CHARACTERIZATION OF *ARABIDOPSIS* TELOMERASE NEGATIVE REGULATOR, A TE-CONTAINING LNCRNA TER2: DEFINING ITS ROLE IN AND BEYOND TELOMERE BIOLOGY

A Dissertation

by

HENGYI XU

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Dorothy Shippen
Committee Members,	Pingwei Li
	Xiuren Zhang
	Matthew Sachs
Head of Department,	Gregory Reinhart

May 2015

Major Subject: Biochemistry

ABSTRACT

In eukaryotes the ends of chromosomes are constituted by nucleoprotein complexes termed telomeres. Telomeres represses a DNA damage response and, more importantly, facilitate the maintenance of the terminal DNA sequence by telomerase. Telomerase activity can be reconstituted by its two core subunits, the catalytic reverse transcriptase TERT and the telomerase RNA TER.

The Shippen lab developed Arabidopsis as a model for telomere studies. Previous work in Shippen lab showed that an alternative copy of telomerase RNA, TER2, serves as a negative regulator of the telomerase in response to DNA damage. In this study I characterized the evolution and function of TER2 and explored its biological significance.

TER2 possesses an intron and analysis of sequences from the 1001 genome project showed that the TER2 intron is derived from a transposable element (TE), specifically long terminal repeat (LTR) of a *Copia*-like retrotransposon. I verified that in most *A. thaliana* accessions the TER2 TE is intact, while in about 10% of accessions it is missing. The TE within TER2 destabilizes this RNA, enabling the plant to down regulate telomerase activity by modulating TER2 abundance. This RNA stability control mechanism contributes to the accumulation of TER2 after DNA damage, and thus links telomerase regulation directly to the DNA damage response.

My results also showed that TER2 is developmentally regulated, but only in accessions that contain the TER2 TE, suggesting the exaptation of the TE endows TER2 with a function in reproductive development. Indeed, plants lacking TER2 have reduced seed production efficiency. In addition, *ter2* mutants have lower pollen viability

than wild type, though not as low as in *tert* mutants. These results defined a novel function of TER2 in plant reproduction.

My work unexpectedly revealed that TER2 processing and/or stability is influenced by the small RNA processing pathway. In plants lacking Dicer-like2 (DCL2), TER2 abundance increases and the expression profile change during flower development. I further discovered that DCL2 affects TER2 in a post-transcriptional manner. Together, these data uncovered unexpected complexity of TER2 RNA processing and its regulation.

Finally, I found that TER1, the canonical TER in *A. thaliana*, has lessons to teach. Single nucleotide polymorphisms (SNPs) in TER1 telomere templating domain were found. I showed that the SNPs do not change the newly synthesized telomere repeats. This observation provided new insight into the mechanisms of template utilization and how this is evolving.

In summary, my research revealed evidence for evolution in two telomerase RNA genes in *A. thaliana,* and provided several novel insights into lncRNA structure, evolution and metabolism that impact telomerase regulation and benefit plant growth and reproduction.

DEDICATION

This dissertation is dedicated to my father Gang Xu and my mother Lan Wang, the most adorable parents in the world. As the only child in my family, I was taken care of with complete love from them. They are always behind me to provide unconditional support, especially, my mother, who made many sacrifices for the family to provide a better environment for my growth. She taught me integrity, braveness and kindness, and ultimately shaped me into who I am today.

ACKNOWLEDGEMENTS

As an amateur video game player, I have found similarities to this activity and the Ph.D. process. Both require high-level skills, and at the beginning a mentor would be very helpful. Many obstacles are encountered during the process. Therefore, endurance and insistence (which is the meaning of my Chinese given name) are essential to overcome the hard situations. Creativity becomes more important. You develop the necessary from uncountable practice, and finally by enriching knowledge and experience. Standing where I am today, I realize that having a Ph.D. is just the beginning of the whole game. It is not an end, but a start. Though achievements cannot be separated from my efforts, it is the help from many people that led me onto the path of success. I owe many thanks to a group of special people.

First among all of them is my PI, Dorothy Shippen. Dorothy is a determined force in the evolution of my becoming a scientist. As a mentor, she taught me how to think, design, and perform research in a professional and scientific way. She encouraged and supported freedom of thinking. When we discussed science, great ideas naturally sparked from the collision of our thoughts. However, she is more than a mentor, she takes care of me in the lab like a mother. She is willing to provide advice for my personal topics and to teach appreciation for the good things in life. Dr. Geoffrey Kapler, the husband of Dorothy, needs to have a special mention here. I gradually learned many philosophies when talking with him in both research and life. Getting to know Dorothy, and Jeff, as well as with their son Simon, has given rise to feelings of family, especially with lots of fun in parties at their house and in sports at gyms. Good memories of Dorothy and as a member in Shippen lab/family will be a great treasure for my entire life.

Next, I want to thank my committee member Drs. Pingwei Li, Matthew Sachs and Xiuren Zhang for providing scientific guidance during each committee meeting. Dr. Xiuren Zhang deserves special thanks for providing priceless research material and helpful advice. I also thank Dr. Geoffrey Kapler, Dr. Feng Qiao, and Dr. Xiuren Zhang and all their lab members for their guidance in our joint lab group meetings every Friday. Beyond useful advice in doing research, these meeting experiences made me confident in front of an audience when showing my research.

Working in Shippen lab for these past years, I have overlapped with many people and shared in their joys, tears and happiness. My “forever” mentor Dr. Andrew Nelson, who was my department assigned mentor (babysitter) at the very beginning, always took care of me as his little brother in lab. His extreme enthusiasm and curiosity in doing research, which is a signature of all Shippen lab members, provided me with a positive force and finally was one important factors in helping me to join in the Shippen lab as a permanent member. After joining the lab, I was lucky to have Dr. Catherine Cifuentes-Rojas as my project mentor. You can read two words in watching her performing experiment and working together with her: perfection and efficiency. Though I had a hard time at the beginning, I now fully appreciate how she transformed me into an aggressive and hardworking researcher.

I also thank older brothers and sisters in the Shippen lab for taking care of me: Dr. Xiangyu Song who helped me get settled in College Station in the very beginning; Dr. Kara Boltz, who was always patient and willing to answer my questions; Dr. Katherine Leehy, who was an amazing cake maker and lab “police-woman” (calling out my misbehaviors in lab, which made me a better lab citizen with carefulness and consideration); Dr. Jungro Lee, who always shared hardworking nights with me in the

lab; Dr. Eugene Shakirov, who always looked “scary”, but is actually very nice to help his Chinese comrade; and Dr. Mark Beilstein, who was a lovely joker and a very impressive swimmer (I believe I will beat him one day). Furthermore, I am grateful to lab members who came after me: Xiaoyuan Xie, for her inclusiveness and willingness to provide much help; Xintao She, for his intelligence in asking very tough questions; Callie Kobayashi, for her responsibilities to shoulder many lab duties and good memories in sharing sports and video games with her; Dr. Amit Arora, for his constant support in helping me solve technical problems and share of his reagents. A special thanks goes to my sweet undergraduate student, Jennifer Tonwley, for her hard work all these years.

I would like to thank all my friends at Texas A&M University for your many years of friendship. Fuqu Hu, Shanshan Yang, Cai Jiang and Cheng Cheng, we are the “forever group of five”. Next, my thanks are also given to all the staff in the Department of Biochemistry and Biophysics. Thank you all for the help in administration.

Finally, I need to give more than just thanks to my family. They are always behind me to encourage me. I cannot count how many frustrating times during these years. My parents comforted me like a baby and pulled me out from bad moods. At Chinese New Year, it is my grandparents, uncles, aunts, and cousins who reminded me that I have a very warm family. Though I am physically single, I am not mentally alone. I love you, all my family members. With your love, I am standing where I am now.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION AND BACKGROUND

The pursuit of knowledge regarding telomeres cannot be separated from an understanding of cell aging and chromosome behavior. The journey to uncover the myth of aging can be traced to ancient times. In Imperial China, a deep and abiding interest in discovering the secret of eternal life was cultivated among the imperial denizens. One fascinating myth tells of how the Empress of the Heaven had possession of an elixir, created in the alchemic furnace of the Eight Diagrams, which would bestow upon the fortunate recipient the gift of life eternal. This elixir was given to the heroic archer Yi, who had pierced the sky with his arrows to bring down nine suns from the heavens, leaving one to circulate the earth. Unfortunately, the elixir was stolen by his wife, Chang'e (the current namesake of China's lunar exploration program), who imbibed of the potion, becoming the embodiment of light and floating to the moon to take up residence. Out of longing for his beloved, Yi baked her favorite cake annually, giving rise to the tradition of the Chinese mid-autumn festival and the eating of moon cakes. Later, the first Chinese emperor, Qin Shi Huang, sent the ancient hero Xufu, along with several virginal boys and girls chosen from the kingdom, to retrieve the elixir. Sailing eastward, which Xufu believed to be the path toward Heaven, Xu and his comrades drifted aimlessly for years before finally landing on an island which he named Danzhou. It was on this island that Xufu decided to establish his own prosperous kingdom. Giving up on completing his quest to retrieve the elixir, Xufu and his child comrades, aided by advanced technologies from glorious China, gave rise to a most powerful kingdom now believed to be the country of Japan. As the ancient chemistry

“science” continued to develop in China, it encouraged most if not all successive emperors to follow in the footsteps of Qin Shi Huang in searching for the elixir of immortality. Using heavy metals such as liquid mercury, alchemists and proto-scientists worked tirelessly to solve the problem of aging. Unfortunately, all experiments failed, a sad truth corroborated by the short lives of Chinese emperors as recorded in ancient history. Though the elixir of immortality was never discovered by the Chinese, the pursuit of this elusive tonic ultimately shaped Chinese history, culture, and politics.

The question of whether organisms could be immortal also captured the attention of modern scientists. It is widely accepted that cells are the basic unit for modern biology. Thus, the pursuit of immortality has been transformed to ask if cellular life span could be eternal. From work undertaken in 1912, Alexis Carrel claimed that cells could continue to grow indefinitely because a fibroblast cell culture could be maintained for over 20 years (Carrel and Ebeling, 1921). People were willing to accept this exciting and fascinating hypothesis of infinite cell growth potential in the early 20th century, although the result was never reproducible. In the 1960s Hayflick and Moore challenged the cell immortality theory by showing a finite cell replication capacity that correlated with cell senescence. These scientists found that fibroblast cells from human skin only undergo 40 to 50 cell divisions before senescence (Hayflick, 1961). Skin cells from younger people could undergo more cell divisions than the cells from older people, indicating that it is the total cell divisions instead of the division time in cell culture that mattered (Hayflick, 1965). This phenomenon is famously known as the “ Hayflick limit”. It seemed that cells contain an intrinsic, but mysterious mechanism that counts the number of times a cell divides. This model raised the question of how the cell defined the limit.

Many excellent scientific discoveries revealed insight into this enigma. Discovery of the double helix structure of DNA indicated a plausible mechanism for its replication (Watson and Crick, 1953). Once the mechanism of DNA replication was proposed, an unavoidable problem, the so-called “end-replication problem”, was uncovered by the Russian theoretical biologist Alexey Olovnikov (Olovnikov, 1971). The conventional DNA polymerases synthesize DNA using the template strand and a preexisting RNA primer with a free 3’ hydroxyl –OH. Therefore, the DNA replication direction is always constrained as 5’ to 3’, causing one strand of synthesis to be continuous (leading strand) and the other to be discontinuous (lagging strand). For lagging strand synthesis, the gap left once the RNA primer is removed is filled in to connect Okazaki fragments. However, when the last RNA primer at the 5’ end of the daughter strand replicated by lagging strand machinery is removed, this gap can never be filled in due to the lack of the 3’-OH at the very end. Thus, a bit of DNA sequence is left unreplicated in each cell cycle. The end replication problem is believed to be the underlying driver for formation of the concatamer genome before replication in phage, thereby minimizing impact from the distal DNA loss by reducing the actual ends (Watson, 1972). Though few details were known about the link between the end replication problem and cell immortality at that time, the end replication problem seemed to explain the “Hayflick limit”: if there is no proper way to replenish the end DNA sequence lost after each round replication, the chromosome becomes shortened, resulting in the loss of genetic information. Olovnikov predicted that the loss of the distal genetic information could serve as a counting mechanism that could limit the cell division. Olovnikov also predicted the existence of an unknown mechanism for replenishing the lost distal DNA to maintain genome integrity. The end of the chromosome must be somehow distinct (Figure 1-1)

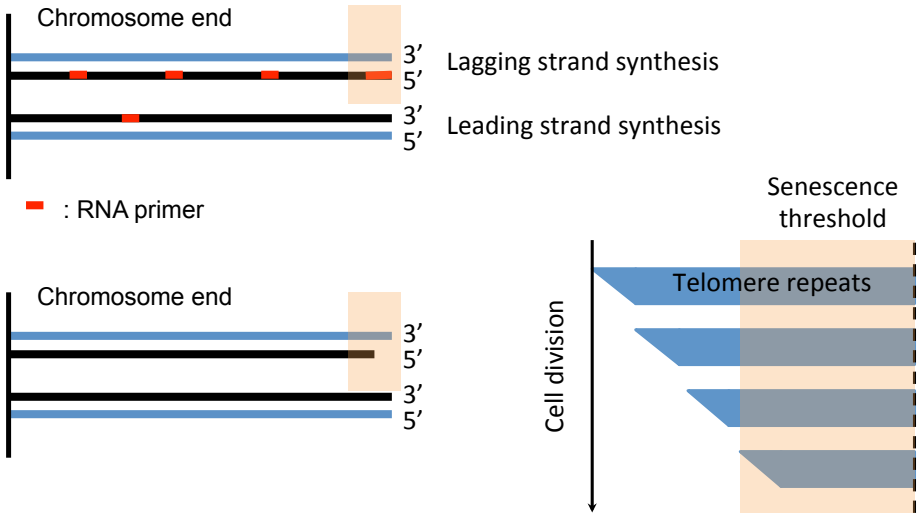


Figure 1-1. End replication problem and senescence. After DNA replication, the newly synthesized DNA (Black line), gapped with the RNA primer (Red line) are indicated. After the ligation of the Okazaki fragments generated by lagging strand synthesis, the 5’ most RNA primer cannot be replaced with DNA, leaving a 3’ overhang structure at the very end of the chromosome. This problem is known as the end replication problem. Without proper replenishment of the chromosome end, the DNA will be shorter after each round of replication. Once the telomere reaches a critical length threshold, the cell stops division and undergoes senescence.

Important conceptual breakthroughs are seldom verified from one discovery. In addition to the cellular and biochemical evidence, results of genetic and biochemical studies supported the idea of a distinct structure at the chromosome end. By using X-ray exposure on *Drosophila* chromosomes, Herman Muller observed several types of chromosome rearrangements. He could not directly see the terminal deletion of the chromosome, but he concluded that the end of chromosome contains a unique structure for stability (Muller coined the term telomere). Correspondingly, genetic studies performed in maize by Barbara McClintock showed that chromosome fusion

occurred at sites of DNA double strand breaks (DSB) (McClintock, 1938), again indicating the chromosome termini are important for genome stability. Another important finding by McClintock was the observation that broken chromosome ends did not fuse in embryos, but were instead “healed” (McClintock, 1941). All of these exciting findings prompted intensive research to characterize chromosome ends in the following decades, opening a new field known as telomere biology (Greek: telo means end and mere means part).

Discovery of telomere structure and telomerase

In 1978 Elizabeth Blackburn worked as a postdoctoral fellow in Joseph Gall’s lab studying the extrachromosomal genes coding for rDNA in *Tetrahymena thermophila*. She successfully sequenced the terminal of rDNA, and discovered that the end of the chromosomes contained a tandem repetitive CCCCAA/GGGGTT sequence ranging from 100 to 400 nucleotides (Blackburn and Gall, 1978). Though the function of this repeat DNA sequence was not understood at that time, similar repeat sequences were found in other organisms. Blackburn proposed that these sequences would be important for rDNA replication.

Blackburn moved to the University of California at Berkeley and collaborated with Jack Szostak to continue her telomere studies. They transformed budding yeast with a plasmid containing the *Tetrahymena* terminal repeat sequence and showed that the artificial mini chromosome was stable in yeast (Szostak and Blackburn, 1982). These results indicate that the telomere has a special structural feature that can stabilize the end of linear chromosomes and this function seems to be conserved in evolution. Blackburn also discovered that a different simple repeat sequence was added

to the *Tetrahymena* telomere repeats in yeast. These new repeats were then identified as yeast telomere sequence, and this data suggested that an active process synthesized and maintained telomere repeats.

If these repetitive DNA sequences had an important function, how were they added to the chromosome ends? Carol Greider, who was a graduate student in Blackburn's lab, answered this question. In 1985, Greider and Blackburn discovered a novel enzyme activity from *Tetrahymena* cell free extracts that could extend the telomeric DNA sequence on a single stranded DNA oligonucleotide. This extension could only occur robustly on a TG rich *Tetrahymena* or yeast telomeric DNA primer, not on an oligonucleotide with C-rich telomere repeats or other non-specific DNA sequence (Greider and Blackburn, 1985). The enzymatic activity was sensitive to protease and heat treatment. Biochemical data showed that it extended oligonucleotide primers by adding one nucleotide at a time instead of preformed telomere repeats.

Two years later, Greider and Blackburn first proposed the term "telomerase" instead of telomere terminal transferase to describe this enzyme. What they found surprising was that telomerase was a ribonucleoprotein and the RNA moiety was required for its enzymatic activity (Greider and Blackburn, 1987). To have a better idea of the function of the RNA, they cloned the RNA from telomerase and found a telomere complementary sequence CAACCCCAA embedded in it, suggesting the telomerase RNA provide a template for *de novo* synthesis of the telomeric repeats (Greider and Blackburn, 1989). Follow up studies demonstrated that telomerase is a specialized reverse transcriptase with a functional RNA template required for its activity (Shippen-Lentz and Blackburn, 1990; Yu et al., 1990).

These discoveries dramatically expanded the foundation of the telomere biology field. In addition, the knowledge acquired from telomere studies resulted in a greater understanding of the basic science in DNA replication (and as discussed below), cell proliferation and genome stability maintenance. Due to their contribution to the telomere field, Drs. Blackburn, Greider and Szostak were awarded the Nobel Prize in Physiology and Medicine in 2009.

Telomeric DNA and its structure

As mentioned above, the DNA component of the telomere is composed of simple G-rich repetitive sequences. These sequences vary slightly among different organisms, from TTTTGGGG in ciliates (Pluta et al., 1982) to TTAGGG in vertebrates (Moyzis et al., 1988). In plants, this sequence is TTTAGGG. This similarity of telomere sequences suggests that the telomeres and their synthesis by telomerase represent a conserved mechanism for genome maintenance, which arose early in evolution.

Probably as an adaptation to the challenges of different cellular environments, telomere length varies greatly among different organisms, from hundreds of base pairs in yeast (Shampay and Blackburn, 1988) to hundreds of kilobases (kb) in tobacco (Murti and Prescott, 2002) and mice (Kipling and Cooke, 1990). In addition to the average length, the range of length also varies, from 2-7kb in *Arabidopsis* (Shakirov and Shippen, 2004) to 2-30kb shown in humans (Zellinger and Riha, 2007) (Meyne et al., 1989). Therefore, different mechanisms evolved and adapted as solutions to satisfy the requirement of chromosome end maintenance. When telomeres are short enough to reach a critical length threshold, they will lose the potential to protect the chromosome end and trigger a DNA damage response and cell cycle arrest (Riha et al., 2006). On

the other end of the spectrum, telomeres that are too long also cause a defect in cell growth (McEachern and Blackburn, 1995), suggesting that telomere length homeostasis is achieved by a finely regulated balance between telomere shortening and lengthening.

The G-overhang

Telomeric DNA also has a distinct structural feature at the very end. The functional telomere cannot be separated from these specific structures, which contribute to its homeostasis. The extreme end of the telomere consists of a single strand protrusion of the G-rich strand, which is named the G-overhang. A 3' overhang is not surprising because it is a natural product from removal of the RNA primer after lagging strand replication (Lingner and Cech, 1998). The G-overhang was first characterized in ciliates in 1981 by analysis of the 3' ends of macronuclear chromosomes (Klobutcher et al., 1981). The G-overhang was later found in many other species including vertebrates, yeast and plants by non-denaturing in-gel hybridization (Wellinger et al., 1993; Makarov et al., 1997; Venkatesan and Price, 1998; Riha et al., 2000). The G-overhang is essential as a substrate for telomerase extension. This single strand DNA can also be inserted into the homologous double strand region by folding back, resulting in the formation of a lasso structure called the T-loop (discussed below). Lastly, the G-overhang can serve as a scaffold for single stand telomere binding proteins, which facilitate the protection of the chromosome end and telomerase recruitment.

Since the G-overhang has multiple functions, several factors are expected to affect the regulation of G-overhang *in vivo.* As mentioned above, the G-overhang is a product of lagging strand synthesis. However, in most organisms G-overhangs are formed on both sides of chromosomes, though the side replicated by leading strand synthesis produces

a blunt end. The blunt end is converted intro 3' overhang by nuclease resection of the C-strand after leading strand synthesis (Wu et al., 2010). The G-overhang is perturbed in mutants that lack single strand telomere binding proteins such as POT1 in mammals and STN1 in plants (Wu et al., 2006; Song et al., 2008). The length of the G-overhang is dependent on the cell cycle. In budding yeast, longer G-overhangs are generated in late S phase mediated by CDK1 activity (Wellinger et al., 1993; Vodenicharov and Wellinger, 2006). CDK1 activity is proposed to control the timing of telomerase extension (Frank et al., 2006): longer G-overhangs are better telomerase primers. Moreover, DNA damage signaling factors such as ATM, ATR and MRN complex (Chai et al., 2006; Denchi and de Lange, 2007; Dimitrova and de Lange, 2009) affect G-overhang length, suggesting a correlation between G-overhang regulation and DNA damage signaling. Hence, regulation of G-overhang length is required for telomerase action and to prevent the telomere from being recognized a DNA double strand break. Homeostasis of the G-overhang length reflects a timely compromise made by different rounds of negotiations among a complicated network of players at different stages of the cell cycle.

The T-loop

As mentioned earlier, insertion of the G-overhang into the double strand telomere region generates a higher order architecture called the t-loop, which was first found in mammals (Griffith et al., 1999). Later, the t-loop was observed in many organisms including trypanosomes, ciliates and plants (Munoz-Jordan et al., 2001; Murti and Prescott, 2002; Zellinger et al., 2007). The t-loop is believed to stabilize the telomere by preventing it from mistakenly being recognized as damaged DNA. Folding

back of the telomere was found even earlier in 1997 in studies of yeast telomeres (Grunstein, 1997), although unlike the t-loop which has the 3' end invade the duplet, the fold-back structure at yeast telomerase is thought to be stabilized by protein-protein interaction (Grunstein, 1997). In yeast, the telomere binding protein Rap1 facilitates the maintenance of the folding back structure (Grunstein, 1997).

The mechanism of strand invasion by the G-overhang to form the t-loop is not well understood. Interestingly, in mammals, TRF2, is involved in the assembly of t-loops (Stansel et al., 2001). TRF2's role in t-loop maintenance was further verified by *in vivo* observation using high-resolution microscopy. These studies showed a lack of t-loops in cell lines deficient in TRF2, but not other telomere binding proteins (Doksani et al., 2013).

Blunt end telomeres

Leading strand DNA synthesis will give birth to blunt end daughter DNA, which must be further processed to generate a G-overhang. In mammals, G-overhang maturation is regulated during the cell cycle, to prevent the inappropriate activation of DSB signaling (Chow et al., 2012). Interestingly, the maintenance of double strand telomeric DNA is dependent on the Ku heterodimer, which is also involved in the non-homologous end joining (NHEJ) double strand break (DSB) repair pathway. The ring structure formed by Ku70-Ku80 encloses the DNA end to facilitate telomere maintenance and DNA break repair (DuBois et al., 2002). Thus, it is unclear how Ku complex distinguishes the telomere ends from DSBs, enabling cells to achieve telomere capping as well as DSB DNA repair. One possibility is through post-translational modification of Ku. The yeast Ku70 (yKu70) C-terminus is sumolyated; mutations

abolishing sumolytion change its DNA binding ability and cause telomere shortening and increased G-overhangs (Hang et al., 2014). Similarly, in mammals, KU80 is ubiquitinated leading to its degradation (Feng and Chen, 2012). Degradation of Ku acts as a switch for dynamic metabolism of telomere ends during the cell cycle. Therefore, interaction of blunt ended telomeres with Ku may have evolved to stabilize telomeres during cell-cycle dependent G-overhang processing to prevent activation of DNA repair machinery at telomeres.

Recent research in *Arabidopsis thaliana* indicates that not all chromosome ends terminate with a 3' single strand overhang. Instead, half of the telomeres have blunt ends and the other half have a 3' G-overhang (Kazda et al., 2012). Interestingly, Ku stabilizes the blunt end telomere in *A. thaliana,* and thus Ku may serve as a switch between G-overhangs and blunt ended telomeres. Although it is still unknown whether blunt ended telomeres exist in other organisms besides plants, the observation in *A. thaliana* provides a novel mechanism for protecting chromosome ends independent of G-overhangs (Figure 1-2).

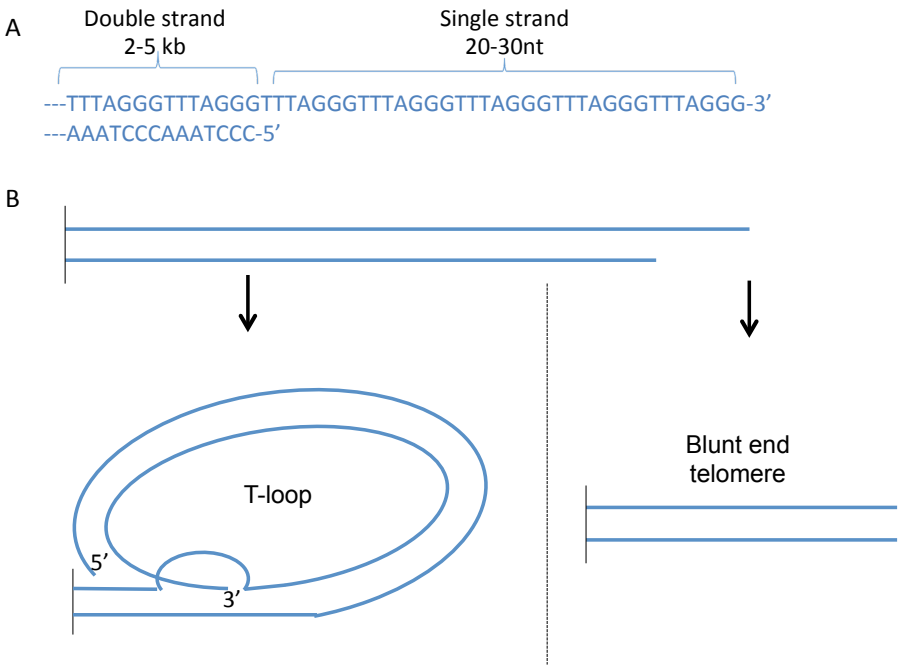


Figure 1-2. Telomere DNA structures in *Arabidopsis*. (A) Telomeres are G-rich repetitive sequences at the chromosome end. In plants, the sequence is TTTAGGG, highly similar to its counterparts in vertebrates, TTAGGG. Telomeres are composed of both double strand (ds) and single strand (ss) regions. The G-rich 3' extension is also named the G-overhang. In *Arabidopsis,* the length of telomere varies and ranges from 2-7kb for ds (Shakirov and Shippen, 2004) with 20-30nt for the G-overhang (Riha et al., 2000). (B) The G-overhang tucks back into the duplet region to form a structure called T-loop. Recent findings suggested the telomere is intrinsically asymmetrical in *A. thaliana* with one end having a G-overhang and the other a blunt end.

Telomere-associated proteins

Telomeric DNA is coated with proteins in both single stranded (ss) and double stranded (ds) regions. The nucleoprotein structures that cap telomeres are believed to protect telomeres from exonucleolytic attack and repress the activation of the DNA damage response. Furthermore, these structures must facilitate the regulation of telomerase access when needed. There are two main classes of telomere binding proteins: the ss binding proteins and the ds binding proteins. The ss group, includes TEBP in ciliates (Gottschling and Zakian, 1986). Cdc13 in budding yeast (Garvik et al., 1995), and POT1 in mammals (Baumann and Cech, 2001), all contain a signature OB fold domain (Oligosacchride/oligonucleotide binding). The OB fold motif makes a five-stranded beta barrel structure used for recognizing nucleic acids as well as proteins (Theobald and Wuttke, 2004). The ds telomere binding proteins include Rap1 in budding yeast (Longtine et al., 1989), TRF1/2 in mammals (van Steensel and de Lange, 1997) and Taz1 in fission yeast (Cooper et al., 1997), and these proteins interact with telomeres via a helix-turn-helix containing motif called the Myb domain. Protection of telomeres is mediated by complexes containing both ss and ds telomere binding proteins. The two major telomere capping complexes are shelterin and CST (CTC1/Cdc13; STN1; TEN1) (See below).

The CST complex

In budding yeast, the ss telomeric DNA is protected by Cdc13-Stn1-Ten1 (CST) (Giraud-Panis et al., 2010). Within this complex, Cdc13 is the main ss binding protein. The N-terminus of this 100kd protein is functional for DNA binding, and multiple OB folds have been found within it. These OB folds are necessary for binding an 11nt core

sequence corresponding to the G-overhang (Mitton-Fry et al., 2002; Zappulla et al., 2009). Cdc13 has dual functions in regulating both telomere protection and telomerase accessibility (Nugent et al., 1996). Temperature sensitive *cdc13* mutants arrested the cell cycle (Garvik et al., 1995) as a result of the inability to repress DDR activation (Lin and Zakian, 1996). A screen for suppressors of *cdc13* mutants uncovered Stn1 as a Cdc13 cofactor required for telomere maintenance (Grandin et al., 1997). Later, Ten1 was found by screening for suppressors of *stn1* mutants (Grandin et al., 2001). Over expression of Stn1 in *cdc13* mutants increases cell viability (Grandin et al., 1997). Cell viability could further be increased by overexpressing Ten1 and Stn1 together (Grandin et al., 2001). The Cdc13-Stn1-Ten1 complex is referred to as CST. The absence of any component of CST is lethal, suggesting the three proteins work in the same genetic pathway.

Another important function of CST is telomerase recruitment to chromosome ends. This process is associated with Cdc13. Cdc13 interact directly with Est1 protein (Evans and Lundblad, 1999), an accessory protein of yeast telomerase (Lundblad and Blackburn, 1990). Deletion of Est1 causes an ever-shorter telomere phenotype (Lundblad and Szostak, 1989). However, fusing the Cdc13 to the telomerase catalytic subunit EST2 (TERT) bypass the requirement of EST1 for telomere maintenance, suggesting that EST1 in budding yeast serves as a recruitment factor (Taggart et al., 2002). This mechanism of telomerase recruitment was verified by telomerase delivery to a DNA ds breakage site when Cdc13 and EST1 were tethered adjacent to the DSB sites (Bianchi et al., 2004). When the genome is intact, the association of Cdc13 and EST1 with telomeres occurs in a cell cycle dependent manner (Taggart et al., 2002).

A functional homolog of the CST complex was recently identified in vertebrates (Miyake et al., 2009)(see below) and plants (Surovtseva et al., 2009). In plants, CST appears to be the major end protection complex. A Stn1 homolog was found in *A. thaliana* using *Schizosaccharomyces pombe* (fission yeast) STN1 as the query for a BLAST search (Song et al., 2008). Plants lacking STN1 display continuous telomere shortening, an increase in G-overhangs, and chromosome fusions (Song et al., 2008). Soon after, a screen for POT1c mutants in EMS mutagenized lines serendipitously identified the *CTC1* gene, which is a functional homolog of Cdc13 in yeast (Surovtseva et al., 2009). The *ctc1* mutant had an obvious defect in telomere capping, similar to *stn1* mutants. Furthermore, CTC1 interacted with STN1 both genetically and physically suggesting they may function together as a complex for telomere maintenance.

The idea that the CTC1-STN1-TEN1 complex always works as a unit has been challenged by characterization of TEN1 in plants. Like *ctc1* and *stn1* mutants, *ten1* mutants have a similar telomere shortening phenotype and end-to-end fusion (Leehy et al., 2013). However, in contrast to CTC1 and STN1, a smaller fraction of telomeres are bound by TEN1 (Leehy et al., 2013). Furthermore, TEN1 is a negative regulator of telomerase activity. Plants lacking TEN1 protein have higher telomerase processivity. Recent results indicate that TEN1 may have a chaperone activity that function to stabilize CTC1 both *in vitro* and *in vivo* (Jungro Lee, unpublished data). As in budding yeast, one component of Arabidopsis CST, CTC1, interacts with a component of telomerase, in this case, POT1a (Renfrew et al., 2014). Furthermore, as in yeast, the binding of the telomerase component appears to display one or more CST. In Arabidopsis, STN1, CTC1 and POT1a can form a tertiary complex, while TEN1 interaction with STN1 is mutually exclusive of POT1a-STN1 binding (Renfrew et al.,

2014). Therefore, an extendible telomere state is established by the formation of a CTC1-STN1-POT1a (CSP complex)-TER1 telomerase RNP, and this can be switched to an unextendible status by the competitive formation of CTC1-STN1-TEN1 (CST complex) (Renfrew et al., 2014). The evolution of CST-CSP gave rise to a binary switch in plants for both telomere capping and telomerase regulation, which contributes to the dynamic telomere end metabolism.

The shelterin complex in vertebrates and fission yeast

Unlike CST in budding yeast and plants, shelterin appears to be the major telomere binding complex in mammals and fission yeast. Human shelterin contains six components: TRF1 and TRF2 for ds binding, POT1 and TPP1 for ss binding, and RAP1 and TIN2 for bridging the ss and ds regions (de Lange, 2010). Shelterin is essential for t-loop formation, telomere length regulation and chromosome end protection (de Lange, 2005).

TRF1 was the first ds telomere protein identified. It was isolated from a nuclear protein extract that had ds telomeric DNA binding activity (Zhong et al., 1992). Later, a search for proteins showing similarity to TRF1 gave rise to TRF2 (Bilaud et al., 1996; Broccoli et al., 1997b; Broccoli et al., 1997a). Both TRF1 and TRF2 contain myb domains for ds telomeric DNA binding. Another conserved domain at the N-terminus of TRF1 and TRF2 allows these proteins to form homodimers (Bianchi et al., 1997). Both TRF1 and TRF2 negatively regulate of telomere length (Smogorzewska et al., 2000; van Steensel and de Lange, 1997). Although TRF1 and TRF2 can both alter the secondary structure of telomeric DNA, only TRF2 promotes the formation of the t-loop *in vitro* (Stansel et al., 2001) . Further biochemical studies showed that TRF2 generates

positive supercoiling and condenses DNA to provide a topological activity for G-overhang invasion into double strand DNA (Amiard et al., 2007).

POT1 is the shelterin component that binds ss telomeric DNA. The identification of fission yeast Pot1 led to the discovery of human POT1 (Baumann and Cech, 2001). This protein shows similarity to the a-subunit of the telomere end binding protein (TEBP) from ciliates. Unlike humans and fission yeast, mice contain two POT1 genes, mPOT1a and mPOT1b (Hockemeyer et al., 2006). Double knock-out cells showed DNA damage response activation and/or aberrant recombination at telomeres. Depletion of POT1 in humans causes loss of G-overhangs, chromosomal fusions and cell death (Wu et al., 2006). Mutations in POT1 are associated with telomere shortening and have been shown to cause chronic lymphocytic leukemia (Ramsay et al., 2013) and melanoma (Shi et al., 2014), indicating that POT1 has a significant role in safeguarding genome integrity. POT1's binding partner TPP1 is another OB-fold containing protein and the homologue to the b-subunit of TEBP (Xin et al., 2007). Its interaction with POT1 is required for telomere end protection (Hockemeyer et al., 2007).

One of the hinge proteins within the shelterin complex, Rap1 regulates telomere length via binding to TRF2 (Li et al., 2000). Rap1 affects the localization of TRF2 at the telomere (Arat and Griffith, 2012). Loss of Rap1 leads to increased telomere recombination and instability. Unexpectedly, ChIP-seq results showed that Rap1 is enriched at extra-telomeric regions and in the vicinity of the coding regions, suggesting this gene like its distant ortholog in budding yeast may be involved in transcriptional regulation function beyond its role at telomeres (Martinez et al., 2010).

Another shelterin hinge protein is TIN2. TIN2 interacts with TPP1 (Takai et al., 2011), but was first identified as a TRF1 interacting protein (Kim et al., 1999). It was

initially believed that the TIN2-TPP1 interaction serves as a telomerase recruitment mechanism to telomeres in human cells (Abreu et al., 2010). However, TIN2 mutations causes a telomerase-independent telomere deregulation (Frescas and de Lange, 2014), and lead to the disorder Dyskeratosis congenita. Interestingly, the localization of TIN2 in mitochondria forges a link between telomere regulation and cellular metabolism (Chen et al., 2012b). Altogether, these findings suggest that the shelterin components are evolving in a distinct path that connects telomere maintenance to other cell behavior.

In *S. pombe* telomere capping is also accomplished by a shelterin-like complex composed of six proteins including Taz1, Rap1, Poz1, Tpz1 Pot1 and Ccq1. Among them, the myb domain containing protein Taz1 binds ds telomeric DNA, while the OB fold containing Pot1 is for ss telomere binding. Rap1, Poz1 and Tpz1 form a bridge (or hinge) to link the ds and ss binding proteins. The Ccq1 is proposed to act as a telomerase recruitment/activation in much the same way as TPP1 in humans (Harland et al., 2014; Moser et al., 2011).

In fission yeast lacking Taz1, telomeres becomes longer, suggesting Taz1 is a negative regulator of telomere length (Cooper et al., 1997). Loss of Taz1 also causes C-strand degradation and increased homologous recombination at telomeres (Ferreira and Cooper, 2001; Miller and Cooper, 2003). Taz1 was found to be cell-cycle regulated to enforce the timing of telomere synthesis (Dehe et al., 2012). Phenotypes similar to those in the *taz1* mutant were detected in yeast cells lacking Rap1 or Poz1 (Miller et al., 2005).

The *S. pombe* Pot1 protein has a complex function. In *pot1* mutants, telomeres immediately shorten, resulting in the deregulation of chromosome segregation and

genome instability (Baumann and Cech, 2001), and suggesting that Pot1 is a positive regulator of telomere length. However, when Pot1 binding to telomeres is reduced, telomeres lengthen, indicating that SpPot1 also plays a negative regulatory role in telomere length (Bunch et al., 2005). Ccq1 is also a positive regulator of telomere length maintenance, but this effect is more likely via its interaction with telomerase. Ccq1 is believed to be a recruiter of telomerase because it interacts with Est1, which is a telomerase RNA accessory protein in *S. pombe* as it is in budding yeast (Webb and Zakian, 2012). Phosphorylation of Ccq1 by Tel1/Rad3 (ATM/ATR) may strengthen this recruitment effect (Moser et al., 2011; Yamazaki et al., 2012).

Lastly, Tpz1 lies in the boundary of these positive and negative telomere groups in fission yeast. In fact, Tpz1 has two interfaces that bind to Poz1 (-) and Ccq1 (+) (Jun et al., 2013). The Tpz1-Ccq1 interaction may enhance phosphorylation of Ccq1 by Tel1/Rad3, while the interaction of Tpz1 to Poz1 repress the phosphorylation of Ccq1 (Harland et al., 2014). These findings suggest that protein interactions within shelterin establish a hierarchy from non-extendible state to a telomerase extendible telomere state during the cell cycle.

CST and shelterin components co-exist

Though shelterin appears to be the major telomere capping complex in humans and fission yeast, CST components were also identified in both model organisms. Stn1 and Ten1, but not Cdc13, were found in *S. pombe* (Martin et al., 2007). Stn1 and Ten1 form a heterodimer essential for telomere stability. Interestingly, Tpz1 coordinates the action of shelterin and STN1-TEN1 for telomere metabolism. Sumolyation of Tpz1

mediates Tpz1 interaction with Stn1-Ten1 to negatively regulate telomerase recruitment (Miyagawa et al., 2014).

The evolution of CST complex seems to generate a different scenario in mammals. Human CST was shown to associate with the DNA polymerasea-primase, which is believed to coordinate C-strand fill in after telomerase elongation of telomeres (Casteel et al., 2009). This is also true for the Arabidopsis CST (Price et al., 2010). Moreover, multiple mutations in CTC1 were found to cause stem cell-releated telomere diseases such as Coat Plus, Dykeratosis congenita and bone marrow failure syndromes (Anderson et al., 2012; Keller et al., 2012; Walne et al., 2013), suggesting a telomeric specific function of mammalian CST complex (Price et al., 2010). However, the function of mammalian CST seems beyond the telomere capping. Recent findings suggest that the mammalian CST terminates telomerase activity at the chromosome end by competing with POT1/TPP1, the positive regulator of telomerase, for binding to the G-overhang at specific periods during cell cycle (Chen et al., 2012a). Moreover, although mammalian CST localizes to telomeres, CST proteins do not display strong specificity for telomeric DNA sequences (Miyake et al., 2009). Depletion of the CST components activates a DDR and leads to chromosome fusions at non-telomeric sites. Only a small portion of CST co-localizes with telomeres, and the rest of the foci in the nucleus are without a clear identity, suggesting the mammalian CST may have a general DNA metabolism function (Surovtseva et al., 2009).

In budding yeast and Arabidopsis where CST complexes plays a dominant role in telomere protection, some components of shelterin were identified too. Rap1 was first discovered in *S. cerevisiae* as a transcriptional regulator (Shore and Nasmyth, 1987) . However, overexpression of Rap1 will cause telomere lengthening and increased

telomere heterogeneity as a result of increased recombination. Thus Rap1 is a positive regulator for telomere length. It is also a negative regulator of telomerase. Rap1 was proposed to be part of a counting mechanism in which long telomeres bind more Rap1 leading to decreased telomerase action, and thus stimulating telomerase synthesis preferentially on short telomeres (Levy and Blackburn, 2004; Marcand et al., 1997; Shore and Bianchi, 2009). A high density of Rap1 at telomeres seems to decrease the accessibility of telomeres to telomerase. This is consistent with the finding that Rap1 facilitates the folding back of the telomere end to establish a telomerase non-extendible state. Furthermore, Rap1 mediates telomere silencing by directly binding to chromatin remodelers Sir3 and Sir4 (Bourns et al., 1998; Hardy et al., 1992). Lack of Rap1 will activate the DDR via the Tel1/Mec1 (ATR/ATM) dependent signaling pathway (Craven and Petes, 2000). Together with Rif1 and Rif2 (Rap interacting factor), Rap1 plays an essential role in maintenance of the telomere homeostasis (Wotton and Shore, 1997).

Efforts to find a plant shelterin-like telomere binding protein complex led to the identification of 12 TRF-like (TRFL) proteins in *Arabidopsis* (Karamysheva et al., 2004). Among them, six proteins (Class I TRFLs) contain a Myb-domain along with adjacent highly conserved “myb-extension” domain. These proteins show binding specificity for ds telomeric DNA sequence (Karamysheva et al., 2004). Their functions may be redundant because telomeres do not show defects in plants lacking one or more of them. However, a ds telomere binding protein TBP1 was subsequently found to be required for telomere length maintenance (Hwang and Cho, 2007). For ss telomeric DNA binding counterparts, *A. thaliana* is unique that it contains three POT1 genes: POT1a, POT1b and POT1c (Shakirov et al., 2005). However, instead of DNA binding, the Arabidopsis POT1 proteins have higher affinity for TERs. Each of them is

associated with a distinct telomerase RNP (Cifuentes-Rojas et al., 2011; Cifuentes-Rojas et al., 2012) (see below). So far, ortholog of the other shelterin components (TPP1, RAP1 or TIN2) have not been identified in plants. In summary, similar components for telomere binding and protection appear in most eukaryotes, but their functions have diverged during evolution, co-opted with distinct mechanisms to promote genome stability (Figure 1-3).

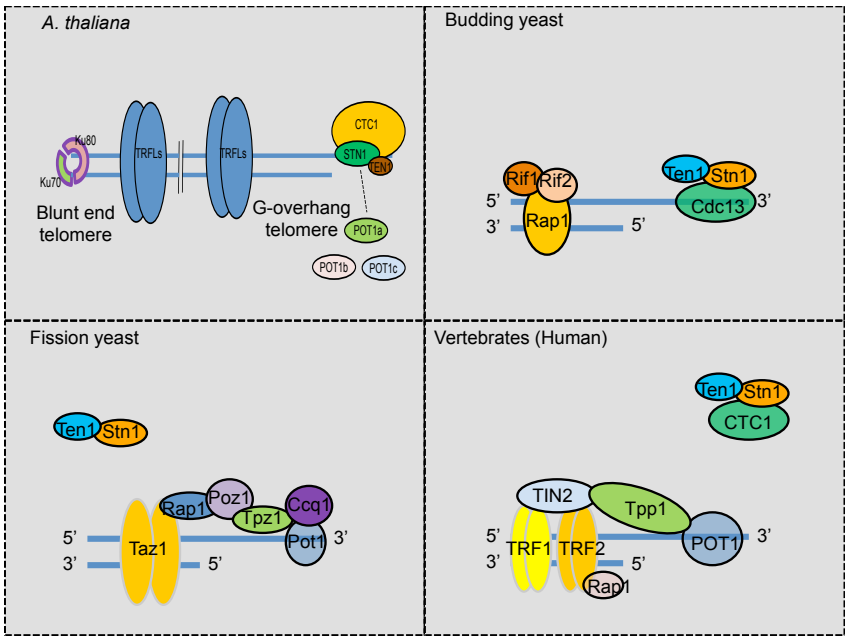


Figure 1-3. Telomere binding proteins in different organisms. Top left, *A. thaliana* telomere binding proteins: CTC1/STN1/TEN1 for the G-overhang; Ku for blunt end telomere binding. TRF-like proteins for double strand telomere binding. Top right, budding yeast telomere binding proteins: Rap1 binds to the ds telomere region with association of Rif1 and Rif2. The Cdc13/Stn1/Ten1 (CST) complex binds ss telomeric DNA. Bottom left, fission yeast telomere binding proteins consist of a shelterin-like complex: Taz1, Rap1, Poz1, Tpz1 Pot1 and Ccq1, which protect the telomere. The CST components Stn1 and Ten1 also exist. Bottom right, vertebrate shelterin complex: TRF1, TRF2, RAP1, TIN2, TPP1 and POT1. CST complex may associate with telomere only transiently.

Telomerase

In the vast majority of eukaryotes, telomerase appears to be the primary mechanism used to replenish telomeric DNA lost as a result of semi-conservative DNA replication (Shore and Bianchi, 2009). Telomerase is a ribonucleoprotein complex with two components essential for its catalytic activity: telomerase reverse transcriptase protein (TERT) and telomerase RNA (TER). Telomerase is a specialized reverse transcriptase. TERT reiteratively copies a template sequence within TER for telomere synthesis. Although telomerase solves the end-replication problem, TERT and TER genes are not found across all eukaryotes. *Drosophila* has lost the TERT and TER genes and instead maintains telomeres using a recombination/transposition based mechanism (Biessmann and Mason, 1997).

Mechanism of telomerase

The telomerase mechanism includes substrate recognition, elongation and translocation (Autexier and Lue, 2006; Harrington, 2003). The substrate recognition mechanism may vary among organisms based on the fact that different proteins are involved in telomerase recruitment to telomeric DNA. When telomerase is recruited to the telomere, ssDNA will be aligned in the RNA template by Watson-Crick base pairing and the 3' end of the DNA will be extended by TERT reverse transcriptase activity. During primer elongation, not all nucleotides in the template form RNA-DNA base pairs with the primer DNA. This relatively unstable complex allows the template to re-align with the elongated DNA after one round of synthesis, which is called translocation (Hammond et al., 1997).

The telomerase mechanism cannot be separated from the flexible structure of its integral long non-coding RNA (lncRNA), TER. Although TER is highly divergent in sequence and length among different organisms (see below), it contains several conserved structures that promote the mechanism of telomere repeat synthesis (Egan and Collins, 2012). The first key structural element of TER is the single stranded template that corresponds to 1.5 copies of telomere complementary sequence. The template not only facilitates substrate recognition, but also indicates the incorporated nucleotide sequence during synthesis.

Next, a pseudoknot is formed as a triple-helix structure adjacent to the template region. Though its function is not well understood, the pseudoknot motif in telomerase RNA is believed to help TER folding and to boost enzymatic activity for primer elongation (Egan and Collins, 2012). Another conserved structure within TER is called the stem-terminus element (STE). The architecture of this element is either a hairpin in ciliates, a hairpin derived from a three-way junction in humans or a three-way junction alone in yeast (Brown et al., 2007). Nucleotide substitution in this motif affects telomerase activity both *in vivo* and *in vitro* (Brown et al., 2007).

An additional important structured motif in TER is the template boundary element (TBE), which directly flanks the template on its 5' side. The TBE stem or hairpin structure is formed by long-range base pairing and prevents non-template nucleotides from being copied during telomeric DNA synthesis. Therefore, the TBE is crucial for stopping each round of repeat synthesis to guarantee telomere repeat sequence fidelity. TBE also facilitates primer translocation. In yeast, the nucleotides that lie between the template and the TBE are proposed to act like an "accordion" for the translocation step (Berman et al., 2011). The precise mechanism of primer translocation is not well

understood in other organisms, however this process must be determined by dynamic interactions at the enzyme active site between the primer, TERT, the template and the TBE. Coordination of these active domains allows telomerase to achieve sequence fidelity as nucleotides are incorporated, and during the translocation step.

Telomerase is a processive enzyme. As with other polymerases, processivity of telomerase refers to the probability that the enzyme will remain associated with the substrate after each nucleotide incorporation event (nucleotide addition processivity) and round of translocation (repeat addition processivity). Telomerase processivity varies among different species. Human and ciliate telomerases can catalyze more than one round of telomere repeat synthesis, while their counterparts in mouse and yeast are not very processive (Cohn and Blackburn, 1995; Greider, 1991; Prowse et al., 1993). Even within the same organism, the processivity of telomerase enzyme varies under different growth conditions (Chang et al., 2007; Zhao et al., 2011) (Figure 1-4). Telomerase processivity can be regulated by telomerase accessory proteins such as p82 in ciliates (Min and Collins, 2010) and POT1a and TEN1 in plants (Leehy et al., 2013; Renfrew et al., 2014). In humans, the telomerase processivity is regulated by the trafficking of telomerase in Cajal bodies (Zhao et al., 2011). The shelterin component TPP1 could also affect telomerase processivity (Zhong et al., 2012)

Telomerase has the capacity to add telomere repeats onto non-telomeric DNA sequences in a process called *de novo telomere formation* (DNTF) (See below) (Diede and Gottschling, 1999; Flint et al., 1994; Kramer and Haber, 1993; Melek et al., 1996). This finding indicates that complementarity of the DNA 3' terminus to the TER template is not required. The promiscuity of telomerase in regards its substrate suggests that non-telomeric DNA can be positioned in the active site of telomerase independent of the

primer-template base pairing. This could be achieved by the dynamic interaction of TERT and TER shaping a unique stereospecific telomerase active site. Indeed previous studies revealed that non-telomeric DNA is aligned at a specific “default” entry site (Fitzgerald et al., 2001; Melek et al., 1996). An anchor site in the TERT subunit was also found to facilitate positioning the primer at the default entry site (Wang and Blackburn, 1997).

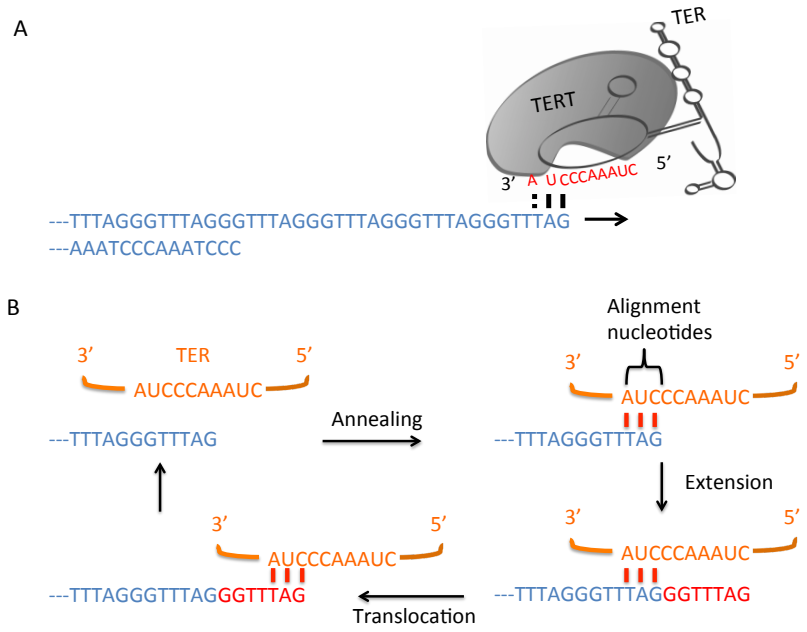


Figure 1-4. Telomerase maintains telomere length. (A) Telomerase, which is composed of the reverse transcriptase TERT and template RNA TER, is the primary mechanism for elongating telomeres. The template sequence within TER (red nucleotides) typically consists of 1.5 copies of the complementary telomere repeat. By the dynamic interaction with TERT, the TER template forms to an active pocket of the enzyme, and this allows the formation of the DNA primer-template complex. Base pairing between the primer 3’ end and the template region facilitates a stable primer interaction with the enzyme. The reverse transcriptase adds telomeric DNA *de novo* by reading the nucleotides at the 5’ template region. (B) Telomerase mechanism. A primer anneals at the alignment nucleotides within the TER template, followed by the extension of the nucleotides. After one round synthesis, the TER template translocates the primer 3’ terminus to the template alignment nucleotides for the next round of synthesis.

Telomerase regulation

Telomerase is a highly regulated enzyme. Hyper-active telomerase is a hallmark of human cancer, while deficient telomerase activity is associated with human stem cell diseases (Armanios and Blackburn, 2012). In single-celled organisms, telomerase activity is restricted to specific cell cycle stages (Bianchi and Shore, 2008). On the contrary, telomerase is active in most of the cells of mice, including somatic cells (Prowse and Greider, 1995), which may account for the fact that mice are prone to have tumors. In humans, telomerase activity is largely constrained to dividing cells (Forsyth et al., 2002). This is similar to the case in plants, where telomerase activity is high in young seedlings, gradually decreases during vegetative growth, and then again peaks in flowers (with many active dividing cells and the germline) (Fitzgerald et al., 1996).

Telomerase activity must be carefully regulated during reproduction to ensure that germline cells have long telomeres to pass to progeny, but also to prevent *de novo telomere formation* (DNTF, see below) during meiotic recombination. To exchange the genetic information, multiple programmed DSBs are introduced during meiosis along the chromatids. At the beginning of this process, a transesterase Spo11 attacks the DNA to generate the DSBs (Cole et al., 2010; Kleckner, 1996; Mahadevaiah et al., 2001). These DSBs DNA ends will then be covalently linked to Spo11 for protection with following steps of recombination. Recent data showed evidence of a Spo11 independent pathway for introducing DSBs (Farah et al., 2005; Storlazzi et al., 2003), raising the possibility that an active process is involved in restricting telomerase access to DSBs during meiosis. As discussed below, it is critical that telomerase must be prohibited from acting on DSBs.

Telomerase activity is modulated during the cell cycle in multiple organisms' studies. Transcription of yeast telomerase RNA TLC1 is controlled by a cell cycle dependent factor (Dionne et al., 2013), and this regulation is coupled with changes in localization, in which the active telomerase is clustered at the elongating telomere during the late S phase stage of the cell cycle. Telomerase engagement of the telomere is dependent on the ss telomere binding factor Cdc13. Phosphorylation of Cdc13 by CDK1 coordinates telomerase recruitment and telomere elongation (Li et al., 2009). Thus, differential recruitment of telomerase to telomeres during the cell cycle serves as another way to control telomerase action (Gallardo et al., 2011).

Cell cycle regulated localization of telomerase has also been found in humans and *Xenopus* oocytes (Li et al., 2010; Venteicher et al., 2009). In humans, this process is associated with the cajal body, which suggests that the assembly of telomerase RNP also plays a role in telomerase regulation. Plant telomerase activity is induced by a cell cycle related auxin hormone dependent factor in both *A. thaliana* and tobacco (*Nicotiana tabacum*) (Ren et al., 2004; Tamura et al., 1999). Therefore, telomerase activity is controlled dynamically through transcription, trafficking, and assembly by the cell cycle dependent signals.

TERT, the reverse transcriptase subunit

The core telomerase component TERT is a highly conserved reverse transcriptase (RT) protein, first found in the ciliate *Euplotes aediculatus* (Lingner and Cech, 1996), with homologs identified in several other organisms including other ciliates (Collins and Greider, 1995), yeast (Cohn and Blackburn, 1995), vertebrates (Greenberg et al., 1998; Meyerson et al., 1997; Nakamura et al., 1997) and plants (Fitzgerald et al.,

1999; Oguchi et al., 1999). In general, TERTs contain four domains: the telomerase N-terminal (TEN) domain, the TER binding domain (TRBD), the RT domain, and the C-terminal extension (CTE) (Lai et al., 2001). TEN and CTE enhance RT domain function in enzyme processivity and nucleic acid association (Autexier and Lue, 2006), though the minimal TERT gene in *C. elegans* only contains the RT domain. Although TERT exists in a single copy gene in most organisms, three different TERT genes were found in *E. crassus* (Karamysheva et al., 2003), and these function to either maintain pre-existing telomeres or to promote DNTF when chromosomes fragment during macronuclear development.

TER, the template containing RNA

Unlike the highly conserved TERT genes, TER varies greatly across species in its sequences and sizes, and ranges from 150nt in ciliates (Greider and Blackburn, 1989) to more than 1000nt in yeast (Dandjinou et al., 2004; Leonardi et al., 2008). However, as discussed earlier TERs harbor a number of conserved discrete structural elements that promote proper folding and telomerase enzymology.

The biogenesis of TER includes its expression and processing, which vary among different species. Transcription is carried out by RNA polymerase II in yeast and humans, and by Pol III in ciliates. In budding yeast and fission yeast, a minor fraction TER contains a polyA tail, while most of the RNA, including the functional TER associated with TERT, does not (Chapon et al., 1997; Leonardi et al., 2008). Interestingly in budding yeast, the polyA tailed-form of TER (TLC1) is not a precursor for the unpolyadenylated mature form (Noel et al., 2012). Instead, the mature 3' end of TLC1 is produced by the Nrd1-Nab3-Sen1 pathway, which is involved in small nuclear

and small nucleolar RNA maturation in yeast. A more surprising mechanism for TER 3' end formation is found in fission yeast. The association of the Sm complex protein at the 3' end of the *S. pombe* TER1 transcript promotes a novel spliceosome-mediated cleavage reaction corresponding to only the first transesterification step of canonical mRNA splicing (Box et al., 2008a). After cleavage, Sm proteins are released and Sm-like proteins (Lsm) subsequently bind to the 3' end to stabilize it (Tang et al., 2012). Dissociation of the first and second transesterification reactions to yield the mature TER1 3' end is mediated by an unconventional long distance between the branch site within intron and the 3' splice site (Kannan et al., 2015; Qi et al., 2015). Little is known about vertebrate telomerase RNA processing. The veterbrate TER (TR) contains a 5' hypermethylated cap, but lack a polyA tail (Zaug et al., 1996). An H/ACA box within the mature RNA, after about 100nt removal from the precursor RNA 3' end, is necessary for processing and maturation of the human telomerase RNP (Fu and Collins, 2003).

Biochemical methods were successfully used to identify *bona fide* TERs in *Arabidopsis*. Unlike other model organisms, *A. thaliana* encodes two TER subunits: TER1 (748nt) and TER2 (784nt) (Cifuentes-Rojas et al., 2011) (Cifuentes-Rojas et al., 2012). TER1 and TER2 share two conserved regions: CR1 (144nt) and CR2 (75nt). In TER1, CR1 and CR2 are contiguous, while in TER2 they are separated by an intervening sequence or intron. Preliminary data indicate that TER2 is subjected to two distinct processing steps: splicing of the intron and cleavage of the 3' end to generate a smaller isoform of TER2 known as TER2s (219nt) (Cifuentes-Rojas et al., 2012). TERT has higher binding affinity to TER2 than TER1, but very low affinity for TER2s. The *A. thaliana* TER2 intron boundary regions are distinct from canonical mRNA splice sites, arguing that the biogenesis of *Arabidopsis* TER2s may be through a novel pathway. So

far, canonical introns have not been correlated with TER in any organism. Chapter II and IV in this dissertation focus on investigating the nature of the ArabidopsisTER2 intervening sequence and its processing.

Telomerase beyond telomere

The lack of stringency in TER sequence and length in different species suggests each TER gene has undergone a distinct evolutionary pathway to meet the need of the host. Interestingly, TERT does not exclusively interact with TER. Human TERT associates with more than 30 RNAs species (Maida et al., 2009), indicating that TERT may have additional functions besides telomere maintenance. For example, human TERT associates with the RNA-dependent RNA polymerase (RdRP), which produces substrates for Dicer processing (Maida et al., 2009). Recent data also indicate that TERT plays a role in heterochromatin maintenance at centromeres and transposons (Maida et al., 2014). Additionally, the search for TERs in plants besides *A. thaliana* discovered the TER-like molecules that lacks completely a template domain (Beilstein et al., 2012). Since plants lacking telomerase are viable for multiple generations (Riha et al., 2001), it is possible that we may witness the scenario where shows evolution of a telomerase RNA like gene to the true telomerase RNA gene.

Telomerase accessory proteins

TERT and TER are sufficient to reconstitute enzymatic activity *in vitro* (Chen and Greider, 2005; Licht and Collins, 1999), but *in vivo* other components are required for full function of the telomerase holoenzyme. In budding yeast, depletion of the core components of telomerase, Est2 (TERT) and Tlc1 (TER), cause a classical EST (ever

shorter telomere) phenotype, which results in the gradual shortening of telomeres over successive generations (Lundblad, 2003). This phenotype was also observed with the loss of Est1 or Est3, two other non-catalytic accessory proteins associated with budding yeast telomerase RNP. Therefore, the requirements for telomerase function *in vivo* are beyond its two core RNP components (Lendvay et al., 1996; Lundblad and Szostak, 1989).

TER is a scaffold for binding telomerase accessory proteins, and contains species-specific RNA elements that recruit protein subunits. These RNA elements together with their accessory proteins play roles in TER maturation, trafficking, RNP assembly and enzymatic activity regulation. For example, in ciliates, stem I recruits the binding of p65, which causes a conformational change in TER to facilitate telomerase holoenzyme assembly (Berman et al., 2010). In vertebrates, the H/ACA motif residing at the 3' of TER serves as the binding site for Dyskerin, a protein essential for small nucleolar RNP maturation (Bellodi et al., 2013), which enhances telomerase RNP assembly as well (Egan and Collins, 2012). In yeast, two stem loops structures distal to the enzymatic core of TLC1 form a binding site for Est1 (Livengood et al., 2002) and Ku (Stellwagen et al., 2003). These interactions contribute to nuclear localization of TLC1 and recruitment of telomerase to chromosome ends.

Among the telomerase accessory proteins in *A. thaliana*, the best characterized so far is POT1a. Mutants lacking POT1a exhibit reduced telomerase activity and continuous telomere shortening (Surovtseva et al., 2007). A *pot1a tert* double mutant does not show further telomere shortening, arguing that POT1a and TERT act in the same genetic pathway (Surovtseva et al., 2007). Unlike POT1 orthologs in other organisms, AtPOT1a binds to TER1 instead of telomeric DNA (Cifuentes-Rojas et al.,

2011; Shakirov et al., 2005). However, POT1a is not required for TER1 RNP assembly or telomerase recruitment (Renfrew et al., 2014). Instead, POT1a facilitates repeat addition processivity of the TER1 RNP (Renfrew et al., 2014). Although POT1a has a distinct function in telomerase regulation, its structure is like other POT1 proteins, retaining two OB folds and a C-terminal domain (Surovtseva et al., 2007). Recent intriguing results showed that the first OB fold of AtPOT1a is sufficient to bind to telomeric DNA (A. Arora and D. Shippen, data not published), suggesting OB2 or the C-terminus of POT1a specify TER1 binding.

Another telomerase-associated protein is Ku. Ku is a heterodimer complex composed of Ku70 and Ku80 subunits and in *Arabidopsis* associates with the TER2 RNP (Cifuentes-Rojas et al., 2012). In humans, Ku heterodimer binds TER (Ting et al., 2005) as well as TERT (Chai et al., 2002), while in budding yeast, Ku associates exclusively with TLC1 (Stellwagen et al., 2003). Ku-TLC1 binding facilitates *de novo* telomere formation (DNTF) at broken chromosomes. Moreover, Ku is proposed to act as a cell-cycle dependent telomerase recruitment factor (Fisher et al., 2004). Yeast lacking Ku has a decreased the telomerase recruitment to telomere in G1 and telomere shortening. Interestingly, yeast Ku also binds telomeric DNA (Lopez et al., 2011), but recent data indicate that Ku does not bind to TLC1 and telomeric DNA simultaneously (Pfingsten et al., 2012). Arabidopsis *ku* muants have longer telomeres with increased telomere recombination (Riha and Shippen, 2003). This extended telomeres in *ku* mutants depends on telomerase, including both TERT and POT1a (Renfrew et al., 2014). Paradoxically, telomerase activity appears to slightly decrease in *ku* mutants. Also, in plants lacking Ku DNTF efficiency is consistently decreased (Nelson et al.,

2011). Thus, it is unclear precisely how Ku impacts telomerase activity and telomere maintenance.

DNA damage and telomere dysfunction

Faithfully transmitting genetic information to offspring is an essential task for life, and is threatened by damage to the DNA. Damaged DNA arises from DNA replication errors including dNTP misincorporation, base deamination, base loss by depurination and alkylation (Ciccia and Elledge, 2010). Another endogenous source contributing to DNA damage is the reactive oxygen species (ROS) produced from normal cellular metabolism. Altogether, an estimate of 10^5/cell DNA lesions are produced everyday (Hoeijmakers, 2009).

DNA damage response (DDR) signal transduction pathways have evolved to sense DNA damage and replication stress and to promote DNA repair. Residing at the center of the signal cascade are the phosphatidylinositol 3-kinase-like protein kinases (PIKK) ATM, ATR, and DNA-PKcs. ATM and DNA-PK are activated by DSBs (Harper and Elledge, 2007). ATR, with its partner ATRIP, is recruited to the RPA-coated ssDNA (Cimprich and Cortez, 2008).

Upon DNA damage, the kinase activity from ATM/ATR/DNA-PKcs is activated leading to post-translational modification of various substrates involved in downstream DNA repair pathways. In the case of DSBs, these kinases phosphorylate the histone variant H2AX surrounding the DSBs identifying this site as a substrate for the repair machinery (Kinner et al., 2008; Szilard et al., 2010). Moreover, ATM/ATR-mediated phosphorylation of CDC25, which is required for CDK activation during the cell cycle

together with other cell cycle regulators further facilitate the precise adoption of the specific repair pathway (Zhou and Elledge, 2000).

There are two main DNA repair pathways for DSBs: homologous recombination (HR) and non-homologous end-joining (NHEJ). HR requires an intact sister chromosome as a template for repair, which thus limits repair to S and G2 phases. In G1, the NHEJ mechanism is utilized. This pathway is an error prone repair pathway, but a strongly preferred option in cells that have not replicated their DNA (Symington and Gautier, 2011). The NHEJ pathway can be further grouped into classical NHEJ (C-NHEJ) and alternative NHEJ (A-NHEJ) (Stracker and Petrini, 2011). C-NHEJ involves Ku-mediated stabilization of the DSB (Riha et al., 2006), followed by processing and ligation of the DNA (Boulton and Jackson, 1996). The downstream steps utilize exonuclease Artemis, which makes the ends of the DSBs a suitable substrate for DNA ligase IV (Moshous et al., 2001).

One important function of telomeres is to prevent the chromosome ends from being mistakenly recognized as DSBs. Failure to do so can result in telomere fusion. In the absence of telomere capping proteins, telomere fusion is rampant (Song et al., 2008; Surovtseva et al., 2009; van Steensel and de Lange, 1997; Wu et al., 2006). In plants, telomere fusion can be temporarily averted by the absence of ATR, which rescues the severe morphological phenotype caused by the lost of CST, the main telomere capping protein complex (Boltz et al., 2012). Ultimately, however, telomeres without CST and proper DDR activation exhibit a drastic shortening of the telomeres (Boltz et al., 2012).

Interestingly, components of the NHEJ pathways are not only associated with normal telomeres and telomerase, but are also required for their proper regulation. Ku is

essential for maintaining telomere ends in human, yeast, and Arabidopsis (Bertuch and Lundblad, 2003; Riha and Shippen, 2003; Wang and Baumann, 2008), suggesting a dynamic interaction between telomere metabolism and DSB repair pathways. Cells with aberrant chromosome ends must be culled out. Therefore, the DDR bridges telomere metabolism and cell cycle regulation via the dynamic interactions in each organism.

Telomerase and *de novo* telomere formation

DSBs are the most lethal lesions on DNA. In addition to the canonical DNA repair pathways, DSBs can also be "repaired" by telomerase action at the break site through *de novo* telomere formation (DNTF). This phenomenon is referred to as "chromosome healing". DNTF was first identified in maize in 1940 (McClintock, 1939, 1941), and later found in yeast and humans (Kramer and Haber, 1993; Pennaneach et al., 2006; Stellwagen et al., 2003). After addition of telomere repeats to the DSB site, the new chromosome end will be stabilized by binding telomere proteins, which will allow the resumption of the cell cycle (Melek et al., 1996). The newly synthesized telomere at the DSB does not allow for accurate DNA repair, however, and causes the formation of an acentric chromosome fragment that will ultimately be lost. The DNTF process therefore reduces fidelity of DNA repair, and decreases chromosome integrity. Failure to prevent DNTF causes genetic diseases in humans (Lamb et al., 1993). As a consequence, mechanisms to inhibit the DNTF at the sites of DNA breaks have evolved to protect the genome (Figure 1-5).

Mechanisms to repress DNTF include disruption of TER/DNA annealing at the breakage site, inhibition of telomerase recruitment to DSBs and telomerase sequestration (Schulz and Zakian, 1994; Zhou et al., 2000). In budding yeast, Pif1, a 5'-

3’ helicase, inhibits telomere addition at DSBs by unwinding the DNA-TLC1 hybrid. In addition, Mec-1 (ATR) suppresses Cdc13 accumulation at the DSB by phosphorylating it, which then abolishes telomerase recruitment to that site (Zhang and Durocher, 2010). Ku also plays an important, yet enigmatic role in yeast DNTF. Ku is required for both NHEJ and protecting telomeres from end joining (DuBois et al., 2002). Ku is also an accessory protein associated with telomerase, and in yeast the association is essential for DNTF (Stellwagen et al., 2003). Notably, DNTF frequency is also lower in plants lacking Ku (Nelson et al., 2011). DNTF has also been studied in human cells (Bae and Baumann, 2007; Barnett et al., 1993; Hanish et al., 1994) (Okabe et al., 2000), but how this process is regulated remains unclear. Immediately after the introduction of DSBs, TERT is transiently sequestered in nucleoli of human cells as well as in yeast as a part of inactive complex lacking TER (Wong et al., 2002). Thus, sequestration of TERT may serve as another layer of regulation to prevent DNTF. Taken together, the findings indicate that the mechanisms to inhibit DNTF are essential, but are not generally conserved across evolution.

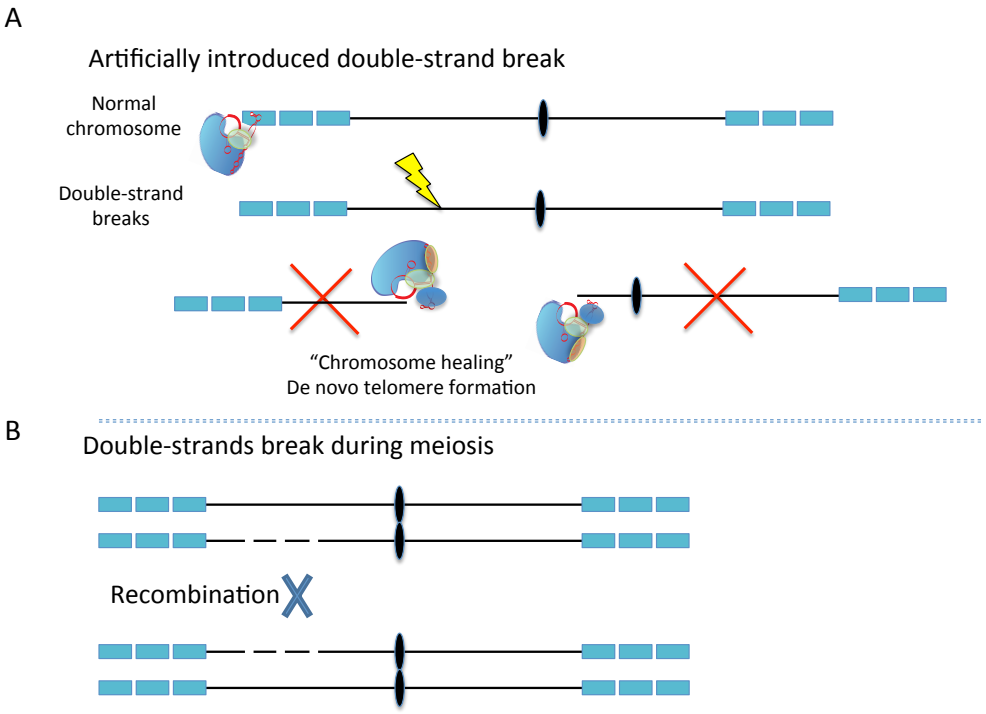


Figure 1-5. DSBs and telomerase (A) Telomerase adds telomere repeats to natural chromosome ends as well as the DSBs formed at internal DSB sites of the chromosome. *De novo telomere formation* (DNTF) is often lethal because the accentromeric fragments will be lost after cell division. Mechanisms have evolved to repress telomerase during DSBs. (B) During meiosis prophase I, the programmed induction of DSBs is required for the homologous recombination on non-sister chromatids to allow for genetic exchange. This can be one potential platform for DNTF, and thus telomerase regulation is needed.

Transposable elements and their exaptation in the genome

Another source of genome instability is through the activity of transposable elements (TE). TEs were first discovered by Barbara McClintock in maize and later found widely in eukaryotes (Fedoroff, 2001). There are two types of TE, Class I (Retrotransposon), which depends on an RNA intermediate for transposition, and Class II (DNA transposon). Class I elements can further be categorized into long terminal repeats (LTR) and non-LTR elements, such as long interspersed nuclear elements (LINEs) and short interspersed nuclear elements (SINEs) (Casacuberta and Gonzalez, 2013).

At one time TEs were regarded as a selfish DNA element that worked to the detriment to the host. However, more evidence suggests that TEs provide important resources for shaping the genome and its evolution. A variety of TEs that cause gene function changes have been discovered. TEs can insert themselves into an open reading frame (ORF) to disrupt the coding sequence for producing proteins. Examples of this were reported in several human diseases (Chen et al., 2005; Kazazian et al., 1988). If the TE resides near the promoter region of a gene, it may introduce an alternative promoter, resulting in a tissue/development-specific transcription profile (Faulkner and Carninci, 2009).

The impact of TEs can also be found at the mRNA processing level, resulting in the introduction of alternative splicing sites and poly(A) tail sites. Moreover, TEs that encode a small RNAs can introduce post-transcriptional regulation of genes (Cowley and Oakey, 2013) (See discussion later).

Benefitting from the flexibility of gene function, TEs are the most abundant component in the human genome, making up half to two thirds of the human genome (de Koning et al., 2011; Lander et al., 2001), and contributing to genetic variation (Huang et al., 2010). Surprisingly, TEs are the major contributing factor for the origin and diversification of long non-coding RNAs (lncRNAs) genes in humans. About 70% of lncRNAs are derived from TEs (Johnson and Guigo, 2014). Because current protein coding gene based models do not fit with the TE impact on lncRNAs, Guigo and Johnson proposed that the subset of TEs residing in lncRNA fold into pre-formed structural domains. Thus, the exaptation of TEs occurs frequently in protein coding regions and lncRNAs as well as non-coding regions to shape evolution of the whole genome (Figure 1-6).

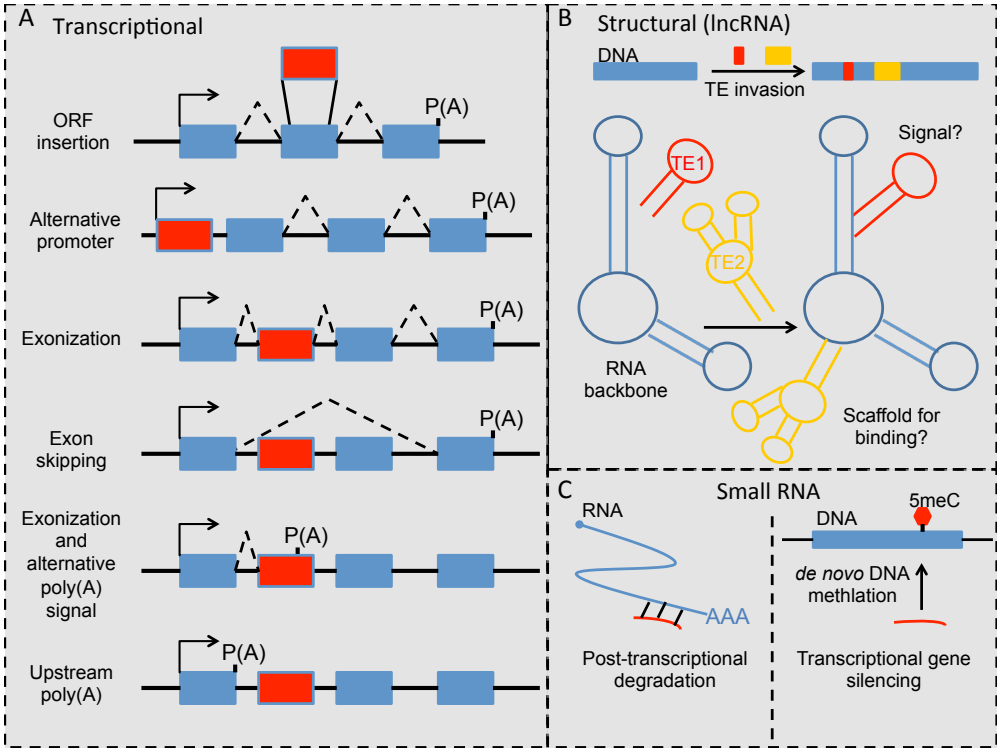


Figure 1-6. Modes for how exaptation of a TE transforms gene function. TEs can alter gene function in three manners: transcriptional, structural and by small RNAs. (A) TE invasion into an ORF can disrupt gene function. TEs can also introduce an alternative promoter to change the expression profile. TEs can affect mRNA splicing by exonization or exon skipping. Lastly, TEs can change the position of the poly(A) site (Figure modified from Cowley and Kakey, 2013). (B) TEs are proposed to have pre-formed structures (yellow and red). Invasion of the TE therefore introduce a structural element into the RNA. TEs are especially represented in long non-coding RNAs. (C) Small RNAs targeted in TEs could introduce a target site in mRNA for small RNA mediated post-transcriptional degradation or *de novo* methylation.

Transposable elements and their regulation via small RNA machinery

TEs must be finely regulated to prevent unwanted mutagenesis. Epigenetic regulation is often associated with TEs through small RNA-dependent DNA methylation (RdDM). The small RNAs derived from TEs guide DNA methylation machinery to incorporate methyl group on cytosines within the TE locus using sequence complementarity. Transcriptional silencing of the TE is the result. RdDM may also affect the expression of nearby genes, thereby contributing to the post-transcriptional regulation of genes adjacent to TEs (Lippman et al., 2004).

At the core of this pathway is the biogenesis of small RNA from TEs. The generation of small RNA is dependent on DICER proteins (Tabara et al., 2002; Vaucheret, 2006). In *A. thaliana,* there are four DICER-LIKE (DCL) RNase III enzymes: DCL1, DCL2, DCL3 and DCL4 (Liu et al., 2009). While some small RNAs are generated by a particular dicer enzyme, biogenesis of other small RNAs, like ta-siRNAs, involves more than one dicer (Catalanotto et al., 2004). The RNA substrate must be converted from a ss precursor to a ds molecule before Dicer can produce small RNAs. This conversion is dependent on RDR2 (RNA dependent RNA polymerase 2) or RDR6 (SGS2) (Figure 1-7). A growing number of functions are emerging for small RNAs in the nucleus, in addition to their well-characterized roles in gene silencing in the cytoplasm (Castel and Martienssen, 2013; Wei et al., 2012). Based on the fact that TEs have a major impact on lncRNA evolution, it is possible that small RNA processing machinery via TEs has an even broader role in lncRNA metabolism. This purpose is further discussed in Chapter IV.

The small RNAs associated with TEs shape the evolution of the TE or the gene that carries it. It is reported that methylation of TEs by small RNA mediated pathways is

under stronger purifying selection near coding regions than at TEs near non-coding regions (Hollister and Gaut, 2009). Increased loss or partial deletion within TEs is associated with coding regions (Wang et al., 2013). Especially relevant to this dissertation, one systematic study reported that RNA-directed DNA methylation silencing of TEs has a lower efficiency in *Arabidopsis lyrata* compared to *Arabidopsis thaliana* (Hollister et al., 2011). This finding sheds light on the cause of differential TE proliferation between *A. thaliana* and *A. lyrata*. Different levels of TE methylation are also detected among *A. thaliana* accessions, with Ler-0 having more TE appearance than Col-0 (Wang et al., 2013) (see Chapter II).

TE epigenetic regulation is involved in environmentally induced responses (Lisch and Bennetzen, 2011). For example, one retrotransposon is responsible for cold dependent accumulation of anthocyanin in blood orange (Butelli et al., 2012). The promoter region of the gene responsible for blood orange pigment metabolism is highly methylated due to the embedment of a TE. Methylation is alleviated by cold stress, allowing gene expression. A similar report shows that a *Copia*-like retrotransposon *ONSEN* is induced and transposes actively upon heat-shock once the sRNA pathway is abolished (Ito et al., 2011). Unlike the cold shock element, this case more likely reflects a evolution of the TE to escape the post-transcriptional regulation by acquiring the ability to respond to stimulants from the environment. There are multiple copies of *ONSEN* copies in the genome, and only the ones capable of binding heat-shock transcriptional factors become active upon heat-shock (Cavrak et al., 2014). Therefore, the silent arms race between host regulation machineries and TE populations involves an optimized host genome to counter environmental challenges.

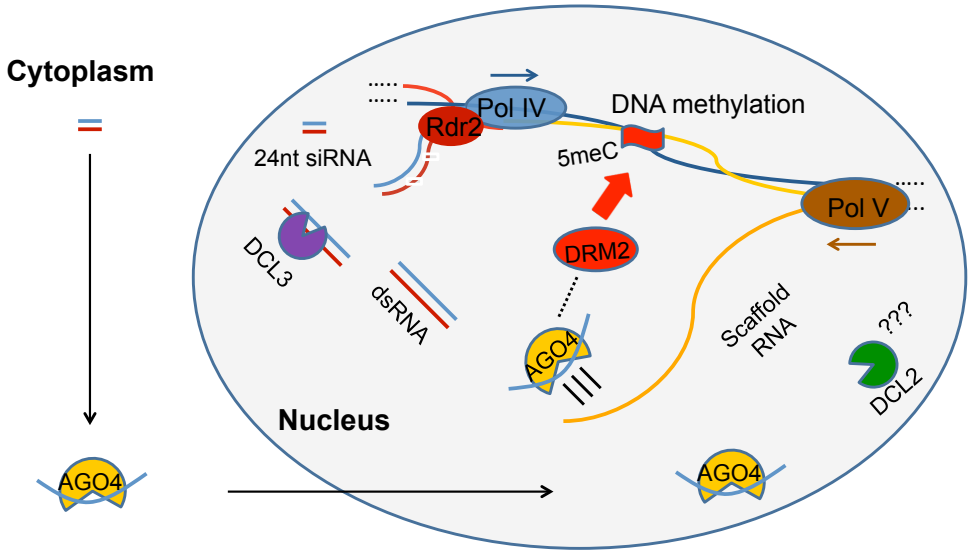


Figure 1-7 Small RNA dependent DNA methylation in *Arabidopsis thaliana.* A transcript (light blue) is generated by RNA polymerase IV at a target locus. This nascent ss will be converted to ds RNA by RNA dependent RNA polymerase 2 (Rdr2). DCL3 is the main Dicer-like protein to recognize the dsRNA as substrate and processes it into 24nt small RNA. The 24nt small RNA is exported to the cytoplasm and loaded into AGO4 followed by import of the RNA-protein complex into nucleus. RNA polymerase V generates a scaffold RNA in the vicinity of the PolIV transcription site. Base pairing between the small RNA and scaffold RNA recruits AGO4. DNA methylation machineries (DRM2 as the major component) interact with AGO4 to specifically modify the DNA site with a 5'methyl group on cytosine (5meC). In this dissertation I provide evidence that the TE bearing TER2 is processed by DCL2 in a novel pathway.

Arabidopsis thaliana as a model system for telomere biology

The Shippen laboratory developed *Arabidopsis thaliana* as a model system to study telomeres and telomerase. Loss-of-function and over-expression mutants are easy to study in *Arabidopsis* because the entire 125MB genome has been sequenced. Adding this to the fact that Arabidopsis is easy to manipulate, has a short growth period, and has a fully developed repertoire of experimental techniques. These facts make this model system ideal to work with. This convenience has been further extended by the recent SALK *A. thaliana* 1001 Genomes project, which organizes fully sequenced *Arabidopsis* genomes from hundreds of different accessions. This work revealed unprecedented insight into how natural selection is shaping the evolution of this important model organism. Genome-wide analysis of multiple accessions showed that the average allele contains seven single nucleotide polymorphisms (SNPs) per kilobase (Nordborg et al., 2005; Schmid et al., 2005). By comparision, *A. thaliana*'s closest evolutionary relatives, *A. lyrata* and its subspecies *A. lyrata ssp Petraea* show about 14 SNP/kb, two fold more than for *A. thaliana*. However, the 7 SNPs/kb ratio is the same as in *Drosophila melanogaster* (Ometto et al., 2005), and roughly an order of magnitude higher than in humans (Akey et al., 2004). This pattern of polymorphisms reflects natural selection of candidate genes for adaptive variation since *A. thaliana* habitats vary dramatically across the world (Agrena et al., 2013). *A. thaliana* is native to Europe and central Asia and naturalized in North America (Mitchell-Olds and Schmitt, 2006). The SNP distribution lies within an optimal range in the *A. thaliana* genome, allowing for fundamental advances in knowledge of gene evolution.

With respect to telomere biology, *Arabidopsis* has many advantages. Telomeres are very short (only 2-5 kb), and bulk telomeres can be easily observed by the Terminal

Restriction Fragment (TRF) assay (Shakirov and Shippen, 2004). Most subtelomeres consist of unique sequences, making it possible to study individual telomeres for their length and their integrity using Primer Extension Telomere Repeat Amplification (PETRA) or Telomere Fusion PCR (TF-PCR) assay (Heacock et al., 2004). Finally, *Arabidopsis* is much more tolerant to telomere dysfunction than mammals, which enables researchers to examine fundamental aspects of telomere biology and its contribution to genome stability. By all these virtues, exciting discoveries were found using plants as a model for telomere studies.

In this dissertation, my main focus is to investigate the function of telomerase RNP in the context of TER evolution. Due to the duplication of TER and the unusual processing events of TER2 in *A. thaliana,* there are three distinct telomerase RNP particles and each RNP consists of a unique RNA and different sets of protein subunits (Cifuentes-Rojas et al., 2011; Cifuentes-Rojas et al., 2012). The TER1 RNP is composed of TERT, POT1a and dyskerin. For the TER2 RNP, TERT is present along with the Ku70/80 heterodimer (Ku), POT1b and dyskerin. *In vitro* and *in vivo* binding studies indicate that TER2s forms a subcomplex containing POT1b and Ku. The function of the TER2s RNP is not well understood. TER1 RNP is responsible for maintaining telomere length like other canonical telomerase RNPs (Cifuentes-Rojas et al., 2011), while the alternative TER2 RNP is required for telomerase repression upon DNA damage (Cifuentes-Rojas et al., 2012) (Figure 1-8). It is possible that hijacking of telomerase by a non-canonical RNA could be a common strategy used to negatively regulate telomerase. Thus, analysis of *A. thaliana* has the potential to reveal new avenues for telomerase regulation that could potentially be important for controlling enzyme activity in human cells to prevent the genome instability that leads to cancer.

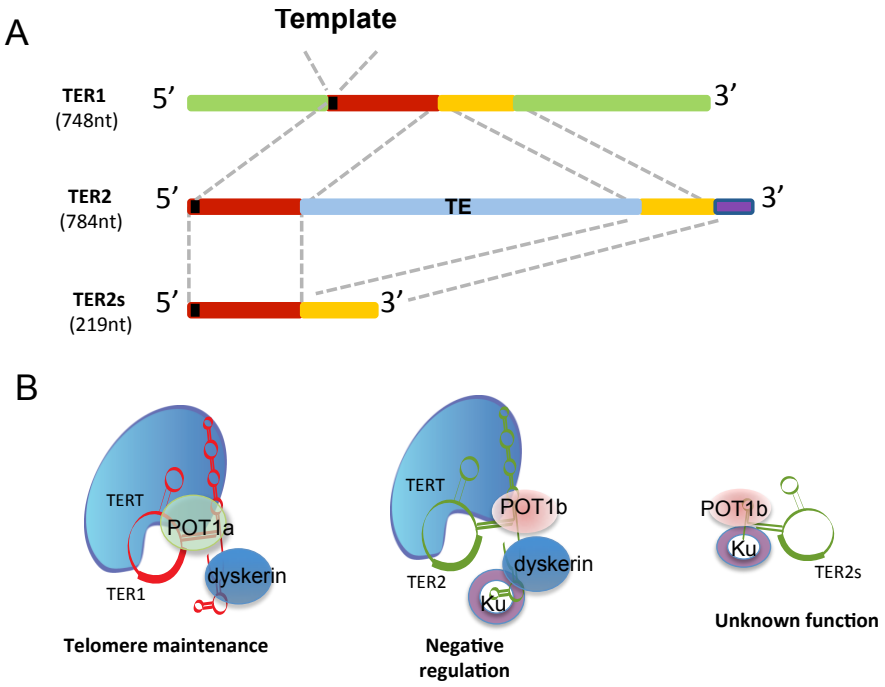


Figure 1-8 Telomerase components in *A. thaliana*. (A) Duplication of TER. Three TERs: TER1, TER2 and TER2s. In TER1, two conserved regions are adjacent. An intervening sequence bisects CR1 and CR2 in TER2. TER also contains a unique 36nt sequence at its 3' end. Removal of the intervening sequence and the 3' end give rise to TER2s. (B) Telomerase RNPs: TER1 RNP, TER2 RNP and TER2s RNP. Each RNP has distinct components as well as functions. TER1 RNP contains dyskerin and POT1a, and is responsible for telomere maintenance. TER2 RNP contains dyskerin, POT1b and Ku, nand serves as a negative regulator of TER1 RNP. TER2s RNP with POT1b and Ku has an unknown function.

Dissertation overview

The overall objective of this dissertation research was to generate a deeper understanding of *A. thaliana* TER molecules. In Chapter II I present a detailed analysis of the evolution and function of the intron within TER2, and specifically of its role in regulating TER2 metabolism and telomerase activity. Here it is shown that the TER2 intron is a TE and this sequence is responsible for many of TER2's unique functions and regulation. Exaptation of a TE in TER2 causes the RNA to become unstable, but this outcome is averted by the stress of DNA damage. I further show that the TE is required for TER2 to serve as a negative regulator of telomerase. I also demonstrate that the TE in TER2 significantly increases TERT binding affinity. Altogether, these findings reveal a powerful example of how exaptation of a TE has provided a novel mode of telomerase regulation using a lncRNA.

Chapter III illustrates the biological significance of TER2 in plant reproduction. I report that TER2 RNA, not TER2D (lacking the TE), peaks in flower buds where the reproductive process is active. I show that *ter2* mutants have increased seed abortion, and reduced pollen viability. I also show that double mutants of *pot1a ter2* exhibit a synergistic effect in seed abortion, suggesting the distinct function of TER2 for reproduction is independent of telomerase regulation. A hypothesis is presented stating that TER2 is a novel signaling molecule that benefits reproductive development. This finding may suggest a common strategy for using the lncRNA in the regulation of the reproduction.

Chapter IV introduces the factors affecting TER2 processing and abundance. This study demonstrates that TER2 processing is not likely to be autocatalytic, in contrast to preliminary results from the Shippen lab (A. Hernandez and D. Shippen

unpublished data). Genetic data show that although TER2 does not contain consensus mRNA splice sites, its abundance is increased in mutants lacking canonical splicing machinery. Finally, data are presented showing that TER2 metabolism is influenced by Dicer-like 2, a key component of the epigenetic small RNA processing pathway. The results indicate a novel mechanism for DCL2 in modulating TER2 that involves post-transcriptional regulation during reproductive development. These findings uncover a new role for DCL2 in the processing of TE-derived lncRNAs. The results of this study will shift the paradigm of Dicer protein function from small RNA processing enzyme to broader role in RNA metabolism.

Chapter V investigates how the template domain of *A. thaliana* TER1 is evolving to promote telomere synthesis. Using bioinformatics analysis data from the 1001 genome project, three types of TER1 template polymorphisms were identified. Single nucleotide polymorphisms at the 5' and 3' sites of the template were verified by sequencing. I show that despite these nucleotide changes, only perfect TTTAGGG repeats are synthesized *in vivo*, and thus these polymorphisms define the functional 5' and 3' boundaries of TER1 in several *A. thaliana* accessions. I also demonstrated that telomerase does not exhibit a stringent requirement for the primer 3' sequence when it positions the primer in the template. Telomerase can work on non-telomeric sequences, suggesting a preferred entry site (default). However, the nucleotide incorporation error rate increases for primers with weak base pairing potential to the template of TER1. These findings suggest two independent modes of primer recognition that 1) contribute to the stability of the initial primer-template complex formation, and 2) determine the fidelity of telomerase. My results provide new insight into mechanism of template

utilization by telomerase and suggest a model for minimizing DNTF to ensure genome stability.

Finally, chapter VI presents conclusions and future directions for the research. Several novel hypotheses for TER2 are proposed to direct future research in the regulation of telomerase and reproduction.

CHAPTER II

A TRANSPOSABLE ELEMENT WITHIN *ARABIDOPSIS* TER2 MODULATES TELOMERASE ENZYME ACTIVITY IN RESPONSE TO DNA DAMAGE

Summary

Long noncoding RNAs (lncRNAs) have emerged as critical factors in many biological processes, but little is known about how their regulatory functions evolved. One of the best-studied lncRNAs is TER, the essential RNA template for telomerase reverse transcriptase. We previously showed that *Arabidopsis thaliana* harbors three TER isoforms: TER1, TER2 and TER2s. TER1 serves as a canonical telomere template, while TER2 is a novel negative regulator of telomerase activity, induced in response to double-strand breaks. TER2 contains a 529 nt intervening sequence that is removed along with 36 nt at the RNA 3' terminus to generate TER2s, an RNA of unknown function. Here we investigate how *A. thaliana* TER2 acquired its regulatory function. Using genomic data from the 1,001 Arabidopsis genomes project, we report that the IS within TER2 is derived from a transposable element termed telomerase regulatory element (TRE). TRE is found in the TER2 loci of most but not all *A. thaliana* accessions. By analyzing accessions with (TER2) and without the TRE (TER2Δ) we demonstrate that this element is responsible for many of the unique properties of TER2, including its enhanced binding to TERT and telomerase inhibitory function. We show that the TRE destabilizes TER2, and further that TER2 induction by DNA damage reflects increased RNA stability and not increased transcription. The TRE-mediated changes in TER2 stability provide a rapid and sensitive switch to fine-tune telomerase enzyme activity. Altogether, our data shows that invasion of the TER2 locus by a small

transposon converted this lncRNA into a DNA damage sensor that modulates telomerase enzyme activity in response to genome assault.

Introduction

The discovery of long noncoding RNA (lncRNA) has challenged the prevailing paradigm of protein-mediated regulation of gene expression and cell behavior. lncRNAs play essential roles in epigenetic regulation, stem cell biology and signal transduction and are emerging as key targets in human disease (Lee et al., 1996; Guttman et al., 2011; Scheuermann and Boyer, 2013). Unlike small regulatory RNAs (e.g. miRNAs, siRNAs), lncRNAs are not subjected to purifying selection, and as a consequence they are very poorly conserved, tending to emerge quickly and evolve swiftly (Ponting et al., 2009). Although transcriptome analyses have uncovered a vast array of lncRNAs, just a tiny fraction of these have an assigned biological function, and fewer still an ascribed molecular mechanism. Little is known about the evolutionary pathways via which lncRNAs gain new functions.

The telomerase RNA subunit TER is a lncRNA and an integral component of the telomerase enzyme. TER functions as template to direct the synthesis of telomeric DNA by the telomerase reverse transcriptase TERT. Telomerase continually synthesizes telomeric DNA in stem and germline cells to avert cellular senescence. However, in cells with limited proliferation programs telomerase activity is repressed, an outcome in vertebrates that may have evolved to avert tumorigenesis (Gunes and Rudolph, 2013; Bernardes de Jesus and Blasco, 2013). Telomerase activity must also be restrained at sites of DNA double-strand breaks to allow faithful DNA repair instead of de novo telomere formation (Kramer and Haber, 1993; Pennaneach et al., 2006). Mechanisms

of telomerase regulation are varied and complex, and include modulation of telomerase localization, recruitment to the telomee and enzymology at the chromosome terminus (Cifuentes-Rojas and Shippen, 2012). Within the telomerase ribonucleoprotein itself, the major target of enzyme regulation is TERT. However, TER is also implicated in telomerase control. In addition, different isoforms of core telomerase components influence telomerase behavior (Karamysheva et al., 2003; Wong et al., 2014).

TER ranges in size from 150 nt in Tetrahymena to >2 kb in certain fungi, and while the nucleotide sequence is highly variable across species, core secondary and tertiary structures are conserved and essential for TER interaction with TERT and for telomerase catalysis (Romero and Blackburn, 1991; Chen et al., 2000; Tzfati et al., 2003; Chappell and Lundblad, 2004; Qi et al., 2013). TER is transcriptionally regulated in mammals (Cairney and Keith, 2008), but the transcript is highly stable with a half-life of several days (Yi et al., 1999). Vertebrate and yeast TERs are Pol II transcribed, acquire a 5' TMG cap and are polyadenylated (Collins, 2006). Recent data show that that 3' terminus of *Schizosaccharomyces pombe* TER is generated by an additional RNA processing step termed slicing, which involves only the first step in mRNA splicing (Box et al., 2008a; Tang et al., 2012). Conventional introns have not been associated with TER.

Arabidopsis thaliana is unusual in that it harbors two *TER* genes, TER1 (784 nt) and TER2 (748 nt) (Cifuentes-Rojas et al., 2011). Within TER1 and TER2, there are two regions of high similarity spanning ~219 nt termed conserved region 1 (CR1) and conserved region 2 (CR2). In TER2, CR1 and CR2 are separated by a 529 nt intervening sequence (IS). An additional unique 36 nts lie at the 3' end of the TER2 CR2. The IS within TER2 is removed *in vivo* to create a truncated isoform of TER called

$TER2_S$ (Cifuentes-Rojas et al., 2012). Sequences flanking the IS do not adhere to consensus splice donor and acceptor sites, suggesting that removal of this element may not proceed via conventional mRNA splicing.

Although the function of TER2s is unclear, TER1 and TER2 play opposing roles in the control of telomerase enzyme activity. TER1 serves as the canonical telomere repeat template necessary for telomere length maintenance *in vivo* (Cifuentes-Rojas et al., 2011). Plants deficient in TER1 exhibit progressive telomere shortening, and mutations in the TER1 template alter the telomere repeat sequence *in vivo*. In contrast, TER2 does not direct telomere repeat incorporation *in vivo*. Instead, this RNA negatively regulates TER1-mediated enzyme activity. Telomerase activity is elevated in plants lacking TER2. Conversely, in plants over-expressing TER2 telomerase activity is decreased and telomeres are shortened (Cifuentes-Rojas et al., 2012).

Under standard growth conditions, the steady state levels of TER1 and TER2s are similar, and 10-20-fold higher than TER2 (Cifuentes-Rojas et al., 2012). However, in response to DNA double-strand breaks (DSBs), TER2 is rapidly induced and becomes the predominant TER isoform. The increase in TER2 is coincident with a reduction in telomerase activity. Telomerase inhibition is dependent on TER2: *ter2* mutants do not down-regulate telomerase in reponse to DNA damage (Cifuentes-Rojas et al., 2012). Telomerase repression is not elicited by replication stress or telomere dysfunction, indicating that TER2-mediated telomerase regulation is specific for DSBs. TERT has a higher affinity for TER2 than for TER1 or TER2s, and preferentially assembles into TER2 containing RNP complexes *in vivo*. Therefore, TER2 is proposed to serve a molecular sponge to sequester TERT in a non-functional RNP *in vivo* (Cifuentes-Rojas et al., 2012).

A search for TER2 orthologs in relatives of *A. thaliana* was unsuccessful (Beilstein et al., 2012). Analysis of sixteen closely related species within the Brassicaceae lineage revealed that these species contain a single locus bearing similarity to the 3' end of TER1 and the 5' end of TER2 from *A. thaliana* (Beilstein et al., 2012), suggesting that TER2 was generated during a massive genome rearrangement that occurred on the branch leading to *A. thaliana* (Samach et al., 2011). The IS within AtTER2 is absent from these loci in Brassicaceae. Moreover, several of these loci lack a template domain altogether, indicating that a functional TER must be encoded elsewhere in the genome. Thus, the TER locus is evolving rapidly in Brassicaceae. The appearance of TER2 and its IS represent very recent events, coinciding with the duplication of TER.

In this study we employ a comparative genomics approach to investigate the regulatory function of TER2. Using data acquired from the 1,001 Arabidopsis genomes project, we show that the IS within TER2 has the characteristics of a solo long terminal repeat (LTR) from a *Copia*-like retrotransposon. The element is associated with most, but not all of the TER2 loci. We report that the unique regulatory function of TER2, including its responsiveness to DSBs, is derived from this transposable element. Consequently, invasion of the TER2 locus by a transposon transformed this lncRNA into a highly sensitive DNA damage sensor that modulates telomerase enzyme activity.

Results

The IS within TER2 is not retained in all A. thaliana *accessions*

Since a clear TER2 ortholog could not be discerned in other members of the Brassicaceae, we analyzed genomic sequence data for *A. thaliana* accessions generated through the 1001 Arabidopsis genomes project (http://signal.salk.edu/atg1001) to learn more about the evolution and function of TER2. TER1 and TER2 loci were retrieved from 853 available accessions and analyzed for variation against Col-0, the primary reference accession and the one in which a regulatory function for TER2 was first described (Cifuentes-Rojas et al., 2012) (Figure 2-1A). The *TER1* locus is highly conserved, including the 5' and 3' regions flanking CR1 and CR2, which lie upstream of the *RAD52* coding region or within a predicted intron (Beilstein et al., 2012; Samach et al., 2011). *TER1* exhibits 99.8% pairwise identity among the sequenced accessions, but a few polymorphisms are scattered across the RNA (Figure 2-1A and 2-1B). The most notable variations lie within the TER1 template domain (Figure 2-2A). A transition of A to C occurred three times while a T-A transversion appeared in 44/853 accessions. In neither instance are the two variations found within the same *TER1* gene. Because the *A. thaliana* TER template is 11 nt in length and encodes one and a half copies of the telomere repeat, these TER1 RNAs retain the potential to direct synthesis of TTTAGGG repeats. More intriguing is C to T mutation in the middle of the template in Bela-1 (Figure 2-2A). Whether this variation reflects a sequencing error or indicates that an alternative TER1 locus is present in this accession is unknown.

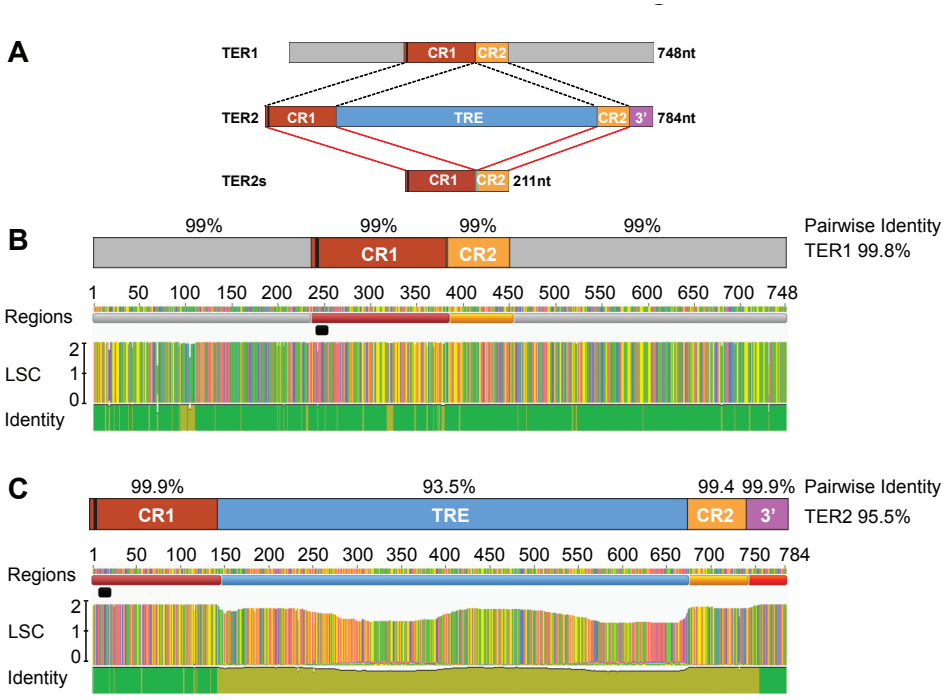


Figure 2-1. Analysis of TER1 and TER2 loci across *A. thaliana* accessions. (**A**) Schematic diagram of TER1, TER2, and $TER2_S$. TER1 and TER2 share a core region of ~219 nt comprised of conserved regions 1 and 2 (CR1 and CR2). The telomere template is denoted by a vertical black bar in CR1. $TER2_S$ is formed by splicing to remove the IS and elimination of the 3' terminus (3' R). (**B**) Analysis of TER1 among 853 *A. thaliana* accessions. Identity shown in green denotes regions 100% conservation whereas mustard yellow indicates variation. Local sequence confidence (LSC) is in a log base 2. LSC of 2 indicates a nucleotide was observed 100% of the time at that location. Reduction in this factor indicates a certain percentage of deletions at this site. A green bar for identity corresponds to a LSC of 2. Pairwise identity for each region is denoted in % above each RNA region or for the entire RNA to the right. The telomere template region is indicated by the horizontal black bar. (**C**) Analysis of TER2 in 853 accessions. Color scheme is the same as in (B). Decreased bar heights indicate the absence of nucleotide sequence in TRE and CR2 of some accessions.

Like *TER1* much of *TER2* is strongly conserved. CR1 retains high identity among the accessions (99.9%) (Figure 2-1C). CR2 and the 3'R are also very highly conserved at 99.4% and 99.9%, respectively. The latter finding was unanticipated since this segment of TER2 is eliminated in the production of $TER2_S$ (Figure 2-1A). Nevertheless, the high degree of conservation in CR1, CR2 and 3'R argues that these regions are important for TER2 function.

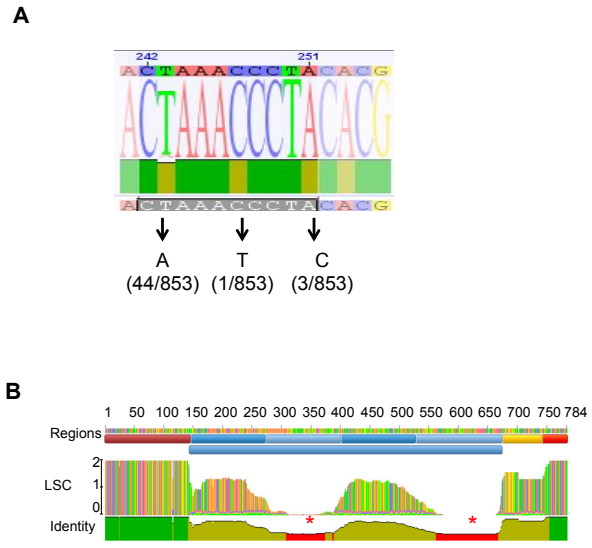


Figure 2-2. Polymorphisms within the TER1 template and TER2 IS alignment across accessions. (A) Screenshot of a Geneious alignment of TER1 template regions amongst 853 *A. thaliana* ecotypes (grey box). Three types of polymorphisms are observed. Arrows denote nucleotide changes and their observed frequency. (B) Alignment of accessions predicted to harbor a partial TER2 IS. Two hyper-variable regions are indicated by red asterisks.

In contrast, the IS within TER2 exhibits remarkable sequence variability. Among *A. thaliana* accessions, two islands of conservation with ≥ 50% identity were identified, one corresponding to 63 nt and a second of 123 nt (Figure 2-1C; Figure 1-2B). Hyper-variable sequences flank these regions within the 65 accessions bearing an incomplete IS (Figure 2-2B). To verify the TER2 sequencing data, we performed PCR genotyping on a sampling of accessions predicted to harbor an intact IS (Col-0, Ws-2), a partial IS (Aa-0, Ang-0, Co-1 and Ei-2) or no IS (Ler-0). PCR primers were positioned within CR1 and 3'R (Figure 2-2A). A 784 bp PCR product is expected for accessions bearing an intact IS, a 255 nt product for accessions completely lacking the IS, and an intermediate size product for accessions with a partial IS. Products of the expected sizes were obtained for loci predicted to contain an intact IS or no IS, but for all TER2 loci predicted to contain a partial IS, the genotyping results indicated IS was completely absent (Figure 2-3B). Genotyping repeated with siblings from accessions predicted to contain a partial IS gave the same result (Figure 2-3C). Genotyping was performed on several additional accessions reported to contain a partial IS. In all cases, the IS was absent. Finally, PCR products were sequenced from TER1 and TER2 reactions, with TER1 polymorphisms serving as a control to ensure that seed stocks were as expected. The sequencing results confirmed the PCR genotyping data (Figure 2-3D). For all partial IS accessions tested, there was complete IS loss. The sequencing data also revealed a substantial deletion (~20 bp) within CR2 in two accessions.

The simplest explanation for these genotyping results is that the TER2 locus was mis-annotated in some of the *A. thaliana* accessions. However, we cannot exclude the possibility that the IS within TER2 is extremely labile and between the time the

genome sequencing was performed and our acquisition of seeds, partially deleted IS elements were completely eliminated.

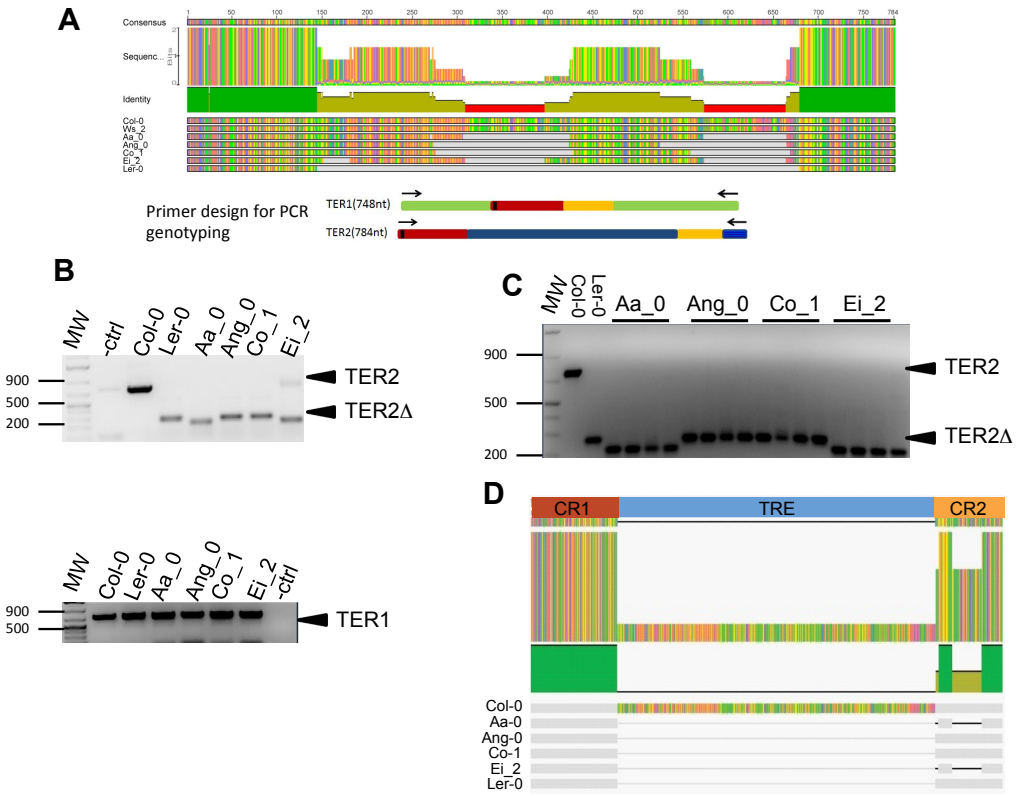


Figure 2-3. Genotyping analysis of *A. thaliana* TER2 loci. (A) Schematic map of the TER2 IS status in different accessions. The positions of PCR primers are indicated by black arrows. (B) Genotyping results for TER1 and TER2 loci in different accessions. TER2 PCR products with the full-length IS are expected to be ~750 bp, and PCR products lacking the IS are ~200 bp. Col-0 was used as a full-length IS control, and Ler-0 used as complete IS loss control. Sequence analysis of all of the TER1 PCR products confirmed accession identity. (C) TER2 intron genotyping results in four siblings of each accession. (D) Sequencing data for TER2 genotyping PCR products in (B). The gap in the IS and CR2 demonstrates sequence loss for these accessions.

The IS bears signatures of a Copia-like solo LTR

For reasons discussed below, we named the IS within TER2 Telomerase Regulatory Element (TRE). BLAST analyses against the *A. thaliana* genome using TRE as a query returned two hits, one on the left arm of chromosome 3 (adjacent to At3G30120) bearing 94.6% identity to TRE_{TER2} termed TRE_{3L} (blue bar, Figure 2-4A), and another on the right arm of chromosome 3 (adjacent to At3G50120) showing 63.4% identity called TRE_{3R} (green bar, Figure 2-4A). Both TRE_{3L} and TRE_{3R} are found within intergenic regions and display a high number of single-nucleotide polymorphisms among accessions.

BLAST was performed to determine if the TRE is present in other species within the Brassicaceae family. *Arabidopsis lyrata* contains 32 copies of TRE dispersed throughout the genome, a significant fraction of which exhibit a high degree of similarity within the 5' 200nt of $TRE_{TER2,}$ and are associated with open reading frames encoding typical retrotransposon proteins (Figure 2-4B and Figure 2-7A). Three TRE elements were also detected in *Capsella rubella,* four in *Brassica rapa*, and ten in *Eutrema salsugineum* (Figure 2-4B). The presence of multiple copies of TRE in *A. thaliana* and its relatives suggested that it is a transposable element (TE).

Consistent with this conclusion, sequences at the 5' and 3' borders of TRE_{TER2} contain a 5 nt tandem inverted repeat (TIR) of TGTTG/ACAAC (Figure 2-4C, brown bar). The TIR at the 5' and 3' boundaries of TRE_{TER2} and TRE_{3L} are highly conserved across the *A. thaliana* accessions and are present at the boundaries of TREs detected in other species. In addition, a target site duplication (TSD) of TCGTC is present at the 3' end of CR1 and the 5' end of CR2 of TER2 (Figure 2-4C, green bar). TSDs flank all three TREs in *A. thaliana*, ranging in length from 5 nt for TRE_{TER2} and TRE_{3L} to 18nt for TRE_{3R}. The TSD sequence varies, consistent with the hypothesis that these insertions represent unique TE insertion events rather than gene duplication. The small size of TRE and its association with TIR and TSD sequences suggest that TRE is derived from a solo LTR of the abundant *Copia* family. Since the large majority of *A. thaliana* accessions apparently harbor an intact TRE within the TER2 locus, it is likely that the element was inserted soon after the TER duplication and was subsequently lost in a small subset of accessions.

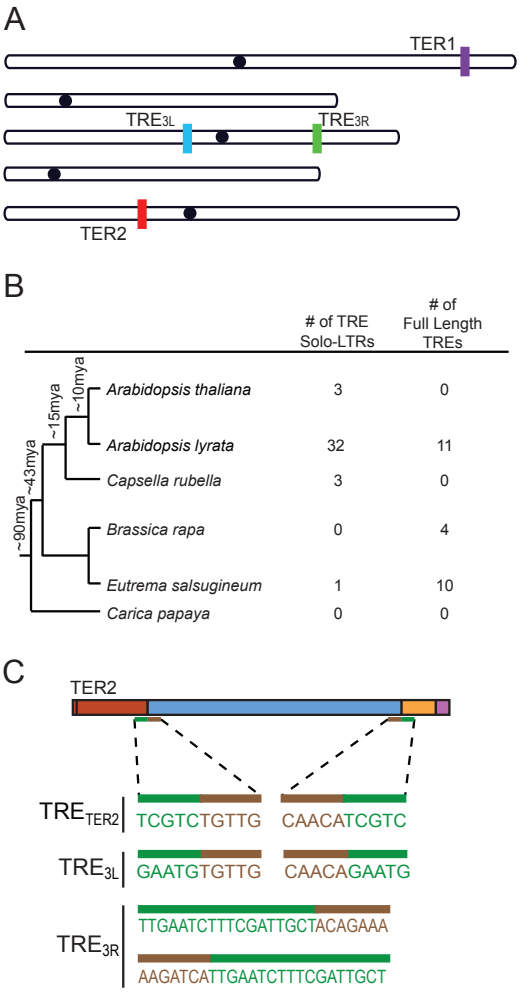


Figure 2-4. The TER2 IS has the properties of a *Copia*-like solo long terminal repeat. (**A**) Schematic of the five chromosomes in *A. thaliana* Col-0 illustrating the locations of TER1, TER2 and Telomerase Regulatory Elements (TRE) on the left arm of chromosome 3 (TRE_{3L}) and the right arm of chromosome 3 (TRE_{3R}) (schematic adapted from TAIR). (**B**) Phylogenetic tree of select Brassicaceae members (including the Brassicales member *Carica papaya*). The number of solo and full-length TREs identified by BLAST are shown to the right. Approximate time of divergence was adapted from [47]. (**C**) Sequences at the 5' and 3' boundary elements of TRE in TER2 (top), TRE_{3L} (middle), and TRE_{3R} (bottom) are shown. Nucleotides within the target site duplication (TSD) are denoted by the green bar and tandem inverted repeats (TIR) of the TRE are represented by the brown bar.

Differential expression of TER2 and TER2Δ

The presence of two distinct TER2 alleles in *A. thaliana* provided us with an opportunity to study the functional impact of TRE. We previously showed that two RNA transcripts are derived from the Col-0 TER2 locus: the primary TER2 transcript and a processed isoform, TER2s, in which TRE_{TER2} is removed along with 3'R (Cifuentes-Rojas et al., 2011; Cifuentes-Rojas et al., 2012). In the Ler-0 accession, the TER2 locus lacks TRE, and thus the primary transcript is predicted to be distinct from that of Col-0. We termed the putative Ler-0 transcript, TER2D. To assay for TER2Δ, RT-PCR was performed on RNA from Ler-0 seedlings using primers directed at CR1 and 3'R, which is unique to TER2 (Figures 2-5A and 2-5B). A product of the expected size was generated (Figure 2-5A), indicating that a Ler-0 transcript containing CR1, CR2 and 3'R is present. Sequence analysis confirmed this conclusion. The CR1/CR2 junction in Ler-0 TER2Δ is distinct from Col-0 TER2s as it contains only a single TSD (Fig 2-5B bottom, underlined sequence). Although a faint signal for TER2 was observed in Col-0 with these PCR conditions, TER2Δ was not (Fig. 2-5A), suggesting that this molecule is either a transient processing intermediate, or is not generated during the conversion of TER2 to TER2s.

Col-0 TER2 is a poorly expressed transcript (Figure 2-5A) and is substantially less abundant than TER1 or TER2s (Cifuentes-Rojas et al., 2012). To assess the relative abundance of Ler-0 TER2Δ, we performed qPCR (Figure 2-5C). The steady state level of TER1 was similar in Ler-0 and Col-0. However, Ler-0 TER2Δ was approximately 6-8 fold more abundant than TER1. By comparison, Col-0 TER2 was 15-20 fold less abundant than TER1 (Figure 2-5C). Thus, Col-0 TER2 and Ler-0 TER2Δ are differentially regulated *in vivo*.

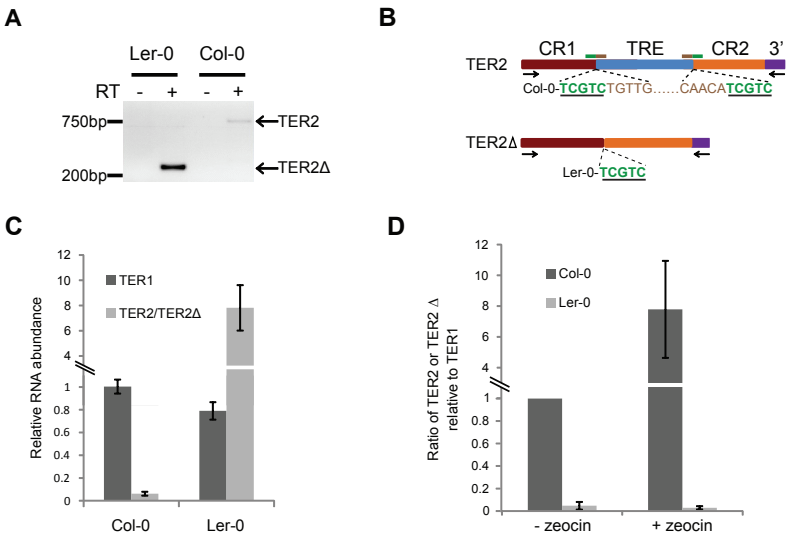


Figure 2-5. Expression of TER2Δ and association with TERT. (**A**) RT-PCR results for TER2Δ in Ler-0 and TER2 in Col-0. Primer positions are indicated by arrows in panel B. (**B**) Schematic showing sequencing results for TER2 and TER2Δ PCR products from Col-0 and Ler-0 obtained from (A). The target site duplication is indicated by the green underlined nucleotides. Tandem inverted repeats are indicated by brown nucleotides. (**C**) qPCR results for TER1, TER2 and TER2Δ in Col-0 and Ler-0. For comparison, the Col-0 TER1 level was set to 1. (**D**) qPCR results following TERT immunoprecipitation in Col-0 and Ler-0 seedling treated with or without zeocin. The TER2:TER1 ratio in Col-0 and the TER2Δ:TER1 ratio in Ler-0 are shown. Values were normalized to Col-0 TER2:TER1 ratio in the absence of zeocin (set to 1).

In Col-0, TER2 but not TER1 or TER2s is rapidly induced by DSBs (Cifuentes-Rojas et al., 2012). Therefore, we asked if Ler-0 TER2Δ abundance changed in response to DSBs. Seven day-old Ler-0 and Col-0 seedlings were treated with 20μM zeocin and qPCR was performed. In control reactions, BRCA1 mRNA was induced in both accessions after 2 hours and peaked at 4 hours, confirming that a DNA damage response was elicited (Figure 2-6A). Furthermore, Ler-0 TER1 levels were unchanged in the presence or absence of zeocin (Data not shown). In contrast, Col-0 TER2 increased 2.5 fold after 2 hours in zeocin relative to untreated seedlings (Figure 2-6B). In contrast, there was no significant change in TER2Δ over the 4 hour zeocin treatment (Figure 2-6B). To test if DSB-mediated regulation of TER2 is a peculiarity of the Col-0 accession, we examined TER2/TER2Δ transcripts in two additional accessions: Ws-2, which contains TRE_{TER2} and Co-1, which lacks it (Figure 2-6D). Consistent with the findings in Ler-0 and Col-0, there was no change in Co-1 TER2D, while Ws-2 TER2 was induced (Figure 2-6D). We conclude that the effect of DSBs on TER2 varies across *A. thaliana* accessions, and correlates with the presence of TRE_{TER2}.

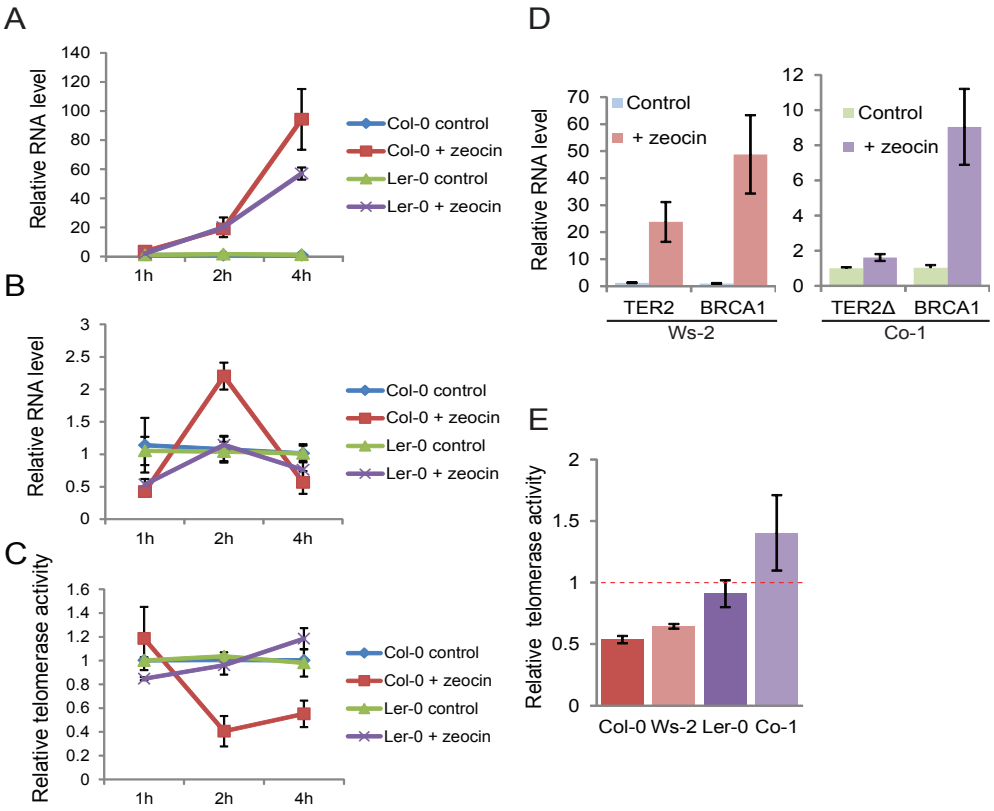


Figure 2-6. DSB-mediated RNA induction and telomerase inhibition are associated with TRE. (**A**) qPCR results for the BRCA1 transcript control. (**B**) qPCR analysis of TER2 (Col-0) or TER2Δ (Ler-0) transcripts from seedlings incubated with or without zeocin for the indicated time points. (**C**) qTRAP results for Col-0 and Ler-0 seedlings with or without zeocin treatment. (**D**) qPCR results for accessions with TER2 (Ws-2) and TER2Δ (Co-1) submitted to the same zeocin regimen as in (**A**). (**E**) qTRAP results for the samples in (**D**). Telomerase activity was normalized to the corresponding untreated controls and set to 1. Red dashed bar indicates no change between treated and untreated samples. The changes in telomerase activity in Col-0 and Ws-2 were statistically significant (p-value< 0.05).

We next asked if transcripts were derived from the other two TRE elements in Col-0, and if so whether they were responsive to DSBs. Based on the alignment (Figure 2-7A), semi-quantitative RT-PCR was performed with primers specific for TRE_{3L} and TRE_{3R} on seedlings in the presence or absence of zeocin. TRE_{3L} transcripts could not be detected under either condition. However, transcripts from TRE_{3R} could be detected in the presence of zeocin (Figure 2-7B). These results suggest that a DNA damage-sensing element may reside within TRE_{TER2} and TRE_{3R}.

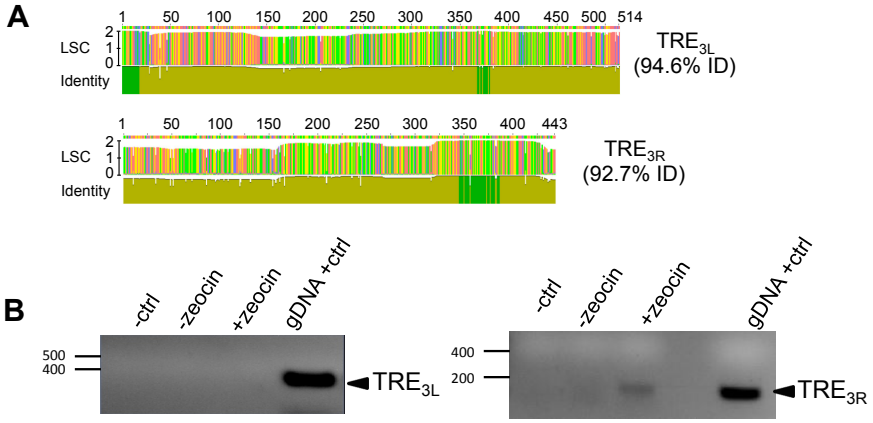


Figure 2-7. Sequence conservation of TRE_{3L} and TRE_{3R} among different *A. thaliana* accessions and their expression in response to DNA damage. (A) TRE_{3L} and TRE_{3R} conservation across *A. thaliana* ecotypes. (B) RT-PCR results for TRE_{3L} and TRE_{3R} with or without zeocin treatment. The expected sizes of the PCR products are highlighted by arrows.

TERT preferentially associates with TER2 over TER2Δ in vivo

In Col-0 TERT displays a hierarchy of binding favoring TER2 > TER1 >> TER2s both *in vitro* and *in vivo* (Cifuentes-Rojas et al., 2012). The higher affinity of TERT for TER2 could reflect the presence of TRE_{TER2} and/or 3'R since both are unique to TER2. To investigate which RNA domain stimulates TERT binding, we examined the relative affinity of TERT for TER2Δ in Ler-0. Col-0 and Ler-0 seedlings were subjected to immunoprecipitation with TERT antibody followed by qPCR (Figure 2-5D). We set the ratio of TER2 to TER1 in the Col-0 TERT IP to 1, and then assessed the change in TERT-bound TER2 following zeocin treatment. The relative abundance of TER2 containing TERT complexes increased ~ 7-fold in response to DNA damage (Figure 2-5D). Since the input level of TER2 increased by only 2.5-fold under these conditions (Figure 2-6B), the data raise the interesting possibility that other DNA damage-induced factors promote TER2 assembly with TERT. In marked contrast to TER2, TER2D is not a preferred binding partner for TERT *in vivo*, and zeocin treatment did not change the relative abundance of TER2D containing TERT complexes (Figure 2-5D). These results indicate that the increased affinity of TERT for TER2 in Col-0 reflects the presence of TRE_{TER2} and not 3'R.

Accessions lacking TRE_{TER2} do not exhibit DSB-induced telomerase inhibition

Quantitative telomere repeat amplification protocol (qTRAP) was used to assess whether Ler-0 modulates telomerase enzyme activity in response to DSBs. Consistent with previous studies (Cifuentes-Rojas et al., 2012), Col-0 seedlings treated with zeocin for 2 hours displayed reduced telomerase activity (70% decrease) compared to untreated seedlings (Figure 2-6C). Although there was a slight alleviation of the

inhibitory effect after 4 hours of treatment, enzyme activity was maintained at 50% of untreated level. Consequently, the net telomerase activity change was inversely proportional to the TER2 transcript level (Figure 2-6B). Under the same treatment regime, telomerase activity was unaltered in Ler-0 (Figure 1-6C). Similar results were obtained with Ws-2 (plus TRE_{TER2}) and Co-1 (minus TRE_{TER2}) accessions, respectively (Figure 1-6E).

We generated two transgenic *A. thaliana* lines to further explore the role of the TRE in telomerase regulation. We first asked if the presence of TER2 was sufficient to alter the level of telomerase activity in Ler-0 by expressing TER2 from its native promoter in this accession. In one of the transformants, the steady state level of transgenic TER2 was higher (2.5 fold) than the basal level of endogenous TER2 in wild type Col-0 (Figure 2-8A). qTRAP revealed a small, but statistically significant decrease in telomerase activity in the transformant (Figure 2-8B), indicating that Ler-0 telomerase can be down-regulated by Col-0 TER2. Next we asked if over-expression of TER2Δ altered telomerase activity in Col-0. TER2Δ expression was driven by the powerful CaMV promoter in wild type Col-0. As expected, there were no change in TER1 or TER2, but the steady state level of transgenic TER2Δ was ~8-fold higher than endogenous TER2D in wild type Ler-0. However, qTRAP showed no change in telomerase activity relative to untransformed Col-0 controls (Figure 2-8A and 2-8B). These data argue that the regulation of telomerase by TER2 is dependent on TRE_{TER2}.

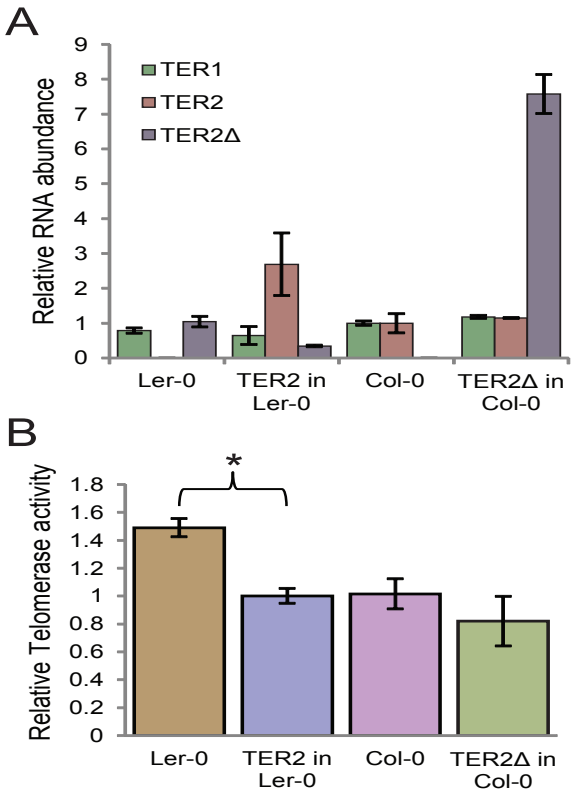


Figure 2-8. TER2 not TER2Δ represses telomerase activity. (A) qPCR results are shown for transgenic seedlings expressing TER2 in Ler-0 or TER2Δ in Col-0. TER1 and TER2 levels were normalized to the values in wild type Col-0 (set to 1). TER2Δ was normalized to the value in wild type Ler-0 (set to 1). GAPDH served as a reference gene. (B) qTRAP results are shown for the seedlings analyzed in (A). Relative telomerase activity was normalized to wild type Col-0. The change in telomerase activity in Ler-0 transformants expressing TER2 relative to wild type Ler-0 is statistically significant (p-value<0.005).

TER2 is an unstable RNA that is stabilized in response to DSBs

The rapid induction of Col-0 TER2 in response to DSBs could occur through increased TER2 transcription or increased RNA stability. Because the sequences upstream of *TER2* genes are highly conserved, we considered the former possibility less likely Indeed, when TER2 transcription was monitored by fusing a GUS reporter to TER2 or TER2Δ promoter (3 kb upstream sequence) in Col-0 and Ler-0 transgenic plants, respectively, approximately the same level of GUS staining was observed in the presence or absence of zeocin (Figure 2-9). Hence, TER2 induction in response to DNA damage is not caused by increased transcription.

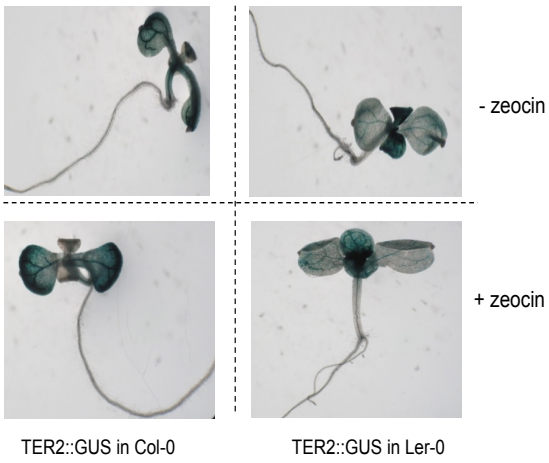


Figure 2-9. TER2 promoter activity in Col-0 and Ler-0. Sequences 3kb upstream of TER2 were cloned into a vector containing the GUS gene as a reporter. The construct was transformed into both Col-0 and Ler-0. Seven day-old seedlings were treated with zeocin for 2 h and then tested for GUS activity.

We assessed TER2 stability using six day-old seedlings treated with the transcription elongation inhibitor cordycepin. TER1 and TER2 RNA levels assessed by qPCR showed that Col-0 and Ler-0 TER1 have similar half-lives, $t_{1/2}$ = 75 and 84 min, respectively (Figure 2-10A). The stability of TER2Δ was even greater with $t_{1/2}$ = 244 min (Figure 2-10B). TER2, on the other hand, had a much shorter half-life than either TER2Δ or TER1: TER2 $t_{1/2}$ = 13 min (Figure 2-10B). Thus, TER2 is an intrinsically unstable transcript.

To test if DSBs reduce TER2 turnover, Col-0 seedlings were treated with cordycepin to pause transcription and then zeocin was added after 90 min to produce DSBs. Although there was a slight change in the abundance of TER1 and BRCA1 mRNA in the presence of zeocin, this change was not statistically significant (Figure 2-10C and 2-10D). In marked contrast, TER2 abundance declined sharply over the 3.5 hour time course, but immediately after the introduction of zeocin, TER2 was stabilized (Figure 2-10E). These data implicate TRE_{TER2} as the causal factor in destabilizing TER2 and in turn negatively regulating telomerase activity during bouts of DNA damage.

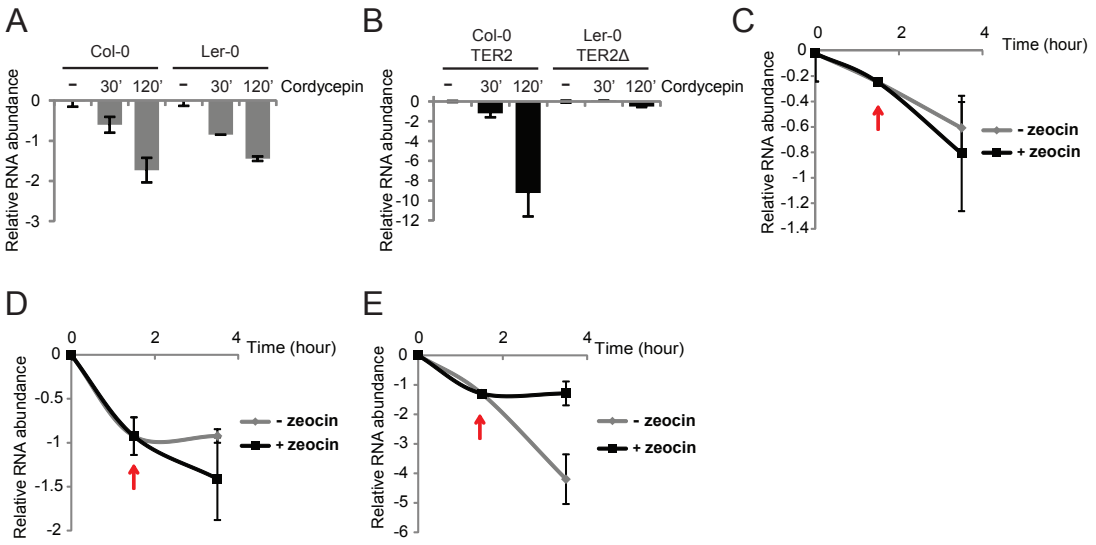


Figure 2-10. TER2 is a labile RNA transcript that is stabilized in response to DNA damage. qPCR results are shown for TER1 and TER2/TER2Δ from Col-0 and Ler-0 in the presence of cordycepin. Col-0 and Ler-0 seedlings were treated with cordycepin (100ug/ul) for the times indicated followed by qPCR to monitor TER1 (**A**) and TER2/ TER2Δ (**B**). The values obtained for untreated RNA samples were set to 0 and the fold decrease is shown. eIF-4a was used as reference gene for normalization. (**C-E**) qPCR results from a time course experiment of Col-0 seedlings treated with cordycepin followed by zeocin. Seedlings were incubated with cordycepin for 1.5 h to shut down transcription, and zeocin was added (red arrows). The incubation continued for 3.5 h. Results for BRCA1 (**C**), TER1 (**D**), TER2 (**E**) are shown.

Discussion

When the insertion of a TE within or adjacent to a gene leads to a change in gene function the process is termed “exaptation” (de Souza et al., 2013). Exaptation can alter gene regulation through myriad different mechanisms. A prominent example in the plants is the multiple TEs inserted adjacent to *teosinte branched1* (tb1) which gave rise to domesticated maize (Studer et al., 2011). One of the TEs disrupts a regulatory region of tb1, leading to increased expression and enhanced apical dominance. In vertebrates, exaptation is more prevalent at lncRNA loci than in protein-coding genes (Sela et al., 2010). Approximately 41% of vertebrate lncRNA sequence is derived from TEs (Kelley and Rinn, 2012; Johnson and Guigo, 2014), leading Johnson and Guigo to propose that TEs can behave as pre-formed functional RNA domains, and further that exaptation is a major driving force in lncRNA evolution (Johnson and Guigo, 2014). A recent systematic survey in vertebrates catalogued multiple instances of TEs altering lncRNA promoters, splice sites, and polyadenylation sites (Kapusta et al., 2013). LncRNAs can also acquire novel interaction partners as a direct result of exaptation (de Souza et al., 2013). For instance, TEs within XIST facilitate interaction with a host of protein complexes including PRC2 and splicing factor ASF2 (Wutz et al., 2002; Jeon and Lee, 2011).

Here we show that invasion of TRE into the TER2 locus in *A. thaliana* profoundly altered the function of this lncRNA. This exaptation event does not appear to be fixed, as the TER2 genes in 9% of the 853 we accessions examined lack TRE. However, by exploiting this genetic heterogeneity, we discovered that the unique functions ascribed to TER2, its responsiveness to DNA damage, ability to inhibit telomerase activity, and enhanced association with TERT *in vivo*, all derive from this element (Figure 2-11).

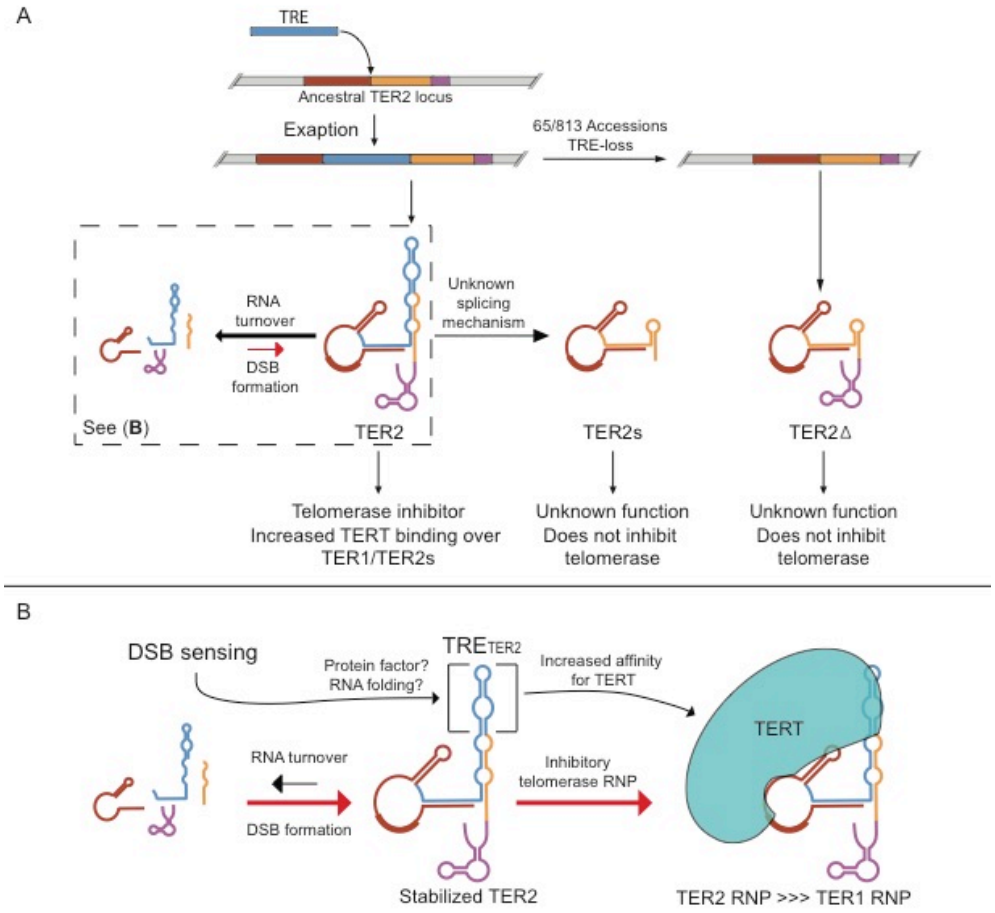


Figure 2-11. Model for exaptation of TER2 and the emergence of a telomerase regulatory lncRNA. (A) Evolution of TER2. Duplication of the single copy ancestral TER gene was followed soon thereafter by exaptation of a telomerase regulatory element (TRE) into the TER2 locus of *A. thaliana*. The majority of *A. thaliana* accessions retain TRE (e.g. Col-0), but a subset have lost it (Ler-0). The TER2 transcript is unstable, and in Col-0 it can be processed to generate TER2s. TER2Δ is produced by accessions lacking TRE. TER2Δ is produced by accessions lacking TRE. Unlike TER2, TER2Δ is a highly stable RNA and cannot regulate telomerase enzyme activity. The functions of TER2s and TER2Δ are unknown. (**B**) Regulation of TER2 stability and telomerase activity by TRE. TRE impacts TER2 function in several ways. First, TRE acts a post-transcriptional DNA damage sensor, destabilizing TER2 under normal physiological conditions, and rapidly stabilizing the RNA in response to DSBs. Second, TRE increases the affinity of TER2 for TERT. In response to DSBs, TER2 accumulates in TERT containing complexes *in vivo*. TER2-mediated telomerase inhibition may reflect competitive inhibition of TER1 for TERT.

First, we find that TRE destabilizes TER2. A survey of ~800 lncRNAs in mouse revealed that only a small fraction are unstable, defined as RNAs with a half-life of less than 60 minutes (Clark et al., 2012). By this criterion, TER2 is a highly unstable transcript with a half-life of only 13 minutes (Figure 2-10A). TER1 (t1/2 =80 min) and TER2D (t1/2=240 min), on the other hand, are categorized as stable RNAs. Unstable lncRNAs, like their unstable mRNA counterparts, are typically associated with regulatory functions, while stable RNAs are thought to serve housekeeping roles (Schwanhausser et al., 2011). With Col-0 *A. thaliana* TER1 and TER2, this paradigm also holds.

A second key observation is that the instability of TER2 arising from TRE is reversed in response to DNA damage (Figure 2-11B). The abundance of TER2, but not TER1 or TER2Δ is elevated in response to DNA damage, and this change is largely, if not entirely, dependent on RNA stabilization rather than new transcription. Exaptation is known to endow host genes with the capacity to respond to environmental cues. For example, a cold-sensitive TE was inserted into the promoter of Ruby, a transcription factor that regulates flesh color in *Citrus sinensis* (blood orange). Cold activates the transposon, which in turn activates Ruby and downstream anthocyanin production (Butelli et al., 2012). In the case of TER2, TRE appears to harbor a DNA damage sensing element that leads to RNA stabilization. Both TRE_{TER2} and TRE_{3R} are induced in response to zeocin treatment, but not TER1 or TER2D.

How TER2 is stabilized in response to DSBs is unknown. One possibility is that TRE_{TER2} carries binding sites for one or more interaction partners responsive to DNA damage, which then stabilize TER2 (Figure 2-11B). Recent work indicates that RNA binding proteins can play a significant role in the DNA damage response by regulating

specific target genes post-transcriptionally (Dutertre et al., 2014). TER2 turnover might also be controlled through the small RNA regulatory pathway. A 24 nt RNA is associated with TRE_{TER2} (Rajagopalan et al., 2006). This finding is particularly intriguing given the recent discovery that small RNAs modulate the response to DSBs in both vertebrates and Arabidopsis (Wei et al., 2012).

A final key observation from this work is that TRE increases the affinity of TER2 for TERT *in vivo* (Figure 2-11B). The presence of TRE may modify TER2 structure in a manner that enhances its inherit affinity for TERT. Alternatively, given the capacity of TEs to serve as discrete functional domains, TRE may independently contact TERT to increase TER2 affinity. Unexpectedly, we also found that zeocin treatment led to an even greater enrichment of TER2 containing TERT complexes than expected based on the fold induction of TER2. This observation raises the interesting possibility that a TER2 associated factor (e.g. POT1b or Ku) is also responsive to DNA damage and drives the assembly of TER2-TERT RNPs.

Altogether, our data are consistent with a model in which the insertion of a TE into the TER2 locus of *A. thaliana* gave rise to a new mode of telomerase regulation. Specifically, we propose that exaptation of TER2 by TRE converted this lncRNA into a DNA damage sensor that controls telomerase enzyme activity through sequestration of TERT. Furthermore, because this regulatory pathway is controlled by changes in RNA stability, it is both rapidly responsive and reversible, allowing *A. thaliana* to fine-tune telomerase activity during changing environmental conditions. These discoveries provide a fresh perspective on the role of exaptation in shaping lncRNA function and evolution.

Materials and methods

Plant material, growth conditions and transformation

For experiments with seedlings, seeds from different accessions (Col-0, Ler-0, Ws-2, etc) were sterilized in 50% bleach with 0.1% Triton X-100 and then stored in 4°C for 2-4 days. Liquid Murashige and Skoog (MS) medium were used for germination and growing. After transferring cold-treated seeds to MS, plants were grown at 22°C under long day light condition for ~7 days. The TER2 gene including 3kb upstream sequence and 300bp downstream sequence was cloned in the pMDC99 vector for transformation in the Ler-0 background. Hygromycin MS plates were used for selection. For Col-0 transformation, TER2Δ together with 300 bp downstream flanking region was cloned into the pBA002 vector with 35S promoter. BASTA MS plates were used for the selection.

Sequence acquisition and analysis

Sequences corresponding to TER2 (Genbank accession number: HQ401285.1) were obtained using the genome browser at http://signal.salk.edu/atg1001. The search query AT5G24660 was used to pinpoint the region of interest, and all available tracks (accessions) were selected. Two sequences were removed from our analysis. Hov 3-2 was removed because it was the only accession with two deletions in the 5' end, corresponding to 20 nt from the 5' start of TER2, and a 100 nt deletion starting at nucleotide #101. The template region was not disturbed in this accession, possibly indicating a functional TER2 is generated. The Tottarp-2 accession was removed because the sequence corresponding to our search region did not contain sequences corresponding to TER2, most importantly, a template region.

Sequences were trimmed in MEGA5, and then analyzed using Geneious v6.0 (Biomatters). Sequence conservation and alignments were performed using Geneious. IS-like sequences were obtained by BLAST searches of the *A. thaliana* (www.arabidopsis.org), *A. lyrata*, *Capsella rubella*, *Brassica rapa*, and *Thellungiella halophila* genomes accessed via www.phytozome.net v9.1 (Hu et al., 2011); (Cheng et al., 2011).

DNA damage treatment and assays

A. thaliana seedlings (5-7 day old) were transferred to fresh MS liquid medium with 20 µM zeocin (Invitrogen) as described (Cifuentes-Rojas et al., 2012). Seedlings were kept in the dark with gentle agitation for 1 h, 2 h or 4 h. Multiple seedlings were combined and flash frozen in liquid nitrogen for RNA extraction or protein extraction for TRAP. The combined sample was treated as a single biological replicate.

Nucleic acid extraction, genotyping and PCR

DNA samples were prepared from the leaves of different accessions. Both TER1 and TER2 loci were used for genotyping. PCR samples were resolved in 1% agrose and gel purified and sequenced. RNA was extracted from seedlings using the Direct-zol RNA MiniPrep kit (Zymo Research, Epigenetics) according to the manufacturer's instructions. 1 µg total RNA was used for preparing cDNA. For RT-PCR, cDNA was synthesized by SuperscriptIII Reverse Transcriptase (Invitrogen). For qRT-PCR, reverse transcription was performed using the Superscript cDNA master mix (Quanta), according to the manufacturer's instructions. 1:5 diluted cDNA was used for qPCR. qPCR was performed on a Bio-Rad CFX-1000 using the following primers: qTER2Δ F:

5’-AGAACGTTGACGGCTAAAGG-3’; qTER2Δ R: 5’- TGTGGCATAAGGCAAACTGA-3’; TER2, BRCA1, TER1 and GAPDH primers are used as described before (Cifuentes-Rojas et al., 2012). Data were analyzed using Bio-Rad’s CFX manager software. ΔΔCT values were obtained by comparing against GAPDH levels.

qTRAP and Immunoprecipitation (IP) qRT-PCR

qTRAP assays were performed as described (Kannan et al., 2008). Data were normalized against untreated Col-0. For immunoprecipitation, TERT antibody (Kannan et al., 2008) was conjugated with Dynabeads Protein A (Invitrogen) then incubated protein extracts in 4°C. RNA was recovered from the IP sample using phenol/chloroform followed by ethanol precipitation. qPCR was performed on TER1 and TER2/TER2Δ. The ΔCT value was used to determine the relative level of TER2 or TER2Δ against TER1.

RNA stability assays

5-7 day old seedlings were treated with cordycepin (100 ng/ml as a working concentration) for 2 h before RNA extraction. RNA was analyzed by qPCR normalized to eIF-4a (Golisz et al., 2013). RNA abundance was converted to the decreased level relative to untreated. RNA half-life was determined by the absolute value of inverse of the slope of the equation plotted by untreated and treated data. For the combined cordycepin/zeocin experiment, seedlings were pre-incubated with cordycepin for 1.5 h followed by zeocin and the incubation was continued for 2 h. RNA extraction and qPCR were used to determine RNA abundance.

GUS staining

3 kb of sequence upstream of the TER2 5' terminus was cloned in a GUS reporter vector pMDC163. The construct was transformed into *A. thaliana* Col-0 and Ler-0 as described (Zhang et al., 2006). After selection in hygromycin, transformants seedlings were treated with zeocin for 2 h and then subjected to GUS histochemical staining as described (Pecinka et al., 2009).

CHAPTER III

CHARACTERIZATION OF TER2 IN REPRODUCTIVE DEVELOPMENT

Summary

To increase genetic diversity, genetic recombination occurs between non-sister chromatids during meiosis. An integral step in this process is the introduction of DNA double strand breaks (DSBs). These DSBs must be hidden from telomerase to prevent *de novo* telomere formation, which could lead to the loss of genetic information. In *Arabidopsis,* an alternative telomerase RNA (TER2) has evolved to negatively regulate telomerase activity in response to DSBs. Here we investigated a potential role for TER2 in plant reproduction. We show that the abundance of TER2 is developmentally regulated, peaking in flower buds and gradually decreasing during flower maturation. Plants lacking TER2 showed decreased seed production with shorter siliques and evidence of seed abortion. Furthermore, pollen viability was reduced relative to wild type plants, indicating the TER2 gene may play a role in meiosis. Previous results indicate a transposable element is embedded in the TER2 gene from some *A. thaliana* accessions. In an accession lacking TE at the TER2 locus, there was no reproductive defect. This finding suggests that the TER2 TE has evolved an important function in reproduction. Finally, we discovered that the seed abortion phenotype was worse in plants doubly deficient in TER2 and POT1a, a telomerase processivity factor required for telomere maintenance. Thus TER2 function during reproduction does not appear to be mediated through conventional telomere maintenance pathways. These data provide a striking example for how exaptation, integration of a transposon, is shaping chromosome biology in Arabidopsis.

Introduction

Telomeres are replicated by telomerase, a reverse transcriptase containing two core subunits: the catalytic moiety TERT and a templating RNA, TER. The main function of telomerase is to replenish DNA lost due to semi-conservative DNA replication. Insufficient telomerase activity caused by mutation of TER leads to a variety of stem cell diseases including dyskeratosis congenita (DC) (Vulliamy et al., 2001). On the other hand, telomerase enzyme activity must be tightly controlled during development, because unrestrained telomerase activity is a hallmark of cancer (Armanios and Greider, 2005).

Another important aspect of telomerase regulation is to prevent it from adding telomere sequence at sites of DNA breaks. This process, termed *de novo* telomere formation (DNTF), causes the permanent loss of the acentromeric DNA fragments, though it transiently appeases the DNA damage signal and stabilizes the broken chromosome ends (Pennaneach et al., 2006). The programmed induction of DNA double strands breaks (DSBs) is required to initiate genetic information exchange between non-sister chromatids during meiosis. At the beginning of this process, the transesterase enzyme Spo11 attacks the phosphodiester bonds of the DNA causing DSB formation. SPO11 is covalently linked to the 5' ends of DNA cutting site as a transient stabilizing structure. Spo11 generated DSBs are mostly repaired via a homologous recombination based mechanism, which then stimulates homologous chromosome pairing (Mahadevaiah et al., 2001; Kleckner, 1996; Cole et al., 2010). However, several studies reveal evidence of a SPO11 independent pathway to initiate meiotic recombination. Such a pathway would require an alternative means to protect site of DNA breaks (Farah et al., 2005; Storlazzi et al., 2003). Recent data from mouse

indicates that there are foci of proteins involved in the processing of DNA damage in the *spo11* knockout meiocytes. The number of these repair foci increases during oocyte development, consistent with the induction of S-phase independent *de novo* DNA damage (Carofiglio et al., 2013). Therefore, there may be a need to decrease telomerase activity during the gamete formation. This notion is supported by studies in mammals showing finely regulated telomerase activity in germ line cells (Riou et al., 2005; Tanemura et al., 2005; Wright et al., 1996).

Telomerase functions as a large RNP complex *in vivo,* although TERT and TER are sufficient to reconstitute telomerase activity *in vitro.* Unlike other model organisms, *Arabidopsis thaliana* has duplicated several telomerase components. Specifically, gene duplication of TER and the telomerase accessory proteins POT1 has given rise to three distinct telomerase RNPs in plants (Cifuentes-Rojas et al., 2011; Cifuentes-Rojas et al., 2012). The TER1 RNP, with TER1 as the scaffold for its specific binding partners POT1a, dyskerin, and TERT, is the canonical telomerase needed to maintain telomere length. The TER1 paralog TER2 forms an alternative TER2 RNP with POT1b, Ku, dyskerin and TERT. The TER2 RNP is a negative regulator of telomerase that is induced in response to DSBs (Cifuentes-Rojas et al., 2012).

TER2 contains a unique 529nt intervening sequence (IS) that disrupts two conserved regions shared between TER1 and TER2 (Cifuentes-Rojas et al., 2011). Removal of the IS as well as cleavage of 36nt from the TER2 3' terminus leads to production of a third TER isoform, TER2s. The TER2s RNP is composed of Ku and POT1b, but not TERT. The function of the TER2s RNP is unknown (Cifuentes-Rojas et al., 2012). Notably, the IS is missing from the TER2 locus in some *Arabidopsis thaliana* accessions (Xu et al, in revision). Comparative bioinformatics revealed multiple copies

of a similar IS element in *A. thaliana* and its relatives. The IS contains sequence signatures of a solo LTR from a *Copia-like* retrotransposon (Xu et al, in revision), and thus appears to be a transposable element (TE). Invasion of the TE (exaptation) into the TER2 locus endowed several unique features to this RNA. Compared to the TE minus transcript TER2Δ (Ler-0), TE embedment significantly increases TER2 affinity for TERT and greatly decrease its RNA stability (Xu et al, in revision). Furthermore, the TE serves as a DNA damage response element. In the presence of DSBs, TER2 becomes stabilized, and the increasing TER2 stability leads to the down regulation of telomerase activity. Finally, overexpression of TER2Δ (lacking the TE) does not decrease telomerase activity. Thus, evolution of TER2 shaped by a TE gives rise to a new strategy for telomerase regulation in Arabidopsis.

In addition to TER, POT1 is duplicated in *A. thaliana*. POT1a acts in the same genetic pathway as TERT to maintain the telomere length (Surovtseva et al., 2007). Recent data show that POT1a stimulates telomerase activity by promoting telomerase repeat addition processivity (Renfrew et al., 2014). Interestingly, plants lacking both POT1a and TER2 undergo faster telomere shortening than *pot1a* single mutants (Kyle Renfrew, unpublished data). This finding reveals an unexpected synergistic role for POT1a and TER2 in telomere maintenance and telomerase regulation.

In this study, we investigate the function of TER2 during reproductive development. We show that TER2 levels peak in the flower buds and decrease during embryogenesis. However, TER2D does not exhibit the same expression pattern, and further show that TER2 regulation in flower development is dependent on its TE. We also demonstrate a role for TER2 in pollen viability and seed production. These findings indicate that TER2 plays a role in *A. thaliana* reproduction. Finally, we discovered that

pot1a ter2 double mutants have an exacerbated seed abortion phenotype. This observation is consistent with a synergistic role for TER2 and POT1a, and argues that TER2 acts independently of the canonical telomerase RNA, perhaps as a signaling lncRNA to promote reproductive development.

Results

TER2 accumulates in the flower bud of the A. thaliana *Col-0 accession*

To test the idea that TER2 negatively regulates telomerase at DSBs formed during meiosis, we firstly asked if TER2 abundance fluctuates during plant reproduction. We predicted that TER2 would peak in meiocytes in response to DSBs. A time course quantitative RT-PCR experiment was performed to monitor TER2 abundance during floral development. (Figure 3-1A). Male and female meiocytes are formed and fertilization takes place upon pollination in the closed flower bud. During the days after pollination (DAP), flowers gradually open and form fruits called siliques containing the zygotes (seeds). In wild type Col-0, TER2 abundance was greatest immediately after pollination and gradually decreased during flower development (Figure 3-1B). This result supports the hypothesis that TER2 functions during meiosis, and furthermore that it is induced by the programmed DSBs introduced during meiotic prophase I.

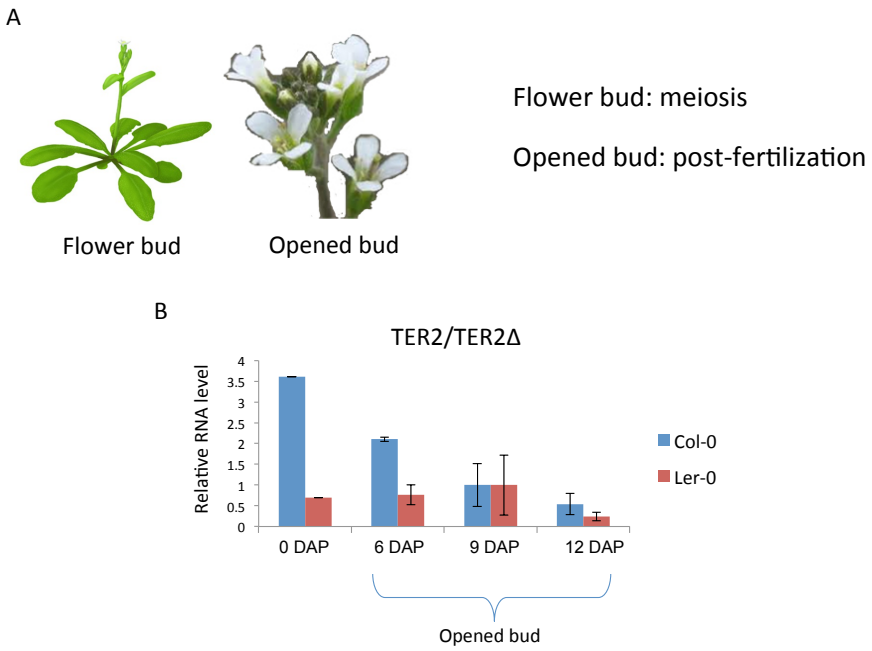


Figure 3-1. Regulation of TER2 during flower development. (A) Two flower stages: Flower bud, where meiosis occurs; Opened bud, where meiosis is completed. (B) RNA was analyzed from flowers at the indicated days after pollination (DAP) qRT-PCR for TER2 (Col-0) or TER2Δ (Ler-0) was normalized to GAPDH. Results from three biological replicates are shown. TER2 abundance at 9 DAP was set to 1 for comparison.

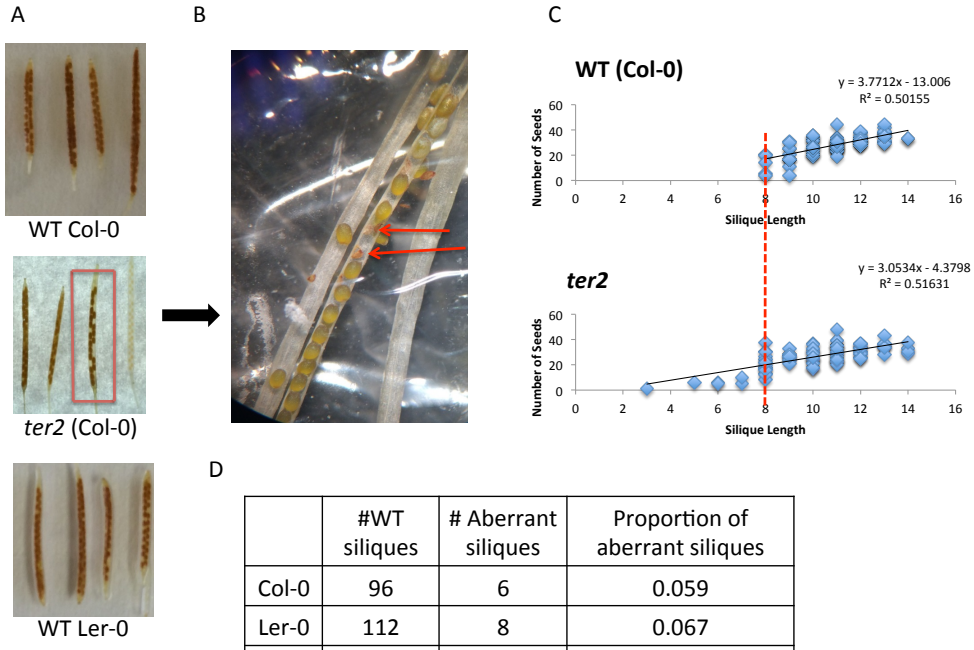
TER2 and TER2Δ are regulated differentially in the Col-0 and Ler-0 accessions

We performed the same experiment in an *A. thaliana* accession lacking the TE in the TER2 locus (Ler-0). If the regulation of TER2 in reproductive development reflects the presence of the TE, we do not expect a change in TER2Δ, the primary TER2-like transcript in Ler-0. qRT-PCR data for TER2Δ showed no significant change during the time course, except a slight decline 12 DAP (Figure 3-1B). This result indicates that the regulation of TER2 during meiosis correlates with the presence of TE.

The frequency of seed abortion is increased in ter2 mutants

If TER2 plays a role in plant reproduction, we expect to see reduced fertility in plants lacking TER2. To test this hypothesis, we examined siliques in *ter2* mutants. In wild type plants, two rows of seeds are aligned parallel to each other giving rise to around 40 embryos (Figure 3-2A, Col-0). However, the siliques of *ter2* mutants showed gaps, indicating no embryo formed or its development was aborted (Figure 3-2A, *ter2*). The seed abortion phenotype was verified by microscopy after opening silique (Figure3-2B, red arrow). Silique length is correlated with seed number. As expected, *ter2* mutants had shorter silique length and fewer seeds comparing to wild type Col-0 (Figure 3C, population on the left of the red lines). These results support the conclusion that TER2 is required for reproductive development, and suggest that *ter2* mutants are defective in meiosis or embryogenesis.

Silique analysis performed with Ler-0 showed no obvious defect in seed formation (Figure 3-2A, Ler-0; Figure 3-2D). Taken together, these data indicate that Col-0 relies on TER2 or a regulation of this RNA for reproductive fitness. The data also indicate that this is an acquired function related to exaptation.



D

	#WT siliques	# Aberrant siliques	Proportion of aberrant siliques
Col-0	96	6	0.059
Ler-0	112	8	0.067
ter2	89	28	0.239

Figure 3-2. Seed abortion in Col-0 *ter2* mutants, but not in wild type Ler-0. (A) Silique analysis. Photos of representative siliques from different genetic backgrounds. The red box highlights seed loss. (B) Analysis of seed abortion by light microscopy. Opened siliques were observed in an optical microscope (4X). Aborted seeds are indicated by red arrows. (C) Seed formation analysis. Seeds were counted in mature siliques. The length of siliques was plotted relative to seed number. Shorter siliques with fewer seeds were observed in *ter2* mutants (left of red line). (D) Summary of the seed abortion frequency. Siliques showing at least 3 empty positions were counted as aberrant.

Pollen viability is reduced in ter2 mutants

To explore if and how TER2 might affect meiosis, we checked the viability of male meiocytes by analyzing pollen with the vital stain fluorescent diacetate (FDA). An esterase in pollen digests the FDA to release a fluorescent signal. The intensity of fluorescence serves as a indicator of pollen fitness (Pinillos and Cuevas, 2008). Pollen grains were freshly prepared in acetone and directly observed under the microscope immediately after the addition of FDA. As expected, plants lacking telomerase reverse transcriptase (TERT) for several generations showed reduced pollen viability (Figure 3-3A, white arrow) (Riha et al., 2001). In fifth generation (G5) *tert* mutants, three distinct types of pollen were observed. Wild type pollen grains were round and the intensity of signal was bright. Type I pollen were also round, but showed decreased signal intensity. Type II pollen were rod-like and the fluorescent signal was undetectable (Figure 3-3B, left). The *ter2* mutant showed an intermediate level of pollen viability between wild type Col-0 and *G5 tert* mutants, with an increasing ratio of type I pollen and appearance of the type II pollen. This result indicates that TER2 contibutes to male meiosis.

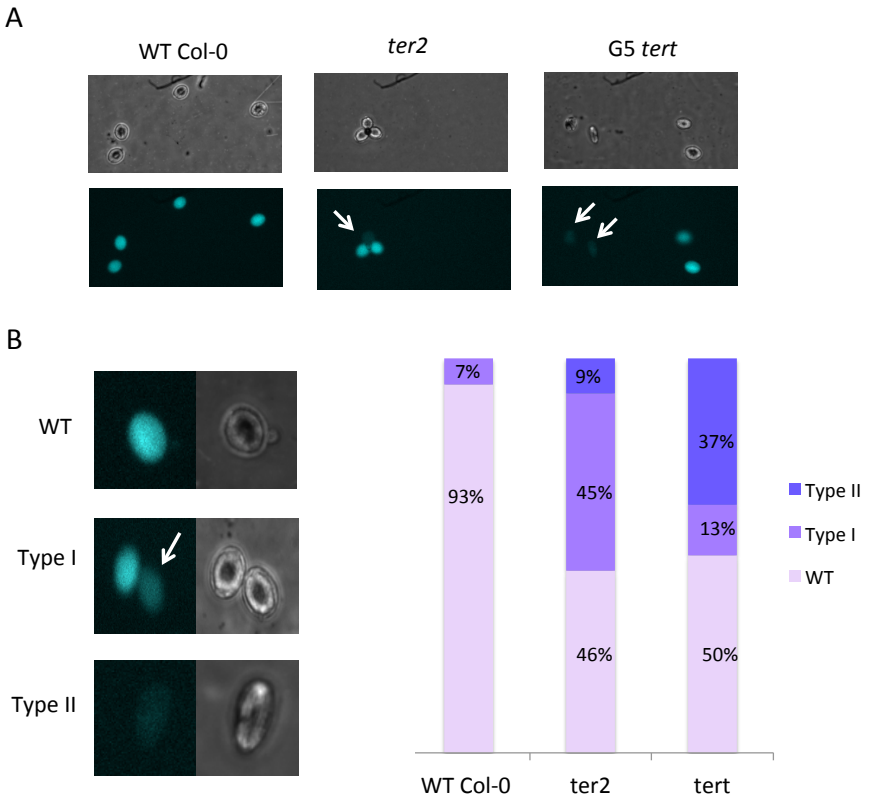


Figure 3-3. Reduced pollen viability in *ter2* mutants. (A) Pollen viability assays in wild type Col-0, *ter2,* and G5 *tert.* Freshly collected pollen grains were incubated in Fluorescent diacetate (FDA) solution and viewed by fluorescent microscopy. Viable pollen produce a blue fluorescent signal using FDA as a substrate. Top, Normarski. Bottom, CFP channel. Decreased intensity of the FDA signal is observed in *ter2* (white arrows) as well as in G5 *tert* . (B) Left, pollen phenotypes. Wild type, round pollen grain that fluorescence bright blue. Type I, round pollen with decreased fluorescence intensity. Type II, rod shaped pollen with little to no fluorescence. Right, summary of phenotypes.

Loss of POT1a exacerbates the ter2 reproductive phenotype

Unpublished results from former Shippen lab graduate student Kyle Renfrew showed that POT1a and TER2 act synergistically to maintain telomeres. Double *pot1a ter2* mutants have much shorter telomeres than pot1 single mutants. Therefore, we asked if the two genes cooperate for reproductive development by examining the siliques in *pot1a ter2* double mutants. Firstly, we confirmed the seed abortion phenotype of the *ter2* mutant in progeny segregated from the double heterozygous parents (equal to the effect of back cross to wild type). The frequency of seed abortion was about same as seen in the previous *ter2* analysis, arguing that the reproductive phenotype is contributed solely by TER2 gene (Figure 3-4 B and C, yellow). F2 *pot1a* mutants did not show a significant seed abortion phenotype, consistent with the fact that their telomeres remain in the wild type range. For *pot1a ter2* double mutants, two types of plant morphologies were observed. Class I plants were slightly shorter height but had a wild type appearance, while Class II were dwarf and had reduced growth robustness (Kyle Renfrew, unpublished data).

Strikingly, the seed abortion ratios for class I and class II *pot1a ter2* double mutant plants were significantly increased (Figure 3-4 B and C, blue). Class I has additive seed abortion ratio of *pot1a* single and *ter2* single mutants, while class II has a even higher ratio of seed abortion than class I. This finding indicates that TER2 acts synergistically with POT1a to promote fertility. It also suggests that TER2 affects reproductive fitness independent of telomere length. In the F3 generation, *pot1a* single mutants began to show a seed abortion phenotype (Figure 3-4C), likely caused by shortened telomeres (Riha et al, Science). Notably, the severity of the phenotype was exacerbated in F3 *pot1 ter2* double mutants, while TER2 still worked synergistically with

POT1a causing the seed abortion (Figure 3-4 B and C, red). These results argue that TER2 directly regulates reproductive fitness, and further that this role is independent of telomere length and telomerase activity status.

Discussion

The TER2 locus is evolving rapidly in *A. thaliana* thanks in part to exaptation of a transposable element. Initial analysis of *ter2* null mutants revealed no obvious phenotypes except the failure to repress telomerase activity upon artificial introduction of DNA damage. However, careful analysis showed a slight increase in the sensitivity of the root apical meristem to DNA damage (Cifuentes-Rojas et al., 2012). The relatively mild phenotype might suggest that TER2 has no important biological role. However in this study, we uncovered a novel function for TER2 in plant reproductive development coupled with its previously defined function in telomerase regulation and DNA damage response; our findings begin to paint a picture of a emergent lncRNA that is rapidly acquiring new functions in *A. thaliana* biology.

A

P: *ter2 -/-* **X** *pot1a -/-*

F1: *pot1a +/- ter2 +/-*

F2: *pot1a -/-* *ter2 -/-* *pot1a -/- ter2 -/-*

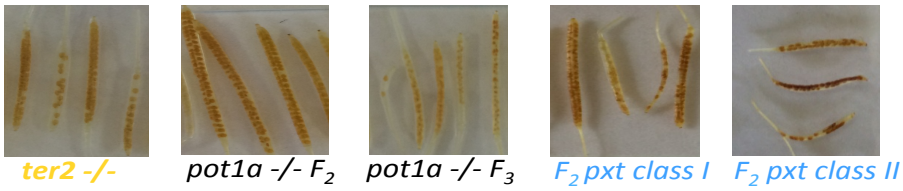
Type I *Type II*

F3: *pot1a -/-* *pot1a -/-ter2 -/-* *pot1a -/-ter2 -/-*

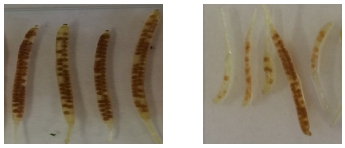
Type I: wilt type morphology

Type II : dwaft plant, much shorter telomeres

B



ter2 -/- *pot1a -/- F_2* *pot1a -/- F_3* *F_2 pxt class I* *F_2 pxt class II*



F_3 pxt class I *F_3 pxt class II*

C

	Normal	Aberrant	Proportion of aberrant
Col-0	96	6	0.059
ter2 -/-	63	32	0.337
pot1a -/- F_2	47	5	0.096
F_2 pxt class I	102	30	0.227
F_2 pxt class II	55	45	0.450
pot1a -/- F_3	15	8	0.348
F_3 pxt class I	39	60	0.606
F_3 pxt class II	10	77	0.885

Figure 3-4. Exacerbated seed abortion phenotype in *pot1a ter2*. (A) Genetic scheme for generating *pot1a ter2* double mutants. (B) Silique analysis. Representative pictures of siliques for each genetic background are shown. (C) Summary of the seed abortion phenotypes of siliques counted in the experiment. Colored boxes correspond to the genetic diagram.

TER2 and de novo telomere formation (DNTF).

One attractive hypothesis for TER2 is that it represses telomerase enzymatic activity to prevent DNTF at the DSBs produced during meiosis. In support of this hypothesis we found that TER2 abundance peaks very early in flower development and declines after fertilization. Our hypothesis is plausible if a primary function of TER2 is as a negative regulator of telomerase. Failure to repress DNTF causes the genome instability, which should lead to decreased fertility. Interestingly, using an artificial DNTF assay (Nelson et al., 2011), graduate student Xiaoyuan Xie observed a slight increase in DNTF in *ter2* mutants. The data presented here showing decrease pollen viability and increased seed abortion in plant lakcing TER2 are consistent with a rol for TER2 as a negative regulator of telomerase to avert DNTF during meiosis.

A potential signaling pathway for TER2 in reproduction fitness

Plants lacking TER2 do not exhibit an obvious defect in telomere maintenance (Cifuentes-Rojas et al., 2012). However, our genetic analysis reveals that *pot1a ter2* double mutants have extra more trouble maintaining telomeres than *pot1a* single mutants. This phenomenon cannot be explained by DNTF model. DNTF requires the active telomerase to add telomere on the DSBs, arising problems in maintaining genome integrity. In the double mutants, if the level of seed abortion reflects the severity of genome stability problem caused by DNTF, we expect to see a less severe seed abortion phenotype in *pot1a ter2* than in *ter2* single, since telomerase is repressed in *pot1a* mutants. However, we find the opposite result. Plants with reduced telomerase activity (*pot1a* mutants) have a more severe seed abortion phenotype when combined

with a TER2 deficiency. This finding indicates that contribution of TER2 in reproductive development may not be dependent on telomerase repression.

Taken together, these results suggest TER2 is playing a more complex role in promoting reproductive fitness than simply repressing telomerase. Further support for a more complex role for TER2 in reproduction comes from analysis of *ter2 ku* double mutants. A similar synergistic seed abortion phenotype result was obtained from *ku70 ter2* double mutants. Plants lacking Ku have ultra long telomeres and extended G-overhangs (Riha and Shippen, 2003), but only a mild seed abortion phenotype (Xiaoyuan Xie, unpublished data). Therefore, the absence of TER2 strongly exacerbates a mild or no seed abortion defect.

Not completely exclusive with the DNTF model, these findings raise the interesting possibility that TER2 acts via a signaling pathway, instead of directly modulating telomere or telomerase during plant reproduction (Figure 3-5). Signaling is one of the four proposed archetypes of long non-coding RNAs (lncRNAs) (Wang and Chang, 2011). In chapter II, we show that TE exaptation in TER2 causes this telomerase lncRNA to become highly unstable. Here, we observed the distinct TER2 abundance change during flower development in Col-0. Interestingly, TER2 without TE (TER2Δ in Ler-0 accession) does not show the same abundance profile, suggesting the RNA may be subjected to post-trancriptional regulation in reproduction. RNA stability may act as a switch to allow this RNA to be a rapid signal during reproduction. However, we cannot rule out a possibility that TER2Δ instead of TER2 is the signaling molecule, and in Ler-0 the unchanged level of this molecule reflect a housekeeping function of this gene. Further supporting this possibility is the high conservation of the RNA domains of TER2Δ across all *A. thaliana* accessions (Chapter II). Whether there

are lncRNA signalling molecules that promote reproduction in other organisms is unknown, but in *A. thaliana* TER2 appears to have evolved into a novel signal, beyond its ability for prevention DNTF, in benefiting the reproduction.

TER2, a new lncRNA with emerging function

Our results suggest that TE seems to have distinct effects in reproduction among different accessions. In Col-0, though seed abortion phenotype caused by the absence of TER2 is relatively mild compared to other meiosis mutants such as *rad51 and dmc1* (Da Ines et al., 2013), suggesting a newly emergent function of this gene. However, the seed abortion in *ter2* mutants suggests that having TE in TER2 has been co-opted by Col-0 to rely this gene's function in reproduction.

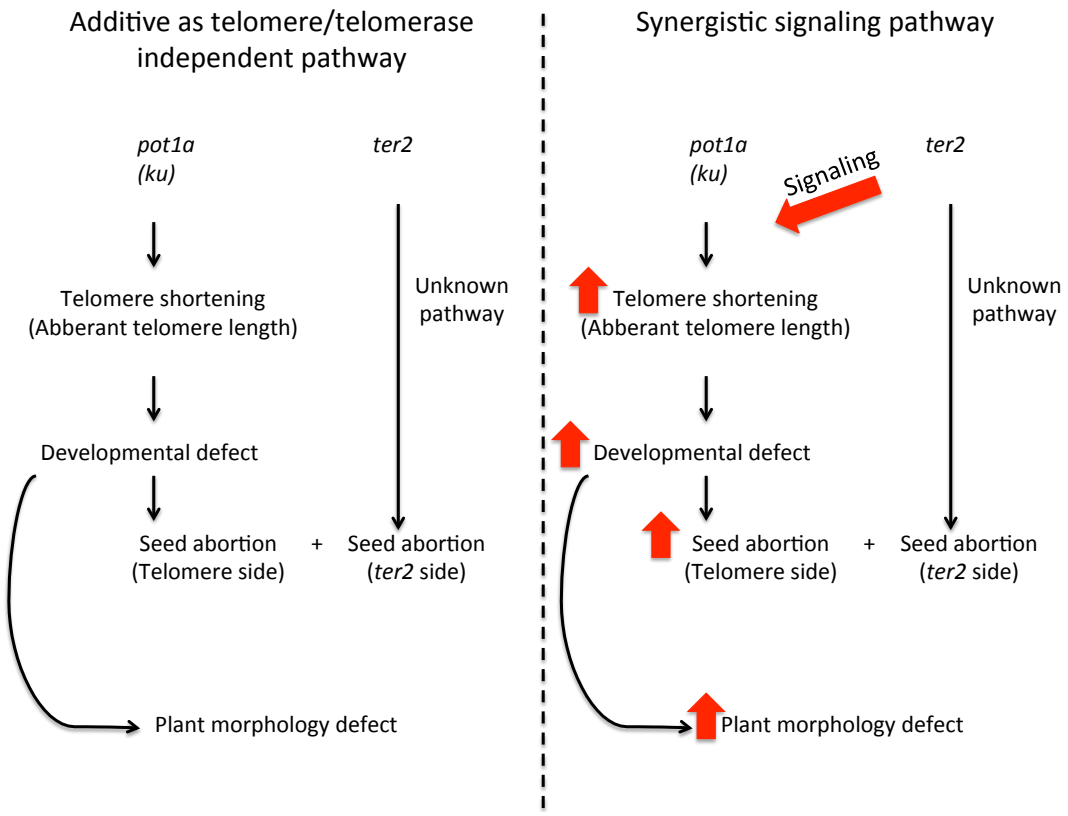


Figure 3-5. TER2 is a novel signaling molecule. Lacking POT1a causes continuous telomere shortening, resulting in the development defects with seed abortion and plant growth morphology. Lacking TER2 causes seed abortion but no telomere shortening. Therefore, the telomere length change is only contributed by the POT1a side, while the seed abortion phenotype is contributed by two sides (1+1=2) (Left part). This mostly happens in class I plants. However, we see a population of plants with exacerbated telomere length shortening and seed abortions in *pot1a ter2* double mutants. Since *ter2* mutant does not show short telomere, the further shortening telomere in this background suggests a deteriorated condition should come from *pot1a* side. This further indicates the lacking of TER2 serve as a signal to cause this deterioration. As expected, we see further shortening of telomere and synergistic exacerbated seed abortion (1(+1)+1=3). This could be most of cases in class II plants.

SNPs are very common with a frequency of 6-7 SNPs per kilo base among *A. thaliana* accessions over a 5 million years evolution (Nordborg et al., 2005; Beilstein et al., 2010). We hypothesize that other pathways must work in Ler-0 to promote reproduction. Unknown loci polymorphisms in Ler-0 could mask/complement the insufficiency in regulation of reproduction caused by TE absence at TER2. This is similar to the natural difference in phenotype of mutations in same loci among ecotypes. One interesting example is the *ago10* mutant required for the stem cell development (Lynn et al., 1999). In the absence of AGO10, the Col-0 accession has a weaker stem cell development defect than to Ler-0. This phenomenon ultimately led to the discovery of additional loci that enhances the *ago10* stem cell defect in Ler-0 (Knauer et al., 2013). Therefore, it is remarkable that the dynamic evolution of the TER2 locus by the exaptation of a TE resulted in a non-canonical telomerase RNA with a novel function to promote reproductive fitness with co-evolution of other unknown loci in *A. thaliana*.

Materials and methods

qRT-PCR

Experiments were performed as decribed in Chapter II.

Seed abortion analysis

Fresh siliques were collected from growing plants. Siliques were bleached in a solution with 1:3 ratio of acetic acid: ethanol at room temperature with shaking for 4 hours to overnight. The bleached siliques were neutralized by overnight incubation in 1N NaOH before analysis by microscopy.

Seed number analysis and pollen viability assay

Brown colored siliques were collected from mature plants. For each silique, the length of the silique was measured and seed number determined.

Pollen grains were collected from fresh flowers and placed on microscope slides. BK S15 buffer was added to the slide surface (To make 50ml, 5ml MOPS (100mM, pH7.5), 7.5g sucrose, 6.35ml 1M Ca(NO)2, 4.05ml (1M) MgSO4, 5ml KNO3, add water up to 50ml). 1ul FDA solution was added from a stock of 2mg/ml acetone. Observations were done immediately with an optical microscope in blue light (wavelength = 495nm). Viable pollen grains show fluorescence.

CHAPTER IV

INVESTIGATION OF THE TER2 PROCESSING PATHWAY

Summary

We previously identified an alternative telomerase RNA in *Arabidopsis thaliana* required for the negative regulation of telomerase in response to DNA damage. This RNA was later subsequently shown to harbor a *Copia*-like solo LTR transposable element (TE), termed telomerase regulateory element TRE, the embedment of which increases the TERT binding affinity relative to the canonical TER, TER1, but significantly decreases RNA stability. A third TER isoform TER2s is derived from TER2 by removal of the TE and a 36nt unique sequence residing at the 3' end. Little is known about the mechanism of TER2 processing and abundance regulation. Here, we investigate the pathways regulating TER2 processing. First, we ruled out the possibility that TER2 is autocatalytic spliced *in vitro.* Second, we showed that TER2 is influenced by the canonical mRNA splicing machinery. This finding is notable since the boundaries of the TE do not contain the consensus mRNA splicing sequences. Most surprinsingly, we find evidence that TER2 is regulated by small RNA processing machineries. We found a 24nt small RNA(sRNA) derived from TER2 TE in an RNA seq database. This sequence of the sRNA corresponds to $(AGA)_8$ tri-nucleotide repeats, and is absent from the two other TRE loci in the *A. thaliana* genome. We examined the role of small RNA machinery in TER2 modulation. We showed that TER2 abundance is increased in plant lacking Dicer-like 2 (DCL2), but not canonical sRNA machinery involved in RNA dependent DNA methylation, the dominant pathway used for TE silencing in plants. These results suggest that the TER2 is not silenced at transcriptional level via DNA

methylation. supported by CHOP-PCR results showing no DNA methylation at the TER2 locus. We also determined that TER2 RNA stability is increased five fold in *dcl2* mutants, indicating DCL2 regulates TER2 post-transcriptionally. Finally, We found that DCL2 regulates TER2 abundance during reproductive development. Altogether, these results indicate that a novel RNA processing pathway regulates the abundance of TER2, by the coordinating machinery involved in canonical RNA splicing and sRNA processings.

Introduction

Telomerase RNA (TER) is a long non-coding RNA that serves both a telomerase template and a protein scaffold for the holoenzyme assembly. Although all TERs shares key structural motifs, such as pseudo-knot and template domain essential for their functions, TERs are evolving much more rapidly than the catalytic subunit TERT. TERs are highly divergent in sequence and length, with distinct and divergent biogenesis pathways, from H/ACA small nucleolar RNP proteins and TCAB1 for vertebrate TER and p50 and p65 for Tetrahymena TER (Mitchell et al., 1999; Venteicher et al., 2009; Fu and Collins, 2007). In addition, divergent groups of TER accessory proteins have become essential for the TER maturation, localization and RNP assembly (Egan and Collins, 2012).

Telomerase RNA processing has been studied in both yeast and vertebrates. In budding yeast, a minor fraction of telomerase RNA Tlc1 contains a polyA tail, which is absent from the functional Tlc1 when it associated with the catalytic subunit (Chapon et al., 1997). Tlc1 3'end formation is dependent on the Nrd1/Nab3 pathway (Noel et al, Wellinger, 2002). In contrast, human telomerase RNA (TR) contains 5' cap structure,

but not a polyA tail (Zaug et al., 1996). The maturation of the TR is determined by the H/ACA domain at its 3' terminus (Mitchell et al., 1999); (Fu and Collins, 2003). In fission yeast, the 3'end of TER contains an intron and TER maturation proceeds through a novel RNA "slicing" mechanism, involving only the first step of splicing. Spliceosome cleavage at the intron 5' splice site releases the first exon by uncoupling the canonical second step splicing reaction (Box et al., 2008a; Tang et al., 2012). Though mechanisms vary, proper processing of telomerase RNA is required for its function.

Arabidopsis thaliana is unusual as it evolved three distinct TERs. TER1 is a canonical telomerase RNA required for synthesis of telomeric DNAs (Cifuentes-Rojas et al., 2011). In contrast, TER2 is a novel regulator of telomerase (Cifuentes-Rojas et al., 2012). Removal of a 529nt intervening sequence within the TER2 along with cleavage of the 3' terminus gives rise to the third isoform, TER2s. The function of TER2s is still elusive (Cifuentes-Rojas et al., 2012). Mechanisms leading to TER2s production are unknown. The sequences flanking the TER2 intervening sequence are not canonical mRNA splicing recognition sites. Notably, the intervening sequence in TER2 appears to be derived from a solo LTR of a *Copia*-like retrotransposon. Thus, its removal may proceed by an unconventional processing mechanism.

One surprising finding that may give some clues about TER2 metabolism is the discovery of a 24nt small RNA that maps to TER2 TE (Rajagopalan et al., 2006). This finding raises the possibility that TER2 is regulated by a small RNA mediated pathway. Transposons are subjected to small RNA dependent epigenetic regulation, leading to transcriptional gene silencing (TGS) (Castel and Martienssen, 2013). In the cases of post-transcriptional gene silencing (PTGS), gene expression is down regulated by cleavage of the target RNA via the RNA-induced silencing complex (RISC). However,

for TGS cases, the small RNA processed from the target gene transcript guides the RNA silencing machinery to the gene locus (DNA level) for *de novo* methylation, known as the RNA dependent DNA methylation (RdDM) (Matzke and Mosher, 2014). In both TGS and PTGS small RNA involved pathways, the Rnase III-like protein Dicers produce the small RNA. In plants, there are four Dicer-like (DCL) genes DCL1, DCL2, DCL3 and DCL4 (Liu et al., 2009). DCL1 is responsible for producing miRNA, leading to the PTGS function by the formation of the RISC complex as aforementioned (Golden et al., 2002). DCL3 is the primary player for producing 24nt siRNA functional for RdDM, while DCL2 and DCL4 act redundantly to complement the loss of DCL3 function (Henderson et al., 2006). In addition, plant specific RNA polymerase IV and V, RNA-dependent RNA polymerase (RDR2), and AGO4 (AGO6 and AGO9) collaborate with Dicer to generate precursor and scaffolding RNA, double-strand RNA conversion, and to orchestrate silencing of targets. The entire process is regarded as a plant defense system (Ding, 2010).

Here we examined the factors involved in TER2 metabolism. We first excluded the possibility that TER2 autocatalytic spliced *in vitro.* Second, we showed that TER2 is increased in a mutant defective in the canonical splicing machinery, indicating that some aspect of TER2 processing proceeds via conventional mRNA splicing. Most importantly, we showed TER2 specifically accumulates in the *dcl2* mutants, but not in other RdDM mutants, suggesting the DCL2 mediated TER2 regulation is not through typical DNA methylation mediated gene silencing. We found that DCL2 affects TER2 in a PTGS manner, and thus TER2 analysis has revealed a novel function of Dicer-like gene product for regulating the long non-coding RNA. Finally, TER2 regulation was disrupted during flower development in *dcl2* mutant, indicating DCL2 is a key gene for

TER2 during reproduction. These results illustrates a complicated and novel processing mechanism for TER2 or TER2-like TE containing lncRNA including collaboration of the canonical mRNA splicing machinery and small RNA processing machinery.

Results

TER2 does not autocatalytically splice in vitro

Preliminary results by previous members of the Shippen lab, A. Hernandez and C. Cifuentes-Rojas, suggested that TER2 might be subjected to autocatalytic splicing to remove the TE. To test this possibility, *in vitro* splicing reactions were carried out with 5' labeled TER2 RNA. A smaller RNA species appeared in the TER2 reaction, but not TER1 reaction, which migrate d at approximately the same size as TER2s (Figure 4-1A). Due to the broken gel during the experiment, the exact size of the product could not be determined. The experiment was repeated, but unfortunately no evidence for splicing *in vitro* was observed after several trials. Numerous changes in pH and RNA folding conditions were tested, but *in vitro* splicing of TER2 was not observed.

To further study TER2 processing *in vitro,* we collaborated with the lab of Dr. Martha Fedor at Scripps in California. As expected a four way helical junction ribozyme HP43 construct acquired from Dr.Fedor's lab was processed robustly *in vitro* (Figure 4-1B). However, Dr. Fedor's lab found no evidence of TER2 self-splicing *in vitro* under any reaction conditions. The previous success of our lab may reflect some mysterious contaminant in our reagents, which helped TER2

splicing. The most probable scenario is a contaminant that facilitates RNA folding *in vitro.* Although it is still unclear whether TER2 splicing is autocatalytic, the data suggests that the TER2 splicing *in vivo* requires additional components.

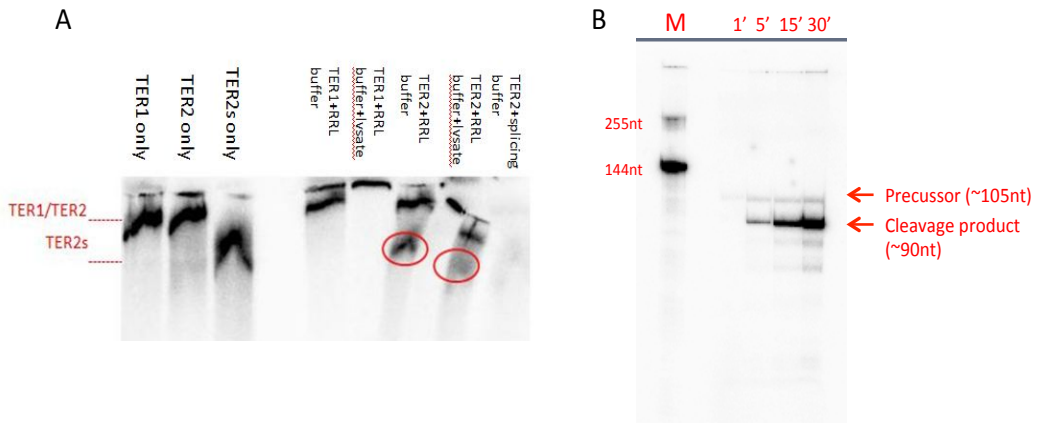


Figure 4-1. Preliminary evidence for TER2 autocatalytic splicing (A). TER2 splicing *in vitro.* RNAs were produced by an *in vitro* transcription system. RNAs were 5' ^{32}P labeled. A promising splicing product is highlighted by the red circles. (B). Positive control of splicing conditions using HP43 ribozyme acquired from Fedor lab.

TER2 abundance is regulated by the canonical splicing machinery

Although the sequences flanking the TER2 TE are not canonical mRNA splicing sites (Figure 4-2A), it is possible that the TE (intron) is removed by conventional splicing machinery (Larkin and Park, 1999). To test this possibility, we studied plants lacking LSM8, which exhibited decreased mRNA splicing (Golisz et al., 2013). Lsm proteins in humans and yeast are components of the U6 small nuclear ribonucleoprotein complex and required for pre-mRNA splicing (Novotny et al., 2012; Spiller et al., 2007). TER2 abundance was measured in *lsm8* mutants. If LSM is required for TE removal, we

expect that the level of TER2 will increase in the absence of LSM8. We first confirmed that the *lsm8* mutant was defective in processing ABL11 mRNA, the positive control. As expected, the precursor ABL11 transcript was higher than in wild type (Figure 4-2 B). Likewise, the TER2 transcript was significantly increased in *lsm8* homozygous mutant, while TER1 did not change (Figure 4-2C). This finding suggests a role for the canonical mRNA splicing machinery in TER2 processing. As shown in Figure 4-2A, the primers used in qPCR could only detect the intron removal. It is possible that LSM8 is required for removal of the 3' tail, not the TE.

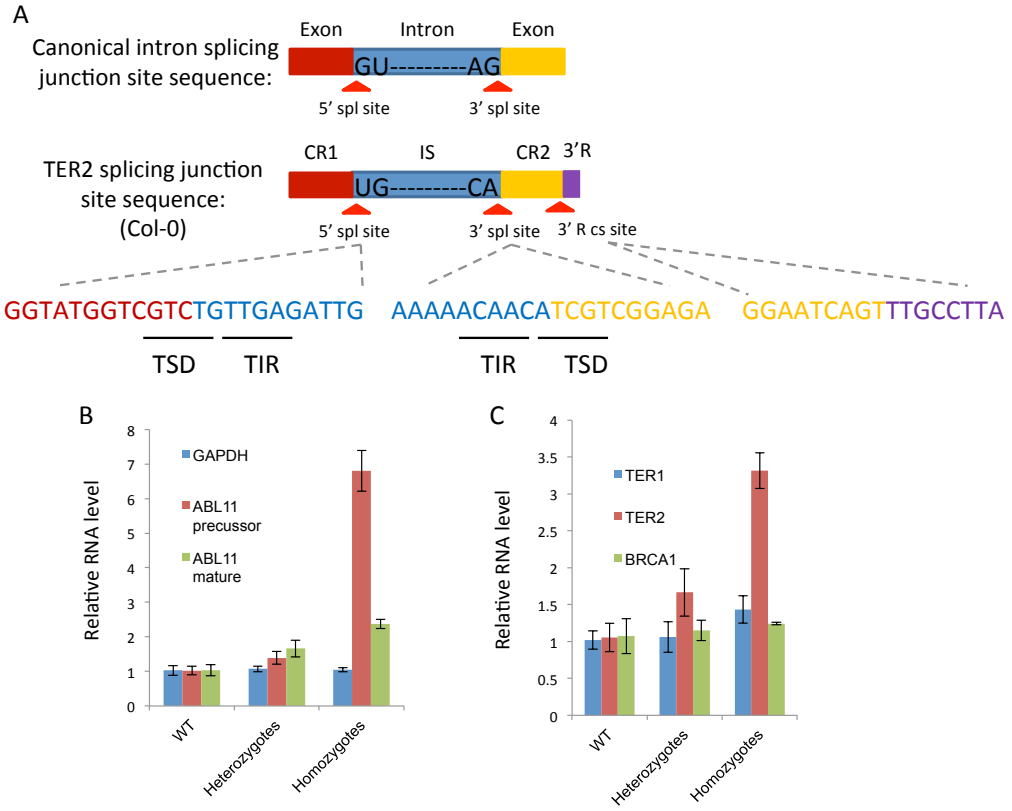


Figure 4-2. LSM8 is necessary to regulate TER2 abundance and/or processing. (A) Diagram of the boundary sequences for canonical consensus splice sites and sequence flanking the TER2 TE. (B). qRT-PCR results for ABL11mRNA in the *lsm8* mutant. (C) qRT-PCR results for TER1, TER2 and BRCA1 in *lsm8* mutant.

TER2 intron harbors a tri-nucleotide repetitive small RNA sequence

A small RNA $[AGA]_8$ mapping to the TER2 TE was identified by the high throughput sequencing (https://www.mcdb.ucla.edu/Research/Jacobsen/LabWebSite/P_Index.shtml) (Figure 4-3). The presence of this 24nt RNA suggests that the TER2 locus might be regulated by the RNA dependent DNA methylation (RdDM) pathway. 24nt siRNAs are typically generated by DCL3, RDR2 and RNA polymerase IV and V orchestrate the *de novo* DNA methylation at viral sequences, leading to the repression of the gene transcription. Therefore, we hypothesized that TER2 locus is under the regulation of 24nt siRNA dependent gene silencing pathway.

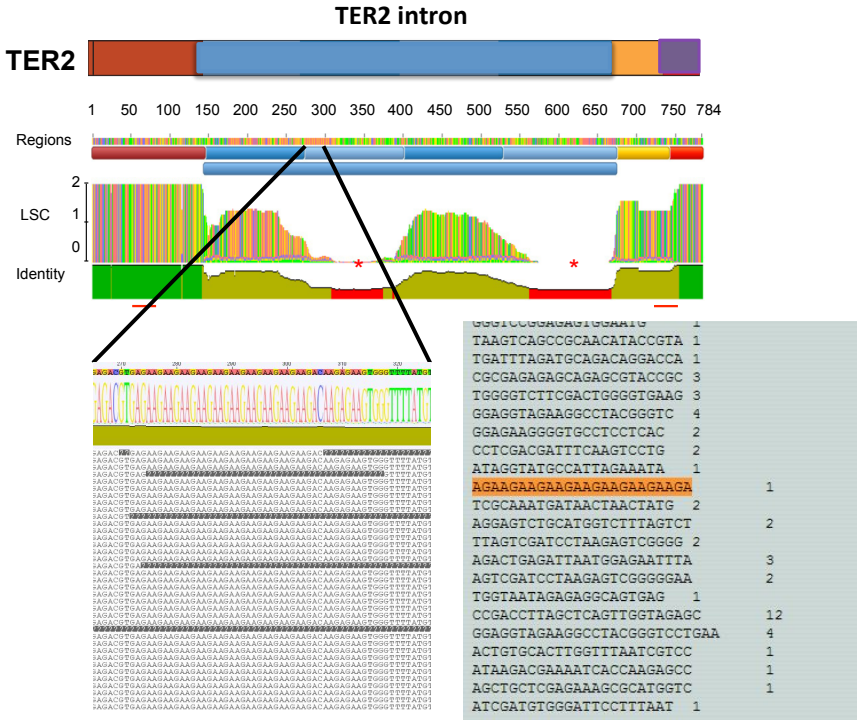


Figure 4-3. A 24nt RNA derived from a tri-nucleotide repetitive sequence embedded in TER2. (AGA) repeats sequence are found in the hypervariable region 1 of the TER2 TE. A small RNA sequence $(AGA)_8$ 24nt small RNA (bottom right), reported by Rajagopalan et al in 2006.

DCL2, not DCL3, is involved in TER2 RNA abundance regulation

To test the hypothesis that TER2 is regulated by a small RNA dependent pathway, we acquired several small RNA processing mutants from Dr. Xiuren Zhang: *dcl2, dcl3 dcl4* and *sgs2*. *Arabidopsis* has four Dicer genes (DCL1-4) used to generate sRNAs. DCL1 is primarily involved in miRNA processing, and plays a role in normal siRNA production. Therefore, we omitted this mutant from our analysis. The dicer mutants we used covered all siRNA biogenesis pathways. qRT-PCR was performed on RNA extracted from mutant flowers. The data showed no change in the TER1 transcript level among different mutants (Figure 4-4A). However, consistent increases of TER2 were observed only in *dcl2* mutant, indicating that TER2 may be under small RNA-mediated regulation specific to DCL2. Interestingly, the increase of TER2 could not be detected in the *sgs2* mutant. SGS2 works in the same genetic pathway with DCL2 for processing transgene small RNAs in plants (Mlotshwa et al., 2008; Borsani et al., 2005). Thus, the genetic data indicate an unexpected and non-canonical role of DCL2 for TER2 regulation.

TER2 transcription is not regulated by the canonical RdDM pathway

DCL2 and DCL4 have a redundant role for trans-gene silencing when DCL3 is absent (Henderson et al., 2006). Therefore, it is possible that DCL2 substitutes for DCL3 to produce a small RNA from TER2 that then feeds into the canonical RdDM pathway for TER2 regulation. To investigate this possibility, we assessed TER2 RNA abundance in other small RNA machinery mutants involved in the RdDM pathway. For upstream mutants of RdDM pathway, there was no change in TER2 abundance (*Pol IV (nrpd1a)* and *rdr2* mutants) (Figure 4-4C, blue circle). Strikingly, however, a mutant in

the Pol V accessory protein DRD1, which acts downstream of the RdDM pathway (Figure 4-4D, red circle), caused a significant increase in TER2. The increase in TER2 was the same as in the *dcl2* mutant (Figure 4-4D).

We checked the DNA methylation at TER2 locus using Chop-PCR (DNA methylation sensitive PCR), where a DNA methylation sensitive enzyme digests the methylated DNA resulting in the failure of the PCR. We performed this experiment using several *A. thaliana* accessions seeding DNAs. Col-0 did not show DNA methylation at TER2 locus, while some other ecotypes suggest DNA methylation. Our result confirmed that there is no DNA methylation in Col-0 wild type plants (Figure 4-4B). Some other ecotypes show the evidence of DNA methylation at TER2, which is not correlating with the TE status. These data indicate that TER2 abundance is modulated by a non-canonical processing pathway independent of conventional gene silencing (TGS) via the RdDM in Col-0. Intriguingly, the TER2 loci have the potential to be methylated by an unknown pathway.

DCL2 regulates TER2 abundance in a post-transcriptional manner

Our genetic analysis indicates that DCL2 does not regulate TER2 at the transcriptional level. An alternative role of DCL2 is to alter TER2 abundance by post-transcriptional modulation. This hypothesis is appealing since TER2 has a short half-life (Chapter II). We checked TER stability in *dcl2* mutants. 14 day old seedlings were treated with cordycepin (150ng/ml) and samples were collected at different incubation times from wild type Col-0 and *dcl2* mutants.

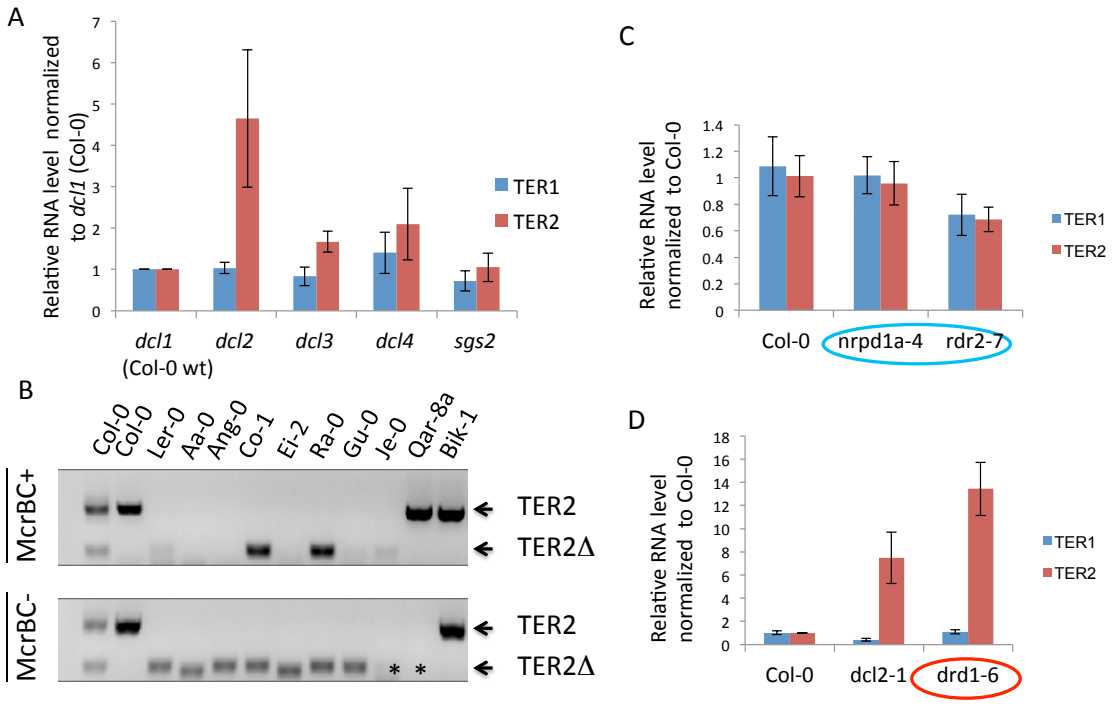


Figure 4-4. TER2 is upregulated in *dcl2* mutants, but not in mutants involved in the canonical RdDM pathway. (A). qRT-PCR result in different small RNA machinery mutant plant including DICER genes in (A), upstream RdDM genes (blue circle) in (C), and downstream RdDM gene DRD1 (red circle) in (D). GAPDH was as a reference. (B) Chop-PCR in different *A. thaliana* accessions. DNA prepared from seedlings was treated with McrBC. PCR was performed after treatment. * indicates the PCR failure.

qRT-PCR showed that TER2 was significantly stabilized in the *dcl2* mutant (Figure 4-5A). The TER2 half-life was ~15 min in wild type and increased ~60min in plants lacking DCL2. In contrast, the half-life of other RNA transcripts including TER1, BRCA1 and ABI1 (Figure 4-5B, C, D) did not change. These results indicate that DCL2 specifically regulates TER2 stability.

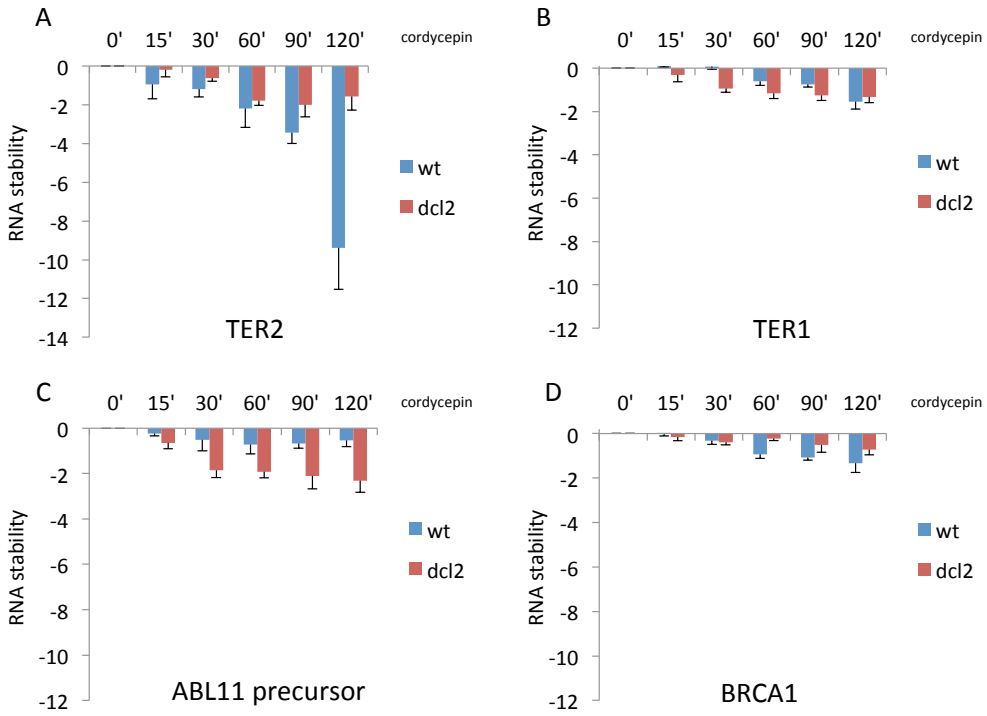


Figure 4-5. TER2 is stabilized in plants lacking DCL2. RNA stability assays were performed in Col-0 and *dcl2* mutant seedlings using cordycepin. RNA abundance was determined by qRT-PCR. eIF-4a was used as a reference gene. The RNA stability results are illustrated as the decreased level of the RNA abundance to the untreated samples. TER2 (A), TER1 (B), ABL11 precursor (C) and BRCA1 (D) were tested.

DCL2 regulates for TER2 during flower development

We previously showed that TER2 accumulates in flower buds and gradually decreases after pollination (Chapter III). To ask whether TER2 regulation during reproductive development requires DCL2, we monitored TER2 at different days post-pollination in *dcl2* mutants (Figure 4-6). The wild type pattern of a gradual decrease in TER2 was grossly disrupted in *dcl2* mutants. TER2 increased in the opened flower, peaking 9 days after pollination, indicating that DCL2 is required for down-regulation of TER2. This result is also consistent with microarray data showing that DCL2 is expressed at a higher level later in floral development (www.arabidopsis.org). Interestingly, this TER2 profile was similar to wild type Ler-0 TER2Δ (Figure 4-6). Since the TER2 promoter is highly conserved across all *A. thaliana* accession (Chapter II), our findings indicate that the TE in TER2 is responsible for TER2 regulation during reproductive development, and further that this sequence is targeted by DCL2 to promote RNA instability.

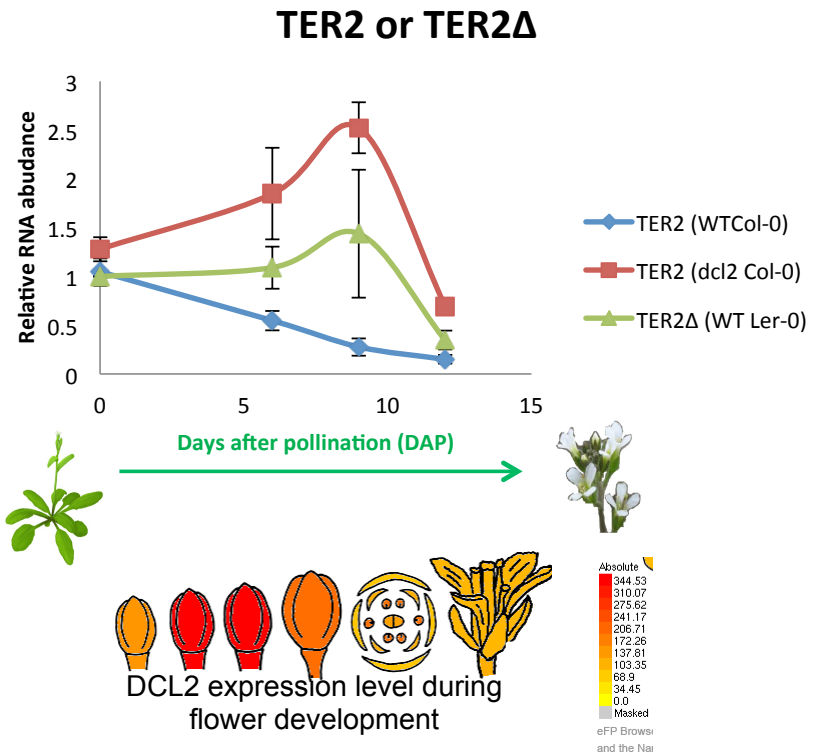
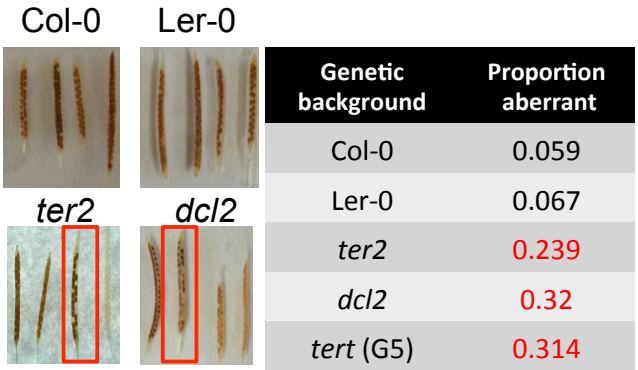


Figure 4-6. DCL2 is required for TER2 regulation during flower development. The experiment was performed as described in Chapter III Figure 3-1. Samples were collected from plants at different days after pollination (DAP). TER2 abundance was determined by qRT-PCR. GAPDH was used as a reference. The data were normalized to flower bud (0 DAP) of the corresponding genetic backgrounds. DCL2 relative expression during floral development was shown below.

DCL2 is required for reproductive development

DCL2 was found highly expressed in the plant egg cell, suggesting an important function in reproductive development (Takanashi et al., 2011). We previously showed that *ter2* mutants are defective in seed formation. Since DCL2 is a candidate for TER2 intron processing, we asked if a seed abortion phenotype was associated with *dcl2* mutants. Strikingly, we found the proportion of seed pods showing abortion is significantly increased in the *dcl2* mutant, and parallels the ratio in *ter2* plants (Figure 4-7). The result raises the intriguing possibility that DCL2 modulation of TER2 and perhaps other lncRNA bearing TEs is important for reproductive fitness.



Genetic background	Proportion aberrant
Col-0	0.059
Ler-0	0.067
ter2	0.239
dcl2	0.32
tert (G5)	0.314

Figure 4-7. Plant lacking of DCL2 exhibit seed abortion in the siique. Seed abortion analysis was performed as described in Chapter III. Figures of representative siliques are at the top. Siliques with aborted seeds are highlighted in the red box. The ratios are summarized in the table.

Discussion

Introns have recently been shown to flank the 3'terminus of fungal TER genes and essential for the functional TER maturation (Box et al., 2008a; Tang et al., 2012). In *A. thaliana* TER2, a solo LTR is embedded within the body of gene, and must be removed to generate TER2s. How TER2 is processed to generate this RNA isoform is unknown. Likewise, it is unclear what factors destabilize TER2. Clearly, the TE is needed, but how the TE cooperates with DNA damage machinery to cause changes in TER2 stability is not known. In this study, we examined TER2 metabolism. TER2 is not subjected to autocatalytic splicing under standard reaction condition. Previous evidence showed that TER2 might be autocatalytically spliced during a telomerase reconstitution assay. The apparent autocatalytic splicing was observed in the presence of magnesium and sodium, and the splicing product was purified and verified by sequencing. However, the *in vitro* splicing reaction was not reproducible in my experiments or those of our collaborator Dr. Matha Fedor. Our current hypothesis is that TE removal might occur in the presence of an unknown contaminant.

Canonical mRNA splicing machinery regulates TER2 abundance

It is still possible that removal of TE occurs by a non-canonical pathway that is largely RNA catalyzed *in vivo. The* failure to see robust TER2 intron splicing *in vitro* implies that the intron removal or TER2 processing *in vivo* relies on additional factors. Interestingly, the intron boundary sites flanking the TE do not contain the consensus sequences required for mRNA intron splicing. Nevertheless, our results showed that TER2 abundance increases when the canonical splicing machinery efficiency decreases. We still do not have a clear idea about TER2 processing mechanism. It is

possible that unknown machinery required for TER2 processing exists. Due to the reduced activity in processing mRNA introns in *lsm8* mutants, it is possible that the effect of this mutant is indirect. Insufficient amount of a TER2 processing factor may prevent TER2 processing. If the canonical splicing machinery directly processes TER2, the reaction proceeds with 5' and 3' boundary sequences that are not optimal. Non-conventional splice sites have been reported in several mRNA splicing cases, revealing the flexibility for intron boundary sequences (Arrisi-Mercado et al., 2004; Burset et al., 2000). It is also possible that LSM8 is needed for 3' end processing of TER2. Base on qPCR primer limitation, we could not determine whether TE removal or 3' end removal was altered in the *lsm8* mutants. Further experiments are required to determine how the loss of LSM8 affects TER2 processing. Although the mechanism of TER2 processing is unclear, these data with LSM8 mutants argue that additional factors are needed for TER2 processing, and these factors may not be confined to conventional mRNA splicing machinery.

DCL2 mediates TER2 abundance in a non-canonical manner

The discovery of a 24nt sRNA associated with the TER2 TE led us to investigate a role for the small RNA processing machinery in TER2 metabolism. Retrotransposons are frequent targets of silencing by small RNA dependent DNA methylation (Castel and Martienssen, 2013). Our result showing the sRNA site is missing in the other two IS loci. Strikingly, the target site for the 24nt RNA is absent from the IS_{3L}, one of two TE in the *A. thaliana* that shows high similarity to the TER2 TE (~95% ID). This result may reflect a purifying selection. To have a better chance to survive, some TEs delete the small RNA target site to alleviate the repression (Wang et al., 2013). If this is true for the TE associated with TER2, it implies a significant role for the 24nt sRNA site in the TER2 TE with respect to TER2 regulation.

One of the most important discoveries here was that TER2 increases in *dcl2* mutants. Currently, we do not know if this increase reflects a role of its derived sRNA. However, the DCL2 mediated regulation of the TER2 lncRNA seems to be novel. Firstly, our results are not consistent with the known DCL2 mechanism. DCL2 produces the sRNA with the size of 21nt (Mlotshwa et al., 2008), while the small RNA derived from TER2 TE are 24nt. In a very rare case, the DCL2 together with SGS2 can produce the 24nt small RNA (Borsani et al., 2005). However, we detect no change in TER2 abundance in *sgs2* mutants, suggesting an SGS2 independent pathway for TER2 processing by DCL2. Secondly, the TER2 regulation is not via the conventional RdDM pathway. We fail to see a TER2 increase in mutants of most RdDM players. Furthermore, we failed to detect DNA methylation at the TER2 locus. These results are consistent with analysis by Daniel Zilberman who found no evidence for methylation of

TER2 (D.Zilberman, personal communication). Thus, a novel mechanism mediated by DCL2 appears to be operating for TER2 processing.

We cannot rule out the possibility that the TER2 locus is under transient/developmental transcriptional regulation. Our result showed that TER2 significantly increases in *drd1* mutants. DRD1 is a chromatin remodeler that facilitates the RdDM pathway (Kanno et al., 2004). Furthermore, our Chop-PCR result showed that TER2 loci are under methylation in some accessions, reflecting a conditional requirement for its methylation. A closer examination of the TER2 methylation status in various tissues and different developmental stages is required to investigate how methylation affects TER2.

We recently discovered that TER2 is an unstable RNA, and DNA damage response stabilizes it (Chapter II). Here we showed that half-life of TER2 significantly increased in *dcl2* mutants. We do not know if the change in TER2 stability in the two settings (DNA damage or depletion of DCL2) reflects the same processing pathway. Furthermore, we do not know if TER2 stabilization reflects a functional processing to generate TER2s or RNA degradation. Perhaps TER2 is depleted by two distinct metabolic processing pathways: one via canonical splicing, the other by degradation via DCL2. Understanding the detailed mechanisms of these pathways will require further investigation. However, the more regulatory pathways that influence TER2 metabolism, the more likely it is that RNA contributes to multiple aspects of *A. thaliana* biology.

A subset of lncRNA may be processed by DCL2 in a developmental and post-transcriptional fashion

Notwithstanding many unknown facts, our result shows a clear function of DCL2 in TER2 post-transcriptional regulation. Furthermore, this happens by the DCL2-TER2 TE interaction. In *dcl2* mutants, the TER2 abundance profile during floral development becomes similar as its counterpart in Ler-0 ecotype where TER2 lacks the TE. This result is consistent with the report that DCL2 is highly expressed in the egg cell (Takanashi et al., 2011), suggesting a unique function of DCL2 for a subset of RNA regulation during a special stage of reproductive development. More importantly, these results may further provide a lesson for lncRNA evolution. Guigo and Johnson recently proposed that TEs with pre-formed structures that increase the RNA structure flexibility (Johnson and Guigo, 2014). In this case of TER2, the TE may simply provide a resource of small RNA sequences to target recognition by DCL2. Having a small RNA site to be processed by a Dicer-like protein may not be for generating small RNA. Instead, the purpose of processing is to destabilize the transcript carrying this small RNA site. In humans, DICER1 can directly process Alu lncRNA and failure to do so causes retinal pigmented epithelium cell degeneration (Kaneko et al., 2011). Thus, Dicer proteins may play a larger role in regulation of lncRNA, and TER2 may represent a general paradigm for how TEs or repetitive sequences evolve and contribute lncRNA regulation together with Dicer like proteins. Transcriptome analysis of the human genome suggests that more than half lncRNAs are derivatives of TE (Kapusta et al., 2013). A detailed transcriptome study in *A. thaliana dcl2* mutants may identify the new important substrate for DCL2. It is conceivable that Dicer proteins have a much broader set of substrates that go well beyond small RNA regulation.

Materials and methods

Plant material, growth conditions and transformation

For experiments with seedlings, seeds were sterilized in 50% bleach with 0.1% Triton X-100 and then stored in 4°C for 2-4 days. Liquid Murashige and Skoog (MS) medium were used for germination and growing. After transferring cold-treated seeds to MS, plants were grown at 22°C under long day light condition for ~7 days. *dcl2-1, dcl3-1, dcl4-2* were acquired from Zhang lab. *rdr2-7, nrpd1a-4* and *drd1-6* were acquired from Riha lab. *lsm8* was acquired from ABRC as the line described in previous publication (Golisz et al., 2013).

RNA labeling and in vitro splicing

Template for making RNA was prepared by PCR. Commercial RNA *in vitro* transcription kit from was used for making RNA (Ambion). Product was resolved after denature PAGE gel. The RNA was recovered from gel and precipitated by ethanol. 1mg RNA was treated with CIP phosphatase. After phenol chloroform purification, RNA was incubated with radioactive ^{32}P-gATP and PNK for labeling. Labeled RNA was isolated by PAGE gel following with gel purifying. 1k cpm RNA was used in *in vitro* splicing reaction. RNA was heated in 95°C for five minutes and folded in 50mM sodium and 20mM magnesium at room temperature for 30 minutes. The magnesium was then added to 50mM to the final concentration. The concentration of sodium and magnesium as well as the folding temperature and time in experiment are under optimization. After splicing reaction, products were resolved in denature PAGE.

Chop-PCR

1mg DNA was treated with 5U McrBC under 37°C for 4 hours. After treatment, DNA was ethanol precipitated and used as the PCR template.

Seed abortion analysis, qRT-PCR and RNA stability assay.

These experiments have been performed as previously described in Chapter II and Chapter III.

CHAPTER V

FLEXIBILITY IN TER1 TEMPLATE UTILIZATION BY *ARABIDOPSIS THALIANA* ACCESSIONS

Summary

Telomerase is a processive enzyme that replenishes telomeric DNA by using a template in its RNA subunit (TER). The template typically consists of 1.5 copies of the telomere repeat sequence, including alignment nucleotides to reposition the template DNA 3' terminus after each round of synthesis. Although its preferred substrate is telomeric DNA, telomerase can also act on non-telomeric DNA by positioning it at a "default" site on the template. How telomerase manipulates the template in TER to achieve processivity, specificity and fidelity is unknown. In *Arabidopsis thaliana,* TER1 is the canonical telomerase RNA template and analysis of the TER1 gene across 853 *A. thaliana* accessions revealed single nucleotide polymorphisms (SNPs) at the 5' and 3' ends of the template region. Here, we examine template utilization in several accessions with variant template sequences. Srikingly, sequence analysis of telomerase products revealed perfect TTTAGGG repeats, indicating that SNPs define the 5' and 3' boundaries of the functional template. These experiments also indicate that for at least one accession (Krot-0) only a single nucleotide is sufficient for primer re-alignment during telomere repeat synthesis. In another accession Qar-8a, 3' SNP precludes use of the same alignment nucleotides, suggesting the enzyme active site is shifted to allow synthesis of perfect TTTAGGG repeats. Using a non-telomeric DNA primer ending with CCC, we observed a uniform primer entry site among different accessions, specifically U9, indicating this is the default entry site for non-telomeric

DNA 3' ends. Previous results showed template utilization varies in different plant species. Our data shows that even after 5 mya evolution within *A. thaliana* populations, enzyme fidelity is maintained by different template utilization strategies.

Introduction

Telomeres are the physical ends of eukaryotic linear chromosomes, which are composed of tandem repetitive G-rich DNA sequences and associated proteins. Telomeres serve two important functions: 1) they stabilize the genome by preventing the cell from recognizing chromosome ends as double-strand breaks (DSB); 2) they allow the cell to solve the end-replication problem (Olovnikov, 1971). Failure to maintain telomeres will activate cellular DNA damage responses leading to nucleolytic degradation by exonucleases and chromosomal end-to-end fusions. When telomere sequences erode below a critical length threshold, the cell will lose its capacity for cell division causing cellular senescence (Abdallah et al., 2009; Anderson et al., 2012).

Telomeric DNA repeats vary among different species (Blackburn and Gall, 1978; Moyzis et al., 1988). The vast majority of plant species harbor perfect TTTAGGG repeats as shown in *Arabidopsis thaliana* (Richards and Ausubel, 1988; Fajkus et al., 2005). The telomeric G-rich strand (G-strand) runs in the 5' to 3' direction relative to the terminus and ends in a short single-stranded 3' overhang (G-overhang). The G-overhang is bound by single-strand specific proteins, and this DNA acts as a primer for telomerase to synthesize telomeric DNA (Henderson and Blackburn, 1989). Telomere binding proteins either directly contact the single-strand or the double-strand region of telomere, or are enriched at the telomere region through protein-protein interactions (Bianchi and Shore, 2008). Telomere proteins promote genome integrity, but they are

not sufficient to overcome the DNA losses caused by the end-replication problem. The problem, which reflects the failure of unidirectional DNA polymerase to fully replicate chromosome termini, causes the linear chromosomes to shorten after each replication cycle. Genome instability arises not because telomere adjacent genes are lost, but because insufficient telomere repeats are available to block a DNA damage response.

The ribonucleoprotein enzyme, telomerase evolved to solve the end replication problem by replenishing telomeric DNA during replication. Telomerase is a specialized reverse transcriptase. At its core, telomerase harbors two essential components for catalytic activity: the protein subunit, TERT, containing a reverse transcriptase activity, and the RNA subunit TER, providing a template for telomeric DNA synthesis. Telomerase adds telomeric DNA *de novo* onto the G-overhang through the cooperation of TERT and TER. After the extension of the G-strand by telomerase, the C-strand is filled in by DNA Polymerase a to maintain the length of telomere (Fan and Price, 1997). Inactivation of telomerase causes stem cell related diseases in humans including Dyskeratosis congenital (Armanios and Blackburn, 2012). Meanwhile, failure to repress telomerase in cells with limited proliferation capacity is associated with cancer (Armanios and Greider, 2005). Therefore, to maintain genome stability and the proper cell behavior, telomerase activity must be under careful surveillance.

Compared to TERT, TER genes are highly divergent in length and nucleotide sequence. Nevertheless, common conserved secondary structures are found in all TER molecules and are necessary to fulfill their function. One conserved domain is the template which contains about 1.5 copies of the complementary (C-rich) telomeric DNA sequence. Most of these residues are used as templating nucleotides to direct synthesis of the TTTAGGG repeat. However, some residues are needed for primer

realignment following translocation during each round of synthesis. These “extra” residues also give flexibility for the extension. Immediately 5’ of the template is a template boundary element (TBE), a stem-loop structure that prevents nucleotide incorporation beyond the template. Together these signature regions, the template domain and TBE, promote fidelity of the newly incorporated sequence (Yu et al., 1990; Chen and Greider, 2003). Another key structure is a pseudoknot domain close to the template region, which stimulates enzymatic activity (Tzfati et al., 2000; Autexier and Greider, 1995; Chen and Greider, 2003). Lastly, the CR4/5 stem loop structure boosts enzymatic activity by a long-range interaction to the pseudoknot (Collins, 2006). The remainder of TER appears to serve as a scaffold for binding accessory proteins, that regulates enzyme assembly (Mitchell et al., 1999), recruitment (Taggart et al., 2002; Wu and Zakian, 2011), and enzymatic activity *in vivo* (Leehy et al., 2013; Renfrew et al., 2014; Min and Collins, 2010).

Telomerase is a promiscuous enzyme and can add telomeric DNA to the sites of double strand breaks (DSB). First observed in 1940s in maize, the “repair” of DSBs by telomerase action through *de novo* telomere formation (DNTF) allows continual cell cycle progression by transiently repressing the DNA damage response. However, the acentromeric chromosome fragment will be lost, leading to genome instability (Ribeyre and Shore, 2013). Therefore, DNTF must be repressed by proper regulation of telomerase. In plants, DNTF capability varies. *In vitro* analysis estabilished three DNTF modes (Fitzgerald et al., 2001). *Glycine max* (Soybean) telomerase can only elongate telomeric DNA sequences, while the *Sorghum bicolor* (sorghum) telomerase utilizes both telomeric and non-telomeric DNA using several alternative entry sites. Interestingly, primer elongation especially the first few repeats is error prone. The *A.*

thaliana and *Zea mays* (maize) telomerases extend both telomeric and non-telomeric DNA using a specific "default" entry site at the C8 position in the TER template. Unlike sorghum, the Arabidopsis and maize telomerase are considerably less error prone under these circumstances (Fitzgerald et al., 2001).

We recently reported that *A. thaliana* encodes two TERs, TER1 and TER2. While TER2 is a novel regulator of telomerase activity (Cifuentes-Rojas et al., 2012), while TER1 is the canonical TER used to maintain telomeres (Cifuentes-Rojas et al., 2011). By leveraging data from the 1001 *A. thaliana* genomes project, we found that TER1 is highly conserved with an overall 99% ID across the 853 accessions analyzed. This value is lower for TER2 due to apparent degeneration of a transposable element embedded in the gene (Xu et al, in revision). Interestingly, single nucleotide polymorphisms (SNPs) are detected at the 5' and 3' regions of the TER1 RNA template domain, raising questions about utilization of the nucleotides in the template domain. Here, we investigate the impact of the TER1 template polymorphisms. Analysis of telomerase products from *A. thaliana* accessions bearing 5' and 3' template SNPs showed only the conventional TTTAGGG repeats. Furthermore, telomere length was not altered, supporting the conclusion that variant nucleotides within the template domain are not copied into telomeric DNA in theses accessions. For at least one accession, Krot-0, only a single nucleotide is used for primer alignment after translocation. Using a series of primers with different 3' sequences, we found a correlation between enzyme fidelity and terminal primer sequence and an important role for the penultimate nucleotide in primer alignment. Finally, we identified a uniform "default" entry site for non-telomeric DNA primers, suggesting SNPs do not affect the capacity of these enzyme to extend non-telomeric DNA 3' ends. Altogether, our results

reveal a dynamic coordination of template nucleotides in annealing, extension and translocation to ensure enzyme fidelity and specificity.

Results

Polymorphisms in the TER1 template region in different A. thaliana accessions

Our previous analysis of TER1 and TER2 from the *A. thaliana* 1001 genome project revealed SNPs in the template of TER1, but interestingly not in TER2 (Xu et al, in revision). The template sequence in *A. thaliana* is 5'-CUAAACCCUA-3', which encodes 1.5 copies of the plant telomere repeat sequence, TTTAGGG (Figure 5-1A). Three types of polymorphisms were identified in the TER1 template. The most frequent was a A to T substitution at the second position (5' polymorphism, A2) (44/853). In addition we find a C to T change in the middle of the template (C6) (1/853) and a A to C transition at the 3' most nucleotide position in the template (3' polymorphism, C10) (3/853) (Figure 5-1A). Unfortunately, seeds from Bela-1, the only accession showing a SNP in the middle of the template domain are not available, preventing us from investigating the function of this particular TER1. Instead, we focused on representatives from *A. thaliana* accessions with 5' or 3' SNPs, 5' SNPs were represented by Krot-0, Nok-3 and Nc-1, and 3' SNPs by Qar-8a and Bik-1 (Figure 5-1B). Genotyping PCR and sequencing performed on several plants from each accession verified the 5' and 3' SNPs in the TER1 template (Figure 5-1C).

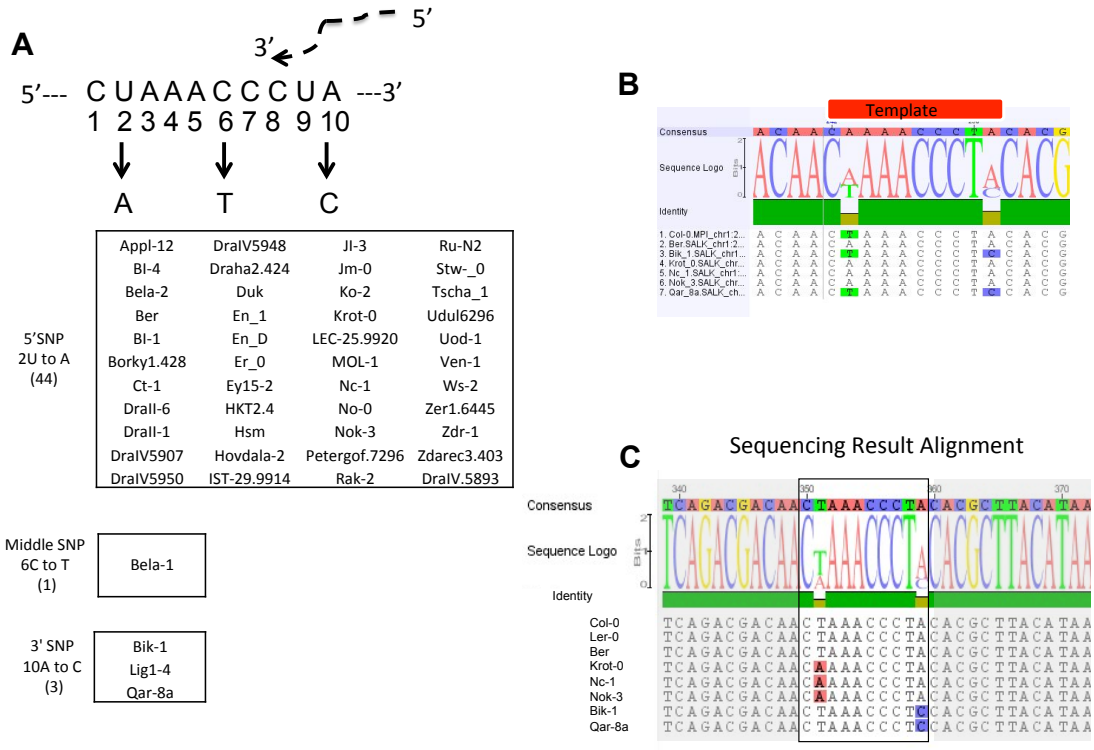


Figure 5-1. Polymorphisms in the TER1 template region in different *A. thaliana* accessions. (A). Diagram of the RNA template for TER1. Nucleotide positions are indicated below. Dashed line represents annealing of the primer to template. Corresponding SNPs are shown by arrows. Accessions with template SNPs are shown in the boxes, bracketed number indicates the total number of accessions with a particular SNP. (B). Bioinformatic analysis of the accessions used in this study. (C) Verification of SNPs by sequencing.

Telomere length is not changed in plants with TER1 template polymorphisms

Telomere length changes can occur when the telomere sequence is altered (Yu et al., 1990). Therefore, we checked telomere length in several accessions with template SNPs. Sequence analysis revealed no subtelomeric polymorphisms on chromosome arm 2R and 4R among those accessions allowing us to perform the primer extension telomere length amplification (PETRA) assay (Figure 5-2A). In all the accessions we monitored, telomeres were within the same length range as in Col-0, the reference accession for functional analysis of TER1 (Cifuentes-Rojas et al., 2011) (Figure 5-2B). The data indicate that if a variant telomere repeat is incorporated into telomere tracts in these accessions, it does not result in a substantial alteration in telomere length.

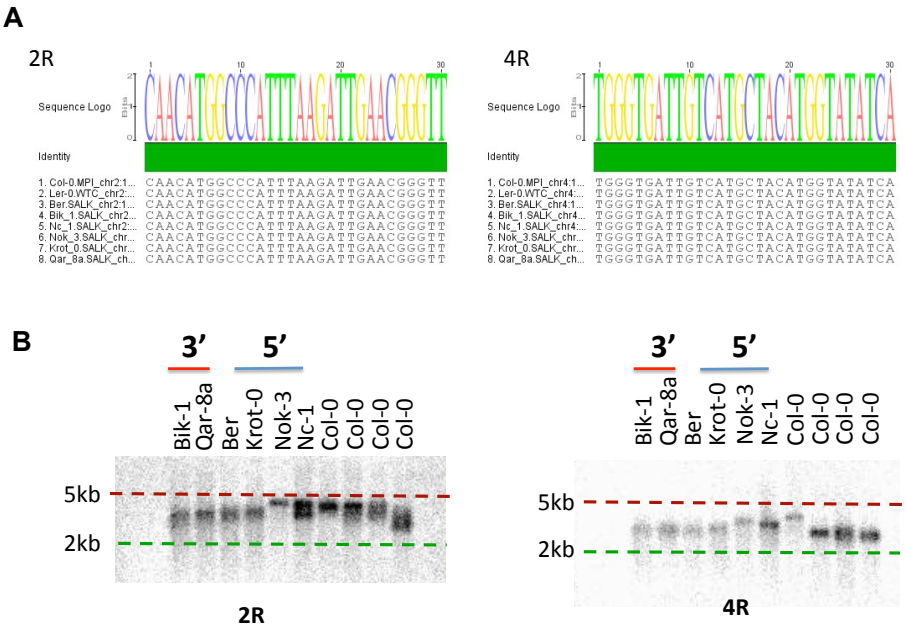


Figure 5-2. No obvious telomere length difference in accessions with TER1 template polymorphisms. (A). Primer annealing site in the subtelomeric regions of right arm of chromosome 2 (2R) and or the right arm of chromosome 4 (4R) chromosome ends. The target sequences are conserved among all accessions tested. (B). Telomere length analysis by PETRA on 2R and 4R.

Telomerase activity levels are unchanged in plants with a 5' TER1 template SNP.

We next asked if the TER1 template polymorphisms alter the level of telomerase enzyme activity. Protein extracts were prepared from seedlings. The Qar-8a and Bik-1 did not grow well in our growth chamber. We were unable to obtain sufficient plant material for extensive telomerase. Therefore, we focused our study on accessions with a 5' polymorphism in the TER1 template using Krot-0 as representative. If the A2 SNP is used as a templating nucleotide, we expect Krot-0 telomerase synthesis TTTTGGG repeats instead of the conventional TTTAGGG repeat. The different telomere repeat sequence might also see cause change in template utilization or enzyme fidelity.

A primer whose 3' terminus ended with CAG was used to monitor telomerase activity in Krot-0. Quantative TRAP (qTRAP) results showed that Krot-0 enzymatic activity levels were the same as in Col-0 (Figure 5-3A). We also examine the profile of elongation by polyacrylamide gel electrophoresis. The banding profile produced by telomerase is indicative of mechanistic "switches" in the elongation process including pausing and primer translocation (Greider and Blackburn, 1989). These "switches" can affect the overall banding pattern pattern of telomerase products from PAGE. If the primer positioning sites are different, we may detect a band shift. However, our results showed that the same banding pattern of telomerase products for Col-0 and Krot-0 (Figure 5-3B). We sequenced the TRAP product to directly assess the sequence synthesized by Krot-0 telomerase. Among 2317 nucleotides sequenced, only TTTAGGG repeats were generated. (Figure 3C). These results indicate that A2 is not used as a template by Krot-0. The data argue that primer translocation occurs after copying A3 in the template.

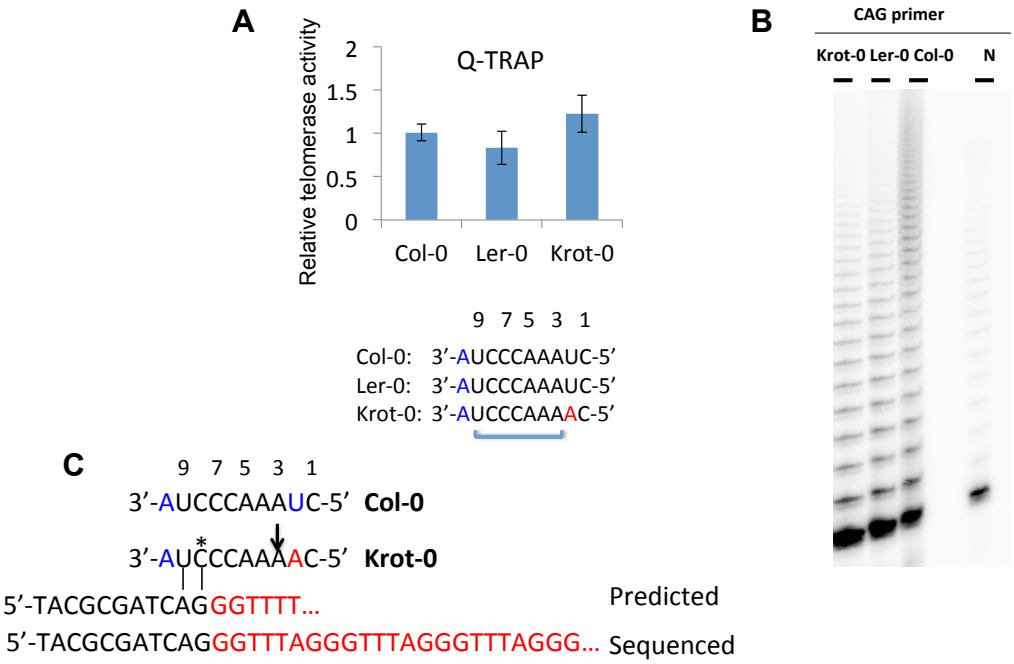
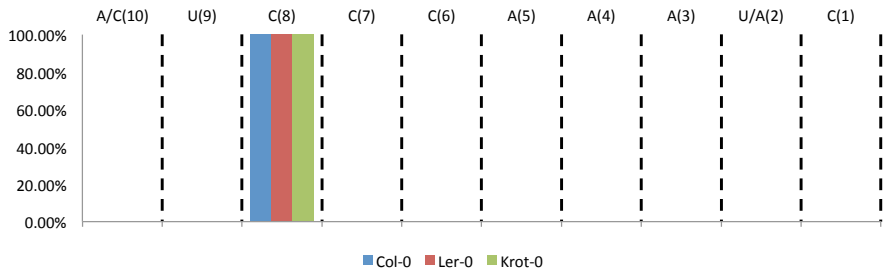


Figure 5-3. Telomerase activity is the same in Krot-0 as in Col-0 using primer ending with CAG. (A). qTRAP results for seedlings are shown. Relative telomerase activity was normalized to Col-0. Data are from 3 biological replicates. The 5' SNP is A2 (red). The predicted alignment nucleotide is A10 (Blue). (Bracket below Krot-0 template indicates the predicted templating nucleotide). (B) Radioactive TRAP results. N, negative control. (C). Illustration of the sequencing result of TRAP product from Krot-0. Arrow indicates the putative translocation position for Krot-0 telomerase.

Analysis of the telomerase entry site and synthesis fidelity

Previous results indicated that the fidelity of telomerase synthesis is determined by the relative stability of the primer-template complex (Fitzgerald et al., 2001). If A2 is not employed for DNA synthesis by Krot-0, 5' boundary of the template is restrained. We tested whether this restraint affects the entry site selection during formation of a primer-template complex by comparing sequence results for Krot-0 with two *A. thaliana* accessions that do not contain template polymorphisms Col-0 and Ler-0. For these experiments, we used a primer terminating in CAG, which has two nucleotides of complementarity to the template, and is expected to align at positions U9 and C8 for the first round of synthesis. Sequence analysis is consistent with this prediction showing that in all three accessions C8 in template is the entry site (Figure 5-4A). Again Krot-0, like Col-0 and Ler-0 gave rise only TTTAGGG repeats synthesized (Figure 5-4B), indicating that the A2 polymorphism does not affect the telomerase fidelity.

A



B

CAG-primer	Entry site (Frequency)	# of Clones	Error/nucleotide sequenced
Col-0	CAGGG(TTTAGGG)$_n$ (100%)	13	0/868
Ler-0	CAGGG(TTTAGGG)$_n$ (100%)	10	0/1323
Krot-0	CAGGG(TTTAGGG)$_n$ (100%)	22	0/2317

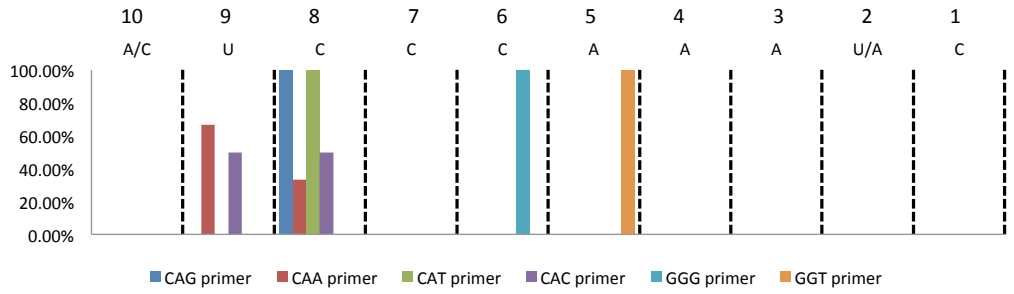
Figure 5-4. The entry site and fidelity of telomerase elongation is not changed in an accession bearing a 5' polymorphism in the TER1 template. (A) The distribution of the entry sites. The percentage indicates the frequency of nucleotides used for primer annealing among all clones. (B) Summary of the entry site sequence and error rate of elongated products.

To further explore the impact of the 5' SNP in the Krot-0 template on telomerase entry site selection and fidelity, TRAP was conducted with a series of primers bearing different 3' sequences. With a primer ending with CAT, the entry site primarily falls into C8 (Figure 5-5A, green bar). Interestingly, in one clone, the first full repeat was GGTTTAGGG (T-slippage), precluding the exact determination of the entry site (Figure 5-5B, highlighted in red in CAT row). In contrast to the CAT primer, primers ending with CAA or CAC showed varied distribution of the entry site. Three out of seven clones in both cases had an entry site of U9 (Figure 5-5A, maroon and purple bar). The CAA 3' terminus can pair with U9, while the CAC primer ending with C should not. Importantly, neither primer has the potential to form two base pairs like the CAG primer. The error rates for all CAN primers were increased relative to CAG (Figure 5-5B). This error rate was also 10 fold higher than the previously reported *A. thaliana* telomerase error rate (Fitzgerald et al., 2001). T-slippage was observed with CAN primers, while G-slippage was associated with CCC extension (Figure 5-5B; Figure 5-6C). These results are consistent with the previous analysis of the Col-0 telomerase error rate using primers ending with A, T and C (Fitzgerald et al., 2001). Altogether, our results indicate that the base pairing potential between the primer 3' terminus and the template impacts synthesis fidelity. Strikingly, we failed to detect evidence that 2A served as a template with any primer sequence. We conclude that this residue is not used as a template for Krot-0 telomerase.

If the absence of primer 3' base pairing is sufficient to cause an increased error rate for telomerase, increased base pairing potential should increase fidelity. Using primers ending with GGG or GGT, we found that the entry site shifted to the expected annealing site in the middle of the template (Figure 5-5A). Unexpectedly, the error rate

was higher with GGG primer (2/490) than GGT primer (0/385), even though the GC base pair at the 3' primer-template junction will be stronger than the AT base pair (Figure 5-5B). These findings suggest that the initial base pairing between primer and template is not sufficient to guarantee synthesis fidelity.

A



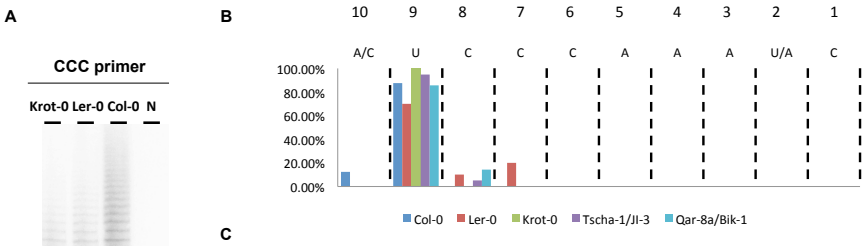
B

	Entry site (Frequency)	# of clones	Error/nucleotide sequenced	Error rate	Error type
CAG	CAGGG(TTTAGGG)$_n$ (100%)	22	0/2317	0%	N.A.
CAA	CAAGG(TTTAGGG)$_n$ (25%) CAAGGG(TTTAGGG)$_n$ (50%) CAAGGT(TTTAGGG)$_n$ (25%)	4	1/329	0.31%	T-Slippage in the first repeat
CAT	CATGG(TTTAGGG)$_n$ (75%) CATGGT(TTTAGGG)$_n$ (25%)	4	1/252	0.40%	T-Slippage in the first repeat
CAC	CACGG(TTTAGGG)$_n$ (33.3%) CACGGG(TTTAGGG)$_n$ (33.3%) CACGGT(TTTAGGG)$_n$ (33.3%)	6	2/364	0.55%	T-Slippage in the first repeat
GGG	GGG(TTTAGGG)$_n$ (100%)	4	2/490	0.41%	G-Slippage after 2 repeats and 4 repeats
GGT	GGTTTAGGG(TTTAGGG)$_n$ (100%)	4	0/385	0%	N.A.

Figure 5-5. The relationship between primer 3' sequence, telomerase entry site and fidelity in Krot-0. (A). Entry site distribution. Primer 3' ends are shown. (B). Summary of the entry site sequence, error rate and types of errors. The red color highlights entry site sequence showing the T-slippage. Because of entry site ambiguity, these sequences were omitted from the analysis in A.

Primers that cannot basepair with the template are extended with decreased fidelity

Since enhanced base paring formed between the primer 3' terminus and the template did not reduce the error rate, we examined how 3' end lacking any complementarity to the template were used. Extracts from several *A. thaliana* accessions were subjected to TRAP using a primer terminating with 3' CCC. The 3' CCC will not base pair with any nucleotide in the template. All of the accessions we tested could extend the 3' CCC primer (Figure 5-6A). Surprisingly, however, instead of the C8 as default entry site for *A. thaliana* (using primer with 3' terminating with CAC) (Fitzgerald et al., 2001), the CCC primer was positioned predominantly at U9 (Figure 5-6B). Comparison of results between the CAC primer and the CCC primer showed different entry site, indicating a contribution of the penultimate nucleotide in initiating primer positioning. Notably, the frequency and nature of telomerase errors varied among the different *A. thaliana* accessions using the 3' CCC primers. For example, G-slippage was associated with Col-0, while nucleotide misincorporation was seen in the Krot-0 and Qar-8a/Bik-1. The type of misincorporation did not correlate with the nucleotide polymorphisms in the template sequence. Instead, the variation in primer utilization appears to reflect natural variation of telomerase fidelity that is influenced by unknown factors (Figure 5-6C; Table 5-1).



C

CCC primer	Entry site (Frequency)	# of clones	Error/ nucleotide sequenced	Error rate	Error type
Col-0	CCCGGG(TTTAGGG)n (7/9) CCCAGGG(TTTAGGG)n (1/9) CCCGGGG(TTTAGGG)n (1/9)	9	1/700	0.14%	G-Sippage in the first repeat
Ler-0	CCCG(TTTAGGG)n (2/11) CCCGG(TTTAGGG)n(1/11) CCCGGG(TTTAGGG)n (7/11) CCCGTGG(TTTAGGG)n (1/11)	11	2/616	0.32%	T-Slippage in the first repeat; misincorporation: GGG(TTTAGGG)4CTTAGGG(TTTAGGG)6
Krot-0	CCCGGG(TTTAGGG)n (23/23)	23	1/1939	0.05%	Misincorporation: GGG(TTTAGGG)CTTAGGG(TTTAGGG)6
Tscha-1/JI-3	CCCGGG(TTTAGGG)n (19/20) CCCGG(TTTAGGG)n(1/20)	20	2/1295	0.16%	Misincorporations: GGG(TTTAGGG)2TTTAGAC(TTTAGGG)15
Qar-8a/ Bik-1	CCCGGG(TTTAGGG)n (18/21) CCCGG(TTTAGGG)n(3/21)	21	0/1127	0.00%	N.A.

Figure 5-6. Analysis of telomerase products generated with a primer ending with CCC. (A). Radioactive TRAP results. (B). Entry site distribution. Tscha-1/JI-3 is an accession with 5' 2A polymorphism. Qar-8a/ Bik-1 are two accessions with 3' 10C SNP. (C). Summary of the entry site sequence, error rate and types of errors.

Accessions	Nucleotides sequenced	Error rate	Clones with errors/total clones	Errors in first 2 repeats/total errors	G-Slippage	T-Slippage	Misincorporation
Col-0	1568	6.38E-04	1/22	0/1	1	0	0
Ler-0	1939	1.03E-03	2/21	1/2	0	1	1
Krot-0	6076	1.15E-03	7/70	6/7	2	4	1
Tscha-1/JI-3	1295	7.72E-04	1/20	0/1	0	0	1
Qar-8a/Bik-1	1127	8.87E-04	1/21	0/1	0	0	1

Table 5-1 Summary of the error distribution in sequenced clones among different accessions using all primers sequences.

Discussion

To guarantee the fidelity, flexibility, and processivity of telomerase, the TER template must be accurately engaged by the DNA primer. Each nucleotide position in the template domain appears to have a designated function. Nucleotides closer to 3' end of the domain are used for primer annealing, while the 5' nucleotides are templating residues used for extension. How these nucleotides cooperate to fulfill the enzyme function of telomerase is not well understood. In this study, we investigated the natural polymorphisms in the TER1 template of *A. thaliana* telomerase. Our analysis increases understanding of the telomerase template function and enzyme fidelity.

A minimized TER1 template in A. thaliana

The TER template is typically 1.5 copies of the complementary telomere repeat sequence. According to this standard, 5'-CUAAACCCUA-3' is expected to be the TER template for *A. thaliana* to synthesizing TTTAGGG repeats (Cifuentes-Rojas et al., 2011). Here, we took a bioinformatics and biochemical approach to examine the contribution of 5' and 3' SNPs within the TER1 template domain. These two naturally occurring polymorphisms provide us with an opportunity to define the boundaries of the functional template. Sequence analysis of telomerase products from accessions with these template SNPs indicate that telomerase does not reverse transcribe the variant residue at the 5' U2 position (e.g. Krot-0, 6076 nt sequenced). The data allow us to conclude that the functional template spans A3 to A10 in Krot-0, and perhaps all *A. thaliana* accessions bearing the 5' polymorphisms. Mutations is bona-fide templating nucleotides cause a dramatic decrease in fidelity (Gilley et al., 1995); (Blasco et al., 1995). The observation that Krot-0 telomerase synthesis perfect TTTAGG repeats

implies that the template boundary of the template is defined by a shift in the relative position of the template boundary element (TBE). The TBE is a helical structure immediately adjacent to the 5' template region among different TERs (Tzfati et al., 2000; Seto et al., 2003; Chen and Greider, 2003). Disrupted base pairing in TBE helix results in nucleotide incorporation beyond the template. Our results predict that the Krot-0 TER1 5' TBE overlaps with the template sequence. Notably, this is the case for the TER1 TBE in fission yeast (Box et al., 2008b). Our preliminary result from SHAPE assay in TER1 structure suggests the nucleotide at 5' template in Col-0 TER1 is restrained, indicative as the TBE. Additional experiments are needed to test the template structure of TER1 from accessions bearing SNPs.

Our sequencing results showed no significant telomerase misincorporation in two ecotypes bearing 3' template SNPs (e.g., Qar-8a/Bik-1, 0%). The proposed 5' translocation position (3A) defined from Krot-0 cannot be used for Qar-8a/Bik-1 since there would be no possibility for primer annealing after translocation. The 5' boundary must be extended further upstream, so that the functional template is shifted to U2-U9 or C1-C8 (Figure 5-7A). Thus, our data suggest that the template boundaries are flexible in *A. thaliana,* and the active site has shifted with the 5 MYA since *A. thaliana* diverged from its last common ancestor (Beilstein et al., 2010) (Figure 5-7A).

Notably, the proposed functional 5' and 3' boundaries make a short TER1 template length with only eight nucleotides, one for annealing and seven for extension. This situation is not without precedent. The TER template domain in mouse is only 7nt long with one annealing nucleotide and six residues for elongation (Blasco et al., 1995). The efficiency for annealing is not highly efficient with the potential to form only one base pair after translocation. Lessons from human telomerase indicate that the

interaction of telomerase with primer substrate is stabilized mainly by contacts with the TERT subunit. Indeed, base paring between primer and template contribute only minimally to its stability ((Wallweber et al., 2003; Wu and Collins, 2014)). Similarly, in budding yeast C-terminus of TERT is crucial for template-primer stability as well as template usage (Hossain et al., 2002). In human TERT, mutations in both the C-terminus and N-terminus alter template (Moriarty et al., 2005). Therefore, TERT may play a key role in setting template boundaries and defining template-primer stability in *A. thaliana* TER1. Since TERT is very highly conserved across *A. thaliana* accessions, it indicates that other polymorphisms within TER1 or telomerase accessory proteins may have co-evolved with the template polymorphisms to refine the highly stereospecific contacts within the active site, and allow high fidelity of telomerase repeat synthesis (Figure 5-7B).

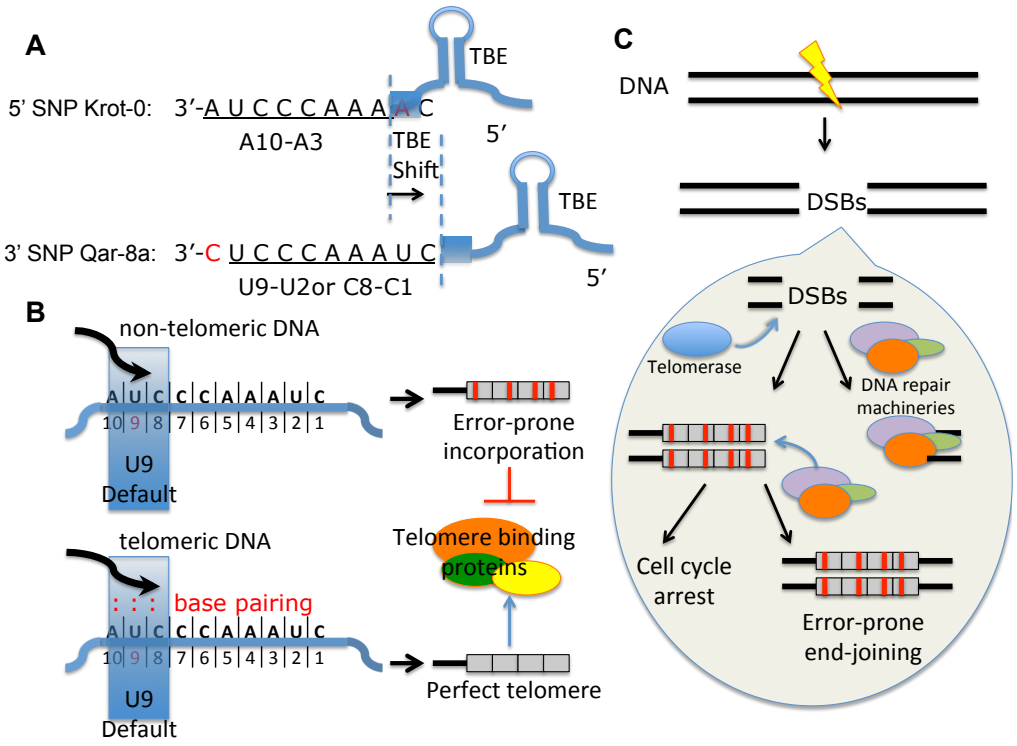


Figure 5-7 Template utilization, telomerase fidelity and DNTF. (A) Telomerase templating nucleotides vary among accessions. Underlined nucleotides indicate the utilized nucleotides, A10-A3 in 5' SNPs accessions and U9-U2 or C8-C1 in 3' SNPs accessions. A potential TBE adjacent to the 5' junction may shift to define the 5' templating nucleotides (Blue block). Boundaries define the functional telomerase RNA template with only eight nucleotides, one for annealing after translocation and seven for elongation. (B) Two mechanisms maintain the stability of the initial primer-template complex. At telomerase active site, the primer including non-telomeric DNA (black arrow line) was positioned by an unknown mechanism with not sufficient stabilizing effect for the primer-template complex, resulting in non-perfect telomere synthesis. (Top). However, the extra base pairings formed alongside with the positioning stabilize the primer-template complex, generating the perfect telomere repeats (Bottom). The extent of the error nucleotides incorporated determines the success in telomere binding proteins association at the newly synthesized products. (C) Telomerase error rate and its significance to DNTF. Upon DNA damage, the DSBs could either processed by DNA repair machineries or serve as substrate as telomerase. Given the DSBs site are typically not telomere like sequence, the stability of the primer-template serves as a mechanism to form the error prone telomeric DNA synthesis by telomerase DNTF. Lacking the telomere binding proteins for stabilization, the newly formed telomeres (with error) prolong the activation of the DNA damage response until getting repaired (error-prone end-joining). Otherwise, cell cycle will be arrested to cull cells with un-repaired DNTFs.

Two mechanisms contribute to primer-template stability

Telomerase is a highly promiscuous enzyme, soon after the discovery of Tetrahymena telomerase, it was realized that the enzyme can elongate primers corresponding to telomere repeats of other organisms (Greider and Blackburn, 1987; Greider and Blackburn, 1989), and more importantly completely non-telomeric DNA whose 3' terminus is not complementary to RNA template (Harrington and Greider, 1991). This enzyme promiscuity was later found in other organisms (Diede and Gottschling, 1999); (Kramer and Haber, 1993; Flint et al., 1994) and argued that a "default" position must exist in the RNA template to initiate elongation on the non-telomeric DNA 3' end. An anchor site within the TERT subunit has been proposed to deliver non-telomeric primers to the default position independent of Watson-crick base pairing (Melek et al., 1996; Wang and Blackburn, 1997). Here, we show that for *A. thaliana* TER1, U9 serves as a default annealing site to initiate the telomere synthesis on non-telomeric 3' ends. This entry site is used by all the accessions we tested, but is not strict, as C8 and A10 can be used less frequently. Interestingly, the extent to which these alternative sites are used differs among different accessions. However, we found that 5' and 3' SNPs in template domain did not impact the default position, suggesting that a structure, not sequence determines how the non-telomeric DNA 3' terminus is positioned.

Notably, the default annealing site uncovered in this study differs from the previous report for *A. thaliana* Col-0 telomerase. One explanation is that two studies used non-telomeric primers with different 3' terminal sequences, CCC in this study and CAC in the former study. There are two considerations to discuss. One is base-pairing (annealing), and the other is sequence-independent positioning of the primer 3'

terminus. When the penultimate nucleotide in the primer 3' terminus is A, it can anneal with U9 in the template. Therefore, CAC primer in Fitzgerald's study may not anneal with the C8 in the template. Instead, U9 anneals with the penultimate A in the primer, together with primer 3' end nucleotide positions in a less optimal site at C8. Fitzgerald's result is reproducible in this study. When using the primer ending with "CAN" (N represents in any one of A, T, C and G) for Krot-0 protein extract, most of the entry sites are C8 as Fitzgerald observed (Figure 5-5).

We hypothesize that primer positioning and base pairing (annealing) both contribute to primer-template complex stability at the active site in *A. thaliana* TER1. Telomerase has a higher error rate when extending non-telomeric DNA sequence (Fitzgerald et al., 2001; this study). Thus, the selection of the entry site for both telomeric and non-telomeric primers has evolved to maximize the primer-complex stability. When the 3' nucleotide in the primer has a weaker or no base pairing potential for the template, the error rate increases (Figure 5-5, 5-6). In this case, the positioning effect alone is not sufficient to stabilize the primer-template complex. Likewise, our data argue that strong base pairing potential alone is not sufficient to initiate a stable primer-template complex that guarantees the synthesis fidelity. Primers ending with GGG and GGT should form stable base pairing at the middle of the template, C6-C8 or C5-C7, respectively. However, extension of the primer GGG primer leads to a similar level error rate as with primers bearing reduced 3' end paring potential (Figure 5-5). These findings indicate that both base pairing and sequence-independent positioning are required for the initial primer-template stability to guarantee the fidelity of synthesis.

Telomerase error rates and genome stability

Previous results showed the telomerase error rate varies significantly (Fitzgerald et al., 2001). Our data indicate that accessions within *A. thaliana* do not show very much variation in error rate, which are averagely $0.8\text{-}0.9 \times 10^{-3}$ ($+/- 0.2 \times 10^{-3}$) (Table 5-1). This value is similar to previous determined *A. thaliana* telomerase error rate, which was 1.8×10^{-3} (Fitzgerald et al., 2001). A similar error rate for telomerase across *A. thaliana* accession is not surprising given a short period evolution. However, the error rate differs when using different primers with various 3' termini. The highest fidelity seen in primer ending with G probably reflects this primer could give a maximum stability of the primer-template complex with the optimal entry site (Figure 5-7B).

Taking non-telomeric synthesis error rate into account, telomerase synthesis at normal chromosome ends should have an even higher fidelity. The increased rate of misincorporation by telomerase in extending non-telomeric DNA may be advantageous, helping the cell to differentiate true telomeres from DNTF at DSBs. The higher error rate associated with DNTF may not allow the association with the sequence-specific telomere binding proteins that would stabilize the new telomere. The less-telomeric like sequence by the error prone non-telomeric synthesis will allow more time for DNA repair, although repair will not be faithful (Figure 5-7C). In the absence of repair it could activate a DNA damage response that finally causes cell cycle arrest. The net outcome would be the promotion of genome stability by culling cells in which DNTF has occurred without a proper handling.

Materials and methods

Sequence acquisition and analysis

Sequences corresponding to TER1 and TER2 were obtained using the genome browser at http://signal.salk.edu/atg1001. Two sequences were removed from the analysis. Hov 3-2 was removed because it was the only accession with two deletions in the 5' end, corresponding to 20nt from the 5' start of TER2, and a 100nt deletion starting at nucleotide #101. The template region was not disturbed in this accession, possibly indicating a functional TER2 is generated. The Tottarp-2 accession was removed because the sequence corresponding to the search region did not contain sequences corresponding to TER2, most importantly, a template region. Sequences were trimmed in MEGA5, and then analyzed using Geneious v6.0 (Biomatters). Sequence conservation and alignments were performed using Geneious.

Nucleic acid extraction, genotyping and PCR

DNA samples were prepared from the leaves of different accessions. PCR samples were resolved in 1% agrose and gel purified and sequenced. Primers for TER1 were Chapter II.

PETRA assays

1mg DNA was used in the assay. The subtelomeric primer for 2R and 4R were used. PETRA was performed as previously described (Heacock et al., 2004).

TRAP assays and sequencing of TRAP products

TRAP assays were resolved using a polyacrylamide gel (Fitzgerald et al., 1996). The product were precipitated by ethanol and resuspended in 4ml nuclease free water for cloning. The ligation reactions were performed using TOPO TA pCR2.1 cloning vector (Invirogen, USA). The products were transformed in TOP10 competent cells. After blue/white screening on kanamycin plates, colonies were selected up and set up for miniprep. The constructs were sent for commercial sequencing in Eton using two different primers M13F-21 and M13R in the backbone of the vector.

CHAPTER VI

CONCLUSIONS AND FUTURE DIRECTIONS

The essential mission of life is to guarantee the fidelity of genetic information as it is passed on to the next generation. While linear chromosomes evolved as a highly successful means of storing massive amounts of genetic information, a disadvantage of linear DNA is that it introduced the end replication/maintenance problem. On one hand, DNA sequences must be continually added to circumvent limitations of the conventional DNA replication mechanism. On the other hand, terminal DNA ends must be differentiated from damaged DNA, specifically DSBs, so the genome can be stable. The machinery of telomeres and telomerase provides an exquisite solution to overcome these challenges.

The cases of human aging and cancer illustrate the importance of telomeres and telomerase. Shortened telomeres cause cell aging and senescence, which can be delayed by the action of telomerase. However, telomerase must be finely regulated. Insufficient telomerase activity limits cell proliferation and decreases stem cell viability, resulting in several severe human genetic disorders. In contrast, hyperactive telomerase is a hallmark of human cancers. Moreover, inappropriate action of telomerase at DSBs to cause DNTF can be lethal. Thus, much effort has been invested in understanding mechanisms of telomerase regulation.

A new mode of telomerase regulation evolved when TER was duplicated in *Arabidopsis thaliana.* TER1 and TER2 have very distinct functions in telomerase metabolism. The TER1 RNP maintains telomere length as the canonical telomerase RNP, while TER2 negatively regulates telomerase upon DNA damage (Cifuentes-Rojas

et al., 2011; Cifuentes-Rojas et al., 2012). TER2 associates with several proteins involved in the DDR and DNA repair pathways including Ku and potentially ATR (Song et al, unpublished data). It is possible that when DNA repair machinery and telomerase meet at a DSB, TER2 could be used as a weapon in the arsenal of DNA repair activities to repress telomerase and thereby prevent DNTF.

Based on these findings, several questions remain. How is TER2 integrated into the DNA damage response? How did TER2 acquire the ability to repress telomerase? How was the regulation of TER2 established? What is the biological significance of TER2? In this dissertation, I examined these questions as well as the evolution of the TER1 templating domain. In this last chapter, I will discuss the conclusions from my research and exciting future directions.

TER2 may be an apical layer regulator coordinating signals from several telomere-related cellular events

One of the important discoveries of this work is a link between TER2 and reproductive fitness (Chapter III). It is unknown if the reproductive regulation by TER2 is caused by the full-length TER2 or TER2D /TER2s, because it remains unclear why some accessions harbor the TER2 TE and others do not. Generating a transgenic plant with full-length TER2 or TER2D expressed from its native promoter will help solve this mystery. Based on the interesting result of *pot1a ter2* (that is, an accelerated rate of telomere shortening) (Chapter III), and some preliminary evidence from analysis of the *ku ter2* mutants (see below), I predict that TER2 serves as a telomerase-independent signaling molecule. Specifically, I propose that TER2 functions as an upper level regulator to integrate and coordinate several related cellular events: telomere

regulation, telomerase regulation, DNA damage sensing and reproduction (meiosis) (Figure 6-1). There are numerous studies showing a correlation between meiosis, telomere capping and telomerase regulation (Arnoult and Karlseder, 2014; Siderakis and Tarsounas, 2007; Tomita and Cooper, 2007). Having a lncRNA such as TER2 as the hub for modulating different aspects of chromosome biology would be an efficient and sensitive approach. TER2 could potentially integrate signals from telomere-related events. Not only does the *pot1a* genetic background sensitize the *ter2* phenotype (increased seed abortion), the telomere length maintenance defect in *pot1a* mutants is more severe in the absence of TER2 (Chapter III). Since telomere length is unperturbed in the *ter2* mutants (Cifuentes-Rojas et al., 2012), these results indicate that the aggravated telomere shortening phenotype in *pot1a ter2* double mutants could be the result of miscommunication (signal feedback) between TER2 and POT1a as well as between TER2 and other pathway players.

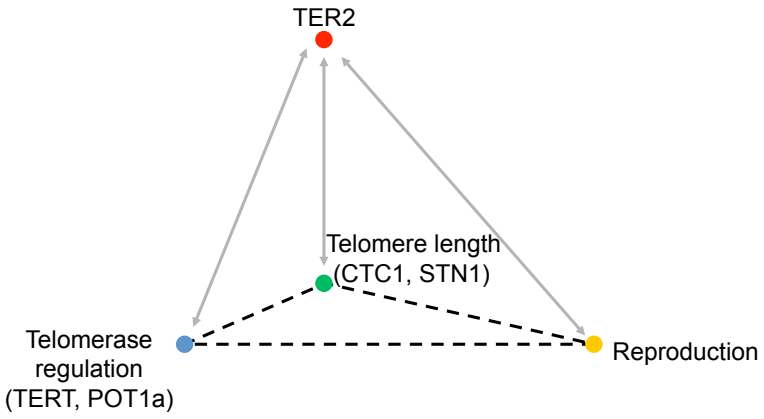


Figure 6-1. Model for TER2 signaling to coordinate several regulatory pathways.
See text.

To test if TER2 has evolved a signaling role for coordinating telomeres, telomerase, and reproduction behaviors, further genetic experiments would be beneficial. For example, if TER2 is a telomerase-independent signaling molecule, *ter2 tert* double mutants should be examined. The cross has been made and the F1 seeds have been collected. Future work should focus on analysis of telomere length as well as seed viability of the F2 and F3 segregants. Synergistic effects as seen in *pot1a ter2* double mutants are expected. If TERT and TER2 work in different genetic pathways, I expect to see further telomere shortening in *tert ter2* double mutants compared to *tert* single mutants. However, if the results show TERT and TER2 work in same genetic pathway, I expect to see telomere shortening at the same rate in *tert* single mutants and *tert ter2* double mutants. If so, this will indicate that a signaling function of TER2 is specific to POT1a. It is very possible that the *tert ter2* double mutants will not be viable. If this is the case, the experiment will still indicate whether TERT and TER2 have a synergistic effect for seed viability. Statistics can be carried out to provide more information on potential haploinsufficiency of the two genes for seed viability. Recent preliminary data indicate that TERT may be haploinsufficient for telomere length maintenance in the absence of TER2 (J. Townley, H. Xu and D. Shippen, unpublished data).

To see if TER2 plays a broader role beyond telomerase, plants lacking components of the CST capping complex should be crossed with *ter2* mutants. *ter2 ctc1* and *ter2 ten1* double mutants have already been made for this purpose. Like *pot1a ter2* and *tert ter2*, telomere length and seed abortion phenotypes will be checked in these double mutants. I predict to see synergistic effects in telomere shortening and seed abortion. If true, it will suggest that TER2 is a general signaling molecule that

monitors very short telomeres, perhaps to ensure that cells with dysfunctional telomeres are culled out from the population.

Another interesting cross to examine is *ter2* X *ku70* (Xiaoyuan Xie, in progress). In plants lacking Ku, telomeres become longer (Riha and Shippen, 2003). It will be interesting to know if TER2 senses the long aberrantly long telomere in *ku* mutants. If so, telomeres may be longer or shorter in *ku70 ter2* mutants. Lastly, a *ter2 ku pot1a* triple mutant, if viable, will be interesting to study to understand if and how much TER2 can integrate a dysfunctional telomere signal and react properly when representatives of two important telomere pathways are perturbed. Results from these experiments will provide new insight into the biological significance of TER2 as a lncRNA signal for Arabidopsis.

The TE gives TER2 its distinct functions

My research showed that TER2 has many functions: RNA destabilizer, sensor for DNA damage, telomerase regulator and promoter of reproductive development (Chapters II and III). Most, if not all, these functions are derived from the transposable element. As discussed above, I hypothesize that TER2 is a master regulator molecule that integrates signals from many telomere-related events (Figure 6-1). To further explore this hypothesis, it would be interesting to dissect which region of TER2 carries each individual function. Elucidating the structure of the TE within TER2 will be helpful to set the stage for detailed mutagenesis studies in the future (Figure 6-2) (Johnson and Guigo, 2014).

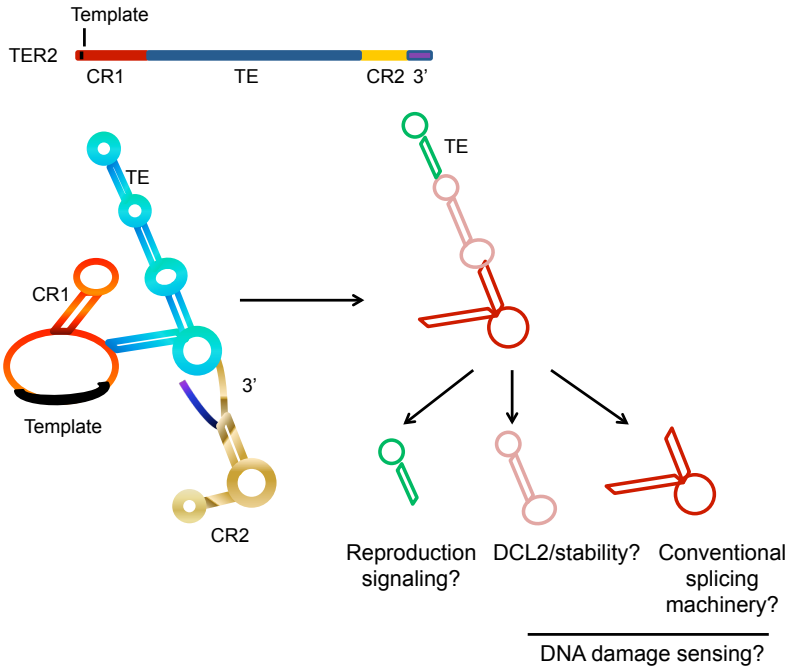


Figure 6-2. The TE is a discrete regulatory domain. The secondary structure of TER2 RNA gives rise to the specific function of this RNA in DNA damage sensing, reproduction regulation and its processing. These functions likely correlate with specific motifs within the TE. Dissecting the functional motifs may reveal insight into RNA function. Therefore, it will be helpful to understand the evolution of the RNA by pinpointing important nucleotides important in the TE.

Based on results with the *pot1a ter2* mutants, it is possible that TER2 can sense shortened telomeres, due to the lack of POT1a as a DNA damage, and initiate downstream events that avert potential harmful effects. I found that TER2 is activated (stabilized) in the presence of DSBs, leading to telomerase down regulation (Cifuentes-Rojas et al., 2012; Chapter III). Therefore, among the different TER2 functions, DNA damage sensing ability could be the major driver of TER2 function.

How did TER2 gain the ability to sense the DNA damage? The first step may have been establishment of a response to regulate TER2 abundance when facing the

proper signal. The basal level of TER2 transcript is very low, caused by the instability of this RNA (Chapter II). Therefore, to understand the DNA damage sensing ability of TER2, a first step should be to define the region of TER2 responsible for RNA instability. Different mutations or deletions in TER2 (specifically in the TE) should be made. A secondary structure model for full-length TER2 and the TE will be informative in designing single nucleotide mutations and sequence deletions. Mutant constructs should be transformed into *ter2* null mutants, followed by RNA stability assays using cordycepin to assess regulatory function. It would also be helpful to perform the TRAP assay in each transgenic line to determine if motifs that affect RNA stability also affect telomerase regulation. Last, the seed abortion rate should be checked to see which constructs rescue the reproduction defect. These experimental strategies could then be used for any mutation construct to define function of each specific region or structure in TER2.

Genetic data from *lsm8* and *dcl2* mutants may provide further insights into TER2 regulation. Genes involved in TER2 processing may contribute to TER2 stability, as shown for mRNA. Among the aforementioned design of mutants that could affect RNA stability, the regions in TER2 that interact with DCL2 or LSM8 could be targets for mutations as a further verification (DCL2 part, discussed below). Based on the analysis of TER2 in *lsm8* mutants (Chapter IV), mutating splicing junctions could be useful to help answer if and how the TER2 TE is processed by canonical mRNA splicing. For specific mutations affecting TER2 intron removal, we expect to see a higher abundance of the TER2 precursor in plants that have mutations in TER2 TE boundary sequences. qTRAP could be performed to determine if telomerase repression is caused by the

build-up of TER2 produced as a result of failed intron splicing (see DCL2 discussed below).

TER2 as a potential target for imprinting

Imprinted gene expression is the biased expression of alleles dependent on their parent-of-origin (Moore and Haig, 1991). During the process of characterizing *ter2 tert* double mutants, I found that the parent-of-origin had an effect on the success of the cross (Chapter III). Almost all crosses using female *ter2* with male *tert* successfully gave F1 offspring seeds. However in striking contrast, the reciprocal cross using male *ter2* and female *tert* produced very few F1 seeds. Those that were produced were tiny and did not germinate

These results raise the interesting possibility that the TER2 locus is subject to imprinting. Imprinted expressed genes are known to be associated with short TEs as a consequence of epigenetic regulation, and probably arise from the genome defense system against TEs (Chung et al., 2008). Furthermore, imprinting is dependent on *de novo* DNA methylation via a 24nt sRNA derived from the TE working with other small RNA machineries in the RNA dependent DNA methylation (RdDM) (Castel and Martienssen, 2013). Although my preliminary data did not show evidence of TER2 expression changes in RdDM mutants, a 24nt sRNA is associated with TER2 (Chapter IV). Therefore, it is possible that the level of TER2 is regulated transiently in only a small group of cells during reproduction events. My experiments were conducted with flowers and so I cannot exclude the possibility that TER2 is regulated via imprinting in a specific stage of embryonic development. It is also possible that other imprinted genes could indirectly affect TER2 in an as yet unknown manner.

An imprinting model for TER2 regulation proposes that TER2 expression will be disproportionately affected by the parental allele in the F1 generation (Figure 6-3). This idea can be tested in a *ter2* mutant and wild type Col-0 reciprocal cross. The F1 generation should all be heterozygotes for TER2. In one set, the wild type allele will be paternal, while the other set F1 contains a maternal wild type allele. Since the allele from *ter2* mutant will not produce TER2 transcript, we then expect to see a difference in the expression level of TER2 in the two experiments. To investigate the contribution of DDM1 (methylation of DNA), the same experimental strategy could be used with a *ddm1* mutant in reciprocal crosses to *ter2* mutants. The result would tell us if the parent-of-origin effect is dependent on DNA methylation.

To determine if the parent-of-origin regulation is via the regulation of the TER2 promoter, a reporter line with the TER2 promoter fused to GUS can be reciprocally crossed with a wild type plant. GUS staining intensity will reveal a potential imprinting effect. If these experiments show a difference depending on the parent-of-origin, it will indicate that the regulation of imprinting is targeted to the promoter region of TER2. Otherwise, the possibility of an imprinting effect targeted to the TER2 gene body would remain.

Next, we could ask if imprinting regulation of TER2 is sufficient for embryogenesis. To achieve this goal, seed abortion analysis will be performed in the lines generated as above (wild type x *ter2,* or *ddm1x ter2*). By comparing the result from *tert ter2,* we can determine if the imprinting effect could be associated with the *TERT* gene or via a synergistic effect of the *TERT* gene loss. Finally, to determine if the imprinting regulation is dependent on regulated telomerase activity, qTRAP assays should be performed in these lines to see if there is a correlation of telomerase activity

and the imprinting modulation of TER2. Altogether, these results will give a better idea of regulatory mechanisms of this newly evolved lncRNA, TER2.

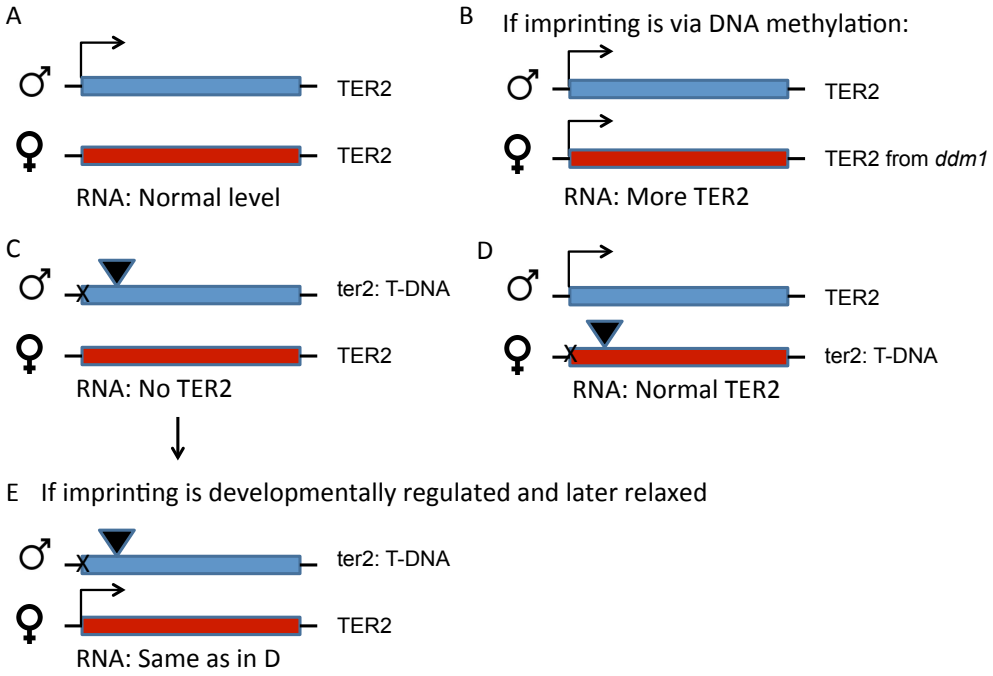


Figure 6-3. Models for imprinting regulation of TER2 during embryogenesis. (A). TER2 is specifically expressed from the paternal locus. The maternal locus is silenced. (B) If the silencing effect is from DNA methylation, the maternal TER2 locus from the DNA methylation maintenance mutants plant (*e.g. ddm1*) will not be silenced and more RNA is expected. (C). In the F1 generation from the wildtype Col-0 and *ter2* (in Col-0) cross, the T-DNA (black triangle) inserted at the paternal TER2 locus will not express TER2. Combined with the silencing effect on the maternal TER2 locus, no TER2 expression is expected. (D). In the reciprocal cross, TER2 will be expected. (E). If imprinting occurs only in embryogenesis, the silenced maternal TER2 locus will become active to express TER2. Therefore, the level of TER2 will be same between the reciprocal cross in the mature F1 plant tissue.

DCL2 regulates TER2 using a novel mechanism

During the course of identifying the pathways involved in TER2 metabolism, I found that the small RNA processing machinery Dicer-like 2 (DCL2) is involved in TER2 regulation (Chapter IV). While exciting, it was not surprising to learn that TER2 is regulated by small RNA machinery, since I found that the TE in TER2 is derived from a transposable element. TEs are widely repressed by small RNA mediated DNA methylation (RdDM) (Castel and Martienssen, 2013). However, my results suggest that DCL2 has an unconventional function in TER2 modulation. The regulation of TER2 by DCL2 may not happen by transcriptional regulation via the RdDM pathway. Instead, I found that TER2 RNA stability was affected in plants lacking DCL2, and this post-transcriptional regulation influenced TER2 abundance during reproduction. Although these exciting results indicate a novel mechanism for DCL2 in RNA processing, little is known about the mechanism of this function. Here, I will discuss several possible roles for DCL2.

DCL2 and telomere metabolism

We previously showed that overexpression of TER2 represses of telomerase activity, resulting in telomere shortening (Cifuentes-Rojas et al., 2012). Since TER2 is elevated in *dcl2* mutants (Chapter IV), one potential effect is that the increase of TER2 causes insufficient telomerase activity, resulting in telomere shortening. To test this idea, telomerase activity should be measured in *dcl2* mutants, where I expect to see a decrease in telomerase activity. If this is true, *dcl2* mutants can be subjected to primer extension telomere repeat amplification (PETRA) or terminal restriction fragment (TRF) analysis to assess telomere length. Plants lacking DCL3 can be used as the control,

which is expected to have no telomerase and telomere changes. It is possible that there will not be a dynamic change of telomerase and telomere length in *dcl2.* One explanation is that the interaction of DCL2-TER2 is only required at a specific time during development. However, if we see a reduction in telomerase activity and telomere shortening in *dcl2,* it will be important to ask if these changes are via TER2. The *dcl2 ter2* double mutants should be made and analyzed by qTRAP and PETRA/TRF assays. These experiments may establish a link between small RNA machinery, telomere metabolism and the processing of TER2.

DDR and TER2 regulation by DCL2

Another important finding from my research is that TER2 abundance increases upon DNA damage and this DDR- induced TER2 increase is caused by stabilizing TER2 (Chapter II). Similarly, insufficient DCL2 causes the up-regulation of TER2 by increasing TER2 stability (Chapter IV).Unexpectedly, I found that the TER2 profile was altered in *dcl2* mutants during reproduction, when a DDR is induced during meiotic recombination. Therefore, dynamic interactions are suggested among TER2, DCL2 and DNA damage. Based on these facts, I hypothesize that DCL2-mediated TER2 regulation is modulated by the DDR. To test this, TER2 abundance should be assessed in *dcl2* mutants treated with DNA damage reagents (e.g. zeocin). I expect to see an increase in TER2 RNA level in *dcl2* mutants with no further increase of TER2 upon DNA damage. In addition, telomerase activity should decrease in the absence of DCL2, with no additional decrease upon DSB induction. If it is determined that TER2 is increased at a higher level in *dcl2* mutants treated with zeocin (and telomerase repression increased), this will indicate that DCL2 is not part of the DDR that regulates TER2. On

the other hand, if TER2 increase without telomerase repression, it will imply that DCL2 is part of the DDR for regulating TER2 by degradation, but not functional processing for its role in telomerase regulation. These results will lay the foundation for future investigations of TER2 induction in response to DNA damage and its role in telomerase regulation.

DCL2-TER2 interaction in vitro

It is not known if DCL2 directly binds to TER2. Biochemistry tests to assess DCL2-TER2 interaction may set a new paradigm of DCL2 (and other Dicer-like proteins) in lncRNA metabolism. Dicer-like proteins contain a PAZ domain (non-specific RNA binding), a dsRNA binding domain (RBD), two RNase III domains, and a Dicer domain (for dimerization) (Taylor et al., 2013). Though Dicer proteins are mostly regarded in small RNA processing, a recent report showed that human DICER I directly processes human Alu RNA, which is associated with human eye diseases (Kaneko et al., 2011). Based on the fact that DCL2 contains RNase activity and that it affects TER2 abundance in a post-transcriptional manner (Chapter IV), it is possible that DCL2 directly contacts TER2 for its processing. To test this possibility, *in vitro* binding between DCL2 and TER2 should be determined, using filter binding assays, for example. DCL3 protein and TER1 RNA could act as negative controls. If we fail to see binding between DCL2 and TER2, it would suggest that DCL2 regulates TER2 indirectly. However, another explanation is that TER2 is degraded upon binding to DCL2. If this is true, a polyacrylamide gel can be used to see the integrity of the product. I expect to see a smear banding pattern (degradation) or smaller products (processing). Sequence analysis of the products can reveal if there is a possibility of

direct interaction. DCL2-TER2 binding affinity can be determined using a nuclease-dead DCL2 mutant that contains substitutions in the highly conserved Rnase III domain (Weinberg et al., 2011; Zhu et al., 2013). I expect to see specific binding of DCL2-TER2. However, if TER2 does not show binding specificity to DCL2 *in vitro,* it may suggest that other proteins are required for the specific binding *in vivo* (Eamens et al., 2009). Altogether, these results can give new insight into the mechanism of DCL2-TER2 interaction and uncover the potential of DCL2 in lncRNA processing.

DCL2-TER2 complex in vivo

To verify the aforementioned *in vitro* results, or to figure out if any *in vivo* factors are required for DCL2-TER2 association/processing, the tagged DCL2 construct and its nuclease dead version can be transformed in *dcl2* mutants to generate transgenic lines. The transgenic lines can be analyzed by qRT-PCR to determine the TER2 RNA abundance. I expect to see TER2 decrease when complemented with wild type DCL2 in a *dcl2* mutant, but not with the nuclease dead DCL2. Another experiment would be to immunoprecipitate DCL2 and determine the TER2 association by RT-PCR in the transgenic lines. A further goal would be to test if there is any change between DCL2-TER2 interactions under DDR. I expect to see a decreased association between DCL2 and TER2 after DCL2 IP in samples with zeocin treatment. If a direct interaction between DCL2 and TER2 can be confirmed, transgenic lines can be subjected to footprinting to pinpoint the motifs of TER2 in DCL2 binding.

If a TER2 binding element for DCL2 can be defined, another interesting experiment would be to engineer TER1 or other RNAs not affected by DCL2 using this determined motif. If the insertion of this motif does not sufficiently decrease the RNA stability, it will

argue that other cis-elements in TER2 are required. Altogether, the results of these studies will broaden our knowledge of DCL2-TER2 interaction mechanism and reinforce the idea that DCL2 has the potential for regulating lncRNA.

Identification of novel lncRNAs associated with DCL2

A final goal should be to test whether DCL2 asssociates with other lncRNAs; and thus test the hypothesis that this is a general paradigm of DCL2's function in lncRNA processing. I suggest using the nuclease dead DCL2 transgenic line to perform cross-linking IP (CLIP) followed by deep sequencing. In chapter II, I showed that TE exaptation in TER2 gave this transcript unique functions. It is possible that TEs are necessary for DCL2 processing of all lncRNAs. To expand the knowledge of DCL2-lncRNA interaction in the context of the TE exaptation, bioinformatic analysis is needed to see if any of the RNAs identified in the RNA-seq experiment contain a TE. Since TEs are proposed to assemble into preformed structures for shaping lncRNA function (Johnson and Guigo, 2014; Figure 1-6), it will be interesting to see if the DCL2 associated RNAs share conserved structures or sequences. Binding assays could be used to test DCL2 binding of these signature sequences. These results can provide direct and insightful evidence of DCL2's function in lncRNA metabolism driven by evolution through a TE.

TER1 template is under dynamic evolution

Analysis of TER2 formed the bulk of my dissertation research. However, I discovered that TER1 is also evolving in unexpected warp. In Chapter V, I showed that TER1 unlike TER2 is well conserved across *A. thaliana* accessions. One interesting

observation was the discovery of polymorphisms in TER1 template domain: one at the 5' end (U2 to A2), and the other at the 3' end (A10 to C10). I found that the 5' polymorphisms did not alter the telomere repeat sequences synthesized in these accessions. Based on this result, I propose that there is a template boundary element (TBE) imposed at the 5' polymorphism site, which prevents variant nucleotide incorporation during the telomerase elongation cycle. Using this TBE requires alignment of the 3' nucleotide in template for the telomerase repeat addition processivity. The TBE is not expected in TER1 genes with the 3' polymorphism. However, the recent preliminary results show a decreased telomerase activity in accessions with the 3' template polymorphism, suggesting that a TBE may indeed be present at the same site in all *A.thaliana* accessions. In organisms with the 3' polymorphism, the TBE may cause reduced telomerase processivity. To test these ideas and verify the data, an *in vitro* reconstitute assay of telomerase is required using TER1 RNA with the 5' and 3' template polymorphisms.

To gain a thorough understanding of the TER1 template polymorphisms, a secondary structures model for TER1 would be helpful. Full-length TER1 genes should be cloned and the RNA expressed *in vitro,* and then applied by selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE). Interestingly, there are other distinct polymorphisms outside the template domain of TER1 molecules that contain template polymorphism. Whether these additional polymorphisms somehow influence template utilization is unknown. Altogether, the results of these experiments could indicate if different strategies are applied in the usage of the TER1 RNA template among *A. thaliana* accessions. Furthermore, these variations may reflect a different requirement of telomerase activity among accessions, which evolved after TER gene duplication.

Conclusion

In this dissertation research, I uncovered new insights into the regulation mechanism of telomerase activity from its RNA subunits in the flowering plant *Arabidopsis thaliana*. Telomerase regulation reflected through both TER1 and TER2 highlights the rapid evolution of these molecules among *A. thaliana* accessions. I characterized the TER1 template polymorphisms and their effects on telomerase activity, revealing new information about the templating function of TER. I defined a transposable element that gives TER2 distinct functions, including RNA stability, telomerase negative regulation and capability to sense DNA damage. In the process of analyzing TER2 RNA function, I found a role of this RNA for reproduction fitness, which is independent of TER2's function in telomerase regulation. A surprising finding from my study was that the processing of TER2 is controlled by small RNA machinery and conventional RNA splicing machinery. Altogether my findings reveal unanticipated diversity in the mechanisms of telomerase regulation through novel roles and processing mechanisms for lncRNA.

REFERENCES

Abdallah, P., Luciano, P., Runge, K.W., Lisby, M., Geli, V., Gilson, E., and Teixeira, M.T. (2009). A two-step model for senescence triggered by a single critically short telomere. Nature Cell Biology *11*, 988-993.

Abreu, E., Aritonovska, E., Reichenbach, P., Cristofari, G., Culp, B., Terns, R.M., Lingner, J., and Terns, M.P. (2010). TIN2-tethered TPP1 recruits human telomerase to telomeres *in vivo*. Molecular and Cellular Biology *30*, 2971-2982.

Agrena, J., Oakley, C.G., McKay, J.K., Lovell, J.T., and Schemske, D.W. (2013). Genetic mapping of adaptation reveals fitness tradeoffs in *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences of the United States of America *110*, 21077-21082.

Akey, J.M., Eberle, M.A., Rieder, M.J., Carlson, C.S., Shriver, M.D., Nickerson, D.A., and Kruglyak, L. (2004). Population history and natural selection shape patterns of genetic variation in 132 genes. PLoS Biology *2*, e286.

Amiard, S., Doudeau, M., Pinte, S., Poulet, A., Lenain, C., Faivre-Moskalenko, C., Angelov, D., Hug, N., Vindigni, A., Bouvet, P.*, et al.* (2007). A topological mechanism for TRF2-enhanced strand invasion. Nature Structural & Molecular Biology *14*, 147-154.

Anderson, B.H., Kasher, P.R., Mayer, J., Szynkiewicz, M., Jenkinson, E.M., Bhaskar, S.S., Urquhart, J.E., Daly, S.B., Dickerson, J.E., O'Sullivan, J.*, et al.* (2012). Mutations in CTC1, encoding conserved telomere maintenance component 1, cause Coats plus. Nature Genetics *44*, 338-342.

Arat, N.O., and Griffith, J.D. (2012). Human Rap1 interacts directly with telomeric DNA and regulates TRF2 localization at the telomere. The Journal of Biological Chemistry *287*, 41583-41594.

Armanios, M., and Blackburn, E.H. (2012). The telomere syndromes. Nature Reviews Genetics *13*, 693-704.

Armanios, M., and Greider, C.W. (2005). Telomerase and cancer stem cells. Cold Spring Harbor Symposia on Quantitative Biology *70*, 205-208.

Arnoult, N., and Karlseder, J. (2014). ALT telomeres borrow from meiosis to get moving. Cell *159*, 11-12.

Arrisi-Mercado, P., Romano, M., Muro, A.F., and Baralle, F.E. (2004). An exonic splicing enhancer offsets the atypical GU-rich 3' splice site of human apolipoprotein A-II exon 3. The Journal of Biological Chemistry *279*, 39331-39339.

Autexier, C., and Greider, C.W. (1995). Boundary elements of the *Tetrahymena* telomerase RNA template and alignment domains. Genes & Development *9*, 2227-2239.

Autexier, C., and Lue, N.F. (2006). The structure and function of telomerase reverse transcriptase. Annual Review of Biochemistry *75*, 493-517.

Bae, N.S., and Baumann, P. (2007). A RAP1/TRF2 complex inhibits nonhomologous end-joining at human telomeric DNA ends. Molecular Cell *26*, 323-334.

Barnett, M.A., Buckle, V.J., Evans, E.P., Porter, A.C., Rout, D., Smith, A.G., and Brown, W.R. (1993). Telomere directed fragmentation of mammalian chromosomes. Nucleic Acids Research *21*, 27-36.

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

Beilstein, M.A., Brinegar, A.E., and Shippen, D.E. (2012). Evolution of the *Arabidopsis* telomerase RNA. Frontiers in Genetics *3*, 188.

Beilstein, M.A., Nagalingum, N.S., Clements, M.D., Manchester, S.R., and Mathews, S. (2010). Dated molecular phylogenies indicate a Miocene origin for *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences of the United States of America *107*, 18724-18728.

Bellodi, C., McMahon, M., Contreras, A., Juliano, D., Kopmar, N., Nakamura, T., Maltby, D., Burlingame, A., Savage, S.A., Shimamura, A.*, et al.* (2013). H/ACA small RNA dysfunctions in disease reveal key roles for noncoding RNA modifications in hematopoietic stem cell differentiation. Cell Reports *3*, 1493-1502.

Berman, A.J., Akiyama, B.M., Stone, M.D., and Cech, T.R. (2011). The RNA accordion model for template positioning by telomerase RNA during telomeric DNA synthesis. Nature Structural & Molecular Biology *18*, 1371-1375.

Berman, A.J., Gooding, A.R., and Cech, T.R. (2010). *Tetrahymena* telomerase protein p65 induces conformational changes throughout telomerase RNA (TER) and rescues telomerase reverse transcriptase and TER assembly mutants. Molecular and Cellular Biology *30*, 4965-4976.

Bernardes de Jesus, B., and Blasco, M.A. (2013). Telomerase at the intersection of cancer and aging. Trends in Genetics *29*, 513-520.

Bertuch, A.A., and Lundblad, V. (2003). The Ku heterodimer performs separable activities at double-strand breaks and chromosome termini. Molecular and Cellular Biology *23*, 8202-8215.

Bianchi, A., Negrini, S., and Shore, D. (2004). Delivery of yeast telomerase to a DNA break depends on the recruitment functions of Cdc13 and Est1. Molecular Cell *16*, 139-146.

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

Bianchi, A., Smith, S., Chong, L., Elias, P., and de Lange, T. (1997). TRF1 is a dimer and bends telomeric DNA. The EMBO Journal *16*, 1785-1794.

Biessmann, H., and Mason, J.M. (1997). Telomere maintenance without telomerase. Chromosoma *106*, 63-69.

Bilaud, T., Koering, C.E., Binet-Brasselet, E., Ancelin, K., Pollice, A., Gasser, S.M., and Gilson, E. (1996). The telobox, a Myb-related telomeric DNA binding motif found in proteins from yeast, plants and human. Nucleic Acids Research *24*, 1294-1303.

Blackburn, E.H., and Gall, J.G. (1978). A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in *Tetrahymena*. Journal of Molecular Biology *120*, 33-53.

Blasco, M.A., Funk, W., Villeponteau, B., and Greider, C.W. (1995). Functional characterization and developmental regulation of mouse telomerase RNA. Science *269*, 1267-1270.

Boltz, K.A., Leehy, K., Song, X., Nelson, A.D., and Shippen, D.E. (2012). ATR cooperates with CTC1 and STN1 to maintain telomeres and genome integrity in *Arabidopsis*. Molecular Biology of the Cell *23*, 1558-1568.

Borsani, O., Zhu, J., Verslues, P.E., Sunkar, R., and Zhu, J.K. (2005). Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in *Arabidopsis*. Cell *123*, 1279-1291.

Boulton, S.J., and Jackson, S.P. (1996). *Saccharomyces cerevisiae* Ku70 potentiates illegitimate DNA double-strand break repair and serves as a barrier to error-prone DNA repair pathways. The EMBO Journal *15*, 5093-5103.

Bourns, B.D., Alexander, M.K., Smith, A.M., and Zakian, V.A. (1998). Sir proteins, Rif proteins, and Cdc13p bind Saccharomyces telomeres *in vivo*. Molecular and Cellular Biology *18*, 5600-5608.

Box, J.A., Bunch, J.T., Tang, W., and Baumann, P. (2008a). Spliceosomal cleavage generates the 3' end of telomerase RNA. Nature *456*, 910-914.

Box, J.A., Bunch, J.T., Zappulla, D.C., Glynn, E.F., and Baumann, P. (2008b). A flexible template boundary element in the RNA subunit of fission yeast telomerase. The Journal of Biological Chemistry *283*, 24224-24233.

Broccoli, D., Chong, L., Oelmann, S., Fernald, A.A., Marziliano, N., van Steensel, B., Kipling, D., Le Beau, M.M., and de Lange, T. (1997a). Comparison of the human and mouse genes encoding the telomeric protein, TRF1: chromosomal localization, expression and conserved protein domains. Human Molecular Genetics *6*, 69-76.

Broccoli, D., Smogorzewska, A., Chong, L., and de Lange, T. (1997b). Human telomeres contain two distinct Myb-related proteins, TRF1 and TRF2. Nature Genetics *17*, 231-235.

Brown, Y., Abraham, M., Pearl, S., Kabaha, M.M., Elboher, E., and Tzfati, Y. (2007). A critical three-way junction is conserved in budding yeast and vertebrate telomerase RNAs. Nucleic Acids Research *35*, 6280-6289.

Bunch, J.T., Bae, N.S., Leonardi, J., and Baumann, P. (2005). Distinct requirements for Pot1 in limiting telomere length and maintaining chromosome stability. Molecular and Cellular Biology *25*, 5567-5578.

Burset, M., Seledtsov, I.A., and Solovyev, V.V. (2000). Analysis of canonical and non-canonical splice sites in mammalian genomes. Nucleic Acids Research *28*, 4364-4375.

Butelli, E., Licciardello, C., Zhang, Y., Liu, J., Mackay, S., Bailey, P., Reforgiato-Recupero, G., and Martin, C. (2012). Retrotransposons control fruit-specific, cold-dependent accumulation of anthocyanins in blood oranges. The Plant Cell *24*, 1242-1255.

Cairney, C.J., and Keith, W.N. (2008). Telomerase redefined: integrated regulation of hTR and hTERT for telomere maintenance and telomerase activity. Biochimie *90*, 13-23.

Carofiglio, F., Inagaki, A., de Vries, S., Wassenaar, E., Schoenmakers, S., Vermeulen, C., van Cappellen, W.A., Sleddens-Linkels, E., Grootegoed, J.A., Te Riele, H.P.*, et al.* (2013). SPO11-independent DNA repair foci and their role in meiotic silencing. PLoS Genetics *9*, e1003538.

Carrel, A., and Ebeling, A.H. (1921). The multiplication of fibroblasts *in vitro*. The Journal of Experimental Medicine *34*, 317-337.

Casacuberta, E., and Gonzalez, J. (2013). The impact of transposable elements in environmental adaptation. Molecular Ecology *22*, 1503-1517.

Casteel, D.E., Zhuang, S., Zeng, Y., Perrino, F.W., Boss, G.R., Goulian, M., and Pilz, R.B. (2009). A DNA polymerase-{alpha}{middle dot}primase cofactor with homology to replication protein A-32 regulates DNA replication in mammalian cells. The Journal of Biological Chemistry *284*, 5807-5818.

Castel, S.E., and Martienssen, R.A. (2013). RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. Nature Reviews Genetics *14*, 100-112.

Catalanotto, C., Pallotta, M., ReFalo, P., Sachs, M.S., Vayssie, L., Macino, G., and Cogoni, C. (2004). Redundancy of the two dicer genes in transgene-induced posttranscriptional gene silencing in *Neurospora crassa*. Molecular and Cellular Biology *24*, 2536-2545.

Cavrak, V.V., Lettner, N., Jamge, S., Kosarewicz, A., Bayer, L.M., and Mittelsten Scheid, O. (2014). How a retrotransposon exploits the plant's heat stress response for its activation. PLoS Genetics *10*, e1004115.

Chai, W., Ford, L.P., Lenertz, L., Wright, W.E., and Shay, J.W. (2002). Human Ku70/80 associates physically with telomerase through interaction with hTERT. The Journal of Biological Chemistry *277*, 47242-47247.

Chai, W., Sfeir, A.J., Hoshiyama, H., Shay, J.W., and Wright, W.E. (2006). The involvement of the Mre11/Rad50/Nbs1 complex in the generation of G-overhangs at human telomeres. EMBO Reports *7*, 225-230.

Chang, M., Arneric, M., and Lingner, J. (2007). Telomerase repeat addition processivity is increased at critically short telomeres in a Tel1-dependent manner in *Saccharomyces cerevisiae*. Genes & Development *21*, 2485-2494.

Chapon, C., Cech, T.R., and Zaug, A.J. (1997). Polyadenylation of telomerase RNA in budding yeast. RNA *3*, 1337-1351.

Chappell, A.S., and Lundblad, V. (2004). Structural elements required for association of the *Saccharomyces cerevisiae* telomerase RNA with the Est2 reverse transcriptase. Molecular and Cellular Biology *24*, 7720-7736.

Chen, J.L., Blasco, M.A., and Greider, C.W. (2000). Secondary structure of vertebrate telomerase RNA. Cell *100*, 503-514.

Chen, J.L., and Greider, C.W. (2003). Determinants in mammalian telomerase RNA that mediate enzyme processivity and cross-species incompatibility. The EMBO Journal *22*, 304-314.

Chen, J.L., and Greider, C.W. (2005). Functional analysis of the pseudoknot structure in human telomerase RNA. Proceedings of the National Academy of Sciences of the United States of America *102*, 8080-8085; discussion 8077-8089.

Chen, J.M., Stenson, P.D., Cooper, D.N., and Ferec, C. (2005). A systematic analysis of LINE-1 endonuclease-dependent retrotranspositional events causing human genetic disease. Human Genetics *117*, 411-427.

Chen, L.Y., Redon, S., and Lingner, J. (2012a). The human CST complex is a terminator of telomerase activity. Nature *488*, 540-544.

Chen, L.Y., Zhang, Y., Zhang, Q., Li, H., Luo, Z., Fang, H., Kim, S.H., Qin, L., Yotnda, P., Xu, J.*, et al.* (2012b). Mitochondrial localization of telomeric protein TIN2 links telomere regulation to metabolic control. Molecular Cell *47*, 839-850.

Cheng, F., Liu, S., Wu, J., Fang, L., Sun, S., Liu, B., Li, P., Hua, W., and Wang, X. (2011). BRAD, the genetics and genomics database for *Brassica* plants. BMC Plant Biology *11*, 136.

Chow, T.T., Zhao, Y., Mak, S.S., Shay, J.W., and Wright, W.E. (2012). Early and late steps in telomere overhang processing in normal human cells: the position of the final RNA primer drives telomere shortening. Genes & Development *26*, 1167-1178.

Chung, W.J., Okamura, K., Martin, R., and Lai, E.C. (2008). Endogenous RNA interference provides a somatic defense against *Drosophila* transposons. Current Biology *18*, 795-802.

Ciccia, A., and Elledge, S.J. (2010). The DNA damage response: making it safe to play with knives. Molecular Cell *40*, 179-204.

Cifuentes-Rojas, C., Kannan, K., Tseng, L., and Shippen, D.E. (2011). Two RNA subunits and POT1a are components of *Arabidopsis telomerase*. Proceedings of the National Academy of Sciences of the United States of America *108*, 73-78.

Cifuentes-Rojas, C., Nelson, A.D., Boltz, K.A., Kannan, K., She, X., and Shippen, D.E. (2012). An alternative telomerase RNA in *Arabidopsis* modulates enzyme activity in response to DNA damage. Genes & Development *26*, 2512-2523.

Cifuentes-Rojas, C., and Shippen, D.E. (2012). Telomerase regulation. Mutation Research *730*, 20-27.

Cimprich, K.A., and Cortez, D. (2008). ATR: an essential regulator of genome integrity. Nature Reviews Molecular Cell Biology *9*, 616-627.

Clark, M.B., Johnston, R.L., Inostroza-Ponta, M., Fox, A.H., Fortini, E., Moscato, P., Dinger, M.E., and Mattick, J.S. (2012). Genome-wide analysis of long noncoding RNA stability. Genome Research *22*, 885-898.

Cohn, M., and Blackburn, E.H. (1995). Telomerase in yeast. Science *269*, 396-400.

Cole, F., Keeney, S., and Jasin, M. (2010). Evolutionary conservation of meiotic DSB proteins: more than just Spo11. Genes & Development *24*, 1201-1207.

Collins, K. (2006). The biogenesis and regulation of telomerase holoenzymes. Nature Reviews Molecular Cell Biology *7*, 484-494.

Collins, K., and Greider, C.W. (1995). Utilization of ribonucleotides and RNA primers by *Tetrahymena* telomerase. The EMBO Journal *14*, 5422-5432.

Cooper, J.P., Nimmo, E.R., Allshire, R.C., and Cech, T.R. (1997). Regulation of telomere length and function by a Myb-domain protein in fission yeast. Nature *385*, 744-747.

Cowley, M., and Oakey, R.J. (2013). Transposable elements re-wire and fine-tune the transcriptome. PLoS Genetics *9*, e1003234.

Craven, R.J., and Petes, T.D. (2000). Involvement of the checkpoint protein Mec1p in silencing of gene expression at telomeres in *Saccharomyces cerevisiae*. Molecular and Cellular Biology *20*, 2378-2384.

Da Ines, O., Degroote, F., Goubely, C., Amiard, S., Gallego, M.E., and White, C.I. (2013). Meiotic recombination in *Arabidopsis* is catalysed by DMC1, with RAD51 playing a supporting role. PLoS Genetics *9*, e1003787.

Dandjinou, A.T., Levesque, N., Larose, S., Lucier, J.F., Abou Elela, S., and Wellinger, R.J. (2004). A phylogenetically based secondary structure for the yeast telomerase RNA. Current Biology *14*, 1148-1158.

de Koning, A.P., Gu, W., Castoe, T.A., Batzer, M.A., and Pollock, D.D. (2011). Repetitive elements may comprise over two-thirds of the human genome. PLoS Genetics *7*, e1002384.

de Lange, T. (2005). Shelterin: the protein complex that shapes and safeguards human telomeres. Genes & Development *19*, 2100-2110.

de Lange, T. (2010). How shelterin solves the telomere end-protection problem. Cold Spring Harbor Symposia on Quantitative Biology *75*, 167-177.

de Souza, F.S., Franchini, L.F., and Rubinstein, M. (2013). Exaptation of transposable elements into novel *cis*-regulatory elements: is the evidence always strong? Molecular Biology and Evolution *30*, 1239-1251.

Dehe, P.M., Rog, O., Ferreira, M.G., Greenwood, J., and Cooper, J.P. (2012). Taz1 enforces cell-cycle regulation of telomere synthesis. Molecular Cell *46*, 797-808.

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

Diede, S.J., and Gottschling, D.E. (1999). Telomerase-mediated telomere addition in vivo requires DNA primase and DNA polymerases alpha and delta. Cell *99*, 723-733.

Dimitrova, N., and de Lange, T. (2009). Cell cycle-dependent role of MRN at dysfunctional telomeres: ATM signaling-dependent induction of nonhomologous end joining (NHEJ) in G1 and resection-mediated inhibition of NHEJ in G2. Molecular and Cellular Biology *29*, 5552-5563.

Ding, S.W. (2010). RNA-based antiviral immunity. Nature Reviews Immunology *10*, 632-644.

Dionne, I., Larose, S., Dandjinou, A.T., Abou Elela, S., and Wellinger, R.J. (2013). Cell cycle-dependent transcription factors control the expression of yeast telomerase RNA. RNA *19*, 992-1002.

Doksani, Y., Wu, J.Y., de Lange, T., and Zhuang, X. (2013). Super-resolution fluorescence imaging of telomeres reveals TRF2-dependent T-loop formation. Cell *155*, 345-356.

DuBois, M.L., Haimberger, Z.W., McIntosh, M.W., and Gottschling, D.E. (2002). A quantitative assay for telomere protection in *Saccharomyces cerevisiae*. Genetics *161*, 995-1013.

Dutertre, M., Lambert, S., Carreira, A., Amor-Gueret, M., and Vagner, S. (2014). DNA damage: RNA-binding proteins protect from near and far. Trends in Biochemical Sciences *39*, 141-149.

Eamens, A.L., Smith, N.A., Curtin, S.J., Wang, M.B., and Waterhouse, P.M. (2009). The *Arabidopsis thaliana* double-stranded RNA binding protein DRB1 directs guide strand selection from microRNA duplexes. RNA *15*, 2219-2235.

Egan, E.D., and Collins, K. (2012). Biogenesis of telomerase ribonucleoproteins. RNA *18*, 1747-1759.

Evans, S.K., and Lundblad, V. (1999). Est1 and Cdc13 as comediators of telomerase access. Science *286*, 117-120.

Fajkus, J., Sykorova, E., and Leitch, A.R. (2005). Telomeres in evolution and evolution of telomeres. Chromosome research : an International Journal on the Molecular, Supramolecular and Evolutionary Aspects of Chromosome Biology *13*, 469-479.

Fan, X., and Price, C.M. (1997). Coordinate regulation of G- and C strand length during new telomere synthesis. Molecular Biology of the Cell *8*, 2145-2155.

Farah, J.A., Cromie, G., Steiner, W.W., and Smith, G.R. (2005). A novel recombination pathway initiated by the Mre11/Rad50/Nbs1 complex eliminates palindromes during meiosis in *Schizosaccharomyces pombe*. Genetics *169*, 1261-1274.

Faulkner, G.J., and Carninci, P. (2009). Altruistic functions for selfish DNA. Cell Cycle *8*, 2895-2900.

Fedoroff, N. (2001). How jumping genes were discovered. Nature Structural Biology *8*, 300-301.

Feng, L., and Chen, J. (2012). The E3 ligase RNF8 regulates KU80 removal and NHEJ repair. Nature Structural & Molecular Biology *19*, 201-206.

Ferreira, M.G., and Cooper, J.P. (2001). The fission yeast Taz1 protein protects chromosomes from Ku-dependent end-to-end fusions. Molecular Cell *7*, 55-63.

Fisher, T.S., Taggart, A.K., and Zakian, V.A. (2004). Cell cycle-dependent regulation of yeast telomerase by Ku. Nature Structural & Molecular Biology *11*, 1198-1205.

Fitzgerald, M.S., McKnight, T.D., and Shippen, D.E. (1996). Characterization and developmental patterns of telomerase expression in plants. Proceedings of the National Academy of Sciences of the United States of America *93*, 14422-14427.

Fitzgerald, M.S., Riha, K., Gao, F., Ren, S., McKnight, T.D., and Shippen, D.E. (1999). Disruption of the telomerase catalytic subunit gene from *Arabidopsis* inactivates telomerase and leads to a slow loss of telomeric DNA. Proceedings of the National Academy of Sciences of the United States of America *96*, 14813-14818.

Fitzgerald, M.S., Shakirov, E.V., Hood, E.E., McKnight, T.D., and Shippen, D.E. (2001). Different modes of *de novo* telomere formation by plant telomerases. The Plant Journal for Cell and Molecular Biology *26*, 77-87.

Flint, J., Craddock, C.F., Villegas, A., Bentley, D.P., Williams, H.J., Galanello, R., Cao, A., Wood, W.G., Ayyub, H., and Higgs, D.R. (1994). Healing of broken human chromosomes by the addition of telomeric repeats. American Journal of Human Genetics *55*, 505-512.

Forsyth, N.R., Wright, W.E., and Shay, J.W. (2002). Telomerase and differentiation in multicellular organisms: turn it off, turn it on, and turn it off again. Differentiation; Research in Biological Diversity *69*, 188-197.

Frank, C.J., Hyde, M., and Greider, C.W. (2006). Regulation of telomere elongation by the cyclin-dependent kinase CDK1. Molecular Cell *24*, 423-432.

Frescas, D., and de Lange, T. (2014). A TIN2 dyskeratosis congenita mutation causes telomerase-independent telomere shortening in mice. Genes & Development *28*, 153-166.

Fu, D., and Collins, K. (2003). Distinct biogenesis pathways for human telomerase RNA and H/ACA small nucleolar RNAs. Molecular Cell *11*, 1361-1372.

Fu, D., and Collins, K. (2007). Purification of human telomerase complexes identifies factors involved in telomerase biogenesis and telomere length regulation. Molecular Cell *28*, 773-785.

Gallardo, F., Laterreur, N., Cusanelli, E., Ouenzar, F., Querido, E., Wellinger, R.J., and Chartrand, P. (2011). Live cell imaging of telomerase RNA dynamics reveals cell cycle-dependent clustering of telomerase at elongating telomeres. Molecular Cell *44*, 819-827.

Garvik, B., Carson, M., and Hartwell, L. (1995). Single-stranded DNA arising at telomeres in cdc13 mutants may constitute a specific signal for the RAD9 checkpoint. Molecular and Cellular Biology *15*, 6128-6138.

Gilley, D., Lee, M.S., and Blackburn, E.H. (1995). Altering specific telomerase RNA template residues affects active site function. Genes & Development *9*, 2214-2226.

Giraud-Panis, M.J., Teixeira, M.T., Geli, V., and Gilson, E. (2010). CST meets shelterin to keep telomeres in check. Molecular Cell *39*, 665-676.

Golden, T.A., Schauer, S.E., Lang, J.D., Pien, S., Mushegian, A.R., Grossniklaus, U., Meinke, D.W., and Ray, A. (2002). SHORT INTEGUMENTS1/SUSPENSOR1/CARPEL FACTORY, a Dicer homolog, is a maternal effect gene required for embryo development in *Arabidopsis*. Plant Physiology *130*, 808-822.

Golisz, A., Sikorski, P.J., Kruszka, K., and Kufel, J. (2013). *Arabidopsis thaliana* LSM proteins function in mRNA splicing and degradation. Nucleic Acids Research *41*, 6232-6249.

Gottschling, D.E., and Zakian, V.A. (1986). Telomere proteins: specific recognition and protection of the natural termini of *Oxytricha* macronuclear DNA. Cell *47*, 195-205.

Grandin, N., Damon, C., and Charbonneau, M. (2001). Ten1 functions in telomere end protection and length regulation in association with Stn1 and Cdc13. The EMBO Journal *20*, 1173-1183.

Grandin, N., Reed, S.I., and Charbonneau, M. (1997). Stn1, a new *Saccharomyces cerevisiae* protein, is implicated in telomere size regulation in association with Cdc13. Genes & Development *11*, 512-527.

Greenberg, R.A., Allsopp, R.C., Chin, L., Morin, G.B., and DePinho, R.A. (1998). Expression of mouse telomerase reverse transcriptase during development, differentiation and proliferation. Oncogene *16*, 1723-1730.

Greider, C.W. (1991). Telomerase is processive. Molecular and Cellular Biology *11*, 4572-4580.

Greider, C.W., and Blackburn, E.H. (1985). Identification of a specific telomere terminal transferase activity in *Tetrahymena* extracts. Cell *43*, 405-413.

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

Greider, C.W., and Blackburn, E.H. (1989). A telomeric sequence in the RNA of *Tetrahymena* telomerase required for telomere repeat synthesis. Nature *337*, 331-337.

Griffith, J.D., Comeau, L., Rosenfield, S., Stansel, R.M., Bianchi, A., Moss, H., and de Lange, T. (1999). Mammalian telomeres end in a large duplex loop. Cell *97*, 503-514.

Grunstein, M. (1997). Molecular model for telomeric heterochromatin in yeast. Current Opinion in Cell Biology *9*, 383-387.

Gunes, C., and Rudolph, K.L. (2013). The role of telomeres in stem cells and cancer. Cell *152*, 390-393.

Guttman, M., Donaghey, J., Carey, B.W., Garber, M., Grenier, J.K., Munson, G., Young, G., Lucas, A.B., Ach, R., Bruhn, L.*, et al.* (2011). lincRNAs act in the circuitry controlling pluripotency and differentiation. Nature *477*, 295-300.

Hammond, P.W., Lively, T.N., and Cech, T.R. (1997). The anchor site of telomerase from *Euplotes aediculatus* revealed by photo-cross-linking to single- and double-stranded DNA primers. Molecular and Cellular Biology *17*, 296-308.

Hang, L.E., Lopez, C.R., Liu, X., Williams, J.M., Chung, I., Wei, L., Bertuch, A.A., and Zhao, X. (2014). Regulation of Ku-DNA association by Yku70 C-terminal tail and SUMO modification. The Journal of Biological Chemistry *289*, 10308-10317.

Hanish, J.P., Yanowitz, J.L., and de Lange, T. (1994). Stringent sequence requirements for the formation of human telomeres. Proceedings of the National Academy of Sciences of the United States of America *91*, 8861-8865.

Hardy, C.F., Sussel, L., and Shore, D. (1992). A RAP1-interacting protein involved in transcriptional silencing and telomere length regulation. Genes & Development *6*, 801-814.

Harland, J.L., Chang, Y.T., Moser, B.A., and Nakamura, T.M. (2014). Tpz1-Ccq1 and Tpz1-Poz1 interactions within fission yeast shelterin modulate Ccq1 Thr93 phosphorylation and telomerase recruitment. PLoS Genetics *10*, e1004708.

Harper, J.W., and Elledge, S.J. (2007). The DNA damage response: ten years after. Molecular Cell *28*, 739-745.

Harrington, L. (2003). Biochemical aspects of telomerase function. Cancer Letters *194*, 139-154.

Harrington, L.A., and Greider, C.W. (1991). Telomerase primer specificity and chromosome healing. Nature *353*, 451-454.

Hayflick, L. (1961). The establishment of a line (WISH) of human amnion cells in continuous cultivation. Experimental Cell Research *23*, 14-20.

Hayflick, L. (1965). The limited *in vitro* lifetime of human diploid cell strains. Experimental Cell Research *37*, 614-636.

Heacock, M., Spangler, E., Riha, K., Puizina, J., and Shippen, D.E. (2004). Molecular analysis of telomere fusions in *Arabidopsis*: multiple pathways for chromosome end-joining. The EMBO Journal *23*, 2304-2313.

Henderson, E.R., and Blackburn, E.H. (1989). An overhanging 3' terminus is a conserved feature of telomeres. Molecular and Cellular Biology *9*, 345-348.

Henderson, I.R., Zhang, X., Lu, C., Johnson, L., Meyers, B.C., Green, P.J., and Jacobsen, S.E. (2006). Dissecting *Arabidopsis thaliana* DICER function in small RNA processing, gene silencing and DNA methylation patterning. Nature Genetics *38*, 721-725.

Hockemeyer, D., Daniels, J.P., Takai, H., and 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., and Hammer, G.D. (2007). Telomere protection by mammalian Pot1 requires interaction with Tpp1. Nature Structural & Molecular Biology *14*, 754-761.

Hoeijmakers, J.H. (2009). DNA damage, aging, and cancer. The New England Journal of Medicine *361*, 1475-1485.

Hollister, J.D., and Gaut, B.S. (2009). Epigenetic silencing of transposable elements: a trade-off between reduced transposition and deleterious effects on neighboring gene expression. Genome Research *19*, 1419-1428.

Hollister, J.D., Smith, L.M., Guo, Y.L., Ott, F., Weigel, D., and Gaut, B.S. (2011). Transposable elements and small RNAs contribute to gene expression divergence between *Arabidopsis thaliana* and *Arabidopsis lyrata*. Proceedings of the National Academy of Sciences of the United States of America *108*, 2322-2327.

Hossain, S., Singh, S., and Lue, N.F. (2002). Functional analysis of the C-terminal extension of telomerase reverse transcriptase. A putative "thumb" domain. The Journal of Biological Chemistry *277*, 36174-36180.

Hu, T.T., Pattyn, P., Bakker, E.G., Cao, J., Cheng, J.F., Clark, R.M., Fahlgren, N., Fawcett, J.A., Grimwood, J., Gundlach, H.*, et al.* (2011). The *Arabidopsis lyrata* genome sequence and the basis of rapid genome size change. Nature Genetics *43*, 476-481.

Huang, C.R., Schneider, A.M., Lu, Y., Niranjan, T., Shen, P., Robinson, M.A., Steranka, J.P., Valle, D., Civin, C.I., Wang, T.*, et al.* (2010). Mobile interspersed repeats are major structural variants in the human genome. Cell *141*, 1171-1182.

Hwang, M.G., and Cho, M.H. (2007). Arabidopsis thaliana telomeric DNA-binding protein 1 is required for telomere length homeostasis and its Myb-extension domain stabilizes plant telomeric DNA binding. Nucleic Acids Research *35*, 1333-1342.

Ito, H., Gaubert, H., Bucher, E., Mirouze, M., Vaillant, I., and Paszkowski, J. (2011). An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. Nature *472*, 115-119.

Jeon, Y., and Lee, J.T. (2011). YY1 tethers Xist RNA to the inactive X nucleation center. Cell *146*, 119-133.

Johnson, R., and Guigo, R. (2014). The RIDL hypothesis: transposable elements as functional domains of long noncoding RNAs. RNA *20*, 959-976.

Jun, H.I., Liu, J., Jeong, H., Kim, J.K., and Qiao, F. (2013). Tpz1 controls a telomerase-nonextendible telomeric state and coordinates switching to an extendible state via Ccq1. Genes & Development *27*, 1917-1931.

Kaneko, H., Dridi, S., Tarallo, V., Gelfand, B.D., Fowler, B.J., Cho, W.G., Kleinman, M.E., Ponicsan, S.L., Hauswirth, W.W., Chiodo, V.A.*, et al.* (2011). DICER1 deficit induces Alu RNA toxicity in age-related macular degeneration. Nature *471*, 325-330.

Kannan, K., Nelson, A.D., and Shippen, D.E. (2008). Dyskerin is a component of the *Arabidopsis* telomerase RNP required for telomere maintenance. Molecular and Cellular Biology *28*, 2332-2341.

Kannan, R., Helston, R.M., Dannebaum, R.O., and Baumann, P. (2015). Diverse mechanisms for spliceosome-mediated 3' end processing of telomerase RNA. Nature Communications *6*, 6104.

Kanno, T., Mette, M.F., Kreil, D.P., Aufsatz, W., Matzke, M., and Matzke, A.J. (2004). Involvement of putative SNF2 chromatin remodeling protein DRD1 in RNA-directed DNA methylation. Current Biology *14*, 801-805.

Kapusta, A., Kronenberg, Z., Lynch, V.J., Zhuo, X., Ramsay, L., Bourque, G., Yandell, M., and Feschotte, C. (2013). Transposable elements are major contributors to the origin, diversification, and regulation of vertebrate long noncoding RNAs. PLoS Genetics *9*, e1003470.

Karamysheva, Z., Wang, L., Shrode, T., Bednenko, J., Hurley, L.A., and Shippen, D.E. (2003). Developmentally programmed gene elimination in Euplotes crassus facilitates a switch in the telomerase catalytic subunit. Cell *113*, 565-576.

Karamysheva, Z.N., Surovtseva, Y.V., Vespa, L., Shakirov, E.V., and Shippen, D.E. (2004). A C-terminal Myb extension domain defines a novel family of double-strand telomeric DNA-binding proteins in *Arabidopsis.* The Journal of Biological Chemistry *279*, 47799-47807.

Kazazian, H.H., Jr., Wong, C., Youssoufian, H., Scott, A.F., Phillips, D.G., and Antonarakis, S.E. (1988). Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. Nature *332*, 164-166.

Kazda, A., Zellinger, B., Rossler, M., Derboven, E., Kusenda, B., and Riha, K. (2012). Chromosome end protection by blunt-ended telomeres. Genes & Development *26*, 1703-1713.

Keller, R.B., Gagne, K.E., Usmani, G.N., Asdourian, G.K., Williams, D.A., Hofmann, I., and Agarwal, S. (2012). CTC1 Mutations in a patient with dyskeratosis congenita. Pediatric Blood & Cancer *59*, 311-314.

Kelley, D., and Rinn, J. (2012). Transposable elements reveal a stem cell-specific class of long noncoding RNAs. Genome Biology *13*, R107.

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

Kinner, A., Wu, W., Staudt, C., and Iliakis, G. (2008). Gamma-H2AX in recognition and signaling of DNA double-strand breaks in the context of chromatin. Nucleic Acids Research *36*, 5678-5694.

Kipling, D., and Cooke, H.J. (1990). Hypervariable ultra-long telomeres in mice. Nature *347*, 400-402.

Kleckner, N. (1996). Meiosis: how could it work? Proceedings of the National Academy of Sciences of the United States of America *93*, 8167-8174.

Klobutcher, L.A., Swanton, M.T., Donini, P., and Prescott, D.M. (1981). All gene-sized DNA molecules in four species of hypotrichs have the same terminal sequence and an unusual 3' terminus. Proceedings of the National Academy of Sciences of the United States of America *78*, 3015-3019.

Knauer, S., Holt, A.L., Rubio-Somoza, I., Tucker, E.J., Hinze, A., Pisch, M., Javelle, M., Timmermans, M.C., Tucker, M.R., and Laux, T. (2013). A protodermal miR394 signal defines a region of stem cell competence in the *Arabidopsis* shoot meristem. Developmental Cell *24*, 125-132.

Kramer, K.M., and Haber, J.E. (1993). New telomeres in yeast are initiated with a highly selected subset of TG1-3 repeats. Genes & Development *7*, 2345-2356.

Lai, C.K., Mitchell, J.R., and Collins, K. (2001). RNA binding domain of telomerase reverse transcriptase. Molecular and Cellular Biology *21*, 990-1000.

Lamb, J., Harris, P.C., Wilkie, A.O., Wood, W.G., Dauwerse, J.G., and Higgs, D.R. (1993). *De novo* truncation of chromosome 16p and healing with (TTAGGG)n in the

alpha-thalassemia/mental retardation syndrome (ATR-16). American Journal of Human Genetics *52*, 668-676.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W.*, et al.* (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.

Larkin, P.D., and Park, W.D. (1999). Transcript accumulation and utilization of alternate and non-consensus splice sites in rice granule-bound starch synthase are temperature-sensitive and controlled by a single-nucleotide polymorphism. Plant Molecular Biology *40*, 719-727.

Lee, J.T., Strauss, W.M., Dausman, J.A., and Jaenisch, R. (1996). A 450 kb transgene displays properties of the mammalian X-inactivation center. Cell *86*, 83-94.

Leehy, K.A., Lee, J.R., Song, X., Renfrew, K.B., and Shippen, D.E. (2013). MERISTEM DISORGANIZATION1 encodes TEN1, an essential telomere protein that modulates telomerase processivity in *Arabidopsis*. The Plant Cell *25*, 1343-1354.

Lendvay, T.S., Morris, D.K., Sah, J., Balasubramanian, B., and Lundblad, V. (1996). Senescence mutants of *Saccharomyces cerevisiae* with a defect in telomere replication identify three additional EST genes. Genetics *144*, 1399-1412.

Leonardi, J., Box, J.A., Bunch, J.T., and Baumann, P. (2008). TER1, the RNA subunit of fission yeast telomerase. Nature Structural & Molecular Biology *15*, 26-33.

Levy, D.L., and Blackburn, E.H. (2004). Counting of Rif1p and Rif2p on *Saccharomyces cerevisiae* telomeres regulates telomere length. Molecular and Cellular Biology *24*, 10857-10867.

Li, B., Oestreich, S., and de Lange, T. (2000). Identification of human Rap1: implications for telomere evolution. Cell *101*, 471-483.

Li, S., Makovets, S., Matsuguchi, T., Blethrow, J.D., Shokat, K.M., and Blackburn, E.H. (2009). Cdk1-dependent phosphorylation of Cdc13 coordinates telomere elongation during cell-cycle progression. Cell *136*, 50-61.

Li, Z.H., Tomlinson, R.L., Terns, R.M., and Terns, M.P. (2010). Telomerase trafficking and assembly in *Xenopus* oocytes. Journal of Cell Science *123*, 2464-2472.

Licht, J.D., and Collins, K. (1999). Telomerase RNA function in recombinant *Tetrahymena* telomerase. Genes & Development *13*, 1116-1125.

Lin, J.J., and 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*. Proceedings of the National Academy of Sciences of the United States of America *93*, 13760-13765.

Lingner, J., and Cech, T.R. (1996). Purification of telomerase from *Euplotes aediculatus*: requirement of a primer 3' overhang. Proceedings of the National Academy of Sciences of the United States of America *93*, 10712-10717.

Lingner, J., and Cech, T.R. (1998). Telomerase and chromosome end maintenance. Current Opinion in Genetics & Development *8*, 226-232.

Lippman, Z., Gendrel, A.V., Black, M., Vaughn, M.W., Dedhia, N., McCombie, W.R., Lavine, K., Mittal, V., May, B., Kasschau, K.D.*, et al.* (2004). Role of transposable elements in heterochromatin and epigenetic control. Nature *430*, 471-476.

Lisch, D., and Bennetzen, J.L. (2011). Transposable element origins of epigenetic gene regulation. Current Opinion in Plant Biology *14*, 156-161.

Liu, Q., Feng, Y., and Zhu, Z. (2009). Dicer-like (DCL) proteins in plants. Functional & Integrative Genomics *9*, 277-286.

Livengood, A.J., Zaug, A.J., and Cech, T.R. (2002). Essential regions of *Saccharomyces cerevisiae* telomerase RNA: separate elements for Est1p and Est2p interaction. Molecular and Cellular Biology *22*, 2366-2374.

Longtine, M.S., Wilson, N.M., Petracek, M.E., and Berman, J. (1989). A yeast telomere binding activity binds to two related telomere sequence motifs and is indistinguishable from RAP1. Current genetics *16*, 225-239.

Lopez, C.R., Ribes-Zamora, A., Indiviglio, S.M., Williams, C.L., Haricharan, S., and Bertuch, A.A. (2011). Ku must load directly onto the chromosome end in order to mediate its telomeric functions. PLoS Genetics *7*, e1002233.

Lundblad, V. (2003). Telomere replication: an Est fest. Current Biology *13*, R439-441.

Lundblad, V., and Blackburn, E.H. (1990). RNA-dependent polymerase motifs in EST1: tentative identification of a protein component of an essential yeast telomerase. Cell *60*, 529-530.

Lundblad, V., and Szostak, J.W. (1989). A mutant with a defect in telomere elongation leads to senescence in yeast. Cell *57*, 633-643.

Lynn, K., Fernandez, A., Aida, M., Sedbrook, J., Tasaka, M., Masson, P., and Barton, M.K. (1999). The PINHEAD/ZWILLE gene acts pleiotropically in *Arabidopsis* development and has overlapping functions with the ARGONAUTE1 gene. Development (Cambridge, England) *126*, 469-481.

Mahadevaiah, S.K., Turner, J.M., Baudat, F., Rogakou, E.P., de Boer, P., Blanco-Rodriguez, J., Jasin, M., Keeney, S., Bonner, W.M., and Burgoyne, P.S. (2001). Recombinational DNA double-strand breaks in mice precede synapsis. Nature Genetics *27*, 271-276.

Maida, Y., Yasukawa, M., Furuuchi, M., Lassmann, T., Possemato, R., Okamoto, N., Kasim, V., Hayashizaki, Y., Hahn, W.C., and Masutomi, K. (2009). An RNA-dependent RNA polymerase formed by TERT and the RMRP RNA. Nature *461*, 230-235.

Maida, Y., Yasukawa, M., Okamoto, N., Ohka, S., Kinoshita, K., Totoki, Y., Ito, T.K., Minamino, T., Nakamura, H., Yamaguchi, S.*, et al.* (2014). Involvement of telomerase reverse transcriptase in heterochromatin maintenance. Molecular and Cellular Biology *34*, 1576-1593.

Makarov, V.L., Hirose, Y., and Langmore, J.P. (1997). Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. Cell *88*, 657-666.

Marcand, S., Wotton, D., Gilson, E., and Shore, D. (1997). Rap1p and telomere length regulation in yeast. Ciba Foundation Symposium *211*, 76-93; discussion 93-103.

Martin, V., Du, L.L., Rozenzhak, S., and Russell, P. (2007). Protection of telomeres by a conserved Stn1-Ten1 complex. Proceedings of the National Academy of Sciences of the United States of America *104*, 14038-14043.

Martinez, P., Thanasoula, M., Carlos, A.R., Gomez-Lopez, G., Tejera, A.M., Schoeftner, S., Dominguez, O., Pisano, D.G., Tarsounas, M., and Blasco, M.A. (2010). Mammalian

Rap1 controls telomere function and gene expression through binding to telomeric and extratelomeric sites. Nature Cell Biology *12*, 768-780.

Matzke, M.A., and Mosher, R.A. (2014). RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. Nature Reviews Genetics *15*, 394-408.

McClintock, B. (1938). The Production of Homozygous Deficient Tissues with Mutant Characteristics by Means of the Aberrant Mitotic Behavior of Ring-Shaped Chromosomes. Genetics *23*, 315-376.

McClintock, B. (1939). The Behavior in Successive Nuclear Divisions of a Chromosome Broken at Meiosis. Proceedings of the National Academy of Sciences of the United States of America *25*, 405-416.

McClintock, B. (1941). The Stability of Broken Ends of Chromosomes in *Zea mays*. Genetics *26*, 234-282.

McEachern, M.J., and Blackburn, E.H. (1995). Runaway telomere elongation caused by telomerase RNA gene mutations. Nature *376*, 403-409.

Melek, M., Greene, E.C., and Shippen, D.E. (1996). Processing of nontelomeric 3' ends by telomerase: default template alignment and endonucleolytic cleavage. Molecular and Cellular Biology *16*, 3437-3445.

Meyerson, M., Counter, C.M., Eaton, E.N., Ellisen, L.W., Steiner, P., Caddle, S.D., Ziaugra, L., Beijersbergen, R.L., Davidoff, M.J., Liu, Q.*, et al.* (1997). hEST2, the putative human telomerase catalytic subunit gene, is up-regulated in tumor cells and during immortalization. Cell *90*, 785-795.

Meyne, J., Ratliff, R.L., and Moyzis, R.K. (1989). Conservation of the human telomere sequence (TTAGGG)n among vertebrates. Proceedings of the National Academy of Sciences of the United States of America *86*, 7049-7053.

Miller, K.M., and Cooper, J.P. (2003). The telomere protein Taz1 is required to prevent and repair genomic DNA breaks. Molecular Cell *11*, 303-313.

Miller, K.M., Ferreira, M.G., and Cooper, J.P. (2005). Taz1, Rap1 and Rif1 act both interdependently and independently to maintain telomeres. The EMBO Journal *24*, 3128-3135.

Min, B., and Collins, K. (2010). Multiple mechanisms for elongation processivity within the reconstituted tetrahymena telomerase holoenzyme. The Journal of Biological Chemistry *285*, 16434-16443.

Mitchell, J.R., Wood, E., and Collins, K. (1999). A telomerase component is defective in the human disease dyskeratosis congenita. Nature *402*, 551-555.

Mitchell-Olds, T., and Schmitt, J. (2006). Genetic mechanisms and evolutionary significance of natural variation in *Arabidopsis*. Nature *441*, 947-952.

Mitton-Fry, R.M., Anderson, E.M., Hughes, T.R., Lundblad, V., and Wuttke, D.S. (2002). Conserved structure for single-stranded telomeric DNA recognition. Science *296*, 145-147.

Miyagawa, K., Low, R.S., Santosa, V., Tsuji, H., Moser, B.A., Fujisawa, S., Harland, J.L., Raguimova, O.N., Go, A., Ueno, M.*, et al.* (2014). SUMOylation regulates telomere length by targeting the shelterin subunit Tpz1(Tpp1) to modulate shelterin-Stn1 interaction in fission yeast. Proceedings of the National Academy of Sciences of the United States of America *111*, 5950-5955.

Miyake, Y., Nakamura, M., Nabetani, A., Shimamura, S., Tamura, M., Yonehara, S., Saito, M., and Ishikawa, F. (2009). RPA-like mammalian Ctc1-Stn1-Ten1 complex binds to single-stranded DNA and protects telomeres independently of the Pot1 pathway. Molecular Cell *36*, 193-206.

Mlotshwa, S., Pruss, G.J., Peragine, A., Endres, M.W., Li, J., Chen, X., Poethig, R.S., Bowman, L.H., and Vance, V. (2008). DICER-LIKE2 plays a primary role in transitive silencing of transgenes in *Arabidopsis*. PloS One *3*, e1755.

Moore, T., and Haig, D. (1991). Genomic imprinting in mammalian development: a parental tug-of-war. Trends in Genetics *7*, 45-49.

Moriarty, T.J., Marie-Egyptienne, D.T., and Autexier, C. (2005). Regulation of 5' template usage and incorporation of noncognate nucleotides by human telomerase. RNA *11*, 1448-1460.

Moser, B.A., Chang, Y.T., Kosti, J., and Nakamura, T.M. (2011). Tel1ATM and Rad3ATR kinases promote Ccq1-Est1 interaction to maintain telomeres in fission yeast. Nature Structural & Molecular Biology *18*, 1408-1413.

Moshous, D., Callebaut, I., de Chasseval, R., Corneo, B., Cavazzana-Calvo, M., Le Deist, F., Tezcan, I., Sanal, O., Bertrand, Y., Philippe, N.*, et al.* (2001). Artemis, a novel DNA double-strand break repair/V(D)J recombination protein, is mutated in human severe combined immune deficiency. Cell *105*, 177-186.

Moyzis, R.K., Buckingham, J.M., Cram, L.S., Dani, M., Deaven, L.L., Jones, M.D., Meyne, J., Ratliff, R.L., and Wu, J.R. (1988). A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes. Proceedings of the National Academy of Sciences of the United States of America *85*, 6622-6626.

Munoz-Jordan, J.L., Cross, G.A., de Lange, T., and Griffith, J.D. (2001). t-loops at trypanosome telomeres. The EMBO Journal *20*, 579-588.

Murti, K.G., and Prescott, D.M. (2002). Topological organization of DNA molecules in the macronucleus of hypotrichous ciliated protozoa. Chromosome Research : an International Journal on the Molecular, Supramolecular and Evolutionary Aspects of Chromosome Biology *10*, 165-173.

Nakamura, T.M., Morin, G.B., Chapman, K.B., Weinrich, S.L., Andrews, W.H., Lingner, J., Harley, C.B., and Cech, T.R. (1997). Telomerase catalytic subunit homologs from fission yeast and human. Science *277*, 955-959.

Nelson, A.D., Lamb, J.C., Kobrossly, P.S., and Shippen, D.E. (2011). Parameters affecting telomere-mediated chromosomal truncation in *Arabidopsis*. The Plant Cell *23*, 2263-2272.

Noel, J.F., Larose, S., Abou Elela, S., and Wellinger, R.J. (2012). Budding yeast telomerase RNA transcription termination is dictated by the Nrd1/Nab3 non-coding RNA termination pathway. Nucleic Acids Research *40*, 5625-5636.

Nordborg, M., Hu, T.T., Ishino, Y., Jhaveri, J., Toomajian, C., Zheng, H., Bakker, E., Calabrese, P., Gladstone, J., Goyal, R.*, et al.* (2005). The pattern of polymorphism in *Arabidopsis thaliana*. PLoS Biology *3*, e196.

Novotny, I., Podolska, K., Blazikova, M., Valasek, L.S., Svoboda, P., and Stanek, D. (2012). Nuclear LSm8 affects number of cytoplasmic processing bodies via controlling cellular distribution of Like-Sm proteins. Molecular Biology of the Cell *23*, 3776-3785.

Nugent, C.I., Hughes, T.R., Lue, N.F., and Lundblad, V. (1996). Cdc13p: a single-strand telomeric DNA-binding protein with a dual role in yeast telomere maintenance. Science *274*, 249-252.

Oguchi, K., Liu, H., Tamura, K., and Takahashi, H. (1999). Molecular cloning and characterization of AtTERT, a telomerase reverse transcriptase homolog in *Arabidopsis thaliana*. FEBS Letters *457*, 465-469.

Okabe, J., Eguchi, A., Masago, A., Hayakawa, T., and Nakanishi, M. (2000). TRF1 is a critical trans-acting factor required for *de novo* telomere formation in human cells. Human Molecular Genetics *9*, 2639-2650.

Olovnikov, A.M. (1971). [Principle of marginotomy in template synthesis of polynucleotides]. Doklady Akademii Nauk SSSR *201*, 1496-1499.

Ometto, L., Glinka, S., De Lorenzo, D., and Stephan, W. (2005). Inferring the effects of demography and selection on *Drosophila melanogaster* populations from a chromosome-wide scan of DNA variation. Molecular Biology and Evolution *22*, 2119-2130.

Pecinka, A., Rosa, M., Schikora, A., Berlinger, M., Hirt, H., Luschnig, C., and Mittelsten Scheid, O. (2009). Transgenerational stress memory is not a general response in *Arabidopsis*. PloS One *4*, e5202.

Pennaneach, V., Putnam, C.D., and Kolodner, R.D. (2006). Chromosome healing by *de novo* telomere addition in *Saccharomyces cerevisiae*. Molecular Microbiology *59*, 1357-1368.

Pfingsten, J.S., Goodrich, K.J., Taabazuing, C., Ouenzar, F., Chartrand, P., and Cech, T.R. (2012). Mutually exclusive binding of telomerase RNA and DNA by Ku alters telomerase recruitment model. Cell *148*, 922-932.

Pinillos, V., and Cuevas, J. (2008). Standardization of the fluorochromatic reaction test to assess pollen viability. Biotechnic & Histochemistry : Official Publication of the Biological Stain Commission *83*, 15-21.

Pluta, A.F., Kaine, B.P., and Spear, B.B. (1982). The terminal organization of macronuclear DNA in *Oxytricha fallax*. Nucleic Acids Research *10*, 8145-8154.

Ponting, C.P., Oliver, P.L., and Reik, W. (2009). Evolution and functions of long noncoding RNAs. Cell *136*, 629-641.

Price, C.M., Boltz, K.A., Chaiken, M.F., Stewart, J.A., Beilstein, M.A., and Shippen, D.E. (2010). Evolution of CST function in telomere maintenance. Cell Cycle *9*, 3157-3165.

Prowse, K.R., Avilion, A.A., and Greider, C.W. (1993). Identification of a nonprocessive telomerase activity from mouse cells. Proceedings of the National Academy of Sciences of the United States of America *90*, 1493-1497.

Prowse, K.R., and Greider, C.W. (1995). Developmental and tissue-specific regulation of mouse telomerase and telomere length. Proceedings of the National Academy of Sciences of the United States of America *92*, 4818-4822.

Qi, X., Li, Y., Honda, S., Hoffmann, S., Marz, M., Mosig, A., Podlevsky, J.D., Stadler, P.F., Selker, E.U., and Chen, J.J. (2013). The common ancestral core of vertebrate and fungal telomerase RNAs. Nucleic Acids Research *41*, 450-462.

Qi, X., Rand, D.P., Podlevsky, J.D., Li, Y., Mosig, A., Stadler, P.F., and Chen, J.J. (2015). Prevalent and distinct spliceosomal 3'-end processing mechanisms for fungal telomerase RNA. Nature Communications *6*, 6105.

Rajagopalan, R., Vaucheret, H., Trejo, J., and Bartel, D.P. (2006). A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*. Genes & Development *20*, 3407-3425.

Ramsay, A.J., Quesada, V., Foronda, M., Conde, L., Martinez-Trillos, A., Villamor, N., Rodriguez, D., Kwarciak, A., Garabaya, C., Gallardo, M.*, et al.* (2013). POT1 mutations cause telomere dysfunction in chronic lymphocytic leukemia. Nature Genetics *45*, 526-530.

Ren, S., Johnston, J.S., Shippen, D.E., and McKnight, T.D. (2004). TELOMERASE ACTIVATOR1 induces telomerase activity and potentiates responses to auxin in *Arabidopsis*. The Plant Cell *16*, 2910-2922.

Renfrew, K.B., Song, X., Lee, J.R., Arora, A., and Shippen, D.E. (2014). POT1a and components of CST engage telomerase and regulate its activity in *Arabidopsis*. PLoS Genetics *10*, e1004738.

Ribeyre, C., and Shore, D. (2013). Regulation of telomere addition at DNA double-strand breaks. Chromosoma *122*, 159-173.

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

Riha, K., Heacock, M.L., and Shippen, D.E. (2006). The role of the nonhomologous end-joining DNA double-strand break repair pathway in telomere biology. Annual review of Genetics *40*, 237-277.

Riha, K., McKnight, T.D., Fajkus, J., Vyskot, B., and Shippen, D.E. (2000). Analysis of the G-overhang structures on plant telomeres: evidence for two distinct telomere architectures. The Plant Journal : for Cell and Molecular Biology *23*, 633-641.

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

Riha, K., and Shippen, D.E. (2003). Ku is required for telomeric C-rich strand maintenance but not for end-to-end chromosome fusions in *Arabidopsis*. Proceedings of the National Academy of Sciences of the United States of America *100*, 611-615.

Riou, L., Bastos, H., Lassalle, B., Coureuil, M., Testart, J., Boussin, F.D., Allemand, I., and Fouchet, P. (2005). The telomerase activity of adult mouse testis resides in the spermatogonial alpha6-integrin-positive side population enriched in germinal stem cells. Endocrinology *146*, 3926-3932.

Romero, D.P., and Blackburn, E.H. (1991). A conserved secondary structure for telomerase RNA. Cell *67*, 343-353.

Samach, A., Melamed-Bessudo, C., Avivi-Ragolski, N., Pietrokovski, S., and Levy, A.A. (2011). Identification of plant RAD52 homologs and characterization of the *Arabidopsis thaliana* RAD52-like genes. The Plant Cell *23*, 4266-4279.

Scheuermann, J.C., and Boyer, L.A. (2013). Getting to the heart of the matter: long non-coding RNAs in cardiac development and disease. The EMBO Journal *32*, 1805-1816.

Schmid, K.J., Ramos-Onsins, S., Ringys-Beckstein, H., Weisshaar, B., and Mitchell-Olds, T. (2005). A multilocus sequence survey in *Arabidopsis thaliana* reveals a

genome-wide departure from a neutral model of DNA sequence polymorphism. Genetics *169*, 1601-1615.

Schulz, V.P., and Zakian, V.A. (1994). The *Saccharomyces* PIF1 DNA helicase inhibits telomere elongation and de novo telomere formation. Cell *76*, 145-155.

Schwanhausser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. Nature *473*, 337-342.

Sela, N., Mersch, B., Hotz-Wagenblatt, A., and Ast, G. (2010). Characteristics of transposable element exonization within human and mouse. PloS One *5*, e10907.

Seto, A.G., Umansky, K., Tzfati, Y., Zaug, A.J., Blackburn, E.H., and Cech, T.R. (2003). A template-proximal RNA paired element contributes to *Saccharomyces cerevisiae* telomerase activity. RNA *9*, 1323-1332.

Shakirov, E.V., and Shippen, D.E. (2004). Length regulation and dynamics of individual telomere tracts in wild-type *Arabidopsis*. The Plant Cell *16*, 1959-1967.

Shakirov, E.V., Surovtseva, Y.V., Osbun, N., and Shippen, D.E. (2005). The *Arabidopsis* Pot1 and Pot2 proteins function in telomere length homeostasis and chromosome end protection. Molecular and Cellular Biology *25*, 7725-7733.

Shampay, J., and Blackburn, E.H. (1988). Generation of telomere-length heterogeneity in *Saccharomyces cerevisiae*. Proceedings of the National Academy of Sciences of the United States of America *85*, 534-538.

Shi, J., Yang, X.R., Ballew, B., Rotunno, M., Calista, D., Fargnoli, M.C., Ghiorzo, P., Bressac-de Paillerets, B., Nagore, E., Avril, M.F.*, et al.* (2014). Rare missense variants in POT1 predispose to familial cutaneous malignant melanoma. Nature Genetics *46*, 482-486.

Shippen-Lentz, D., and Blackburn, E.H. (1990). Functional evidence for an RNA template in telomerase. Science *247*, 546-552.

Shore, D., and Bianchi, A. (2009). Telomere length regulation: coupling DNA end processing to feedback regulation of telomerase. The EMBO Journal *28*, 2309-2322.

Shore, D., and Nasmyth, K. (1987). Purification and cloning of a DNA binding protein from yeast that binds to both silencer and activator elements. Cell *51*, 721-732.

Siderakis, M., and Tarsounas, M. (2007). Telomere regulation and function during meiosis. Chromosome Research : an International Journal on the Molecular, Supramolecular and Evolutionary Aspects of Chromosome Biology *15*, 667-679.

Smogorzewska, A., van Steensel, B., Bianchi, A., Oelmann, S., Schaefer, M.R., Schnapp, G., and de Lange, T. (2000). Control of human telomere length by TRF1 and TRF2. Molecular and Cellular Biology *20*, 1659-1668.

Song, X., Leehy, K., Warrington, R.T., Lamb, J.C., Surovtseva, Y.V., and Shippen, D.E. (2008). STN1 protects chromosome ends in Proceedings of the National Academy of Sciences of the United States of America *105*, 19815-19820.

Spiller, M.P., Boon, K.L., Reijns, M.A., and Beggs, J.D. (2007). The Lsm2-8 complex determines nuclear localization of the spliceosomal U6 snRNA. Nucleic Acids Research *35*, 923-929.

Stansel, R.M., de Lange, T., and Griffith, J.D. (2001). T-loop assembly in vitro involves binding of TRF2 near the 3' telomeric overhang. The EMBO Journal *20*, 5532-5540.

Stellwagen, A.E., Haimberger, Z.W., Veatch, J.R., and Gottschling, D.E. (2003). Ku interacts with telomerase RNA to promote telomere addition at native and broken chromosome ends. Genes & Development *17*, 2384-2395.

Storlazzi, A., Tesse, S., Gargano, S., James, F., Kleckner, N., and Zickler, D. (2003). Meiotic double-strand breaks at the interface of chromosome movement, chromosome remodeling, and reductional division. Genes & Development *17*, 2675-2687.

Stracker, T.H., and Petrini, J.H. (2011). The MRE11 complex: starting from the ends. Nature Reviews Molecular Cell Biology *12*, 90-103.

Studer, A., Zhao, Q., Ross-Ibarra, J., and Doebley, J. (2011). Identification of a functional transposon insertion in the maize domestication gene tb1. Nature Genetics *43*, 1160-1163.

Surovtseva, Y.V., Churikov, D., Boltz, K.A., Song, X., Lamb, J.C., Warrington, R., Leehy, K., Heacock, M., Price, C.M., and Shippen, D.E. (2009). Conserved telomere

maintenance component 1 interacts with STN1 and maintains chromosome ends in higher eukaryotes. Molecular Cell *36*, 207-218.

Surovtseva, Y.V., Shakirov, E.V., Vespa, L., Osbun, N., Song, X., and Shippen, D.E. (2007). Arabidopsis POT1 associates with the telomerase RNP and is required for telomere maintenance. The EMBO Journal *26*, 3653-3661.

Symington, L.S., and Gautier, J. (2011). Double-strand break end resection and repair pathway choice. Annual Review of Genetics *45*, 247-271.

Szilard, R.K., Jacques, P.E., Laramee, L., Cheng, B., Galicia, S., Bataille, A.R., Yeung, M., Mendez, M., Bergeron, M., Robert, F.*, et al.* (2010). Systematic identification of fragile sites via genome-wide location analysis of gamma-H2AX. Nature Structural & Molecular Biology *17*, 299-305.

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

Tabara, H., Yigit, E., Siomi, H., and Mello, C.C. (2002). The dsRNA binding protein RDE-4 interacts with RDE-1, DCR-1, and a DExH-box helicase to direct RNAi in *C. elegans*. Cell *109*, 861-871.

Taggart, A.K., Teng, S.C., and Zakian, V.A. (2002). Est1p as a cell cycle-regulated activator of telomere-bound telomerase. Science *297*, 1023-1026.

Takai, K.K., Kibe, T., Donigian, J.R., Frescas, D., and de Lange, T. (2011). Telomere protection by TPP1/POT1 requires tethering to TIN2. Molecular Cell *44*, 647-659.

Takanashi, H., Ohnishi, T., Mogi, M., Hirata, Y., and Tsutsumi, N. (2011). DCL2 is highly expressed in the egg cell in both rice and *Arabidopsis*. Plant Signaling & Behavior *6*, 604-606.

Tamura, K., Liu, H., and Takahashi, H. (1999). Auxin induction of cell cycle regulated activity of tobacco telomerase. The Journal of Biological Chemistry *274*, 20997-21002.

Tanemura, K., Ogura, A., Cheong, C., Gotoh, H., Matsumoto, K., Sato, E., Hayashi, Y., Lee, H.W., and Kondo, T. (2005). Dynamic rearrangement of telomeres during spermatogenesis in mice. Developmental Biology *281*, 196-207.

Tang, W., Kannan, R., Blanchette, M., and Baumann, P. (2012). Telomerase RNA biogenesis involves sequential binding by Sm and Lsm complexes. Nature *484*, 260-264.

Taylor, D.W., Ma, E., Shigematsu, H., Cianfrocco, M.A., Noland, C.L., Nagayama, K., Nogales, E., Doudna, J.A., and Wang, H.W. (2013). Substrate-specific structural rearrangements of human Dicer. Nature Structural & Molecular biology *20*, 662-670.

Theobald, D.L., and Wuttke, D.S. (2004). Prediction of multiple tandem OB-fold domains in telomere end-binding proteins Pot1 and Cdc13. Structure *12*, 1877-1879.

Ting, N.S., Yu, Y., Pohorelic, B., Lees-Miller, S.P., and Beattie, T.L. (2005). Human Ku70/80 interacts directly with hTR, the RNA component of human telomerase. Nucleic Acids Research *33*, 2090-2098.

Tomita, K., and Cooper, J.P. (2007). The telomere bouquet controls the meiotic spindle. Cell *130*, 113-126.

Tzfati, Y., Fulton, T.B., Roy, J., and Blackburn, E.H. (2000). Template boundary in a yeast telomerase specified by RNA structure. Science *288*, 863-867.

Tzfati, Y., Knight, Z., Roy, J., and Blackburn, E.H. (2003). A novel pseudoknot element is essential for the action of a yeast telomerase. Genes & Development *17*, 1779-1788.

van Steensel, B., and de Lange, T. (1997). Control of telomere length by the human telomeric protein TRF1. Nature *385*, 740-743.

Vaucheret, H. (2006). Post-transcriptional small RNA pathways in plants: mechanisms and regulations. Genes & Development *20*, 759-771.

Venkatesan, R.N., and Price, C. (1998). Telomerase expression in chickens: constitutive activity in somatic tissues and down-regulation in culture. Proceedings of the National Academy of Sciences of the United States of America *95*, 14763-14768.

Venteicher, A.S., Abreu, E.B., Meng, Z., McCann, K.E., Terns, R.M., Veenstra, T.D., Terns, M.P., and Artandi, S.E. (2009). A human telomerase holoenzyme protein required for Cajal body localization and telomere synthesis. Science *323*, 644-648.

Vodenicharov, M.D., and Wellinger, R.J. (2006). DNA degradation at unprotected telomeres in yeast is regulated by the CDK1 (Cdc28/Clb) cell-cycle kinase. Molecular Cell *24*, 127-137.

Vulliamy, T., Marrone, A., Goldman, F., Dearlove, A., Bessler, M., Mason, P.J., and Dokal, I. (2001). The RNA component of telomerase is mutated in autosomal dominant dyskeratosis congenita. Nature *413*, 432-435.

Wallweber, G., Gryaznov, S., Pongracz, K., and Pruzan, R. (2003). Interaction of human telomerase with its primer substrate. Biochemistry *42*, 589-600.

Walne, A.J., Bhagat, T., Kirwan, M., Gitiaux, C., Desguerre, I., Leonard, N., Nogales, E., Vulliamy, T., and Dokal, I.S. (2013). Mutations in the telomere capping complex in bone marrow failure and related syndromes. Haematologica *98*, 334-338.

Wang, H., and Blackburn, E.H. (1997). De novo telomere addition by *Tetrahymena* telomerase *in vitro*. The EMBO Journal *16*, 866-879.

Wang, K.C., and Chang, H.Y. (2011). Molecular mechanisms of long noncoding RNAs. Molecular Cell *43*, 904-914.

Wang, X., and Baumann, P. (2008). Chromosome fusions following telomere loss are mediated by single-strand annealing. Molecular Cell *31*, 463-473.

Wang, X., Weigel, D., and Smith, L.M. (2013). Transposon variants and their effects on gene expression in *Arabidopsis*. PLoS Genetics *9*, e1003255.

Watson, J.D. (1972). Origin of concatemeric T7 DNA. Nature: New Biology *239*, 197-201.

Watson, J.D., and Crick, F.H. (1953). The structure of DNA. Cold Spring Harbor Symposia on Quantitative Biology *18*, 123-131.

Webb, C.J., and Zakian, V.A. (2012). *Schizosaccharomyces pombe* Ccq1 and TER1 bind the 14-3-3-like domain of Est1, which promotes and stabilizes telomerase-telomere association. Genes & Development *26*, 82-91.

Wei, W., Ba, Z., Gao, M., Wu, Y., Ma, Y., Amiard, S., White, C.I., Rendtlew Danielsen, J.M., Yang, Y.G., and Qi, Y. (2012). A role for small RNAs in DNA double-strand break repair. Cell *149*, 101-112.

Weinberg, D.E., Nakanishi, K., Patel, D.J., and Bartel, D.P. (2011). The inside-out mechanism of Dicers from budding yeasts. Cell *146*, 262-276.

Wellinger, R.J., Wolf, A.J., and Zakian, V.A. (1993). Saccharomyces telomeres acquire single-strand TG1-3 tails late in S phase. Cell *72*, 51-60.

Wong, J.M., Kusdra, L., and Collins, K. (2002). Subnuclear shuttling of human telomerase induced by transformation and DNA damage. Nature Cell Biology *4*, 731-736.

Wong, M.S., Wright, W.E., and Shay, J.W. (2014). Alternative splicing regulation of telomerase: a new paradigm? Trends in Genetics *30*, 430-438.

Wotton, D., and Shore, D. (1997). A novel Rap1p-interacting factor, Rif2p, cooperates with Rif1p to regulate telomere length in *Saccharomyces cerevisiae*. Genes & Development *11*, 748-760.

Wright, W.E., Piatyszek, M.A., Rainey, W.E., Byrd, W., and Shay, J.W. (1996). Telomerase activity in human germline and embryonic tissues and cells. Developmental Genetics *18*, 173-179.

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

Wu, P., van Overbeek, M., Rooney, S., and de Lange, T. (2010). Apollo contributes to G overhang maintenance and protects leading-end telomeres. Molecular Cell *39*, 606-617.

Wu, R.A., and Collins, K. (2014). Human telomerase specialization for repeat synthesis by unique handling of primer-template duplex. The EMBO Journal *33*, 921-935.

Wu, Y., and Zakian, V.A. (2011). The telomeric Cdc13 protein interacts directly with the telomerase subunit Est1 to bring it to telomeric DNA ends *in vitro*. Proceedings of the National Academy of Sciences of the United States of America *108*, 20362-20369.

Wutz, A., Rasmussen, T.P., and Jaenisch, R. (2002). Chromosomal silencing and localization are mediated by different domains of Xist RNA. Nature Genetics *30*, 167-174.

Xin, H., Liu, D., Wan, M., Safari, A., Kim, H., Sun, W., O'Connor, M.S., and Songyang, Z. (2007). TPP1 is a homologue of ciliate TEBP-beta and interacts with POT1 to recruit telomerase. Nature *445*, 559-562.

Yamazaki, H., Tarumoto, Y., and Ishikawa, F. (2012). Tel1(ATM) and Rad3(ATR) phosphorylate the telomere protein Ccq1 to recruit telomerase and elongate telomeres in fission yeast. Genes & Development *26*, 241-246.

Yi, X., Tesmer, V.M., Savre-Train, I., Shay, J.W., and Wright, W.E. (1999). Both transcriptional and posttranscriptional mechanisms regulate human telomerase template RNA levels. Molecular and Cellular Biology *19*, 3989-3997.

Yu, G.L., Bradley, J.D., Attardi, L.D., and Blackburn, E.H. (1990). In vivo alteration of telomere sequences and senescence caused by mutated *Tetrahymena* telomerase RNAs. Nature *344*, 126-132.

Zappulla, D.C., Roberts, J.N., Goodrich, K.J., Cech, T.R., and Wuttke, D.S. (2009). Inhibition of yeast telomerase action by the telomeric ssDNA-binding protein, Cdc13p. Nucleic Acids Research *37*, 354-367.

Zaug, A.J., Linger, J., and Cech, T.R. (1996). Method for determining RNA 3' ends and application to human telomerase RNA. Nucleic Acids Research *24*, 532-533.

Zellinger, B., Akimcheva, S., Puizina, J., Schirato, M., and Riha, K. (2007). Ku suppresses formation of telomeric circles and alternative telomere lengthening in *Arabidopsis*. Molecular Cell *27*, 163-169.

Zellinger, B., and Riha, K. (2007). Composition of plant telomeres. Biochimica et Biophysica Acta *1769*, 399-409.

Zhang, W., and Durocher, D. (2010). *De novo* telomere formation is suppressed by the Mec1-dependent inhibition of Cdc13 accumulation at DNA breaks. Genes & Development *24*, 502-515.

Zhang, X., Henriques, R., Lin, S.S., Niu, Q.W., and Chua, N.H. (2006). *Agrobacterium*-mediated transformation of *Arabidopsis thaliana* using the floral dip method. Nature Protocols *1*, 641-646.

Zhao, Y., Abreu, E., Kim, J., Stadler, G., Eskiocak, U., Terns, M.P., Terns, R.M., Shay, J.W., and Wright, W.E. (2011). Processive and distributive extension of human telomeres by telomerase under homeostatic and nonequilibrium conditions. Molecular Cell *42*, 297-307.

Zhong, F.L., Batista, L.F., Freund, A., Pech, M.F., Venteicher, A.S., and Artandi, S.E. (2012). TPP1 OB-fold domain controls telomere maintenance by recruiting telomerase to chromosome ends. Cell *150*, 481-494.

Zhong, Z., Shiue, L., Kaplan, S., and de Lange, T. (1992). A mammalian factor that binds telomeric TTAGGG repeats in vitro. Molecular and Cellular Biology *12*, 4834-4843.

Zhou, B.B., and Elledge, S.J. (2000). The DNA damage response: putting checkpoints in perspective. Nature *408*, 433-439.

Zhou, J., Monson, E.K., Teng, S.C., Schulz, V.P., and Zakian, V.A. (2000). Pif1p helicase, a catalytic inhibitor of telomerase in yeast. Science *289*, 771-774.

Zhu, H., Zhou, Y., Castillo-Gonzalez, C., Lu, A., Ge, C., Zhao, Y.T., Duan, L., Li, Z., Axtell, M.J., Wang, X.J.*, et al.* (2013). Bidirectional processing of pri-miRNAs with branched terminal loops by Arabidopsis Dicer-like1. Nature Structural & Molecular Biology *20*, 1106-1115.