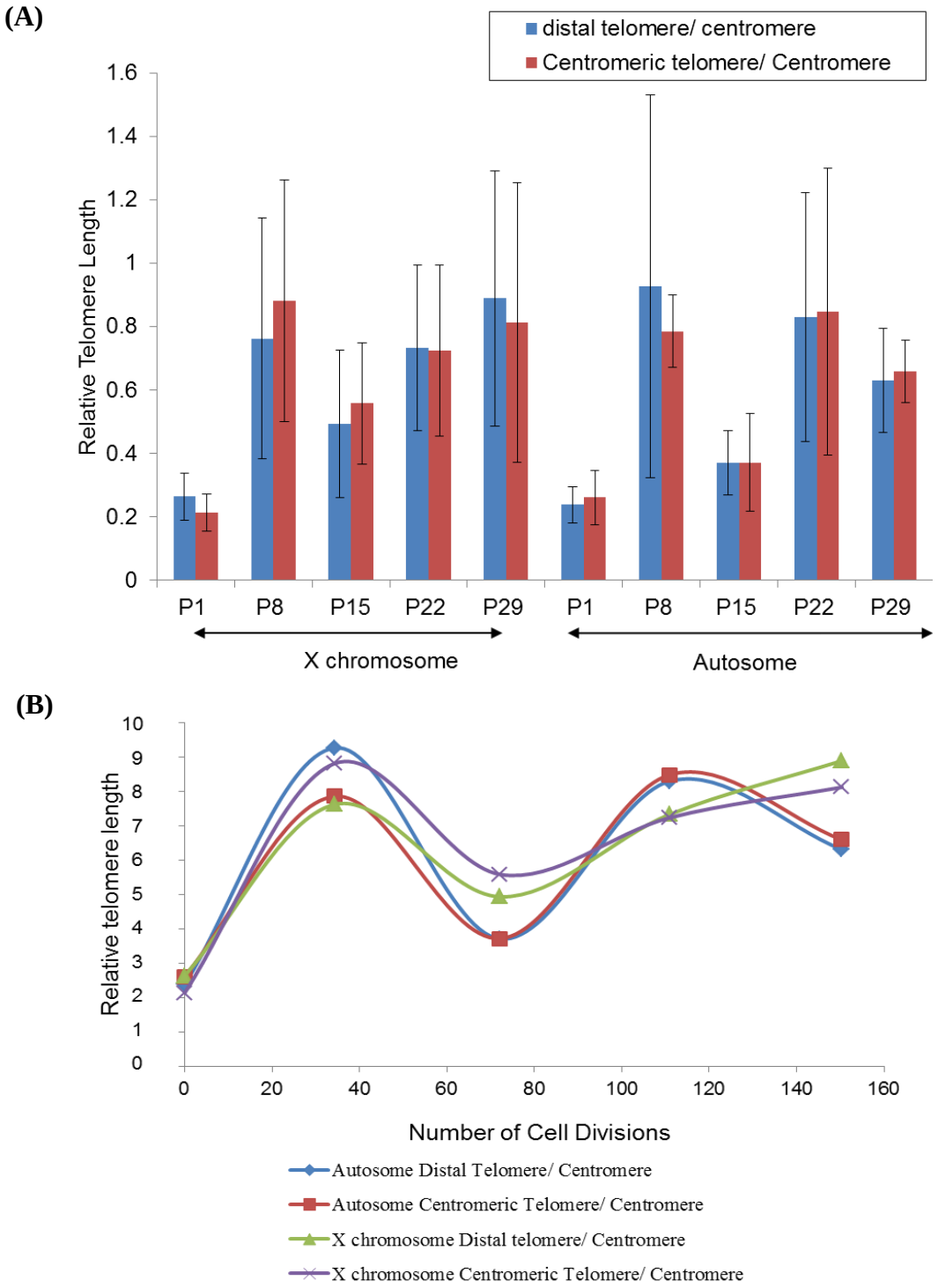


Figure 4.16. To monitor telomere length changes in wild type J1 cells during long term cell culture. (A) Bar graph displaying relative telomere length variation between centromeric and distal end over ~29 passages (P=passage number). (B) Graph displaying relative telomere length variation across 100 cell divisions (N=20). Autosome telomere length was measured as mean telomere length from a random sampling of autosomes

Additionally, there wasn't any significant difference in telomere lengths between X chromosome and autosomes in J1 cells. Hence, we did not observe a negative effect of *Terra* on telomerase-mediated telomere elongation *in vivo*. This could be due to technical reasons. Like our assay could be not quantitative enough to detect the effect. Another reason could be attributed to the period of cell culture which was not long enough to reveal the effect. Moreover, there could be other unknown mechanisms that are involved in the regulation of *Terra*, possibly inhibiting its function. One more scenario could be the maintenance of telomere length by an alternative mechanism by homologous recombination which is telomerase independent and therefore, not affected by *Terra* association.

4.2.5. Chromosome fusion in the late passages of F19 cells

In this study, we made an interesting side observation. Chromosome fusions occurred at high frequencies on sex chromosomes in the late passages of F19 cells. In compliance with previous reports, F19 ES cells fell in growth arrest at the end of our culture period (Niida et al., 1998). The colony morphology of growth arrested F19 cells could be easily distinguished from normal ES cell colonies (Fig. 4.17). Normal ES cells form colonies that are oval, shiny, with distinct boundaries. Upon growth arrest, F19 cells showed colonies with irregular shape and no distinct boundaries.

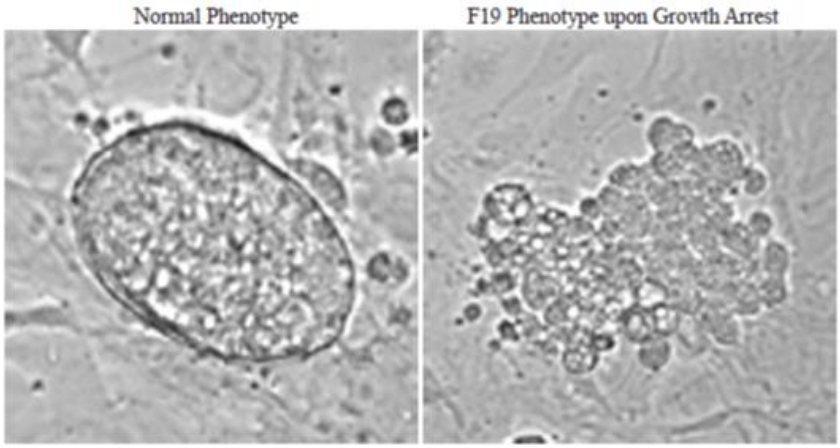
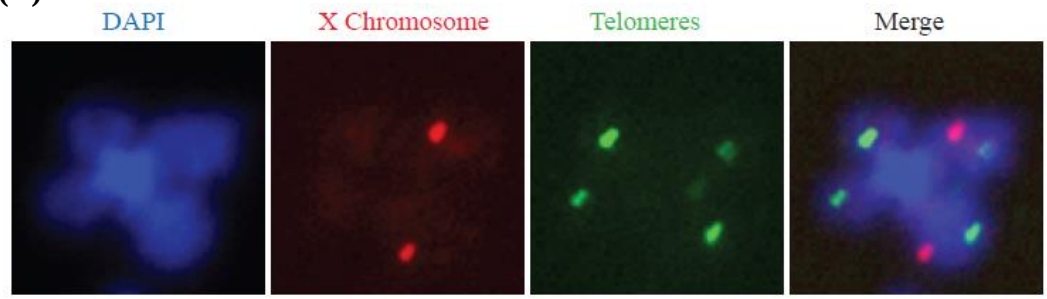


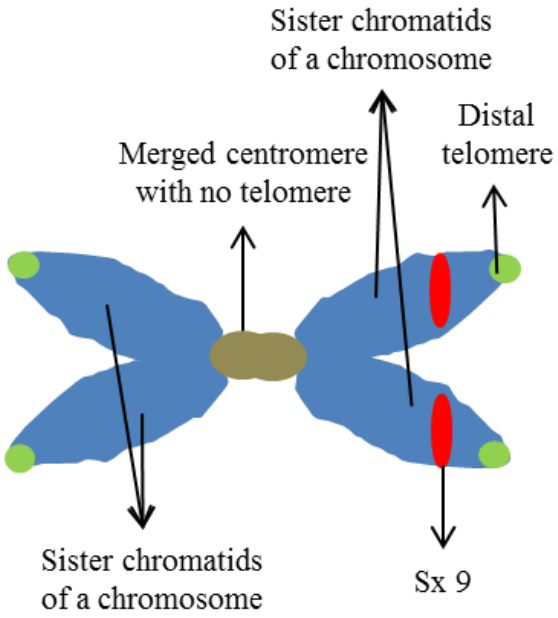
Figure 4.17. Morphological studies on wild type and telomerase knockout cell line. The morphology of a normal ES cell colony from J1 cells is compared with an ES cell colony from F19 cells upon growth arrest, using DIC.

Similar to the observation made by Niida et al, we also detected chromosome fusion in the late passages of F19 cells. Chromosome fusion was an expected phenotype for chromosomes with extremely shortened telomeres. In the metaphase chromosome spreads, true chromosome fusion could be clearly distinguished from accidental head-to-head localization of two chromosomes. As shown in Figure 4.18A and B, in a true chromosome fusion, centromeric telomere signals were missing from the fused chromosomes. In an "accidental head-to-head localization" (Fig. 4.18C & D), the centromeric telomere signals could be clearly recognized in between the two chromosomes.

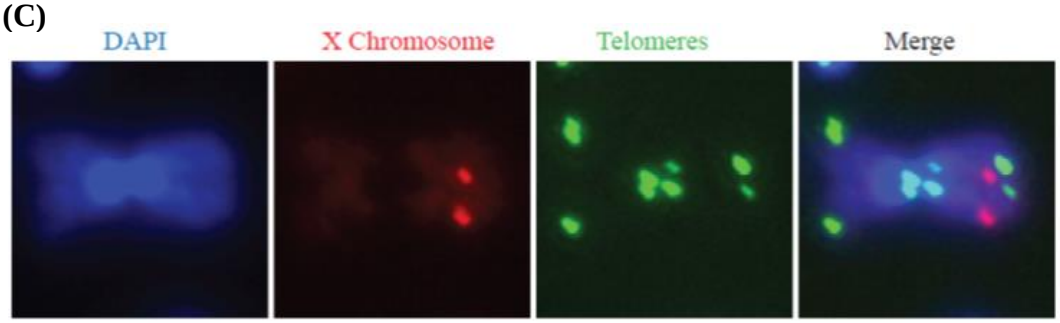
(A)



(B)



(C)



(D)

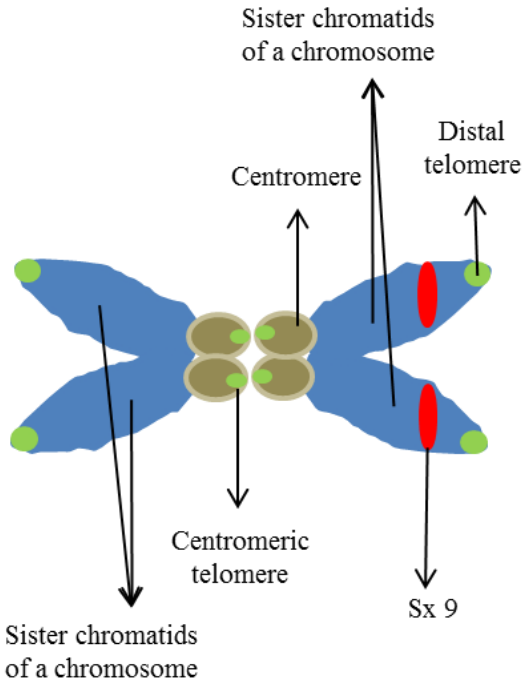


Figure 4.18. Study of chromosome fusion in F19 cells. (A) An example of X chromosome fused with an autosome in F19 Passage 22. (B) Diagram representing chromosomal fusion. (C) An example of two chromosomes being positioned closely but not fused. (D) Diagram representing two chromosomes that were closely positioned together but not fused.

As shown in Figure 4.19C, chromosome fusion started to appear in passage 8 population. At passage 22, F19 cells fell in growth arrest, and about 20% of the chromosomes are fused in each chromosome spread and more than 70% of the X chromosomes are fused. For our curiosity, we also checked the Y chromosome using DNA FISH. Interestingly, almost all Y chromosomes are fused in passage 22 population of F19 cells (Fig. 4.19A). Both X chromosome and Y chromosome were fused with an autosome partner (Fig 4.19A & B). We did not observe an X-Y

chromosome fusion. Most of the chromosome fusions occurred between two centromeric ends. Fusions at the distal ends were rarely seen.

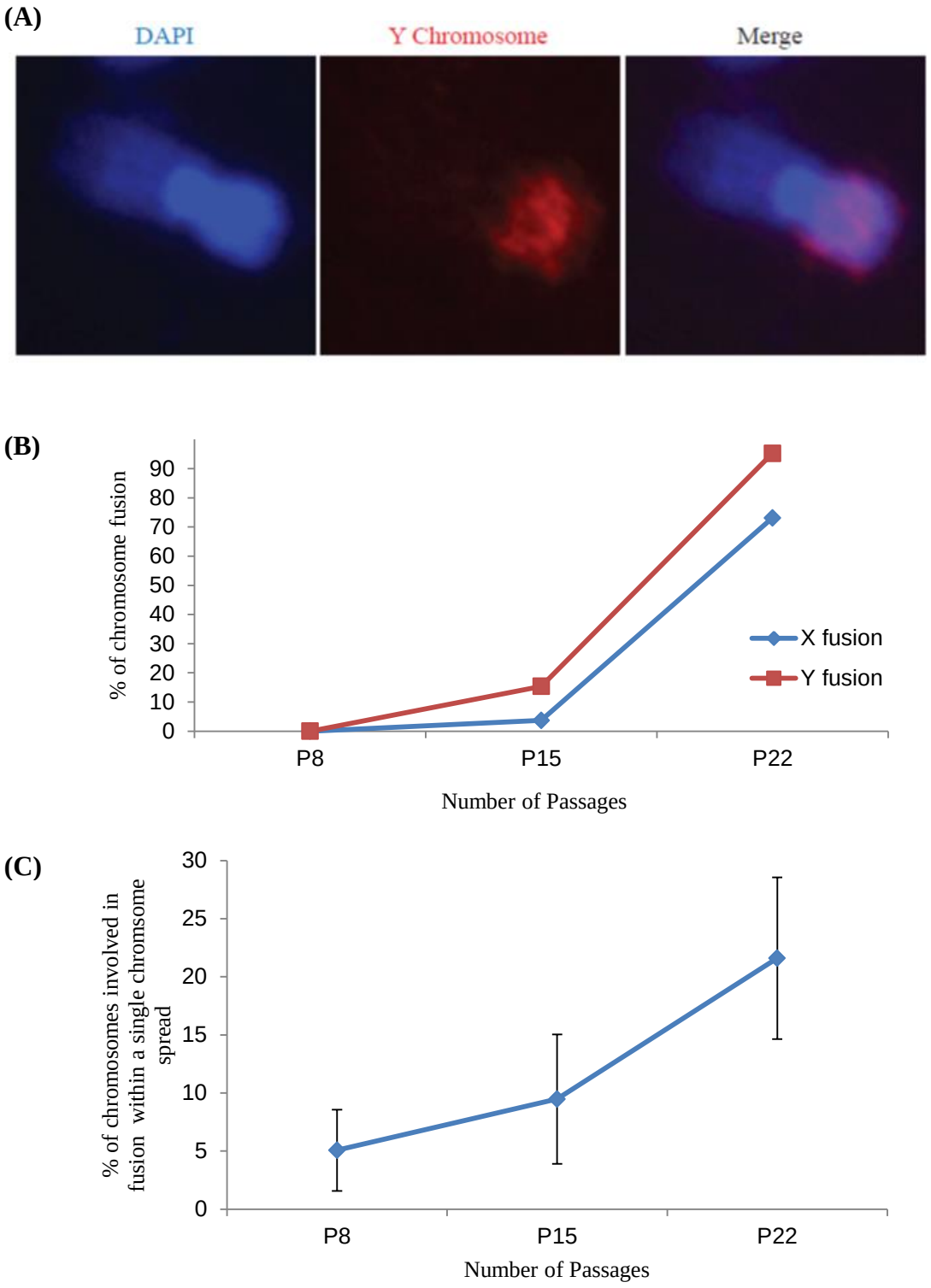


Figure 4.19. Study of sex chromosome fusions in late passages of F19 cells. (A) DNA FISH using Y-chromosome paint detected fusion of Y chromosome with autosomes. (B) Percentage of X and Y chromosomes being fused among total number of X and Y chromosomes. (N=25) (C) Average percentage of chromosomes fused in a chromosome spread. (N=25)

To summarise, in this section we examined the telomere length change of X chromosome over a long-term cell culture in a mouse ES cell-line. This study provides a method to detect effects of *Terra* on telomere *in vivo*. With prior knowledge that one of the two ends of X chromosome is specifically associated with a bright and large *Terra* signal, we compared the telomere length at the two ends of X chromosome**.** Our observations here do not support *Terra's* role as a telomerase inhibitor *in vivo*. This is in agreement with the results in our previous section where association of *Terra* was not dependent on the length of telomeres. This suggests that *Terra* could perform functions other than telomere length regulation. In the next section we look more into the various biological roles of *Terra.*

Another interesting outcome of this study was the finding that X chromosome is under high pressure to maintain its telomere length. In a telomerase knockout ES cell line, X chromosome is one of the earliest chromosomes to show chromosome fusion, a sign of a chromosome suffering severe telomere length shortening. These observations are in fact consistent with a new hypothesis put forth by a recent live cell imaging study that *Terra* may recruit telomerase onto a telomere to maintain its length (Cusanelli et al., 2013). With this hypothesis, our observations can be reasonably explained. X chromosome is under high pressure to maintain its telomere length; therefore, *Terra* is transcribed at a higher level on its telomere to recruit telomerase. Therefore, the association of *Terra* with sex chromosomes could be involved in recruitment of telomerase enzyme to prevent telomere shortening and chromosome fusion which otherwise lead to growth arrest.

4.3. *Terra* and stress response

Apart from its possible role in telomere length regulation, *Terra* may have other functions in the cell as its levels are governed by several mechanisms, which include developmental status, cellular stress, and epigenetic state of the cell (Azzalin et al., 2007; Caslini et al., 2009; Deng et al., 2012a; Schoeftner and Blasco, 2008; Zhang et al., 2009). The effect of cellular stress can lead to either upregulation or downregulation of *Terra* transcript levels. In this section we take a look at the effect of heat induced stress on *Terra* expression pattern and make an attempt to elucidate some factors which are involved in the regulation of *Terra* transcription in mouse under normal conditions and in response to heat shock.

4.3.1. Heat shock and *Terra*

Among various types of cellular stresses present, for e.g. UV exposure, gamma irradiation and heat stress, we observed an upregulation in *Terra* transcription in response to heat shock. On incubating mouse ES cells at 45°C for 45 minutes we were able to detect a visible increase in *Terra* transcription by RNA fluorescence *in-situ* hybridization (Fig. 4.20).

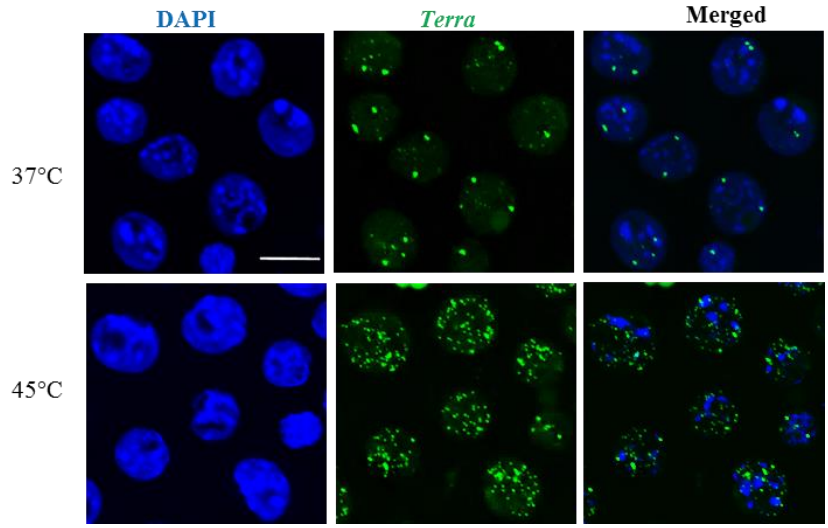


Figure 4.20. *Terra* RNA FISH on the mouse embryonic stem cell-line J1 at 37°C (control) and 45°C (heat-shock). A FITC labeled DNA oligonucleotide probe was used for detecting *Terra*. Scale bar, 10 µm.

To quantitatively measure the transcription upregulation of *Terra* on heat shock we compared the volume of the *Terra* signal in stressed and non-stressed cells. Volume measurement was made using Nikon NIS-Elements microscope imaging software. A threshold value was set for signal fluorescence intensity and the software measured the volume of *Terra* signals in a cell equal to or above the threshold value. We observed a three-fold increase in the levels of *Terra* transcript upon heat shock (Fig. 4.21). Even though we observe an upregulation of expression, distribution pattern of *Terra* exhibiting two bright foci and few smaller pin point signals was similar in normal and heat-shocked mouse ES cells with a marked increase in the number of smaller pinpoint signals on heat shock which led to an increase in volume of *Terra* signal.

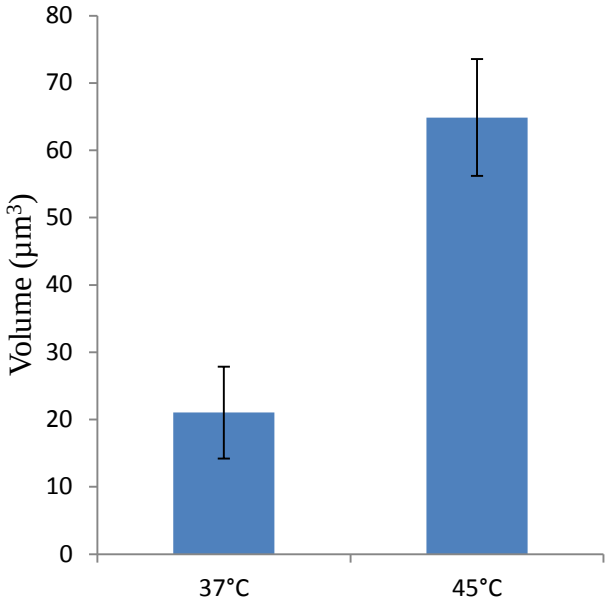


Figure 4.21. Comparison of the volume of *Terra* signal between J1cells at 37°C (control) and 45°C (heat-shock). Error bars show standard deviation (n=50 cells).

We further tested the levels of *Terra* transcription by q-PCR from different telomeric ends in a mouse ES cell line (J1). Mouse chromosomes are telocentric with their centromere being located at one end of the chromosome. The telomere extends from both ends of the chromosome, i.e. from the end of the 'q arm' (long arm) and the centromeric end of the telocentric chromosome. The designed primer set detected *Terra* being transcribed from telomeric end of the 'q arm' of chromosome 2, 11 and 18

as well the telocentric end of mouse chromosomes. We observed that different telomeric ends respond differently to heat shock. Among the telomeric ends tested in this study, *Terra* transcription from 2q and TeloCentric end did not show a big difference before and after heat shock. While *Terra* transcription levels were seen to increase from the 11q and 18q arms of the mouse chromosomes. The increase in expression level was highest at 18q telomeric end (Fig. 4.22).

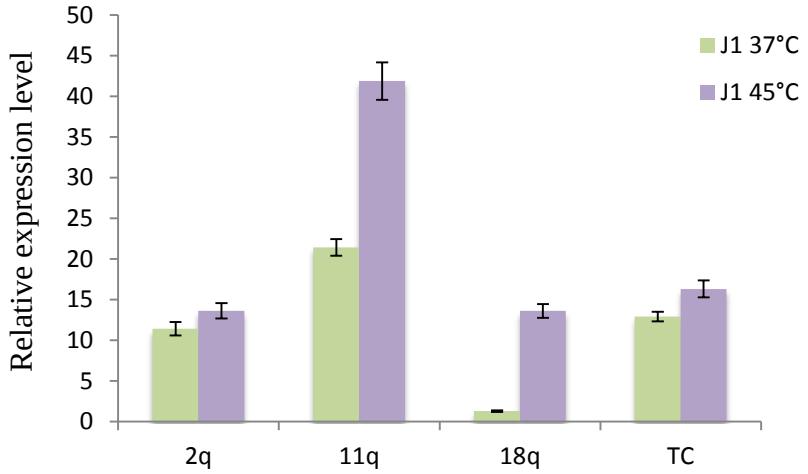


Figure 4.22. Effect of heat shock on *Terra* expression in J1 cells. *Terra* expression from the 2q, 11q, 18q chromosomal ends and TeloCentric (TC) region was studied by qPCR. The error bars represent standard deviation. qPCR was performed in set of triplicates.

Next we looked at the distribution pattern of *Terra* expression. *Terra* foci have been demonstrated previously to be either overlapping or non-overlapping with the telomeric ends of a chromosome. The non-overlapping *Terra* signals can further be distinguished into signals located adjacent to the telomeric end or signals located separately (Fig. 4.23) (Lai et al., 2013).

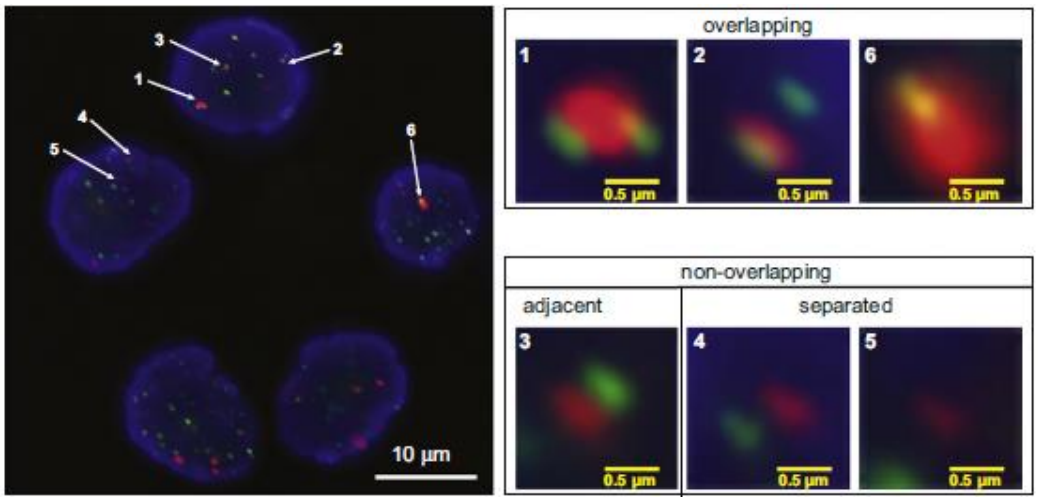


Figure 4.23. *Terra* RNA-DNA FISH. Staining was performed using a Cy3-labeled PNA probe (red) for RNA FISH. The RNA signals were post-fixed, before DNA FISH was performed using a FITC-labeled PNA probe (green) carrying the G-rich telomeric sequence. The panel on the left is a deconvolved single focal plane of a cross section of cells. The right panel demonstrates enlarged views of 6 selected *Terra* signals together with each signal's corresponding number and its surrounding telomere DNA signal. Signals numbered 1, 2 and 6 show overlapping *Terra* and telomeric DNA. In 3, the signal is adjacent to the telomere DNA; while *Terra* and telomere DNA signals are completely separated in 4 and 5. (Lai et al., 2013)

To test the distribution pattern of *Terra* with telomere DNA upon heat shock we performed a simultaneous RNA-DNA FISH. (Fig. 4.24)

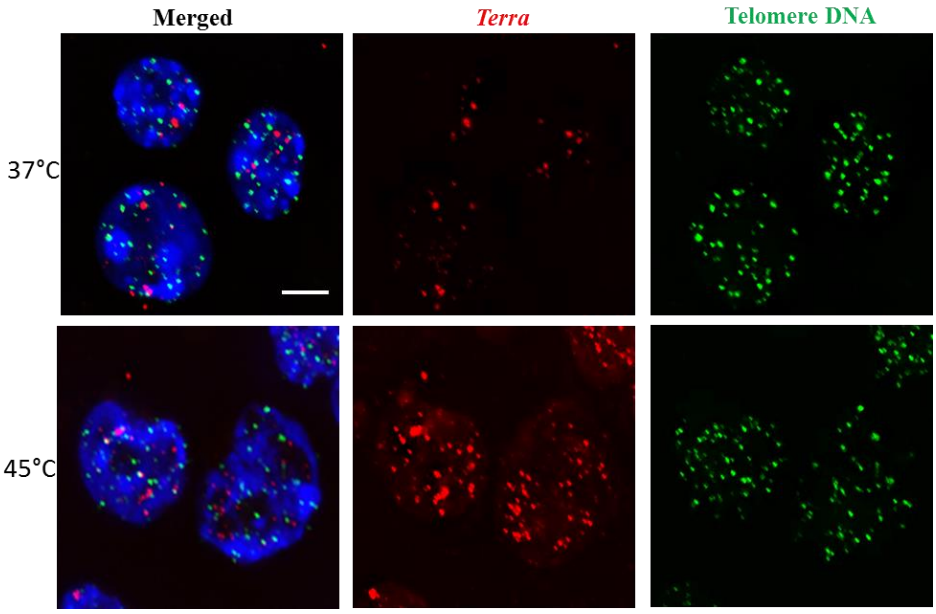


Figure 4.24. Simultaneous RNA-DNA FISH performed on non-stressed (37°C) and heat-stressed (45°C) J1 cells. RNA FISH was performed by using Cy3 (red) labelled probe and DNA FISH by a FITC (green) labelled probe. Scale bar, 5 µm.

We counted the number of *Terra* signals on each plane of a z-series image in 100 cells and examined the interaction of these signals with the telomere DNA. The number of *Terra* signals increased to 395 in cells post-heat-shock as compared to 273 in the absence of heat shock. Although the number of signals showed an obvious change, we can see from Figure 4.25 that the association pattern of the signals with the telomeric ends was similar in the presence and absence of heat-shock.

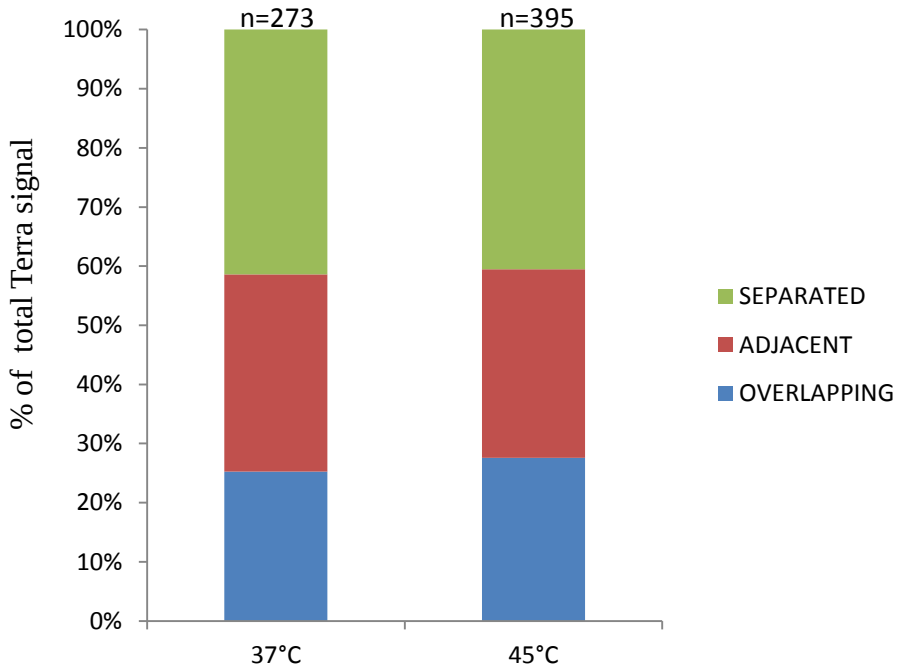


Figure 4.25. ***Terra*** **expression upregulation and its association with telomeric ends in J1cells in response to heat shock.**

Hence, from our experiments we concluded that there is an obvious increase in *Terra* expression upon heat shock. As seen from our qPCR data, this increase could be a result of higher sensitivity of certain telomeric ends to heat shock. Even though the number of signals increased, their interaction with telomeric ends remained same. This observation further strengthens our hypothesis that *Terra* has functions in the cell which are independent of its interaction with the telomeres. Our next aim was to look for cellular factors regulating *Terra* expression.

4.3.2. NFκB inhibition and *Terra* expression

Nuclear factor-kappaB (NF-κB) is a stress-regulated transcription factor belonging to the Rel family and is made up of p65 (RelA), RelB, cRel, p50 and p52 subunits. It plays a key role in regulating over 150 target genes including the expression of inflammatory cytokines, chemokines, immunoreceptors, and cell adhesion molecules. It has been demonstrated that NF-κB is activated during the recovery period after heat shock and is essential for cell survival after heat shock by activating mechanisms that probably helps the cell to cope with hyperthermic stress through clearance of damaged proteins. Heat shock induces many cellular alterations such as protein misfolding and aggregation (depending on severity) (Pinto et al., 1991), cell cycle arrest (Kuhl and Rensing, 2000), modulation of membrane fluidity (Balogh et al., 2005), redox state (Davidson and Schiestl, 2001) and transcription modulation (Park et al., 2005). During heat stress, heat shock genes are preferentially transcribed (Morimoto and Santoro, 1998). The resulting heat shock proteins (Hsps) are molecular chaperones that help the cell to cope with misfolded proteins by either refolding them or addressing them to the degradative systems (Young et al., 2004). As the transcription regulator NFκB is involved in upregulating expression of a number of genes during the period of heat stress, it would be interesting to find out if NFκB regulates *Terra* transcription as well.

To test this idea, we decided to use BAY-11-7082 (Sigma Aldrich #B5556), a known NFκB inhibitor. In the cytoplasm, inactive NFκB complexes are bound to its endogenous inhibitor IκB. However, in response to certain stimuli, IκB is phosphorylated, ubiquitinated, and degraded, enabling the translocation of NFκB to the nucleus. BAY-11-7082 selectively and irreversibly inhibits NFκB activation by blocking phosphorylation of IκB-α.

The goal of our experiment was to inhibit NFκB in cells which are heat stressed and observe its effects on *Terra* expression. We performed RNA FISH to detect *Terra* in control samples (no heat stress) and heat stressed samples in the presence or absence of the inhibitor molecule (Fig. 4.26). On analyzing the RNA FISH results, we did not observe a change in *Terra* transcription in the presence of NFκB inhibitor. This could be due to an absence of any involvement of NFκB in *Terra* transcription in

response to heat shock. In the next section we look at a few more transcription factors that may play a role in *Terra* transcription regulation.

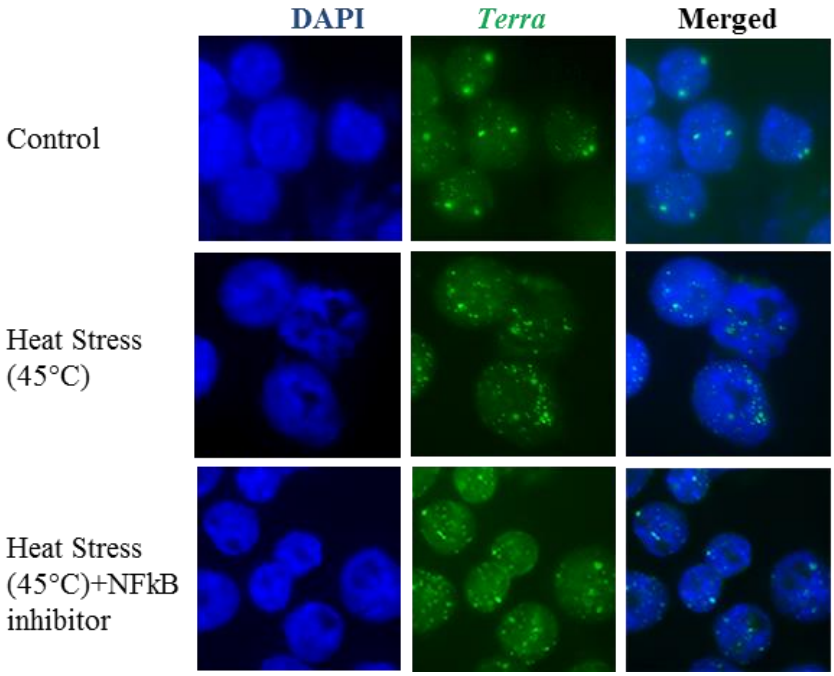


Figure 4.26. *Terra* RNA FISH on mouse ES cell-line J1 in control (37°C), heat-shock (45°C) and heat shock samples in the presence of BAY-11-7082, an NFκB inhibitor molecule. FITC labelled DNA oligo probe was used for targeting *Terra*.

4.3.3. Effect of shRNA knockdown of *CTCF* and *Rad21* on *Terra* expression

The chromosome region immediately adjacent to the terminal repeats have been referred to as subtelomere. The contribution of human subtelomeric DNA and chromatin organization to telomere integrity and chromosome end protection is not yet understood in complete molecular detail. It has however been shown by ChIP-Seq that most human subtelomeres contain a CTCF (CCCTC Binding factor) and cohesin-binding site within ~1–2 kb of the TTAGGG repeat tract and adjacent to the CpG-islands implicated in *TERRA* transcription control (Deng et al., 2012b). In human cells, the distal subtelomeres consist of a variety of degenerate repeat elements with a few discrete gene transcripts interspersed at various distances from the terminal TTAGGG repeat tracts (Ambrosini et al., 2007; Linardopoulou et al., 2007; Riethman, 2008a, b;

Riethman et al., 2005). *TERRA* transcription in human cells initiates from within the subtelomeres, and a promoter containing a CpG-island and subtelomeric 29- and 61-bp repeat element has been identified in plasmid reconstitution assays (Nergadze et al., 2009). ChIP-Seq also revealed that RNA polymerase II (RNAPII) was enriched at sites adjacent to the CTCF sites and extending towards the telomere repeat tracts.

Mouse subtelomeric regions have not been studied in such detail, hence we are unable to identify a definite *Terra* transcription start site in mouse cells. If we were to assume that the transcription control of *Terra* in mouse is similar to that in human, then the factors CTCF and Rad21 could be involved in *Terra* transcriptional regulation in mouse cells also.

To test this possibility, we decided to knockdown (k/d) *CTCF* and *Rad21* in mouse embryonic stem cells and study the transcription of *Terra* under normal and heat stressed conditions. By using the pSUPER RNAi system we were able to knockdown the levels of *CTCF* and *Rad21* substantially (Fig. 4.27A and B) by shRNA based method. The effect of knockdown on *CTCF* being more pronounced than that on *Rad21*.

(A) **(B)**

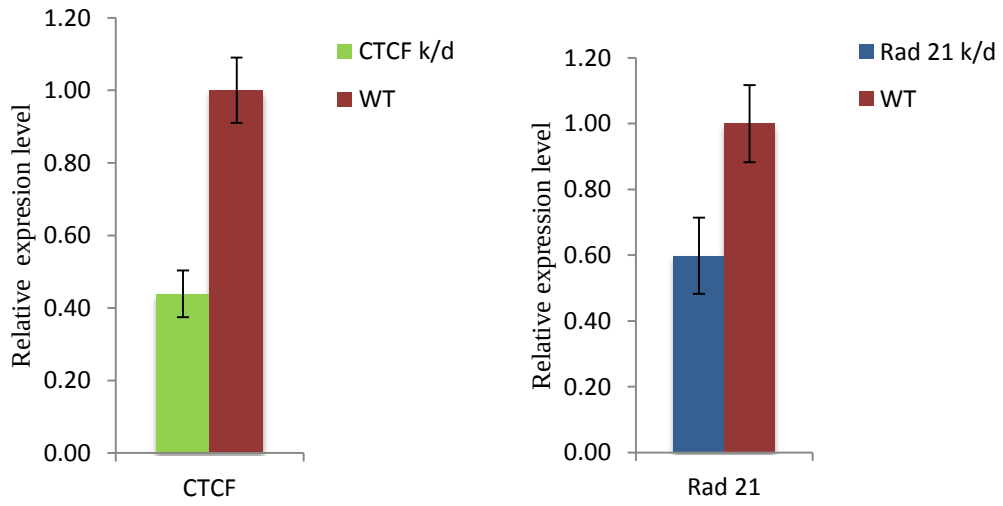
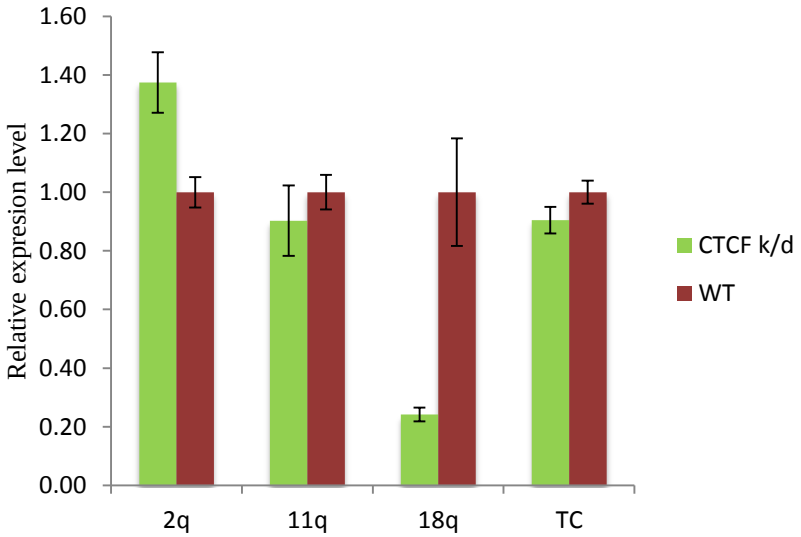


Figure 4.27. Knockdown of *CTCF* and *Rad21* transcripts by shRNA. (A) Comparison of *CTCF* levels in *CTCF* knockdown (k/d) cell-line and wild type (WT) J1 mouse embryonic stem cells by qPCR. (B) Comparison of *Rad21* levels in *Rad21* knockdown (k/d) cell-line and wild type (WT) J1 mouse embryonic stem cells by qPCR. Expression levels of genes were normalized to beta-actin mRNA. Error bars represent standard deviation of mean.

Our next goal was to quantify the levels of *Terra* transcribed from different chromosomal ends in the cell-lines in which we knocked down *CTCF* and *Rad21* expression. We looked at 2q, 11q, 18q and TeloCentric (explained previously) ends for *Terra* transcription levels. From the quantitative PCR results we could deduce that among the studied chromosomal ends, *Terra* transcription responded differently in *CTCF* and *Rad21* k/d cell lines (Fig. 4.28).

In the *CTCF* k/d cell-line, except for *Terra* transcribed from 18q telomeric end the rest of the chromosomal ends studied did not show a change in transcription as compared to a wild type (WT) cell line (Fig. 4.28A). When we analyzed the *Terra* transcription in *Rad21* k/d cell-line, we found that all the ends studied showed a decrease in *Terra* transcription in comparison to the WT cell-line. The effect was maximal on the 18q telomeric end in this set also (Fig. 4.28B). Therefore, from our study we observed that *Terra* transcription from 18q telomeric end was sensitive to *CTCF* and *Rad21* levels and could be regulated by these factors.

(A)



(B)

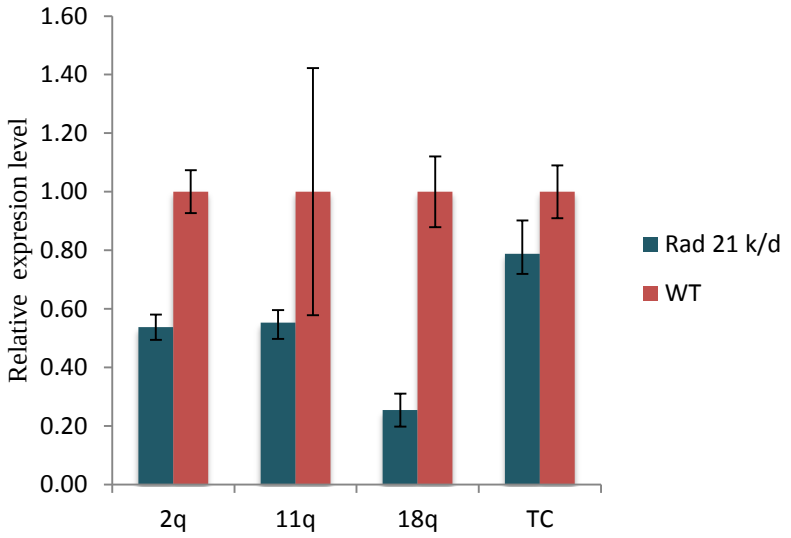


Figure 4.28. Comparison of *Terra* transcription at the given chromosomal ends. (A) *Terra* transcription is quantified by qPCR at the given chromosomal ends in *CTCF* k/d cell-lines and WT (wild type J1 cells). (B) *Terra* transcription is quantified by qPCR at the given chromosomal ends in mouse *Rad21 k/d* cell-lines and J1 cells. Expression levels of genes have been normalized to beta-actin mRNA. The error bars show standard deviation of mean.

Since we were able to observe a pronounced effect of *CTCF* and *Rad21* k/d on 18q chromosomal end transcription, we decided to test their role in *Terra* transcription regulation under heat shock. Both cell lines were exposed to heat shock and RNA was extracted and reverse transcribed. After analysis of the qPCR results we found that on comparison of fold change in *Terra* transcript levels between heat shocked and non-heat shocked cells in *CTCF* k/d and WT cell-line, the regulatory effect of lower CTCF levels was not observed on 18q *Terra* transcription (Fig. 4.29A). Similarly when *Terra* transcription was studied from 18q telomere end post heat shock, an increase in *Terra* levels was observed in both WT and *Rad21* k/d cell-lines (Fig. 4.29B).

This suggests that *Terra* transcription from mouse chromosomal ends varies in sensitivity to different transcription regulators. In our study *CTCF* and *Rad21* had an effect on *Terra* transcription from 18q telomeric end. Under cellular stresses like heat shock, there might be other heat response genes which regulate *Terra* transcription. Our data also suggests that regulation of mouse *Terra* transcription may involve factors that are different from the ones in human *TERRA* transcription.

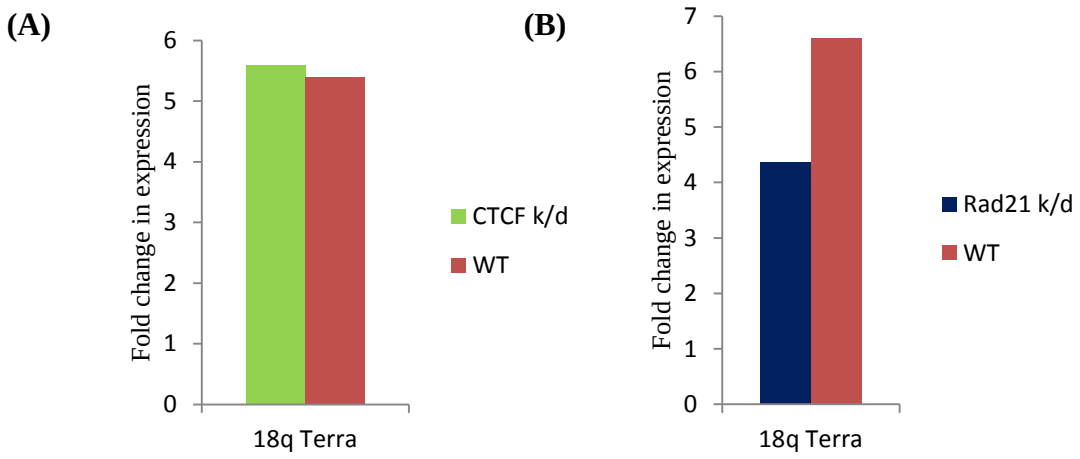


Figure 4.29. Comparison of fold change in *Terra* transcription from mouse 18q chromosomal end upon heat shock. (A) Fold change in *Terra* transcription quantified by qPCR from 18q telomeric end in *CTCF* k/d cell-line and wild type (J1 cells) after heat shock. (B) Fold change in *Terra* transcription quantified by qPCR from 18q telomeric end in *Rad21* k/d cell-line and wild type (J1 cells) after heat shock.

Next we performed RNA FISH on the two cell-lines post heat shock to examine global *Terra* expression *in situ* (Fig. 4.30). In both the cases we saw upregulation of *Terra* transcription upon heat shock and expression pattern was similar to that of the WT (J1) cells on heat shock. This could be a result of difference in sensitivity of the telomeres to CTCF and Rad21. So transcription from a handful of telomeres may have been affected in the knock-down, thereby not leading to any overall change in expression. Secondly, as I mentioned above, upon heat shock the factors controlling *Terra* transcription could be different. Hence, *Terra* transcription levels showed an increase even with CTCF and Rad21 being present at low levels.

In this section we looked at the effect of heat shock on *Terra*. Next, we investigated the role of certain factors in *Terra* transcription under normal and stressful conditions. Knockdown of transcription factors *CTCF* and *Rad21* affected *Terra* expression in cells, but the effect was not seen to be present uniformly at all the telomeric ends. Among the telomeres studied, *Terra* transcription from 18q was most susceptible to *CTCF* and *Rad21*. But, sensitivity to *CTCF* and *Rad21* was lost on heat shock. We also tested the role of transcription factor NFκB in *Terra* transcription as NFκB is the master regulator of stress response in cells. *Terra* upregulation in response to heat shock was seen to be independent of NFκB inhibition suggesting the

role of other factors controlling *Terra* transcription. In the future we intend to find other transcription factors involved in *Terra* transcription.

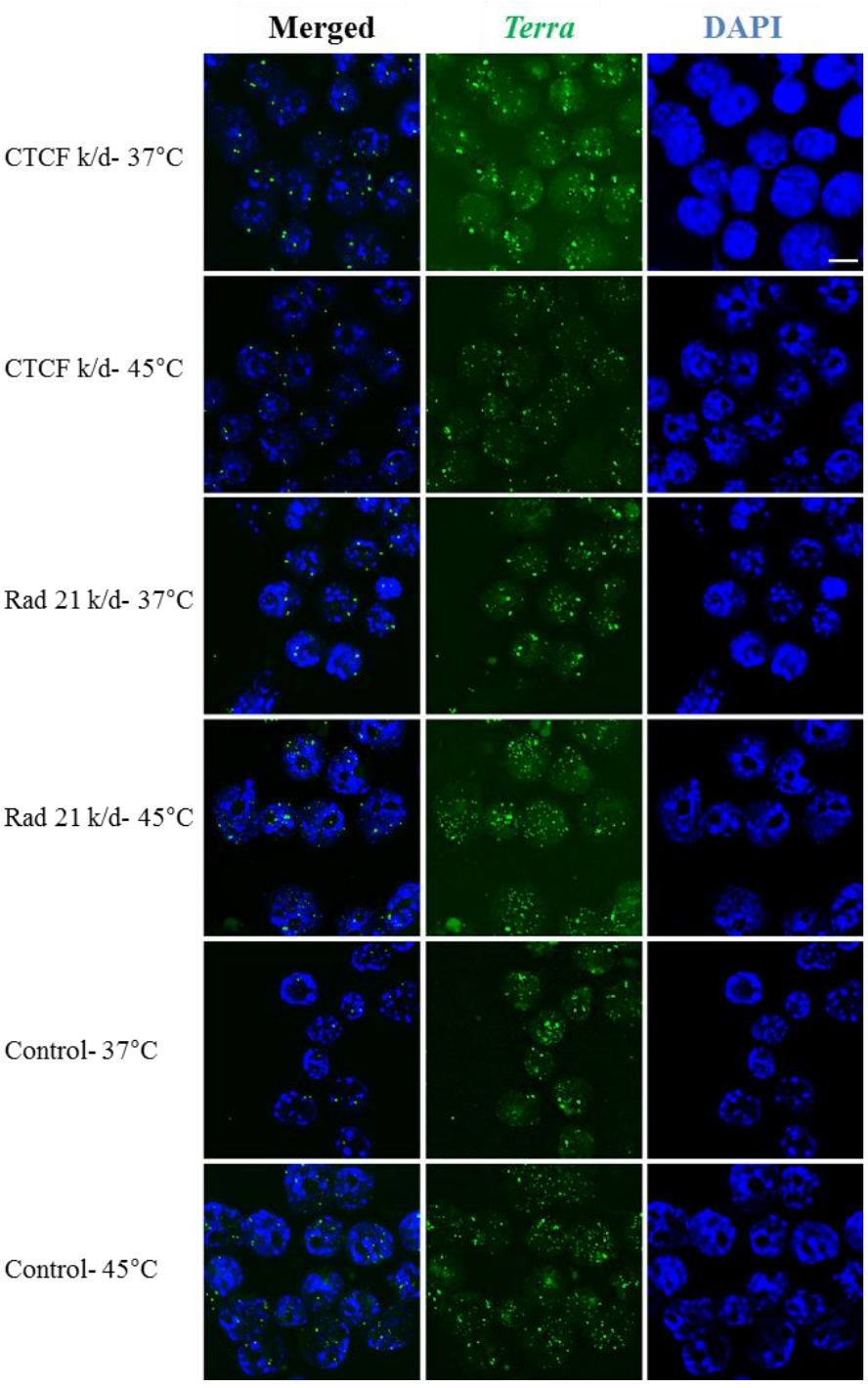


Figure 4.30. *Terra* RNA FISH performed on *CTCF* k/d cell line and *Rad21* k/d cell line in control (37°C) and heat-shock (45°C) conditions. The nucleus is stained in blue by DAPI. RNA detection probe is FITC tagged 42nt probe with a sequence complimentary to *Terra* RNA. Scale bar, 10 µm.

4.4. Manipulating *Terra* expression

Previously we have looked at methods to detect and study *Terra*, tried to understand its role in telomere maintenance and study its transcriptional regulation. To further our study on *Terra* biology, in this section we aim to directly manipulate *Terra* expression in mouse cells by using different techniques to knockdown or overexpress it. By studying the phenotype of a mutant cell-line with modified *Terra* expression we can gather information about the biological importance of *Terra.*

4.4.1. Knockdown of *Terra* expression

As the the regulatory elements in *Terra* transcription in mouse are still unclear, we could not modify it at genomic level. We chose lentiviral vector mediated short hairpin RNA (shRNA) transduction system to knockdown *Terra* transcripts in the cell. shRNA transduction provides a stable expression system in contrast to the conventional method of transient siRNA transfection, allowing us study the effects of *Terra* knockdown over a longer period of time.

4.4.1.1. Generation of *Terra* knockdown in male mouse ES cell-line:

For generating a knockdown cell-line we ordered the lentiviral particles from MISSION RNAi (Sigma Aldrich). The lentiviral particles contain pLKO.1-puro transfer vector comprising target shRNA insert (Fig. 4.31). Three sets of inserts were designed having 3 repeat units of the telomeric repeat TTAGGG in 3 reading frames (Table 4.3).

Table 4.3. Sequences of the shRNA inserts targeting *Terra*.

pLKO.1 puro-Zhang1 (Z1)	TTAGGGTTAGGGTTAGGGTTA
pLKO.1 puro-Zhang2 (Z2)	GTTAGGGTTAGGGTTAGGGTT
pLKO.1 puro-Zhang3 (Z3)	GGTTAGGGTTAGGGTTAGGGT

Lentiviral transduction and Puromycin selection was carried out according to the protocol mentioned earlier. The schematic of the vector backbone is shown in Figure 4.31.

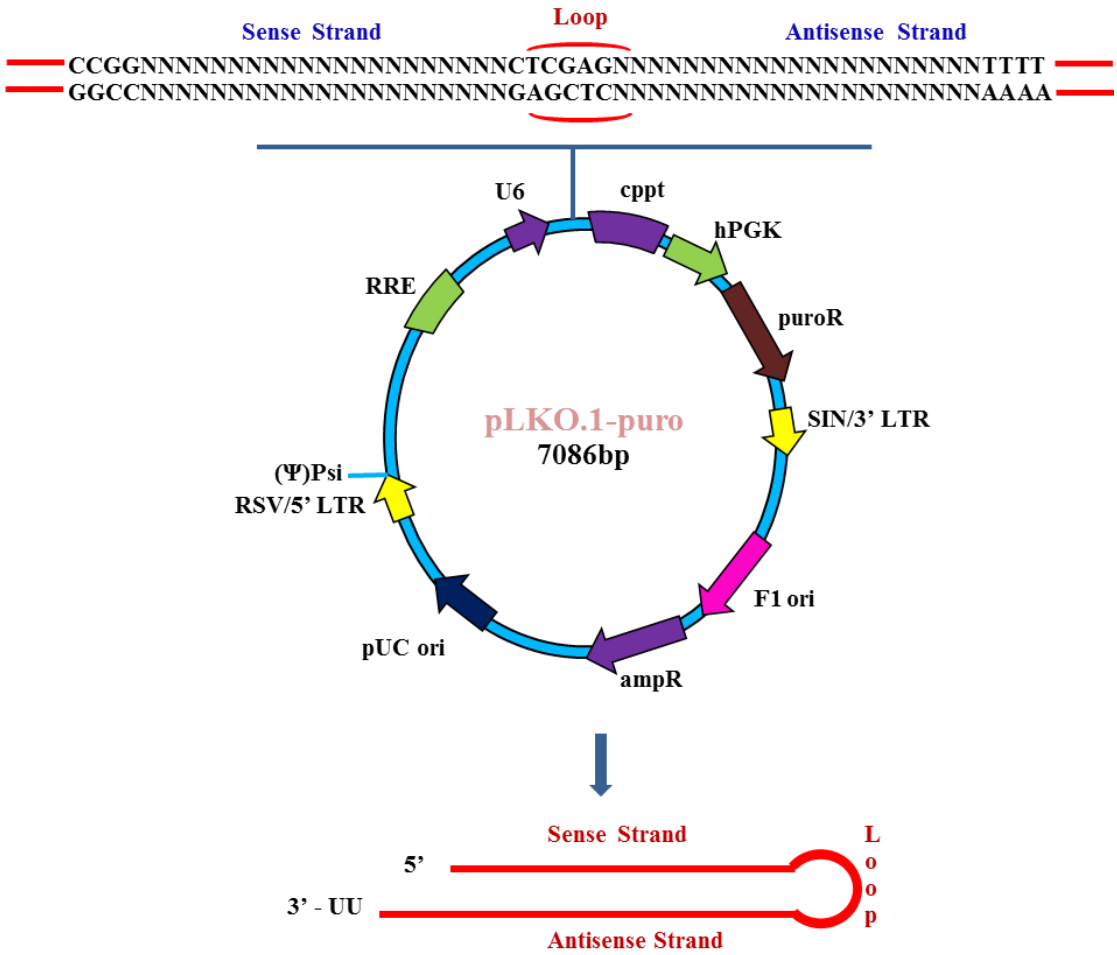


Figure 4.31. Lentiviral transfer vector design containing the insert for the target gene. The different components and their orientations are shown on the vector map.

The different components of the pLKO.1 puro vector shown in Figure 4.31 are explained in brief in Table 4.4. shRNAs are capable of DNA integration and consist two 19-22 bp complementary RNA sequences separated by a 4-11 nt loop (Fig. 4.31). Transcription of the 21mer shRNA construct was under U6 (human) promoter regulation. The transcribed short hairpin RNA is exported to the cytoplasm where it is processed further by Dicer and loaded onto the RISC complex to target the RNA of interest.

Table 4.4. Important genetic elements of the lentiviral transfer vector design.

Element	Function
RSV/5' LTR	Chimeric Long Terminal Repeat (LTR) involved in replication and integration.
Psi	Packaging dimerization signal
RRE	Involved in nuclear export of viral genome transcript during packaging
U6	Robust Pol III promoter with precise initiation and termination of shRNA trancription at high levels
cPPT	Role in cDNA synthesis and nuclear import in target cell
hPGK	human Phosphoglycerate kinase promoter
puroR	Allows stable selection with puromycin
SIN/3' LTR	LTR involved in replication and integration.

4.4.1.2. *Terra* knockdown was detected by Northern blot

Post transductions, 5 colonies per construct were selected based on their resistance to Puromycin. They were named as Z1a, Z1b, Z1c, Z1d and Z1e for the Z1 construct and similarly for the Z2 and Z3 construct. The colonies were expanded and total RNA was isolated from them using TRIzol, as described in methods. The total RNA from the samples were analyzed by small RNA Northern (15% PAGE) followed by autoradiography using γP32-labeled C-rich telomeric probe, G-rich telomeric probe and miR292 probe.

Terra ranges from 100 bp to 9 kb in length and is detected as a smear signal in Northern blot (Azzalin et al., 2007; Schoeftner and Blasco, 2008; Zhang et al., 2009). Its sequence is G (Guanine)-rich, hence, the C (Cytosine)-rich probe targets *Terra* RNA isolated from the cells. In all the Northern experiments, J1 was the control cell-line and *Terra* expression in the transduced cell-lines was compared to the expression in untransduced J1 cells.

In Figure 4.32, the Northern blot results are summarized. The three panels contain RNA samples from cell lines selected after transduction with Z1, Z2 and Z3 shRNA, which showed a lowered expression of *Terra* in comparison to the other clones that were isolated (details presented in Fig. S1, Appendix).

miR292, is a miRNA expressed constitutively in ES cells and functions as our positive control and loading control for Northern blot. In Figure 4.32C, miR292 bands can be visualized as a doublet around 23 mer in length. A faint signal for the miR292 precursor (~82 mer) can also be observed. Since the signal with miR292 probe is approximately of the same intensity in all the lanes, we can conclude that an equal amount of RNA was loaded in each lane. In Figure 4.32B, with the G-rich probe we detected the signal for shRNA transcript (highlighted with a red circle). The shRNA targeting *Terra* has a C-rich sequence; therefore the G-rich probe targets the shRNA transcript. A difference was observed in the detection of the shRNA transcript among the four samples, with Z2a having a very faint signal. This is possibly due to the recruitment of most of the shRNA of the cell-line Z2a in targeting *Terra*. Figure 4.32A has the results for hybridization of the Northern blot membrane with the C-rich probe. We observed that among the cell lines selected, **Z2a** showed a consistent downregulation of *Terra* expression in repeated rounds of Northern hybridization. Thus, from our Northern blot experiments, Z2a was our *Terra* knockdown cell-line. A faint signal originating from the shRNA precursor was also seen with the C-rich probe (Fig. 4.32, red arrow).

Table 4.5. Probes designed for Northern hybridization

Probe	Sequence
C-rich Probe	TAACCCTAACCCTAACCCTAA
G-rich Probe	TTAGGGTTAGGGTTAGGGTTA
miR292-as Probe	ACACTCAAAACCTGGCGGCACTT

When the samples were probed with the G-rich probe (Fig. 4.32B), a faint smear-pattern similar to the C-rich probe is observed for all the samples. The intensity

of the signal varied among different samples. This smear could be indicative of the fact that an antisense transcript of *Terra* may also be transcribed. As discussed earlier previous literature suggests that *Terra* transcription is unidirectional, in an outward orientation in mammals (Azzalin et al., 2007; Schoeftner and Blasco, 2008; Zhang et al., 2009). The Northern data from all these studies consistently showed a strong signal with the C-rich probe and a weak, but, detectable signal with the G-rich probe, from which it was concluded that the transcription is taking place in one direction. However, in light of recent studies documenting the expression of an antisense of *TERRA*, known as *ARRET*, in some organisms like *Arabidopsis* (Vrbsky et al., 2010) and *S. cerevisiae* (Luke et al., 2008), and our consistent observation of a RNA smear with the G-rich probe, we started to speculate that antisense of *Terra* may exist in mammals also. We attempted to test the presence of *Terra* antisense with RNase protection assay but were not successful. Further studies are required to satisfactorily prove the presence of an antisense *Terra* transcript in mouse ES cells and the role of this transcript in the organism.

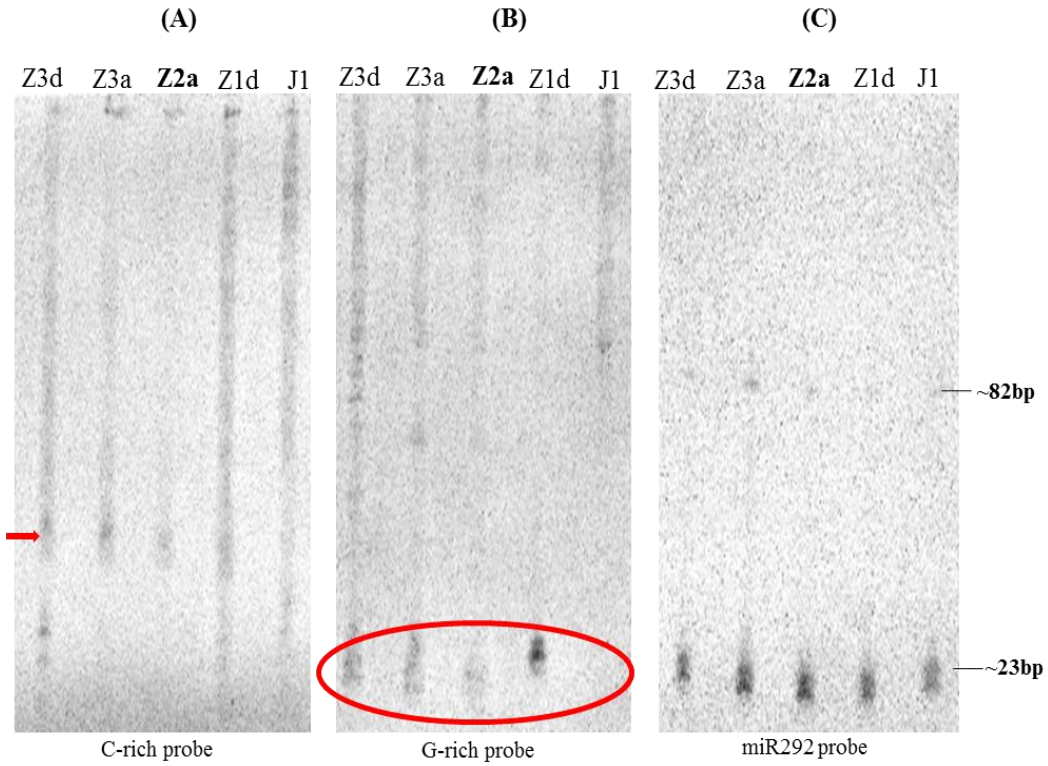


Figure 4.32. Northern blot to study the knock-down of *Terra* expression upon lentiviral transduction. The probes used are mentioned below each panel. The red arrow in (A) marks the faint shRNA precursor signal. The highlighted part in (B) represents the shRNA signal.

4.4.1.3. To validate the effect of *Terra* knockdown by RNA FISH

In order to verify knockdown of *Terra* expression in Z2a cell line, we performed *Terra* RNA FISH. We expected a decrease in *Terra* signal upon RNA FISH. However, contrary to this view, RNA FISH couldn't detect any visible change in *Terra* signal intensity in Z2a cells when compared to wild type J1 cells (Fig. 4.33).

Two possible explanations can be provided for this contradictory observation. One reason could be that *Terra* is partially knocked down, such that only the lower molecular weight *Terra* fraction detected by Northern blot is targeted. This could explain the absence of *Terra* smear in the Northern blot but the presence of signal in the RNA FISH. Another reason could be that the shRNA system is effective in the cytoplasm. Hence, we could have knocked down cytoplasmic fraction of *Terra*. Although, previous work has shown that *Terra* is localized in the nucleus but, from the observations in this section and the next few sections, we speculate that a fraction of *Terra* could be present in the cytosol.

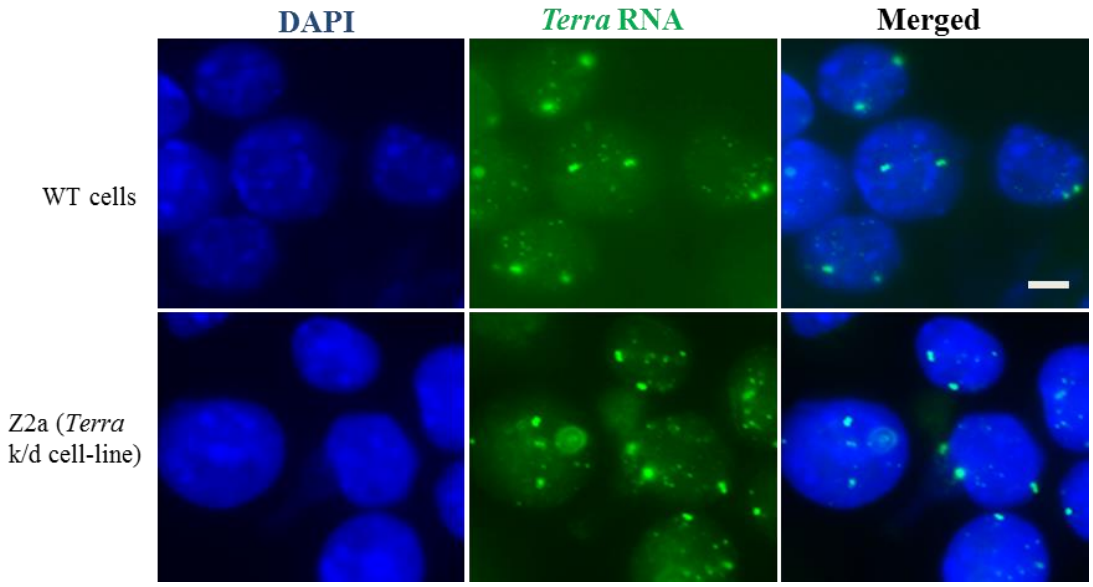


Figure 4.33. RNA FISH for *Terra* on wild type cells (J1 cells) and Z2a (transduced cell-line). The C-rich *Terra* probe was tagged with FITC. Scale bar, 5 µm.

We also tested the effect of this partial knockdown of *Terra* on proliferative ability and phenotype of the cell. Culturing Z2a cell-line over an extended period did

not have any effect on growth rate or colony morphology. This could be a result of the partial knockdown of *Terra* generated here not being able affect the general *Terra* functions or absence of any role played by *Terra* in cellular proliferation and phenotype maintenance.

4.4.2. Overexpression of *Terra*

In this section we attempted to overexpress *Terra* in cells. We utilized CRISPR technique for this purpose. CRISPR (Clustered, regularly interspaced, short palindromic repeats) systems of bacteria have enabled the development of RNA-guided technology for targeted genome editing (Mali et al., 2013b), and for repression and activation of genes (Maeder et al., 2013). CRISPR associated 9 nuclease (Cas9) from *Streptococcus pyogenes* can be directed by a ~100 nt guide RNA (gRNA) to target a genomic DNA sequence that is complementary to the first 20 nt of the gRNA and is followed by a protospacer adjacent motif (PAM) sequence of the form -NGG where N is any nucleotide (Jinek et al., 2012; Mali et al., 2013a). RNA guided recruitment of catalytically inactive form of Cas9 (Cas9m4) can induce repression of endogenous genes in bacteria and human cells (Gilbert et al., 2013). In addition, fusion of Cas9m4 to effector domains have been used to activate reporter genes in *E. coli* and human cells (Gilbert et al., 2013) (Fig. 4.34). Here we tried to determine whether a catalytically inactive Cas9 construct fused to synthetic VP64 activation domain (Mali et al., 2013b) at its C-terminus (Cas9m4-VP64), could enhance the transcription of *Terra* in the presence of a guide RNA binding to the telomere region of mouse chromosomes.

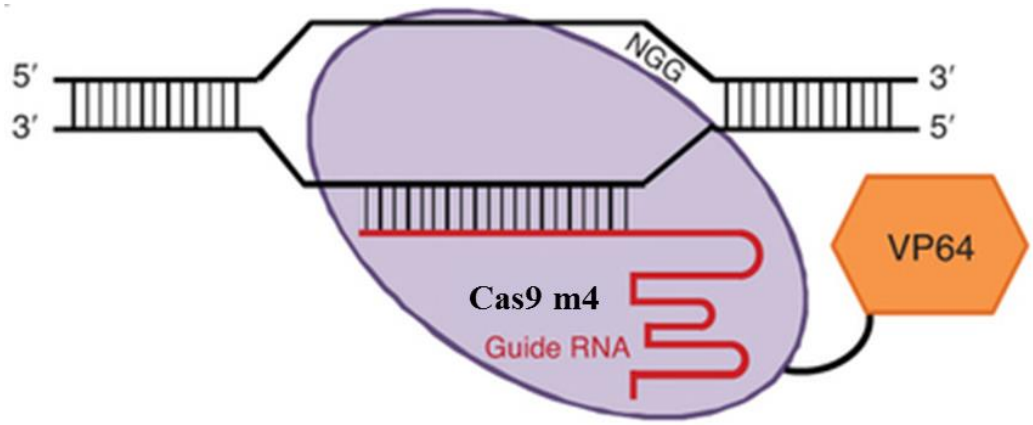


Figure 4.34. Schematic depicting recruitment of Cas9m4-VP64 to a genomic target by a guide RNA (gRNA). (Maeder et al., 2013)

For targeting *Terra* transcription, we introduced 19 nt of the G-rich telomere repeats ending with the -NGG PAM sequence (**5'-**AGGGTTAGGGTTAGGGTTA*GGG*-**3')** into the gRNA vector backbone Our objective was to target the catalytically inactive Cas9m4 coupled with VP64 transcription activator (Cas9m4-VP64) to the telomere region by the gRNA containing insert sequence recognizing telomere (*Terra* gRNA). *Terra* promoter and start site have not been characterized in mouse. So we targeted our CRISPR gene activator system to the telomeric repeats instead of a promoter region for over expression of *Terra.*

We transfected J1 cells with catalytically inactive Cas9 plasmids and gRNA. In Figure 4.35, we report the results of *Terra* RNA FISH post transfection. First row represents wild type J1 cells transfected with blank gRNA. Second row represents cells that were transfected with blank gRNA and Cas9m4-VP64. As per our expectations, we did not see an increase in *Terra* expression upon transient transfection as blank gRNA does not carry the insert for guiding Cas9m4-VP64 to the telomere. This showed that Cas9m4-VP64 could not single handedly upregulate *Terra* transcription. In the third panel we transfected catalytically inactive Cas9 without transactivator (Cas9m4) with *Terra*-gRNA for guiding Cas9m4 to telomeres. Since the transactivator domain is absent in Cas9m4 plasmid we did not expect any upregulation of *Terra* transcription. Our result was according to our expectation. This showed that

the Cas9m4 plasmid was unable to upregulate *Terra* in the absence of VP64 transactivator.

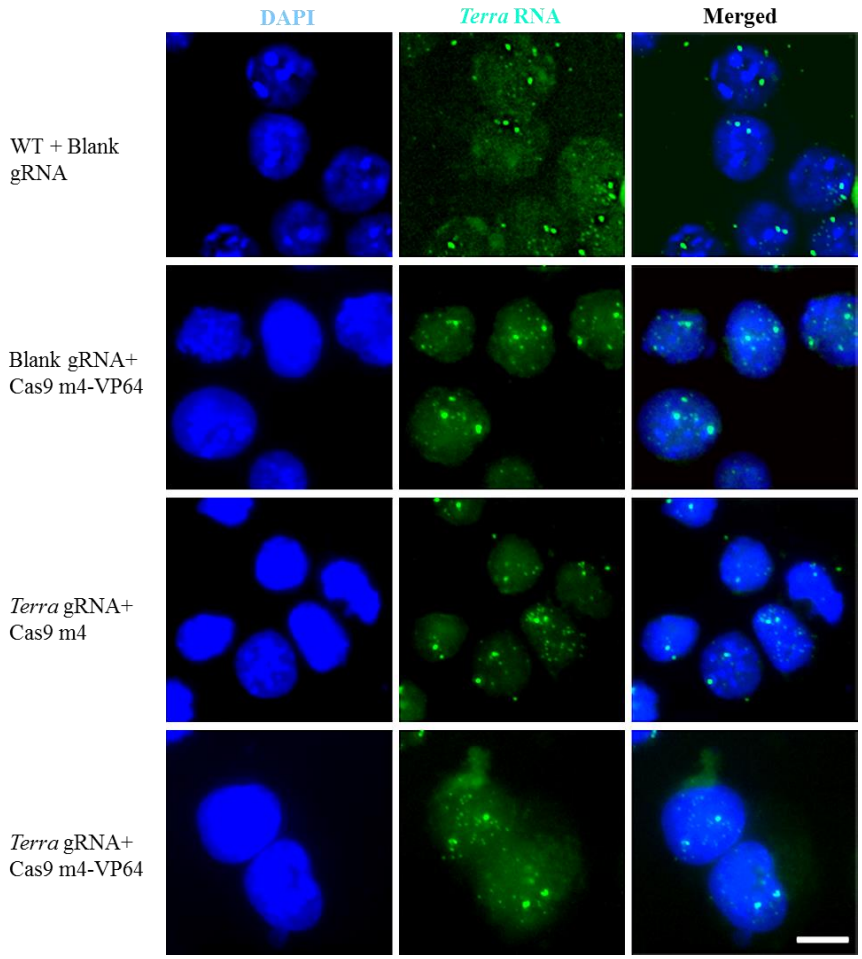


Figure 4.35. *Terra* RNA FISH on wild-type mouse ES cells transfected with modified Cas9 plasmids. The transfected plasmids are listed on the left of each panel. FISH was performed using FITC labeled C-rich probe. Three sets of plasmids are: Cas9m4 (catalytically inactive Cas9), Casm4-VP64 (catalytically inactive Cas9 with VP64 transactivator), *Terra* gRNA (guide RNA with insert) and blank gRNA (guide RNA without insert). Scale bar 5 µm.

Fourth panel displays the results from cells transfected with Cas9m4-VP64 and guide RNA containing insert directing the CRISPR system to the telomeres (*Terra* gRNA). We did not observe a change in *Terra* expression here. This could be due to the inability of this modified CRISPR-Cas9 system to initiate transcription from areas other than defined promoter sites. The transactivator VP64 may not be able to induce transcription in the absence of RNA Polymerase binding. As mentioned above we have very limited knowledge of *Terra* transcription promoters making it difficult to

target the promoter region. Also it is not certain if transcription of *Terra* in mouse always originates from subtelomers or it can originate spontaneously from the telomeric region. With better knowledge about the transcription start site and the subtelomeres of mouse chromosome, we could use this system in our future work to conduct studies on *Terra.*

4.5. Study of *Terra* in the malaria parasite *P. falciparum*

Malaria in humans is caused by protozoan pathogen *Plasmodium falciparum*. Till date no vaccines have been discovered against the disease due to the antigenic variation of the pathogen. This variability is caused by a set of Variable Surface Antigenic (VSA) genes located in the subtelomeric region (Scherf et al., 2008a). Our aim here is to study if the telomeres of malaria are transcribed into *Terra* and if so, the effects of *Terra* on the antigenicity determining genes of the organism could be investigated in future studies.

4.5.1. Detection of *Terra* in *Plasmodium falciparum* by Northern blot

To confirm the presence of *Terra* in the malaria parasite, we first attempted Northern blot on *P. falciparum* samples. RNA was isolated from Dd2 strain of *P. falciparum* from trophozoite and schizont stage parasites. The RNA samples were run on a 15% PAGE gel followed by autoradiography using γ^{32}P-ATP labeled probes.

Table 4.6. Probes used for Northern blot.

Probe	Sequence
Pf C-rich	AACCCTAAACCCTAAACCCTA
Mouse C-rich	TAACCCTAACCCTAACCCTAA
miR292	ACACTCAAAACCTGGCGGCACTT

As shown in Figure 4.36C, the Northern blot membrane was probed for miR292. miR292 is a miRNA expressed in mouse ES cells. In our Northern blot experiments, we selected miR292 as the control to validate the experimental quality. miR292 probe detected a "doublet" signal with a size of ~23 nt, and a faint signal with a size of ~82 nt. These are the expected signals from the mature microRNA and its precursor.

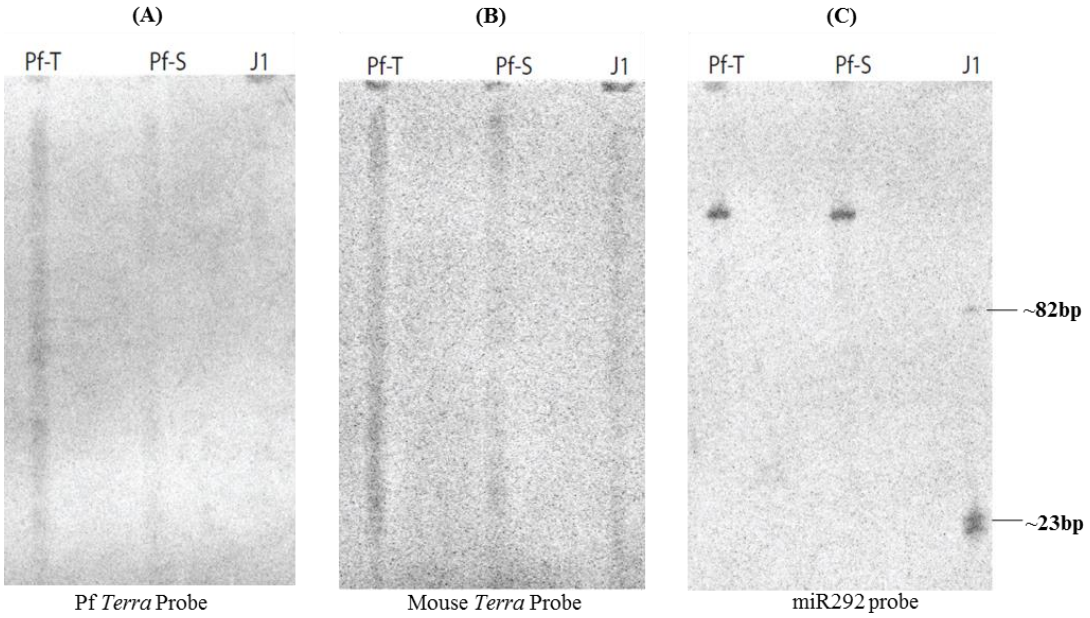


Figure 4.36. Northern blot to study the presence of *Terra* in *Plasmodium falciparum*. Pf-T= *Plasmodium falciparum* trophozoite stage; Pf-S= *Plasmodium falciparum* schizont stage; J1= wild type mouse ES cell-line.

From *Plasmodium falciparum* (Pf) RNA samples, we detected a smear RNA signal (Fig. 4.36A). The smear signal detected by Northern blot is similar to the smear signals observed for the mouse *Terra* (Fig. 4.36B). As already mentioned, studies have shown that *Terra* has no fixed length and ranges from 100bp to 9kb (Azzalin et al., 2007; Zhang et al., 2009), hence, the smear observed from the parasite sample could be indicative of the fact that *Plasmodium falciparum* telomeres are transcribed into *Terra*. The RNA used for the Northern are from two different stages of the parasite. The difference of *Terra* expression level between the Pf-T and Pf-S stages of the parasite (Fig. 4.36A) could be the result of a possible stage dependent expression of *Terra* in the parasite.

4.5.1.1. Presence of *TERRA* in human blood sample

We made two side observations in our Northern blot experiment (Fig.4.36). First, we detected a "small" RNA signal from the parasite RNA samples when the membrane was probed with miR292 probe (Fig. 4.36C). Ideally we should not detect

any signals in the parasite RNA samples as they lack an RNAi system. The size of the parasite RNA signal is around 100 nt. Since the parasites were cultured in human RBC, one possibility is the detected signal could come from human blood. Preliminary bioinformatics analysis showed that the miR292 probe should not detect any signals from the RNA sample of human RBC (data not shown). Therefore, mir292 probe detected a clear small RNA signal from the malaria parasite RNA sample. This observation is interesting; however, it is beyond the focus of this study.

The second side observation and a more interesting one, is that a smear signal was detected from the parasite RNA by the mouse telomeric probe (Fig. 4.36B). This result was a cause of concern as the mouse telomeric probe should not hybridize with the *Plasmodium* telomeric RNA due to the difference in sequence of the telomeric repeat.

We hypothesized that RNAs from human red blood cells (*Plasmodium* culture/growth media) could be present in the isolated RNA sample. Since the telomeric sequence for humans is same as the mouse, human *TERRA* signal could be detected with mouse telomeric RNA probe. To confirm the presence of human RNA in parasite sample, we used a 21mer probe to detect miR16, a human specific miRNA expressed in blood cells (Calin et al., 2002). This probe detected a strong signal of miR16 in the parasite RNA sample (Fig. 4.37). This result clearly confirmed our hypothesis that contaminating human RNA is present in the parasite RNA extract. Hence, the smear signal detected is human *TERRA*. An interesting notion can be further deduced from this result. A red blood cell does not contain a nucleus. Therefore, human *TERRA* signal that we detected in this experiment is likely to be *TERRA* RNA from the cytoplasm of human RBCs. Previous work on *Terra* states that, *Terra* is localized in the nucleus. However, in the light of our present results we would like to reconsider the cellular distribution pattern of *Terra* and perform further studies to prove the presence of *Terra* in cytoplasm of cells. We elaborate more on this topic in the next section.

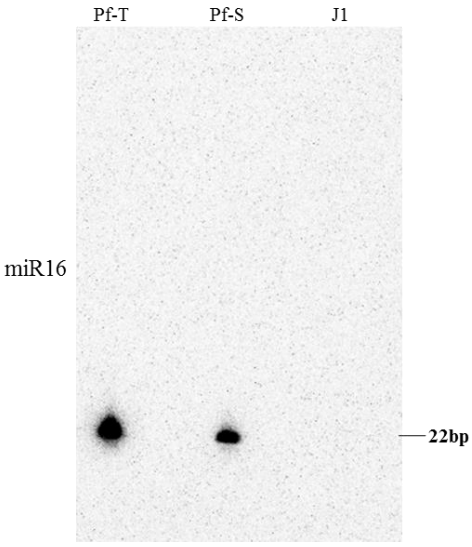


Figure 4.37. Northern blot results for the miR16 probe. Pf-T= *Plasmodium falciparum* trophozoite stage; Pf-S= *Plasmodium falciparum* schizont stage; J1=wild type mouse ES cells.

4.5.2. Detection of *Terra* in malaria parasite by RNA FISH

As we could detect the expression of *Terra* in *Plasmodium falciparum* by Northern blot, we decided to further confirm its *in-situ* expression level using RNA FISH. Besides being a technique to validate *Terra* expression in the parasite, RNA FISH experiment can also reveal the sub-cellular localization of *Terra* in the parasite cells.

For RNA FISH, we initially designed a 200 bp long oligo containing the telomeric repeats with adaptor sequences at both ends for PCR amplification. Probe was synthesized by PCR amplification of the oligo in the presence of fluorescently labeled nucleotide, Cy3-dUTP. However, the probe did not give satisfactory results. This could be due to inability of the probe to bind strongly with the target RNA.

Therefore, we ordered a peptide nucleic acid (PNA) probe for the experiment. PNA probes have extraordinary thermal stability and are resistant to nucleases. The PNA probe was FITC tagged on the 5' end (Sequence given in Section 2.28).

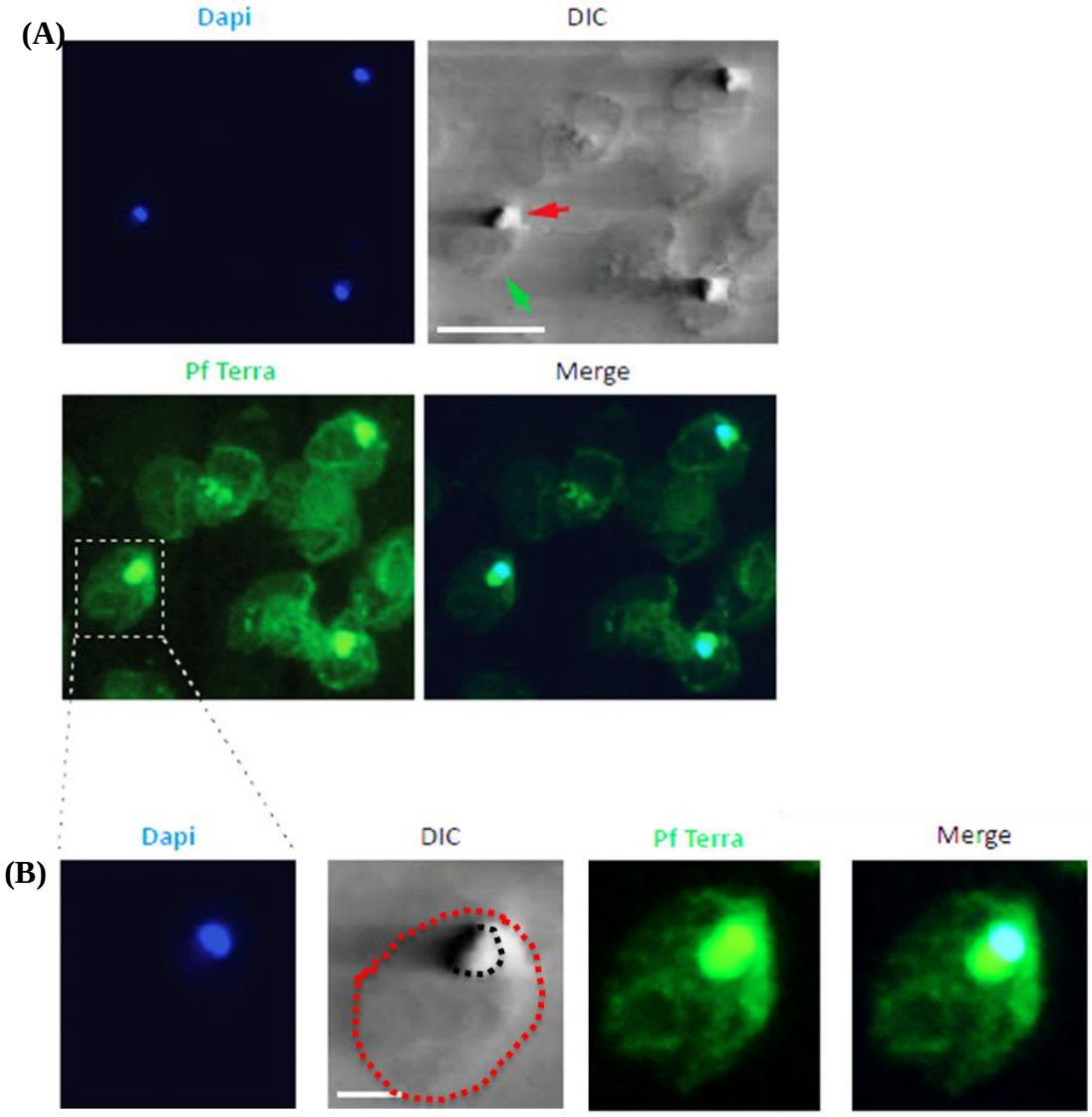


Figure 4.38. *Terra* RNA FISH on trophozoite phase cells. (A) Image depicts malaria parasites inside the RBCs. The parasite nucleus was stained in DAPI (blue). The RBCs and parasites can be observed in the DIC channel. The red arrow points to a malaria parasite. The green arrow points to a RBC. The Pf *Terra* probe was labeled by FITC (green). Scale bar, 10 µm. (B) An image of a single RBC infected by one parasite indicated by a square with white dashed lines in (A) is magnified for a detailed view. The red dotted line demarcates the RBC boundary and the black dotted line marks the parasite inside the RBC. Scale bar, 2.5 µm.

Figure 4.38 displays the results of *Terra* RNA FISH on *P. falciparum* trophozoite stage cells. The RBCs (nucleus-free) are not fluorescently labeled in the experiment. A DIC (differential interference contrast) channel was added to observe the unstained RBCs. DIC produces contrast by visually displaying the refractive index gradients of different areas of a specimen. The sensitivity of DIC depends on the

sample thickness and how different its refractive index is from the slide background. The cytoplasm of RBCs were extracted to allow the permeabilization of the RNA FISH probe into the parasites, therefore the RBCs were flattened during this process and suffered significant cytoplasm content loss. Despite this, the RBCs can still be visualized in the DIC channel (Fig. 4.38A, the green arrow). The malaria parasites, probably due to its thickness and different refractive index from its surroundings, can be very clearly seen in the DIC channel (Fig. 4.38A, the red arrow). We detected strong fluorescent signals in the malaria parasite and significant amount of *Terra* signals were detected outside the "DAPI-defined" nucleus of the parasite.

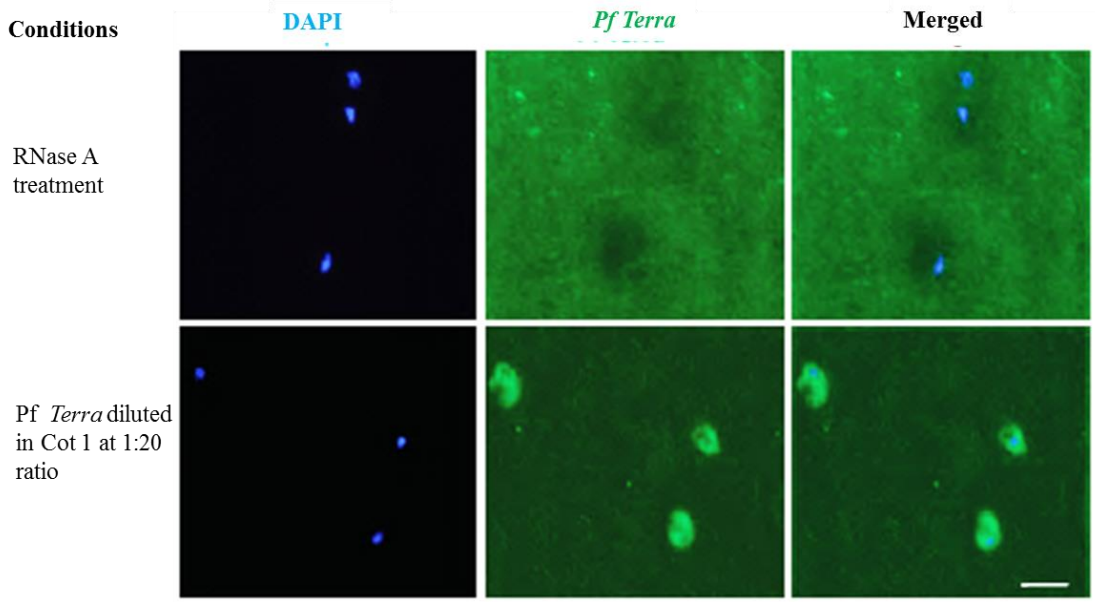


Figure 4.39. *Terra* RNA FISH in *Plasmodium falciparum* trophozoites. Top panel shows Pf *Terra* RNA FISH on RNaseA treated slides. Bottom panel shows Pf *Terra* RNA FISH wherein Pf *Terra* PNA probe was diluted with unlabeled Cot 1 at 1:20 ratio. Scale bar, 5 μm.

To further confirm the signals detected are RNA FISH signals, we treated the slides with RNaseA (Fig. 4.39). The treatment completely abolished the previously detected signal confirming the RNA nature of the signal. This result confirmed that the signals detected in RNA FISH are RNA signals. To further confirm that the RNA signals were detected by the PNA probe with high sequence specificity, we tried combining the PNA probe with high concentration of unlabeled Cot1 DNA (1:20 ratio of the PNA probe to mouse Cot-1 DNA) (Fig. 4.39). Cot-1 DNA is repetitive DNAs isolated from the mouse genome. Unlabeled Cot-1 DNA acts as the blocking agent to

prevent non-specific hybridization. PNA probe combined with high concentration of unlabeled Cot-1 was still able to detect clear signals. This indicates that the detected RNA signal is sequence-specific to the PNA probe. In addition, the hybridization temperature of RNA FISH experiment was raised from 37°C (Fig. 4.38) to 42°C (Fig.4.39). At 42°C, the green background noise from RBCs, recognizable in Figure 4.38, disappeared, while the bright RNA signals from the parasite were still clearly visible. These results confirmed the sequence-specific detection of *Terra* with the PNA probe. Taken together, the RNA FISH experiments show that *Terra* is present in the malaria parasite and *Terra* RNA is highly enriched in the cytoplasm of the parasite.

The results here, coupled with our observations in the previous experiments, which point to the presence of *Terra* in RBC, a cell type devoid of nucleus, further suggests that *Terra* RNA could be distributed in the cytoplasm in different cell types. However, further experiments need to be performed involving cellular RNA fractionation assays for confirming this hypothesis.

The detection of *Terra*, typically known as a nuclear RNA, outside of the nucleus of the parasite is also in-line with the recent observation on histone modifications of the malaria parasite (Luah et al., 2010). It was observed that many histone modifications showed perinuclear localization with a significant fraction being detected outside the DAPI-defined nucleus. One dramatic example is H3K9me1, which showed clear localization in the cytoplasm of the parasite.

4.5.3. Detection of *Terra* in malaria parasite by RNase protection assay

To further confirm transcription of *Terra* in malaria parasite we decided to detect *Terra* by Ribonuclease protection assay (RPA). RPA provides a sensitive method for detecting and quantifying RNA. To generate the probe used in RPA, we PCR amplified template oligo Pf_Telo-C using primer set Pf_Telo-C_Forward and Reverse mentioned in Table 2.2. Following PCR amplification it was cloned into the

vector backbone pGEM-3Zf(+) at the EcoRI and HindIII restriction sites. The vector backbone has a T7 and a SP6 promoter site (Fig. 4.40A). The established plasmid construct was named as "pGEM_Pf_Telo". Its structure is shown in Figure 4.40B. The plasmid was used for *in vitro* transcription reaction to generate RNA probes used in RPA experiment. Similar to our previously designed RPA experiment, *in vitro* transcription produces a 188 nt RNA transcript, which carries 153 nt sequence complementary to Pf-*Terra*. Therefore, a successful RNA protection assay should result in a 153 nt RNA signal for the Pf-*Terra* probes.

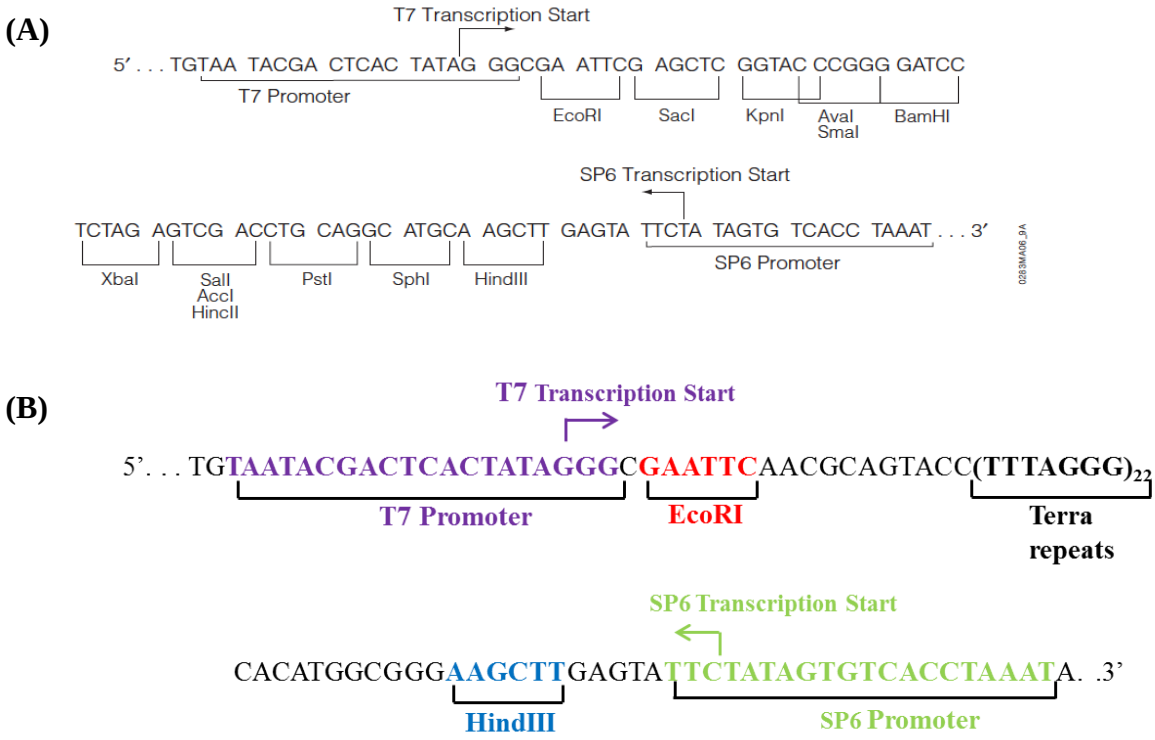


Figure 4.40. Generation of pGEM_Pf_Telo for RPA. (A) pGEM-3Zf (+) vector multiple cloning region and circle map. (Promega, Technical Bulletin No. 086) (B) pGEM_Pf_Telo vector map showing the region with *Terra* repeats inserted between HindIII and EcoRI sites under SP6 and T7 promoters.

We used mouse *β-actin* as a positive control for our RPA experiment. A successful RNA protection assay should result in a 245 nt RNA signal for the *β-actin* probe. As shown in Figure 4.41, the *β-actin* RNA probe generated by *in vitro* transcription showed as a sharp band on PAGE gel (Fig. 4.41, lane 1). Under RNase treatment, the probe was protected by mouse RNA sample and generated a smaller

RNA signal on the PAGE gel (Fig. 4.41, lane 5). The signal intensity of the protected RNA signal is weaker compared to the original RNA probe. This is because the radiolabeled probe was added in excess of the cellular *β-actin* levels. These results show that the RPA experiment is working.

As shown in Figure 4.41, lanes 2 and 3 contain unprotected C-rich and G-rich probes without RNase treatment. The C-rich probe is anti-sense to *Terra* and is used to detect parasite *Terra*, while the G-rich has the same sequence as *Plasmodium Terra* transcript and is used as the control probe in the experiment. Both probes showed up as sharp signals with the correct sizes (188 nt) on the PAGE gel. The unprotected C-rich probe was completely destroyed by the RNase treatment (Fig. 4.41, lane 4). With the protection of 20 µg of total RNA isolated from the parasite, a smaller RNA signal (~153 nt) was detected (Fig. 4.41, lane 6). The protected RNA signal is weak. This could be due to the low level of *Terra* presented in the parasite RNA sample. Supplementary figure 2 shows an image from a separate round of RPA with stronger *Terra* signal although *β-actin* probe did work in that round due to some technical error. When the C-rich probe and the G-rich probe were combined before RNase treatment, the two probes protected each other and resulted a RNA signal of ~188 nt. These results show that a *Terra* RNA signal was detected by RPA from the parasite RNA sample.

Another observation made from lane 6 was the ladder-like pattern shown by the blue arrow (Refer to Fig. S2 also). Malaria parasite telomere DNA contains degenerate G-rich tandem repeats, with **GGGTT(T/C)A** as the most frequently repeated sequence. Our designed probe is anti-sense to the $(GGGTTTA)_n$ sequence. It is possible that due to the presence of two different kinds of repeat sequences in the telomere DNA we observe the ladder-like pattern. RNase A present in the RNase cocktail used in RPA has previously been observed to cause double stranded cleavage if a nick is present in the binding region (Lenk and Wink, 1997). Since the ladder size is smaller than the band, it could be due to cleavage by RNase A at mismatch regions between the radiolabelled probe and telomere DNA.

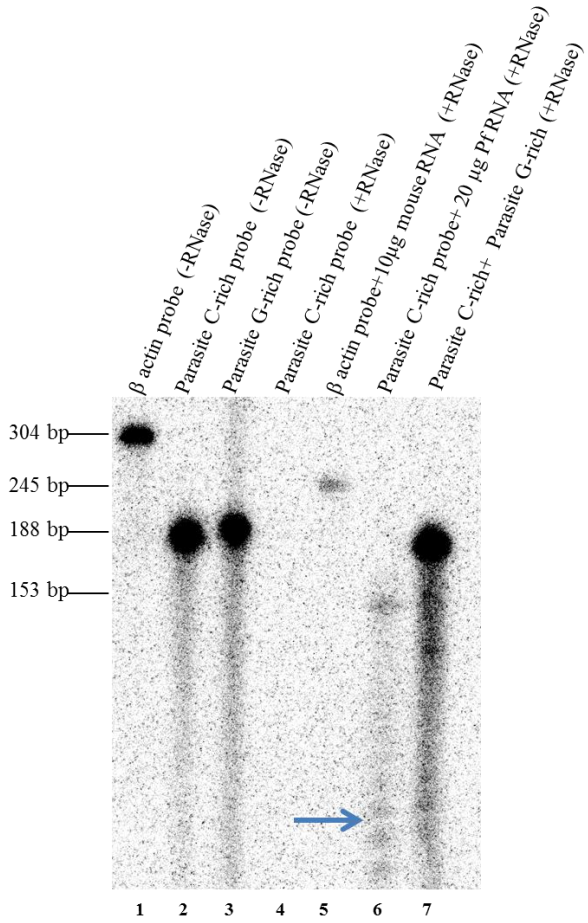


Figure 4.41. RNase Protection assay to test the presence of *Terra* in the malaria parasite *P. falciparum*. Blue arrow points to the ladder like pattern we observe in the RNase protection experiment in lane 6.

To test this hypothesis we designed another 200 nt probe called 'Pf-Chimeric'. This probe contains a mixture of GGGTTTA and GGGTTCA in anti-sense orientation (Table 4.7). The adaptor regions for primer binding are marked in red (Table 4.7). We PCR amplified this probe and inserted it in the pGEM3Zf plasmid in a similar fashion as mentioned above. Our next aim is to perform RPA with the *in vitro* transcription product of Pf-Chimeric with (a) malaria parasite RNA and (b) the G-rich probe *in vitro* transcribed from plasmid pGEM_Pf_Telo. If we observe a similar ladder-pattern in the

above RPA reactions, it would further validate our findings on *Terra* transcription in the malaria parasite.

Table 4.7. Sequence of Pf-Chimeric probe.

Probe	Sequence
Pf-Chimeric	GGTAACAATTTTCCCGCCATGTGaaccctgaac cctgaaccctgAACCCTAAACCCTAaaccctgaaccctg AACCCTAaaccctgaaccctgAACCCTAAACCCTAA ACCCTAaaccctgaaccctgaaccctgaaccctgAACCC TAAACCCTAAACCCTAaaccctgaaccctgGGTAC TGCGTTCCGTTTATTTCC

To summarize, we have tried to detect *Terra* in malaria parasite. We saw a smear on a Northern hybridization blot probed for malaria parasite *Terra* which is similar to *Terra* pattern seen from mouse and humans. By RNA FISH we showed that the nature of the signal detected by probes targeting malaria *Terra* is indeed RNA as the signal was destroyed upon RNase treatment. We found that *Terra* is present in the cytoplasm of the parasite by RNA-FISH. We further detected *Terra* in malaria parasite by RNase protection assay. These observations are in line with the recent discovery of lncRNA families in *Plasmodium* and the possibility of transcription from its telomere associated regions. Our study provides evidence regarding the presence of *Terra* in malaria parasite.

CHAPTER 5: CONCLUSIONS

5. CONCLUSIONS

Observations of *Terra:*

- Developed a simple and robust method for simultaneous RNA-DNA FISH and successfully performed RNA-DNA FISH at single molecule and single cell level.
- The average length of telomeres associated and not associated with *Terra* does not have a statistically significant difference suggesting *Terra* association may not always be for the purpose of inhibiting telomerase activity.
- *Terra* is not always associated with the telomere DNA hinting towards the possibility that *Terra* may have functional roles apart from telomerase regulation.
- Performed RNase protection assay (RPA), an extremely sensitive technique for RNA detection, to detect *Terra* in a mouse RNA sample.
- Previously we observed a faint smear signal when we probed our Northern blot for antisense *Terra* so we attempted to detect antisense *Terra* by RPA but were not successful.

Terra and telomere length:

- Comparing the length of *Terra*-associated distal end and non-associated centromeric end of X chromosome, we did not recognize any significant difference in their telomere length shortening in J1 cells on long term culture suggesting *Terra* association may not affect telomere length maintenance *in vivo*.
- In a telomerase knock out cell-line, telomere at the centromeric end of the X chromosome was significantly shorter than the telomere at the distal end and upon long term cell culture telomere shortening was observed at both ends.
- When the *Terc*-/- telomerase knock out cell-line underwent long term cell culture, chromosome fusion occurred when telomere ends reached their critically short length. The sex chromosomes are the earliest chromosomes

showing chromosome fusion during this process suggesting sex chromosomes are under high pressure in maintaining their telomere length.

Terra and stress response:

- *Terra* expression is upregulated in cells upon heat shock showing that *Terra* expression is regulated in response to certain stress conditions.
- *Terra* expression from different telomeric ends showed a difference in response to heat shock, with transcription from 18q showing maximum upregulation upon heat shock.
- Although *Terra* expression is significantly increased upon heat shock, the association pattern of *Terra* with telomeres was similar.
- Nuclear factor-kappaB (NF-κB), a master transcriptioin factor related to stress response, was not seen to be involved in regulating *Terra* expression change on heat shock.
- Two other transcription factors CTCF and Rad21 were studied and it was seen that they control transcription from 18q end among the ends we investigated.
- CTCF and Rad21 are not involved in 18q telomeric end's response to heat shock suggesting other transcription factors may be involved in stress regulated transcriptional changes.

Manipulating *Terra* expression:

- We used shRNA to knock down *Terra* and observed lowered expression in Northern blot but not in RNA FISH.
- We detected the presence of *Terra* in nucleus-free human RBC by Northern blot.
- We also detected the presence of *Terra* in the cytoplasm of *Plasmodium* parasites by RNA FISH.
- Taken together, these results suggest the presence of *Terra* in the cytoplasm. We believe the decreased Terra level shown in Northern blot from the shRNA knockdown sample, showed that we knocked down the cellular fraction of *Terra* by our shRNA system.

- We attempted to over express *Terra* using a modified CRISPR-Cas9 system but were unable to see a significant change in expression by RNA FISH.

Study of *Terra* in malaria parasite:

- We attempted to detect *Terra* in *Plasmodium* by Northern blot and observed a smear similar to previously observed *Terra* signal suggesting presence of *Terra* in *Plasmodium. Terra* expression levels vary in different developmental stages of the parasite.
- In our Northern blot we also detected presence of *Terra* in nucleus-free RBC suggesting that *Terra* may not be localized only in the nucleus.
- We detected *Terra* in *Plasmodium* by RNA FISH and observed that *Terra* signal is detected from both nucleus and cytoplasm in the parasite giving further proof to the idea regarding the presence of *Terra* outside nucleus.
- We further detected *Terra* from *Plasmodium* by RNase protection assay.

.

CHAPTER 6:

DISCUSSION

6. DISCUSSION

A lot of progress has been made in characterizing biogenesis and regulation of *Terra* while knowledge on biological functions of *Terra* is still limited. For the purpose of studying functions of lncRNAs like *Terra,* we have established a RNA-DNA FISH system using amino labeled probes and a post-fixation approach. Besides the post-fixation approach (using formaldehyde fixation to protect RNA FISH signals from the damage of the DNA FISH), an alternative approach for RNA-DNA FISH is to carry out the hybridization steps of the RNA FISH and the DNA FISH simultaneously. The presumption of this approach is that the RNA target survives the denaturation conditions used for DNA FISH. Two disadvantages should be considered for this approach. Firstly, it requires the FISH probes to be able to recognize the DNA and the RNA targets specifically, which is difficult to achieve in some cases, for example bi-directionally transcribed RNA transcripts coating the DNA region in *cis*. Secondly, a reliable denaturation condition, which is strong enough to denature the DNA target and mild enough to be harmless to the RNA target, may be difficult to find. These difficulties can be circumvented by our established method.

From the growing list of nuclear lncRNAs in the current research many are involved in epigenetic regulations. The established method can be used to study the long list of regulatory lncRNAs and their DNA targets in the nucleus. In addition, the method may also be applied to clinical investigations. For example, the increased gene copy number of FGFR1 is used as a biomarker to predict the response of squamous-cell lung cancer patients to FGFR tyrosine kinase inhibitor in clinical trials (Wynes et al., 2014). However, high level of FGFR1 gene expression is not always directly related with the increased gene copy number (Wynes et al., 2014). Our RNA-DNA FISH method can provide direct observation at the single cell level for this topic.

Using the established method, we show that the average length of telomeres associated with *Terra* is slightly longer than those non-associated, but the difference is not statistically significant (Fig. 4.6). This observation does not lend strong support to the hypothesis that *Terra* is associated with long telomeres to inhibit telomerase activity and to concentrate the telomerase activity onto short telomeres (Schoeftner

and Blasco, 2008). So far studies on biological functions of *TERRA* have ascribed the role of telomere length homeostasis, compaction and replication to it. They were mostly based on *in vitro* experiments performed using short *TERRA*-mimetic oligonucleotides. These studies assigned a telomerase inhibitory activity to *TERRA.*

On testing this hypothesis *in vivo* we found that *Terra* association with the distal telomere of X chromosome did not have a negative effect on its length when compared to the centromeric telomere over the course of 100+ cell divisions (Fig. 4.16). This is contrary to the observations made from *in vitro* studies. Another observation in this study was the occurrence of chromosome fusions undergone by sex chromosomes during the late passages of telomerase null F19 cells (Fig. 4.18). Chromosome fusion occurs on chromosomes with extremely short telomeres (Niida et al., 1998). Based on our results, X chromosome inherently possessed a shorter centromeric telomere length than distal telomere and the telomeres of the autosomes. Therefore, the centromeric end telomere in X chromosome reached its critically short length earlier following which X chromosome was among the earliest chromosomes to undergo fusion. In corroboration with this notion, a previous study had shown that the inactive X chromosome had an accelerated telomere shortening, compared with autosomes (Surralles et al., 1999). In comparison, Y chromosome fused at a higher frequency than X chromosome. Taken together, these observations reveal that sex chromosomes are under high pressures in maintaining their telomere length. These observations do not support the telomerase inhibitory role of *Terra* based on *in vitro* studies but, are in fact consistent with a new hypothesis put forth by a recent live cell imaging study that *Terra* may recruit telomerase onto a telomere to maintain its length (Cusanelli et al., 2013). Cusanelli and Chartrand group observed accumulation of *TERRA* in a nuclear focus followed by its association with telomerase during S phase (Cusanelli and Chartrand, 2014). Most intriguingly, as a *TERRA* cluster moved around the nucleus, it was ten times more likely to associate with the telomere from which it originated than with another telomere. These results suggested that *TERRA* plays a positive role in telomere elongation—first, by acting as a scaffold that sequesters telomerase, and second, by directing preformed telomerase clusters to the short telomere from which the *TERRA* cluster originated. With this hypothesis, our observations can be reasonably explained. X chromosome is under high pressure to maintain its telomere

length; therefore *Terra* is transcribed at high level on its telomere to recruit telomerase. This could be the reason of a *Terra* foci being associated with the sex chromosomes (Zhang et al., 2009).

In light of these observations it is tempting to speculate that *TERRA* transcripts could serve as unique identifiers of the chromosome end from which they were transcribed, thereby aiding the recruitment of telomerase to the shortest telomeres. These transcripts might contain distinct sequences for directing them to the telomere of their origin and recognition could be mediated by DNA/RNA interactions. The sequence differences in the subtelomeric section of *TERRA* may serve to mediate targeting, a possibility that needs to be tested. Alternatively, structural changes could occur at short telomeres which increase their affinity for a *TERRA*-telomerase cluster. A plethora of protein-protein interactions and posttranslational modifications have already been implicated in telomerase recruitment and activation, and placing *TERRA* into this network will be a necessary and exciting next step. The observation that in mouse cells *Terra* can favor telomere length elongation to prevent chromosome fusion when telomerase is present suggests one possible mechanism by which *Terra* may contribute to telomere elongation.

To gain better knowledge about the functional roles of *Terra*, we need to develop a robust system to modify *Terra* expression *in vivo*. Recent studies in yeast and human cells have used transcriptionally inducible telomeres (tiTELS) to enhance *TERRA* expression from certain telomeres (Maicher et al., 2012). One drawback of using inducible telomeres is the contradiction in observations made regarding the effect of induced telomere transcription in different organisms. This may either be an effect of differing telomere length homeostasis mechanisms in organisms, or it could also be due to the differential levels of transcription achieved by the inducible promoters used in the respective systems. There is a necessity to develop a global approach to modify *Terra* transcription where efficiency does not fluctuate between organisms so that it can be applied broadly. We can consider use of transgenes or design plasmids from which *Terra* can be transcribed. The limited knowledge available on *Terra* promoters hinders our attempts to knockdown *Terra* expression by directly targeting the promoters. A better understanding of the subtelomeric region could help us in this direction.

One interesting observation made through our experiments is the presence of *Terra* outside nucleus (Fig. 4.38). Studies in mouse and other organisms report that *Terra* is localized to the nucleus (Azzalin et al., 2007; Porro et al., 2010; Schoeftner and Blasco, 2008). But, the presence of *Terra* signal in the cytoplasm of *Plasmodium*, coupled with the observation that *Terra* may be present in human RBCs (Fig 4.36B), a nucleus-free cell, led to the question- does *Terra* fraction exists in the cytoplasm? This idea can reasonably explain why our *Terra* knock down could be detected by Northern blot but not by RNA-FISH. Our shRNA system could have targeted the *Terra* fraction in the cytoplasm leading to the generation of partial knock down of *Terra*. The presence of *Terra* in cytoplasm has many implications on the role performed by *Terra* in cells. With the help of subcellular fractionation we hope to establish and confirm the presence of *Terra* in the cytoplasm, which would further help us in analyzing the function of this RNA in the parasite and its possible role in malarial pathology.

From our studies we concur that other than being a regulator of the telomerase, *Terra* may play other functional roles in the cell. It has been noted that *Terra* expression sensitively responds to cellular stresses like heat shock (Fig. 4.21). Besides, *Terra* is not always associated with the telomere DNA. Even on its upregulation in response to stresses the distribution of *Terra* in the nucleus is not affected (Fig. 4.25). The lengths of telomeres associated and non-associated with *Terra* does not differ significantly (Fig. 4.6). We could detect *TERRA* from human RBC and also in the cytoplasm of malarial parasite suggesting that *Terra* may exist outside the nucleus. Therefore, although *Terra* is transcribed from the telomere DNA, it may not always associate with telomere DNA in *cis* to function as a telomerase inhibitor. Future studies need to provide strong *in vivo* evidence to fully elucidate the functional roles of *Terra*. Nonetheless, our established RNA-DNA FISH method provides the first direct observation on the telomere length of individual telomeres associated and non-associated with *Terra*.

CHAPTER 7: FUTURE DIRECTIONS

7. FUTURE DIRECTIONS

Our initial analyses of *Terra* functions have revealed that it may not perform an inhibitory role in telomere length regulation. To further confirm our hypothesis we need to knockdown *Terra* levels and observe its effect on telomere maintenance and regulation. As *Terra* is mostly localized to the nucleus it limits the use of RNAi for loss-of-function studies due to the low prevalence of the requisite enzymes in the nucleus (Guttman and Rinn, 2012). Therefore, there is a need to use systems that target and knock down nuclear RNAs. LNA™ gapmers are potent antisense oligonucleotides used for highly efficient inhibition of mRNA and lncRNA function. Antisense oligonucleotides (ASOs) are short, synthetic 14–22 nt oligonucleotides that localize to the nucleus. They include phosphorothioate (PS) linkages that confer nuclease resistance, thus enhancing intracellular stability. A gapmer is a chimeric antisense oligonucleotide that contains a central block of deoxynucleotide monomers sufficiently long to induce RNase H cleavage. RNase H, an endogenous enzyme found in all cells, specifically cleaves RNA linkages in a double-stranded RNA:DNA heteroduplex (Walder and Walder, 1988), making RNase H cleavage of the target the most effective mechanism for antisense activity. Thus, RNase H–dependent antisense silencing using gapmer ASOs is a powerful tool for specific modulation of nuclear noncoding RNAs and can be used to knockdown *Terra* for studying its functions.

From our experimental study we observed that in addition to being associated with telomere ends, telomere non-associated *Terra* is present in the nucleoplasm. Coupled with the observation that distribution pattern of *Terra* does not vary under heat stress, even though there is a signal upregulation, led us to speculate that *Terra* may perform roles apart from telomere length regulation. For exploring this possibility it would be interesting to analyze the interacting protein partners of *Terra*. This can either be performed *in silico* by utilizing catRAPID algorithm which uses the sequence of proteins and RNA to compute likely interaction pairs (Bellucci et al., 2011) or by using pull-down techniques. ChIRP (chromatin isolation by RNA purification), CHART (capture hybridization analysis of RNA targets) and RAP (RNA antisense

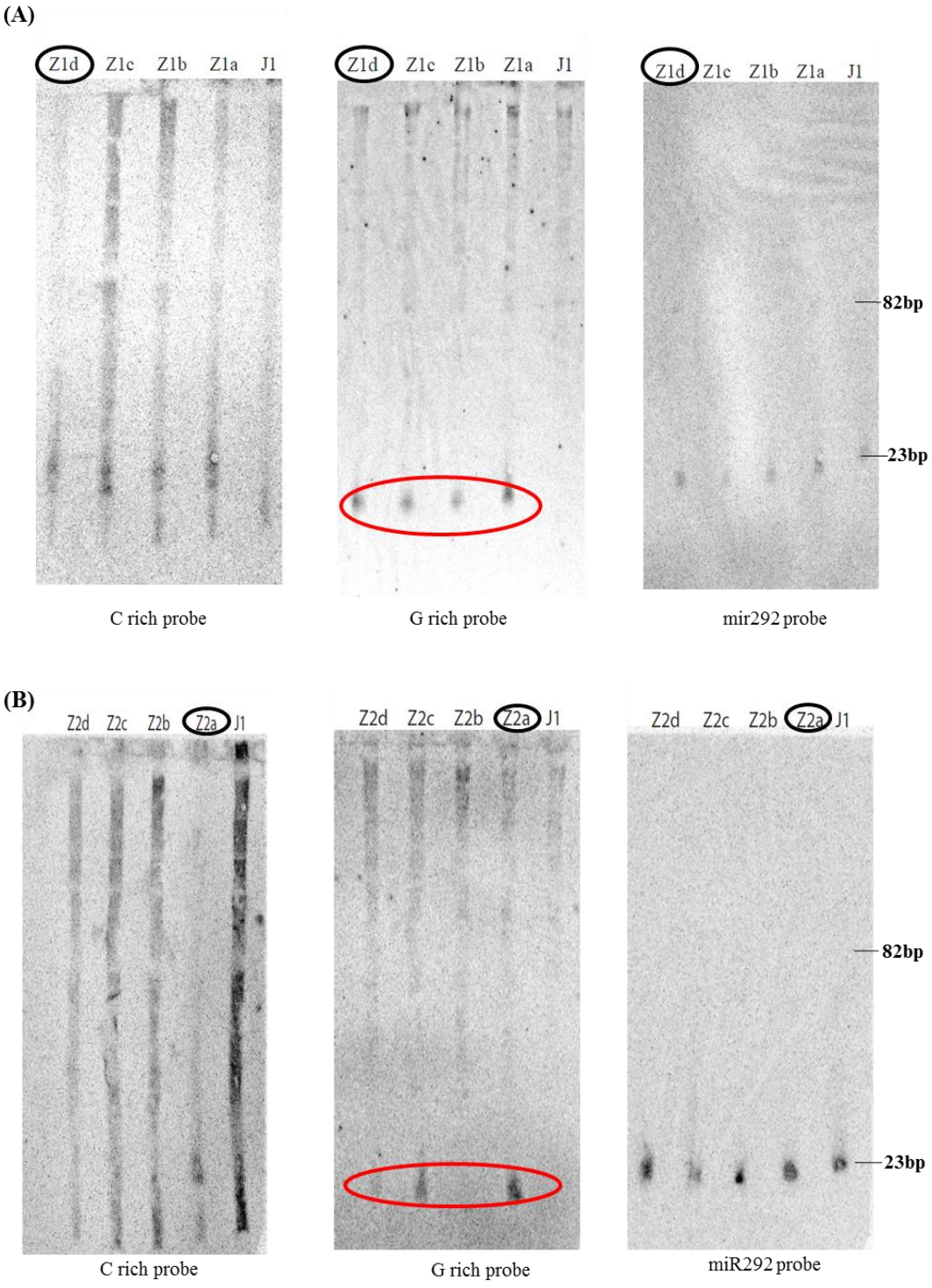
purification) are techniques based on the same basic idea of using biotinylated oligonucleotides complementary to the RNA of interest as a handle to pull down associated proteins or chromatin. This can be followed by next-generation sequencing and/or mass spectrometry to identify the proteins associated with the RNA and the genomic locations at which those interactions occurred (Chu et al., 2011; Engreitz et al., 2013; Simon et al., 2011). In one pair of recent studies Mitchell Guttman of the California Institute of Technology and Jeannie Lee of the University of Pennsylvania, used RAP and CHART (respectively) to probe the mechanism by which the *Xist* lncRNA spreads over and inactivates the silent X chromosome in female mammalian cells (Engreitz et al., 2013; Simon et al., 2013).

In this study, *Terra* was detected from RBCs which are nucleus-free cells and the cytoplasm of malaria parasite *Plasmodium*. It is not uncommon for lncRNAs to be translocated to the cytoplasm. Studies have shown that a substantial proportions of lncRNAs reside within, or are dynamically shuttled, to the cytoplasm where they regulate protein localization (Willingham et al., 2005), mRNA translation (Carrieri et al., 2012) and stability (Gong and Maquat, 2011). For example, the NFAT (Nuclear factor of activated T-cells) transcription factor is trafficked from the cytoplasm to the nucleus to activate target genes in response to calcium-dependent signals. A lncRNA, NRON (ncRNA repressor of the nuclear factor of activated T cells), complexes with importin-β proteins and regulates the trafficking of NFAT. Notably, NRON inhibits the trafficking of NFAT to the nucleus specifically, such that other proteins also trafficked by importin-β proteins, for e.g. NF-κB, are unaffected. *Terra* could be another such lncRNA that is trafficked to the cytoplasm to perform various functions. Further studies can be performed to prove the presence of *Terra* in cytoplasm. One line of experiments can be done by using subcellular fractionation to separate the nucleus and cytoplasm followed by detection of *Terra* from the cytoplasmic fraction by various techniques like Northern hybridization, qPCR and RNA pull-down. Different cell types can be analyzed to show that translocation into cytoplasm is a general phenomenon for the localization of *Terra*. Also the possible role performed by *Terra* in the cytoplasm is an interesting question. Looking at the protein interacting partners of *Terra* in cytoplasm would give us information about its functional roles in the cytoplasm.

Apart from identifying its functions it would be interesting to identify the promoter region for *Terra* transcription in mouse cells. It could be achieved by doing 5' RACE (Rapid amplification of cDNA ends) followed by cloning different regions upstream of the main 5′RACE products into promoter-free luciferase reporter vectors. The heterogeneity in *Terra* length stems at least in part, from the use of multiple transcription starts and by these studies we can narrow down the promoter regions associated with *Terra* transcription and the transcription start sites involved. Further the subtelomere ends of individual chromosomes should be examined for *Terra* transcription to get a wholesome idea about its transcriptional regulation.

In our study we detected *Terra* in malaria parasite *Plasmodium falciparum* by Northern hybridization, RNA FISH and RNase protection assay. It is known that the antigenic determinant *var* genes which help the pathogen evade human immune response are transcribed from the region adjacent to telomeres. With the known role of *Terra* in heterochromatization and chromatin modification (Deng et al., 2009)**,** it is possible that *Terra* may play a part in regulation of *var* gene expression, possibly through chromatin remodeling. Future studies will be important in establishing a direct molecular link between *Terra* transcripts and parasite proteins. We also observed a variation in *Terra* expression over different intra-erythrocytic stages of the parasite. We can carry out experiments to identify if the blood stage expression variation of *Terra* is associated with change in parasite protein expression. Recent studies in model eukaryotes suggest that lncRNAs represent a new paradigm in genome regulation and chromatin remodeling. Hence, profiling the non-coding transcriptome of drug-resistant parasites, parasites with mis-regulated virulence gene phenotypes, and hyper-virulent clinical isolates and studying *Terra* expression variation is an exciting new research direction in the quest to study the role of *Terra* in the pathology of malaria parasite.

8. APPENDIX


(A)
Z1d Z1c Z1b Z1a J1
Z1d Z1c Z1b Z1a J1
Z1d Z1c Z1b Z1a J1
82bp
23bp
C rich probe
G rich probe
mir292 probe
(B)
Z2d Z2c Z2b Z2a J1
Z2d Z2c Z2b Z2a J1
Z2d Z2c Z2b Z2a J1
82bp
23bp
C rich probe
G rich probe
miR292 probe

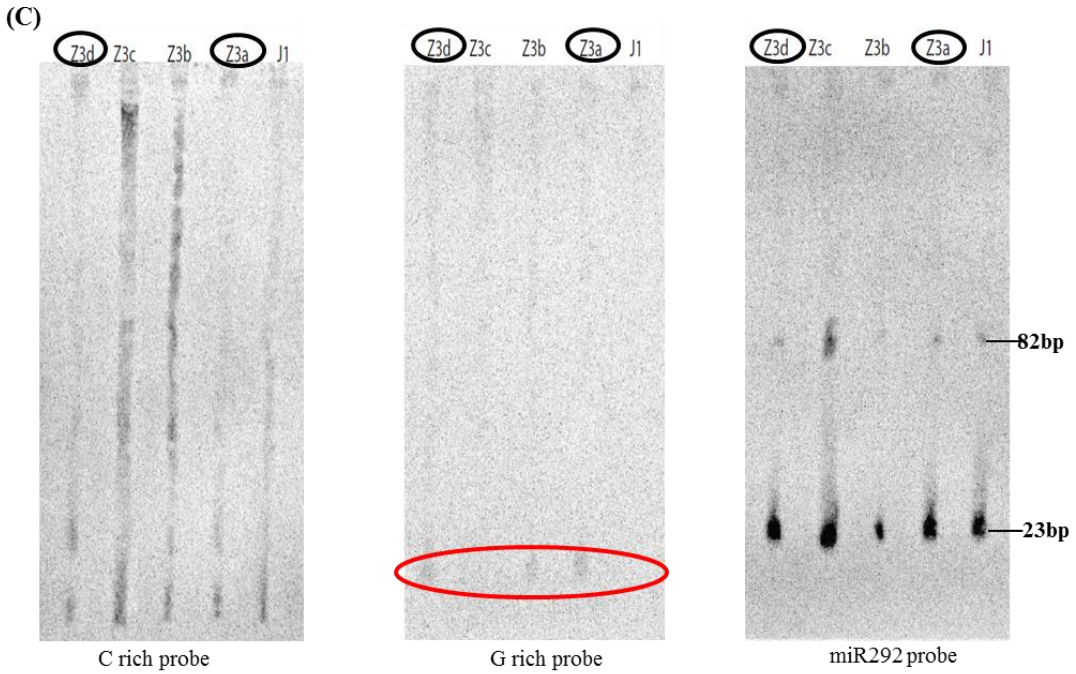


Figure S1 (A, B and C). Results for Northern Hybridization for cell lines selected after Z1, Z2 and Z3 shRNA transduction respectively. The RNA samples are labeled on top of each Blot. The area highlighted with a red circle denotes the shRNA expression in the cell lines. The highlighted black circles on top of each panel denote the candidate cell lines suspected to have lowered *Terra* expression. The probe used is labeled below each blot.

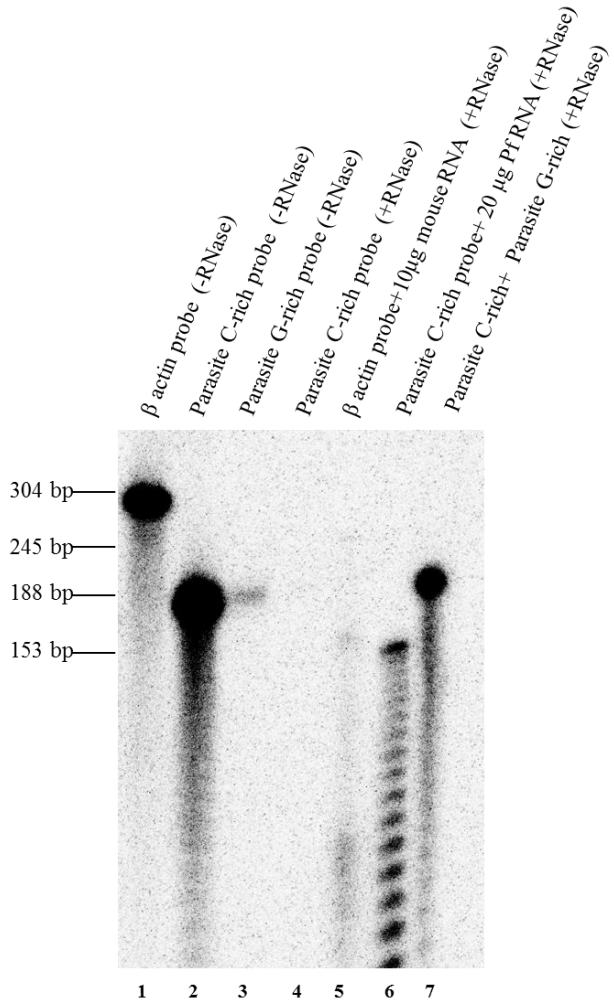


Figure S2. RNase protection assay to test the presence of *Terra* in *P. falciparum*.

Table S1. List of primers used for qPCR

Transcript	**Primer**	**Sequence(5′-3′)**
2q-*Terra*	5′ primer	TTTCCAGTGATGGCCGACTAG
2q- *Terra*	3′ primer	CCCCGGAGCTCTTGACTCT
5q- *Terra*	5′ primer	ATTAACAAGCACAAGAGGGTAGCA
5q- *Terra*	3′ primer	CAACCATACCTGAAATGCCTAGATC
11q- *Terra*	5′ primer	TGCCATTGGAACACAGCAA
11q- *Terra*	3′ primer	CGTCTGCTGAGGTCCACAGA
TeloCen- *Terra*	5′ primer	CCAAAGTTTCTGCAAGGCAAA
TeloCen- *Terra*	3′ primer	CCCAATCTGTTGGTGGTCTTTT
18q- *Terra*	5′ primer	CAGGCCAAAGAAGGGACAGA
18q- *Terra*	3′ primer	GCTTCCTCACTGATCCACAGTACA
β-actin	5′ primer	ACTGCCGCATCCTCTTCCTC
β-actin	3′ primer	CCGCTCGTTGCCAATAGTGA
CTCF	5′ primer	TTGTCACGCTCGGTTTACCC
CTCF	3′ primer	CGCAAGTGGACACCCAAATC
Rad21	5′ primer	AGATGACAATGGCTCACTGGG
Rad21	3′ primer	AATGGGCTCCAACGCAAAAG

Sequences of *Terra* primers taken from (Deng et al., 2012a)

9. REFERENCES

Amaral, P., Mattick, J., 2008. Noncoding RNA in development. Mammalian Genome 19, 454-492.

Ambros, V., 2004. The functions of animal microRNAs. Nature 431, 350-355.

Ambros, V., Lee, R.C., Lavanway, A., Williams, P.T., Jewell, D., 2003a. MicroRNAs and other tiny endogenous RNAs in C. elegans. Curr Biol 13, 807 - 818.

Ambros, V., Bartel, B., Bartel, D.P., Burge, C.B., Carrington, J.C., Chen, X., Dreyfuss, G., Eddy, S.R., Griffiths-Jones, S.a.M., Marshall, M., Matzke, M., Ruvkun, G., Tuschl, T., 2003b. A uniform system for microRNA annotation. RNA 9, 277-279.

Ambrosini, A., Paul, S., Hu, S., Riethman, H., 2007. Human subtelomeric duplicon structure and organization. Genome biology 8, R151.

Anno, Y.-N.L., Myslinski, E., Ngondo-Mbongo, R.P., Krol, A., Poch, O., Lecompte, O., Carbon, P., 2010. Genome-wide evidence for an essential role of the human Staf/ZNF143 transcription factor in bidirectional transcription. Nucleic acids research 39, 3116-3127.

Aravin, A.A., Naumova, N.M., Tulin, A.V., Vagin, V.V., Rozovsky, Y.M., Gvozdev, V.A., 2001. Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the D. melanogaster germline. Current Biology 11, 1017-1027.

Aravin, A.A., Sachidanandam, R., Bourc'his, D., Schaefer, C., Pezic, D., Toth, K.F., Bestor, T., Hannon, G.J., 2008. A piRNA Pathway Primed by Individual Transposons Is Linked to De Novo DNA Methylation in Mice. Molecular Cell 31, 785-799.

Aravin, A.A., Lagos-Quintana, M., Yalcin, A., Zavolan, M., Marks, D., Snyder, B., Gaasterland, T., Meyer, J., Tuschl, T., 2003. The small RNA profile during Drosophila melanogaster development. Dev Cell 5, 337 - 350.

Arnoult, N., Van Beneden, A., Decottignies, A., 2012. Telomere length regulates TERRA levels through increased trimethylation of telomeric H3K9 and HP1alpha. Nature structural & molecular biology 19, 948-956.

Arora, R., Brun, C.M., Azzalin, C.M., 2012. Transcription regulates telomere dynamics in human cancer cells. Rna 18, 684-693.

Artandi, S.E., Depinho, R.A., 2010. Telomeres and telomerase in cancer. Carcinogenesis 31, 9-18.

Azzalin, C.M., Lingner, J., 2008. Telomeres: the silence is broken. Cell cycle (Georgetown, Tex.) 7, 1161-1165.

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

Balaji, S., Babu, M.M., Iyer, L.M., Aravind, L., 2005. Discovery of the principal specific transcription factors of Apicomplexa and their implication for the evolution of the AP2-integrase DNA binding domains. Nucleic acids research 33, 3994-4006.

Balogh, G., Horvath, I., Nagy, E., Hoyk, Z., Benko, S., Bensaude, O., Vigh, L., 2005. The hyperfluidization of mammalian cell membranes acts as a signal to initiate the heat shock protein response. FEBS Journal 272, 6077-6086.

Bartel, D.P., 2004. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. 116, 281-297.

Batenburg, N.L., Mitchell, T.R., Leach, D.M., Rainbow, A.J., Zhu, X.D., 2012. Cockayne Syndrome group B protein interacts with TRF2 and regulates telomere length and stability. Nucleic acids research 40, 9661-9674.

Batista, P.J., Ruby, J.G., Claycomb, J.M., Chiang, R., Fahlgren, N., Kasschau, K.D., Chaves, D.A., Gu, W., Vasale, J.J., Duan, S., Conte, D., Luo, S., Schroth, G.P., Carrington, J.C., Bartel, D.P., Mello, C.C., 2008. PRG-1 and 21U-RNAs Interact to Form the piRNA Complex Required for Fertility in C. elegans. Molecular Cell 31, 67-78.

Baum, J., Papenfuss, A.T., Mair, G.R., Janse, C.J., Vlachou, D., Waters, A.P., Cowman, A.F., Crabb, B.S., De Koning-Ward, T.F., 2009. Molecular genetics and comparative genomics reveal RNAi is not functional in malaria parasites. Nucleic acids research 37, 3788-3798.

Baumann, P., Cech, T.R., 2001. Pot1, the putative telomere end-binding protein in fission yeast and humans. Science 292, 1171-1175.

Baur, J.A., Zou, Y., Shay, J.W., Wright, W.E., 2001. Telomere position effect in human cells. Science 292, 2075-2077.

Bekaert, S., Koll, S., Thas, O., Van Oostveldt, P., 2002. Comparing telomere length of sister chromatids in human lymphocytes using three-dimensional confocal microscopy. Cytometry 48, 34-44.

Bellucci, M., Agostini, F., Masin, M., Tartaglia, G.G., 2011. Predicting protein associations with long noncoding RNAs. Nature methods 8, 444-445.

Benetti, R., Garcia-Cao, M., Blasco, M.A., 2007. Telomere length regulates the epigenetic status of mammalian telomeres and subtelomeres. Nat Genet 39, 243-250.

Bianchi, A., Shore, D., 2008. How telomerase reaches its end: mechanism of telomerase regulation by the telomeric complex. Mol Cell 31, 153-165.

Biffi, G., Tannahill, D., Balasubramanian, S., 2012. An intramolecular G-quadruplex structure is required for binding of telomeric repeat-containing RNA to the telomeric protein TRF2. J Am Chem Soc 134, 11974-11976.

Brennecke, J., Malone, C.D., Aravin, A.A., Sachidanandam, R., Stark, A., Hannon, G.J., 2008. An Epigenetic Role for Maternally Inherited piRNAs in Transposon Silencing. Science 322, 1387-1392.

Broadbent, K.M., Park, D., Wolf, A.R., Van Tyne, D., Sims, J.S., Ribacke, U., Volkman, S., Duraisingh, M., Wirth, D., Sabeti, P.C., Rinn, J.L., 2011. A global transcriptional analysis of Plasmodium falciparum malaria reveals a novel family of telomere-associated lncRNAs. Genome biology 12, R56.

Calin, G.A., Dumitru, C.D., Shimizu, M., Bichi, R., Zupo, S., Noch, E., Aldler, H., Rattan, S., Keating, M., Rai, K., Rassenti, L., Kipps, T., Negrini, M., Bullrich, F., Croce, C.M., 2002. Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proceedings of the National Academy of Sciences 99, 15524-15529.

Campbell, T.L., De Silva, E.K., Olszewski, K.L., Elemento, O., Llinas, M., 2010. Identification and genome-wide prediction of DNA binding specificities for the ApiAP2 family of regulators from the malaria parasite. PLoS pathogens 6, e1001165.

Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shimokawa, K., Bajic, V.B., Brenner, S.E., Batalov, S., Forrest, A.R.R., Zavolan, M., Davis, M.J., Wilming, L.G., Aidinis, V., Allen, J.E., Ambesi-Impiombato, A., Apweiler, R., Aturaliya, R.N., Bailey, T.L., Bansal, M., Baxter, L., Beisel, K.W., Bersano, T., Bono, H., Chalk, A.M., Chiu, K.P., Choudhary, V., Christoffels, A., Clutterbuck, D.R., Crowe, M.L., Dalla, E., Dalrymple, B.P., De Bono, B., Gatta, G.D., Di Bernardo, D., Down, T., Engstrom, P., Fagiolini, M., Faulkner, G., Fletcher, C.F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M., Georgii-Hemming, P., Gingeras, T.R., Gojobori, T., Green, R.E., Gustincich, S., Harbers, M., Hayashi, Y., Hensch, T.K., Hirokawa, N., Hill, D., Huminiecki, L., Iacono, M., Ikeo, K., Iwama, A., Ishikawa, T., Jakt, M., Kanapin, A., Katoh, M., Kawasawa, Y., Kelso, J., Kitamura, H., Kitano, H., Kollias, G., Krishnan, S.P.T., Kruger, A., Kummerfeld, S.K., Kurochkin, I.V., Lareau, L.F., Lazarevic, D., Lipovich, L., Liu, J., Liuni, S., Mcwilliam, S., Babu, M.M., Madera, M., Marchionni, L., Matsuda, H., Matsuzawa, S., Miki, H., Mignone, F., Miyake, S., Morris, K., Mottagui-Tabar, S., Mulder, N., Nakano, N., Nakauchi, H., Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O., Okazaki, Y., Orlando, V., Pang, K.C., Pavan, W.J., Pavesi, G., Pesole, G., Petrovsky, N., Piazza, S., Reed, J., Reid, J.F., Ring, B.Z., Ringwald, M., Rost, B., Ruan, Y., Salzberg, S.L., Sandelin, A., Schneider, C., Schã¶Nbach, C., Sekiguchi, K., Semple, C.a.M., Seno, S., Sessa, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D., Sinclair, B., Sperling, S., Stupka, E., Sugiura, K., Sultana, R., Takenaka, Y., Taki, K., Tammoja, K., Tan, S.L., Tang, S., Taylor, M.S., Tegner, J., Teichmann, S.A., Ueda, H.R., Van Nimwegen, E., Verardo, R., Wei, C.L., Yagi, K., Yamanishi, H., Zabarovsky, E., Zhu, S., Zimmer, A.,

Hide, W., Bult, C., Grimmond, S.M., Teasdale, R.D., Liu, E.T., Brusic, V., Quackenbush, J., Wahlestedt, C., Mattick, J.S., Hume, D.A., Group, R.G.E.R., Genome Science, G., Kai, C., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori-Katayama, M., Suzuki, M., Aoki, J., Arakawa, T., Iida, J., Imamura, K., Itoh, M., Kato, T., Kawaji, H., Kawagashira, N., Kawashima, T., Kojima, M., Kondo, S., Konno, H., Nakano, K., Ninomiya, N., Nishio, T., Okada, M., Plessy, C., Shibata, K., Shiraki, T., Suzuki, S., Tagami, M., Waki, K., Watahiki, A., Okamura-Oho, Y., Suzuki, H., Kawai, J., Hayashizaki, Y., 2005. The Transcriptional Landscape of the Mammalian Genome. Science 309, 1559-1563.

Carrieri, C., Cimatti, L., Biagioli, M., Beugnet, A., Zucchelli, S., Fedele, S., Pesce, E., Ferrer, I., Collavin, L., Santoro, C., Forrest, A.R., Carninci, P., Biffo, S., Stupka, E., Gustincich, S., 2012. Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. Nature 491, 454-457.

Caslini, C., Connelly, J.A., Serna, A., Broccoli, D., Hess, J.L., 2009. MLL Associates with Telomeres and Regulates Telomeric Repeat-Containing RNA Transcription. Molecular and Cellular Biology 29, 4519-4526.

Celli, G.B., De Lange, T., 2005. DNA processing is not required for ATM-mediated telomere damage response after TRF2 deletion. Nature cell biology 7, 712-718.

Chaal, B.K., Gupta, A.P., Wastuwidyaningtyas, B.D., Luah, Y.H., Bozdech, Z., 2010. Histone deacetylases play a major role in the transcriptional regulation of the Plasmodium falciparum life cycle. PLoS pathogens 6, e1000737.

Chen, K., Rajewsky, Nikolaus, 2007. The evolution of gene regulation by transcription factors and microRNAs. Nature reviews. Genetics 8, 11.

Chen, L.-L., Carmichael, G.G., 2009. Altered Nuclear Retention of mRNAs Containing Inverted Repeats in Human Embryonic Stem Cells: Functional Role of a Nuclear Noncoding RNA. Molecular cell 35, 467-478.

Chen, L.-L., Decerbo, J.N., Carmichael, G.G., 2008. Alu element-mediated gene silencing. The EMBO journal 27, 1694-1705.

Chu, C., Qu, K., Zhong, F.L., Artandi, S.E., Chang, H.Y., 2011. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. Mol Cell 44, 667-678.

Clemson, C.M., Hutchinson, J.N., Sara, S.A., Ensminger, A.W., Fox, A.H., Chess, A., Lawrence, J.B., 2009. An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. Mol Cell 33, 717-726.

Collins, F.S., Patrinos, A., Jordan, E., Chakravarti, A., Gesteland, R., Walters, L., Doe, T.M.O.T., Groups, N.P., 1998. New Goals for the U.S. Human Genome Project: 1998-2003. Science 282, 682-689.

Consortium, T.F., Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shimokawa,

K., Bajic, V.B., Brenner, S.E., Batalov, S., Forrest, A.R.R., Zavolan, M., Davis, M.J., Wilming, L.G., Aidinis, V., Allen, J.E., Ambesi-Impiombato, A., Apweiler, R., Aturaliya, R.N., Bailey, T.L., Bansal, M., Baxter, L., Beisel, K.W., Bersano, T., Bono, H., Chalk, A.M., Chiu, K.P., Choudhary, V., Christoffels, A., Clutterbuck, D.R., Crowe, M.L., Dalla, E., Dalrymple, B.P., De Bono, B., Gatta, G.D., Di Bernardo, D., Down, T., Engstrom, P., Fagiolini, M., Faulkner, G., Fletcher, C.F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M., Georgii-Hemming, P., Gingeras, T.R., Gojobori, T., Green, R.E., Gustincich, S., Harbers, M., Hayashi, Y., Hensch, T.K., Hirokawa, N., Hill, D., Huminiecki, L., Iacono, M., Ikeo, K., Iwama, A., Ishikawa, T., Jakt, M., Kanapin, A., Katoh, M., Kawasawa, Y., Kelso, J., Kitamura, H., Kitano, H., Kollias, G., Krishnan, S.P.T., Kruger, A., Kummerfeld, S.K., Kurochkin, I.V., Lareau, L.F., Lazarevic, D., Lipovich, L., Liu, J., Liuni, S., Mcwilliam, S., Babu, M.M., Madera, M., Marchionni, L., Matsuda, H., Matsuzawa, S., Miki, H., Mignone, F., Miyake, S., Morris, K., Mottagui-Tabar, S., Mulder, N., Nakano, N., Nakauchi, H., Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O., Okazaki, Y., Orlando, V., Pang, K.C., Pavan, W.J., Pavesi, G., Pesole, G., Petrovsky, N., Piazza, S., Reed, J., Reid, J.F., Ring, B.Z., Ringwald, M., Rost, B., Ruan, Y., Salzberg, S.L., Sandelin, A., Schneider, C., Schönbach, C., Sekiguchi, K., Semple, C.a.M., Seno, S., Sessa, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D., Sinclair, B., Sperling, S., Stupka, E., Sugiura, K., Sultana, R., Takenaka, Y., Taki, K., Tammoja, K., Tan, S.L., Tang, S., Taylor, M.S., Tegner, J., Teichmann, S.A., Ueda, H.R., Van Nimwegen, E., Verardo, R., Wei, C.L., Yagi, K., Yamanishi, H., Zabarovsky, E., Zhu, S., Zimmer, A., Hide, W., Bult, C., Grimmond, S.M., Teasdale, R.D., Liu, E.T., Brusic, V., Quackenbush, J., Wahlestedt, C., Mattick, J.S., Hume, D.A., Group, R.G.E.R., Group, G.S., Kai, C., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori-Katayama, M., Suzuki, M., Aoki, J., Arakawa, T., Iida, J., Imamura, K., Itoh, M., Kato, T., Kawaji, H., Kawagashira, N., Kawashima, T., Kojima, M., Kondo, S., Konno, H., Nakano, K., Ninomiya, N., Nishio, T., Okada, M., Plessy, C., Shibata, K., Shiraki, T., Suzuki, S., Tagami, M., Waki, K., Watahiki, A., Okamura-Oho, Y., Suzuki, H., Kawai, J., Hayashizaki, Y., 2005. The Transcriptional Landscape of the Mammalian Genome. Science 309, 1559-1563.

Coulson, R.M., Hall, N., Ouzounis, C.A., 2004. Comparative genomics of transcriptional control in the human malaria parasite Plasmodium falciparum. Genome Res 14, 1548-1554.

Crick, F., 1970. Central dogma of molecular biology. Nature 227, 561-563.

Cui, L., Miao, J., 2010. Chromatin-mediated epigenetic regulation in the malaria parasite Plasmodium falciparum. Eukaryotic cell 9, 1138-1149.

Cusanelli, E., Chartrand, P., 2014. Telomeric noncoding RNA: telomeric repeat-containing RNA in telomere biology. Wiley interdisciplinary reviews. RNA 5, 407-419.

Cusanelli, E., Romero, C.A., Chartrand, P., 2013. Telomeric noncoding RNA TERRA is induced by telomere shortening to nucleate telomerase molecules at short telomeres. Mol Cell 51, 780-791.

Czech, B., Malone, C.D., Zhou, R., Stark, A., Schlingeheyde, C., Dus, M., Perrimon, N., Kellis, M., Wohlschlegel, J.A., Sachidanandam, R., Hannon, G.J., Brennecke, J., 2008. An endogenous small interfering RNA pathway in Drosophila. Nature 453, 798-802.

Davidson, J.F., Schiestl, R.H., 2001. Mitochondrial Respiratory Electron Carriers Are Involved in Oxidative Stress during Heat Stress inSaccharomyces cerevisiae. Molecular and Cellular Biology 21, 8483-8489.

Denchi, E.L., De Lange, T., 2007. Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1. Nature 448, 1068-1071.

Deng, Z., Norseen, J., Wiedmer, A., Riethman, H., Lieberman, P.M., 2009. TERRA RNA Binding to TRF2 Facilitates Heterochromatin Formation and ORC Recruitment at Telomeres. Molecular cell 35, 403-413.

Deng, Z., Wang, Z., Xiang, C., Molczan, A., Baubet, V., Conejo-Garcia, J., Xu, X., Lieberman, P.M., Dahmane, N., 2012a. Formation of telomeric repeat-containing RNA (TERRA) foci in highly proliferating mouse cerebellar neuronal progenitors and medulloblastoma. Journal of Cell Science.

Deng, Z., Wang, Z., Stong, N., Plasschaert, R., Moczan, A., Chen, H.S., Hu, S., Wikramasinghe, P., Davuluri, R.V., Bartolomei, M.S., Riethman, H., Lieberman, P.M., 2012b. A role for CTCF and cohesin in subtelomere chromatin organization, TERRA transcription, and telomere end protection. The EMBO journal 31, 4165-4178.

Dieci, G., Fiorino, G., Castelnuovo, M., Teichmann, M., Pagano, A., 2007. The expanding RNA polymerase III transcriptome. Trends Genet 23, 614-622.

Dinger, M.E., Amaral, P.P., Mercer, T.R., Pang, K.C., Bruce, S.J., Gardiner, B.B., Askarian-Amiri, M.E., Ru, K., Soldã , G., Simons, C., Sunkin, S.M., Crowe, M.L., Grimmond, S.M., Perkins, A.C., Mattick, J.S., 2008. Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. Genome Research 18, 1433-1445.

Dong, J., Jiang, G., Asmann, Y.W., Tomaszek, S., Jen, J., Kislinger, T., Wigle, D.A., 2010. MicroRNA Networks in Mouse Lung Organogenesis. PloS one 5, e10854.

Engreitz, J.M., Pandya-Jones, A., Mcdonel, P., Shishkin, A., Sirokman, K., Surka, C., Kadri, S., Xing, J., Goren, A., Lander, E.S., Plath, K., Guttman, M., 2013. The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. Science 341, 1237973.

Ewan Birney, J.a.S., Anindya Dutta, Roderic Guigó, Thomas R. Gingeras, Elliott H. Margulies, Zhiping Weng, Michael Snyder, Emmanouil T. Dermitzakis, 2007. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature 447, 17.

Faghihi, M.A., 2010. Evidence for natural antisense transcript-mediated inhibition of microRNA function. GenomeBiology.com 11, R56.

Farnung, B.O., Brun, C.M., Arora, R., Lorenzi, L.E., Azzalin, C.M., 2012. Telomerase efficiently elongates highly transcribing telomeres in human cancer cells. PloS one 7, e35714.

Fazi, F., Nervi, C., 2008. MicroRNA: basic mechanisms and transcriptional regulatory networks for cell fate determination. Cardiovascular Research 79, 553-561.

Feng, J., Bi, C., Clark, B.S., Mady, R., Shah, P., Kohtz, J.D., 2006. The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. Genes & Development 20, 1470-1484.

Figueiredo, L.M., Pirritt, L.A., Scherf, A., 2000. Genomic organisation and chromatin structure of Plasmodium falciparum chromosome ends. Molecular and biochemical parasitology 106, 169-174.

Filipowicz, W., Bhattacharyya, Suvendra N., Sonenberg, Nahum, 2008. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? Nature reviews. Genetics 9, 13.

Freitas-Junior, L.H., Bottius, E., Pirrit, L.A., Deitsch, K.W., Scheidig, C., Guinet, F., Nehrbass, U., Wellems, T.E., Scherf, A., 2000. Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of P. falciparum. Nature 407, 1018-1022.

Fvorstemann, K., Tomari, Y., Du, T., Vagin, V.V., Denli, A.M., Bratu, D.P., Klattenhoff, C., Theurkauf, W.E., Zamore, P.D., 2005. Normal microRNA Maturation and Germ-Line Stem Cell Maintenance Requires Loquacious, a Double-Stranded RNA-Binding Domain Protein. PLoS biology 3, e236.

Gao, H., Cervantes, R.B., Mandell, E.K., Otero, J.H., Lundblad, V., 2007. RPA-like proteins mediate yeast telomere function. Nature structural & molecular biology 14, 208-214.

Garcia-Cao, M., Gonzalo, S., Dean, D., Blasco, M.A., 2002. A role for the Rb family of proteins in controlling telomere length. Nat Genet 32, 415-419.

Garvik, B., Carson, M., Hartwell, L., 1995. Single-stranded DNA arising at telomeres in cdc13 mutants may constitute a specific signal for the RAD9 checkpoint. Mol Cell Biol 15, 6128-6138.

Ghildiyal, M., Zamore, P.D., 2009. Small silencing RNAs: an expanding universe. Nature reviews. Genetics 10, 15.

Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., Lim, W.A., Weissman, J.S., Qi, L.S., 2013. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell 154, 442-451.

Girard, A., Hannon, G.J., 2008. Conserved themes in small-RNA-mediated transposon control. Trends in cell biology 18, 136-148.

Gong, C., Maquat, L.E., 2011. lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. Nature 470, 284-288.

Gonzalez, S., Pisano, D.G., Serrano, M., 2008. Mechanistic principles of chromatin remodeling guided by siRNAs and miRNAs. Cell cycle (Georgetown, Tex.) 7, 2601-2608.

Gonzalo, S., Blasco, M.A., 2005. Role of Rb family in the epigenetic definition of chromatin. Cell cycle (Georgetown, Tex.) 4, 752-755.

Goodrich, J.A., Kugel, Jennifer F., 2006. Non-coding-RNA regulators of RNA polymerase II transcription. Nat Rev Mol Cell Biol 7, 5.

Greider, C.W., Blackburn, E.H., 1987. The telomere terminal transferase of Tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity. Cell 51, 887-898.

Grimson, A., Srivastava, M., Fahey, B., Woodcroft, B. J., Chiang, H. R., King, N., Degnan, B. M., Rokhsar, D. S., and Bartel,D. P., 2008. Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. Nature 455, 5.

Grivna, S.T., Beyret, E., Wang, Z., Lin, H., 2006. A novel class of small RNAs in mouse spermatogenic cells. Genes & Development 20, 1709-1714.

Guttman, M., Rinn, J.L., 2012. Modular regulatory principles of large non-coding RNAs. Nature 482, 339-346.

Guttman, M., Amit, Ido, Garber, Manuel, French, Courtney, Lin, Michael F., Feldser, David, Huarte, Maite, Zuk, or, Carey, Bryce W., Cassady, John P., Cabili, Moran N., Jaenisch, Rudolf, Mikkelsen, Tarjei S., Jacks, Tyler, Hacohen, Nir, Bernstein, Bradley E., Kellis, Manolis, Regev, Aviv, Rinn, John L., Lander, Eric S., 2009. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature 458, 5.

Hansen, T.B., 2011. miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA. EMBO journal 30, 4414-4422.

Hicks, J., Tembhurne, P., Liu, H.-C., 2009. Identification of microRNA in the developing chick immune organs. Immunogenetics 61, 231-240.

Hockemeyer, D., Daniels, J.P., Takai, H., De Lange, T., 2006. Recent expansion of the telomeric complex in rodents: Two distinct POT1 proteins protect mouse telomeres. Cell 126, 63-77.

Hockemeyer, D., Palm, W., Else, T., Daniels, J.P., Takai, K.K., Ye, J.Z., Keegan, C.E., De Lange, T., Hammer, G.D., 2007. Telomere protection by mammalian Pot1 requires interaction with Tpp1. Nature structural & molecular biology 14, 754-761.

Horwich, M.D., Li, C., Matranga, C., Vagin, V., Farley, G., Wang, P., Zamore, P.D., 2007. The Drosophila RNA Methyltransferase, DmHen1, Modifies Germline piRNAs and Single-Stranded siRNAs in RISC. Current Biology 17, 1265-1272.

Huarte, M., 2010. A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. Cell 142, 409-419.

Huarte, M., Rinn, J.L., 2010. Large non-coding RNAs: missing links in cancer? Human Molecular Genetics 19, R152-R161.

Huarte, M., Guttman, M., Feldser, D., Garber, M., Koziol, M.J., Kenzelmann-Broz, D., Khalil, A.M., Zuk, O., Amit, I., Rabani, M., Attardi, L.D., Regev, A., Lander, E.S., Jacks, T., Rinn, J.L., 2010. A Large Intergenic Noncoding RNA Induced by p53 Mediates Global Gene Repression in the p53 Response. Cell 142, 409-419.

Hung, T., 2011. Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. Nature Genetics 43, 621-629.

Hüttenhofer, A., Schattner, P., Polacek, N., 2005. Non-coding RNAs: hope or hype? Trends in genetics : TIG 21, 289-297.

Iglesias, N., Redon, S., Pfeiffer, V., Dees, M., Lingner, J., Luke, B., 2011. Subtelomeric repetitive elements determine TERRA regulation by Rap1/Rif and Rap1/Sir complexes in yeast. EMBO reports 12, 587-593.

Inui, M., Martello, Graziano, Piccolo, Stefano, 2010. MicroRNA control of signal transduction. Nat Rev Mol Cell Biol 11, 10.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., Charpentier, E., 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337, 816-821.

Jolly, C., Metz, A., Govin, J., Vigneron, M., Turner, B.M., Khochbin, S., Vourc'h, C., 2004. Stress-induced transcription of satellite III repeats. The Journal of cell biology 164, 25-33.

Jopling, C.L., Yi, M., Lancaster, A.M., Lemon, S.M., Sarnow, P., 2005. Modulation of Hepatitis C Virus RNA Abundance by a Liver-Specific MicroRNA. Science 309, 1577-1581.

Kapranov, P., 2007. RNA maps reveal new RNA classes and a possible function for pervasive transcription. Science 316, 1484-1488.

Kapranov, P., Cheng, J., Dike, S., Nix, D.A., Duttagupta, R., Willingham, A.T., Stadler, P.F., Hertel, J., Hackermã¼Ller, J.R., Hofacker, I.L., Bell, I., Cheung, E., Drenkow, J., Dumais, E., Patel, S., Helt, G., Ganesh, M., Ghosh, S., Piccolboni, A., Sementchenko, V., Tammana, H., Gingeras, T.R., 2007. RNA Maps Reveal New RNA Classes and a Possible Function for Pervasive Transcription. Science 316, 1484-1488.

Karlseder, J., Hoke, K., Mirzoeva, O.K., Bakkenist, C., Kastan, M.B., Petrini, J.H., De Lange, T., 2004. The telomeric protein TRF2 binds the ATM kinase and can inhibit the ATM-dependent DNA damage response. PLoS biology 2, E240.

Kawamura, Y., Saito, K., Kin, T., Ono, Y., Asai, K., Sunohara, T., Okada, T.N., Siomi, M.C., Siomi, H., 2008. Drosophila endogenous small RNAs bind to Argonaute 2 in somatic cells. Nature 453, 793-797.

Khalil, A.M., Guttman, M., Huarte, M., Garber, M., Raj, A., Rivea Morales, D., Thomas, K., Presser, A., Bernstein, B.E., Van Oudenaarden, A., Regev, A., Lander, E.S., Rinn, J.L., 2009. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proceedings of the National Academy of Sciences.

Kim, D.H., Sætrom, P., Snøve, O., Rossi, J.J., 2008. MicroRNA-directed transcriptional gene silencing in mammalian cells. Proceedings of the National Academy of Sciences 105, 16230-16235.

Kim, V.N., 2005. MicroRNA biogenesis: coordinated cropping and dicing. Nat Rev Mol Cell Biol 6, 376-385.

Kim, V.N., 2006. Small RNAs just got bigger: Piwi-interacting RNAs (piRNAs) in mammalian testes. Genes & Development 20, 1993-1997.

Kim, V.N., Han, J., Siomi, M.C., 2009. Biogenesis of small RNAs in animals. Nat Rev Mol Cell Biol 10, 126-139.

Koering, C.E., Pollice, A., Zibella, M.P., Bauwens, S., Puisieux, A., Brunori, M., Brun, C., Martins, L., Sabatier, L., Pulitzer, J.F., Gilson, E., 2002. Human telomeric position effect is determined by chromosomal context and telomeric chromatin integrity. EMBO reports 3, 1055-1061.

Kohlmaier, A., Savarese, F., Lachner, M., Martens, J., Jenuwein, T., Wutz, A., 2004. A Chromosomal Memory Triggered by <italic>Xist</italic> Regulates Histone Methylation in X Inactivation. PLoS biology 2, e171.

Kuhl, N., Rensing, L., 2000. Heat shock effects on cell cycle progression. Cellular and Molecular Life Sciences CMLS 57, 450-463.

Lachner, M., O'carroll, D., Rea, S., Mechtler, K., Jenuwein, T., 2001. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature 410, 116-120.

Lagos-Quintana, M., Rauhut, R., Lendeckel, W., Tuschl, T., 2001. Identification of novel genes coding for small expressed RNAs. Science 294, 853 - 858.

Lai, L.T., Lee, P.J., Zhang, L.F., 2013. Immunofluorescence protects RNA signals in simultaneous RNA‚ÄìDNA FISH. Experimental Cell Research 319, 46-55.

Lansdorp, P.M., Verwoerd, N.P., Van De Rijke, F.M., Dragowska, V., Little, M.T., Dirks, R.W., Raap, A.K., Tanke, H.J., 1996. Heterogeneity in telomere length of human chromosomes. Hum Mol Genet 5, 685-691.

Lau, N.C., Lim, L.P., Weinstein, E.G., Bartel, D.P., 2001. An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. Science 294, 858 - 862.

Le, P.N., Maranon, D.G., Altina, N.H., Battaglia, C.L., Bailey, S.M., 2013. TERRA, hnRNP A1, and DNA-PKcs Interactions at Human Telomeres. Frontiers in oncology 3, 91.

Lee, J.T., 2002. Homozygous Tsix mutant mice reveal a sex-ratio distortion and revert to random X-inactivation. Nat Genet 32, 195-200.

Lee, J.T., 2005. Regulation of X-chromosome counting by Tsix and Xite sequences. Science 309, 768-771.

Lee, J.T., 2012. Epigenetic regulation by long noncoding RNAs. Science 338, 1435-1439.

Lee, R.C., Ambros, V., 2001. An extensive class of small RNAs in Caenorhabditis elegans. Science 294, 862 - 864.

Lee, R.C., Feinbaum, R.L., Ambros, V., 1993a. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75, 843 - 854.

Lee, R.C., Feinbaum, R.L., Ambros, V., 1993b. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75, 843-854.

Lenk, P., Wink, M., 1997. A RNA/RNA heteroduplex cleavage analysis to detect rare mutations in populations. Molecular ecology 6, 687-690.

Lin, J.J., Zakian, V.A., 1996. The Saccharomyces CDC13 protein is a single-strand TG1-3 telomeric DNA-binding protein in vitro that affects telomere behavior in vivo. Proc Natl Acad Sci U S A 93, 13760-13765.

Linardopoulou, E.V., Parghi, S.S., Friedman, C., Osborn, G.E., Parkhurst, S.M., Trask, B.J., 2007. Human subtelomeric WASH genes encode a new subclass of the WASP family. PLoS genetics 3, e237.

Lipinski, M.M., Jacks, T., 1999. The retinoblastoma gene family in differentiation and development. Oncogene 18, 7873-7882.

Luah, Y.-H., Chaal, B.K., Ong, E.Z., Bozdech, Z., 2010. A Moonlighting Function of <italic>Plasmodium falciparum</italic> Histone 3, Mono-Methylated at Lysine 9? PloS one 5, e10252.

Luke, B., Panza, A., Redon, S., Iglesias, N., Li, Z., Lingner, J., 2008. The Rat1p 5' to 3' exonuclease degrades telomeric repeat-containing RNA and promotes telomere elongation in Saccharomyces cerevisiae. Mol Cell 32, 465-477.

Maciej Szymañski, M.Z.B., Marek Zywicki, Jan Barciszewski, 2003. Noncoding RNA transcripts. J. Appl. Genet 44, 19.

Maeder, M.L., Linder, S.J., Cascio, V.M., Fu, Y., Ho, Q.H., Joung, J.K., 2013. CRISPR RNA-guided activation of endogenous human genes. Nature methods 10, 977-979.

Maicher, A., Kastner, L., Dees, M., Luke, B., 2012. Deregulated telomere transcription causes replication-dependent telomere shortening and promotes cellular senescence. Nucleic acids research.

Majid, S., Dar, Altaf A., Saini, Sharanjot, Yamamura, Soichiro, Hirata, Hiroshi, Tanaka, Yuichiro, Deng, Guoren, Dahiya, Rajvir, 2010. MicroRNA-205–directed transcriptional activation of tumor suppressor genes in prostate cancer. Cancer 116, 5637-5649.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., Dicarlo, J.E., Norville, J.E., Church, G.M., 2013a. RNA-guided human genome engineering via Cas9. Science 339, 823-826.

Mali, P., Aach, J., Stranges, P.B., Esvelt, K.M., Moosburner, M., Kosuri, S., Yang, L., Church, G.M., 2013b. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat Biotechnol 31, 833-838.

Malone, C.D., Hannon, G.J., 2009. Small RNAs as Guardians of the Genome. Cell 136, 656-668.

Mansfield, K.D., 2009. The ribonome: a dominant force in co-ordinating gene expression. Biology of the cell 101, 169-181.

Martianov, I., Ramadass, A., Serra Barros, A., Chow, N., Akoulitchev, A., 2007. Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. Nature 445, 666-670.

Mason, J.M., Frydrychova, R.C., Biessmann, H., 2008. Drosophila telomeres: an exception providing new insights. BioEssays : news and reviews in molecular, cellular and developmental biology 30, 25-37.

Matera Ag, T.R., Terns Mp, 2007. Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. Nat Rev Mol Cell Biol 8, 12.

Mattick, J.S., 2004. RNA regulation: a new genetics? Nature reviews. Genetics 5, 316-323.

Mcclintock, B., 1941. The stability of broken ends of chromosomes in Zea mays. Genetics 26, 234.

Mercer, T.R., Wilhelm, D., Dinger, M.E., Soldã , G., Korbie, D.J., Glazov, E.A., Truong, V., Schwenke, M., Simons, C., Matthaei, K.I., Saint, R., Koopman, P.,

Mattick, J.S., 2011. Expression of distinct RNAs from 3â€² untranslated regions. Nucleic acids research 39, 2393-2403.

Merrick, C.J., Duraisingh, M.T., 2010. Epigenetics in Plasmodium: what do we really know? Eukaryotic cell 9, 1150-1158.

Michael, M.Z., O' Connor, S.M., Van Holst Pellekaan, N.G., Young, G.P., James, R.J., 2003. Reduced Accumulation of Specific MicroRNAs in Colorectal Neoplasia11Note: Susan M. O' Connor and Nicholas G. van Holst Pellekaan contributed equally to this work. Molecular Cancer Research 1, 882-891.

Michishita, E., Mccord, R.A., Berber, E., Kioi, M., Padilla-Nash, H., Damian, M., Cheung, P., Kusumoto, R., Kawahara, T.L., Barrett, J.C., Chang, H.Y., Bohr, V.A., Ried, T., Gozani, O., Chua, K.F., 2008. SIRT6 is a histone H3 lysine 9 deacetylase that modulates telomeric chromatin. Nature 452, 492-496.

Morimoto, R.I., Santoro, M.G., 1998. Stress-inducible responses and heat shock proteins: New pharmacologic. Nature biotechnology 16.

Muller, H.J., 1932. Further studies on the nature and cause of gene mutation. In: Jones, D.F. (Ed.), Proceedings of the Sixth International Congress of Genetics, Brooklyn Botanic Gardens, Menasha, Wisconsin, pp. 213-255.

Munoz, P., Blanco, R., Flores, J.M., Blasco, M.A., 2005. XPF nuclease-dependent telomere loss and increased DNA damage in mice overexpressing TRF2 result in premature aging and cancer. Nat Genet 37, 1063-1071.

Nagano, T., Fraser, P., 2011. No-nonsense functions for long noncoding RNAs. Cell 145, 178-181.

Nakayama, J., Rice, J.C., Strahl, B.D., Allis, C.D., Grewal, S.I., 2001. Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. Science 292, 110-113.

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

Niida, H., Matsumoto, T., Satoh, H., Shiwa, M., Tokutake, Y., Furuichi, Y., Shinkai, Y., 1998. Severe growth defect in mouse cells lacking the telomerase RNA component. Nat Genet 19, 203-206.

Nishikura, K., 2010. Functions and Regulation of RNA Editing by ADAR Deaminases. Annual Review of Biochemistry 79, 321-349.

O'sullivan, R.J., Karlseder, J., 2010. Telomeres: protecting chromosomes against genome instability. Nat Rev Mol Cell Biol 11, 171-181.

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y.,

Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., Mckenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E., Hayashizaki, Y., 2002. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. Nature 420, 11.

Orom, U.A., Derrien, T., Beringer, M., Gumireddy, K., Gardini, A., Bussotti, G., Lai, F., Zytnicki, M., Notredame, C., Huang, Q., Guigo, R., Shiekhattar, R., 2010. Long Noncoding RNAs with Enhancer-like Function in Human Cells. Cell 143, 46-58.

Osada, H., Takahashi, T., 2007. MicroRNAs in biological processes and carcinogenesis. Carcinogenesis 28, 2-12.

Palm, W., De Lange, T., 2008. How shelterin protects mammalian telomeres. Annual review of genetics 42, 301-334.

Palm, W., Hockemeyer, D., Kibe, T., De Lange, T., 2009. Functional dissection of human and mouse POT1 proteins. Mol Cell Biol 29, 471-482.

Park, K.-J., Krishnan, V., O,Äômalley, B.W., Yamamoto, Y., Gaynor, R.B., 2005. Formation of an IKKŒ±-Dependent Transcription Complex Is Required for Estrogen Receptor-Mediated Gene Activation. Molecular cell 18, 71-82.

Peters, A.H., O'carroll, D., Scherthan, H., Mechtler, K., Sauer, S., Schofer, C., Weipoltshammer, K., Pagani, M., Lachner, M., Kohlmaier, A., Opravil, S., Doyle, M., Sibilia, M., Jenuwein, T., 2001. Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. Cell 107, 323-337.

Pinto, A., Stewart, D., Van Rooijen, N., Morahan, P., 1991. Selective depletion of liver and splenic macrophages using liposomes encapsulating the drug dichloromethylene diphosphonate: effects on antimicrobial resistance. Journal of leukocyte biology 49, 579-586.

Plasterk, R.H.A., 2006. Micro RNAs in Animal Development. 124, 877-881.

Plath, K., Fang, J., Mlynarczyk-Evans, S.K., Cao, R., Worringer, K.A., Wang, H., De La Cruz, C.C., Otte, A.P., Panning, B., Zhang, Y., 2003. Role of Histone H3 Lysine 27 Methylation in X Inactivation. Science 300, 131-135.

Ponting, C.P., Oliver, P.L., Reik, W., 2009. Evolution and Functions of Long Noncoding RNAs. Cell 136, 629-641.

Porro, A., Feuerhahn, S., Reichenbach, P., Lingner, J., 2010. Molecular dissection of telomeric repeat-containing RNA biogenesis unveils the presence of distinct and multiple regulatory pathways. Mol Cell Biol 30, 4808-4817.

Prasanth, K.V., Prasanth, S.G., Xuan, Z., Hearn, S., Freier, S.M., Bennett, C.F., Zhang, M.Q., Spector, D.L., 2005. Regulating Gene Expression through RNA Nuclear Retention. Cell 123, 249-263.

Qureshi, I.A., Mehler, M.F., 2012. Emerging roles of non-coding RNAs in brain evolution, development, plasticity and disease. Nature reviews. Neuroscience 13, 528-541.

Rackham, O., Shearwood, A.-M.J., Mercer, T.R., Davies, S.M.K., Mattick, J.S., Filipovska, A., 2011. Long noncoding RNAs are generated from the mitochondrial genome and regulated by nuclear-encoded proteins. RNA.

Raj, A., Van Den Bogaard, P., Rifkin, S.A., Van Oudenaarden, A., Tyagi, S., 2008. Imaging individual mRNA molecules using multiple singly labeled probes. Nature methods 5, 877-879.

Ray, A., Norden, B., 2000. Peptide nucleic acid (PNA): its medical and biotechnical applications and promise for the future. FASEB J 14, 1041-1060.

Redon, S., Reichenbach, P., Lingner, J., 2010. The non-coding RNA TERRA is a natural ligand and direct inhibitor of human telomerase. Nucleic acids research 38, 5797-5806.

Reinhart, B.J., Slack, F.J., Basson, M., Pasquinelli, A.E., Bettinger, J.C., Rougvie, A.E., Horvitz, H.R., Ruvkun, G., 2000. The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature 403, 901 - 906.

Richards, E.J., Ausubel, F.M., 1988. Isolation of a higher eukaryotic telomere from Arabidopsis thaliana. Cell 53, 127-136.

Riethman, H., 2008a. Human telomere structure and biology. Annual review of genomics and human genetics 9, 1-19.

Riethman, H., 2008b. Human subtelomeric copy number variations. Cytogenetic and genome research 123, 244-252.

Riethman, H., Ambrosini, A., Paul, S., 2005. Human subtelomere structure and variation. Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology 13, 505-515.

Riha, K., Mcknight, T.D., Griffing, L.R., Shippen, D.E., 2001. Living with genome instability: plant responses to telomere dysfunction. Science 291, 1797-1800.

Rinn, J.L., Kertesz, M., Wang, J.K., Squazzo, S.L., Xu, X., Brugmann, S.A., Goodnough, L.H., Helms, J.A., Farnham, P.J., Segal, E., Chang, H.Y., 2007. Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Noncoding RNAs. Cell 129, 1311-1323.

Rizzi, N., Denegri, M., Chiodi, I., Corioni, M., Valgardsdottir, R., Cobianchi, F., Riva, S., Biamonti, G., 2004. Transcriptional activation of a constitutive heterochromatic domain of the human genome in response to heat shock. Molecular biology of the cell 15, 543-551.

Rong, Y.S., 2008. Telomere capping in Drosophila: dealing with chromosome ends that most resemble DNA breaks. Chromosoma 117, 235-242.

Rusche, L.N., Kirchmaier, A.L., Rine, J., 2002. Ordered nucleation and spreading of silenced chromatin in Saccharomyces cerevisiae. Molecular biology of the cell 13, 2207-2222.

Saito, K., Sakaguchi, Y., Suzuki, T., Suzuki, T., Siomi, H., Siomi, M.C., 2007. Pimet, the Drosophila homolog of HEN1, mediates 2â€²-O-methylation of Piwi- interacting RNAs at their 3â€² ends. Genes & Development 21, 1603-1608.

Saito, K., Nishida, K.M., Mori, T., Kawamura, Y., Miyoshi, K., Nagami, T., Siomi, H., Siomi, M.C., 2006. Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome. Genes & Development 20, 2214-2222.

Sandell, L.L., Gottschling, D.E., Zakian, V.A., 1994. Transcription of a yeast telomere alleviates telomere position effect without affecting chromosome stability. Proc Natl Acad Sci U S A 91, 12061-12065.

Scheibe, M., Arnoult, N., Kappei, D., Buchholz, F., Decottignies, A., Butter, F., Mann, M., 2013. Quantitative interaction screen of telomeric repeat-containing RNA reveals novel TERRA regulators. Genome Research.

Scherf, A., Lopez-Rubio, J.J., Riviere, L., 2008a. Antigenic Variation in Plasmodium falciparum. Annual review of microbiology 62, 445-470.

Scherf, A., Lopez-Rubio, J.J., Riviere, L., 2008b. Antigenic variation in Plasmodium falciparum. Annual review of microbiology 62, 445-470.

Schoeftner, S., Blasco, M.A., 2008. Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. Nature cell biology 10, 228-236.

Schoeftner, S., Blasco, M.A., 2010. Chromatin regulation and non-coding RNAs at mammalian telomeres. Semin Cell Dev Biol 21, 186-193.

Schoeftner, S., Sengupta, A.K., Kubicek, S., Mechtler, K., Spahn, L., Koseki, H., Jenuwein, T., Wutz, A., 2006. Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing. The EMBO journal 25, 3110-3122.

Schotta, G., Sengupta, R., Kubicek, S., Malin, S., Kauer, M., Callen, E., Celeste, A., Pagani, M., Opravil, S., De La Rosa-Velazquez, I.A., Espejo, A., Bedford, M.T., Nussenzweig, A., Busslinger, M., Jenuwein, T., 2008. A chromatin-wide transition to H4K20 monomethylation impairs genome integrity and programmed DNA rearrangements in the mouse. Genes Dev 22, 2048-2061.

Shampay, J., Szostak, J.W., Blackburn, E.H., 1984. DNA sequences of telomeres maintained in yeast. Nature 310, 154-157.

Simon, M.D., Wang, C.I., Kharchenko, P.V., West, J.A., Chapman, B.A., Alekseyenko, A.A., Borowsky, M.L., Kuroda, M.I., Kingston, R.E., 2011. The genomic binding sites of a noncoding RNA. Proc Natl Acad Sci U S A 108, 20497-20502.

Simon, M.D., Pinter, S.F., Fang, R., Sarma, K., Rutenberg-Schoenberg, M., Bowman, S.K., Kesner, B.A., Maier, V.K., Kingston, R.E., Lee, J.T., 2013. High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation. Nature 504, 465-469.

Solovei, I., Gaginskaya, E.R., Macgregor, H.C., 1994. The arrangement and transcription of telomere DNA sequences at the ends of lampbrush chromosomes of birds. Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology 2, 460-470.

Storz, G., Altuvia, S., Wassarman, K.M., 2005. AN ABUNDANCE OF RNA REGULATORS*. Annual Review of Biochemistry 74, 199-217.

Su, X.-Z., Ferdig, M.T., Huang, Y., Huynh, C.Q., Liu, A., You, J., Wootton, J.C., Wellems, T.E., 1999. A Genetic Map and Recombination Parameters of the Human Malaria Parasite Plasmodium falciparum. Science 286, 1351-1353.

Surralles, J., Hande, M.P., Marcos, R., Lansdorp, P.M., 1999. Accelerated telomere shortening in the human inactive X chromosome. American journal of human genetics 65, 1617-1622.

Szostak, J.W., Blackburn, E.H., 1982. Cloning yeast telomeres on linear plasmid vectors. Cell 29, 245-255.

Tang, G., Reinhart, B.J., Bartel, D.P., Zamore, P.D., 2003. A biochemical framework for RNA silencing in plants. Genes & Development 17, 49-63.

Teixeira, M.T., Arneric, M., Sperisen, P., Lingner, J., 2004. Telomere length homeostasis is achieved via a switch between telomerase- extendible and -nonextendible states. Cell 117, 323-335.

Telser, J., Cruickshank, K.A., Morrison, L.E., Netzel, T.L., 1989. Synthesis and Characterization of DNA Oligomers and Duplexes Containing Covalently Attached Molecular Labels - Comparison of Biotin, Fluorescein, and Pyrene Labels by Thermodynamic and Optical Spectroscopic Measurements. J Am Chem Soc 111, 6966-6976.

Thomson, T., Lin, H., 2009. The Biogenesis and Function of PIWI Proteins and piRNAs: Progress and Prospect. Annual Review of Cell and Developmental Biology 25, 355-376.

Tiffany Hung, Y.W., Michael F Lin,Ashley K Koegel,Yojiro Kotake, Gavin D Grant, Hugo M Horlings, Nilay Shah, Christopher Umbricht, Pei Wang, Yu Wang, Benjamin Kong, Anita Langerød, Anne-Lise Børresen-Dale, Seung K Kim, Marc Van De Vijver, Saraswati Sukumar, Michael L Whitfield, Manolis Kellis, Yue Xiong, David J Wong & Howard Y Chang, 2011. Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. Nat Genet 43, 9.

Tyagi, A.R.a.S., 2010. Detection of Individual Endogenous RNA Transcripts *In Situ* using Multiple Singly Labeled Probes, *Methods in Enzymology*. Elsevier Inc., pp. 365-386.

Ucar, A., Vafaizadeh, Vida, Jarry, Hubertus, Fiedler, Jan, Klemmt, Petra a B, Thum, Thomas, Groner, Bernd, Chowdhury, Kamal, 2010. miR-212 and miR-132 are required for epithelial stromal interactions necessary for mouse mammary gland development. Nat Genet 42, 8.

Ugarkovic, D., 2011. Long non-coding RNAs. Springer.

Vagin, V.V., Sigova, A., Li, C., Seitz, H., Gvozdev, V., Zamore, P.D., 2006. A Distinct Small RNA Pathway Silences Selfish Genetic Elements in the Germline. Science 313, 320-324.

Valencia-Sanchez, M.A., Liu, J., Hannon, G.J., Parker, R., 2006. Control of translation and mRNA degradation by miRNAs and siRNAs. Genes & Development 20, 515-524.

Van Bakel, H., Nislow, C., Blencowe, B.J., Hughes, T.R., 2010. Most "dark matter" transcripts are associated with known genes. PLoS biology 8, e1000371.

Vargas, D.Y., Raj, A., Marras, S.A., Kramer, F.R., Tyagi, S., 2005. Mechanism of mRNA transport in the nucleus. Proc Natl Acad Sci U S A 102, 17008-17013.

Vega, L.R., Mateyak, M.K., Zakian, V.A., 2003. Getting to the end: telomerase access in yeast and humans. Nat Rev Mol Cell Biol 4, 948-959.

Vrbsky, J., Akimcheva, S., Watson, J.M., Turner, T.L., Daxinger, L., Vyskot, B., Aufsatz, W., Riha, K., 2010. siRNA-mediated methylation of Arabidopsis telomeres. PLoS genetics 6, e1000986.

Walder, R.Y., Walder, J.A., 1988. Role of RNase H in hybrid-arrested translation by antisense oligonucleotides. Proc Natl Acad Sci U S A 85, 5011-5015.

Wang, F., Podell, E.R., Zaug, A.J., Yang, Y., Baciu, P., Cech, T.R., Lei, M., 2007. The POT1-TPP1 telomere complex is a telomerase processivity factor. Nature 445, 506-510.

Wang, Kevin c., Chang, Howard y., 2011. Molecular Mechanisms of Long Noncoding RNAs. Molecular Cell 43, 904-914.

Wang, S.S., Zakian, V.A., 1990. Sequencing of Saccharomyces telomeres cloned using T4 DNA polymerase reveals two domains. Mol Cell Biol 10, 4415-4419.

Wapinski, O., Chang, H.Y., 2011. Long noncoding RNAs and human disease. Trends in cell biology 21, 354-361.

Watson, J.D., 1972. Origin of concatemeric T7 DNA. Nature: New biology 239, 197-201.

White, R.J., 2011. Transcription by RNA polymerase III: more complex than we thought. Nature reviews. Genetics 12, 5.

Wightman, B., Ha, I., Ruvkun, G., 1993. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell 75, 855-862.

Willingham, A.T., Orth, A.P., Batalov, S., Peters, E.C., Wen, B.G., Aza-Blanc, P., Hogenesch, J.B., Schultz, P.G., 2005. A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. Science 309, 1570-1573.

Willis, M.C., Hicke, B.J., Uhlenbeck, O.C., Cech, T.R., Koch, T.H., 1993. Photocrosslinking of 5-iodouracil-substituted RNA and DNA to proteins. Science 262, 1255-1257.

Wilusz, J.E., Sunwoo, H., Spector, D.L., 2009. Long noncoding RNAs: functional surprises from the RNA world. Genes Dev 23, 1494-1504.

Wilusz, J.E., Jnbaptiste, C.K., Lu, L.Y., Kuhn, C.-D., Joshua-Tor, L., Sharp, P.A., 2012. A triple helix stabilizes the 3′ ends of long noncoding RNAs that lack poly (A) tails. Genes & development 26, 2392-2407.

Wu, J., 2006. Comparative sequence analysis reveals an intricate network among REST, CREB and miRNA in mediating neuronal gene expression. GenomeBiology.com 7, R85.

Wu, L., Multani, A.S., He, H., Cosme-Blanco, W., Deng, Y., Deng, J.M., Bachilo, O., Pathak, S., Tahara, H., Bailey, S.M., Deng, Y., Behringer, R.R., Chang, S., 2006. Pot1 deficiency initiates DNA damage checkpoint activation and aberrant homologous recombination at telomeres. Cell 126, 49-62.

Wu, L., Zhou, Huanyu, Zhang, Qingqing, Zhang, Jianguang, Ni, Fangrui, Liu, Chang, Qi, Yijun, 2010. DNA Methylation Mediated by a MicroRNA Pathway. Molecular Cell 38, 465-475.

Wutz, A., 2011. Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. Nature reviews. Genetics 12, 12.

Wynes, M.W., Hinz, T.K., Gao, D., Martini, M., Marek, L.A., Ware, K.E., Edwards, M.G., Bohm, D., Perner, S., Helfrich, B.A., Dziadziuszko, R., Jassem, J., Wojtylak, S., Sejda, A., Gozgit, J.M., Bunn, P.A., Jr., Camidge, D.R., Tan, A.C., Hirsch, F.R., Heasley, L.E., 2014. FGFR1 mRNA and Protein Expression, not Gene Copy Number, Predict FGFR TKI Sensitivity across All Lung Cancer Histologies. Clin Cancer Res 20, 3299-3309.

Xu, M., You, Y., Hunsicker, P., Hori, T., Small, C., Griswold, M.D., Hecht, N.B., 2008a. Mice Deficient for a Small Cluster of Piwi-Interacting RNAs Implicate Piwi-Interacting RNAs in Transposon Control. Biology of reproduction 79, 51-57.

Xu, Y., Kimura, T., Komiyama, M., 2008b. Human telomere RNA and DNA form an intermolecular G-quadruplex. Nucleic acids symposium series (2004), 169-170.

Xu, Y., Suzuki, Y., Ito, K., Komiyama, M., 2010. Telomeric repeat-containing RNA structure in living cells. Proc Natl Acad Sci U S A 107, 14579-14584.

Yan, Z., Hu, H.Y., Jiang, X., Maierhofer, V., Neb, E., He, L., Hu, Y., Hu, H., Li, N., Chen, W., Khaitovich, P., 2011. Widespread expression of piRNA-like molecules in somatic tissues. Nucleic acids research 39, 6596-6607.

Yao, M.C., Blackburn, E., Gall, J., 1981. Tandemly repeated C-C-C-C-A-A hexanucleotide of Tetrahymena rDNA is present elsewhere in the genome and may be related to the alteration of the somatic genome. The Journal of cell biology 90, 515-520.

Ye, J.Z., De Lange, T., 2004. TIN2 is a tankyrase 1 PARP modulator in the TRF1 telomere length control complex. Nat Genet 36, 618-623.

Ye, J.Z., Hockemeyer, D., Krutchinsky, A.N., Loayza, D., Hooper, S.M., Chait, B.T., De Lange, T., 2004. POT1-interacting protein PIP1: a telomere length regulator that recruits POT1 to the TIN2/TRF1 complex. Genes Dev 18, 1649-1654.

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

Young, J.C., Agashe, V.R., Siegers, K., Hartl, F.U., 2004. Pathways of chaperone-mediated protein folding in the cytosol. Nat Rev Mol Cell Biol 5, 781-791.

Zamore, P.D., Haley, B., 2005. Ribo-gnome: The Big World of Small RNAs. Science 309, 1519-1524.

Zhang, L.F., Ogawa, Y., Ahn, J.Y., Namekawa, S.H., Silva, S.S., Lee, J.T., 2009. Telomeric RNAs mark sex chromosomes in stem cells. Genetics 182, 685-698.

Zhang, Y., 2005. miRU: an automated plant miRNA target prediction server. Nucl. Acids Res. 33, W701-704.

10. LIST OF PUBLICATIONS

1 X chromosome inactivation: a silence that needs to be broken. Basu R, Zhang LF. Genesis. 2011, Nov; 49(11):821-34. doi: 10.1002/dvg.20792. Epub 2011 Oct 13.

2 Using amino-labeled nucleotide probes for simultaneous single molecule RNA-DNA FISH. **Basu R**., Lai L.T., Meng Z., Wu J., Shao F. & Zhang LF. PloS one, 2014. 9(9): p. e107425.