TELOMERE-ASSOCIATED PROTEINS IN *Arabidopsis thaliana*

A Dissertation

by

YULIA V. SUROVTSEVA

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

DOCTOR OF PHILOSOPHY

May 2008

Major Subject: Biochemistry

TELOMERE-ASSOCIATED PROTEINS IN *Arabidopsis thaliana*

A Dissertation

by

YULIA V. SUROVTSEVA

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

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Dorothy E. Shippen
Committee Members,	Thomas McKnight
	David Peterson
	Gregory D. Reinhart
Head of Department,	Gregory D. Reinhart

May 2008

Major Subject: Biochemistry

ABSTRACT

Telomere-Associated Proteins in *Arabidopsis thaliana*.

(May 2008)

Yulia V. Surovtseva, B.S., Novosibirsk State University, Russia

Chair of Advisory Committee: Dr. Dorothy E. Shippen

Telomeres comprise the physical ends of chromosomes. Essential functions of telomeres include protecting the terminus from being recognized as a DNA double-strand break and facilitating the complete replication of the physical end of the DNA. Telomere functions are mediated by a large array of telomere-associated proteins. Mutations in telomere-related genes cause immediate telomere dysfunction, activation of DNA damage response, and accumulation of end-to-end chromosome fusions. In addition, changes in telomere complex composition may affect the ability of the telomerase enzyme to maintain telomeres *in vivo*.

Here, we describe the characterization of telomere-associated proteins in the flowering plant, *Arabidopsis thaliana*. Using a bioinformatics approach, we identified twelve proteins with sequence similarity to vertebrate duplex telomere DNA binding proteins TRF1 and TRF2. We showed that, like their vertebrate counterparts, some of the *Arabidopsis* TRFL (TRF-LIKE) proteins can homodimerize and bind telomeric DNA *in vitro*, indicating that *Arabidopsis* encodes a large family of double-strand telomeric DNA binding proteins. We have also characterized three *Arabidopsis* POT1 proteins whose homologs in yeast and vertebrates associate with the single-stranded portion of telomeric DNA. Unexpectedly, we found that unlike POT1 protein in other organisms, *Arabidopsis* AtPOT1a protein associates with telomeres only in the S phase of the cell

cycle and is a physical component of the active telomerase RNP complex, providing positive telomere length regulation. Our data implicated AtPOT1b, another *Arabidopsis* POT1 protein, in chromosome end protection. Finally, we showed that *Arabidopsis thaliana* has evolved a third POT1 protein, AtPOT1c, which contributes to both telomere length regulation and telomerase activity, and maintenance of the structure of the chromosome terminus. Thus, *Arabidopsis* has evolved a set of POT1 proteins that make distinct and novel contributions to telomere biology.

Finally, we describe the identification and characterization of a novel *Arabidopsis* protein CIT1 (Critical for Integrity of Telomeres 1), and show that CIT1 deficiency leads to an immediate and profound telomere dysfunction and chromosome end deprotection. Altogether, these data provide new insight into plant telomere-associated factors and significantly improve our understanding of the overall architecture and evolution of telomeric complex in *Arabidopsis*.

DEDICATION

To my family members, for their love and for being so supportive and unreasonably proud of me.

To my husband, who knows me better than anybody else. Thank you for your never-ending love and support.

To my son, who screams with delight when I come back home from work, and who loves me without any questions asked.

I could not have done this without all of you.

ACKNOWLEDGEMENTS

Most importantly, I would like to thank my committee chair, Dr. Dorothy Shippen. She has taught me not only how to conduct good research experiments, but also how to think and see a general picture. Dr. Shippen has become a role model for me over these years. Her advice and guidance have impacted both my research career and my life outside the lab.

I would like to thank my committee, Dr. David Peterson, Dr. Thomas McKnight, and Dr. Patricia LiWang, for their thoughtful advice and help throughout my graduate career. I would also like to thank Dr. Gregory Reinhart for agreeing to join my committee for the final exam.

I am very grateful to Tom McKnight who always had excellent ideas and whose expertise has helped a lot with the design and interpretation of my experiments. I am also very grateful to Dr. Geoffrey Kapler for his sincere interest in my work and for all his suggestions and ideas. I would also like to thank Dr. Allan Pepper for his advice and help on map-based cloning.

I would like to thank Dr. Zemfira Karamysheva and Dr. Eugene Shakirov. They both helped me tremendously when I first joined the lab, and taught me to be a good scientist. Special thanks go to Dr. Eugene Shakirov. We have collaborated on the AtPOT1a and AtPOT1b projects. We are the first co-authors on both MCB and EMBO papers reprinted in Chapters III and IV. These papers would have been impossible to publish without Dr. Shakirov. I also thank Andrew Nelson for his enthusiasm for the AtPOT1c project and for the development of quantitative TRAP assay.

I would also like to thank all the co-authors on papers reprinted in Chapters II, III and IV. For Chapter II, Zemfira Karamysheva is thanked for the identification of twelve TRFL genes in *Arabidopsis* and for the characterization of the myb-extension domain. Laurent Vespa and Eugene Shakirov are thanked for cloning of some of the TRFL genes. For Chapter III, Eugene Shakirov is thanked for making constructs and for his work on telomere analysis in POT1a and POT1b mutants. Nathan Osbun is also thanked for all his help. For Chapter IV, Eugene Shakirov is thanked for the genetic characterization of *pot1* mutants, Laurent Vespa is thanked for help with chromatin immunoprecipitation, and Xiangyu Song is thanked for her work with *pot1* callus.

I thank all the rotation students and undergraduate students who have helped me a lot over these years. Special thanks go to Yixiao Liu for her help on the AtPOT1c project, and to Xin Tao and Kara Boltz for their hard work on the CIT1 project.

I was very fortunate to work in the Department of Biochemistry and Biophysics. I am very grateful to Juanita, Pat, Tillie, and Sherry, who are always there for us and who made my life easier on so many occasions. I am also grateful to BGA for their guidance and help. Their help is extremely valuable for international students who know nothing about life in the US.

Finally, I am in deep gratitude to all the members of the Shippen lab. They have become a family for me in College Station. I am especially grateful to Matt Watson, who was always able to find time to answer my questions and help me. Special thanks are also owed to Michelle Heacock and Kalpana Kannan, who have suffered and laughed through graduate school with me. Dr. Laurent Vespa, Dr. Rachel Idol, Xiangyu Song and Dr. Jon Lamb are also thanked for making my graduate school an enjoyable and successful experience.

TABLE OF CONTENTS

CHAPTER Page

LIST OF FIGURES

Page

Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

In the 1920s, Herman Muller realized that the ends of the chromosomes have properties distinct from the rest of the chromosome. Working with X-ray mutagenized *Drosophila*, he noticed that flies with deletions or inversions at the chromosome terminus could never be recovered. He concluded that something must seal the end of the chromosome. He coined the term telomere ("end part" in Greek) for the chromosome terminus (Muller, 1938).

In 1930s, cytogeneticist Barbara McClintock was working with maize mutants that contained chromosomes bearing two centromeres. Such dicentric chromosomes can be pulled apart in anaphase, eventually resulting in chromosome breakage. The broken ends can be repaired by the cell by fusion to another broken end, thus creating a dicentric chromosome again. The process is then repeated. Barbara McClintock called this process the breakage-fusion-bridge (BFB) cycle. Interestingly, while BFB cycles were readily observed in the endosperm, it never occurred during the development of the embryo. She called this phenomenon "chromosome healing", because the chromosomes had become stabilized (McClintock, 1939; McClintock, 1941).

This dissertation follows the style and format of *EMBO Journal*.

Starting nearly 70 years ago by these two biologists, the telomere field has been greatly expanded. The chromosome healing activity that McClintock first discovered was shown to be telomerase. It has been experimentally proven that telomeres are essential for chromosome end protection and for maintaining genome integrity. Linear ends of chromosomes need to be protected from various deleterious activities such as degradation by nucleases, recognition as a DNA break by DNA damage repair machinery, and inappropriate recombination. Telomeres achieve this by forming a protective nucleoprotein cap at the chromosome terminus.

Telomeres

Telomeres are the physical ends of linear eukaryotic chromosomes. Telomeric DNA is usually composed of long arrays of simple G-rich repeats. The first telomere composition was identified by direct sequencing in *Tetrahymena*, where chromosomes end in tandem copies of TTGGG repeats (Blackburn and Gall, 1978). In vertebrates, telomeres consist of TTAGGG repeats (Moyzis *et al*., 1988), while in *Arabidopsis* and most other plants this telomere repeat contains ine additional thymine (TTTAGGG) (Burr *et al*., 1992; Higashiyama *et al*., 1995; Kilian *et al*., 1995; Richards and Ausubel, 1988). Although chromosomes end with telomere repeats in most organisms with linear genomes, there are a few examples of alternative strategies for chromosome end maintenance. For example, poxvirus chromosomes end in a hairpin structure (reviewed in Kobryn and Chaconas, 2001). In eukaryotes, *Drosophila* chromosome ends are represented and maintained by retrotransposons. Occasionally, the new retrotransposon is added to the very end of the chromosome to solve the end replication problem (se below) (Biessmann *et al*., 1992).

G-overhangs

While bulk telomeric DNA is double-stranded (ds), the G-rich strand runs in the 5' to 3' direction relative to the terminus and ends in a short single-stranded (ss) 3' overhang (G-overhang) (Figure 1). The first evidence for the presence of G-overhangs came from the ciliates *Oxytricha nova* and *Euplotes crassus.* When the telomeric DNA from these organisms was sequenced, the 3' terminal sequence appeared to be complementary to the 5' terminal sequence, but was 16 nucleotides longer (Klobutcher *et al*., 1981). Later, ss TG_{1-3} tails were found in *Sacharomyces cerevisiae* by non-denaturing in-gel Southern hybridization (Wellinger *et al*., 1993). G-overhangs have now been detected in other model systems, including humans (Makarov *et al*., 1997) and plants (Riha *et al*., 2000; Riha and Shippen, 2003a).

The length of the G-overhang varies among different species. For example, in vertebrates the G-overhang is ~250 nt (Makarov *et al*., 1997), while in wild type *Arabidopsis* G-overhangs are ~20 nt long (Riha *et al*., 2000). The exact mechanism of G-overhang generation is unclear. During lagging strand synthesis, G-overhangs are naturally produced when the most 3' RNA primer is removed (Figure 2). Therefore, a ss overhang corresponding to the size of the primer (or longer, if the primer was not positioned at the very end of DNA strand) is expected. In contrast, the leading strand replication machinery produces a blunt end (Figure 2). Therefore, another mechanism for G-overhang generation must exist to make the terminus symmetrical (Makarov *et al*., 1997; Wellinger *et al*., 1996; Wright *et al*., 1997). It is currently thought that the G-overhangs are formed through a combination of telomerase action (see below) and a nucleolytic resection of the C-strand (Chai *et al*., 2006; Jacob *et al*., 2003; Larrivee *et al*., 2004). Interestingly, it has recently been reported that in addition to G-overhangs,

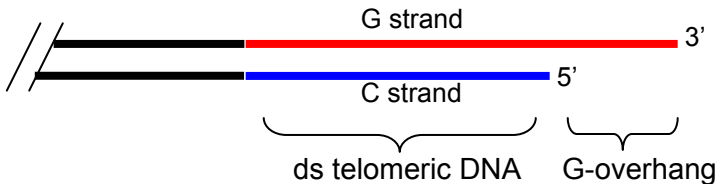


Figure 1. Telomeres – the ends of linear chromosomes. Telomeric DNA consists of ds and ss portions. The G-rich strand (red) ends in a 3’ G-overhang. The C-rich strand is shown in blue.

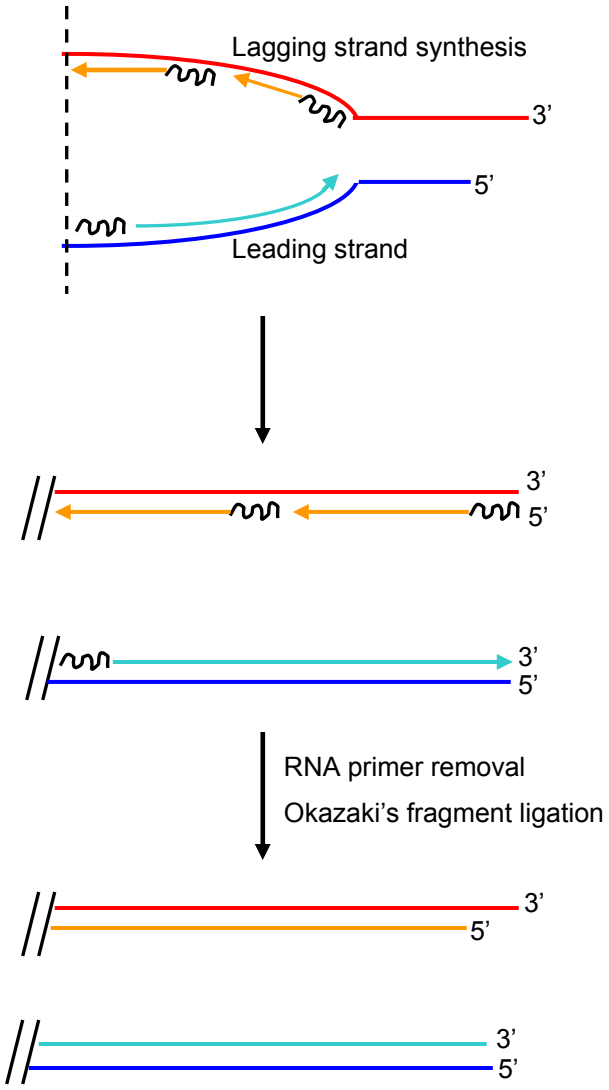


Figure 2. Telomere replication. The replication fork contains a leading (blue) and lagging (orange) strand. Lagging strand synthesis is primed by RNA oligonucleotides (wavy black lane). Following RNA primer removal, the 3’ end G-overhangs are formed on the telomere replicated by the lagging strand machinery. In contrast, leading strand replication machinery produces a blunt end.

C. elegans possess telomeric C-overhangs (Raices *et al*., in press). Moreover, the G- and C-overhangs are bound by different ss telomere binding proteins. The relevance of similar structures in other organisms remains to be determined. The G-overhang is an important structural and functional feature of the chromosome. Although the exact role of telomeric G-overhangs in cell viability is unknown, exposure of the G-overhang has been implicated in senescence in human cells (Li *et al*., 2003; Stewart *et al*., 2003). G-overhangs are also essential for telomere replication *in vivo* and to protect telomeres from end-to-end chromosome fusions (Zhu *et al*., 2003).

T-loops

G-overhangs are also essential for the formation of a t-loop at the chromosome end. Using a telomeric DNA model, Jack Griffith observed lariat-like molecules *in vitro* by electron microscopy (Griffith *et al*., 1999). Psoralen cross-linking of the DNA strands was required to preserve the lariat structures. Moreover, the presence of the G-overhang was necessary for loop formation *in vitro*. Altogether, these data imply that loop formation includes invasion of the G-overhang into duplex DNA (Griffith *et al*., 1999). In the current model, the G-overhang folds back and invades the duplex region of the telomere, creating a displacement loop (D-loop) consisting of ss G-rich repeats (Figure 3). The resulting complex secondary structure of the telomere is called a t-loop (Cesare *et al*., 2003; Griffith *et al*., 1999; Munoz-Jordan *et al*., 2001). T-loops have been observed *in vitro* in many eukaryotes, including plants (Cesare *et al*., 2003; de Lange, 2004). Interestingly, t-loops have not been found in budding yeast. Since budding yeast telomere repeats are highly irregular, the limited base-pairing between

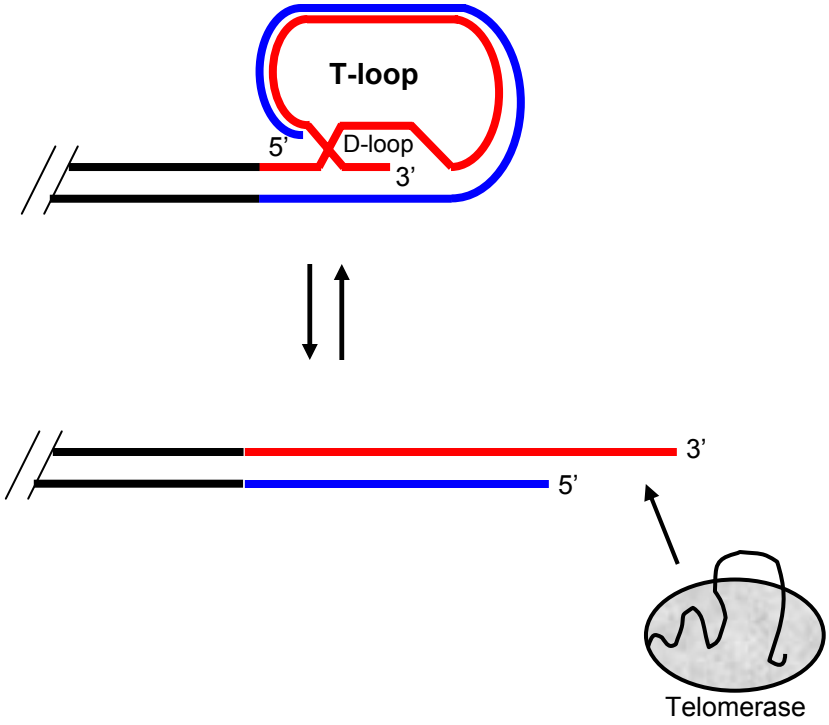


Figure 3. Telomeres form a t-loop. In the t-loop structure, the G-overhang invades into the duplex telomeric DNA, sequestering the chromosome end. T-loops unfold in S phase, allowing telomere replication by telomerase.

the G-overhang and a C-rich strand of duplex DNA may preclude t-loop formation in *S. cerevisiae* (Tomaska *et al*., 2004). The precise function of the t-loop *in vivo* is not known, but it appears to play an important role in chromosome end protection by sequestering the G-overhang and the telomeres from deleterious activities (reviewed in Wei and Price, 2003). Notably, although *S. cerevisiae* telomeres do not form a t-loop, a fold-back structure is thought to be formed at the chromosome end (de Bruin *et al*., 2001; Grunstein, 1997), demonstrating the importance of having higher order structure at the chromosome terminus. During S-phase, the t-loop is proposed to unfold, making telomeres accessible for telomere repeat addition by telomerase (Lebel and Wellinger, 2005) (Figure 3).

Telomerase

In the early 1970s, Alexey Olovnikov and James Watson independently predicted that chromosome ends cannot be replicated completely (Olovnikov, 1971; Watson, 1972). The removal of the very 3' RNA primer during lagging strand synthesis would result in the progressive shortening of the chromosome terminus with each cell division (Figure 2). This phenomenon was named the end replication problem. Olovnikov hypothesized that this problem would ultimately lead to the cellular senescence (Olovnikov, 1971). Therefore, a mechanism must exist to overcome the end replication problem in dividing cells.

In 1980's, Jack Szostack and Elizabeth Blackburn initiated a series of experiments in yeast using a linear plasmid that contained *Tetrahymena* telomere sequence at both ends. Remarkably, this linear plasmid could be maintained in *S. cerevisiae*, suggesting that *Tetrahymena* telomeric DNA could function as a telomere in

yeast (Szostak and Blackburn, 1982). Upon propagation, yeast telomere sequences were added onto the ends of *Tetrahymena* telomeric DNA repeats (Shampay *et al*., 1984), indicating that the cells possess a terminal transferase-like activity capable of maintaining telomeres *in vivo*. This activity was subsequently shown to be RNaseA sensitive, implying that the enzyme was a ribonucleoprotein (Greider and Blackburn, 1985; Greider and Blackburn, 1987). Purification of the activity followed by RNA cloning and sequencing confirmed this prediction and showed that the RNA subunit of the enzyme contained a template sequence complementary to the telomere repeat (Greider and Blackburn, 1989). Altogether, these data demonstrated that the telomere terminal transferase is a reverse transcriptase containing its own RNA template which directs the addition of telomeric repeats onto chromosome ends *in vivo*. This enzyme is now known as telomerase.

Telomerase biogenesis, structure and function has been studied extensively in many different model organisms (reviewed in Autexier and Lue, 2006; Collins, 2006). The essential core components of telomerase include an integral telomerase RNA (TER) and a catalytic telomerase reverse transcriptase (TERT) protein component.

TER is transcribed by RNA polymerase II in yeast and vertebrates and by RNA polymerase III in ciliated protozoa. The size of TER varies widely, ranging from 148 nucleotides in ciliates, to 400-600 nt in vertebrates and to ~1300 nt in yeast (Feng *et al*., 1995; Greider and Blackburn, 1989; Singer and Gottschling, 1994). The primary sequence of TER is also highly divergent in different species (reviewed in Chen and Greider, 2004), making bioinformatics approach for identification of telomerase RNA impractical. However, the secondary structure has some common elements in ciliates, yeast and vertebrates (reviewed in Theimer and Feigon, 2006). One conserved feature

of TER is a species-specific ss template region corresponding to ~1.5 telomere repeats complementary to the G-rich telomere strand. In humans, the TER template region contains 11 nt (5' CUAACCCUAAC) complementary to the human telomere repeat sequence (5' TTAGGG) (Feng *et al*., 1995).

The catalytic subunit, TERT, was initially identified in budding yeast in a genetic screen for EST (Ever Shorter Telomere) mutants (Lendvay *et al*., 1996), and subsequently in the ciliate *Euplotes aediculatus* by peptide sequencing of a purified telomerase fraction (Lingner *et al*., 1997). Sequence conservation allowed identification of TERT homologs in a wide variety of other model organisms. TERT proteins harbor a central RT-like domain containing seven universally conserved RT motifs (1, 2, A, B', C, D, E). In addition, TERT also possess a large N-terminal extension (NTE, ~400 amino acids) and a small C-terminal extension (CTE, ~150-200 amino acids), both of which are required for proper enzyme function (reviewed in Autexier and Lue, 2006).

Telomerase is capable of extending telomeres using the TER as a template and the G-overhang in the chromosome terminus as a primer. Extension of telomeric DNA involves several steps (Figure 4). During the first step, the telomerase holoenzyme binds to the ss G-rich telomeric DNA. The 3' end of the G-overhang forms Watson-Crick base-pairs with the template sequence. In the second step, nucleotides are added to the 3' end of the G-overhang in a template-directed manner. The extension occurs through reverse transcription by TERT and continues until the end of the template is reached. In the third step, the RNP translocates, resulting in the repositioning of primer 3' end back at the beginning of the RNA template for another round of nucleotide addition (Autexier and Lue, 2006).

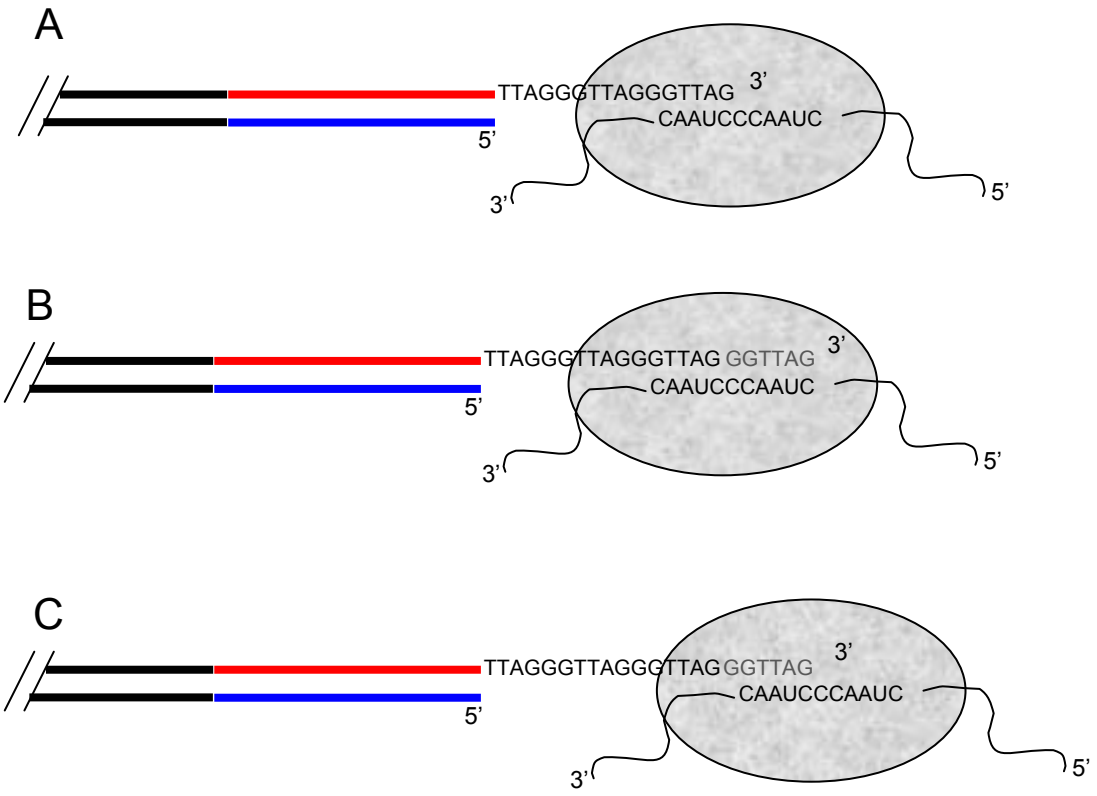


Figure 4. The mechanism of telomere extension by vertebrate telomerase. (A) Binding of the telomerase RNP to the G-overhang. Watson-Crick base pairs are formed between the TER template sequence and the 3' terminus of the G-overhang. (B) Extension of the G-overhang. Six nucleotides that are added are shown in red. (C) Translocation of telomerase. The RNA template is repositioned for the next round of synthesis. TERT is shown as a blue oval; wavy black line represents TER.

Even though TERT and TER expressed *in vitro* are sufficient to reconstitute enzyme activity (Autexier and Greider, 1994; Weinrich *et al*., 1997), the telomerase holoenzyme contains additional components that are needed for its function *in vivo*. Multiple proteins required for telomerase holoenzyme biogenesis and stability have been identified (reviewed in Collins, 2006; Harrington, 2003). For example, in mammals, the H/ACA box of the TER is bound by dyskerin (Dragon *et al*., 2000; Mitchell *et al*., 1999). Mutations in the H/ACA box destabilize TER, leading to telomere shortening (Dragon *et al*., 2000; Lukowiak *et al*., 2001; Mitchell *et al*., 1999). The examples of telomerase-associated proteins interacting with TERT include 14-3-3 protein required for telomerase nuclear localization and chaperons p23 and hsp90 required for efficient telomerase assembly (Forsythe *et al*., 2001; Holt *et al*., 1999; Seimiya *et al*., 2000). Other telomerase-associated components are required for telomerase action at the chromosome terminus, likely mediating the recruitment and/or activation of the enzyme at the telomere. In yeast, these additional components include Est1p and Est3p (Lendvay *et al*., 1996; Lundblad and Szostak, 1989). Both of these proteins are required for telomere extension by telomerase *in vivo*. While the precise role of Est3p is not known, Est1p is necessary for recruitment of telomerase to the telomere end. This function is achieved through direct interaction of Est1p with Cdc13p, a protein that binds G-overhang (Evans and Lundblad, 1999). In addition, it was recently suggested that Est1p can modulate telomerase activity *in vivo*, but the mechanism is not known (Evans and Lundblad, 2002; Taggart *et al*., 2002). Est1 sequence homologs were identified in other organisms, including humans, where one of them (hEst1A) associates with telomerase *in vivo* and causes telomere length alterations and chromosome fusions when over-expressed (Reichenbach *et al*., 2003;

Snow *et al*., 2003). Although there are two putative Est1 orthologs in *Arabidopsis*, neither is important for telomere biology (Riha *et al*., unpublished).

In most organisms, telomerase activity is restricted to highly proliferative tissues and is low in somatic cells (Wright *et al*., 1996). Telomeres in somatic cells therefore shorten with each cell division, which eventually leads to telomere dysfunction, cell cycle arrest and cell death (Harley *et al*., 1990). Telomere length and the rate of telomere shortening therefore determines a Hayflick limit, or the maximal number of times the cell can divide (Hayflick and Moorhead, 1961). In mammals, the inactivation of telomerase in somatic cells and the resulting telomere shortening has been proposed to act as a biological clock, preventing cells from becoming immortalized and therefore preventing cancer formation (Cech, 2004).

In highly proliferative tissues, telomeres are replicated by telomerase. Extension of the 3' end of the G-overhang by telomerase is followed by fill-in synthesis of the C-strand by conventional replication machinery. Even when telomerase is active in the cell, it does not act at each telomere every time the cell divides. In fact, studies from different model organisms demonstrated that while short telomeres are preferentially elongated by the enzyme, telomerase does not extend long telomeres. This was originally shown in mice (Hemann *et al*., 2001), and later confirmed in plants (Shakirov and Shippen, 2004). In these organisms, crosses between individuals with long and short telomeres result in a preferential elongation of short telomeres. In yeast, an assay that allows analysis of the single telomere elongation confirmed the preference of telomerase for short telomeres (Teixeira *et al*., 2004). It was proposed that telomeres switch between telomerase-extendable and non-extendable states, allowing establishment of telomere length homeostasis (Teixeira *et al*., 2004).

Telomere length homeostasis

Although telomere length varies dramatically between different species, a species-specific length set point is set and maintained within a strict range. For example, yeast telomeres are short (~300 bp), while in humans telomeres range from 10 to 15 kb in length (reviewed in Lebel and Wellinger, 2005). *Arabidopsis* telomeres are 2-5 kb long (Richards and Ausubel, 1988), whereas in tobacco telomeres reach 150 kb (Fajkus *et al*., 1995). Telomere length homeostasis is achieved through a competition between multiple forces that shorten and lengthen telomeres (Figure 5A). Critically short telomeres can no longer provide the protective cap for the chromosome terminus, which leads to the initiation of DNA damage checkpoint and cell cycle arrest (reviewed in Riha *et al*., 2006). Grossly elongated telomeres are not tolerated by the cell either. For example, in *K. lactis*, mutations in TER resulting in the dramatic elongation of telomeres impair cell growth (McEachern and Blackburn, 1995). Therefore, it is of critical importance to maintain the equilibrium of telomere shortening and lengthening *in vivo*.

Telomere shortening

As described above, telomeres inevitably shorten with each cell division due to the end replication problem. In addition, at least in mammals, telomeres are also thought to be subjected to a nuclease attack routinely (reviewed in Verdun and Karlseder, 2007). This was proposed since the rate of telomere shortening in dividing cells is actually greater than the rate expected solely from the end replication problem (Counter *et al*., 1992; Harley *et al*., 1990; Levy *et al*., 1992).

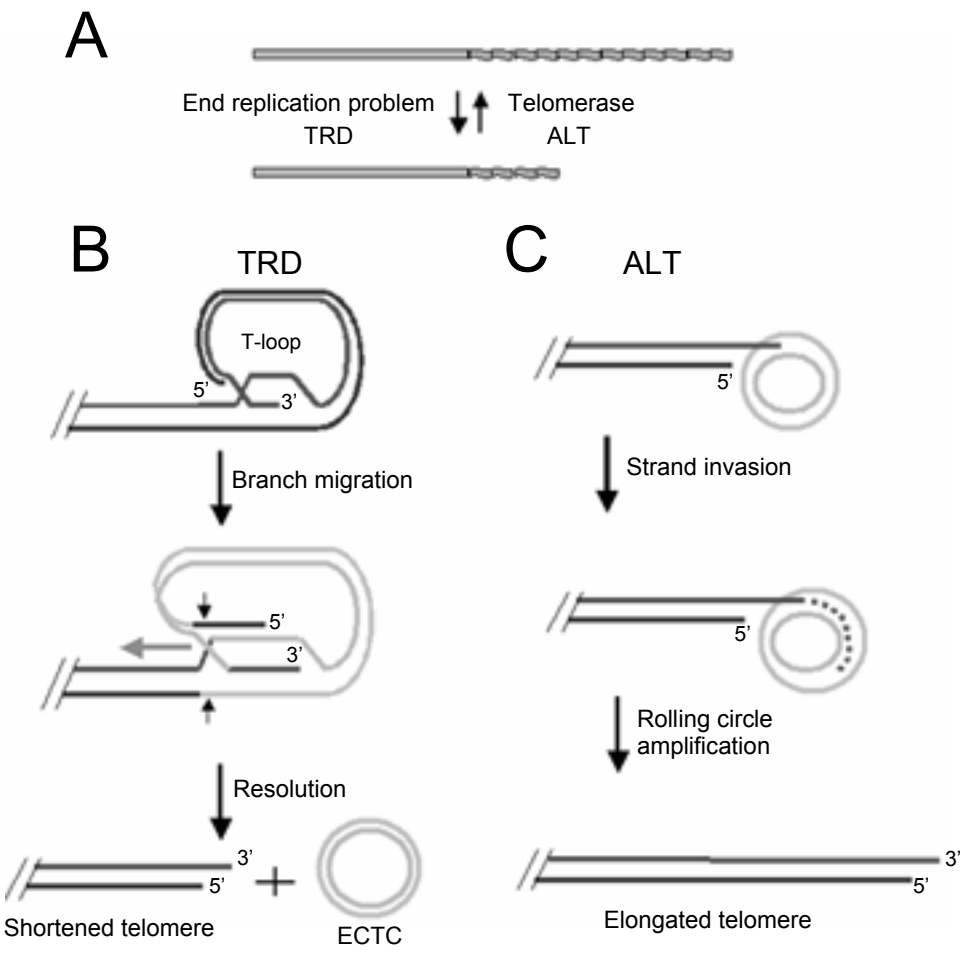


Figure 5. Multiple activities contribute to telomere length homeostasis. (A) Forces that shorten and lengthen telomeres. While extension by telomerase and ALT are the major forces contributing to telomere lengthening, telomeres shorten due to the end replication problem and through TRD. (B) Schematic of TRD. The t-loop can undergo branch migration (green arrow) resulting in the formation of a Holliday junction. Resolution of this structure (black arrows) results in a shortened telomere and an extrachromosomal telomeric circle (ECTC). (C) Schematic of ALT. Telomere elongation in the absence of telomerase via rolling circle amplification is shown.

Another mechanism recently shown to result in a single-step catastrophic loss of telomeric sequences is telomere rapid deletion (TRD). First described in yeast (Kyrion *et al*., 1992; Li and Lustig, 1996), this mechanism has also been recently observed in *Arabidopsis* (Watson and Shippen, 2007). TRD shortens elongated telomeres to a wild-type size (or a size of most of the telomeres in the cell). TRD is a recombination-based process that in yeast is dependent on Rad52 (essential recombination protein) and Mre11/Rad50/XRS2 complex (an important regulator of homologous recombination) (Bucholc *et al*., 2001). Current models propose that TRD is a result of branch migration in the t-loop structure, followed by Holliday junction resolution and t-loop cleavage (Figure 5B) (reviewed in Lustig, 2003). The products of TRD are a shortened telomere and an extra-chromosomal telomeric circle (ECTC). In plants, TRD was shown to function at grossly elongated telomeres as well as telomeres within the wild type size range, with the frequency of TRD decreasing as telomeres shorten (Watson and Shippen, 2007). Importantly, ECTCs, the expected byproduct of TRD, have been found in plants undergoing TRD, suggesting that the TRD mechanism in plants is similar to that of yeast (Zellinger *et al*., 2007). In human cells, one of the mutant alleles of mammalian duplex telomeric DNA binding protein TRF2 induces generation of dramatically shortened telomeres and t-loop-sized telomeric circles (Wang *et al*., 2004). Remarkably, ECTCs were also detected in wild type human cells (Wang *et al*., 2004). Thus, TRD is an evolutionary conserved mechanism contributing to telomere length homeostasis.

Telomerase-independent telomere lengthening

Activities that extend telomeres have also been identified. Telomerase-mediated extension of the G-overhang followed by the C-strand synthesis by conventional replication machinery is the major mechanism to maintain telomeres in dividing cells. However, telomerase-independent alternative telomere lengthening (ALT) pathways for telomere maintenance have been described. In budding yeast, loss of telomerase activity leads to a gradual loss of telomeric DNA, which eventually results in genome instability and cell death. However, survivors arise frequently, and in these cells telomere tracts are maintained by homologous recombination-based pathways (Lundblad and Blackburn, 1993; Teng and Zakian, 1999). Type I survivors utilize a Rad51-dependent mechanism to amplify repetitive subtelomeric regions. Type II survivors rely on Rad50-dependent recombination between telomeres. In the latter case, telomeres use another chromosome end as a template for extension (reviewed in Lundblad, 2002). Rolling circle amplification of ECTC is thought to be another mechanism for type II recombination and for telomerase-independent telomere length maintenance (reviewed in de Lange, 2004).

In human cells, ALT is an efficient way to maintain telomeres in the absence of telomerase. Notably, while ~90% of cancer cells re-activate the telomerase enzyme, the rest utilize ALT to escape telomere dysfunction (reviewed in Bryan *et al*., 1997; Henson *et al*., 2002; Muntoni and Reddel, 2005). Telomeres in ALT cells are extremely long and heterogeneous. Although the exact mechanism of ALT in human cells is unknown, it is thought to include elongation of telomeres through homologous recombination with another telomere, t-loop mediated extension, and rolling circle amplification (reviewed in Henson *et al*., 2002) (Figure 5C). Notably, ECTCs were

detected by 2D gel electrophoresis and by FISH in human ALT cells (Cesare and Griffith, 2004; Hande *et al*., 2001; Ogino *et al*., 1998; Regev *et al*., 1998). These data suggest that ALT and TRD might be mechanistically linked through ECTCs, predicted intermediates for both of these processes. Interestingly, work on the essential telomere capping protein STN1 in *K. lactis* demonstrated that ALT can occur even in the presence of active telomerase (Iyer *et al*., 2005). This report implicated STN1 in the regulation of both ALT and TRD at telomeres, validating the notion that TRD and ALT could act together to maintain telomere length balance.

Telomere-binding proteins

Telomeres are nucleoprotein complexes. Specific non-nucleosomal proteins bind both double-stranded and single-stranded regions of telomeric DNA, and additional proteins associate with the chromosome terminus via protein-protein interactions (Figure 6).

Telomere-binding proteins can be categorized into two classes based on their affinity for either duplex or single-stranded telomeric DNA. The best studied ss telomeric DNA binding proteins include the telomere end binding protein (TEBP) from the ciliate *O. nova* (Gottschling and Zakian, 1986; Price and Cech, 1987), Cdc13 protein from *S. cerevisiae* (Garvik *et al*., 1995; Lendvay *et al*., 1996), and the recently identified Protection Of Telomeres (Pot) proteins from *S. pombe* and humans (SpPot1 and hPot1, respectively) (Baumann and Cech, 2001). Even though sequence similarity is limited between these proteins, they all share a conserved structural domain termed the oligosacharide/ oligonucleotide binding-fold (OB-fold). The OB-fold consists of a five-stranded β barrel and is found in many single-strand nucleic acid binding proteins (Theobald *et al*., 2003). *In vitro,* all of these proteins display specificity for the telomeric

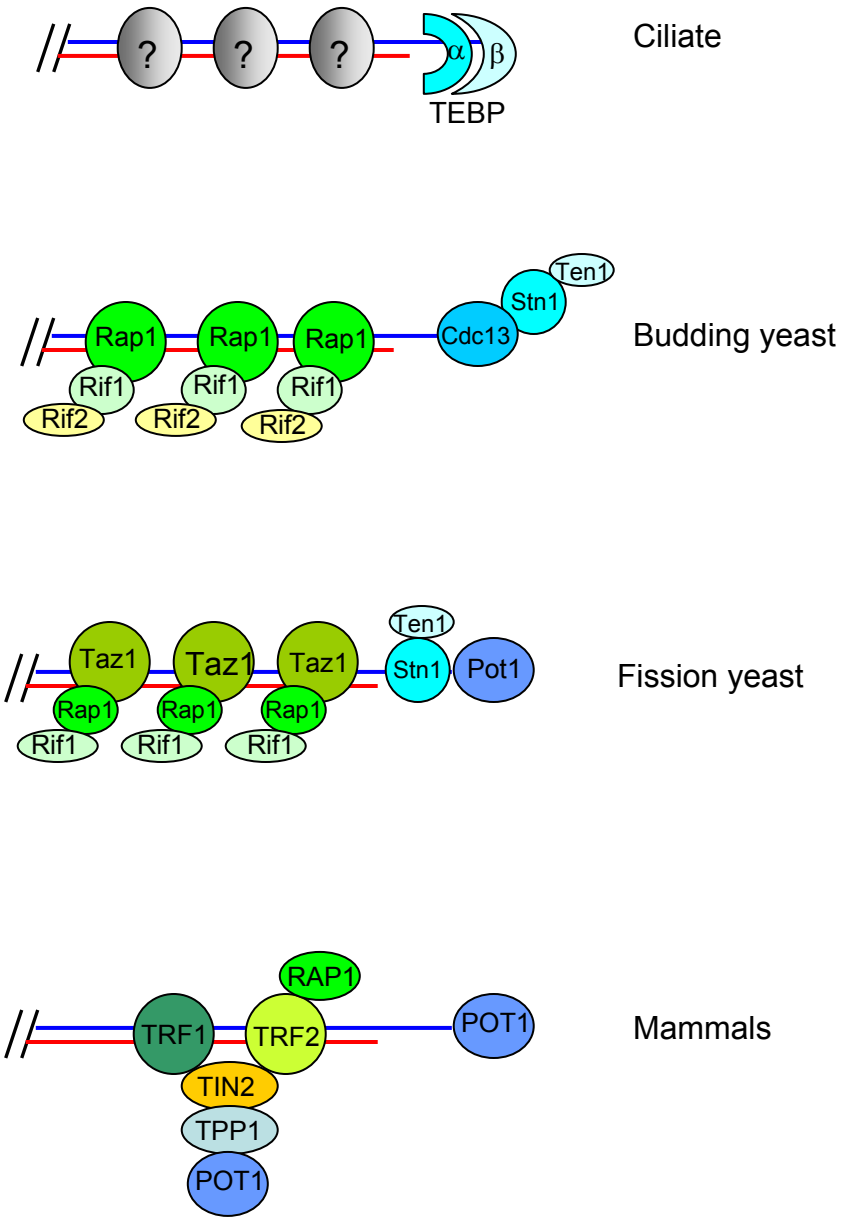


Figure 6. Telomere composition in different organisms. Telomere proteins include ds telomeric DNA binding proteins and their associated factors (shown in shades of green) and ss telomeric DNA binding proteins and their associated factors (shown in shades of blue). Proteins that bind duplex telomeric DNA have not been identified in ciliates yet (gray ovals). In mammals, six core telomere proteins form a protective structure called shelterin.

G-rich strand (Baumann and Cech, 2001a; Lei *et al*., 2002). Double-strand telomeric DNA binding proteins include Taz1 in fission yeast (Cooper *et al*., 1997a), Rap1 in budding yeast (Shore and Nasmyth, 1987), and TRF1 and TRF2 in vertebrates (Bilaud *et al*., 1997; Broccoli *et al*., 1997; Chong *et al*., 1995). These proteins associate with DNA via a Myb-like helix-turn-helix DNA binding motif (Bilaud *et al*., 1996; Broccoli *et al*., 1997; Konig *et al*., 1998; Nishikawa *et al*., 1998; Nishikawa *et al*., 2001). Rap1 possess two Myb motifs. Taz1, TRF1 and TRF2 proteins encode a single Myb domain, however, these proteins homodimerise and this interaction is required for binding to telomeric DNA (Cooper *et al*., 1997; Evans and Lundblad, 2000; Fairall *et al*., 2001; Shore and Nasmyth, 1987). A specific telobox sequence within the Myb domain is conserved in yeast and vertebrates and is thought to be used for telomeric DNA recognition (Bilaud *et al*., 1996).

One of the major functions of telomere-associated proteins is telomere length regulation. As described above, telomere length is strictly regulated. Telomerase action and recombinational mechanisms are the main forces contributing to telomere length changes *in vivo*. Telomere proteins control both of these activities to maintaine telomere length in a wild type species-specific range. It is currently believed that the control of telomere length is achieved via counting of ds telomeric DNA binding proteins. In this model, longer telomeres contain more binding sites for telomere proteins, resulting in negative regulation of telomerase and telomere shortening. This negative feedback mechanism was first discovered in yeast (Marcand *et al*., 1997) and later found in many other model systems (reviewed in Smogorzewska and de Lange, 2004). It is thought that the information on the length of telomeres is transduced from duplex telomeric DNA to the telomere terminus where telomerase acts by G-overhang

binding proteins (Loayza and de Lange, 2003). Ss telomeric DNA binding proteins appear to directly regulate telomerase by affecting its recruitment and/or activation at the chromosome end (reviewed in Wei and Price, 2003).

In addition to telomere length regulation, telomere-associated proteins are essential for chromosome end protection. As described above, ends of chromosomes need to be protected from deleterious nuclease activities, inappropriate recombination, and activation of DNA damage machinery. Protection is achieved by a large array of both double-strand and single-strand telomeric DNA binding proteins. Perturbations in the protein composition or telomere shortening to the level where necessary binding sites for telomeric proteins are lost leads to telomere dysfunction and to the loss of the protective cap at the chromosome terminus (reviewed in de Lange, 2002; de Lange, 2005). The composition of telomeres in different model organisms is described below (Figure 6).

Telomere protein composition in ciliates

Although the ds telomeric DNA binding proteins have not been yet identified in ciliates, the G-overhang binding protein from *O. nova* was the founding member of this class of proteins (Gottschling and Zakian, 1986; Price and Cech, 1987) (Figure 6). The *O. nova* telomere binding protein (TEBP) is composed of two protein subunits, α and β (Hicke *et al*., 1990). The α subunit binds ssDNA via two N-terminal OB-folds and interacts with the OB-fold of the β subunit via a C-terminal OB-fold (Fang and Cech, 1993; Gray *et al*., 1991; Peersen *et al*., 2002). In the presence of telomeric ss DNA, a very stable α-β-ssDNA complex is formed (Fang and Cech, 1993). Although no genetic studies are possible in this organism, biochemical studies and crystal structure data for the TEBP

bound to the ssDNA suggest that TEBP tightly binds the extreme 3' terminus of the *O. nova* G-overhang, thus forming a protective cap at the chromosome end (Horvath *et al*., 1998).

Telomere protein composition in budding yeast

Duplex telomeric DNA in budding yeast is bound by Rap1 (repressor/activator protein 1) (Figure 6). Rap1 was first identified as a transcriptional regulator (Huet *et al*., 1985; Shore and Nasmyth, 1987), and later shown to associate with telomeric DNA (Conrad *et al*., 1990). Analysis of temperature-sensitive *rap1* alleles and over-expression mutants demonstrated that Rap1 is involved in both telomere length regulation and chromosome end protection in *S. cerevisiae* (Conrad *et al*., 1990; Lustig *et al*., 1990). Rap1 negatively regulates telomere length by recruitment of Rif1 and Rif2 (Rap1-interacting factors 1 and 2) proteins to telomeres (Hardy *et al*., 1992; Wotton and Shore, 1997) (Figure 6). Deletion of either Rif1 or Rif2 results in slight telomere elongation, while double mutants display dramatic elongation of telomeric DNA. Elegant studies from David Shore's laboratory demonstrated that an increase in the number of Rap1proteins results in shorter telomeres (Marcand *et al*., 1997). The extent of the telomere shortening was proportional to the number of Rap1 molecules, suggesting that Rap1, Rif1 and Rif2 negatively regulate telomere length in cis.

The G-overhang in *S. cerevisiae* telomeres is bound by Cdc13p (Garvik *et al*., 1995; Lendvay *et al*., 1996; Nugent *et al*., 1996) (Figure 6). Cdc13p is a multifunctional protein essential for proper telomere function and cell viability. Cdc13p plays several important roles at telomeres via dynamic interactions with distinct protein complexes (reviewed in Lustig, 2001). Cdc13p positively regulates telomere length by direct

recruitment of the telomerase to the chromosome end. This is achieved through an interaction of Cdc13p with the Ets1p component of yeast telomerase. Mutant alleles of Cdc13p that abolish the interaction with Est1p lead to an ever shorter telomeres phenotype similar to the telomerase deficiency, despite the presence of biochemically active telomerase (Chandra *et al*., 2001; Evans and Lundblad, 1999; Pennock *et al*., 2001). Cdc13p alleles resulting in telomere lengthening have also been recovered, suggesting that this protein has a separate role in the negative regulation of telomere length (Chandra *et al*., 2001; Grandin *et al*., 2000). Cdc13p also coordinates G-strand and C-strand synthesis by interaction with the catalytic subunit of DNA pol α (Qi and Zakian, 2000). It is now thought that Cdc13p regulates telomere length by first recruiting telomerase to the chromosome terminus and subsequently by limiting telomerase action, thus allowing the C-strand replication (Chandra *et al*., 2001).

In addition to telomere length maintenance, Cdc13p is essential for chromosome end protection. Temperature-sensitive *cdc13* mutants display extensive degradation of the C-strand and cell cycle arrest (Garvik *et al*., 1995). The essential telomere capping function of Cdc13p is likely to be mediated through interactions with the Stn1p (suppressor of cdc thirteen) and Ten1p (telomeric pathways in association with Stn1 number 1) proteins (figure 6). Stn1 was identified in a screen for *cdc13* suppressors (Grandin *et al*., 1997), while Ten1 was subsequently found in the screen for *stn1* suppressors (Grandin *et al*., 2001). As in *cdc13* mutants, *stn1* and *ten1* mutations result in C-strand resection, accumulation of ss G-rich telomeric DNA, and cell cycle arrest (Grandin *et al*., 2001; Grandin *et al*., 1997). Cdc13, Stn1 and Ten1 proteins physically interact with each other, forming a heterotrimeric complex that provides chromosome end protection. Moreover, like Cdc13 (Mitton-Fry *et al*., 2002), both Stn1

and Ten1 proteins harbor an OB-fold and are capable of binding to ss telomeric DNA *in vitro* (Gao *et al*., 2007). Thus, it is proposed that Cdc13-Stn1-Ten1 complex acts as an RPA-like complex that is specific for single-strand telomeric DNA (Gao *et al*., 2007).

Telomere protein composition in fission yeast

In *S. pombe*, Taz1 binds to double-strand telomeric DNA (Figure 6) (Cooper *et al*., 1997). Like Rap1 from *S. cerevisiae*, *S. pombe* Taz1 negatively regulates telomere length; *taz1* mutants display extremely elongated telomeres (Cooper *et al*., 1997). In addition, Taz1 is implicated in chromosome end protection. Loss of Taz1 leads to C strand degradation and increased homologous recombination at telomeres (Cooper *et al*., 1997; Ferreira and Cooper, 2001; Miller and Cooper, 2003). Rap1 and Rif1 homologs have also been identified in *S. pombe* (Figure 6) (Kanoh and Ishikawa, 2001), but in contrast to *S. cerevisiae* Rap1, fission yeast Rap1 does not bind telomeric DNA directly. Instead, both Rap1 and Rif1 are recruited to telomeres through interactions with Taz1. *S. pombe* Rap1 and Rif1 negatively regulate telomere length, likely via the protein counting mechanism (Kanoh and Ishikawa, 2001; Miller *et al*., 2005). It is not known whether the functional homolog of Rif2 exists in fission yeast.

A single-strand telomeric DNA binding protein has only recently been identified in *S. pombe* (Figure 6). Peter Bauman and Tom Cech found spPot1 (Protection of telomeres) based on very limited sequence similarity to TEBP from ciliates (Baumann and Cech, 2001). Genetic studies demonstrated that *S. pombe* Pot1 is essential for chromosome end protection. Knockouts display immediate loss of telomeric and subtelomeric DNA, chromosome mis-segregation, and profound genome instability. Although the null mutation is lethal, survivors that circularized all their chromosomes

arise upon Pot1 loss (Baumann and Cech, 2001). In addition to chromosome end protection, it is possible that *S. pombe* Pot1 regulates telomerase access to telomeres like other G-overhang binding proteins. However, this function may be hidden by the extreme telomere uncapping and end deprotection phenotypes observed in *pot1*$^{-}$ fission yeast. Consistent with this prediction, recent studies implicate *S. pombe* Pot1 in negative regulation of telomere length: reduction of telomere-bound SpPot1 results in dramatic telomere lengthening. However, upon further reduction, loss of telomeric DNA and chromosome fusions are observed, suggesting that cells must carefully regulate the amount of telomere-bound Pot1 (Bunch *et al*., 2005).

Stn1 and Ten1 homologs have recently been identified in S*. pombe* (Figure 6). SpStn1 was discovered in a search for OB-fold containing proteins in the fission yeast genome sequence, while SpTen1 was found in a BLAST search using scTen1 sequence as the query (Martin *et al*., 2007). Both of these proteins co-localize with Pot1 at telomeres. Genetic analysis demonstrates that both SpStn1 and SpTen1 are essential for chromosome end protection. Like POT1 deficiency, loss of Stn1 or Ten1 results in extensive degradation of telomeric and subtelomeric DNA and genome instability. As with deletion of POT1, some yeast cells survive by circularizing their chromosomes (Martin *et al*., 2007). Thus, although the details of deprotection phenotypes differ between *S. cerevisiae* and *S. pombe stn1* and *ten1* mutants, this complex appears to be evolutionary conserved.

Telomere protein composition in mammals

In mammals, there are six core proteins that associate with telomeres to form the protective complex called shelterin (reviewed in de Lange, 2005). Shelterin

components are defined as proteins present at the telomeres throughout the cell cycle that function exclusively at telomeres in contrast to non-shelterin proteins that may transiently associate with telomeres and which primarily function off the telomeres. Shelterin subunits include the double-strand telomeric binding proteins TRF1 and TRF2, the single-strand telomeric DNA binding protein Pot1, and three bridging proteins, TPP1, TIN2 and RAP1, that associate with telomeres via protein-protein interactions (Figure 6).

TRF1 was isolated from HeLa cells based on its *in vitro* binding to ds TTAGGG repeat (Chong *et al*., 1995). TRF1 negatively regulates telomere length: over-expression of TRF1 results in progressive telomere shortening, while a dominant negative allele of TRF1 displaces endogenous TRF1 from telomeres and leads to telomere elongation (van Steensel and de Lange, 1997). TRF2 was identified based on sequence similarity to TRF1, but it plays a different role at telomeres (Bilaud *et al*., 1997; Broccoli *et al*., 1997). The primary function of TRF2 is chromosome end protection and maintaining of the protective cap on the chromosome terminus. TRF2 knockouts are lethal. Therefore, most work with TRF2 has been performed using a dominant-negative TRF2 allele that displaces the wild type protein from telomeres. Loss of telomere-bound TRF2 causes immediate degradation of the G-overhang and end end-to-end chromosome fusions (van Steensel *et al*., 1998). In addition, TRF2 protects telomeres from homologous recombination. Over-expression of the dominant-negative allele of TRF2 results in rapid telomere shortening and the accumulation of extrachromosomal telomeric circles (Wang *et al*., 2004).

TIN2 and Rap1 were identified in a yeast-two-hybrid search for TRF1 and TRF2-interacting proteins, respectively (Kim *et al*., 1999; Li *et al*., 2000) (Figure 6). TIN2 was

later shown to bind to TRF2 as well, forming a link between two duplex DNA –binding proteins (Kim *et al*., 2004). Like TRF1, TIN2 negatively regulates telomere length in a telomerase-dependent manner (Kim *et al*., 1999). Moreover, like TRF2, TIN2 is required for telomere protection from DNA damage response initiation (Kim *et al*., 2004). Since RNAi-mediated depletion of TIN2 results in a decreased presence of TRF2 at the telomeres, the end protection function of TIN2 is thought to depend on its role in the stabilization of TRF2 complex at the chromosome terminus (Ye *et al*., 2004a). Mammalian Rap1 protein, similar to *S. pombe* Rap1, does not bind telomeric DNA directly (Figure 6). However, the telomere function of Rap1 is evolutionary conserved: vertebrate Rap1 negatively regulates telomere length (Li *et al*., 2000; O'Connor *et al*., 2004).

The G-overhang is bound by Pot1 in vertebrates (Figure 6). As with *S. pombe* Pot1, human Pot1 was identified based on limited sequence similarity to TEBP α from ciliates (Baumann and Cech, 2001). The crystal structure of human Pot1 reveals two OB-folds (Lei *et al*., 2004). Interestingly, however, immunolocalization experiments demonstrate that Pot1 protein lacking the OB-fold still localizes to telomeres (Loayza and de Lange, 2003). It was shown that hPot1 localizes to telomeres primarily through its interaction with the duplex telomeric DNA binding protein TRF1 (Loayza and de Lange, 2003). hPot1 binds to TRF1 through an interaction with TPP1 (Liu *et al*., 2004; Ye *et al*., 2004b), and then binds to the G-overhang.

The co-crystal structure of human Pot1 bound to its telomeric DNA substrate also revealed that two OB-folds in Pot1 protein make an extensive contact with the G-overhang and bury the 3’ terminal guanine. This configuration allows the 3’ terminus to be protected from deleterious activities and provides a means to regulate telomerase

access (Lei *et al*., 2004). It was therefore speculated that like other G-overhang binding proteins, human Pot1 must be essential for chromosome end protection. Indeed, RNAi-mediated Pot1 deficiency in human cells results in a chromosome end deprotection phenotype characterized by a DNA damage response, end-to-end fusions, a decrease in the G-overhang signal, and cellular senescence or apoptosis (Churikov *et al*., 2006; Hockemeyer *et al*., 2005; Veldman *et al*., 2004; Yang *et al*., 2005). Interestingly, mice encode two Pot1 related proteins, Pot1a and Pot1b, which display ~75% similarity and appear to be partially redundant for chromosome end protection (He *et al*., 2006; Hockemeyer *et al*., 2006; Wu *et al*., 2006).

In addition to chromosome end protection, over-expression and dominant negative mutations of human Pot1 implicate this protein in both positive and negative regulation of telomere length (Armbruster *et al*., 2004; Colgin *et al*., 2003). Accordingly, data from the crystal structure of hPot1 and several *in vitro* experiments suggest that hPot1 can both recruit and negatively regulate access of telomerase (Kelleher *et al*., 2005; Lei *et al*., 2004; Lei *et al*., 2005).

POT1 functions are influenced by TPP1, another OB-fold containing shelterin component (Figure 6). TPP1 was initially found as a POT1-interacting protein that recruits POT1 to the TIN2/TRF1 complex at telomeres and negatively regulates telomere length (Houghtaling *et al*., 2004; Liu *et al*., 2004; Ye *et al*., 2004b). While POT1 was found based on its similarity to the TEBP α subunit, TPP1 was shown to harbor an OB fold that is structurally similar to the TEBP β subunit (Wang *et al*., 2007; Xin *et al*., 2007). Disruption of POT1-TPP1 interaction resulted in a DNA damage response and telomere length changes (Xin *et al*., 2007). Remarkably, TPP1 was found to associate with active telomerase, and the TPP1-POT1 complex was shown to

increase the activity and processivity of telomerase *in vitro*. The TPP1-POT1 complex appears to both inhibit telomerase access to the telomere and increase telomerase activity during telomere extension (Wang *et al*., 2007; Xin *et al*., 2007).

An extensive search for shelterin-associated factors by gel filtration and co-immunoprecipitation failed to identify additional proteins (Liu *et al*., 2004; O'Connor *et al*., 2004; Ye *et al*., 2004a). However, the possibility exists that some sub-stoichiometric shelterin components were missed by this biochemical approach. In agreement with this, position-specific iterated BLAST (PSI-BLAST) search using the recently characterized *S. pombe* STN1 as a query identified a putative human STN1 homolog (Martin *et al*., 2007). As for yeast STN1, this protein contains an N-terminal OB-fold.

The capping function of telomeres

Perturbations in the protein composition of telomeres often lead to telomere dysfunction and chromosome uncapping (reviewed in Ferreira *et al*., 2004). One cause of telomere dysfunction is a replicative attrition of telomeric DNA in the absence of telomerase. In this case, progressive telomere shortening ultimately triggers telomere dysfunction when the minimal functional telomere length is reached. Telomere dysfunction can also be an immediate outcome of disruption of essential telomere proteins. In this case, even long telomeres can become uncapped.

At the molecular level, telomere dysfunction can have various outcomes (reviewed in de Lange, 2005) (Figure 7). Since uncapped telomeres can no longer protect the chromosome terminus from being recognized by the DNA damage machinery, initiation of DNA damage checkpoints is usually an immediate result of telomere uncapping. DNA damage response proteins accumulate at dysfunctional

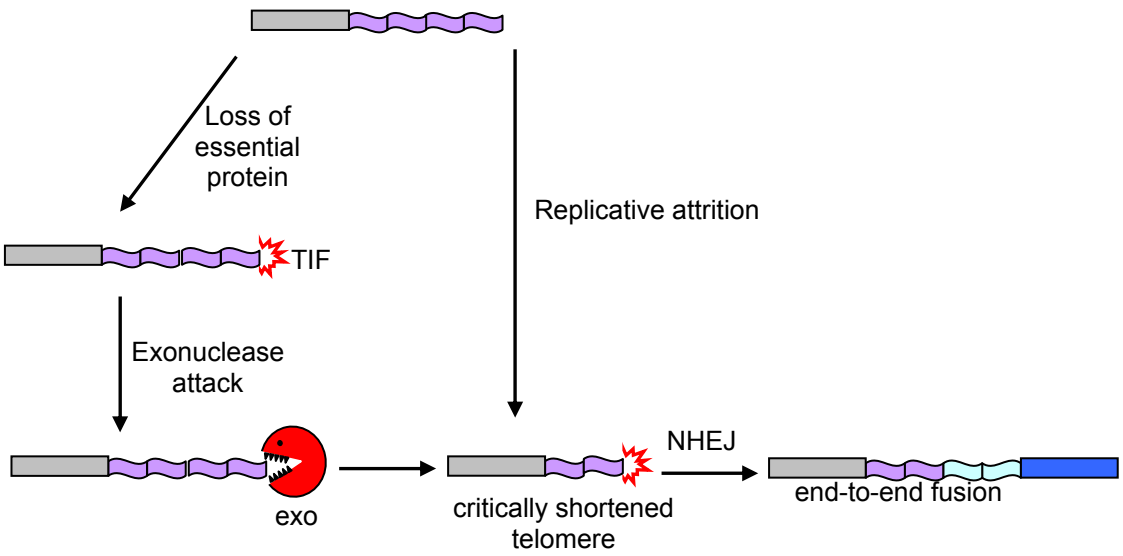


Figure 7. Telomere uncapping. Loss of essential telomeric proteins or telomere shortening due to telomerase inactivation leads to telomere dysfunction. DNA damage repair proteins accumulate at telomeres forming TIFs. Chromosome ends are also subjected to nuclease attack, resulting in the loss of telomeric sequences. Critically shortened telomeres are recruited into end-to-end fusions by NHEJ.

telomeres forming structures known as TIFs (telomere dysfunction-induced foci) (Takai *et al*., 2003). Growth arrest is the likely response to DNA damage arising at uncapped telomeres. Cells that escape the arrest are subjected to nuclease attack and undergo inappropriate repair with telomeres recruited into end-to-end chromosome fusions by nonhomologous end joining (NHEJ) (Figure 7). Moreover, loss of some telomere-associated proteins results in increased levels of recombination.

How telomere-associated proteins mask telomeres from being recognized by DNA damage machinery is still unclear. One model suggests that telomere-associated proteins physically form a protective cap at the chromosome terminus that blocks its ability to engage in repair activities. The best-studied example of such telomere protection mechanism is provided by TEBPα/β heterodimer from ciliates. The two subunits tenaciously associate with each other and with the single-stranded telomeric DNA, sequestering the G-overhang and protecting the chromosome termini from the digestion by Bal31 exonuclease (Gottschling and Zakian, 1986; Price and Cech, 1987). Similarly, the 3' terminal guanine of the G-overhang is buried in the POT1 protein, likely rendering the chromosome terminus inaccessible for various harmful activities (Lei *et al*., 2004).

In addition to an efficient capping by G-overhang binding proteins, the t-loop may make the chromosome terminus inaccessible to various harmful activities. As described above, loss of telomere-bound TRF2 results in immediate telomere uncapping, loss of the G-overhang and telomere fusions (Broccoli *et al*., 1997), and these phenotypes may be caused by t-loop destabilization. TRF2 binds at the junction of double-stranded and single-stranded DNA and can facilitate formation of t-loop like

structures *in vitro* (Fouche *et al*., 2006; Stansel *et al*., 2001). Thus, the integrity of telomere-associated proteins is required for proper telomere structure and function.

Role of DNA damage response proteins at telomeres

Dysfunctional telomeres trigger a cellular response remarkably similar to the response to double-strand breaks. Paradoxically, however, even functional telomeres require DNA damage response proteins for their maintenance. The uncapping phenotype is observed not only when shelterin components are disturbed, but also when DNA damage response (DDR) proteins are inactivated (reviewed in Verdun and Karlseder, 2007). Notably, many DDR proteins localize to telomeres. Examples include the DNA damage signaling kinases ATM and ATR, the double-strand break repair MRN repair complex, and KU, a key component of NHEJ (reviewed in Verdun and Karlseder, 2007). In yeast deletion of both ATM and ATR homologs lead to dramatic and progressive telomere loss and chromosome fusions (Chan *et al*., 2001). In human cells, inactivation of the NBS1 component of the MRN complex or ATM results in accelerated telomere shortening and end-to-end fusions (Ranganathan *et al*., 2001; Vaziri *et al*., 1997). Loss of KU70 in *S. cerevisiae* leads to C strand resection and hyper-recombination at telomeres (Fisher *et al*., 2004; Fisher and Zakian, 2005). Altogether, these data indicate that DDR proteins play a central role in telomere maintenance and chromosome end capping.

Recent data indicate that DDR proteins act at newly replicated telomeres, allowing them to assume a protective state. It has been demonstrated that DNA damage response machinery is recruited to mammalian telomeres in late S/ early G2 of the cell cycle (Verdun *et al*., 2005; Verdun and Karlseder, 2006). The current model

suggests that DNA damage machinery is required for exonucleolytic processing of the C-strand after telomere replication (Chai *et al*., 2006; Larrivee *et al*., 2004; Verdun and Karlseder, 2006). C-strand processing is, in turn, obligatory for the formation of the t-loop and thus the formation of the protective telomere state (Dionne and Wellinger, 1996; Jacob *et al*., 2003; Jacob *et al*., 2001; Wellinger *et al*., 1992; Wellinger *et al*., 1993). It therefore appears that telomere processing by DDR proteins is an essential step for chromosome end protection.

Although DDR proteins are important for telomere maintenance, they must be kept in check. Interestingly, recent data suggest that TRF2 controls ATM, while POT1 independently limits the ATR pathway at functional telomeres (Churikov and Price, 2008; Denchi and de Lange, 2007; Guo *et al*., 2007). It was suggested that TRF2 blocks activation of ATM kinase via direct interaction with ATM (Denchi and de Lange, 2007; Karlseder *et al*., 2004). POT1 is thought to block the ATR signaling pathway by competing with RPA for G-overhang binding (Denchi and de Lange, 2007). As a result, several independent mechanisms provide protection of fully capped telomeres from DNA damage response initiation.

Epigenetic regulation of telomere function

Increasing evidence indicates the importance of the telomeric and subtelomeric chromatin state for telomere function (reviewed in Blasco, 2007). Both telomeric and subtelomeric regions in mammals contain nucleosomes (Makarov *et al*., 1993; Tommerup *et al*., 1994). Telomeric chromatin has been studied extensively using the mouse model system (reviewed in Blasco, 2007). Mammalian telomeres have properties characteristic of heterochromatin; heterochromatin epigenetic marks such as

trimethylation of lysine 9 at histone H3 (H3K9) and lysine 20 of histone H4 (H4K20) are found at telomeres (Garcia-Cao *et al*., 2004). In addition, both H3 and H4 are hypoacetylated at mammalian telomeres (Benetti *et al*., 2007). Although the telomeric DNA cannot be methylated due to lack of CpG sequences, DNA at subtelomeres is methylated (Gonzalo *et al*., 2006). Both maintenance DNA methyltransferase DNMT1 and de novo methyltransferases DNMT3a and DNMT3b are required for subtelomeric DNA methylation (Gonzalo *et al*., 2006).

The relationship between epigenetic chromatin modifications and mammalian telomere maintenance has been investigated in some detail. It was demonstrated that chromatin modifications affect telomere length homeostasis. For example, decreased level of trimethylation of H3K9 at telomeres upon loss of SUV histone methyltransferase results in dramatic telomere elongation (Garcia-Cao *et al*., 2004). Similarly, DNA methylation at subtelomeres negatively regulates telomere length; a decrease in DNA methylation leads to telomere elohgation (Gonzalo *et al*., 2006). In addition, loss of subtelomeric DNA methylation results in increased levels of homologous recombination at telomeres (Gonzalo *et al*., 2006).

Not only is telomere function influenced by chromatin status, but also the telomere length affects the epigenetic status of mammalian telomeres. Progressive telomere shortening in telomerase-deficient mice results in epigenetic changes at telomeres. While heterochromatin-specific histone modifications are decreased, an increase in H3 and H4 acetylation characteristic of active chromatin is observed (Benetti *et al*., 2007). Therefore, while histone and DNA epigenetic modifications are critical for telomere integrity, telomere length affects the epigenetic status of telomeres. Further

studies are needed to elucidate the exact relationship between chromatin-modifying activities and telomere-associated factors.

Plant telomeres

In most plant species, telomeric DNA consists of TTTAGGG repeats. The exceptions include plants from the *Asparagales* clade, whose telomeres contain vertebrate-type TTAGGG repeats (Sykorova *et al*., 2006), and some of the *Allium* (for example, onion), where chromosome termini contain unknown sequences that do not hybridize with either vertebrate or plant type telomere repeats (Pich *et al*., 1996). Telomere length in plants varies dramatically. For example, telomeres in unicellular green alga *Chlorella vulgaris* are only 0.5 kb (Higashiyama *et al*., 1995), while tobacco telomeres are 150 kb in length (Fajkus *et al*., 1995).

As demonstrated for *Arabidopsis thaliana* and *Silene latifolia*, plant telomeres harbor ss G-overhangs (Riha and Shippen, 2003a). Interestingly, in *S. latifolia* G-overhangs were found only on half of telomeres. One explanation is that a portion of telomeres in this plant harbor G-overhangs shorter than the detection limit (Riha *et al*., 2000). Moreover, as in other eukaryotes, plant telomeres form t-loops: these structures were observed in tobacco and in garden pea (Cesare *et al*., 2003; de Lange, 2004). The telomere structure in plants is therefore very similar to the telomere structure in vertebrates. Although different aspects of telomere biology have been investigated in different plant species, the most knowledge about plant telomeres comes from studies of telomere in the flowering plant *Arabidopsis thaliana*.

Arabidopsis as a model for telomere biology

Several laboratories have been developing *Arabidopsis* as a model system to study telomere structure and function. This work is facilitated by sequence analysis of the 125 Mb *Arabidopsis* genome, and by the genetic tractability of this system, which allows one to study loss-of-function and over-expression mutations. Large collections of T-DNA insertion lines, activation-tagged lines and EMS-mutagenized lines as well as protocols for efficient transformation and transgenic studies are available for genetic studies. Moreover, crosses between different mutants can be easily done, allowing for the generation of plants carrying mutations in several genes.

With respect to telomere biology, the *Arabidopsis* model has advantages over other systems. Telomeres are very short (only 2-5 kb) and therefore telomere length can be accurately measured by conventional agarose gel electrophoresis (Shakirov and Shippen, 2004). Additionally, 8 out of 10 subtelomeric sequences in *Arabidopsi*s are unique, making it possible to follow the fate of individual telomeres (Heacock *et al*., 2004). One extremely important feature of *Arabidopsis* is its remarkable tolerance to telomere dysfunction, which allows us to examine fundamental aspects of genome stability (Riha *et al*., 2001). Many mutations in telomere-related proteins that are lethal in mammals are tolerated by plants. The examples include ATM (Vespa *et al*., 2005), ATR (Culligan *et al*., 2004), MRE11(Bundock and Hooykaas, 2002) and RAD50 (Gallego *et al*., 2001) proteins, whose *Arabidopsis* mutants are viable. This greatly facilitates the investigation of the role of such proteins in telomere biology and genome stability. The role of KU, a component of NHEJ repair pathway that binds and stabilizes the ends of ds breaks, has been extensively studied in *Arabidopsis*. In plants, KU is required for telomeric C-strand protection from degradation (Riha and Shippen, 2003a).

In addition, KU is a negative regulator of telomerase: *ku* mutants display dramatic telomerase-dependent telomere elongation (Riha and Shippen, 2003a). Finally, KU regulates homologous recombination and suppresses formation of ECTCs in *Arabidopsis* (Zellinger *et al*., 2007).

Arabidopsis telomerase

Arabidopsis TERT was identified based on its sequence similarity to human telomerase reverse transcriptase (Fitzgerald *et al*., 1999). Although the identification of *Arabidopsis* TER has not been yet reported, our laboratory has recently found a putative telomerase RNA component through a biochemical purification of active telomerase complex (K. Kannan, C. Cifuentes-Rosias and D. Shippen, unpublished). In addition to TERT and TER, dyskerin has been recently shown to associate with telomerase in *Arabidopsis* (Kannan *et al*., 2008).

Similar to vertebrates, telomerase expression is tightly regulated and correlates with cell proliferation in *Arabidopsis*. Telomerase is active in dividing tissues such as flowers, seedlings or cell cultures, and suppressed in leaves and stems (Fitzgerald *et al*., 1999). Our laboratory previously examined the consequences of telomerase deficiency in plants (Fitzgerald *et al*., 1999; Riha *et al*., 2001). Mutants carrying a transfer DNA (T-DNA) insertion in *TERT* were identified and propagated for several generations. Telomere length, G-overhang status and genome stability were monitored in each generation. Terminal restriction fragment analysis (TRF) showed that approximately 200-500 bp of telomeric DNA was lost each generation in *tert* plant mutants. When the shortest telomeres in the population reached ~1 kb in the sixth generation (G6), anaphase bridges started to appear (Riha *et al*., 2001). Anaphase

bridges arise when end-to-end fusions of chromosomes produce dicentric molecules that cannot be efficiently segregated during mitosis. With successive generations, TERT mutants displayed worsening genome instability. In addition, cytogenetic damage correlated with defects in vegetative organs and reproductive tissues. G9 mutants were arrested in a terminal vegetative state (Riha *et al*., 2001).

Composition of *Arabidopsis* telomeres

Although telomeric DNA structure, telomerase function, and the role of several DNA damage repair proteins at *Arabidopsis* telomeres are fairly well characterized, almost nothing is known about the protein composition in *Arabidopsis*. TRF1 or TRF2 homologs have not yet been identified in plants. Although several proteins have been shown to bind ds telomeric DNA *in vitro*, their contribution to telomere maintenance *in vivo* remains unclear (Chen *et al*., 2001; Hwang *et al*., 2001; Marian *et al*., 2003; Schrumpfova *et al*., 2004; Yu *et al*., 2000). In addition, although two putative POT1 proteins have been found based on their sequence similarity to the α subunit of the *O. nova* TEBP (Baumann and Cech, 2001), their role in telomere biology has not been investigated yet. Finally, key components of the shelterin complex in mammals, RAP1, TPP1 or TIN2, can not be discerned in *Arabidopsis* genome, arguing that the sequence of these proteins highly diverged, or that *Arabidopsis* has evolved new proteins to maintain the integrity of telomeres.

Overview

In this dissertation, I explore *Arabidopsis* telomere-associated proteins. In Chapter II, I describe the identification of putative *Arabidopsis* ds telomeric DNA binding proteins. A

bioinformatics approach allowed us to identify twelve Myb-containing TRFL proteins in *Arabidopsis* that fall into two categories based on their sequence and their biochemical properties. We found that TRFL family 1 possess a highly conserved Myb-extension domain absent from the TRFL family 2. We also demonstrated that proteins from TRFL family 1 can bind ds telomeric DNA and form homo- and heterodimers *in vitro*. Finally, we found that the Myb-extension is required for binding plant telomeric DNA *in vitro* (Karamysheva *et al*., 2004).

In Chapters III, IV and V, I describe the characterization of *Arabidopsis* POT proteins. It should be noted that In May 2007, the nomenclature of *Arabidopsis* POT1 proteins changed (Surovtseva *et al*., 2007). AtPOT1 was designated AtPOT1a, while AtPOT2 was designated AtPOT1b. Therefore, while I refer to *Arabidopsis* POT1 proteins as POT1a, POT1b and POT1c throughout the dissertation, in Chapters II-IV I refer to AtPOT1a as AtPOT1 and AtPOT1b as AtPOT2 (since papers reprinted at these chapters were published before May 2007).

In Chapter III, I report the characterization of AtPOT1a and AtPOT1b dominant negative mutants. In transgenic experiments, I show a role for AtPOT1a in telomere length regulation, and a role for AtPOT1b in chromosome end protection (Shakirov *et al*., 2005).

In Chapter IV, I describe the consequences of a AtPOT1a null mutation. Genetic analysis of *pot1a* mutants demonstrated that AtPOT1a is a positive regulator of telomere length. We also showed that AtPOT1a acts in the telomerase pathway for telomere maintenance. Finally, and most unexpectedly, biochemical analysis demonstrated that AtPOT1a is a telomerase component that only transiently associates with *Arabidopsis* telomeres (Surovtseva *et al*., 2007).

In Chapter V, I investigate the role of AtPOT1c. I describe the expression and alternative splicing of *AtPOT1c*. I also demonstrate that over-expression of AtPOT1c leads to deregulation of telomere length. Finally, I show that AtPOT1c works in concert with telomerase in telomere length regulation and defining the architecture of the chromosome terminus.

In Chapter VI, I describe the identification and characterization of CIT1, a novel *Arabidopsis* protein essential for chromosome end protection. I show that, consistent with an essential capping function, *cit1* mutants display severe telomere dysfunction, loss of telomeric sequences, end-to-end fusions and genome instability. I also demonstrate that CIT1 is conserved in plants and vertebrates, and discuss potential models for the role of this protein at telomeres.

Finally, in Chapter VII, I discuss the contribution of our research into the understanding of telomere biology in plants and other eukaryotes. I also present a model for the function of *Arabidopsis* telomere components, and propose future experiments that will provide new insights into *Arabidopsis* telomere structure and function.

CHAPTER II

A C-TERMINAL MYB-EXTENSION DOMAIN DEFINES A NOVEL FAMILY OF DOUBLE-STRAND TELOMERIC DNA BINDING PROTEIN IN *Arabidopsis**

Summary

Little is known about the protein composition of plant telomeres. We queried the *Arabidopsis thaliana* genome database in search of genes with similarity to the human telomere proteins hTRF1 and hTRF2. hTRF1/hTRF2 are distinguished by the presence of a single Myb-like domain in their C-terminus that is required for telomeric DNA binding *in vitro.* Twelve *Arabidopsis* genes fitting this criterion, dubbed TRF-like (TRFL), fell into two distinct gene families. Notably, TRFL family 1 possessed a highly conserved region C-terminal to the Myb domain called Myb-extension (Myb-ext) that is absent in TRFL family 2 and hTRF1/hTRF2. Immunoprecipitation experiments revealed that recombinant proteins from TRFL family 1, but not those from family 2, formed homo- and heterodimers *in vitro*. DNA binding studies with isolated C-terminal fragments from TRFL family 1 proteins, but not family 2, showed specific binding to double-stranded plant telomeric DNA *in vitro*. Removal of the Myb-ext domain from TRFL1, a family 1 member, abolished DNA binding. However, when the Myb-ext

---

*Reprinted with permission from 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.

domain was introduced into the corresponding region in TRFL3, a family 2 member, telomeric DNA binding was observed. Thus, Myb-ext is required for binding plant telomeric DNA and defines a novel class of proteins in *Arabidopsis*.

Introduction

Telomeres are the specialized nucleoprotein structures that comprise the natural ends of linear eukaryotic chromosomes and ensure their complete replication and stability (Blackburn, 1991; Zakian, 1996). In most eukaryotes, telomeric DNA is composed of tandem arrays of simple G-rich repeat sequences terminating in single-strand 3' overhang, which is maintained through the action of the telomerase reverse transcriptase (Blackburn, 1991). Both the double- and single-strand region of the telomere are coated by non-histone proteins that provide protection for telomeric DNA and regulate telomerase access to the chromosome terminus. Proteins that bind double-strand telomeric DNA are typified in vertebrates by TRF1 and TRF2, and in budding and fission yeast by Rap1 and Taz1, respectively (Bilaud *et al*., 1997; Broccoli *et al*., 1997; Chong *et al*., 1995; Cooper *et al*., 1997; De Rycker *et al*., 2003; Shore and Nasmyth, 1987).

Human TRF1 behaves as a negative regulator of telomere length; over-expression results in telomere shortening, while a dominant negative allele induces telomere elongation (Smogorzewska *et al*., 2000; van Steensel and de Lange, 1997). hTRF1 mediates telomere length control through interactions with other telomere-associated factors including tankyrase (Smith *et al*., 1998), TIN2 (Kim *et al*., 1999), PinX1 (Zhou and Lu, 2001), and Pot1 (Loayza and de Lange, 2003). Although hTRF2 contributes to telomere length regulation (Smogorzewska *et al*., 2000), its major

function is to conceal telomere ends from detection as double-strand breaks (Karlseder *et al*., 1999; van Steensel *et al*., 1998). Inhibition of hTRF2 in cultured human cells results in loss of the 3' overhang and formation of covalently fused telomeres (van Steensel *et al*., 1998). In addition, compromised hTRF2 function culminates in cell cycle arrest and ATM/p53-mediated apoptosis (Karlseder *et al*., 1999).

The functional domains of vertebrate TRF1 and TRF2 have been studied in some detail (Rhodes *et al*., 2002). The two proteins have similar molecular masses (50-60kD), and resemble each other in domain structure. Although the N-terminus is highly acidic in hTRF1 and highly basic in hTRF2, both proteins harbor a centrally located flexible hinge region called the TRF homology (TRFH) domain that is required for homodimer formation and interactions with other telomere-associated proteins (Fairall *et al*., 2001). The most strongly conserved feature is a Myb/homeodomain type helix-turn-helix DNA binding motif near the C-terminus (Bilaud *et al*., 1996; Broccoli *et al*., 1997; Konig *et al*., 1996; Nishikawa *et al*., 1998; Nishikawa *et al*., 2001). The Myb domain is sufficient for specific binding to the telomeric DNA *in vitro,* but dimerization through the TRFH domain is required for TRF protein association with DNA *in vivo* (Fairall *et al*., 2001). Biluad *et al*. (1996) noted a sequence within the Myb domain of hTRF1 and hTRF2, VDLKDKWRT, that is also conserved in yeast double strand telomere binding proteins and dubbed it the telobox consensus. An NMR structure of the Myb motif from hTRF1 revealed specific contacts between amino acid residues within the telobox consensus and the human telomere repeat sequence TTAGGG (Nishikawa *et al*., 2001).

Although the plant telomere repeat sequence, TTTAGGG, is closely related to that of humans, almost nothing is known about the protein composition of plant

telomeres. Several proteins have been shown to bind double-stranded telomeric DNA *in vitro* (Chen *et al*., 2001; Hwang *et al*., 2001; Schrumpfova *et al*., 2004; Yang *et al*., 2003; Yu *et al*., 2000; Zentgraf, 1995; Zentgraf *et al*., 2000). From *Arabidopsis* these include several relatively small (~30kD) proteins with N-terminal Myb domains (Marian *et al*., 2003; Schrumpfova *et al*., 2004). Two other telomeric DNA binding proteins, TRP1 and TBP1, more closely resemble vertebrate TRF1 and TRF2 in size (65kD and 70kD, respectively) and in architecture as they harbor a single Myb domain in their C-terminus. *TRP1* from *Arabidopsis* was identified in a yeast one-hybrid screen for proteins that bind double-strand telomeric DNA (Chen *et al*., 2001). Like vertebrate TRF1 and TRF2, full-length TRP1 shows a strong *in vitro* preference for extended telomeric DNA tracts with a minimum binding site of five TTTAGGG repeats. Another *Arabidopsis* gene that harbors a single Myb domain at its C-terminus is *TBP1*. *TBP1* encodes a homolog of the rice RTBP1, which has been shown to specifically bind plant telomeric DNA *in vitro* (Hwang *et al*., 2001; Yu *et al*., 2000).

In this study we employed a BLAST search to identify *Arabidopsis* homologs of hTRF1 and hTRF2 using their Myb domains as the query. Although *Arabidopsis* harbors more than 100 genes with Myb domains (Kranz *et al*., 1998), we found only 12 with a single Myb domain in their C-terminus that contains the telobox consensus motif. We designated this group of genes TRF-like (TRFL). Here we provide a molecular characterization of the TRFL proteins. Our data reveal that the TRFL genes encode two distinct families of proteins that differ dramatically in their amino acid sequences, DNA binding properties and protein interactions. Furthermore, we define a novel functional domain C-terminal to the Myb domain that is required for specific binding to duplex plant telomeric DNA.

Materials and methods

Computer search for Myb-containing genes and phylogenetic analysis

The Myb-domains of hTRF1, hTRF2, and *Arabidopsis* TRP1 were used in separate NCBI Blast searches to identify *Arabidopsis thaliana* genes predicted to encode proteins with a single Myb repeat at the C-terminus. The telobox consensus motif was used as an additional criterion (Bilaud *et al*., 1996). Twelve TRFL genes were identified. Multiple protein alignments were conducted using Oxford Molecular Group's sequence analysis software MacVector 7.0 (Accelrys, San Diego, CA). A phylogenetic tree was constructed using the neighbor-joining method (Saitou and Nei, 1987) with bootstrap mode.

Expression analysis, molecular cloning and production of recombinant proteins in vitro RT-PCR was performed for 35 cycles using SuperScript III reverse transcriptase (Sigma) to determine whether TRFL genes were expressed in different *Arabidopsis* tissues. For most of the TRFL genes, primers flanked intron junctions, ruling out the possibility of genomic DNA contamination. Full-length cDNAs were obtained by RT-PCR using total RNA from flowers. Sequence analysis of the cloned cDNAs verified the NCBI annotations. For protein interaction studies, the full-length cDNAs of TRFL genes were subcloned into pET-28A and pCITE 4A. For DNA binding studies, PCR was used to amplify either full-length coding regions or the Myb domains and adjacent C-terminal region using plasmids that contained full-length cDNAs. Overlapping PCR (Ho *et al*., 1989) was applied to create the fusion constructs. A T7 phage promoter was introduced into all the constructs by PCR and the products were used as templates for *in vitro* transcription by T7 RNA polymerase (Stratagene). Transcripts were translated

in either rabbit reticulocyte lysate (Promega) or a wheat germ translation system (Promega) in the presence or absence of ^{35}S- methionine (Amersham). An aliquot of the labeled protein was used to verify the expected apparent molecular weight of translated protein products by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and autoradiography.

Electrophoresis mobility shift assay (EMSA)

Telomeric DNA probes and competitors were as follows: *Arabidopsis* AGCATGCAGC$(TTTAGGG)_8$, AGCATGCAGC$(TTTAGGG)_6$, AGCATGCAGC$(TTTAGGG)_4$, AGCATGCAGC$(TTTAGGG)_2$, *Paramecium* AGCATGCAGC$(TTTGGG)_8$, human AGCATGCAGC$(TTAGGG)_8$. Double-strand DNA probes were obtained by annealing the corresponding complementary oligonucleotides. For gel shift assays, peptides consisting of the Myb domain and the remainder of the C-terminus (see above) were tested for DNA binding activity. Preparation of the probes and binding reactions were carried out as described previously (Hall and Milner, 1997). Protein-DNA complexes were loaded on a 4% PAGE in 0.5xTBE and subjected to electrophoresis for 4 h at 120 V. Dried gels were exposed for autoradiography.

Co-immunoprecipitation

For each TRFL protein analyzed, two constructs were made: one with a T7 protein tag and one without. ^{35}S- methionine labeled non-tagged proteins or T7 tagged unlabeled proteins were synthesized in a TNT-coupled rabbit reticulocyte lysate translation system following the manufacturer's recommendations (Promega). Translation of T7 tagged proteins was verified in the presence of ^{35}S- methionine on a small aliquot from the

same master mix. T7 tagged and untagged radiolabeled proteins were combined and subjected to immunoprecipitation using agarose beads (Novagen) containing T7 monoclonal antibody as described (Bryan *et al*., 2000). Precipitate and supernatant fractions were analyzed by SDS-PAGE and autoradiography.

Results

Identification and expression of TRFL genes in Arabidopsis

We performed a BLAST search to identify TRF-like (TRFL) genes in *Arabidopsis thaliana*. Three query sequences were employed consisting of the Myb domains from human TRF1 and TRF2 and from *Arabidopsis* TRP1. Twelve genes were uncovered that encode proteins with a single Myb domain at their C-terminus (Figure 8, Table 1). As expected, the searches found *TRP1* and *TBP1* (Ho *et al*., 1989; Hwang *et al*., 2001), along with ten uncharacterized genes that were designated *TRFL1-10*. With the hTRF1 Myb domain as the query, the most closely related sequence was the Myb domain for *TRFL6* (E-value=$3e^{-05}$), while for hTRF2, the Myb domain of *TRFL3* displayed the most similarity (E-value=$3e^{-06}$).

A conceptual protein alignment of the C-terminal Myb-containing region of TRFL proteins with the corresponding regions in hTRF1 and hTRF2 is shown in Figure 8. The NMR structure of the Myb domain of hTRF1 bound to the human telomeric DNA repeat sequence reveals that the N-terminal arm of the first helix interacts with the TT sequence in the minor groove, while the third helix recognizes TAGGG in the major groove (Nishikawa *et al*., 2001). The corresponding regions are well conserved among the *Arabidopsis* TRFL proteins and display extensive sequence similarity to

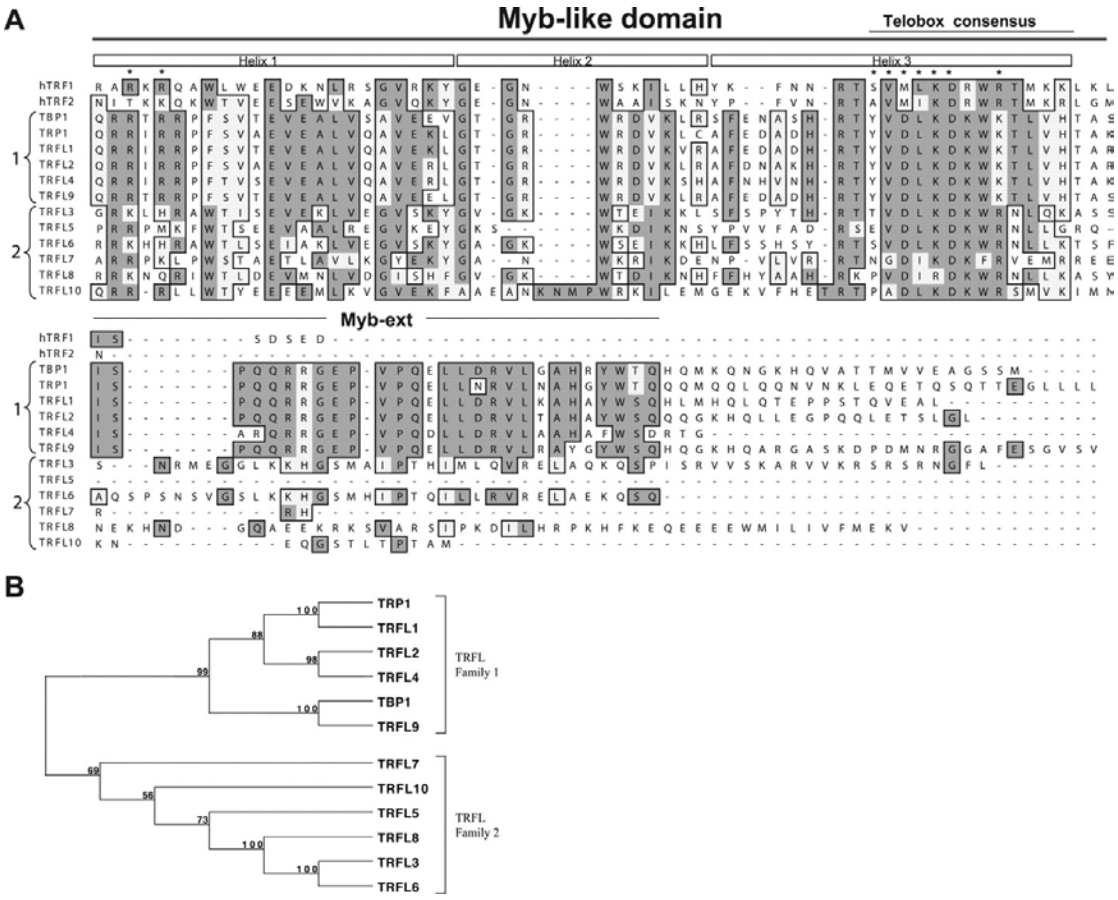


Figure 8. Identification of two TRFL gene families in *Arabidopsis*. (A) Comparison of the C-terminal region of TRFL proteins with the corresponding region of hTRF1 and hTRF2. Amino acid sequences were obtained by conceptual translation of corresponding cDNAs. Sequences used for alignment include Myb domains and adjacent C-terminal region. Identical residues are shown in dark gray and similar residues in light gray. The likely positions of the three helices, which comprise the Myb domain of hTRF1, including the telobox consensus motif, are shown in boxes above the hTRF1 sequence. Asterisks indicate amino acid residues in hTRF1 that make specific contacts with human telomeric repeat DNA in the NMR structure (Nishikawa *et al*., 2001). Brackets denote TRFL family 1 and 2 members. The Myb-ext domain in TRFL family 1 proteins is shown. (B) Phylogenetic analysis of TRFL proteins from *Arabidopsis*. The tree was constructed by the neighbor-joining method after sequence alignment of the two TRFL protein families. The number at each node indicates the percentage of 10,000 bootstrap replicates for statistical support.

Table 1. DNA binding properties and protein interactions of TRFL proteins

Locus	Accession number	Telomere protein	Regions expressed for DNA binding	Min. telo repeats bound	Homodimer
At5g13820	NP_196886	TBP1	ND	2[1]	ND
At5g59430	NP_200751	TRP1	ND	4[2]	Yes
At3g46590	NP_190243	TRFL1	449-553	6	Yes
At1g07540	NP_172234	TRFL2	516-622	6	Yes
At3g53790	NP_190947	TRFL4	311-400	4	Yes
At3g12560	NP_187862	TRFL9	505-620	4	Yes
At1g17460	NP_564025	TRFL3	483-605	No binding	No
At1g15720	NP_173024	TRFL5	303-390; 334-390	No binding	No
At1g72650	NP_565045	TRFL6	530-624	No binding	No
At1g06910	NP_683280	TRFL7	331-390	No binding	ND
At2g37025	NP_671929	TRFL8	297-404	No binding	ND
At5g03780	NP_195998	TRFL10	344-420	No binding	ND

[1]Hwang *et al*., 2001
[2]Chen *et al*., 2001
Min. telo, minimum telomere
ND=not determined.

hTRF1 and hTRF2 (Figure 8A). The most highly conserved region within all of the TRFL proteins, VDLKDKWRT, lies within the telobox consensus.

Phylogenetic analysis indicated the presence of two distinct TRFL gene families (Figure 8B). TRFL family 1 includes *TBP1*, *TRP1*, *TRFL1*, *TRFL2*, *TRFL4*, and *TRFL9*, while TRFL family 2 contains *TRFL3*, *TRFL5-8* and *TRFL10.* The similarity within TRFL family 1 or family 2 members extends throughout their entire sequence, suggesting the DNA binding domains co-evolved with the remainder of the genes. Phylogenetic analysis further suggests that all of the TRFL genes arose from duplication events (Figure 8B). Four of the TRFL genes reside in regions of the *Arabidopsis* genome known to be duplicated. These include *TRP1* and *TRFL1,* which display 59% identity; 70% similarity overall and reside on chromosomes 5 and 3, respectively, and *TRFL3* and *TRFL6,* which exhibit 52% identity; 62% similarity and are found in a duplicated region on chromosome 1. Although *TBP1* and *TRFL9* are not located in a duplicated region, they display a remarkably high degree of similarity (54% identity; 66% similarity) and hence may have been derived from a recent gene duplication.

Closer inspection of the predicted amino acid sequence for the TRFL proteins provided further evidence for distinct gene families. In all of the TRFL family 1 proteins, the region of amino acid conservation in the Myb domain extends further into the C-terminus, creating a Myb-extension (Myb-ext) domain (Figure 8A). Interestingly, Myb-ext was absent from TRFL family 2 and from hTRF1 and hTRF2, which terminate immediately adjacent to the Myb domain. An additional region of substantial sequence conservation was detected in the central portion of TRFL family 1 members called the central domain (CD) (Figure 9A and B). This region bears no significant similarity to the

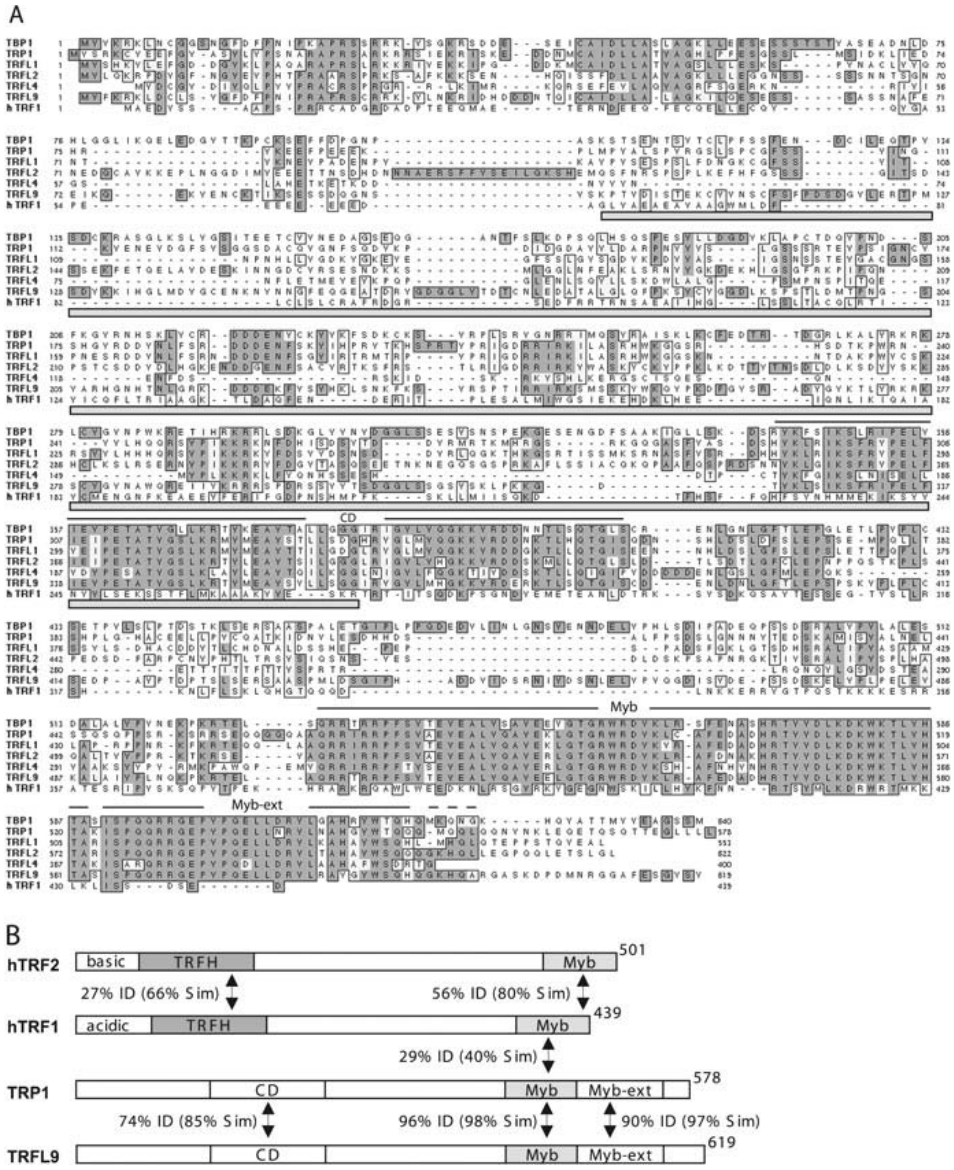


Figure 9. Sequence alignment and domain structure of TRFL family 1 members. (A) Alignment of full-length TRFL family 1 proteins with hTRF1. The CD, Myb and Myb-ext domains are indicated by an over-line. The TRFH domain in hTRF1 is denoted by a wide gray box under the hTRF1 sequence. (B) Comparison of the domain structures of hTRF1 and TRF2 with TRP1 and TRFL9.

TRFH domain in vertebrate TRF1/TRF2 proteins and is also absent in TRFL family 2. Aside from TRFL3 and TRFL6, which clearly represent a recent gene duplication, the remaining members of TRFL family 2 display no obvious sequence similarity outside their Myb domains.

RT-PCR analysis revealed that all of TRFL genes are expressed in *Arabidopsis* (Figure 10). Transcripts for each of the TRFL genes could be detected in all of the organs we examined (Figure 10; data not shown). Moreover, all of the genes were expressed at a relatively high level, with the exception of *TRFL2* and *TRFL4*, which are both members of TRFL family 1. While *TRFL2* and *TRFL4* were ubiquitously expressed, their transcripts were scarce and could only be observed by nested PCR (Figure 10). Nevertheless, the constitutive expression of TRFL genes is consistent with a structural role in the telomere complex.

Proteins in TRFL family 1 can form homo- and heterodimers

Human TRF1 and TRF2 bind telomeric DNA as homodimers *in vivo* and *in vitro* (Bianchi *et al*., 1997; Fairall *et al*., 2001). Therefore, we tested whether recombinant TRFL proteins could form homodimers *in vitro* using co-immunoprecipitation experiments. For these studies, untagged ^{35}S-methionine labeled TRFL proteins were subjected to immunoprecipitation in the presence of the corresponding unlabeled T7-tagged protein using a T7 antibody (Figure 11A and B). Following immunoprecipitation, a homomeric interaction with tagged protein will allow a radiolabeled, but untagged protein to precipitate with the beads. Control reactions performed in the absence of tagged proteins showed no interaction between the T7 antibody and untagged proteins (Figure 11A and B, lane 2 in all panels). However, specific homomeric protein interactions were

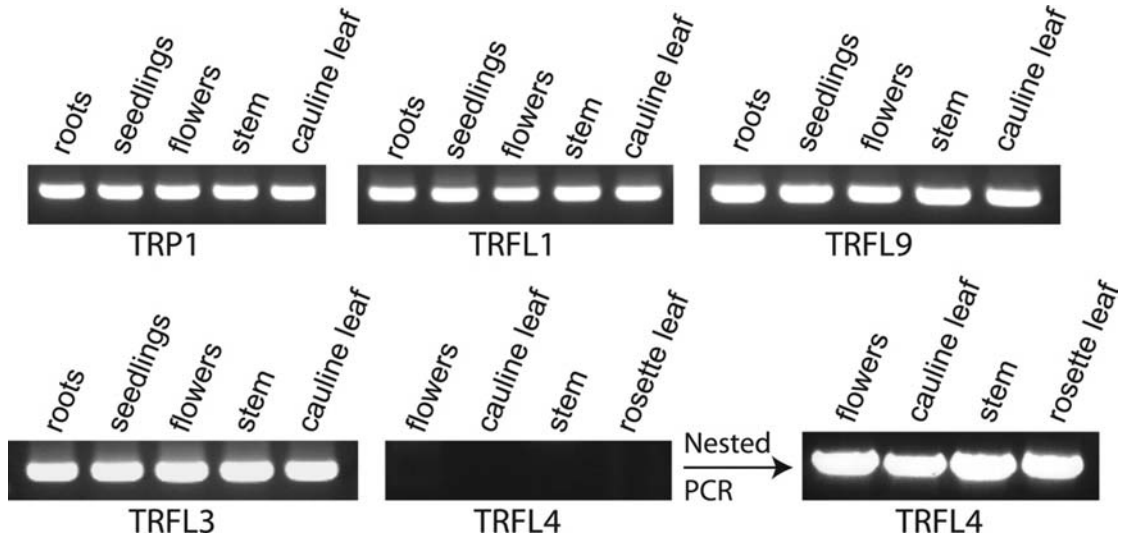


Figure 10. Expression of TRFL genes in *Arabidopsis*. RT-PCR analysis was performed on RNA isolated from the sources indicated. All of the TRFL genes analyzed were found to be expressed in all of the organs examined. Results from *TRP1*, *TRFL 1, 3, 4* and *9* are shown as examples. Transcripts from *TRFL4* were detected only after nested PCR.

detected for TRP1, TRFL1, TRFL2, TRFL4 and TRFL9 (Figure 11A, upper panels; Table 1). All of these proteins are members of TRFL family 1. In contrast, TRFL3, TRFL5 and TRFL6, members of TRFL family 2, did not exhibit the capacity for homodimerization *in vitro* (Figure 11A, lower panels; Table 1).

Despite significant amino acid similarity in the TRFH domains of hTRF1 and hTRF2, steric hinderence prevents the formation of heterodimers (Fairall *et al*., 2001). To determine whether TRFL proteins have the capacity to heterodimerize *in vitro* additional co-immunoprecipitation experiments were performed with family 1 members TRP1, TRFL1 and TRFL9 (Figure 11B). These data showed that TRP1 and TRFL1 can heterodimerize, as can TRP1 and TRFL9. Interestingly, although TRP1 is closely related to TRFL1, it bears only limited sequence similarity to TRFL9 (38.2% identity, 46.3% similarity). Thus, it is conceivable that functionally distinct TRFL proteins directly interact *in vivo*.

DNA binding properties of TRFL proteins

We used electrophoretic mobility shift assays (EMSA) to ask whether TRFL proteins bind telomeric DNA *in vitro*. Initially, several full-length recombinant proteins expressed in rabbit reticulocyte lysate were tested for DNA binding. As expected from a previous study, full-length TRP1 bound a duplex telomeric DNA probe consisting of eight TTTAGGG repeats (AtTR8) and formed a discrete protein-DNA complex that migrated into the gel as well as complexes that stayed in the well (data not shown). Assays with other full-length TRFL proteins, including TRFL4 (Figure 12A), produced complexes that failed to exit the well. However, such complexes were sequence-specific (Figure 12A, left panel) and displayed length dependence in DNA binding (Figure 12A, right panel).

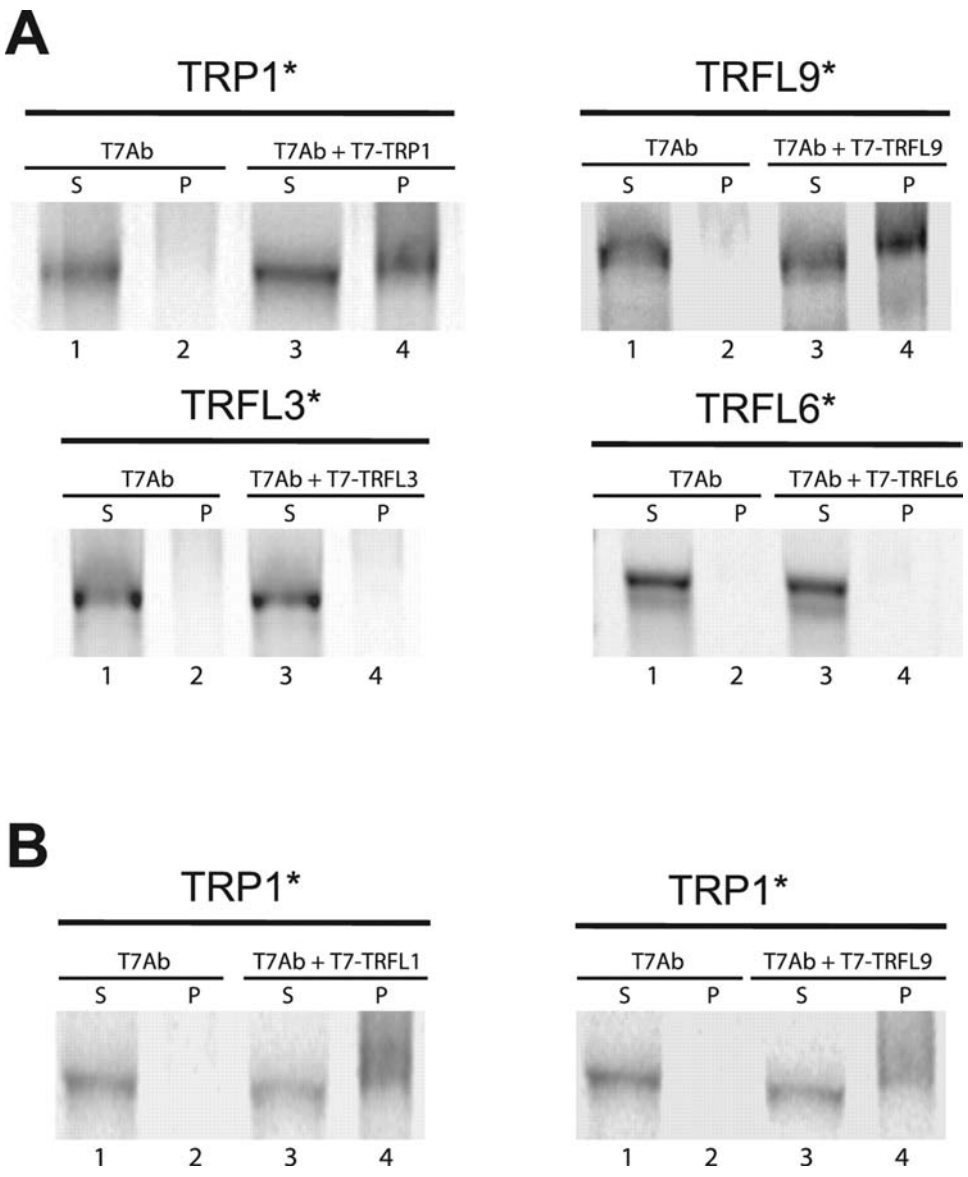


Figure 11. Dimerization of TRFL proteins *in vitro*. Co-immunoprecipitation experiments were conducted with full-length TRFL proteins expressed in rabbit reticulocyte lysate. In each experiment, one protein was labeled with ^{35}S-Met (asterisk) and the other was unlabeled, but contained a T7-tag. Control reactions performed in the absence of T7-tagged proteins showed no non-specific binding of ^{35}S-Met labeled protein to the beads. Examples of co-immunoprecipitation assay results for homodimeric (A) and heterodimeric (B) interactions are shown.

To alleviate the concern of protein aggregation and to further analyze protein-DNA interactions, we tested whether the isolated Myb domains from TRFL proteins would bind telomeric DNA in EMSA. Previous studies showed that the corresponding regions of hTRF1 and TRP1 and TBP1 were sufficient for specific interaction with duplex telomeric DNA *in vitro* (Chen *et al*., 2001; Hwang *et al*., 2001; Konig *et al*., 1998). We expressed the corresponding region of TRFL proteins (Myb domain and the adjacent C-terminal region) in wheat germ extract, which we found yielded a higher amount of recombinant TRFL protein than rabbit reticulocyte lysate. A summary of DNA binding results for the truncated TRFL proteins is shown in Table 1.

All members of TRFL family 1 bound specifically to plant telomeric DNA (Table 1; Figure 12A and B). Surprisingly, the minimal number of repeats required for binding varied among the different proteins. For example, TRFL1 $^{449\text{-}553}$ required a minimum of six telomere repeats for optimal binding (Figure 12B, lane 6). Although a weak interaction was detected with probes consisting of four telomere repeats (Figure 12B, lane 4), the complex appeared to be unstable as no discrete band was observed. In contrast, binding assays with TRFL9$^{505\text{-}620}$ yielded a more discrete DNA-protein complex with a probe containing four telomeric repeats (Figure 12C, lane 4; data not shown). Competition experiments revealed that the association of TRFL family 1 proteins with plant telomeric DNA is specific (Figure 12B, lanes 8-12 and Figure 12C, lanes 10-13; data not shown). Competition was observed with probes consisting of double-stranded plant telomeric DNA, but not with non-specific duplex DNA or duplex DNA comprised of human or *Paramecium* telomeric repeats (Figure 12A-C; data not shown). Assays with single-stranded telomeric DNA competitor also failed to compete with duplex plant telomeric DNA for binding (data not shown).

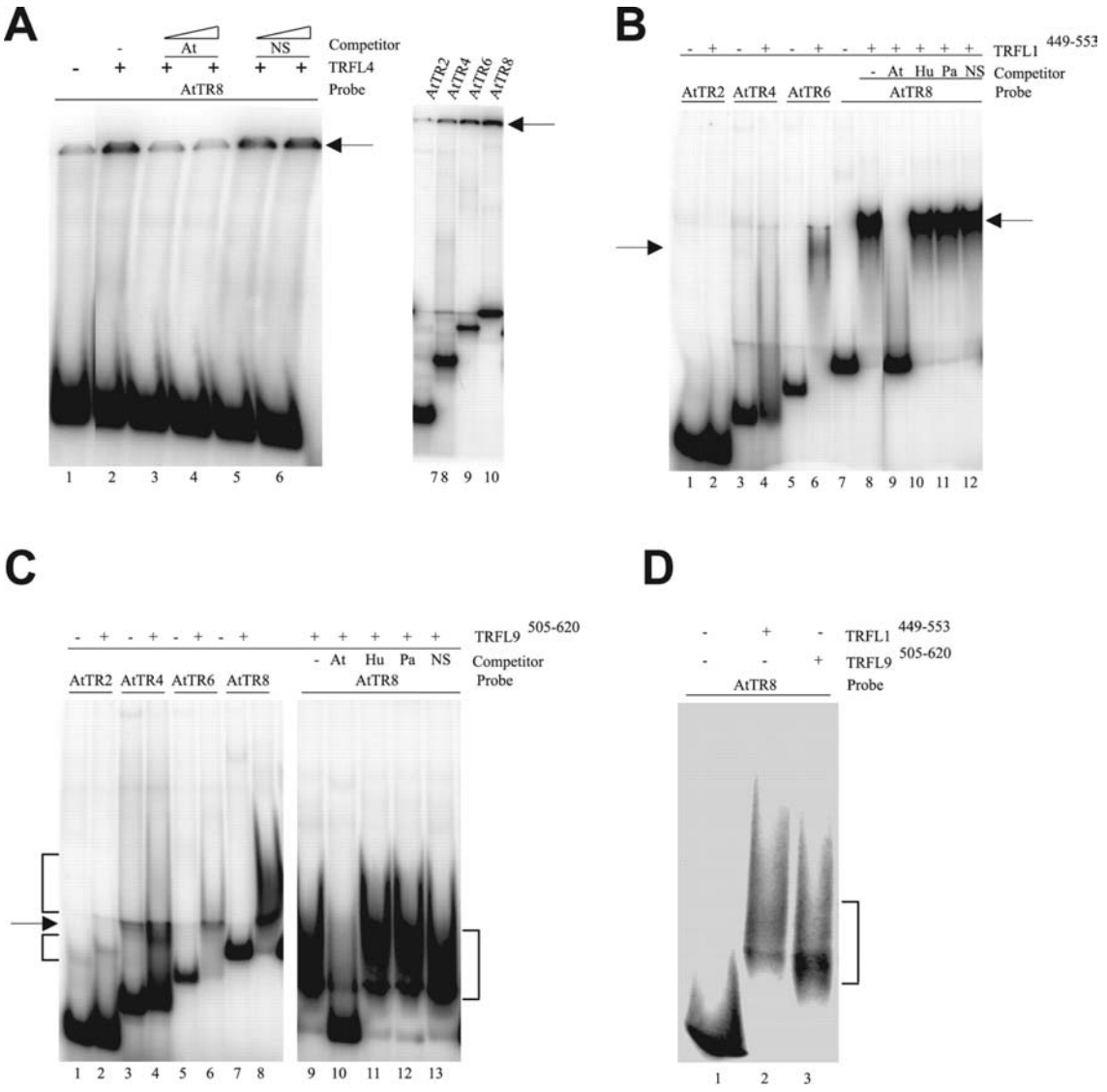


Figure 12. DNA binding properties of *Arabidopsis* TRFL proteins. EMSA was conducted with full-length TRFL4 expressed in rabbit reticulocyte lysate (A) or isolated C-terminal regions of TRFL1 and TRFL9 (B and C) expressed in wheat germ extract. ^{32}P-labeled double-stranded DNA probes consisted of two (AtTR2), four (AtTR4), six (AtTR6) or eight (AtTR8) *A. thaliana* telomeric repeats. EMSA controls with unprogrammed translation extract are shown in lane 1 of panel A, lanes 1, 3, 5 and 7 of panel B, and lanes 1, 3, 5 and 7 of panel C. Competition experiments were carried out with a 50X or 100X excess of unlabeled competitor DNA containing eight telomeric DNA repeats of *A. thaliana* (At), human (Hu), *Paramecium* (Pa) DNA, or a non-specific (NS) DNA sequence. Arrows denote position of protein-DNA complexes.

Strikingly, no DNA binding was detected in EMSA with TRFL family 2 members using either full-length proteins or isolated Myb domains (plus the remaining C-terminus) (Table 1; See Figure 13). As for TRFL family 1 members, all of the recombinant proteins generated from TRFL family 2 genes expressed well in the wheat germ translation system and were soluble (data not shown). Hence, the lack of telomeric DNA binding is likely to reflect inherent differences in the properties of the two TRFL families rather than protein misfolding. We conclude that, in contrast to the situation with human TRF1 (Konig *et al*., 1998), the capacity to bind to duplex plant telomeric DNA *in vitro* is not conveyed solely by the presence of a C-terminal Myb domain, since all of the TRFL proteins we examined have this feature.

The Myb-ext domain is necessary for telomeric DNA binding in vitro

One intriguing difference between TRFL family 1 and 2 members is the presence of the Myb-ext domain in family 1 (Figure 8B). To examine the contribution of this region in telomeric DNA binding, we performed EMSA on several truncated versions of TRFL1 (Figure 13A). TRFL1$^{449\text{-}540}$ (construct C-2) contains the Myb domain and Myb-ext, but lacks the remainder of the C-terminus. TRFL1$^{449\text{-}509}$ (construct C-3) includes only the Myb domain. EMSA showed that TRFL1$^{449\text{-}540}$ formed a complex with telomeric DNA that resembled that of the TRFL1$^{449\text{-}553}$ control (Figure 13B, lanes 2 and 3). In contrast, no shifted complex was detected in the assay with TRFL1$^{449\text{-}509}$ (Figure 13B, lane 4). One explanation for the inability of TRFL1$^{449\text{-}509}$ to bind telomeric DNA is that the C-terminus is required for proper protein folding. To address this concern, we generated a construct in which the Myb domain of TRFL1 was fused directly to the C-terminus of TRFL3, a member of TRFL family 2 (Figure 13A; construct C-4). No binding to

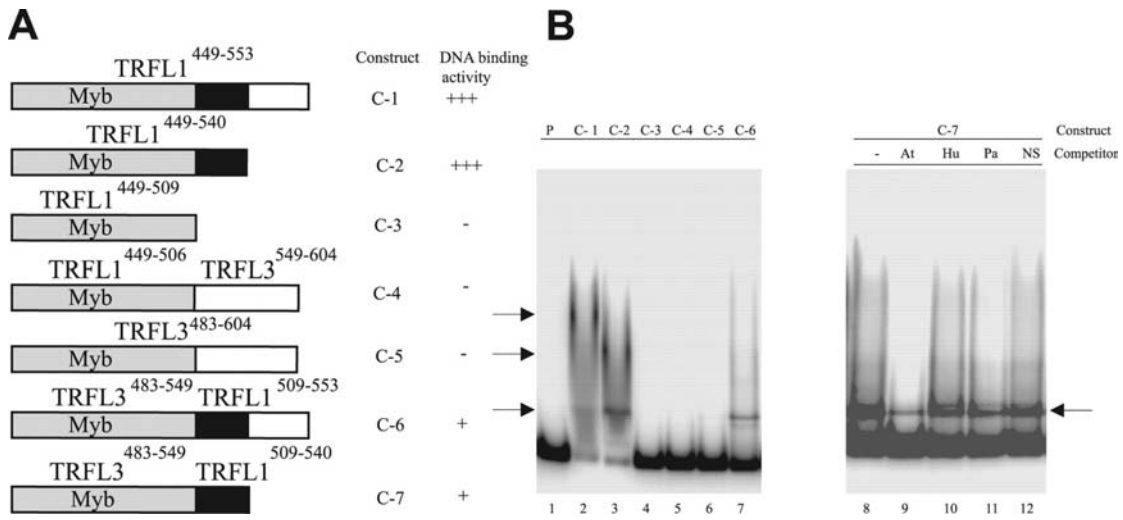


Figure 13. Analysis of the Myb-ext domain in telomeric DNA binding. (A) Schematic representation of truncation and chimeric fusion constructs. The Myb domain is highlighted in gray, the Myb-ext domain in black and the adjacent C-terminal region in white. A summary of results from EMSA is shown on the right. (B) EMSA reactions were performed with the constructs shown in panel A as described in Figure 12. The probe was AtTR8. Arrows denote the position of protein-DNA complexes.

telomeric DNA was detected with this construct (Figure 13B, lane 5). Taken together, these data argue that the Myb-ext domain is required for binding to plant telomeric DNA.

We next asked whether the Myb-ext domain was sufficient to confer telomeric DNA binding to a TRFL family 2 protein. For this experiment, a chimeric protein was generated that consisted of the TRFL3 Myb domain fused to the TRFL1 Myb-ext and C-terminal region (Figure 13A, construct 6). Strikingly, a DNA-protein complex was observed with this construct (Figure 13B, lane 7). The capacity to bind telomeric DNA was not influenced by the residues C-terminal to the Myb-ext domain, as another chimeric construct (C-7) lacking this region also specifically bound telomeric DNA (Figure 13B, lanes 8-12). This result not only indicates that the Myb-ext domain is required for binding to plant telomeric DNA, but it also demonstrates that when placed next to a Myb domain of a TRFL family 2 protein it is sufficient to convey telomeric DNA binding specificity.

Discussion

TRF-like proteins from Arabidopsis

The integrity of the nucleoprotein complex known as the telomere cap is crucial for genome stability and for cell proliferation in eukaryotes. Recent data indicate that plants display an extraordinary tolerance to telomere dysfunction, as well as distinct mechanisms of telomere length regulation (McKnight *et al*., 2002; McKnight and Shippen, 2004). For example, mutations in the telomere-associated protein complex Ku70/80, which result in telomere shortening in yeast (Boulton and Jackson, 1996), lead to dramatic telomere elongation in *Arabidopsis* (Riha and Shippen, 2003a; Riha *et al*.,

2002). Unraveling these interesting evolutionary distinctions will require a greater understanding of telomere architecture in plants.

In this study we used a genomic approach to search for *Arabidopsis* homologs of the vertebrate TRF1 and TRF2 genes. The defining feature of this class of proteins is a single C-terminal domain reminiscent of the c-Myb DNA binding domain in oncoproteins (Fairall *et al*., 2001; Nishikawa *et al*., 2001). Among Myb-containing proteins, TRF1 and TRF2 are unique in that their Myb domains harbor an additional telobox consensus motif, which contributes to the specific recognition of human telomeric DNA. Therefore, we searched for genes harboring a single C-terminal Myb domain with a telobox consensus motif and identified 12 TRF-like genes.

Since TRF1 and TRF2 are the only Myb-containing proteins known to bind directly to double-strand telomeric DNA in vertebrates, it was surprising to discover that *Arabidopsis* harbors so many TRFL genes. At least four TRFL genes (*TRP1/TRFL1* and *TRFL3/TRFL6*) reside in regions of the *Arabidopsis* genome known to be duplicated. Another pair, *TBP1* and *TRFL9*, also appear to reflect a recent duplication. Evidence for functional redundancy was obtained when we examined T-DNA insertion lines that disrupted the coding region in seven of the twelve TRFL genes (*TRP1, TRFL1, 2, 3, 4, 5, 8*). None of the single gene disruption mutants displayed defects in growth and development, nor showed perturbations in telomere length or genome stability (Vespa, Surovtseva, Karamysheva and Shippen, unpublished data). These data argue that at least a subset of TRFL genes may have overlapping function.

Nevertheless, the remarkable sequence divergence outside the DNA binding domain for many of the TRFL genes raises the distinct possibility of unique roles at telomeres or elsewhere in the *Arabidopsis* genome. The telomeric DNA binding protein

Rap1 from budding yeast exhibits pleiotropic functions and is required for telomere maintenance as well as regulation of gene expression at internal sites in the genome (Pina *et al*., 2003). Similarly, while the primary role of hTRF1 is associated with chromosome termini, TRF1 associates with interstitial blocks of telomeric DNA in some mammalian species (Krutilina *et al*., 2001). *Arabidopsis* also contains short stretches of interstitial telomeric DNA (typically one to two TTTAGGG repeats), predominantly located in promoter regions (Regad *et al*., 1994) that function in the control of gene expression (Tremousaygue *et al*., 2003; Tremousaygue *et al*., 1999). Thus, some TRFL proteins may act in transcriptional regulation in a manner similar to Rap1. In this regard, it is noteworthy that TRFL family 1 members exhibit distinct minimal length requirements for telomeric DNA binding. For efficient *in vitro* binding, the Myb domain of TBP1 requires only two telomeric repeats while TRP1, TRFL4 and TRFL9 require a minimum of four repeats, and TRFL1 and TRFL2 need six repeats (this study; Chen *et al*., 2001; Hwang *et al*., 2001). These apparent distinctions in telomere sequence recognition could potentially impact *in vivo* function. Proteins that bind efficiently to shorter stretches of telomeric DNA may preferentially act at promoters, while those that favor longer telomere tracts would have a higher probability of binding the long tracts of telomeric DNA associated with chromosome ends. Preliminary localization experiments indicate that all of the TRFL proteins accumulate in the *Arabidopsis* nucleus (Vespa, Kato, Lam and Shippen, unpublished data), however more in-depth analysis will be required to determine which of these proteins associate specifically with chromosome ends *in vivo*.

While all of the TRFL genes were identified on the basis of a C-terminal Myb-like domain, phylogenetic analysis and inspection of the deduced protein sequence

alignment indicates that they encode two distinct classes of proteins (Figure 8B). In addition to the Myb domain, all of the proteins in family 1 possess two other highly conserved regions, a Myb-ext domain and another centrally located CD domain. The CD domain was previously noted by Yaung and colleagues, who suggested that it might serve as a substrate for ubiquitination as it contains a region similar to the ubiquitin domain (Yang *et al*., 2003). For vertebrate TRF1 and TRF2, the only conserved region outside the Myb-domain is the TRFH domain, which not only facilitates formation of homodimers, but also interacts with other telomere-associated factors such human Rap1 and Tin2 (Kim *et al*., 1999; Li and de Lange, 2003). Homodimerization of hTRF1 and hTRF2 is required for association with telomeric DNA *in vivo,* but these proteins cannot form heterodimers (Fairall *et al*., 2001). In contrast, *Arabidopsis* TRFL family 1 proteins form both homo- and heterodimers *in vitro* (Figure 11). Whether the CD domain in TRFL family 1 proteins serves as an interaction interface analogous to the TRFH domain in hTRF1 and hTRF2 remains to be determined. However, the capacity to form both homo-and heterodimers *in vitro* suggests that TRFL family 1 members could participate in complex structural and functional regulation *in vivo*.

A novel domain is necessary for TRFL proteins to bind plant telomeric DNA

One of the most striking differences among TRFL family 1 and family 2 proteins is the presence of a highly conserved Myb-ext domain in family 1. Interestingly, the maize initiator-binding protein (IBP1) and the parsley BoxP-binding factor (BPF1) also harbor a Myb domain with an adjacent region with striking similarity to Myb-ext (Karamysheva and Shippen, unpublished data). IBP1 is known to interact at the transcription start site of the Shrunken promoter containing a perfect telomeric repeat AGGGTTT (Lugert and

Werr, 1994), while BPF1 binds a series of GT-rich motifs (da Costa e Silva *et al*., 1993). Based on their similarity to the TRFL family 1 from *Arabidopsis*, we predict that these proteins may have the capacity to efficiently bind longer stretches of plant telomeric DNA.

It is intriguing that the Myb domain of hTRF1 is sufficient for binding human telomeric repeat DNA *in vitro*, but this is not the case for TRFL proteins from *Arabidopsis*. Our data demonstrate that this latter class of proteins requires a more extended domain for telomeric DNA interactions. We found that Myb-ext is both necessary and sufficient to allow the Myb domains of TRFL proteins to bind double strand plant telomeric DNA. Precisely how Myb-ext contributes to telomeric DNA binding is unclear. One possibility is that it stabilizes the essential protein contacts between the Myb domain and telomeric DNA.

In eukaryotes a Myb-like domain is a defining feature of double-strand telomeric DNA binding proteins. The large number of such proteins in *Arabidopsis* confounds the identification of bona-fide telomere-associated factors. Although the Myb domains of TRFL family 2 proteins display even greater sequence similarity to the corresponding region of hTRF1 and hTRF2 than TRFL family 1 proteins, TRFL family 2 proteins fail to bind plant telomeric DNA *in vitro.* We cannot exclude the possibility that TRFL family 2 proteins associate with telomeres *in vivo* through protein interactions in a manner similar to human Rap1. Furthermore, it is evident that TRFL family 1 proteins are not the only factors capable of specific binding to double-strand plant telomeric DNA *in vitro*. While the architecture of TRFL family 1 proteins closely resembles the known double-strand telomere binding proteins in higher eukaryotes, the SMH-like proteins from maize and *Arabidopsis* plant harbor a single Myb domain in their N-terminus, but

lack an associated Myb-ext domain, yet still bind telomeric DNA *in vitro* (Marian *et al*., 2003; Schrumpfova *et al*., 2004). Thus, TRFL family 1 and SMH-like proteins apparently represent two distinct classes of telomeric DNA binding proteins. Deciphering the functions and interactions of such proteins may provide new insight into the structure and complex regulation of the telomere cap.

CHAPTER III

THE *Arabidopsis* POT1 AND POT2 PROTEINS FUNCTION IN TELOMERE LENGTH HOMEOSTASIS AND CHROMOSOME END PROTECTION*

Summary

Pot1 (Protection of telomeres 1) is a single-strand telomere binding protein that is essential for chromosome end protection and telomere length homeostasis. *Arabidopsis* encodes two Pot1-like proteins, dubbed AtPot1 and AtPot2. Here we show that telomeres in transgenic plants expressing a truncated *AtPot1* allele lacking the N-terminal OB-fold (P1ΔN) are 1 1.5kb shorter than in wild type, suggesting that AtPot1 contributes to the positive regulation of telomere length control. In contrast, telomere length is unperturbed in plants expressing the analogous region of AtPot2. A strikingly different phenotype is observed in plants over-expressing the AtPot2 N-terminus (P2ΔC), but not the corresponding region in AtPot1. Although bulk telomeres in P2ΔC mutants are 1–2kb shorter than in the wild-type, these plants resemble late generation telomerase-deficient mutants with severe growth defects, sterility and massive genome instability, including bridged chromosomes and aneuploidy. The genome instability associated with P2ΔC mutants implies that AtPot2 contributes to chromosome end

---

*Reprinted with permission from 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.

protection. Thus, *Arabidopsis* has evolved two Pot genes that both function in telomere biology. These findings provide unanticipated information about the evolution of single-strand telomere binding proteins.

Introduction

Telomeres are the essential protein-DNA structures at the ends of linear eukaryotic chromosomes whose primary functions are to facilitate complete replication of the chromosome terminus and to sequester it from DNA repair machinery and exonucleolytic attack (Blackburn, 2001; Chan and Blackburn, 2002; de Lange, 2002). In most organisms, the DNA component of telomeres consists of tandem repeats of simple G-rich sequences that terminate in a single-strand 3' extension. The telomere can fold back on itself to form a t-loop, where the 3' G-overhang invades the duplex region of the telomere to create a displaced loop (D-loop) consisting of single-stranded G-rich repeats (Griffith *et al*., 1999). During S phase, the t-loop is thought to unfold, allowing telomerase access to the G-overhang for telomere length maintenance (Wei and Price, 2003).

The G-overhang associates with single-strand specific proteins (Smogorzewska and de Lange, 2004; Wei and Price, 2003). The first G-strand binding protein identified, telomere end binding protein (TEBP), was found in the hypotrichous ciliate *Oxytricha nova.* TEBP is a heterodimer of α and β subunits that binds tenaciously to the 3' terminus of the G-overhang (Price and Cech, 1987) via four oligonucleotide/oligosaccharide binding folds (OB folds) (Horvath *et al*., 1998). The OB fold is a structurally conserved feature also associated with single-strand telomere binding proteins in fungi and vertebrates (Theobald and Wuttke, 2004). The

Saccharomyces cerevisiae protein, Cdc13p, is the best characterized of this class of proteins (Evans and Lundblad, 2000; Smogorzewska and de Lange, 2004; Wei and Price, 2003). A multifunctional protein, Cdc13p binds the single-strand G-overhang and provides telomere end protection, facilitates telomerase recruitment and repression, and coordinates telomeric G- and C-strand synthesis through interactions with lagging strand replication machinery (Wei and Price, 2003).

A distant relative of TEBP called Pot1 was recently found in *Schizosaccharomyces pombe* (Baumann and Cech, 2001). SpPot1 shares weak sequence similarity with TEBP at its N-terminus and, like Cdc13p, assumes an OB-fold that facilitates specific recognition of telomeric DNA (Lei *et al*., 2002; Lei *et al*., 2003). Deletion of *SpPot1* leads to immediate and catastrophic loss of telomeric repeats and, to some extent, erosion of subtelomeric DNA. *SpPot1* mutants mis-segregate their chromosomes and ultimately fail to divide. Cells that survive the loss of *SpPot1* undergo chromosome circularization (Baumann and Cech, 2001). In a genetic screen for mitotic mutants, a sequence homologue of Pot1 was identified in *Aspergillus nidulans* (Pitt *et al*., 2004). As with *SpPot1* mutants, *AnPot1* deficiency results in severe mitotic defects, as well as chromosome mis-segregation. These data imply that Pot1 from fungi is involved in chromosome end protection.

Pot1 orthologs have also been identified in vertebrates (Baumann *et al*., 2002; Wei and Price, 2004) and are implicated in both telomere length regulation and chromosome end protection (Colgin *et al*., 2003; Loayza and de Lange, 2003; Veldman *et al*., 2004). Both human and chicken Pot1 proteins localize to telomeres *in vivo* (Baumann *et al*., 2002; Wei and Price, 2004). Although hPot1 binding is reduced in cells that have lost the G-overhang, hPot1 appears to associate with telomeres primarily

through interactions with proteins that bind along the length of the duplex (Loayza and de Lange, 2003). The C-terminal domain of hPot1 interacts with the Pip1/PTOP protein linking hPot1 to the TRF1/TIN2 complex (Liu *et al*., 2004; Ye *et al*., 2004b). In several studies, over-expression of full-length hPOT1 resulted in lengthened telomeres in some, but not all human tumor cell types, suggesting that Pot1 is a positive regulator of telomere length (Armbruster *et al*., 2004; Colgin *et al*., 2003; Liu *et al*., 2004). However, in another study over-expression of a dominant negative C-terminal fragment of hPot1 led to extensive telomerase-dependent telomere lengthening (Loayza and de Lange, 2003), implying that hPot1 serves as a negative regulator of telomere length. In support of this conclusion, RNAi-mediated knock down of hPot1 resulted in longer telomeres (Veldman *et al*., 2004; Yang *et al*., 2005; Ye *et al*., 2004b), and recent *in vitro* studies suggest that hPot1 negatively regulates telomerase activity (Kelleher *et al*., 2005). hPot1 has also been implicated in chromosome end protection (Veldman *et al*., 2004). Together, these data indicate that hPot1, like Cdc13p, is a multifunctional protein.

Here we investigate the role of Pot proteins in *Arabidopsis thaliana*. *Arabidopsis* has emerged as a powerful system for telomere biology (Riha and Shippen, 2003b). A genetically tractable higher eukaryote, *Arabidopsis* harbors short telomere tracts (2-5 kb in the Columbia ecotype) (Shakirov and Shippen, 2004) and displays an unusually high tolerance to the genome instability that accompanies telomere dysfunction (Riha *et al*., 2001). Although relatively little is known about telomere-associated proteins in plants, *Arabidopsis* is distinguished from many other model organisms in that it harbors two Pot1-like genes (Baumann *et al*., 2002). In this study we show that AtPot1 and AtPot2 are ubiquitously expressed at a low level. We also present data from transgenic plants establishing a function for both proteins in telomere length maintenance, and a role for

AtPot2 in chromosome end protection. These findings argue that the known functions of single-strand telomere binding proteins have been segregated into two distinct polypeptides in *Arabidopsis*.

Materials and methods

Plant materials, construction of AtPot1 *and* AtPot2 *mutant alleles and transformation*

Wild type *Arabidopsis* seeds were purchased from the *Arabidopsis* Biological Resource Center (Ohio State University, Columbus), cold-treated overnight at 4°C, then placed in the environmental growth chamber and grown under a 16/8-hr light/dark photoperiod at 23°C. To obtain constructs for over-expression, regions of *AtPot1* and *AtPot2* cDNAs corresponding to the full-length, ΔC (amino acids 1-175), and ΔN (amino acids 176-467 for AtPot1 and 176-454 for AtPot2) polypeptides were amplified by PCR and then inserted into a binary vector pCBK05 (Riha *et al*., 2002) to allow expression from a 35S CaMV promoter. The constructs were introduced into *Agrobacterium tumefaciens* strain GV3101. Transformation of wild type plants was performed by the *in planta* method (Riha *et al*., 2002). T1 primary transformants were selected on 0.5 BM medium supplemented with 20 mg/l of phosphinothricine (Crescent Chemical), genotyped and analyzed by RT-PCR for transgene expression.

cDNA synthesis

Total RNA was extracted from 0.1–0.5 g of plant tissue using Tri Reagent solution (Sigma). *AtPot1* and *AtPot2* cDNAs were synthesized from total leaf RNA using Superscript III reverse transcriptase (Gibco). Primers complementary to the stop codons of each cDNA were incubated with 2 μg of total RNA in the supplied buffer at

65°C for 5 min. Reverse transcription was carried with 100U of Superscript III at 55°C for 50 min. RNA was degraded with RNase H (USB). The coding regions of *AtPot1* and *AtPot2* were then amplified with Turbo-Pfu polymerase (Stratagene). PCR products were cloned and verified by sequencing. The sequences of *AtPot1* and *AtPot2* cDNA have been deposited into GenBank, accession numbers AY884593 and AY884594, respectively.

TRF analysis, TRAP assays, cytogenetics and fusion PCR

DNA from individual whole plants was extracted as described (Cocciolone and Cone, 1993). TRF analysis was performed with *Tru*1I (Fermentas) restriction enzyme and ^{32}P 5' end-labeled $(T_3AG_3)_4$ oligonucleotide as a probe (Fitzgerald *et al*., 1999). TRAP assays were performed on plant tissues as described in (Fitzgerald *et al*., 1996). Anaphase spreads were prepared from pistils and stained with DAPI as described (Riha *et al*., 2001). Telomere fusion PCR was performed as previously described (Heacock *et al*., 2004). Briefly, DNA from mutant or wild-type plants was PCR-amplified using primers for unique subtelomeric sequences that were directed towards telomeres. PCR products were resolved on an agarose gel, transferred to a nitrocellulose membrane and hybridized with ^{32}P 5' end-labeled $(T_3AG_3)_4$ oligonucleotide probe.

Results

Identification of two Pot genes in Arabidopsis

Two genes encoding putative Pot1 orthologs from *Arabidopsis thaliana* were previously identified using the *S. pombe* Pot1 protein sequence (Baumann *et al*., 2002) as the query in a BLAST search of the *Arabidopsis* genome database. *Arabidopsis* genes

At2g05210 and At5g06310 showed strong similarity with the N-terminal Telo_bind_N domain, pfam02307, (Marchler-Bauer and Bryant, 2004) of the *Oxytricha nova* TEBP α subunit and with the corresponding regions in *S. pombe* and human Pot1 proteins (Figure 14A). The *Arabidopsis* genes were designated *AtPot1* and *AtPot2*, respectively. RT-PCR analysis of mRNA from leaves, flowers, roots, stems and callus revealed that *AtPot1* and *AtPot2* mRNAs are ubiquitously expressed, although both transcripts are present in low amounts (Figure 14B). This expression pattern differs from *AtTERT* mRNA, which is confined to proliferating tissues (Fitzgerald *et al*., 1996). Human Pot1 mRNA is alternatively spliced (Baumann *et al*., 2002) and recent RT-PCR analysis suggests that AtPot1 and AtPot2 are also subject to alternative splicing (Tani and Murata, 2005). However, the variant splicing in *Arabidopsis* is not tissue-specific and all but one alternatively spliced forms of the *AtPot1* and *AtPot2* mRNAs are predicted to result in prematurely truncated proteins. The functional relevance of these truncated isoforms is unclear.

The N-termini of both AtPot1 and AtPot2 are predicted to assume two OB-folds (M. Lei, personal communication). The OB-fold domains in AtPot1 and AtPot2 are 31% identical and 47% similar to each other (Figure 14C). Outside this domain, the identity is 39% (55% similarity). As expected, both *Arabidopsis* proteins are more closely related to the mammalian orthologs than to Pot1 proteins from fungi. The N-terminus of AtPot1 is 17% identical and 33% similar to the corresponding region of hPot1, while the N-terminus of AtPot2 is 17% identical and 29% similar. Outside of this region, the similarity is lower (Figure 14C). Notably, both *Arabidopsis* genes encode proteins that are significantly smaller in size than their mammalian and fungal counterparts.

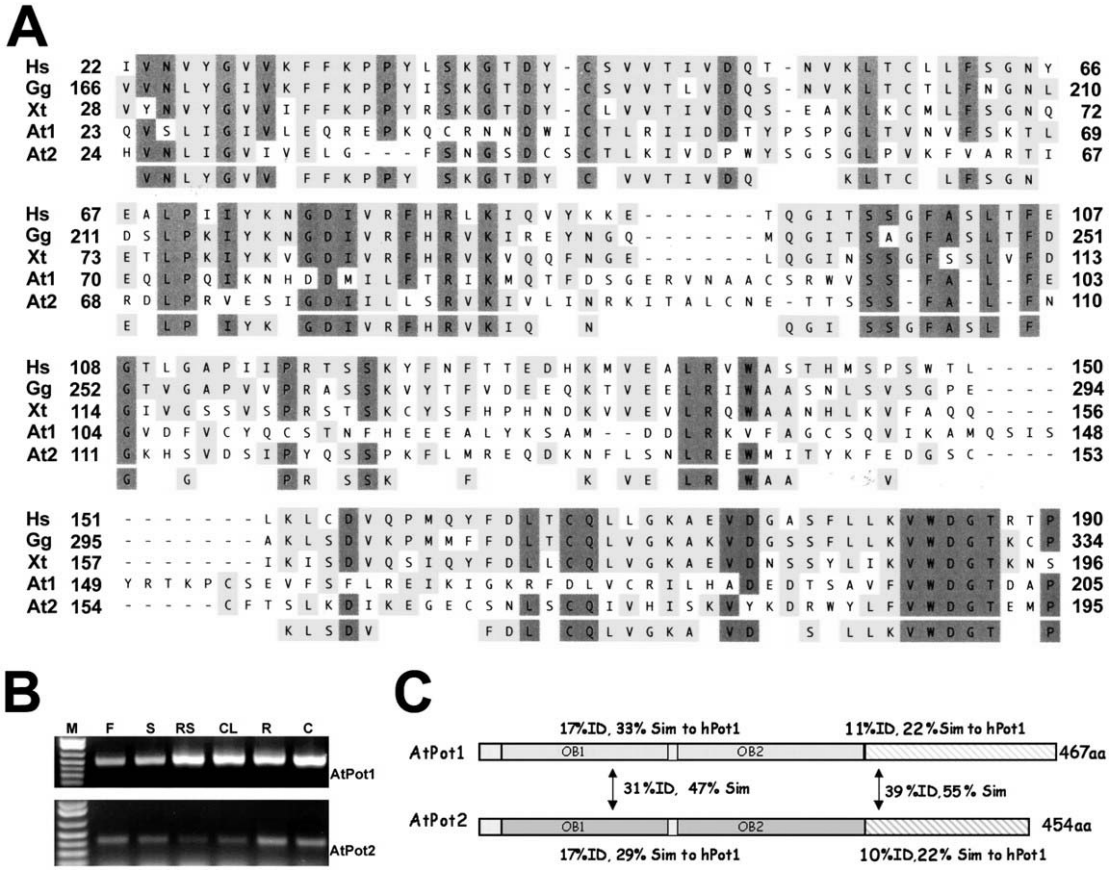


Figure 14. Two *Pot* genes in *Arabidopsis*. (A) Amino acid alignment of N-terminal portions of Pot1 proteins. Hs, *Homo sapiens*; Gg, *Gallus gallus* (chicken), Xt, *Xenopus tropicalis* (accession NP_998876); At1, AtPot1; At2, AtPot2. Residues conserved in at least two sequences are highlighted in light gray. Residues conserved in at least four sequences are highlighted in dark gray. Residues conserved in all five sequences are highlighted in pink. Consensus sequence is shown below alignment. (B) RT-PCR of *AtPot1* (upper panel) and *AtPot2* (lower panel) mRNAs in different *Arabidopsis* tissues. M, molecular weight markers; F, flowers; S, stems; RL, rosette leaves; CL, cauline leaves; R, roots; C, callus. (C) Sequence similarity between AtPot1 and AtPot2 proteins and human Pot1 protein.

Over-expression of AtPot1 and AtPot2 in transgenic Arabidopsis

Pot1 proteins from vertebrates and fission yeast specifically bind telomeric DNA *in vitro* (Baumann and Cech, 2001; Loayza *et al*., 2004). To determine whether AtPot1 and AtPot2 display affinity for the plant telomeric DNA sequence, we expressed recombinant forms of the *Arabidopsis* proteins in *E. coli* and in baculovirus. Most of the protein was insoluble, despite the application of many different expression conditions; however, the small fraction of soluble AtPot1 and Atpot2 protein that could be obtained displayed specific, but very weak affinity for single-strand plant telomeric DNA (data not shown). This observation suggests that AtPot1 and AtPot2 have the capacity to associate with telomeric DNA. Unfortunately, the inability to obtain sufficient quantities of soluble recombinant protein has stymied further biochemical analysis.

To investigate the function of AtPot1 and AtPot2, we employed a transgenic approach. When this study was initiated, no T-DNA disruption lines were available for either gene in the annotated databases of several major *Arabidopsis* insertional mutagenesis facilities. Moreover, our attempts to reduce AtPot1 and AtPot2 protein levels by RNAi were unsuccessful. Therefore, we examined the consequences of over-expressing full-length *AtPot1*, *AtPot2* or C-terminal and N-terminal truncation derivatives. This strategy has been previously employed to investigate the function of human Pot1 (Colgin *et al*., 2003; Loayza and de Lange, 2003).

The C-terminal fragments of AtPot1 and AtPot2, dubbed P1ΔN (residues 176-467) and P2ΔN (residues 176-454), respectively, roughly correspond to the hPot1$^{\Delta OB}$ construct (Loayza and de Lange, 2003). The N-terminal AtPot1 and AtPot2 fragments, P1ΔC and P2ΔC (residues 1-175), span the remainder of the proteins and contain the

first OB-fold. Previous studies indicate that the corresponding region in *S. pombe* Pot1 is sufficient for telomeric DNA binding *in vitro* (Lei *et al*., 2002).

To ensure robust expression in transgenic plants, the AtPot1 and AtPot2 constructs were cloned downstream of the CaMV *35S* promoter and were transformed into *Arabidopsis*. Over-expression of transgene mRNA was verified in all the transformants by RT-PCR with N- and C-terminal primer pairs (see below and data not shown). Ectopic expression of full-length AtPot1 or AtPot2 proteins yielded no defects in growth or development, and telomeres were indistinguishable from wild type controls as assayed by terminal restriction fragment (TRF) analysis (Figure 15A and B). The modest variability in telomere length associated with over-expression of full-length AtPot1 and AtPot2 falls within the wild type range of 2-5kb for *Arabidopsis* plants of the Columbia ecotype (Shakirov and Shippen, 2004). The longest telomeres in wild-type typically span 4.0-5.5 kb and the shortest 1.6-2.5 kb (Richards and Ausubel, 1988; Shakirov and Shippen, 2004). The precise range is largely determined by the size of the telomere tract in the parent (Shakirov and Shippen, 2004). Transgenic plants over-expressing the AtPot1 N-terminus (P1ΔC) or the AtPot2 C-terminus (P2ΔN) also appeared wild type and showed no telomere perturbations when subjected to TRF analysis (Figure 15C and D). Although the mRNAs for these constructs were highly expressed in transgenic plants, we could not verify that this was true for the corresponding polypeptides. Therefore, the lack of a phenotype in transgenic lines expressing full-length AtPot1, AtPot2, P1ΔC or P2ΔN may reflect poor expression of these proteins. Alternatively, over-expression of these constructs may simply have no detrimental consequences for *Arabidopsis*.

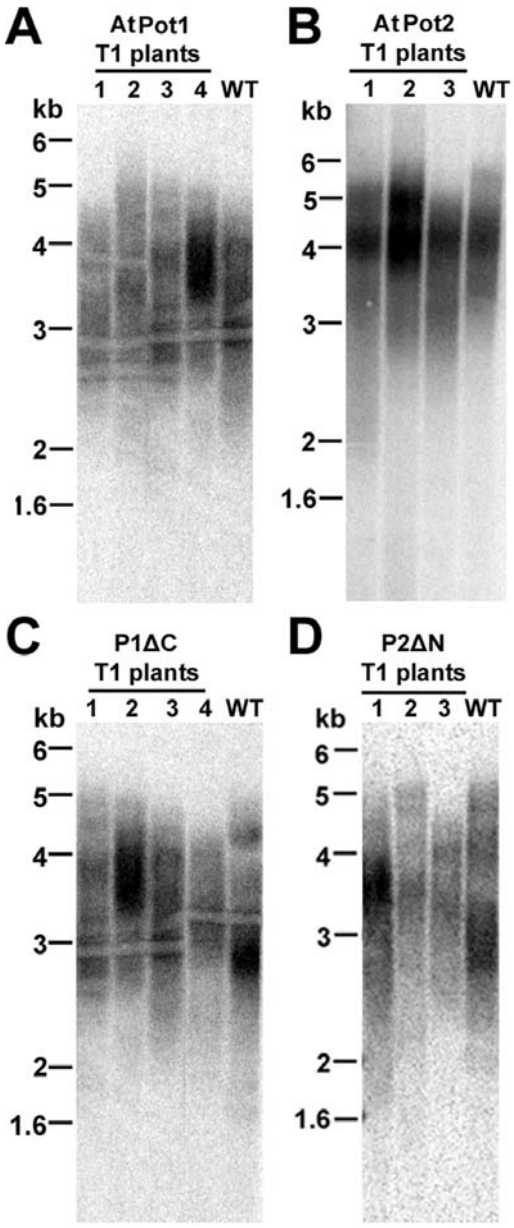


Figure 15. Over-expression of full-length AtPot1 and AtPot2 or the P1ΔC and P2ΔN derivatives in *Arabidopsis* does not alter telomere length. TRF analysis was performed on DNA extracted from whole primary transformants (T1) shown by RT-PCR to express full-length AtPot1 (A), full-length AtPot2 (B), P1ΔC (C) or P2ΔN (D). Telomere length is 2-5 kb in wild-type (WT) *Arabidopsis* plants of Columbia ecotype (Shakirov and Shippen, 2004).

Plants over-expressing P1ΔN and P2ΔC, by contrast, displayed reproducible, aberrant phenotypes, and these plants were evaluated further. All of the plants over-expressing P1ΔN (n = 14) (Figure 16A) were fertile and morphologically indistinguishable from wild type. While TRF analysis revealed that many primary transformants (T1) had wild type telomere length (Figure 16B, lanes 2 and 4), a significant percentage (~21%) displayed telomeres that were 1-1.5kb shorter than wild type. In such cases, the longest telomeres averaged 3.5 kb (Figure 16B, lanes 1 and 3).

Subsequent generations of self-pollinated mutants displayed either very limited (≤300 bp) or no additional telomere shortening (Figure 16C and data not shown), indicating that a new set point for telomere length was reached. Interestingly, in a subset of second generation progeny (T2), telomeres occupied a much broader size range, suggesting that a fraction of telomeres had returned to the wild type length (Figure 16C, lanes 4, 5, 6 and 8), even though RT-PCR indicated that these plants continued to express the transgene mRNA (data not shown). A somewhat similar unstable phenotype is associated with clonal human cell lines over-expressing full-length hPot1 (Colgin *et al*., 2003). The high degree of sibling-to-sibling variation in telomere length observed with P1ΔN mutants in T2 is not seen in wild type *Arabidopsis* (Shakirov and Shippen, 2004), and implies that the loss of telomeric DNA is reversible. We conclude that AtPot1 contributes to telomere length homeostasis, and since telomere shortening was detected, our data suggest that AtPot1 is a positive regulator of telomere length.

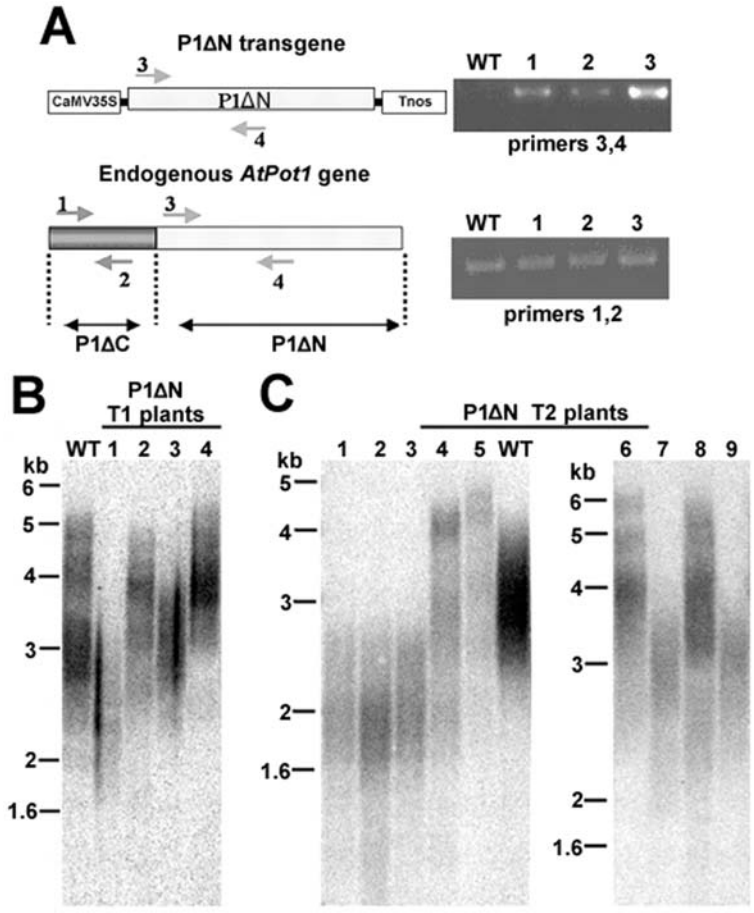


Figure 16. Over-expression of P1ΔN leads to telomere shortening. (A) Schematic diagram of the P1ΔN transgene and the corresponding endogenous *AtPot1* allele (left); RT-PCR results of P1ΔN expression in transgenic plants 1-3 from panel B (right). Primers 1 and 2 are specific for the endogenous *AtPot1* mRNA. After 40 cycles, *AtPot1* mRNA is amplified in all plants equally well (lower panel, lanes 1-3 and WT). Primers 3 and 4 amplify both the P1ΔN transgene and the endogenous *AtPot1* allele. After 20 PCR cycles, P1ΔN mRNA is amplified only in the transgenic plants (upper panel, lanes 1-3), but not in the wild type plant, (upper panel, WT), confirming that the P1ΔN transgene is over-expressed. (B) and (C) TRF analysis of P1ΔN mutants. (B) Examples of P1ΔN primary transformants (T1) displaying wild type telomere lengths (lanes 2 and 4), or shortened telomeres (lanes 1 and 3) are shown. (C) Analysis of the second generation (T2) progeny from transformants #1 (lanes 1-5) and #3 (lanes 6-9) from panel B. For most T2 plants, no additional telomere shortening was observed (lanes 1-3, 7 and 9), but in some siblings telomere tracts were extended to resemble wild type (lanes 4 and 5, 6 and 8).

Over-expression of P2ΔC results in severe morphological defects and telomere dysfunction

As with P1ΔN mutants, the majority of P2ΔC transformants (n = 52) (Figure 17A) were normal in appearance. However, a substantial fraction (~10%) showed distinct morphological defects at two-three weeks of age (Figure 17B and C). Such plants exhibited a "terminal" morphological phenotype similar to that of late generation (G_{7-9}) telomerase-deficient plants (Riha *et al*., 2001), with delayed growth (Figure 17B) and flowering time, and abnormally small rosette leaves that were wrinkled and curled-down (Figure 17C). Although numerous small siliques formed, they were sterile and produced no seeds (data not shown). Several attempts to make reciprocal crosses of the mutants to wild type failed, consistent with severe anomalies in both male and female reproductive systems.

Over-expression of P2ΔC also resulted in an altered telomere phenotype in plants that displayed morphological defects. TRF analysis of such mutants revealed telomeres that were significantly shorter than in wild type, with some transformants losing between 1 and 2 kb of telomeric sequence in a single generation (Figure 17D, lanes 1, 3, and 5). As with P1ΔN, not all primary P2ΔC transformants had shortened telomeres (Figure 17D, lanes 2, 4 and 6); however, this phenotype was reproducible and exclusively associated with mutants that displayed morphological defects and sterility. The telomere shortening associated with P2ΔC and P1ΔN over-expression is not due to inactivation of telomerase. TRAP assays revealed that telomerase is biochemically active in both mutants (Figure 18). Furthermore, relative to first generation telomerase mutants, which lose 200-500bp of telomeric DNA (Riha *et al*., 2001a), depletion of telomeric DNA in P2ΔC and P1ΔN transgenic plants is greater by

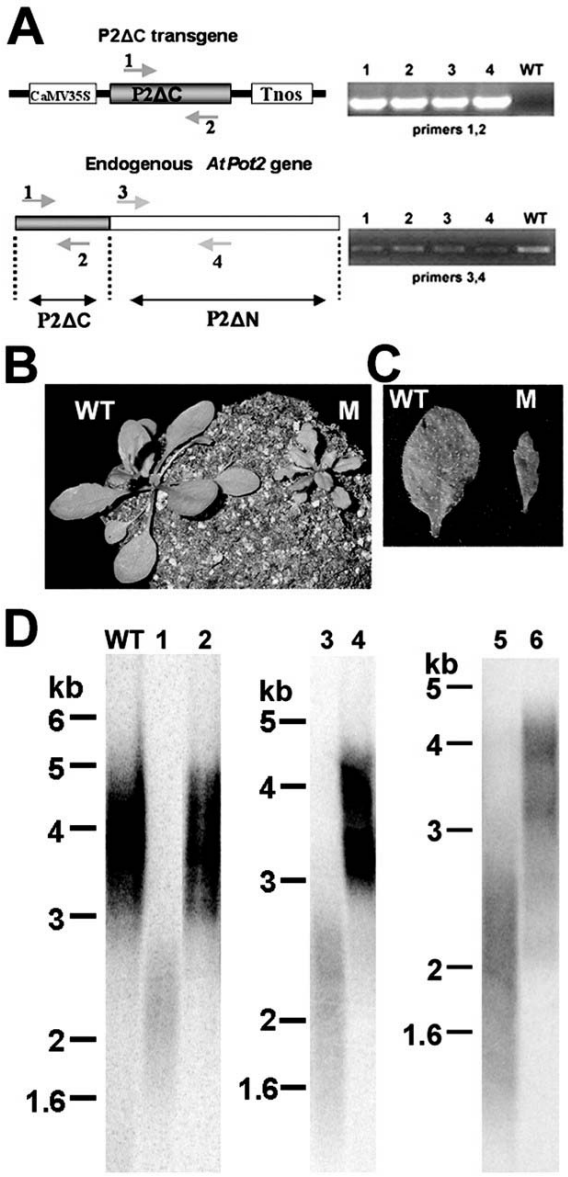


Figure 17. Morphological defects and telomere shortening in mutants over-expressing P2ΔC. (A) Schematic of the P2ΔC transgene and the corresponding endogenous *AtPot2* allele (left); RT-PCR analysis of P2ΔC expression in transgenic plants 1-4 from panel D (right). RT-PCR reactions were conducted as described in Figure 16. (B) and (C) Growth and developmental defects in P2ΔC mutants. (B) At 2-3 weeks of age, growth of several independent transgenic mutants (M) was delayed relative to wild type (WT). (C) Such mutants displayed severe defects in leaf development. (D) TRF analysis of P2ΔC transformants. Primary transformants (T1) with a wild type appearance displayed wild type telomere length (lanes 2, 4 and 6), whereas plants that exhibited morphological defects had shortened telomeres (lanes 1, 3 and 5).

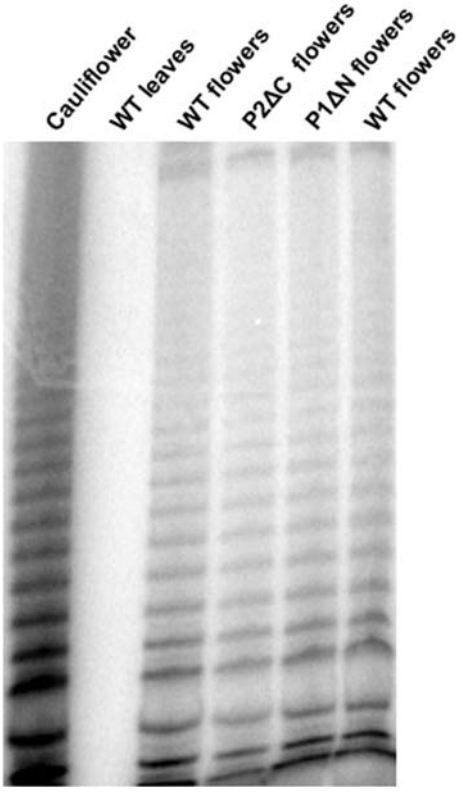


Figure 18. TRAP assays of wild type and mutant *Arabidopsis*. Extracts from cauliflower inflorescence and wild type *Arabidopsis* flowers serve as positive controls and wild type *Arabidopsis* leaves as a negative control for telomerase activity. P2ΔC and P1ΔN mutants have wild-type levels of telomerase activity *in vitro*.

2-3 fold and 1.5 fold, respectively (Figures 16 and 17). Thus, the telomeres in both mutants appear to suffer additional replication defects or are exposed to nuclease attack.

In *Arabidopsis* the severe morphological abnormalities associated with dysfunctional telomeres are accompanied by genome instability (Riha *et al*., 2001). Therefore, transgenic plants were subjected to cytogenetic analysis of mitotic cells in the pistil. No cytogenetic defects were observed in plants over-expressing full-length AtPot1, full-length AtPot2, P1ΔC, P2ΔN or P1ΔN (Table 2 and data not shown), consistent with their wild type appearance. Cytogenetic analysis was also performed on three of the P2ΔC mutants with morphological defects that were capable of producing flowers (Figure 19). In each case, we detected profound genome instability. Up to 8% of all mitotic figures contained anaphase bridges (Figures 19B and C; Table 2), consistent with the formation of dicentric chromosomes as a result of end-to-end telomere fusions. Remarkably, 50% of all anaphases with bridged chromosomes contained two or more fusions (Figure 19D and E), and in several cells a majority of the chromosomes appeared to be fused (Figure 19F). In some instances, large DNA fragments, possibly entire chromosomes, lagged behind (Figure 19G). Many mitotic cells displayed other defects, including unequal chromosome segregation that would result in aneuploidy in daughter cells (Figure 20H). Similar mitotic abnormalities, including increased chromosome instability and segregation errors, are associated with *A. nidulans Pot1* mutants (Pitt *et al*., 2004) and human cells with knocked-down hPot1 expression (Veldman *et al*., 2004).

To ask whether the cytological defects associated with P2ΔC mutants involve dysfunctional telomeres, we performed telomere fusion PCR using primers directed

Table 2. Frequency of chromosome fusions in *AtPot1* and *AtPot2* mutants

Genotype	Number of analyzed pistils	Number of anaphases		% bridges
		with bridges	total scored	
WT	4	0	291	0
P1ΔN	4	0	307	0
P2ΔC				
Plant 1	2	9	122	7.4
Plant 2	2	11	145	7.6
Plant 3	1	4	51	7.8

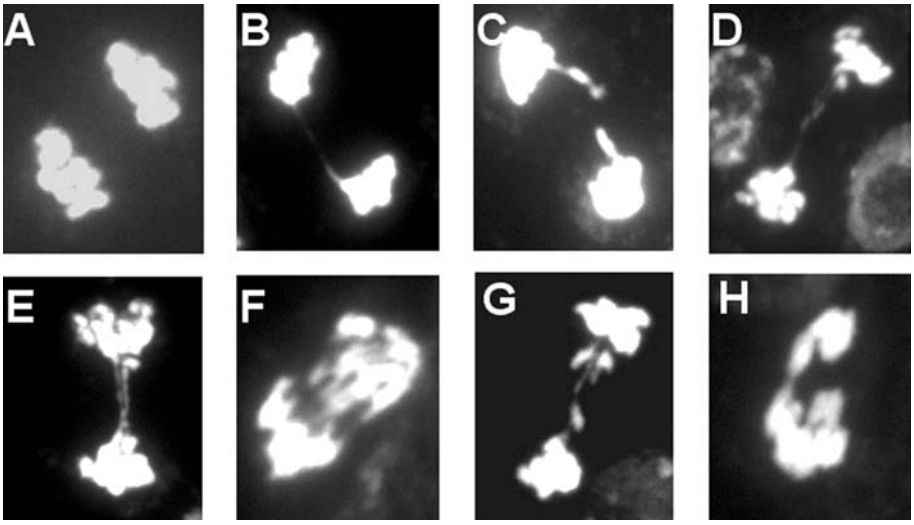


Figure 19. Cytogenetic defects in mutants over-expressing P2ΔC. Cytogenetic analysis was performed on actively dividing mitotic tissues of pistils from wild type (A) or P2ΔC mutants (B-H) using DAPI. Anaphase bridges, lagging chromosomes and aneuploidy are evident in P2ΔC mutants.

outward from the unique subtelomeric sequences found on *Arabidopsis* chromosomes (Heacock *et al*., 2004). In this assay, PCR products are generated when telomeres form covalent associations (Figure 20A). PCR reactions were performed with seven different primer combinations directed at subtelomeric sequences. Primer binding sites ranged from 200bp (5R) to 950bp (3R) upsteam of the telomere repeats. Strikingly, in contrast to PCR reactions with DNA from G_7 *tert* mutants, which display a comparable number of anaphase bridges to the most severely affected P2ΔC mutants (Heacock *et al*., 2004) (Figure 20B), only a few PCR products were generated from P2ΔC DNA with a subset of primer combinations (Figure 20B). The rare products that could be cloned from these reactions (i.e. 3L + 3R reaction) contained DNA derived from centromeric regions, which harbor a short stretch of telomeric DNA sequence. This observation implies that anaphase bridges in P2ΔC mutants do not involve intact telomeres. We speculate that the failure to amplify chromosome fusion junctions using subtelomeric primers reflects extensive nucleolytic processing of telomeres prior to end joining. Chromosomal termini in *S pombe pot1* mutants lose more than 5kb of DNA prior to circularization (Baumann and Cech, 2001). If this occurs in *Arabidopsis* P2ΔC mutants, the binding sites for the primers used in our assay would be absent. Alternatively, chromosomes in *Arabidopsis* P2ΔC mutants could be held together by non-covalent interactions. Taken together, our data argue that AtPot2 contributes to the protective cap on the chromosome terminus, and that loss of this function triggers telomere shortening, chromosome fusions and massive genome instability.

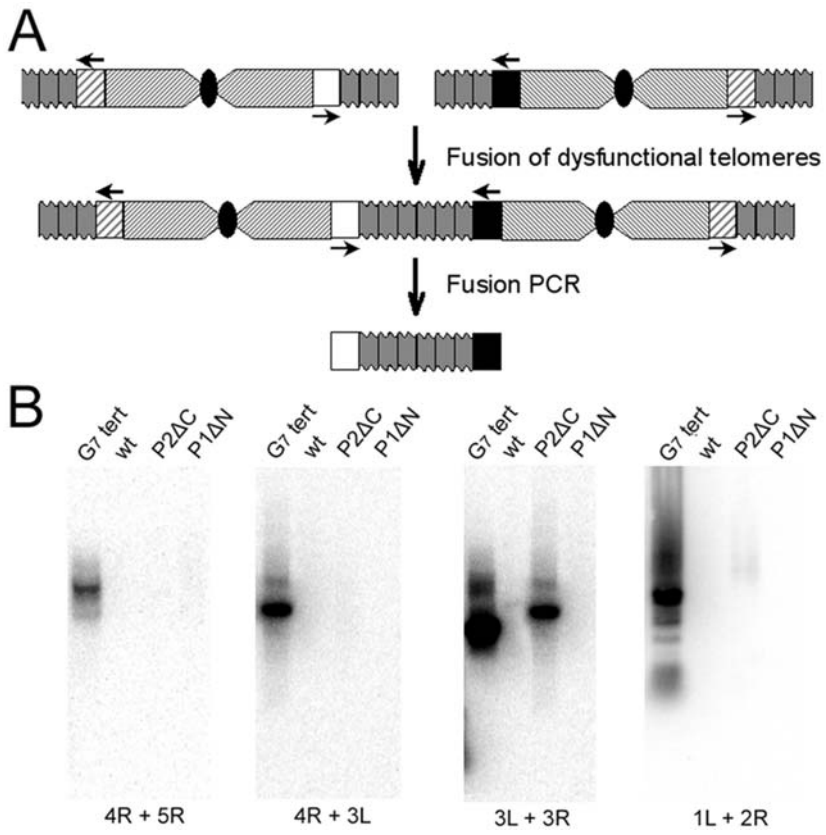


Figure 20. Telomere fusion PCR analysis of P2ΔC and *tert* mutants. (A) Schematic diagram of telomere fusion PCR. Primers specific to unique subtelomeric sequences (white, black and diagonally striped boxes) are directed towards telomeres and will only amplify a product if telomeres form covalent attachments with each other. (B) Primers specific for the right arm of chromosome 4 (4R) and left arm of chromosome 3 (3L), 4R and 5L, 4R and 5R, 3L and 3R, 3L and 5R, 5R and 5L, and 1L and 2R were used in the assay to amplify chromosome fusion products from DNA extracted from *tert,* wild-type, P2ΔC and P1ΔN plants.

Discussion

Single-strand telomere binding proteins are found in a diverse array of single-celled and multicellular eukaryotes. Cdc13p from *S. cerevisiae* provides the best studied example. Its functions include essential contributions to telomere length homeostasis and chromosome end protection, roles that are mediated through a plethora of protein interactions at the telomere (Evans and Lundblad, 2000; Wei and Price, 2003). Other proteins from this group include TEBP from *Oxytricha nova* and the Pot1 proteins from *S. pombe*, *A. nidulans,* chicken and humans. Although it is not possible to explore the role of TEBP *in vivo* since *Oxytricha* is not a genetically tractable organism, TEBP's exquisite specificity for the 3'OH on the G-overhang (Horvath *et al*., 1998) argues strongly for a function in telomere end protection. Like TEBP, *S. pombe* Pot1 specifically binds single-strand G-rich telomeric DNA, albeit not with a correspondingly strong preference for the 3'OH (Baumann and Cech, 2001; Lei *et al*., 2002). Analysis of *S. pombe* and *A. nidulans Pot1* deletion mutants have established a critical role for these proteins in chromosome end protection (Baumann and Cech, 2001; Pitt *et al*., 2004). Moreover, recent data suggest that SpPot1 may also contribute to telomere length regulation (Trujillo *et al*., 2005), and *in vivo* studies show that it is capable of binding differentially to its telomeric DNA substrate in a manner that can expose or block the terminus from elongation by telomerase (Lei *et al*., 2004). hPot1 has also been implicated in both telomere length regulation (Colgin *et al*., 2003; Loayza and de Lange, 2003) and chromosome end protection (Veldman *et al*., 2004). Here, we establish that *Arabidopsis* harbors two Pot proteins that both contribute to telomere biology.

AtPot1 contributes to telomere length homeostasis

Our data provide strong evidence that the AtPot1 protein plays a role in telomere length regulation. First generation transgenic plants over-expressing the AtPot1 C-terminus (P1ΔN) harbor telomeres that are up to 1.5 kb shorter than wild type. In contrast to telomerase mutants, telomere length does not decline further in subsequent generations. Instead, in several second generation transformants, many telomeres returned to wild type length, indicating that AtPot1 influences telomere length homeostasis. Since the truncated AtPot1 polypeptide (P1ΔN) lacks the OB fold-containing N-terminus, its effect is unlikely to be mediated by direct DNA binding. One possibility is that P1ΔN dislodges AtPot1-interacting proteins involved in telomerase recruitment, thereby disturbing telomere length homeostasis and tipping the balance in favor of telomere shortening. Alternatively, as has been shown for the human $P1^{\Delta OB}$ protein (Loayza and de Lange, 2003), this truncated *Arabidopsis* polypeptide may remain associated with telomeres via protein-protein interactions, directly influencing the telomere-counting mechanism proposed to regulate the length of the telomere tract (Marcand *et al*., 1997b; Smogorzewska *et al*., 2000b). Although one can envision ways in which AtPot1 could act as a negative regulator of telomere homeostasis, the simplest explanation is that AtPot1 is a positive regulator of telomere length. AtPot2 may also contribute to the positive regulation of telomere length, as P2ΔC mutants display shorter telomeres. However, our data suggest that AtPot1 may be specialized to serve this role.

One surprising outcome of our study is that telomeres in plants over-expressing P1ΔN suffer the opposite fate of telomeres in human cells that over-express the corresponding region of hPot1 ($P1^{\Delta OB}$) (Colgin *et al*., 2003; Kelleher *et al*., 2005; Loayza

and de Lange, 2003). Disparate behavior of telomere proteins is not without precedent. The yeast Rif1 protein regulates telomere length in wild type cells (Hardy *et al*., 1992), while human Rif1 is only present at dysfunctional telomeres (Silverman *et al*., 2004). Similarly, Ku deficiency in yeast leads to telomere shortening (Boulton and Jackson, 1998; Gravel *et al*., 1998; Polotnianka *et al*., 1998), but in *Arabidopsis* culminates in significant telomere elongation (Gallego *et al*., 2003; Riha *et al*., 2002). Most notably, no sequence homologues for hPot1 interacting factors Pip1/PTOP and TIN2 exist in the *Arabidopsis* and *S. pombe* genomes. Thus, a different ensemble of Pot-associated proteins remains to be discovered in these organisms.

AtPot2 functions in chromosome end protection

Our data implicate AtPot2 in chromosome end protection. A significant fraction of primary transformants over-expressing the AtPot2 N-terminus (P2ΔC) resemble terminal generation telomerase-deficient mutants with severe morphological defects, sterility and a high incidence of anaphase bridges (Riha *et al*., 2001). It is possible that P2ΔC competes with the endogenous AtPot2 for telomeric DNA, dislodging AtPot2 altogether or reducing the affinity of AtPot2-interacting factors involved in chromosome end protection. A more definitive understanding of the mechanism of both AtPot1 and AtPot2 dominant-negative alleles will require localization of these proteins and their full-length counterparts to telomeres *in vivo*.

The phenotype of P2ΔC mutants is remarkably similar to loss-of-function *pot1* mutants in *S. pombe* (Baumann and Cech, 2001; Pitt *et al*., 2004). Both mutants experience an immediate onset of cytogenetic defects including chromosome missegregation and chromosome fusion. Both lose telomeric DNA. However, unlike

the situation in the *S. pombe* deletion strain, where the entire telomere tract and several kilobases of subtelomeric DNA are lost within 10 cell generations (Baumann and Cech, 2001), bulk telomeres decline by only 1-2 kb in *Arabidopsis* P2ΔC mutants and do not reach the critically shortened threshold defined for *Arabidopsis tert* mutants (Heacock *et al*., 2004). Interestingly, however, we failed to amplify chromosome fusion junctions containing telomeric DNA in P2ΔC mutants. One interpretation of these data is that the subset of chromosomes with dysfunctional telomeres that are targeted for end-joining reactions are first subjected to extensive exonucleolytic degradation.

The relatively modest decline of bulk telomere length in P2ΔC mutants may reflect a limited opportunity for nuclease action on *Arabidopsis* telomeres, as telomeres are confined to the nucleolus throughout most of the mitotic and meiotic cell cycles (Armstrong *et al*., 2001; Fransz *et al*., 2002). Alternatively, the presence of endogenous AtPot1 and AtPot2 may offer partial protection to the chromosome terminus. A third possibility is that AtPot2 exerts its capping function in concert with a TRF2-like protein. hPot1 associates with TRF2, a protein essential for chromosome end-protection in human cells (Ye *et al*., 2004a). A number of TRF-like proteins have been described in *Arabidopsis* (Chen *et al*., 2001; Hwang *et al*., 2001; Karamysheva *et al*., 2004). One or more of these could act in the same pathway with AtPot2 to facilitate chromosome capping.

Evolution of Pot proteins

The long delay in the identification of single-strand telomere binding proteins in multicellular eukaryotes is due to the rapid divergence of this class of proteins. *Arabidopsis* may not be unique in possessing more than one *Pot1* gene. Database

searches reveal the presence of two relatively similar *Pot1* genes in mouse, and the ciliate *Euplotes crassus* genome encodes two highly divergent Pot1-like proteins (Wang *et al*., 1992). Moreover, analysis of plant EST databases showed that several species harbor multiple *Pot* genes, which like the *AtPot1* and *AtPot2*, exhibit marked sequence divergence (E. Shakirov and D. Shippen, unpublished data). The striking sequence divergence of the *Arabidopsis Pot1* and *Pot2* genes, coupled with the genetic data presented here, argue that these proteins make distinct contributions to telomere biology. Hence, further analysis of these proteins should provide useful insight into the functions and evolution of single-strand telomere binding proteins.

CHAPTER IV

Arabidopsis POT1 ASSOCIATES WITH THE TELOMERASE RNP AND IS REQUIRED FOR TELOMERE MAINTENANCE*

Summary

POT1 is a single-copy gene in yeast and humans that encodes a single-strand telomere binding protein required for chromosome end protection and telomere length regulation. In contrast, *Arabidopsis* harbors multiple, divergent POT-like genes that bear signature N-terminal OB-fold motifs, but otherwise share limited sequence similarity. Here we report that plants null for *AtPOT1* show no telomere deprotection phenotype, but rather exhibit progressive loss of telomeric DNA. Genetic analysis indicates that *AtPOT1* acts in the same pathway as telomerase. *In vitro* levels of telomerase activity in *pot1* mutants are significantly reduced and are more variable than wild type. Consistent with this observation, AtPOT1 physically associates with active telomerase particles. Although low levels of AtPOT1 can be detected at telomeres in unsynchronized cells and in cells arrested in G2, AtPOT1 binding is significantly enhanced during S-phase when telomerase is thought to act at telomeres. Our findings indicate that AtPOT1 is a novel accessory factor for telomerase required for positive telomere length regulation,

---

*Reprinted by permission from Macmillan Publishers Ltd: *EMBO J* **26**, 3653-3661, 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. Copyright 2004.

and underscore the coordinate and extraordinarily rapid evolution of telomere proteins and the telomerase enzyme.

Introduction

Telomeres stabilize eukaryotic genomes by facilitating the complete replication of the chromosome terminus, and sequestering the ends from recognition by DNA damage checkpoint machinery that would otherwise lead to inappropriate engagement of recombination and DNA repair activities. Telomeres typically consist of simple G-rich repeats that terminate in a single-strand 3' extension, termed the G-overhang. The G-overhang serves as the substrate for the telomerase ribonucleoprotein (RNP) reverse transcriptase, which replenishes telomeric DNA. How telomerase engages the telomere is unknown, but its activity is modulated in cis by cell cycle-regulated interactions with resident telomeric DNA binding proteins (reviewed in Collins, 2006).

G-strand binding proteins play a crucial role in regulating telomerase access to telomeres, and in controlling other telomere-associated activities. The best characterized of these is Cdc13p from budding yeast (Nugent *et al*., 1996). Typical of this class of proteins, Cdc13p binds telomeric DNA via an oligonucleotide-oligosaccharide binding fold (OB-fold) (Mitton-Fry *et al*., 2002). Cdc13p is a multifunctional protein (reviewed in Lustig, 2001). It dynamically interacts with other constituents of the chromosome terminus and contributes to positive and negative regulation of telomerase, coupling of leading and lagging strand DNA synthesis, and protection of the C-rich strand of the chromosome terminus against nucleolytic attack.

In *Schizosaccharomyces pombe* and in higher eukaryotes the presumed ortholog of Cdc13p is POT1 (Protection of telomeres) (Baumann and Cech, 2001).

Although most organisms harbor only a single *POT1* gene, ciliates, mouse and *Arabidopsis* possess at least two of these (Hockemeyer *et al*., 2006; Jacob *et al*., 2006; Shakirov *et al*., 2005; Wang *et al*., 1992; Wu *et al*., 2006). POT1 binds telomeric DNA *in vitro* (Baumann and Cech, 2001), but its attachment to the chromosome terminus *in vivo* is mediated primarily through protein interactions in the shelterin complex (de Lange, 2005; Loayza and de Lange, 2003). Recent studies indicate that POT1 function is conveyed through its association with TPP1, another OB-fold containing protein (Houghtaling *et al*., 2004; Liu *et al*., 2004; Wang *et al*., 2007; Xin *et al*., 2007; Ye *et al*., 2004b).

The co-crystal structure of human POT1 bound to its DNA substrate indicates that the 3' terminal residues of the DNA are sequestered within the protein binding pocket, implying that hPOT1 functions to protect against nucleases and limit accessibility to telomerase (Lei *et al*., 2004). Consistent with this prediction, hPOT1 negatively regulates telomerase activity *in vitro*; this inhibition requires the DNA binding activity of hPOT1 (Kelleher *et al*., 2005; Lei *et al*., 2005). *In vitro* studies suggest that hPOT1 could also promote telomerase action at the chromosome terminus. Disruption of G-quartet structures by hPOT1 facilitates elongation by telomerase *in vitro* (Zaug *et al*., 2005). Moreover, hPOT1 stimulates unwinding of telomeric DNA by WRN and BLM helicases (Opresko *et al*., 2005), and depending on the location of POT1 binding site, hPOT1 can improve telomerase activity and processivity *in vitro* (Lei *et al*., 2005).

Genetic analysis of POT1 in fission yeast and vertebrates reveals a complex role for this protein in telomere length control. Human cells with reduced levels of POT1 display telomere elongation (Veldman *et al*., 2004; Yang *et al*., 2005; Ye *et al*., 2004b) as do mice conditionally null for POT1a (Wu *et al*., 2006). Similarly, reduction of

telomere-bound POT1 in *S. pombe* results in dramatic telomere elongation (Bunch *et al*., 2005). In contrast, over-expression studies implicate *S. pombe* and human POT1 in the positive regulation of telomere length (Armbruster *et al*., 2004; Bunch *et al*., 2005; Colgin *et al*., 2003; Liu *et al*., 2004). Thus, like Cdc13p, POT1 may contribute to both positive and negative regulation of telomere length.

POT1 is also necessary for chromosome end protection. *S. pombe pot1*$^{-}$ mutants suffer immediate and catastrophic loss of telomeric repeats, erosion of subtelomeric DNA and chromosome mis-segregation (Baumann and Cech, 2001). Depletion of vertebrate POT1 leads to a DNA damage response at telomeres (Churikov *et al*., 2006; Hockemeyer *et al*., 2005), and in chicken cells results in a rapid G2 cell cycle arrest (Churikov *et al*., 2006). Other studies of POT1 depleted mammalian cells reveal genome instability, senescence, and apoptosis (Veldman *et al*., 2004; Yang *et al*., 2005). The mouse POT1a and POT1b genes appear to be partially redundant for chromosome end protection. Although POT1b mutants are viable (Hockemeyer *et al*., 2006), conditional knockout of POT1a results in embryonic lethality (Hockemeyer *et al*., 2006; Wu *et al*., 2006). Single POT1a or double POT1a POT1b mutants exhibit a strong telomere DNA damage response, low levels of telomere fusions and endoreduplication, along with proliferative arrest and senescence (Hockemeyer *et al*., 2006). A second study implicated POT1a and POT1b in repression of non-homologous end joining and homologous recombination at telomeres (He *et al*., 2006; Wu *et al*., 2006).

Arabidopsis encodes two POT-like proteins, AtPOT1 and AtPOT2 (Shakirov *et al*., 2005), and possibly a third, AtPOT3 (Surovtseva *et al*., in preparation). In contrast to the mouse POT1a and POT1b proteins, which share 72% similarity (Hockemeyer *et*

al., 2006), the plant POT proteins are more divergent and display only 49% sequence similarity overall. AtPOT2 is implicated in chromosome end protection as over-expression of the N-terminal portion of the protein leads to severe growth and developmental defects, telomere shortening, and a high incidence of anaphase bridges and chromosome mis-segregation. AtPOT1, by contrast, contributes to telomere length regulation. Over-expression of a C-terminal fragment of AtPOT1 lacking the OB-fold motifs results in modest telomere shortening, but plants are wild type in appearance and show no signs of genome instability (Shakirov *et al*., 2005).

In this study, we examined the fate of *Arabidopsis* mutants null for *AtPOT1*. We found no evidence that AtPOT1 contributes to chromosome end protection or genome stability. Instead, AtPOT1 is required for positive regulation of telomere length: *pot1* mutants display progressive telomere shortening at the same rate as telomerase null plants. Notably, *in vitro* telomerase activity levels are reduced in *pot1* mutants, but not abolished. Finally, we show that AtPOT1 physically associates with the telomerase RNP and is enriched at telomeres during S-phase. Thus, AtPOT1 appears to be a novel telomerase accessory factor that promotes its activity *in vitro* and *in vivo.*

Materials and methods

Mutant lines

The *pot1-1* allele was identified in ALPHA population of T-DNA insertion lines at the University of Wisconsin *Arabidopsis* Knock-out facility. The collection was screened using primers 5'-TTTGTACTGGCCTCTCCAAGGTTCACCAT-3' and 5'-CATTTTATAATAACGCTGCGGACATCTAC-3' according to the protocol available at http://www.biotech.wisc.edu/Arabidopsis/Index2.asp. The *pot1-2* line was identified by

screening a pooled genomic DNA collection (ABRC #CD10-A) from the D. Weigel T-DNA lines using primers 5'-CGGGATCCCACCCAGAAGATACTAAGATG-3' and 5'-TTGACCATCATACTCATTGCTG-3'. *tert* and *ku70* mutants and plant growth conditions are as described (Riha *et al*., 2001; Riha *et al*., 2002). All crosses were made between plants heterozygous for the desired mutations. Double- and triple-heterozygous F1 plants were identified by PCR genotyping and then self-propagated to F2 to obtain single-, double- and triple-homozygous mutants and their wild-type siblings. F2 plants (G1) were self-propagated for several generations. Independent lines from at least two F2 plants were established and analyzed for each genotype.

Complementation was performed as described in Figure 24. Wild type and G3 *pot1-1* were used to establish callus. Callus initiation and maintenance were performed as described (Watson *et al*., 2005) with slight modifications. Seeds were germinated on 0.5X Murashige and Scoog (MS) medium plates supplemented with 3% sucrose and 2.8 g/liter phytagel. Roots were harvested at three weeks, finely chopped, and placed on 1X MS medium plates supplemented with 2 mg/liter 2,4-D, 0.05 mg/liter kinetin, 3% sucrose and 2.8 g/liter phytagel (CIM). Callus was grown on CIM at 25^{o}C in the dark and transferred to fresh medium every four weeks.

RT-PCR analysis, telomere analysis, TRAP assays and cytogenetics

Total RNA was extracted from plant tissue using Tri Reagent solution (Sigma). Reverse transcription was performed using Superscript III reverse transcriptase (Invitrogen) as described (Shakirov *et al*., 2005). To evaluate expression of the regions flanking the T-DNA insertion in the *pot1-1* allele, primer 1 (5'-GGATCCATGGCGAAGAAGAGAGAGAGTCCCAAGCTCATCA-3') and primer 2 (5'

GCTCTAGACTTGATCTCTCTCAAGAAGGA-3') were used. To analyze the *pot1-2* allele, we used primer 1 and primer 3 (5'-TACCTCGAGCTAGATTAGGCTATCAGAGA-3'). DNA from individual whole plants was extracted as described (Cocciolone and Cone, 1993). TRF analysis was performed using *Tru1*I (Fermentas, Hanover, MD) restriction enzyme and a ^{32}P 5' end–labeled $(T_3AG_3)_4$ oligonucleotide probe (Fitzgerald *et al*., 1999). Subtelomeric TRF analysis was conducted using a 2R probe (Shakirov and Shippen, 2004) or a probe for 1L generated with 5'-ACGCTTGTCATCTCATCTCT-3' and 5'-CGGGATCTTTGGTGTTTCTC-3'. Telomere fusion PCR and PETRA was performed as described (Heacock *et al*., 2004). TRAP assays were conducted on plant tissues as in (Fitzgerald *et al*., 1996) using protein extracts prepared from a single wild type or *pot1* inflorescence unless otherwise indicated. Anaphase spreads were prepared from pistils and stained with DAPI (4',6'-diamidino-2-phenylindole) as discussed in (Riha *et al*., 2001).

Antibodies, western blotting and immunoprecipitation

The P1-R polyclonal antibody was raised in rabbits against full-length recombinant AtPOT1 protein expressed in *E. coli* (Covance). The P1-P1 and P1-P2 peptide antibodes were raised in rabbits against the N'-CSDENRRHHQVLLTLEDST and N'-AAYPWQVEDFCSDENRRHHQVLLT peptides, respectively, and affinity purified (Covance). Western blots were conducted with primary antibodies and peroxidase-conjugated light chain-specific mouse anti-rabbit secondary antibodies (Jackson Immunoresearch). For immunoprecipitation of endogenous proteins from suspension culture, protein extracts were prepared in buffer containing 50 mM Tris-HCl, pH7.4, 10 mM NaCl, 10 mM $MgCl_2$, 1 mM EDTA, 10% glycerol, 1 mM PMSF, 1 mM DTT, protease

inhibitors (Roche). For immunoprecipitation of endogenous proteins from four-day-old seedlings, protein extracts were prepared according to (Fitzgerald *et al*., 1996). Extracts were diluted 1:5 in buffer W100 (20 mM TrisOAc, pH 7.5, 10% glycerol, 1 mM EDTA, 5 mM $MgCl_2$, 200 mM NaCl, 100 mM KGlu, 1% NP40, 0.5 mM Na Deoxycholate, 1 mM DTT), precleared with protein A agarose (Pierce) and subjected to immunoprecipitation with POT1 antibodies. Following immunoprecipitation, the beads were washed three times with buffer W300 (same as buffer W100, but 300 mM KGlu) and two times with TMG buffer (10 mM TrisOAc, pH 8.0, 1 mM $MgCl_2$, 10% glycerol, 1 mM DTT) and used for TRAP assay. In peptide competition experiments, 100X molar excess of the P1-P1 peptide or AtPOT2 peptide (N'-DDYKFLRIQDAFKALHLHVNC) was added during immunoprecipitation step. For salt stability experiments, the NaCl concentration in the input was adjusted by addition of 5M NaCl to suspension culture extracts. For immunoprecipitation of recombinant proteins, AtPOT1 and AtPOT2 were expressed in rabbit reticulocyte lysate (Promega) in the presence of ^{35}S-methionine. Following immunoprecipitation, the signal was quantified using ImageQuant Software. Immunoprecipitation efficiency was calculated as a ratio of the immunoprecipitated signal versus input.

Cell culture syncronization and FACS

MM2d *Arabidopsis* cell suspension culture was maintained as described in (Menges and Murray, 2002). For synchronization, the original protocol was employed with the following modifications. 100 ml of four-day-old culture was divided into five 20 ml aliquots, diluted 1:5 with fresh medium and then blocked in G1/early S phase for 21 h with 12 µg/ml of aphidicolin (A.G. Scientific). Cells were filtered through Miracloth

(Calbiochem), washed twice with 500 ml of fresh media, and resuspended in 100 ml of fresh media. Aliquots were taken at various time points for DNA content analysis and immediately frozen. DNA content was analyzed by flow cytometry. The cultured cells were chopped with a razor blade, resuspended in homogenization buffer, filtered through a 20 µm nylon mesh, treated with 10 µg/ml of RNase A, and stained with 50 µg/ml of propidium iodide. Samples were run on a Becton-Dickinson FACSCalibur at 488 nm and analyzed using CellQuest (Becton-Dickinson) and ModFit LT (Verity) programs.

Chromatin immunoprecipitation

ChIP was performed essentially as described by (Leibfried *et al*., 2005) with some modifications. The equivalent of 100 ml of four-day-old unsynchronized or synchronized cell suspension culture (~3 g of dry material) was fixed for 1 h in 1% formaldehyde followed by quenching for 10 min in 125 mM glycine. After vacuum filtration through Miracloth, cells were resuspended and washed in 500 ml of PBS, and filtered again before storage at -80°C. Cells were ground in liquid N_2 and then ChIP was performed using P1-R or P1-P2 antibody in a 1:100 dilution. The elution products were subjected to slot blot (Hybond N+, Amersham) and hybridized using a $(TTTAGGG)_4$ telomeric probe. Blots were stripped and rehybridized with a combination of radiolabeled 5S (5'-TTGCAGAATCCCGTGAACCATCGAGT-3') and 18S rDNA (5'-TGGAGCCTGCGGCTTAATTTGACTCA-3') oligo-probes to monitor the specificity of immunoprecipitation. For quantification, the fold enrichment was determined as a ratio of the hybridization signal obtained with the POT1 antibody versus the pre-immune sera control. rDNA sequences were used as a negative control.

Results

Plants null for AtPOT1 do not exhibit genome instability

To identify an *Arabidopsis* line null for *AtPOT1*, we screened T-DNA collections from the University of Wisconsin *Arabidopsis* Knock-out Facility. Analysis of the ALPHA population uncovered a mutant with an insertion in the first intron of *AtPOT1* (Figure 21 and Figure 22A). This line was designated *pot1-1*. In the D. Weigel collection, we found a second *AtPOT1* allele (*pot1-2*), bearing a T-DNA in the seventh exon (Figure 21 and Figure 22A). To determine if these insertions disrupt *AtPOT1* gene expression, RT-PCR was performed using primers flanking the insertion sites (Figure 21 and Figure 22B). No PCR products were generated in reactions with cDNA from the mutant plants (Figure 22B, lanes 2 and 4), confirming that expression of the full-length *AtPOT1* mRNA was abolished in *pot1-1* and *pot1-2* mutants.

To monitor AtPOT1 protein, antibodies were raised against two peptides corresponding to a segment in the C-terminus of AtPOT1 protein (P1-P1 and P1-P2) (Figure 22A) and against a full-length recombinant AtPOT1 protein (P1-R). All three antibodies detected recombinant AtPOT1 by western blotting (data not shown), and each immunoprecipated the recombinant protein with ~2% IP efficiency (Figure 22C; data not shown). Importantly, P1-P1 and P1-P2 detected a 55 kDa protein that corresponds to endogenous AtPOT1 protein in extracts from wild type seedlings and callus, but not from *pot1-1* mutants (Figure 22D; data not shown). We conclude that *pot1-1* and likely *pot1-2* (see below) are null for *AtPOT1*.

In striking contrast to yeast and vertebrate cells deficient in POT1, *Arabidopsis pot1* mutants appeared morphologically indistinguishable from wild type and showed no

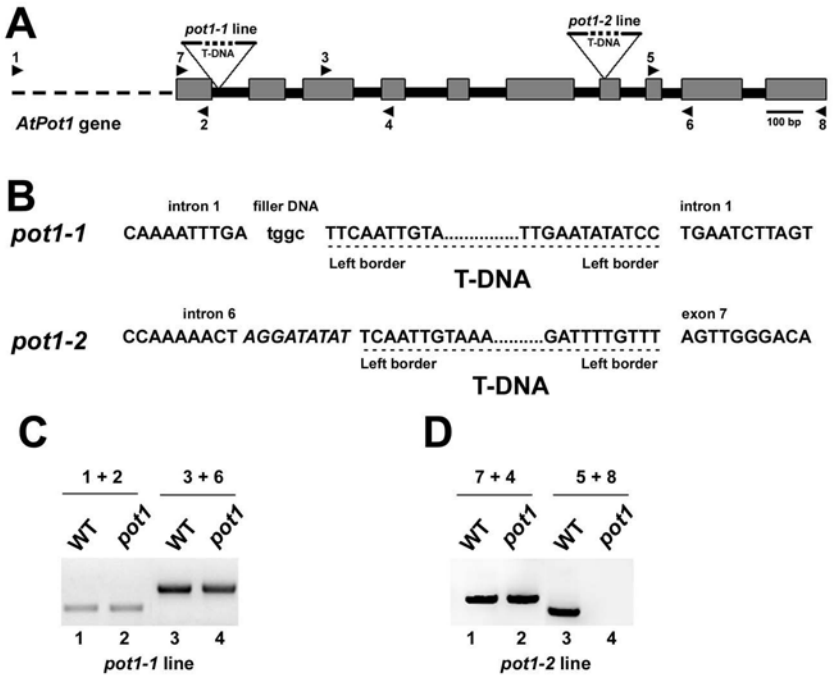


Figure 21. Analysis of T-DNA insertions in the *AtPOT1* gene. (A) Position of *pot1-1* and *pot1-2* insertions. Primers used for RT-PCR analysis are shown by arrowheads. Dashed line represents the 5' UTR. (B) Sequence analysis of *pot1-1* and *pot1-2* T-DNA insertion sites. For both mutants, the insertion sites are larger than 10 kb with two T-DNAs inserted in a head-to-head orientation, left borders facing outward. Lower case letters represent filler DNA, presumably added during T-DNA integration. Italics denote a region of microhomology shared by both *AtPot1* genomic sequence and the left border of T-DNA. (C and D) RT-PCR analysis of the disrupted *AtPOT1* gene in *pot1-1* and *pot1-2* mutants. Although *AtPOT1* transcripts from a region upstream of the T-DNA can be amplified in wild type (WT) (C, lane 1) and *pot1-1* mutant (D, lane 2) samples, it encodes only a small (42 amino acid) polypeptide which is likely to be non-functional. For *pot1-1*, a 3' end transcript is also detected (C, lane 4). Further analysis demonstrated that this transcript in *pot1-1* retained intron 1 and portions of T-DNA, resulting in the generation of in-frame stop codons prior to exon 2 (data not shown). No 3' end *AtPOT1* transcripts were detected in RT-PCR reactions with *pot1-2* mutants (D, lane 4). Although a 5' transcript is generated in *pot1-2*, the phenotype of the mutant indicates that this is also a null allele (see text for details).

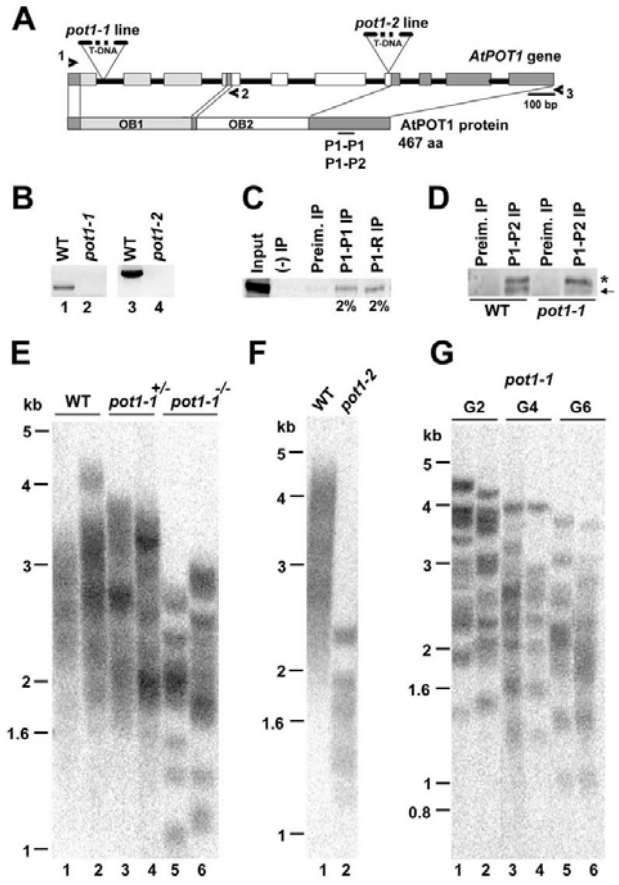


Figure 22. Telomere phenotypes in AtPOT1-deficient *Arabidopsis*. (A) Genomic map and coding region of the *AtPOT1* locus. The position of T-DNA insertions in the *pot1-1* and *pot1-2* alleles are shown. The position of the peptide used to raise P1-P1 and P1-P2 antibodies is indicated. (B) RT-PCR analysis of the *AtPOT1* gene expression in *pot1-1* and *pot1-2* mutants. Primer pairs 1-2 and 1-3 (shown as arrowheads in panel A) were used to analyze gene expression. (C) Immunoprecipitation of recombinant ^{35}S-labeled AtPOT1 protein. Immunoprecipitation efficiencies are indicated. (D) Detection of endogenous AtPOT1 protein in wild type and *pot1-1* callus. Arrow indicates the 55 kDa endogenous AtPOT1 protein immunoprecipitated from wild type callus. Asterisk indicates a nonspecific cross-reacting protein. (E) TRF analysis of DNA from six siblings segregating from a heterozygous *pot1-1* parent. (F) TRF analysis of a *pot1-2* mutant. (G) Multi-generational TRF analysis of *pot1-1*. DNA samples from two individual *pot1-1* plants from the second (G2), fourth (G4) and sixth (G6) generation of self-pollination were analyzed. Blots shown in E, F and G were hybridized with a radiolabeled telomeric DNA probe.

decrease in fertility or perturbation in growth and development for the six generations they were propagated. Furthermore, chromosome ends were refractory to nuclease attack and non-homologous end-joining in the absence of AtPOT1. No anaphase bridges were observed in first (G1) or second (G2) generations of *pot1-1* mutants (Table 3; data not shown). The more sensitive telomere fusion PCR assay (Heacock *et al*., 2004) also failed to detect an increased frequency in chromosome end-joining reactions in *pot1-1* (data not shown). Thus, AtPOT1 is dispensable for chromosome end protection in *Arabidopsis*.

AtPOT1 is required for telomere length maintenance in vivo

To examine telomere length in *pot1* mutants, terminal restriction fragment (TRF) analysis was performed on plants segregated from self-pollination of a heterozygous *pot1-1* parent. As expected, telomeres in wild type siblings appeared as a homogeneous smear of products ranging from 1.6 to 4.5 kb (Figure 22E, lanes 1, 2). As for *AtTERT* (Fitzgerald *et al*., 1999b), *AtPOT1* is not haploinsufficient for telomere maintenance in *Arabidopsis*; plants heterozygous for the T-DNA insertion exhibited a wild type telomere profile (Figure 22E, lanes 3 and 4).

Strikingly, telomere tracts in *pot1-1* were much shorter than in wild type and showed a more discrete banding pattern (Figure 22E, lanes 5 and 6). To determine whether disruption of the *AtPOT1* gene was responsible for the telomere phenotypes, TRF analysis was performed on *pot1-2* mutants. Telomeres were significantly shorter in *pot1-2* than in wild type, or even *pot1-1* (Figure 22F, lane 2). Since the *pot1-2* mutant was homozygous when we identified it, we suspect that this line had already been propagated at the *Arabidopsis* stock center for several generations in the absence of

Table 3. Incidence of anaphase bridges in *pot1-1* mutants.

Genotype		Phenotype*	No. of anaphases scored	With bridges	Without bridges	% Bridges
WT		WT	315	0	315	0
pot1-1	G2	WT	353	0	353	0
pot1-1 tert	G2	WT	314	0	314	0
pot1-1 ku70	G2	II	322	64	258	20
	G3	T	450	194	256	43
ku70 tert	G2	II	387	74	313	19
	G3	T	442	180	262	41
pot1-1 ku70 tert	G2	II	340	58	282	17
	G3	T	434	155	279	36

* The severity of morphological phenotypes was assayed as described (Riha *et al*., 2001). T=terminal phenotype.

AtPOT1 prior to our analysis, leading to more substantial loss of telomeric DNA than in *pot1-1*. Complementation experiments provided further verification that AtPOT1 depletion caused telomere shortening. Plants heterozygous for *pot1-1* were transformed with the wild type *AtPOT1* coding region under control of the constitutive CaMV *35S* promoter. In plants expressing the *35S::AtPOT1* transgene telomeres, particularly the shortest ones in the population, were returned to the wild type length (Figure 23A, lanes 2 and 3). A second complementation experiment performed with *pot1 ku70* double mutants confirmed that AtPOT1 is required for telomere maintenance (Figure 23B; see below).

Telomeres in pot1 *mutants shorten at the same rate as in* tert *mutants*

We followed the fate of telomeres in *pot1* mutants for several generations and found that telomere length in *pot1-1* decreased progressively with each generation (Figure 22G). The ever-shorter-telomere phenotype and sharp TRF banding profile were strikingly similar to the phenotype associated with *tert* mutants, which lose 200-500 bp of telomeric DNA per plant generation (Fitzgerald *et al*., 1999; Riha *et al*., 2001).

To determine if the rate of telomere shortening in *pot1-1* was the same as in *tert*, we used a parent-progeny analysis to measure the rate of bulk telomere loss. DNA extracted from first generation (G1) parents (P) homozygous for the *pot1-1* or *pot1-2* allele and their progeny (G2) was subjected to TRF analysis. For *pot1-1*, bulk telomeres declined by approximately 200-500 bp, while for *pot1-2*, a loss of approximately 200 bp was observed (Figure 24A and B). To obtain a more accurate estimate of the telomere shortening rate, individual telomeres were examined using subtelomeric TRF analysis (Shakirov and Shippen, 2004). Using probes specific for the

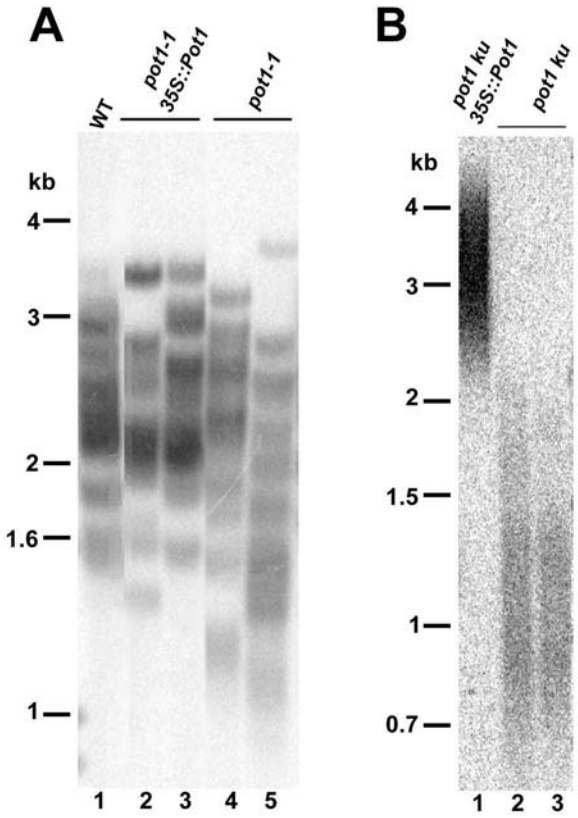


Figure 23. Complementation analysis of *pot1-1* and *pot1-1 ku70* mutants. (A) Complementation of *pot1-1* mutants. *AtPOT1* was transformed into heterozygous *pot1-1* plants. Homozygous *pot1-1* segregants containing the *35S::AtPOT1* transgene, as well as their untransformed wild type and *pot1-1* siblings, were subjected to TRF analysis. Similar to the situation in segregating WT plants (lane 1), homozygous *pot1-1* progeny carrying the *35S::AtPOT1* transgene showed no signs of telomere shortening (lanes 2-3), while the shortest telomeres in the untransformed homozygous *pot1-1* siblings were significantly diminished relative to the shortest telomeres in wild type (lanes 4-5). (B) Complementation of *pot1-1 ku70* mutants. *AtPOT1* was transformed into plants homozygous for *pot1-1* and heterozygous for *ku70*. Double homozygous *pot1-1 ku70* segregants containing the *35S::AtPOT1* transgene, as well as their untransformed *pot1-1 ku70* siblings, were subjected to TRF analysis. While telomeres in untransformed homozygous *pot1-1 ku70* plants remained short and exhibited a very heterogeneous profile (lanes 2 and 3), significant telomere elongation was restored in homozygous *pot1-1 ku70* seedlings carrying the *35S::AtPOT1* transgene (lane 1).

South (right) arm of chromosome two (2R), or the North (left) arm of chromosome one (1L) (Figure 24C), only a single discrete band was detected in the parent and its progeny, likely reflecting the coordinate regulation of telomere length on homologous chromosomes throughout plant development (Shakirov and Shippen, 2004). For both telomeres, the length decreased by approximately 200-300 bp relative to the parent (Figure 24C).

To further examine the rate of telomere shortening in *pot1* mutants, we performed primer extension telomere repeat amplification (PETRA) (Heacock *et al*., 2004). In this assay, telomeres are amplified in a PCR reaction using primers directed at the G-overhang and a unique subtelomeric sequence. PETRA showed a decline of approximately 200 bp on the 2R and 5R telomeres in both *pot1-1* and *tert* mutants (Figure 24D and E; data not shown). The same degree of shortening occurred in *pot1-2* (data not shown). We conclude that disruption of *AtPOT1* leads to a progressive loss of telomeric DNA that proceeds at the same rate as in *tert* mutants.

Telomeres become critically shortened in G6 *tert* mutants giving rise to end-to-end fusions and genome instability (Riha *et al*., 2001). Because our *pot1-1* mutants were derived from the WS ecotype, which naturally has longer telomeres than the Columbia ecotype (Shakirov and Shippen, 2004), from which the *tert* mutant was obtained, it is not surprising that G6 *pot1-1* mutants do not yet show signs of genome instability. Assuming a telomere shortening rate of 200 bp/plant generation, we expect two or three additional generations are required for some *pot1-1* telomeres to become critically shortened (Heacock *et al*., 2004).

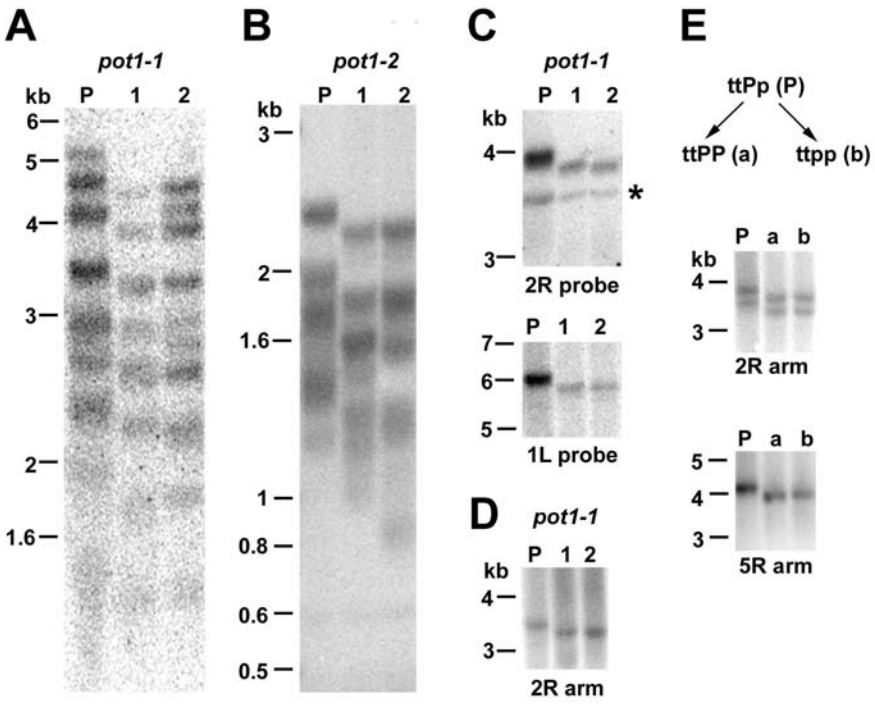


Figure 24. Parent-progeny analysis reveals the same rate of telomere shortening in *pot1-1*, *pot1-2* and *tert* mutants. (A and B) TRF analysis of bulk telomeric DNA from *pot1-1* and *pot1-2* parents (P) and two progeny (1 and 2) using a telomeric probe. (C) Subtelomeric TRF analysis of DNA from a *pot1-1* parent and two progeny. DNA blots were hybridized with a probe corresponding to unique subtelomeric regions on 2R and 1L chromosome arms. The asterisk indicates a cross-hybridizing band. (D) PETRA analysis of the 2R telomere in a *pot1-1* parent and two progeny. (E) PETRA analysis of the 2R and 5R telomeres in a parent homozygous for *tert* and heterozygous for *pot1-1* (ttPp) and its *tert* (ttPP) (a) and *pot1-1 tert* (ttpp) (b) progeny. The two PETRA bands detected in the 2R reaction may represent different size telomeres on homologous chromosomes or two populations of cells (Shakirov and Shippen, 2004). A telomeric probe was used to detect PETRA products.

AtPOT1 *and* AtTERT *act in the same genetic pathway*

To determine whether *AtPOT1* and *AtTERT* act in the same genetic pathway, plants heterozygous for *pot1-1* were crossed to plants heterozygous for *tert*. Double heterozygous mutants from F1 were allowed to self-pollinate to generate F2 progeny. As shown in Figure 25A, telomeres of the same length and sharp banding profile were found in *pot1-1 tert* as in their *tert* and *pot1-1* siblings. PETRA and TRF parent-progeny analysis confirmed that telomeres in *pot1-1 tert* mutants shortened at the same rate as in either single mutant (Figures 24E and 25B).

If AtPOT1 is required for telomerase function *in vivo*, its contribution should be especially obvious in a genetic background where telomerase generates ultra-long telomere tracts. KU is a negative regulator of telomere length in *Arabidopsis*, and mutants deficient in KU70 or KU80 undergo telomerase-dependent expansion to more than twice the normal length in a single generation (Gallego *et al*., 2003; Riha *et al*., 2002) (Figure 25C, lanes 3 and 4). In contrast, *ku70 tert* double mutants display accelerated telomere shortening and a precocious onset of genome stability (Riha and Shippen, 2003a). To further investigate the role of AtPOT1, we generated *pot1 ku70* mutants. In contrast to their *ku70* siblings, *pot1 ku70* mutants failed to elongate their telomeres (Figure 25C, lanes 7 and 8), and exhibited a heterogeneous profile of TRF products similar to that seen in *ku70 tert* (Figure 26B, lane 3). Notably, telomeres in *pot1 ku70* mutants were significantly elongated when an exogenous copy of *AtPOT1* was introduced (Figure 23B), confirming that AtPOT1 is required for telomere elongation in the absence of KU.

We also found that degree of telomere shortening was the same in *pot1 ku70* and *ku70 tert* double mutants as in triple *pot1 ku70 tert* mutants. Parent-progeny

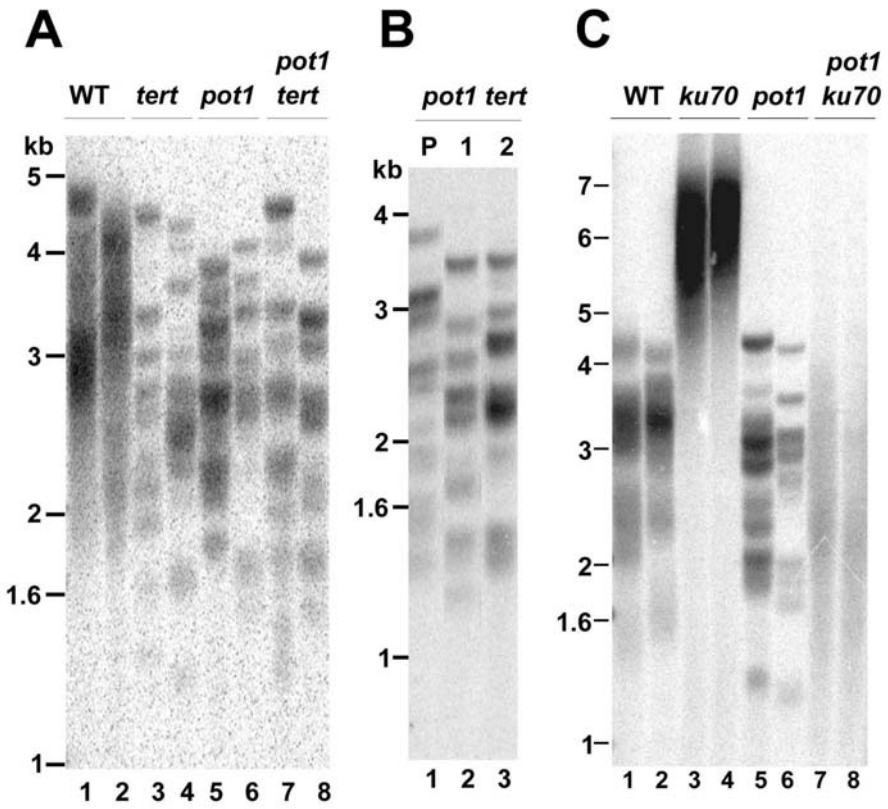


Figure 25. AtPOT1 functions in the telomerase pathway. (A) TRF analysis of *pot1-1 tert* mutants. Results for eight progeny (two for each genotype) that were segregated from a parent heterozygous for *pot1-1* and *tert* are shown. (B) TRF analysis of telomeres in *pot1-1 tert* parent (P) and its two progeny (1 and 2). (C) TRF analysis of *pot1-1 ku70* mutants. Results for progeny segregated from a parent heterozygous for *pot1-1* and *ku70* are shown. Two different progeny were analyzed for each genotype. The blot was hybridized with a telomeric DNA probe.

analysis confirmed that plants from all three genotypes lost approximately the same amount of telomeric DNA from G1 (Figure 26B, lanes 3, 4 6) to G2 (Figure 26C). All three mutants reached the terminal phenotype in G3. The incidence of anaphase bridges in *pot1 ku70, tert ku70* and *pot1 ku70 tert* mutants in G3 were the same (Table 3). These genetic data reinforce the notion that AtPOT1, like AtTERT, does not contribute to chromosome end protection. We conclude that *AtPOT1* acts in the same genetic pathway as telomerase and is specialized for telomere length maintenance *in vivo.*

POT1 is a component of the telomerase RNP required for maximal activity in vitro

We considered the possibility that AtPOT1 is required for telomerase enzyme activity. TRAP assays were performed in parallel with extracts from wild type and *pot1* seedlings. As shown in Figure 27A (lanes 1, 3-5), robust telomerase activity was reproducibly detected in extracts from wild type plants and from suspension culture. In contrast, telomerase levels were reduced and more variable in both *pot1-1* and *pot1-2* mutants (Figure 27A, lanes 6-12). Extract mixing experiments indicated that the reduction in enzyme activity was not due to the presence of a soluble PCR inhibitor (data not shown). Interestingly, although the *in vitro* levels of telomerase activity varied among *pot1* mutants, the progressive telomere shortening phenotype observed *in vivo* was extremely consistent among the dozens of *pot1-1* and *pot1-2* mutant plants we examined.

Most TRAP reactions carried out with *pot1* mutants showed reduced, but detectable levels of enzyme activity levels (Figure 27A, lanes 6-8 and 10-12). Titration of such samples suggested that telomerase activity was decreased by approximately

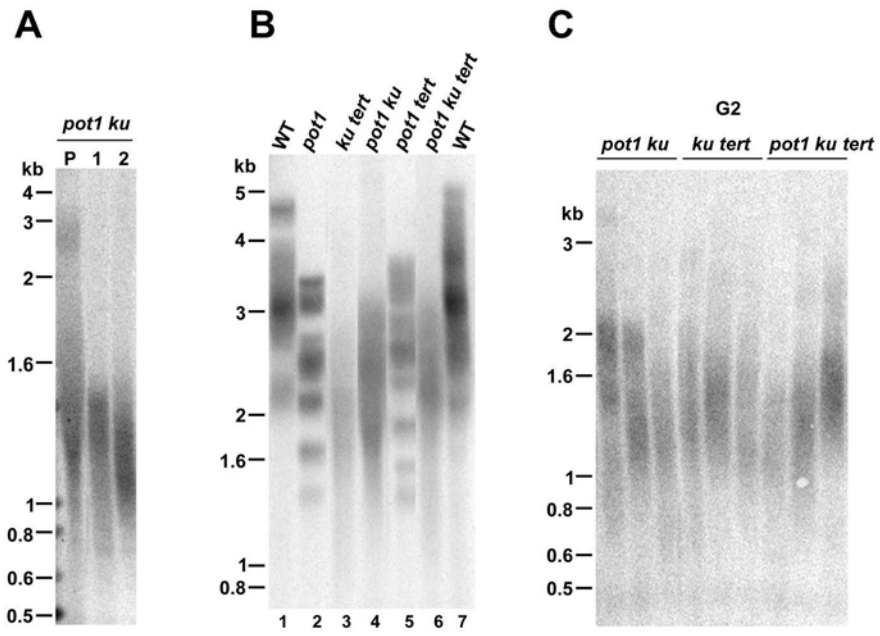


Figure 26. AtPOT1 is required for telomerase-mediated telomere elongation in the absence of KU70. (A) Parent (G1) and two progeny (G2) TRF analysis of *pot1-1 ku70* plants. Telomeres in *pot1-1 ku70* mutants decline at the same accelerated rate as in *ku70 tert* (Riha and Shippen, 2003a). (B) TRF analysis of triple *pot1-1 ku70 tert* mutants and the three double mutant combinations segregating from a parent heterozygous for *pot1-1*, *ku70* and *tert* alleles. Telomere length is approximately the same in G1 for all three double mutants and the triple. (C) TRF analysis of individual second generation (G2) *pot1 ku70*, *ku70 tert* and *pot1 ku70 tert* progeny of G1 plants shown in panel B. Telomere shortening is not accelerated in triple *pot1 ku70 tert* mutants relative to double *pot1 ku70* or *ku70 tert* mutants. All blots were hybridized with a telomeric DNA probe.

10-fold (Figure 27B, compare lanes 3 and 6). A more detailed understanding of AtPOT1's role in *Arabidopsis* telomerase biochemistry will require the development of a conventional (non-PCR based) primer extension assay to monitor the catalytic properties of the enzyme, a goal that has thus far proven elusive. Nevertheless, our current data allow us to conclude that AtPOT1 promotes telomerase action *in vitro*, but is not absolutely essential for its biochemical activity.

Next we looked for a physical interaction between AtPOT1 and the telomerase RNP. We could not detect a direct association between recombinant AtPOT1 and TERT by co-immunoprecipitation or by yeast two-hybrid assay (Y. Surovtseva, M. Jasti and D. Shippen, unpublished data). However, since TERT is the only *Arabidopsis* telomerase subunit isolated so far, AtPOT1 could contact another component of the RNP. To test this idea, immunoprecipitations were performed with POT1 antibody on extracts from wild type seedlings and asynchronous *Arabidopsis* cell culture (Menges and Murray, 2002). We confirmed that telomeres in this cell line fall within the wild type range (Figure 28A and B) and chromosome ends are protected against end-joining reactions (Figure 28C). Moreover, a 55 kDa band corresponding to endogenous AtPOT1 was observed by western blotting following immunoprecipitation of cell culture extracts (Figure 28D).

TRAP assays conducted on immunoprecipitates in the absence of antibody or with pre-immune serum did not generate PCR products (Figure 27C, lanes 2 and 3; Figure 27D, lanes 2 and 5). However, telomerase activity could be immunoprecipitated from both seedlings and cell culture using all three of the POT1 antibodies (Figure 27C, lanes 4 and 5; Figure 27D, lane 3). The specificity of the AtPOT1 interaction with telomerase was demonstrated in three ways. First, immunoprecipitation of extracts

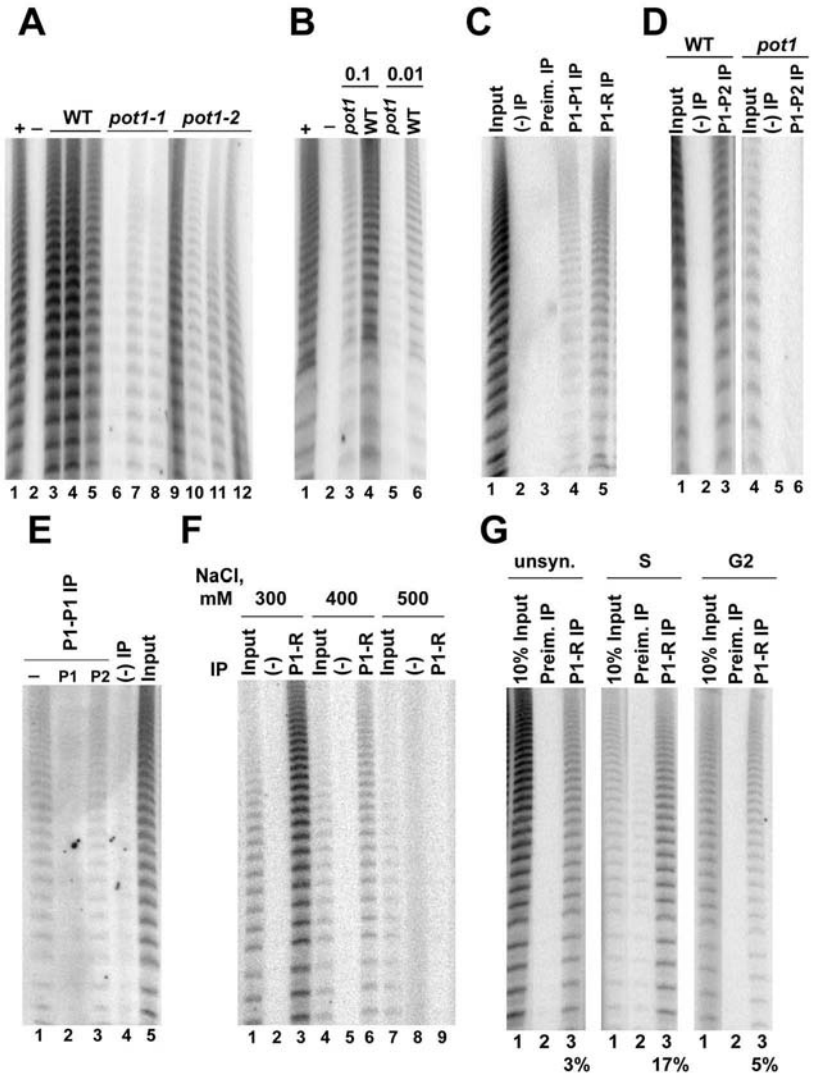


Figure 27. AtPOT1 interacts with the telomerase RNP. (A) TRAP assay results for wild type (WT), *pot1-1* and *pot1-2* flowers. (B) TRAP assay results for wild type and *pot1-2* mutant flowers. 10X and 100X dilutions of protein extracts were used for TRAP as indicated. +, positive control. (C) TRAP assay following P1-P1 or P1-R AtPOT1 immunoprecipitation. (D) TRAP assays with P1-P2 immunoprecipitates from wild type and four-day-old *pot1-2* seedlings extracts. (E) TRAP assay results for P1-P1 antibody immunoprecipitates with no peptide added (-), 100X excess of P1-P1 peptide (P1), or 100X excess of a nonspecific AtPOT2 peptide (P2). (F) TRAP assays with eluates from P1-R immunoprecipitation in the presence of NaCl. (G) TRAP assays with P1-R immunoprecipitates from unsynchronized suspension culture (unsyn.) or S-phase and G2-phase synchronized cells. The relative amount of precipitated active telomerase is indicated. Immunoprecipitation with no antibody added ((-) IP) or with pre-immune serum (Preim. IP) was used as a negative control.

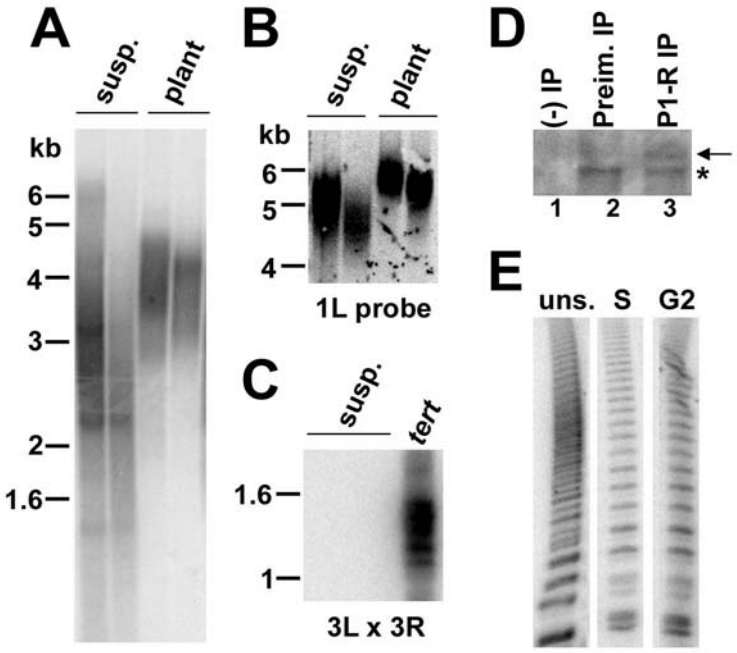


Figure 28. Characterization of telomeres and telomerase in *Arabidopsis* suspension culture. (A) TRF analysis of DNA from *Arabidopsis* suspension culture cells and *Arabidopsis* plant tissue. Telomeres in suspension culture display more heterogeneous profile, but bulk telomere length is similar to telomeres in plant tissue. (B) Subtelomeric TRF analysis of DNA from suspension culture cells and plant tissue. The blot was hybridized with a probe corresponding to a unique subtelomeric region on 1L chromosome arm. (C) Telomere fusion PCR analysis of *Arabidopsis* suspension culture. Primers corresponding to subtelomeric regions on the 3L and 3R chromosome arms were used to amplify telomeric fusions. DNA from G6 *tert* mutants was used as a positive control. (D) Detection of endogenous AtPot1 in wild type plants and in *Arabidopsis* suspension culture. AtPOT1 was immunoprecipitated using P1-R, and was detected by Western blot analysis using P1-P1 antibody. Arrow indicates the 55 kDa endogenous AtPOT1 protein. Asterisk indicates a nonspecific cross-reacting protein. (E) TRAP assay on protein extracts from unsynchronized suspension culture cells and cells synchronized in S and G2. Telomerase activity levels are approximately the same in all samples.

from *pot1* mutant seedlings failed to bring down telomerase activity (Figure 27D, lane 6). Second, addition of a 100-fold excess of the P1-P1 peptide to cell culture extracts during the immunoprecipitation with the P1-P1 antibody dramatically decreased the TRAP signal, while a nonspecific AtPOT2 peptide of a similar length failed to compete (Figure 27E, lanes 2 and 3). Third, the AtPOT1 interaction with telomerase was stable in high salt; the association persisted in up to 400 mM NaCl (Figure 27F, lane 6). Since telomerase activity is strongly inhibited in salt concentrations greater than 450 mM (Figure 27F, lane 7), the AtPOT1 interaction may be even more robust.

To investigate whether AtPOT1 association with telomerase is cell cycle regulated, immunoprecipitation was performed on synchronized cell extracts (Menges and Murray, 2002). The level of telomerase activity is approximately the same in unsynchronized cultured cells as in cells arrested in S or G2 (Figure 28E). While POT1 antibodies immunoprecipitated telomerase activity at all of the time points examined (Figure 27G, lane 3 in all panels), in three separate experiments TRAP products precipitated from S phase cell extracts were significantly increased relative to unsynchronized cells (avg = 4.5-fold). Hence, AtPOT1 shows a dynamic interaction with the telomerase RNP.

AtPOT1 dynamically associates with telomeres in vivo

Mammalian and fission yeast POT1 bind telomeric DNA *in vitro* (Baumann and Cech, 2001; Wu *et al*., 2006). To investigate AtPOT1 interactions with telomeric DNA, gel-shift experiments were performed. AtPOT1 is extremely insoluble when expressed in *E. coli*, but soluble full length AtPOT1 or the N-terminal domain containing the OB-folds can be obtained from rabbit reticulocyte lysate (data not shown). Under the same

conditions that the OB-fold containing amino terminus of mouse POT1a binds its cognate telomere sequence (Wu *et al*., 2006), AtPOT1 failed to bind *Arabidopsis* telomeric DNA (Figure 29). Thus, under these *in vitro* assay conditions, AtPOT1 does not interact with telomeric DNA in the same manner as its mammalian counterpart.

Human POT1 binds telomeres throughout the cell cycle, showing a transient decrease in binding late in G2 (Verdun *et al*., 2005). To investigate AtPOT1 interaction with telomeres *in vivo*, chromatin immunoprecipitation (ChIP) was employed. Since we failed to detect AtPOT1 binding to telomeres in chromatin preparations from plant cell extracts, ChIP assays were performed on suspension culture cells. Slight (1.8-fold) enrichment of AtPOT1 at telomeres was observed in unsynchronized cells relative to the pre-immune sera control (Figure 30B and C). Therefore, we asked whether AtPOT1 localization at telomeres was regulated during the cell cycle. Our synchronization protocol did not allow us to examine cells blocked in G1, however more than 75% of the unsynchronized cells are in this phase of the cell cycle (Figure 30A). Using aphidicolin, we could enrich for cells in S phase and in G2. In four separate experiments, AtPOT1 interaction with telomeres significantly increased in S phase cells; the enhancement ranged from 3.4 to 7.1-fold (avg = 5.4-fold) over the pre-immune sera control (Figure 30B and C). As an additional control, we monitored the ratio of the rDNA signal immunoprecipitated by the POT1 antibody relative to the pre-immune control. As expected, no significant enrichment in S phase was observed (Figure 30B and C). The AtPOT1 association with telomeres decreased dramatically as cells transitioned into G2 (down to 1.8-fold), indicating that AtPOT1 association with telomeres in S phase does not reflect an increased number of binding sites after telomere replication. Furthermore,

h	5'-GGTTAGGGTTAG-3'
1	5'-TTAGGGTTTAGGGT-3'
2	5'-TAGGGTTTAGGGTT-3'
3	5'-AGGGTTTAGGGTTT-3'
4	5'-GGGTTTAGGGTTTA-3'
5	5'-GGTTTAGGGTTTAG-3'
6	5'-GTTTAGGGTTTAGG-3'
7	5'-TTTAGGGTTTAGGG-3'

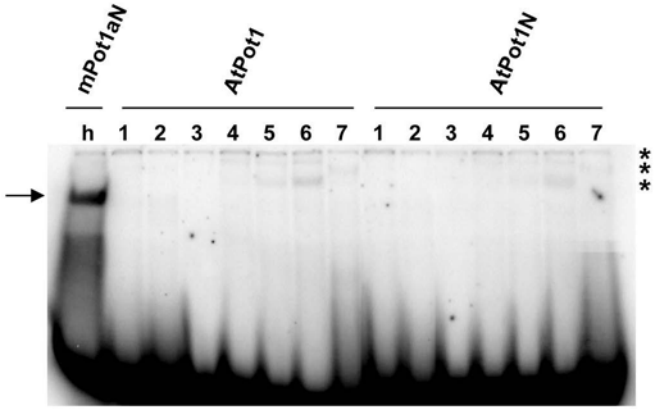


Figure 29. AtPOT1 does not bind telomeric DNA *in vitro*. Oligonucleotides containing vertebrate Pot1 minimal binding site (h) or two *Arabidopsis* telomeric repeats in all seven different permutations of the $(TTTAGGG)_2$ (1-7) (Top panel) were {^{32}P} 5' end-labeled with a T4 polynucleotide kinase (Fermentas). AtPOT1 full length cDNA or a 5' end fragment encoding predicted OB folds (AtPot1N) were cloned into pET28a vector (Novagen). Corresponding proteins were translated using coupled rabbit reticulocyte lysate expression system (Promega). EMSA experiments were performed as described (Baumann and Cech, 2001; Wu *et al*., 2006) (Bottom panel). Mouse Pot1aN protein binding to vertebrate telomeric sequence was monitored as a positive control (Wu *et al*. 2006). Although mouse POT1aN protein bind the mammalian telomere sequence GGTTAGGGTTAG (arrow), no binding of AtPOT1 to any of the seven permutations of the *Arabidopsis* telomeric repeat sequence was observed. Asterisks indicate non-specific bands.

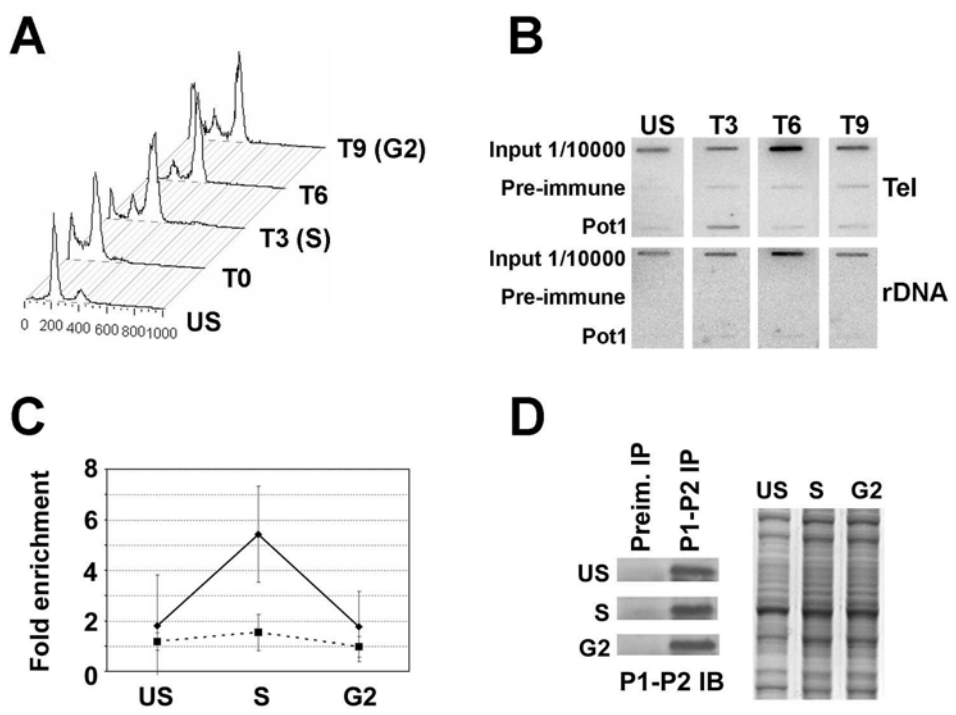


Figure 30. AtPOT1 is associated with telomeric chromatin in S-phase. (A) FACS analysis of *Arabidopsis* suspension culture cells synchronized with aphidicolin. Data are shown for unsynchronized (US) and synchronized cells at 0, 3, 6, and 9 hours after release from aphidicolin arrest. (B) Example of ChIP analysis on synchronized suspension culture extracts using P1-P2 antibody or preimmune serum. Immunoprecipitated DNA was monitored on slot blot using a radiolabeled telomeric or rDNA probe. (C) Quantitation of AtPOT1 association with telomeric DNA. The average of results from four independent experiments is shown. Solid black line indicates the ratio of the telomeric DNA signal obtained with the POT1 antibody relative to the pre-immune sera control. As a negative control, the rDNA signal obtained with the POT1 antibody relative to the preimmune sera (gray dashed line) is shown. (D) Western blot analysis of AtPOT1 protein. Extracts from synchronized cells were precipitated with P1-P2 antibody, followed by P1-P2 western blot analysis. Commassie stained inputs (right) are shown as loading controls.

AtPOT1 protein levels were unchanged during the cell cycle (Figure 30D), arguing that its interaction with telomeres is dynamic and peaks in S phase.

Discussion

Cdc13p and POT1 from yeast and vertebrates are multifunctional gatekeepers at the chromosome terminus, performing the crucial functions of distinguishing the ends from double-strand breaks, protecting against inappropriate recombination and nucleolytic attack, and controlling telomerase activity (Baumann, 2006; de Lange, 2005). Here we demonstrate that AtPOT1 exhibits distinctly different interactions with telomeres. Unlike human POT1, disruption of the *AtPOT1* gene is not lethal and *Arabidopsis pot1* mutants display no evidence of chromosome end deprotection. Even in a *ku70 tert* background where *Arabidopsis* telomeres are severely compromised (Riha and Shippen, 2003a), the loss of AtPOT1 does not exacerbate telomere erosion or increase the frequency of chromosome end-joining. While it is conceivable that AtPOT1 acts redundantly with another component of the telomere complex to protect the terminus, the dynamic interaction of AtPOT1 with telomeres is inconsistent with a primary role in this pathway.

Further distinguishing AtPOT1 from the previously described POT1 proteins is the ever-shorter-telomere phenotype displayed by *Arabidopsis* null mutants. Depletion of POT1 in mammals leads to telomere elongation (Veldman *et al*., 2004; Wu *et al*., 2006; Yang *et al*., 2005; Ye *et al*., 2004b), implying a role in the negative regulation of telomere length. Although it is possible that the mammalian POT1 contribution to positive telomere length regulation is obscured by the other more severe phenotypes associated with POT1 depletion (Churikov *et al*., 2006; Hockemeyer *et al*., 2006; Wu *et al*., 2006), our data argue that the AtPOT1 protein has evolved a pivotal and highly

specialized role in promoting telomerase action at telomeres by working in the context of the telomerase RNP.

Four lines of genetic and biochemical evidence strongly implicate AtPOT1 in the telomerase pathway. First, telomeres shorten at the same rate in *tert* and *pot1* mutants, and depletion of telomere tracts is not accelerated in plants with a deficiency in both genes. Second, AtPOT1 is required for the telomerase-dependent elongation of telomeres in *ku70* mutants. Third, *in vitro* telomerase activity levels are significantly reduced in *pot1* mutants. Fourth, enzymatically active telomerase is specifically immunoprecipitated with AtPOT1 antibodies.

How AtPOT1 interacts with the telomerase RNP to facilitate telomere maintenance is unknown. Transient transfection experiments in tobacco with GFP-tagged AtPOT1 and TERT show that the two proteins co-localize in the nucleolus (N. Kato, E. Lam, E. Shakirov and D. Shippen, unpublished data), where telomerase RNP biogenesis occurs in both yeast and mammals (Etheridge *et al*., 2002; Teixeira *et al*., 2002). Intriguingly, we found that AtPOT1 association with enzymatically active telomerase is regulated in the cell cycle, increasing by an average of ~4.5 fold in S phase relative to unsynchronized cells (predominantly G1) and cells arrested in G2. Thus, AtPOT1 may stabilize an enzymatically active form of the RNP. In support of this model, telomerase activity levels are more variable in the absence of AtPOT1. Although *Arabidopsis* shows no haploinsufficiency with respect to *TERT* (Fitzgerald *et al*., 1999b) or *AtPOT1* (this study), the more compromised telomerase enzyme found in *pot1* null mutants may be unable to solve the end replication problem.

It is also possible that AtPOT1 functions to promote telomerase action on its telomeric DNA substrate. Notably, the variability of *in vitro* telomerase activity levels in

pot1 mutants is incongruent with the highly consistent ever-shorter-telomere phenotype displayed by these plants. By virtue of its two OB-folds, AtPOT1 is predicted to directly bind telomeric DNA. However, under conditions where *S. pombe* and mammalian POT1 associate with telomeric DNA *in vitro* (Baumann and Cech, 2001; Wu *et al*., 2006), AtPOT1 showed no binding. While AtPOT1 may simply need different biochemical reaction conditions to associate with telomeric DNA, another more interesting possibility is that AtPOT1 requires a binding partner. Recent studies reveal that the mammalian POT1 binding partner, TPP1, which cannot bind telomeric DNA on its own (Wang *et al*., 2007; Xin *et al*., 2007), not only greatly enhances the affinity of POT1 for telomeric DNA *in vitro*, but also stimulates telomerase activity and processivity (Wang *et al*., 2007). Moreover, like AtPOT1, TPP1 can assume a canonical OB-fold (Wang *et al*., 2007; Xin *et al*., 2007) and physically interacts with the telomerase RNP (Xin *et al*., 2007). Mice deficient in TPP1 show profound developmental defects and animals that survive to adulthood are infertile (Keegan *et al*., 2005). Thus, TPP1 contrasts with *Arabidopsis* POT1 in that it appears to possess additional functions besides stimulating telomerase activity (Xin *et al*., 2007). Altogether our observations underscore the extraordinarily rapid evolution of the telomeric complex, and indicate that OB-fold bearing proteins, such as AtPOT1, are co-evolving with the telomerase RNP.

CHAPTER V

IDENTIFICATION AND CHARACTERIZATION OF THE THIRD *Arabidopsis* POT1 PROTEIN, AtPOT1c

Summary

POT1 is a protein that in yeast and mammals binds ss telomeric DNA and plays an essential role in chromosome end protection and telomere length regulation. While most organisms encode a single *POT1*, *Arabidopsis* possess two *POT1* like genes, dubbed *AtPOT1a* and *AtPOT1b*. Here, we describe identification and characterization of the third *Arabidopsis POT1* gene, *AtPOT1c*. We show that *AtPOT1c* reflects a very recent gene duplication of *AtPOT1a* that is specific for *Arabidopsis thaliana*. We demonstrate that AtPOT1c is expressed in proliferative tissues and is subjected to alternative splicing. Similar to AtPOT1a, AtPOT1c does not bind telomeric DNA *in vitro* despite the presence of the OB-fold. Moreover, like AtPOT1a, transgenic analysis suggests that AtPOT1c acts in a telomerase pathway for telomere length maintenance, but unlike AtPOT1a, may act as a negative regulator of telomerase action *in vivo*. Furthermore, unlike AtPOT1a, AtPOT1c contributes to chromosome end protection and maintenance of the telomere architecture. Altogether, our data demonstrate that AtPOT1c is a novel protein that has evolved unique functions at *Arabidopsis* telomeres.

Introduction

The ends of linear chromosomes in eukaryotes are protected by telomeres, special nucleoprotein structures that preclude activation of DNA damage repair pathways and

exonucleolytic attack. Telomere extension by telomerase also facilitates the complete replication of chromosome termini and thus is necessary to overcome the end replication problem in actively dividing cells. Telomeres typically consist of simple G-rich repeats. While telomeric DNA is mostly double-stranded (ds), the G-rich strand running in the 5' to 3' direction toward the chromosome end forms a single-stranded (ss) protrusion termed G-overhang. The G-overhang can fold back and invade the duplex region, forming t-loop, a complex secondary structure that sequesters chromosome ends from deleterious activities. The G-overhang associates with ss telomere-specific proteins. Genetic and biochemical analysis of G-overhang binding proteins from different organisms has revealed essential roles in chromosome end protection and in regulation of telomere length (reviewed in Wei and Price, 2003).

The first telomere end binding protein (TEBP) was discovered in the ciliate *Oxytricha nova*, where a high abundance of telomeres allowed purification of DNA-protein complexes through gel-filtration (Gottschling and Zakian, 1986; Price and Cech, 1987). The *Sacharomyces cerevisiae* G-overhang binding protein Cdc13 was initially identified from mutants causing Rad9-dependent cell cycle arrest (Garvik *et al*., 1995), and was subsequently found in a genetic screen for mutations that lead to an ever-shorter-telomere (est) phenotype (Lendvay *et al*., 1996; Nugent *et al*., 1996). The identification of ss telomere proteins from multicellular organisms lagged due to the low abundance of these proteins and to poor conservation of the primary sequence. However, the development of more advanced alignment tools allowed identification of the *S. pombe* and human POT1 proteins, the presumed orthologs of *O. nova* TEBP and *S. cerevisiae* Cdc13 (Baumann and Cech, 2001). POT1 homologs have subsequently been found in other species, including plants. Interestingly, while most eukaryotes

encode only a single G-overhang binding protein, there are few exceptions. In *O. nova*, TEBP consists of two subunits, α and β, which form a heterodimer and display different interactions with the ss telomeric DNA (Fang and Cech, 1993; Gray *et al*., 1991; Hicke *et al*., 1990). Similarly, in another ciliate protozoon *Euplotes*, there are two G-overhang binding proteins (TP and rTP) that are highly divergent in sequence and in function (Price, 1990; Skopp *et al*., 1996). Mice also encode two POT1 proteins that share 72% similarity and appear to be partially redundant for telomere length regulation and chromosome end protection (He *et al*., 2006; Hockemeyer *et al*., 2006; Wu *et al*., 2006).

Structural studies of G-overhang binding proteins demonstrate that despite very limited sequence similarity, these proteins share conserved oligosacharide/ oligonucleotide binding-fold (OB-fold). The OB-fold is a five-stranded β barrel found in many single-strand nucleic acid binding proteins (Theobald *et al*., 2003). *In vitro* all of them bind their corresponding single-strand telomeric DNA, displaying specificity for the G-rich strand mediated through this domain (Baumann and Cech, 2001; Lei *et al*., 2002). Since telomeres form a t-loop structure, it is thought that ss telomere proteins can bind both the G-overhang and the displaced G-rich single-stranded repeats at the base of the t-loop *in vivo.*

G-overhang binding proteins are essential for chromosome end protection. Cdc13p is required to protect C-strand from nucleolytic degradation (Garvik *et al*., 1995) through its interaction with other OB-fold containing proteins, Stn1p and Ten1p. Mutations in *Cdc13*, *Stn1* and *Ten1* result in cell cycle arrest due to accumulation of long stretches of ss G-rich telomeric DNA (Grandin *et al*., 2001; Grandin *et al*., 1997). In *S. pombe*, deletion of POT1 leads to exonucleolytic degradation of telomeric as well as subtelomeric sequences, chromosome mis-segregation, and cell death (Baumann

and Cech, 2001). Intriguingly, deletion of Stn1 and Ten1 in fission yeast also leads to a phenotype identical to POT1 deficiency (Martin *et al*., 2007), demonstrating that multiple proteins associate with the G-overhang to provide chromosome end protection.

In vertebrates, genetic studies of chicken POT1 reveal the essential role of this protein in protecting the chromosome end from a DNA damage response (Churikov *et al*., 2006). Similarly, knockdown of human POT1 results in accumulation of DNA damage response proteins at telomeres and cell cycle arrest (Hockemeyer *et al*., 2005; Veldman *et al*., 2004; Yang *et al*., 2005). Mammalian POT1 controls the ATR-mediated DNA damage response pathway at telomeres (Churikov and Price, 2008; Churikov *et al*., 2006; Denchi and de Lange, 2007; Guo *et al*., 2007). Since activation of ATR pathway requires binding of RPA to single-stranded DNA (Zou and Elledge, 2003), POT1 is thought to block ATR by competing with RPA for G-overhang binding (Denchi and de Lange, 2007). Human POT1 has been also implicated in blocking end-to-end fusions and control of 5' exonucleolytic processing of telomeres (Hockemeyer *et al*., 2005; Veldman *et al*., 2004; Yang *et al*., 2005). Mice POT1 orthologs similarly abrogate a DNA damage response at telomeres and are implicated in protecting the C strand from nuclease attack and blocking homologous recombination at telomeres (He *et al*., 2006; Hockemeyer *et al*., 2006; Wu *et al*., 2006).

Studies of ss telomere binding proteins from different organisms have uncovered a complex role for these proteins in telomere length regulation. A role for Cdc13p in *S. cerevisiae* telomere maintenance is clear. Cdc13 directly recruits telomerase to the chromosome terminus via an interaction with Est1, a telomerase holoenzyme component. While mutations that abolish the Cdc13-Est1 interaction lead to an ever-shorter-telomere phenotype, $Cdc13_{DBD}$-Est1 fusion protein rescues the

progressive telomere shortening phenotype of both *cdc13* and *est1* mutants (Chandra *et al*., 2001; Evans and Lundblad, 1999; Pennock *et al*., 2001). Mutant Cdc13 alleles that result in telomere elongation have also been identified, suggesting that Cdc13 plays a role in both positive and negative regulation of telomere length (Chandra *et al*., 2001; Grandin *et al*., 2000). Similarly, *in vitro* and *in vivo* studies of human POT1 have revealed a complex role of this protein in telomere length control and demonstrated that hPOT1 contributes to both positive and negative regulation (Kelleher *et al*., 2005; Lei *et al*., 2005; Veldman *et al*., 2004; Yang *et al*., 2005; Ye *et al*., 2004b). Unlike Cdc13, the exact mechanism for telomere length regulation by hPOT1 is unknown. It has been recently shown that mammalian POT1 exerts its function in collaboration with TPP1. TPP1 physically interacts with both POT1and active telomerase enzyme (Xin *et al*., 2007). Although TPP1 harbors a canonical OB-fold, it cannot bind telomeric DNA on its own (Wang *et al*., 2007; Xin *et al*., 2007). Instead, the OB-fold in TPP1 is thought to mediate its association with telomerase in an interaction that is required for telomere length control (Xin *et al*., 2007). It is therefore possible that TPP1 fulfils the role of Est1 protein in mammals.

We previously showed that *Arabidopsis* encodes two POT1-like proteins, AtPOT1a and AtPOT1b (Shakirov *et al*., 2005). These proteins display only 49% overall sequence similarity. Moreover, AtPOT1a and AtPOT1b play non-redundant roles at *Arabidopsis* telomeres (Shakirov *et al*., 2005). AtPOT1a positively regulates telomere length through a physical interaction with telomerase RNP. This protein stably associates with active telomerase enzyme, but dynamically associates with telomeres in a cell-cycle dependent manner (Surovtseva *et al*., 2007). In contrast, AtPOT1b is

implicated in chromosome end protection, as dominant-negative AtPOT1b mutants display telomere degradation and end-to-end fusions (Shakirov *et al*., 2005).

Here, we describe the identification and characterization of a third POT1 protein in *Arabidopsis*, AtPOT1c. *AtPOT1c* is the result of very recent gene duplication of *AtPOT1a*. We found that AtPOT1c is present only in *Arabidopsis thaliana* and not in other closely related plant species. Despite its recent appearance, AtPOT1c has evolved functions distinct from AtPOT1a and AtPOT1b. We show that over-expression of AtPOT1c leads to deregulation of telomere length at individual chromosome arms, a phenotype consistent with telomerase dysfunction *in vivo*. We also show that AtPOT1c is involved in the maintenance of telomere architecture, as AtPOT1c over-expression leads to a dramatic reduction in G-overhang. Notably, AtPOT1c exerts the end protection function in concert with telomerase: over-expression of AtPOT1c in the absence of telomerase leads to a further reduction in G-overhang length. Finally, we show that telomeres remain refractory to end-to-end fusions despite a reduction in the G-overhang in AtPOT1c mutants. Thus, AtPOT1c appears to be a novel protein that has evolved unique functions at *Arabidopsis* telomeres.

Materials and methods

Plant materials and plant transformation

The *tert* mutants and plant growth conditions are as described (Riha *et al*., 2001; Shakirov *et al*., 2005). *Agrobacterium tumefaciens* strain GV3101 was used for plant transformation. Transformation was performed by the *in planta* method (Riha *et al*., 2002). T1 primary transformants were selected as described (Shakirov *et al*., 2005).

RNA extraction, RT-PCR, and molecular cloning

RNA was extracted from ~0.1 g of plant tissue using Plant RNA Purification Reagent (Invitrogen). cDNA was synthesized with Superscript III reverse transcriptase (Invitrogen). *AtPOT1c* transcripts were obtained by PCR with POT1c_nested_fwd (5-CACTGTCTAATGACGTCAGGAT) and POT1c_nested_rev (5'-CCTGAATGAATGACTCAGCTC) primers, followed by PCR with POT1c_cDNAF (5'-ATGGCGAAGAAGAGAGACAGTC) and POT1c_nested_rev primers. For sequencing, AtPOT1c transcripts were cloned into the pDrive vector (Quiagen). For plant transformation, the AtPOT1c locus was amplified from genomic DNA using primers 5-CGGGATCCATGGCGAAGAAGAGAGACAGTC and 5'-CGACTAGTTCACTTAGCCAGAATAGATGCTGG, and cloned into the pCBK05 expression vector (Shakirov *et al*., 2005).

Gel shift

AtPOT1cS and *AtPOT1cL* were cloned into a pET28a expression vector (Novagen). Proteins were translated using rabbit reticulocyte lysate (Promega). Gel-shifts were performed as described (Surovtseva *et al*., 2007; Wu *et al*., 2006).

Telomere analysis, quantitative TRAP and DNA damage sensitivity assay

Total plant DNA was extracted as described (Cocciolone and Cone, 1993). TRF analysis was performed on *Tru1I* (Fermentas) digested DNA samples (Fitzgerald *et al*., 1999). PETRA and telomere fusion PCR were performed as described (Heacock *et al*., 2004). TRF, PETRA and fusion PCR blots were hybridized with a ^{32}P 5' end-labeled $(T_3AG_3)_4$ probe. Quantitative real-time TRAP was conducted on protein extracts

prepared from two to three influorescences as described (Kannan *et al*., 2008). For DNA damage sensitivity assay, seedlings were incubated in methyl methanesulfonate (Aldrich) as discussed (Heacock *et al*., 2007).

Results

Identification of the third POT gene in Arabidopsis

A BLAST search of the *Arabidopsis thaliana* genome database with the *AtPOT1a* sequence as a query unexpectedly identified a second gene (At2g04395) on chromosome 2 located ~450 kb upstream of the *AtPOT1a* transcribed in the opposite orientation (Figure 31A). We termed this gene *AtPOT1c*. Analysis of the *AtPOT1c* sequence revealed ~90% identity at the nucleotide level to the 5' region of *AtPOT1a*. *AtPOT1c* was annotated as an expressed gene in the database, and this was confirmed by reverse transcription-PCR analysis. The *AtPOT1c* transcript was not abundant, however, and could only be detected by nested PCR following reverse transcription. In contrast to AtPOT1a and AtPOT1b, which are ubiquitously expressed, RT-PCR analysis of mRNA from different plant organs revealed *AtPOT1c* mRNA only in proliferative tissues, such as flowers and seedlings, but not in rosette leaves or stems. Furthermore, two major bands were amplified in the PCR reaction, suggesting that *AtPOT1c* is subjected to alternative splicing (Figure 31B). Cloning and sequencing of *AtPOT1c* transcripts confirmed this prediction and revealed two major isoforms of *AtPOT1c*, dubbed *AtPOT1cL* and *AtPOT1cS*. These findings indicate that the transcription profile and presence of mRNA of *AtPOT1c* is distinct from *AtPOT1a* and *AtPOT1b*.

The predicted translation products of both *AtPOT1cL* and *AtPOT1cS* encode a single OB-fold (D. Wuttke, personal communications), which display 39% identity (40%

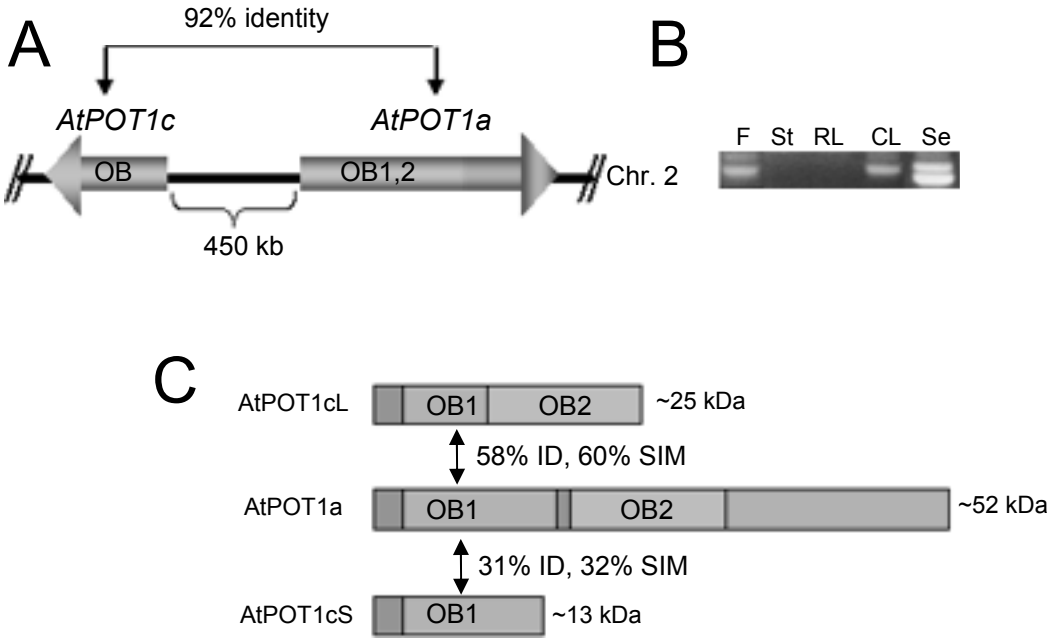


Figure 31. *AtPOT1c* undergoes alternative splicing. (A) Schematic of the *Arabidopsis* genome locus on chromosome 2 encoding *AtPOT1a* and *AtPOT1c*. *AtPOT1c* is a product of a recent duplication of the 5' end of *AtPOT1a*. (B) RT-PCR analysis of *AtPOT1c* expression in different plant tissues. (C) Schematic of AtPOT1cL and AtPOT1cS proteins. Identity (ID) and similarity (SIM) to the OB-fold region of AtPOT1a is shown.

similarity) to each other. Interestingly, while the OB-fold sequence in *AtPOT1cS* is similar to the first OB-fold in AtPOT1a, *AtPOT1cL* harbors an OB-fold which has sequence similarity to both OB folds of *AtPOT1a* (Figure 31C). AtPOT1cS and AtPOT1cL show only 31% and 58% identity, respectively, to the corresponding region of AtPOT1a protein, indicating that alternative splicing of *AtPOT1c* may result in protein isoforms that are functionally distinct from each other and from AtPOT1a.

To investigate AtPOT1c interactions with telomeric DNA *in vitro*, gel shift experiments were performed. AtPOT1cS and AtPOT1cL proteins were expressed in rabbit reticulocyte lysate and incubated with a ss oligonucleotide corresponding to two *Arabidopsis* telomeric repeats. Under conditions that allowed mouse POT1a binding to its cognate telomeric sequence (Wu *et al*., 2006) (Figure 32), we failed to detect AtPOT1c binding to any of the seven permutations of the *Arabidopsis* telomeric DNA repeat sequence (Figure 32). We conclude that despite the presence of an OB-fold, neither of the AtPOT1c isoforms directly interacts with telomeric DNA under these *in vitro* conditions. In this regard, AtPOT1c is similar to AtPOT1a and AtPOT1b, neither of which binds to telomeric DNA *in vitro*.

Over-expression of AtPOT1c results in telomere length deregulation

To study the function of AtPOT1c *in vivo*, we sought to identify a null mutant in this gene. Unfortunately, there is no T-DNA disruption line available for *AtPOT1c* in any of the annotated T-DNA collections. Moreover, a PCR screen of the DNA pools from D. Weigel and T. Jack collections failed to identify an AtPOT1c mutant (data not shown). Therefore, we undertook a transgenic approach to examine the function of AtPOT1c *in vivo* by over-expressing the wild type protein. The genomic sequence of *AtPOT1c* was

h	5'-GGTTAGGGTTAG-3'
1	5'-TTAGGGTTTAGGGT-3'
2	5'-TAGGGTTTAGGGTT-3'
3	5'-AGGGTTTAGGGTTT-3'
4	5'-GGGTTTAGGGTTTA-3'
5	5'-GGTTTAGGGTTTAG-3'
6	5'-GTTTAGGGTTTAGG-3'
7	5'-TTTAGGGTTTAGGG-3'

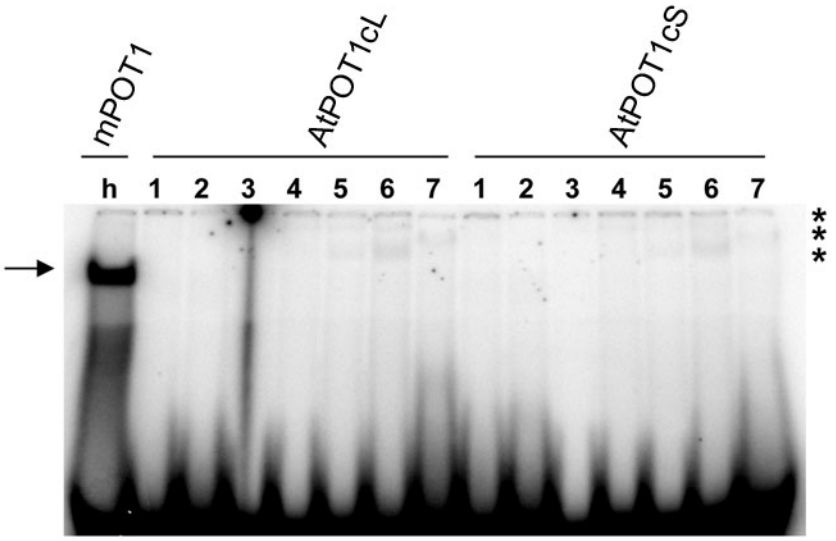


Figure 32. AtPOT1c does not bind telomeric DNA *in vitro*. Oligonucleotides containing vertebrate POT1 minimal binding site (h) or two *Arabidopsis* telomeric repeats in all seven different permutations of the $(TTTAGGG)_2$ are shown on the top. Gel-shift experiment results are shown in the bottom panel. While mouse POT1a protein binds the vertebrate telomere sequence (arrow), no binding of AtPOT1cL or AtPOT1cS to any permutations of the *Arabidopsis* telomeric repeat sequence is observed. Asterisks indicate non-specific bands.

cloned downstream of the strong cauliflower mosaic virus 35S promoter and transformed into wild type *Arabidopsis*. Over-expression of *AtPOT1c* mRNA was verified by RT-PCR (Figure 33A). Sequencing of the RT-PCR products confirmed that both *AtPOT1cL* and *AtPOT1cS* isoforms were expressed at a similar, high level in transgenic plants (data not shown).

Plants over-expressing *AtPOT1c* did not display any growth or developmental defects and were completely fertile (data not shown). To analyze bulk telomere length, terminal restriction fragment (TRF) analysis was performed on *AtPOT1c* primary transfromants. In wild type *Arabidopsis* plants, bulk telomeres appear as a smear ranging from 2 to 5 kb (Shakirov and Shippen, 2004). Since the level of *AtPOT1c* expression is likely to vary in different transgenic plants, several AtPOT1c transformants were subjected to TRF analysis. Interestingly, while in most primary transformants telomeres fell into wild type range (Figure 33B, lanes 4,6 and 7), some plants displayed telomeres that were either shorter (Figure 33B, lane 3) or significantly longer (Figure 33B, lane 5) than wild type. These data indicate that bulk telomere length regulation is perturbed in AtPOT1c OE mutants. We have previously shown that over-expression of the corresponding regions of AtPOT1a and AtPOT1b does not result in a similar TRF profile (Shakirov *et al*., 2005), suggesting that AtPOT1c plays a distinct role in telomere length regulation.

To examine telomere tracts on individual chromosome arms, we used primer extension telomere repeat amplification (PETRA). In this PCR assay, individual telomeres are amplified using a primer that anneals to the G-overhang and a primer that recognizes a unique sequence in the subtelomeric region of each chromosome arm (Heacock *et al*., 2004). We have previously shown that the length of telomere tracts on

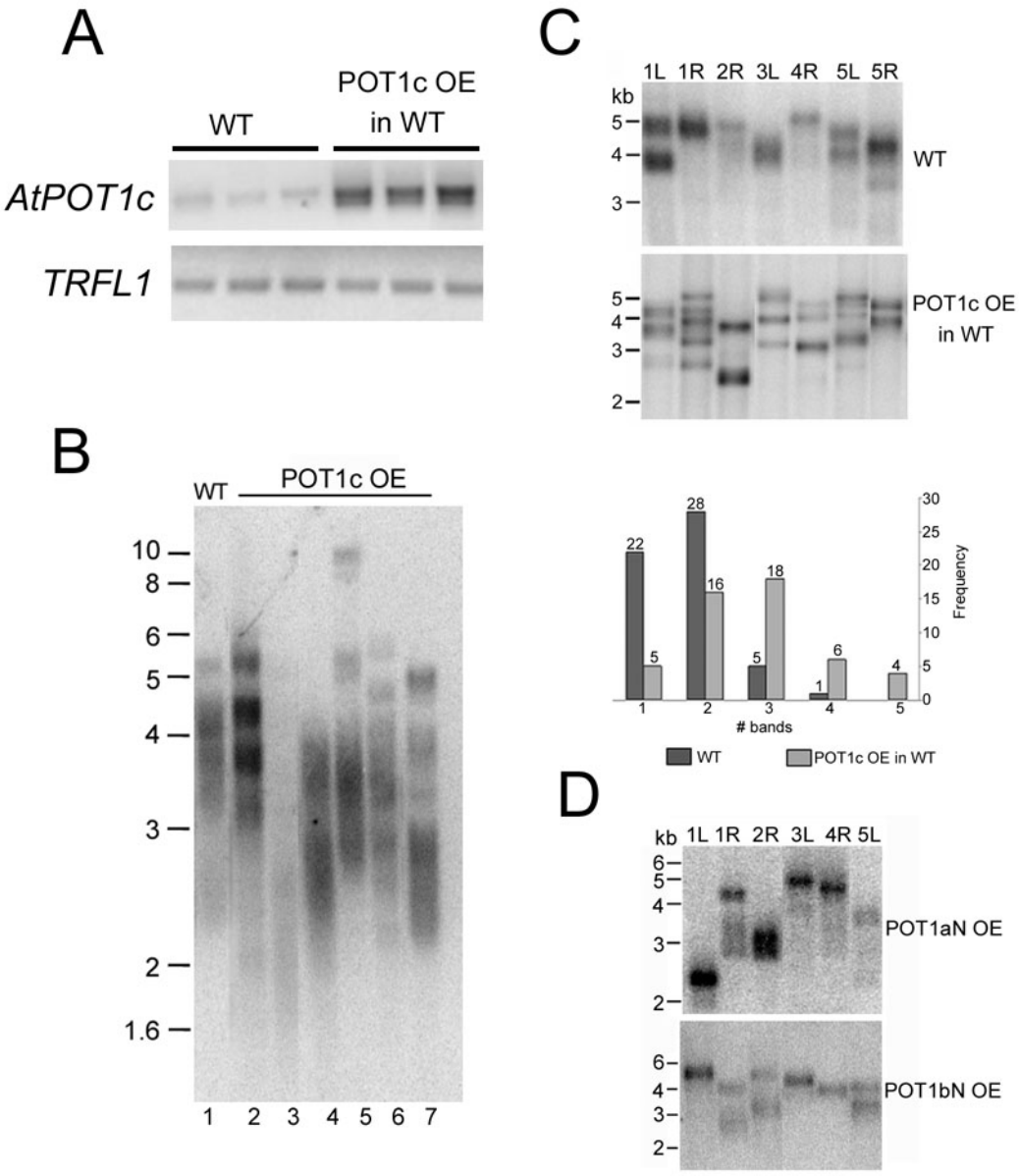


Figure 33. AtPOT1c over-expression results in deregulation of telomere length. (A) RT-PCR analysis of *AtPOT1c* transcripts in wild type plants and plants over-expressing AtPOT1c. (B) TRF analysis of primary transformants over-expressing AtPOT1c. (C) Top panel, PETRA analysis of telomeres on seven individual chromosome arms in wild type and AtPOT1c over-expressing mutants. Bottom panel, quantification of the number of PETRA bands. (D) PETRA analysis of plants over-expressing AtPOT1a OB-folds (AtPOT1aN) or AtPOT1b OB-folds (AtPOT1bN). In panels B, C and D, blots were hybridized with the radiolabeled telomeric probe. Molecular weight markers are indicated.

homologous chromosomes is tightly and coordinately regulated, giving rise to only one or two hybridization signals (Shakirov and Shippen, 2004). In striking contrast, PETRA revealed multiple telomeric hybridization signals in DNA samples obtained from AtPOT1c OE mutants (Figure 33C). Quantification of the number of hybridization bands for multiple chromosome arms in wild type (n=56) and AtPOT1c OE (n=49) plants demonstrated that on average 1.5 telomeric bands are generated for each chromosome arm in wild type plants. In contrast, an average of three telomeric bands is detected in plants over-expressing AtPOT1c. In addition to an increased number of PETRA signals, the PETRA profile consisted of much sharper bands than in wild type (Figure 33C). Therefore, over-expression of AtPOT1c results in deregulation of telomere length on individual chromosome ends. To test whether this phenotype is specific to AtPOT1c over-expression, we performed PETRA on plants over-expressing the corresponding regions of AtPOT1a and AtPOT1b. In contrast to AtPOT1c, PETRA generated diffuse, smeary telomeric bands similar to wild type for POT1a-N-terminal fragment over-expression. Interestingly, however, although the bands were more heterogeneous than for POT1c over-expression, over-expression of POT1b-N-terminal fragment generated a complex PETRA banding pattern for some chromosome arms (Figure 33D). The significance of this observation for AtPOT1b function is unknown. Nevertheless, these data argue strongly that AtPOT1c plays a different role at telomeres than AtPOT1a.

AtPOT1c controls telomere length by working in the telomerase pathway

In addition to an increased number of PETRA signals, the PETRA profile consisted of much sharper bands than in wild type (Figure 33C). In *Arabidopsis*, telomerase is the major force contributing to the heterogeneity of individual telomere tracts (Riha *et al*.,

2001). As a consequence, when telomerase is inactivated, the diffuse PETRA signals obtained in wild type plants become much sharper, consistent with a reduction in length heterogeneity at telomeres on homologous chromosomes (Heacock *et al*., 2004; Riha *et al*., 2001). Therefore, the sharpness of the bands in AtPOT1c over-expression mutants could result from telomerase dysfunction. To examine telomerase activity levels in AtPOT1c over-expressing plants, quantitative real time TRAP (telomere repeat amplification protocol) was performed with extracts from wild type plants and several plants over-expressing AtPOT1c (Figure 34A). Telomerase activity levels in plants over-expressing AtPOT1c were decreased by approximately two-fold relative to wild type, arguing that AtPOT1c is required for maximal telomerase activity.

We next asked whether AtPOT1c acts in collaboration with telomerase for telomere maintenance *in vivo*. AtPOT1c was over-expressed in telomerase-deficient (*tert*) mutants, and PETRA analysis was performed on primary transformants. As a control, PETRA was performed on untransformed *tert* mutants. As expected, a single sharp band was generated from each PETRA reaction when *tert* DNA was used (Figure 34B, right panel). Interestingly, the PETRA profile was the same in the *tert* mutants over-expressing AtPOT1c as in telomerase single mutants (Figure 34B, left panel), suggesting that AtPOT1c works in the telomerase pathway to maintain telomere tracts.

Over-expression of AtPOT1c results in a decrease in the G-overhang signal.

Since telomerase acts on the ss G-overhang, we examined the status of this structure in plants over-expressing AtPOT1c using in-gel hybridization. Non-denatured genomic DNA was hybridized with a C-rich telomeric probe to evaluate the relative amount of ss G-rich telomeric DNA. The G-overhang signal was decreased by ~5-fold in AtPOT1c

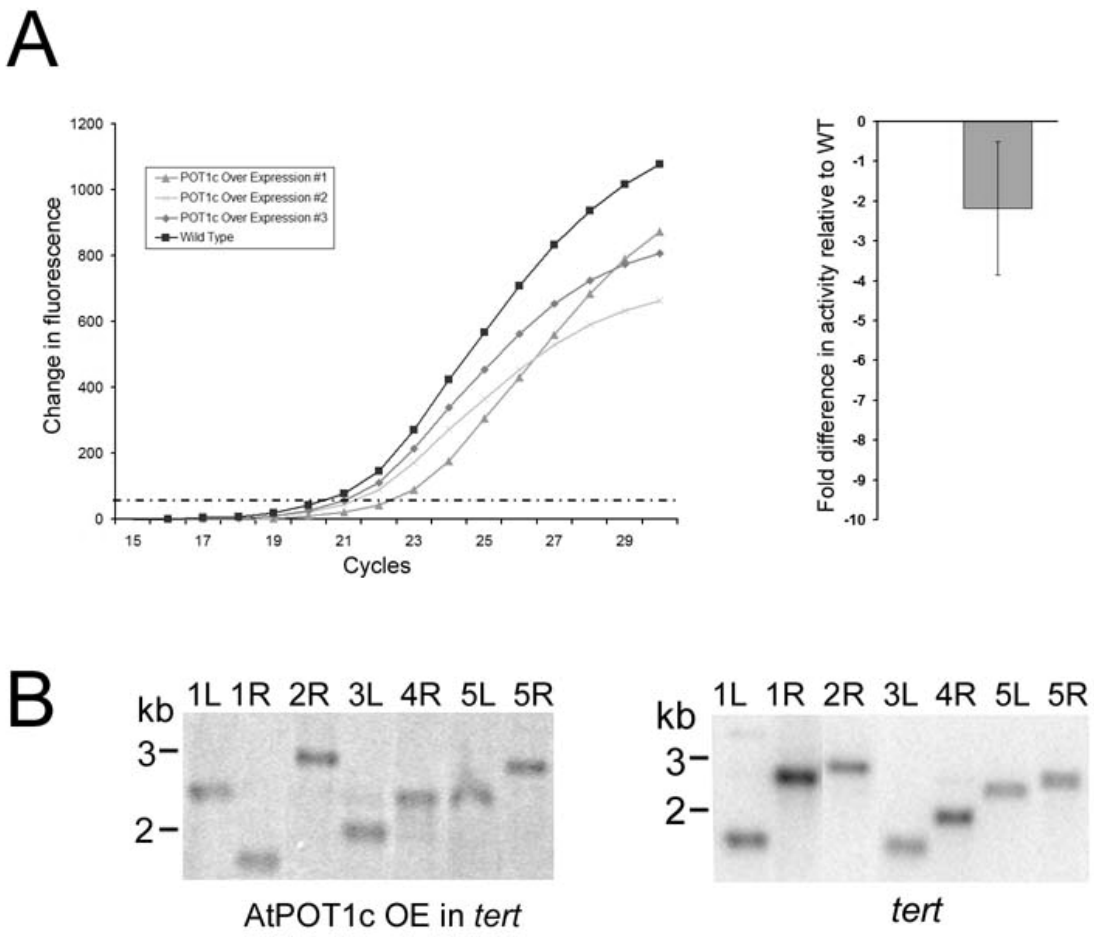


Figure 34. AtPOT1c is required for normal telomerase activity *in vitro*, and works in concert with telomerase to maintain telomeres *in vivo*. (A) Results of real time TRAP on AtPOT1c over-expressing mutants. Data for three primary transformants is shown. Dashed line represents the threshold change in fluorescence. Quantification of the telomerase activity levels in AtPOT1c over-expression mutants relative to wild type is shown on the right. (B) PETRA analysis of *tert* mutants over-expressing AtPOT3 (left panel) and single *tert* mutants (right panel). Telomere length was measured on individual chromosome arms as indicated.

over-expression mutants compared to wild type (Figure 35A), arguing that AtPOT1c contributes to proper maintenance of telomere architecture.

We next examined the contribution of telomerase in G-overhang length regulation in the context of AtPOT1c over-expression. In-gel hybridization was performed on *tert* mutants over-expressing AtPOT1c. Remarkably, the G-overhang signal was reduced by more than 30-fold (Figure 35A), which represents a further six-fold decrease in the G-overhang signal relative to POT1c over-expression alone. We conclude that AtPOT1c acts in concert with telomerase for G-overhang maintenance.

Chromosome ends remain protected in plants over-expressing AtPOT1c

The G-overhang is an essential feature of telomeres, as its loss leads to end-to-end chromosome fusions (Celli and de Lange, 2005). Since G-overhang signals are ~5-fold reduced when AtPOT1c is over-expressed in wild type plants and barely detectable when AtPOT1c is over-expressed in *tert* mutants, we asked whether chromosome end were deprotected in these settings using telomere fusion PCR. In this assay, PCR products are generated when telomeres are fused end-to-end (Heacock *et al*., 2004). As expected, a strong hybridization signal was generated with DNA samples from plants deficient in the essential telomere capping protein CIT1 (Y. Surovtseva and D. Shippen, unpublished). In striking contrast, no telomere fusion PCR products were obtained when AtPOT1c was over-expressed in wild type plants or in *tert* mutants (Figure 35B). Thus, perturbation in the G-overhang structure does not lead to genome instability associated with AtPOT1c over-expression.

Although we failed to detect evidence for telomere fusions in AtPOT1c mutants, it is possible that telomeres in these plants elicit a DNA damage that does not

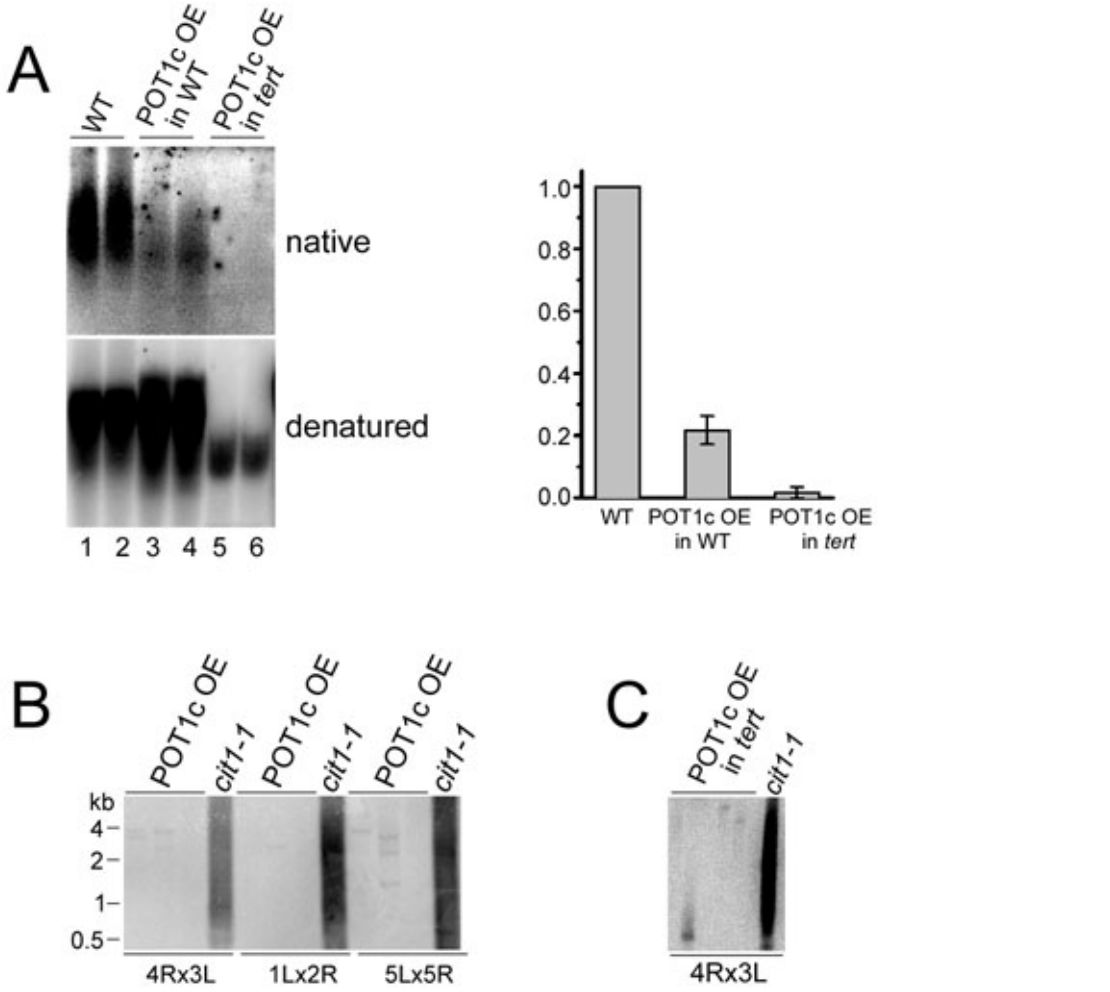


Figure 35. AtPOT1c over-expression alters telomere architecture but does not lead to genome instability. (A) Left panel, in-gel hybridization analysis of the G-overhang signal in wild type and POT1c over-expressing mutants. Native and denatured gels are shown. Right panel, quantification of the G-overhang signal. Ratio of the hybridization signal intensity relative to wild type is shown. Wild type is set to one. (B) Fusion PCR results for AtPOT1c over-expression in wild type. Primers specific for subtelomeric regions on 1L, 2R, 3L, 4R, 5L and 5R chromosome arms were used. (C) Fusion PCR results for AtPOT1c over-expression in *tert*. In panels B and C, *cit1-1* DNA was used as a positive control. All blots were hybridized with a radiolabeled telomeric probe.

result in end joining. To test this hypothesis, we investigated the sensitivity of AtPOT1c over-expressing mutants to double-strand break inducing agent MMS. Cells with dysfunctional telomeres are typically hypersensitive to DNA damage-inducing agents. Intriguingly, however, we failed to detect MMS hypersensitivity in plants over-expressing AtPOT1c in either wild type or *tert* backgrounds (data not shown). These data argue that despite the altered architecture of the chromosome terminus, telomeres remain fully capped and functional in AtPOT1c over-expressing mutants.

Discussion

The key functions of telomeres are mediated by a large array of proteins that bind both ds and ss portions of telomeric DNA. Single-strand telomeric DNA binding proteins are essential as they provide both telomere length regulation and chromosome end protection. We previously showed that *Arabidopsis* encodes two highly diverged POT1 paralogs that play distinct roles in telomere biology (Shakirov *et al*., 2005). Here we describe the identification and characterization of a third *Arabidopsis thaliana* POT1 protein, AtPOT1c.

AtPOT1c reflects a very recent gene duplication of *AtPOT1a*. It displays 92% identity to the 5' region of AtPOT1a at the nucleotide level. *In silico* searches through the available genome database of *Arabidopsis lyrata* failed to identify *POT1c* (A. Nelson and D. Shippen, unpublished). Furthermore, PCR based approaches to identify *POT1c* in *A. lyrata* or other *Brassicaceae* species closely related to *A. thaliana* were unsuccessful (A. Nelson and D. Shippen, unpublished). It is notable that *A. lyrata* diverged from *A. thaliana* only five million years ago. By comparison, this evolutionary

distance in mammals is equivalent to the divergence between humans and chimps. Therefore, *POT1c* appears to be a very recent invention of *Arabidopsis thaliana*.

Several lines of evidence indicate that despite its very recent emergence, *AtPOT1c* has evolved a novel function distinct from its ancestor *AtPOT1a*. First, the expression profile of these two genes is different. While *AtPOT1a* is ubiquitously expressed (Shakirov *et al*., 2005), *AtPOT1c* is transcribed only in proliferative tissues. Second, although both *AtPOT1a* and *AtPOT1c* mRNA are subjected to alternative splicing, *AtPOT1a* splice variants are ubiquitous and not tissue-specific (Tani and Murata, 2005). In contrast, *AtPOT1c* shows an alternative splicing pattern that varies in different tissues and that is distinct from the splicing pattern observed in the corresponding region of AtPOT1a. Although the importance of the two *AtPOT1c* isoforms is unclear, the differential expression and splicing might reflect a regulatory role. Finally, consequences of AtPOT1c over-expression differ from over-expression of the corresponding region of AtPOT1a (Shakirov *et al*., 2005). Altogether, these data indicate that AtPOT1c has evolved a function that is distinct from AtPOT1a.

Our data implicate AtPOT1c in telomere length regulation. In contrast to the PETRA profile generated with wild type plants, we observed a cluster of several sharp bands when the DNA from AtPOT1c over-expressing plants was examined. A similar PETRA profile has recently been observed in plants bearing a mutation in dyskerin, a newly identified component of telomerase RNP (Kannan *et al*., 2008). In these mutants, telomerase activity *in vitro* and *in vivo* is decreased, but not abolished, suggesting that the distinctive banding pattern produced by PETRA reflects a reduction in telomerase function. As for dyskerin mutants, quantitative *in vitro* TRAP assay performed with AtPOT1c over-expression mutants revealed a decrease in telomerase activity relative to

wild type, arguing that the distinct PETRA profile observed in these two different genetic backgrounds reflects the relative level of telomerase activity *in vivo*.

Genetic experiments support the conclusion that AtPOT1c acts in concert with telomerase. Over-expression of AtPOT1c in *tert* mutants results in PETRA profile that is identical to that of single *tert* mutants, and is consistent with the complete loss of telomerase action on all chromosome ends. This finding is consistent with the interpretation that multiple PETRA signals generated for each chromosome end in AtPOT1c over-expressing plants reflect reduced telomerase activity. Furthermore, as discussed below, perturbations of the G-overhang in AtPOT1c over-expressing plants is strongly exacerbated in mutants also deficient in TERT.

The mechanism for AtPOT1c role in telomerase function remains unclear. Despite the presence of the OB-fold, AtPOT1c does not interact with telomeric DNA *in vitro*. One possibility is that the OB-fold in AtPOT1c mediates its interaction with a telomere protein that contributes to telomerase recruitment to the chromosome terminus. In this scenario, over-expression of AtPOT1c may result in titration of this factor from the terminus, thus decreasing telomerase action. Alternatively, AtPOT1c may bind to one of the newly discovered telomerase RNAs in *Arabidopsis*, as has been shown for AtPOT1a (C. Cifuentes-Rojas and D. Shippen, unpublished). Indeed, very recent data from our lab shows a direct interaction between AtPOT1c and the *Arabidopsis* TER1G7 RNA. Thus, it is possible that AtPOT1c is a natural dominant negative regulator of telomerase, perhaps by competing with AtPOT1a for a binding site on TER. In this scenario, over-expression of AtPOT1c would displace AtPOT1a, a positive regulator of telomerase, from the enzyme, leading to a reduction in the overall level of telomerase activity *in vitro* and *in vivo*.

Unexpectedly, in addition to a role for AtPOT1c in telomere length regulation, we found evidence that this protein is important for the proper architecture of the chromosome end. In plants over-expressing AtPOT1c, the G-overhang signal is reduced by ~5-fold. Telomeric G-overhangs are generated through a combination of telomerase action and exonucleolytic processing of the C strand (Chai *et al*., 2006; Jacob *et al*., 2003; Larrivee *et al*., 2004). It is therefore possible that over-expression of AtPOT1c reduces ability of telomerase to extend G-overhangs, or alternatively reduces the ability of nuclease to process the C-strand. In *Arabidopsis*, in-gel hybridization reveals no difference in G-overhang status in telomerase-deficient mutants relative to wild type (M. Heacock and D. Shippen, unpublished), suggesting that the C-strand processing is the major factor contributing to the length of ss telomeric DNA. Thus, we favor the model that AtPOT1c controls susceptibility of the C-strand to nuclease attack. Interestingly, over-expression of the corresponding region of AtPOT1b results in chromosome end deprotection, and mutants display severe morphological defects and loss of terminal sequences (Shakirov *et al*., 2005). Although the G-overhang status has not been analyzed in AtPOT1b mutants, these data argue that AtPOT1b and AtPOT1c play distinct roles in chromosome end protection.

Although *tert* mutants do not display any change in the G-overhang signal, AtPOT1c over-expression in telomerase-deficient mutants resulted in a dramatic reduction (~30-fold) in the G-overhang signal relative to wild type and 6-fold reduction relative to that in wild type plants over-expressing AtPOT1c. This observation implies that AtPOT1c collaborates with telomerase for G-overhang maintenance. It has been hypothesized in yeast and mammals that TERT itself may play a direct role in chromosome end protection (Blackburn, 2005; Hsu *et al*., 2007; Kim *et al*., 2003; Zhu *et*

al., 1999). Thus, the combined dysfunction of AtPOT1c and TERT may abrogate C-strand protection. Our data further imply AtPOT1c contacts an additional factor on the chromosome terminus that is involved in C-strand protection since over-expression of AtPOT1c in the telomerase background leads to a much greater loss of G-strand relative to *tert* mutant alone. In this view, over-expression of POT1c results in the displacement of an essential capping protein from the chromosome terminus. Finally, despite the significant perturbations in the architecture of the chromosome terminus in plants over-expressing AtPOT1c, telomeres are resistant to end-to-end fusions. Since G-overhangs are only ~20 nucleotides in length in *Arabidopsis* (Riha *et al*., 2000), the 30-fold reduction of the signal intensity argues that G-overhangs at most of the telomeres are dramatically shortened or even lost when AtPOT1c is over-expressed in *tert* background. Despite this, plants are completely fertile at least in the first generation and display no morphological defects. While it is possible that telomeres initiate DNA damage response in *tert* mutants over-expressing AtPOT1c, these plants do not show an increased sensitivity to DNA damaging agents, arguing that the telomere perturbations in this background do abolish the normal global DNA damage response.

CHAPTER VI

IDENTIFICATION AND CHARACTERIZATION OF A NOVEL *Arabidopsis* PROTEIN CRITICAL FOR THE INTEGRITY OF TELOMERES

Summary

The protective telomere function is mediated by proteins that associate with either ds or ss telomeric DNA. Although telomere composition has been examined in detail in yeast and in mammals, little is known about the proteins that protect plant telomeres. Here, we describe identification and characterization of the novel factor essential from chromosome end protection in *Arabidopsis* termed Critical for the Integrity of Telomeres (CIT1). The *CIT1* gene was identified by map-based cloning of a mutation that resulted in a severe telomere deprotection phenotype. Analysis of two additional null alleles of CIT1 verified that CIT1 is required for chromosome end protection in *Arabidopsis*. Disruption of CIT1 leads to a severe telomere deprotection phenotype reminiscent of a deficiency in the ds or ss telomere capping proteins found in other organisms. Telomeres in *cit1* mutants are subjected to massive nuclease attack, resulting in both extensive C-strand resection and loss of telomeric and subtelomeric sequences. Moreover loss of CIT1 leads to end-to-end chromosome fusions and profound genome instability. Finally, CIT1 appears to protect telomeres from homologous recombination. Notably, orthologs of CIT1 are present in a wide variety of plants and high eukaryotes, indicating that CIT1 represents a novel component of the telomere cap.

Introduction

One challenge faced by linear chromosomes is the need to protect their termini from being recognized as double strand-breaks, which can lead to deleterious nuclease activities and recruitment into inappropriate repair reactions. Telomeres solve this problem by forming a protective cap on the chromosome end, which consists of an elaborate architecture and a suite of specific telomere proteins. The formation of a t-loop regulates the end protection (reviewed in de Lange, 2004). Telomeric proteins that associate with either ds or ss telomeric DNA provide further protection for the chromosome terminus. In mammals, the six core telomere associated proteins are termed shelterin (de Lange, 2005) and are comprised of the double-strand telomeric DNA binding proteins TRF1 and TRF2, the G-overhang binding protein POT1, and three proteins associated with telomeres via protein-protein interactions (TPP1, TIN2 and RAP1). Perturbations in individual shelterin components or telomere shortening due to telomerase deficiency lead to telomere dysfunction (reviewed in Ferreira *et al*., 2004).

There are two major classes of proteins that are required for chromosome end protection. The first class is represented by proteins that bind ds telomeric DNA, including Taz1 in *S. pombe* and TRF2 in mammals (Broccoli *et al*., 1997; Cooper *et al*., 1997). Deletion of Taz1 leads to MRN-dependent elongation of the G-overhang, implicating Taz1 in the protection of the C strand from nucleolytic processing (Cooper *et al*., 1997; D'Amours and Jackson, 2002). While no telomere fusions is observed in G2 where homologous recombination (HR) is a prevalent double-strand break (DSB) repair mechanism , G1 arrest leads to NHEJ-mediated end-to-end fusions between long telomeres and cell death upon Taz1 loss (Ferreira and Cooper, 2001). Taz1 is also

required to prevent telomere entanglement, accumulation of DNA DSBs, and checkpoint activation after telomere replication in the cold (Miller and Cooper, 2003).

In mammals, the functional homolog of Taz1 is TRF2 (Broccoli *et al*., 1997). Telomere uncapping phenotype has been most extensively studied using TRF2$^{\Delta B\Delta M}$ dominant negative mutant allele that displaces the wild type TRF2 as well as portion of TRF2-associated shelterin components from telomeres (Loayza and de Lange, 2003; van Steensel *et al*., 1998). Such inhibition of TRF2 results in recruitment of numerous DNA damage response factors to the chromosome terminus and formation of telomere dysfunction-induced foci (TIFs) (Takai *et al*., 2003). The DNA damage checkpoint arising at telomeres ultimately triggers cell cycle arrest and cell death (van Steensel *et al*., 1998). Since TRF2 can directly bind ATM kinase *in vitro* and *in vivo*, the current model proposes that TRF2 inhibits ATM signaling pathway by blocking activation of ATM at functional telomeres (Denchi and de Lange, 2007; Karlseder *et al*., 2004).

Uncapped telomeres in TRF2$^{\Delta B\Delta M}$ mutant cells or in conditional TRF2 knockout mouse cells are subjected to ERCC1/XPF flap endonuclease dependent G-overhang cleavage and are immediately recruited into NHEJ-mediated fusions without prior degradation of telomeric repeats (van Steensel *et al*., 1998; Zhu *et al*., 2003). TRF2 facilitates formation of t-loop like structures *in vitro* (Fouche *et al*., 2006; Stansel *et al*., 2001), and hence the t-loop destabilization and G-overhang exposure that occur with TRF2 inhibition are thought to be responsible for telomere fusions and genome instability that arise in these mutants.

Analysis of a dominant negative TRF2 allele that lacks the N-terminal basic domain (TRF2$^{\Delta B}$) demonstrated that in addition to protection from NHEJ, TRF2 is essential to prevent homologous recombination at telomeres. Expression of TRF2$^{\Delta B}$

results in dramatic recombination-dependent telomere shortening and accumulation of t-loop sized extrachromosomal telomeric circles (ECTCs) (Wang *et al*., 2004). Remarkably, telomere deletions upon TRF2 inhibition are dependent on XRCC3, a protein involved in Holiday junction resolution (Wang *et al*., 2004). It is currently believed that the mechanism of t-loop cleavage in $TRF2^{\Delta B}$ cells is similar to TRD in yeast (Lustig, 2003), where branch migration of the displaced loop leads to the formation of Holliday junction, resolution of which produces a shortened telomere and an extra-chromosomal telomeric circle (ECTC). Although homologous recombination is one of the best solutions for telomere length maintenance in the absence of telomerase, TRF2 appears to suppress it in normal cells (Wang *et al*., 2004). Altogether, these data demonstrate an essential role of TRF2 in the protection of chromosome end from DNA damage response, G-overhang loss, NHEJ and inappropriate homologous recombination.

A second major class of proteins involved in chromosome end protection are the ss telomeric DNA binding proteins that associate with the extreme terminus of the chromosome via an OB fold binding domain. In budding yeast, this protein is Cdc13p. Cdc13p is required to protect C-strand from nucleolytic degradation (Garvik *et al*., 1995) through its interaction with Stn1p and Ten1p. *stn1* and *ten1* mutants display accumulation of extremely long G-overhangs similar to *cdc13-1* (Grandin *et al*., 2001; Grandin *et al*., 1997). Like Cdc13p, Stn1p contains an OB-fold, and both Stn1p and Ten1p bind telomeric DNA *in vitro* (Gao *et al*., 2007). Thus, it is proposed that Cdc13-Stn1-Ten1 complex acts as an RPA-like complex that is specific for single-strand telomeric DNA (Gao *et al*., 2007). Recent data indicate that STN1p and TEN1p can

provide telomere capping independent of Cdc13p (Larrivee and Wellinger, 2006; Petreaca *et al*., 2006; Zubko and Lydall, 2006).

In fission yeast, the OB-fold bearing protein POT1 functions in chromosome end protection: SpPOT1 deletion results in immediate erosion of telomeric and subtelomeric DNA and chromosome mis-segregation (Baumann and Cech, 2001). The surviving cells circularize all three chromosomes, thus achieving protection from further fusion reactions and genome instability (Baumann and Cech, 2001). Intriguingly, Stn1 and Ten1 homologs have been recently identified in *S. pombe*, and both proteins are essential for chromosome end protection (Martin *et al*., 2007).

Vertebrate POT1 also contributes to chromosome end protection. RNAi-mediated depletion of human POT1 activates a DNA damage response and triggers growth arrest (Churikov *et al*., 2006; He *et al*., 2006; Hockemeyer *et al*., 2006; Wu *et al*., 2006). Similarly, conditional knockout of chicken POT1 results in γH2AX accumulation at telomeres and cell cycle arrest (Churikov *et al*., 2006). It is currently believed that while TRF2 protects telomeres from ATM signaling, POT1 independently controls initiation of the ATR-mediated DNA damage response pathway (Churikov and Price, 2008; Churikov *et al*., 2006; Denchi and de Lange, 2007; Guo *et al*., 2007).

Vertebrate POT1 is also implicated in the control of 5' exonucleolytic processing of telomeres (Hockemeyer *et al*., 2005). However, the role of vertebrate POT1 in protection of telomeres from end-to-end fusions is controversial. While conditional knockout of chicken POT1 and RNAi-mediated depletion of human POT1 do not lead to telomere fusions (Churikov *et al*., 2006; Hockemeyer *et al*., 2005), some studies demonstrate end-to-end fusions and genome instability upon hPOT1 knockdown (Veldman *et al*., 2004; Yang *et al*., 2005).

Notably, mice encode two POT1 orthologs, which display ~75% similarity and appear to be partially redundant for chromosome end protection, with each contributing to abrogating a DNA damage response at telomeres. Furthermore, although somewhat conflicting results were obtained in different studies, mouse POT1a and POT1b are also implicated in protecting the C strand from nuclease attack and blocking homologous recombination at telomeres (He *et al*., 2006; Hockemeyer *et al*., 2006; Wu *et al*., 2006).

Due to extreme tolerance of plants to genome instability, the flowering plant *Arabidopsis thaliana* provides an attractive model system for studying of fundamental aspects of telomere capping (Riha *et al*., 2001). However, essential components of the shelterin complex have not been thoroughly elucidated in plants. In rice, deletion of the putative TRF2 homolog, RTBP1, results in telomere fusions and genome instability (Hong *et al*., 2007). Similarly, knockdown of NgTRF1, a RTBP1 homolog in tobacco, leads to reduction of cell viability (Yang *et al*., 2004). In contrast, the *Arabidopsis* homolog AtTBP1 appears to be dispensable for chromosome end protection and genome stability (Hwang *et al*., 2001; Karamysheva *et al*., 2004; Yu *et al*., 2000). Although several other *Arabidopsis* Myb-containing TRF-like proteins bind telomeric dsDNA *in vitro*, analysis of corresponding null mutants did not reveal chromosome end deprotection, likely due to genetic redundancy (reviewed in Zellinger and Riha, 2007). *Arabidopsis* encodes three POT1-like proteins. Although dominant-negative and over-expression studies implicate AtPOT1b and AtPOT1c in telomere capping (Shakirov *et al*., 2005; Chapters III and V), analysis of the null AtPOT1b mutant did not reveal any significant telomere deprotection defects, suggesting that this protein does not make a major contribution to chromosome end protection (Shakirov and Shippen, unpublished). Furthermore, AtPOT1a is a telomerase accessory factor that is dispensable for end

protection and is specialized for positive regulation of telomere length instead (Surovtseva *et al*., 2007). Finally, while all three POT1-like proteins harbor an OB-fold, none of them binds single-stranded telomeric DNA *in vitro* under conditions in which mammalian and yeast POT1 bind (Surovtseva *et al*., 2007); Y. Surovtseva and D. Shippen, unpublished).

Here we describe the identification and characterization of a novel chromosome capping protein in *Arabidopsis thaliana* critical for the integrity of telomeres (CIT1). *cit1* mutants display a profound telomere deprotection phenotype characterized by telomere length deregulation, extreme nuclease digestion, end-to-end chromosome fusions, and increased recombination. Intriguingly, *CIT1* encodes a novel protein whose only distinguishing feature is a putative OB-fold. Orthologs of CIT1 can be discerned in a wide variety of high eukaryotes, indicating that CIT1 represents a novel component of the telomere cap.

Materials and methods

Mutant lines, genotyping, DNA and RNA extraction, and RT-PCR

The *cit1-1* line was identified in the TILLING collection (Till *et al*., 2003). For genotyping, a genomic region flanking the *cit1-1* point mutation was amplified with CIT1_M2 fwd (5'-GTAATGCCCATCTCAAGTTTTG) and CIT1_M2_rev (5'-CAGCACACGCATAGCACTATG) primers and sequenced with the CIT1_M2 rev primer. *cit1-2* and *cit1-3* lines were found in the SALK database (stock lines SALK_114032 and SALK_083165, respectively). Genotyping was performed with T-DNA and gene-specific primers. The *tert* line was previously described (Riha *et al*., 2001). A genetic cross was performed between plants heterozygous for *tert* and for *cit1-1*. Plants were

grown under 16 hours light/ 8 hours dark conditions at 23^{o}C. DNA from plants was extracted as previously described (Cocciolone and Cone, 1993). RNA samples were prepared using Plant RNA Purification Reagent (Invitrogen). Reverse transcription was performed using 2 μg of RNA, as described (Shakirov *et al*., 2005). *AtCIT1* cDNA was amplified in the PCR reaction with primers CIT1_start_fwd (5'-ATGGAGAACACCACAATTCTCAC) and CIT1_stop_rev (5'-TCAGCTATTTAGCAAACCTTGGAG). To evaluate expression of the region flanking the T-DNA insertion in the *cit1-2* allele, primers 5'-GTCACGCTTTTGAGAGGTCTG and CIT1_M2_rev were used. For the *cit1-2* allele, primers CIT1_M2_fwd and 5'-CACTTGAGGAACTTATCCTCTG were utilized.

Map-based cloning

Map-based cloning was performed essentially as described (Lukowitz *et al*., 2000). Briefly, a mutant line (*Columbia* (Col) ecotype) was out-crossed to wild type *Arabidopsis Landsberg erecta* (Ler) ecotype. F1 plants were self-propagated to F2. Pools of wild type and mutant plants were generated (~50 plants in each pool) for bulked segregant analysis. CIW5 and CIW6 markers were identified as markers linked to the mutation. 150 individual mutant plants were subsequently used to find recombinants in the genomic interval between CIW5 and CIW6. The region containing the mutation was mapped by creating and analyzing new markers. Primer sequences of mapping markers are available upon request.

TRF, PETRA, quantitative TRAP, and in-gel hybridization

TRF was conducted as previously described (Fitzgerald *et al*., 1999). Subtelomeric TRF analysis was performed using a 1L probe (Surovtseva *et al*., 2007), or 5R probe (Shakirov and Shippen, 2004). For PETRA, 2 μg of DNA was used. PETRA conditions are described in (Heacock *et al*., 2004). Quantitative TRAP assay was performed as described (Kannan *et al*., 2008). For in–gel hybridization, HaeIII- and HinfI-digested DNA was separated on a native agarose gel. The dried gel was hybridized with a ss $(C_3TA_3)_3$ ^{32}P 5' end labeled oligonucleotide as discussed (Churikov *et al*., 2006).

Telomere fusion PCR, cytogenetics, and FISH

Telomere fusion PCR was performed as previously described (Heacock *et al*., 2004). PCR products were purified, cloned into pDrive vector (Quiagen), and sequenced. For cytological analysis, spreads were prepared from pistils as described (Riha *et al*., 2001). Chromosomes were stained with DAPI (4',6'-diamidino-2-phenylindole) for cytogenetic analysis or denatured and hybridized with the cocktail of subtelomeric probes for FISH.

Telomeric circle assays

For TCA and bubble trapping, DNA was digested with Alu1. TCA was performed using 50 μg of DNA as described in (Zellinger *et al*., 2007). For the bubble trapping technique (Mesner *et al*., 2006), 100 μg of DNA was used. Equal volumes of DNA and 1% low-melt agarose were equilibrated at 45^{o}C, mixed, and loaded on 0.6% agarose gel. The gel was run at 20 V at 4^{o}C for 16 hours. DNA was then transferred to the nylon membrane and hybridized with a G-rich telomeric probe.

Methylation analysis of pericentromeric repeats

The methylation status of 180-bp centromere repeats was analyzed as previously described (Vongs *et al*., 1993). Genomic DNA was digested with HpaII and subjected to Southern blotting as discussed.

Results

Identification of a novel gene critical for integrity of telomeres

As part of the characterization of AtPOTc, we screened a TILLING collection of EMS-mutagenized *Arabidopsis* plants, and identified a line bearing a C to T mutation in the third intron of the *AtPOT1c* gene. A fraction of these mutants showed a profound telomere uncapping phenotype (described below), but genetic analysis demonstrated that the phenotype did not segregate with the At*POT1c* genotype. Therefore, we used map-based cloning to identify the gene responsible for telomere uncapping. The mutation mapped to the ~2 mbp region between the markers NGA8 and CIW6 on chromosome 4. Subsequent analysis of 150 mutants narrowed the region to ~30 kb. Since the *AtPOT1c* gene is found on chromosome 2, these data confirm that the telomere uncapping phenotype is unlinked to *AtPOT1c*.

Sequence analysis of candidate genes within this region identified a single-nucleotide G to A mutation in the At4g09680 gene (Figure 36A). The gene was named *CIT1* (Critical for Integrity of Telomeres 1), and the mutant line was designated *cit1-1*. Consistent with its essential role in chromosome end protection, *CIT1* mRNA is expressed in all of the tissues examined, although at lower level in suspension culture cells (Figure 36B). Sequence analysis of *CIT1* cDNA revealed a large ORF with 16

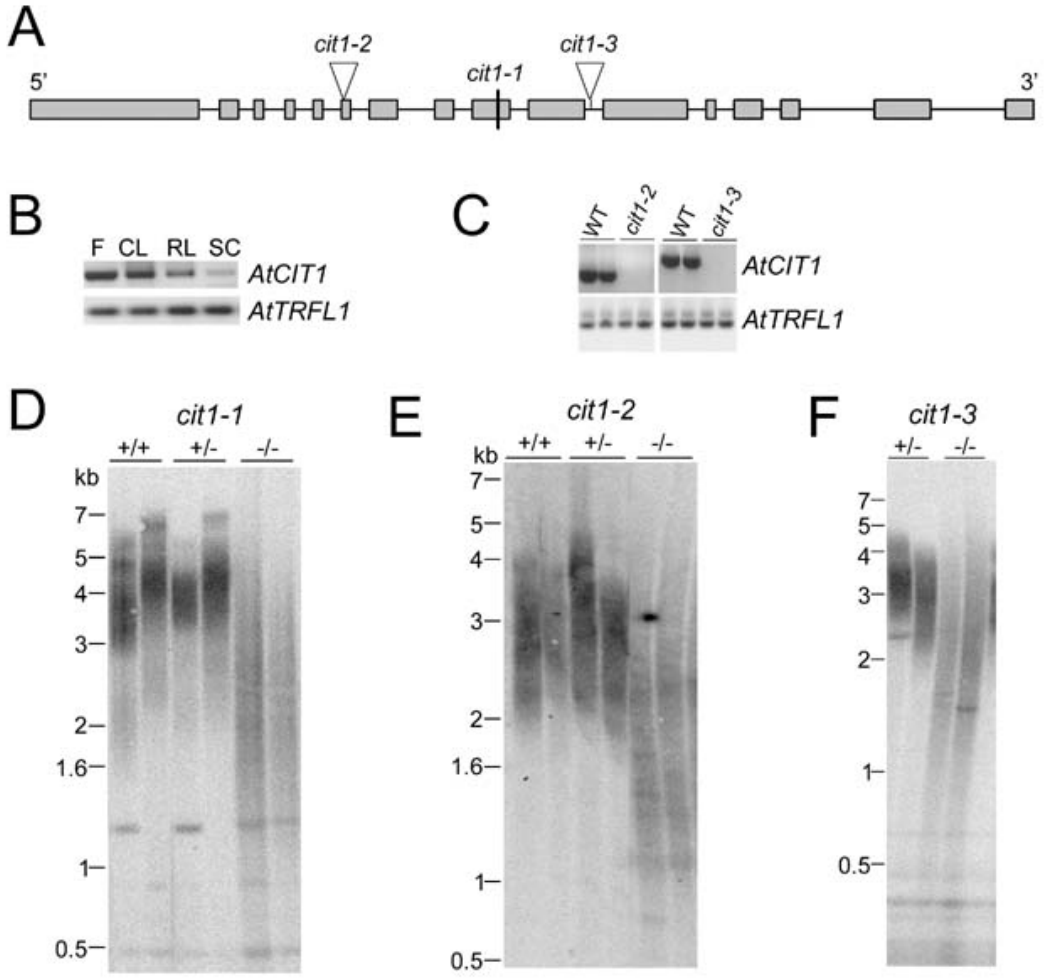


Figure 36. Telomere length deregulation in CIT1-deficient *Arabidopsis*. (A) Schematic of the *AtCIT1* gene locus. Rectangles represent exons; horizontal black lines are introns. The positions of a point mutation (*cit1-1*) and T-DNA insertions (*cit1-2* and *cit1-3*) are shown. (B) RT-PCR analysis of the *AtCIT1* gene expression in different plant tissues. F, flowers; CL, cauline leaves; RL, rosette leaves; SC, suspension culture. (C) RT-PCR analysis of *AtCIT1* gene expression in *cit1-2* and *cit1-3* mutants. Primers flanking the insertion were used in both cases. (D) TRF analysis of a *cit1-1* mutant. (E) TRF analysis of a *cit1-2* mutant. (F) TRF analysis of a *cit1-3* mutant. In panels D, E and F, results for progeny segregated from a parent heterozygous for *cit1* are shown. Blots were hybridized with a radiolabeled telomeric DNA probe. Molecular weight markers are indicated.

exons that encodes an ~142 kDa protein. In *cit1-1* mutants, a G(1935)A mutation resides in the 9th exon and produces a stop codon. From the SALK database, two additional *CIT1* alleles bearing T-DNA insertions in the sixth exon or tenth intron were identified and designated *cit1-2* and *cit1-3*, respectively (Figure 36A). RT-PCR analysis revealed that no *CIT1* full length mRNA is produced in *cit1-2* or *cit1-3* mutants, arguing that these plants are null for AtCIT1 (Figure 36C).

Terminal restriction fragment (TRF) analysis was used to examine bulk telomere length in *cit1* mutants derived from a self-pollinated heterozygous parent. TRF analysis of multiple *cit1-1* individuals revealed gross deregulation of telomere length in homozygous *cit1-1* mutants relative to their heterozygous and wild type siblings (Figure 36D). Homozygous *cit1-2* and *cit1-3* mutants showed an aberrant telomere phenotype similar to *cit1-1* (Figure 36E and F). Altogether, these data indicate that At4g09680 encodes *Arabidopsis CIT1*.

cit1 *mutants display severe defects in telomere architecture*

To determine how individual telomeres were affected by CIT1 depletion, subtelomeric TRF analysis was performed using probes directed at specific chromosome termini. As expected (Shakirov and Shippen, 2004), a relatively sharp single band is detected for wild type telomeres. In contrast, telomeres in *cit1* mutants gave rise to an extremely heterogeneous hybridization signal that spanned ~ 1.5 kb (Figure 37A). Individual telomere tracts were also examined by primer extension telomere repeat amplification (PETRA), a PCR based technique that monitors telomere length on individual chromosome ends (Heacock *et al*., 2004). As with subtelomeric TRF, the PETRA

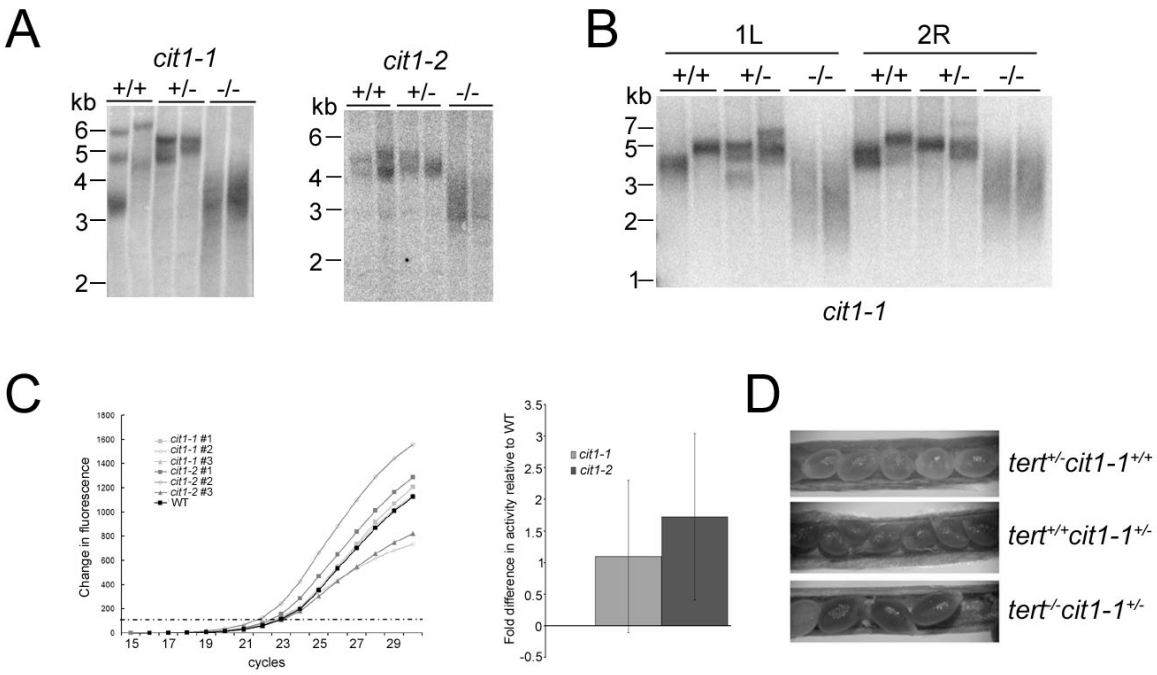


Figure 37. *cit1* mutants display telomere length deregulation on individual chromosome arms despite wild type levels of telomerase activity *in vitro*. (A) Subtelomeric TRF analysis of DNA from *cit1-1* and *cit1-2* mutants. DNA blots were hybridized with a probe corresponding to subtelomeric regions on the right arm of chromosome 5 (left panel) or the left arm of chromosome 1 (right panel). (B) PETRA analysis of DNA from *cit1-1* mutants. Results for telomeres on the left arm of chromosome arm and the right arm of chromosome 2 are shown. (C) Results of real time TRAP on *cit1-1* and *cit1-2* mutants. Left panel shows raw results. Dashed line represents the threshold change in fluorescence. Quantification of the telomerase activity levels in *cit1* mutants relative to wild type is shown on the right. (D) Siliques (seed pods) from plants heterozygous for *tert* (top), heterozygous for *cit1-1* (middle), or double $tert^{-/-}$ $cit1\text{-}1^{+/-}$ mutants. In contrast to uniform seed size observed in single mutants, double mutants displayed a reduced seed set and embryonic lethality.

products generated from *cit1* mutants were broad smears, confirming that telomere length is grossly deregulated (Figure 37B).

The unusual heterogeneity of telomere tracts in *cit1* mutants could be caused by deregulation of telomerase. However, quantitative TRAP demonstrated that *in vitro* telomerase activity is unchanged in *cit1* mutants compared to wild type (Figure 37C). To investigate whether telomerase contributes to *in vivo* telomere length deregulation in *cit1* mutants, plants heterozygous for *cit1-1* were crossed to homozygous *tert* mutants (Riha *et al*., 2001). Double heterozygous mutants were identified in F1 and allowed to self-pollinate to generate F2 progeny. Unexpectedly, we failed to recover any *tert cit1-1* double mutants (0/96 plants). Examination of developing embryos in the siliques of $tert^{-/-}$ $cit1\text{-}1^{+/-}$ mutants indicated that the double deficiency of TERT and CIT1 is embryonic lethal (Figure 37D). Thus, telomerase is required for the viability of *cit1* mutants.

To determine whether telomeres in *cit1* mutants are subjected to increased recombination, we looked for the presence of extra-chromosomal telomeric circles (ECTC) using TCA (t circle amplification) analysis (Zellinger *et al*., 2007). In this procedure, telomere sequences are amplified by phi29, a polymerase with strand displacement activity that generates high molecular weight ssDNA products from a circular template. As a positive control, TCA was performed on DNA from *ku70* mutants which have previously been shown to accumulate ECTCs (Zellinger *et al*., 2007). A high molecular weight DNA band was detected in *ku70* and *cit1* DNA samples, but not in wild type (Figure 38A). To verify the presence of ECTCs in *cit1* mutants, we employed a bubble trapping technique (Mesner *et al*., 2006). For this procedure, DNA is loaded on a agarose gel in a low-melt agarose plug and then subjected to hybridization with a telomeric probe. Linear fragments enter the gel, but the circular

DNA does not (Mesner *et al*., 2006). As anticipated, a telomeric signal was detected in the well with DNA from *cit1* and *ku70* mutants, but not with wild type (Figure 38B). These data confirm that ECTCs accumulate in the *cit1* background and argue that loss of CIT1 results in elevated rates of homologous recombination at telomeres.

Since PETRA requires the presence of a G-overhang, we can conclude that this structure is present at telomeres in *cit1* mutants. However, to more precisely monitor the status of the G-overhang, in-gel hybridization was employed (Figure 38C). Strikingly, the G-overhang signal was ~five times stronger in *cit1* mutants than in wild type (5.9+/-2.5), indicating that telomere architecture is grossly perturbed in the absence of AtCIT1. Although we currently can not rule out the possibility that the increased G-overhang signal is dependent on telomerase, we favor the model that CIT1 protects the telomeric C-strand from nucleolytic degradation similar to *S. cerevisiae* Cdc13, Stn1 and Ten1 proteins (Garvik *et al*., 1995; Grandin *et al*., 2001; Grandin *et al*., 1997).

cit1 *mutants display end-to-end chromosome fusions*

In *Arabidopsis*, telomeres below 1 kb in length are prone to end-to-end chromosome fusions (Heacock *et al*., 2007). Since bulk telomere length analysis of *cit1* mutants demonstrated that a substantial fraction of telomeres dropped below this critical length threshold, we looked for evidence of mitotic abnormalities. Anaphase bridges were scored in four different individual *cit1-1* mutants and in their wild type siblings. As expected, there was no evidence of genome instability in wild type plants, but in all four *cit1-1* mutants a high fraction of mitotic cells (up to 40%) displayed anaphase bridges (Figure 39A). Many anaphases contained multiple bridged chromosomes as well as

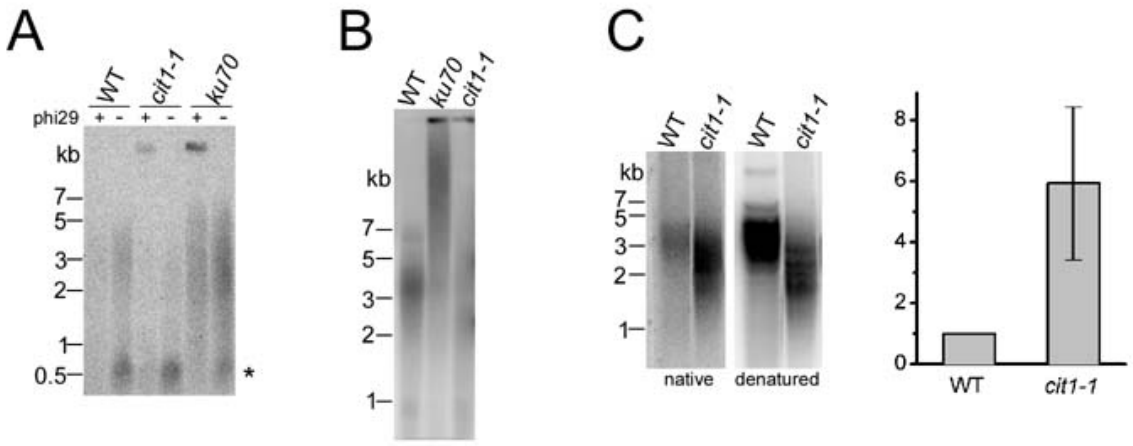


Figure 38. *cit1-1* mutants display elevated telomere recombination and G-overhang lengthening. (A) TCA analysis of *cit1-1* DNA. Reactions were performed in the presence or absence of phi29 polymerase. *ku70* DNA was used as a positive control. The asterisk indicates an interstitial telomeric repeats signal. (B) Bubble trapping results for *cit1-1* and *ku70* mutants. (C) G-overhang analysis using in-gel hybridization. Native and denatured gels are shown on the left. Quantification of the *cit1-1* signal relative to wild type (the average of results from six independent experiments) is shown on the right. All blots were hybridized with a radiolabeled telomeric probe. In A and B panels, the probe hybridizes to both circular and linear telomeric DNA products.

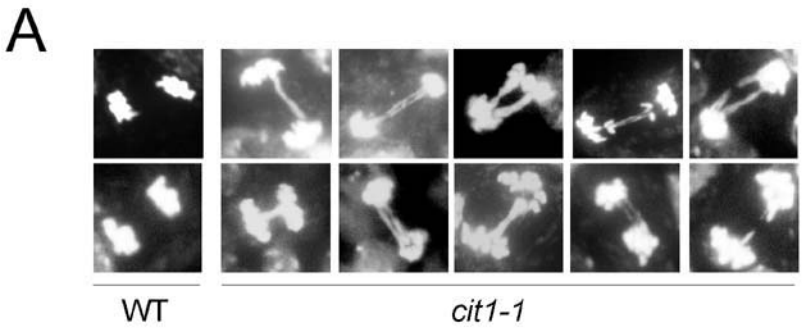
instances of unequal chromosome segregation. Such profound genome instability is comparable to the ninth generation (G9) of telomerase mutants (Riha *et al*., 2001).

The DNA composition of anaphase bridges was examined by fluorescence *in situ* hybridization (FISH). A cocktail of eight different subtelomeric BACs was used for hybridization of mitotic *cit1-1* cells. Hybridization signals were detected at all anaphase bridges, suggesting that end-to-end fusions are prevalent (Figure 39B) and that the telomere dysfunction is the primary cause of genome instability in *cit1* mutants.

To examine the nucleotide sequence at the fusion junction, telomere fusion PCR was performed with subtelomeric primers directed at individual chromosome arms. Abundant fusion products were obtained with DNA extracted from *cit1* homozygous plants, but not from heterozygous or wild type siblings (Figure 39C and data not shown). To characterize the structure of fusion junctions, the fusion PCR products were cloned and sequenced. Unexpectedly, among 17 clones analyzed, we failed to detect junctions involving direct fusion of telomere repeats. Instead, telomere-subtelomere fusions (5%) and subtelo-subtelo fusions (95%) were recovered. The fusions were characterized by extensive erosion of subtelomere sequences (752 bp average). For comparison, in G9 telomerase mutants, telomere-subtelomere fusions are the most prevalent (78%), and the average erosion of subtelomeric DNA sequences is only 290 bp (Heacock *et al*., 2004). Altogether, these data suggest that chromosome ends are subjected to extensive exonucleolytic degradation prior to fusion in *cit1* mutants.

cit1 *mutants display a fasciated phenotype consistent with epigenetic alterations*

In late generation telomerase mutants, loss of telomere integrity and the accompanying genome instability leads to morphological abnormalities (Riha *et al*., 2001). Defects in



	# of analyzed pistils	# of anaphases		% anaphase bridges
		with bridges	total scored	
cit1-1 #1	4	50	127	39
cit1-1 #2	6	95	395	24
cit1-1 #3	3	80	278	29
cit1-1 #4	4	54	191	28
WT	4	1	140	0

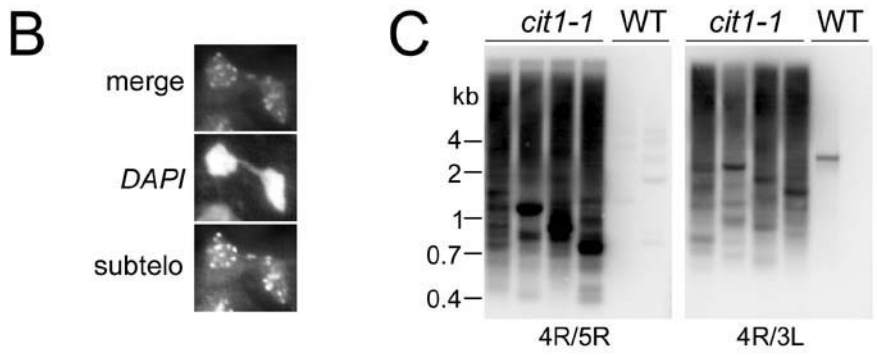


Figure 39. Telomere fusions in *cit1-1* mutants. (A) Cytogenetic analysis of *cit1-1* mutants. Top panel shows an analysis of DAPI-stained anaphase figures. Bridges are detected in *cit1-1* but not in wild type. The frequency of anaphase bridges in four *cit1-1* mutants is shown in the bottom panel. The total number of scored anaphases and the percentage of anaphases containing bridges are shown. (B) FISH analysis of an anaphase bridge from *cit1- 1* using probes for subtelomeric DNA sequences. Hybridization signal is detected in the bridge. (C) Telomere fusion PCR analysis of *cit1-1* mutants. Primers specific for 4R and 5R (left) or 4R and 3L were used.

both vegetative and reproductive organs first appear in G6 and worsen in successive generations. Ultimately, *tert* mutants reach a terminal phenotype in G8-10, where they arrest in a dedifferentiated vegetative state without producing an influorecsence bolt, or with a short bolt bearing sterile flowers (Riha *et al*., 2001) (Figure 40A). In contrast, *cit1* mutants displayed an immediate onset of severe morphological defects and were nearly completely sterile in the first generation. Moreover, the phenotypical abnormalities associated with CIT1 deficiency are distinct from terminal generation *tert* mutants (Figure 40A). Although most *cit1* mutants produced an influorescence bolt, it was highly variable in size, ranging from very short to indistinguishable from wild type (Figure 40B, compare right and left bottom panels). In addition, mutants exhibit grossly distorted phyllotaxy with an irregular branching pattern, fused flowers and siliques, and fasciated (thick and broad) main and lateral stems and siliques (Figure 40B). The vegetative defects were accompanied by abnormalities in reproductive organs. Flowers were often fused, anthers and ovules were undeveloped. The germination efficiency of the few seeds that could be recovered was extremely low, making propagation to the next generation almost impossible.

The fasciated phenotype is a characteristic feature of mutations that alter the epigenetic status of chromatin in *Arabidopsis* (Takeda *et al*., 2004). Consistent with epimutations, the severity of phenotypical defects varied widely between different *cit1* mutants (Figure 40B) as well as between different branches of the same mutant plant (Figure 40C). Therefore, we tested whether mutations in general chromatin-modifying proteins or DNA methylation factors would lead to a similar chromosome end deprotection phenotype. Telomere fusion PCR and TRF analysis performed on several known *Arabidopsis* epigenetic mutants (Cao and Jacobsen, 2002; Kaya *et al*., 2001;

Takeda *et al*., 2004; Vaillant *et al*., 2006; Vongs *et al*., 1993) failed to reveal any signs of telomere dysfunction (Figure 41A and B). We next asked the reciprocal question, whether inactivation of CIT1 leads to global epigenetic alterations. While loss of DNA methylation is associated with *ddm* mutants (Vongs *et al*., 1993), this is not the case for *cit1-1* mutants (Figure 41C). Furthermore, loss of CIT1 does not lead to repression of silencing of transposable elements (data not shown). These data argue that CIT1 inactivation does not result in global epigenetic perturbations in the *Arabidopsis* genome. However, further analysis will be needed to determine whether CIT1 plays a role in modulating chromatin structure at telomeric and subtelomeric regions.

CIT1 *is found in other plants and vertebrates*

TBLASTN analysis using the CIT1 protein sequence as a query revealed *CIT1* homologs in rice (E-value=4e-64) and grape (E-value=9e-60). Searches against the EST databases identified *CIT1* homologs in numerous other plant species, including *Physcomitrella*, the oldest living plant. These findings suggest that CIT1 is conserved in the plant kingdom. A putative *CIT1* ortholog was also found in the chicken genome (*GgCIT1*) (E-value=0.94). TBLASTN with *GgCIT1* as the query uncovered putative homologs in numerous other vertebrates, including primates, rodents, birds, amphibians and fish. Notably, we failed to detect a CIT1 homolog in yeast, *C. elegans* or in *Drosophila melanogaster*, suggesting that CIT1 is confined to vertebrates and plants. None of the putative CIT1 homologs encode a conserved functional domain that provides a possible clues for the mechanism of CIT1 action. We therefore used the 3D-PSSM program (Kelley *et al*., 2000) to predict a secondary structure of AtCIT1 protein. The analysis predicted the presence of the OB-fold in the middle portion of the protein

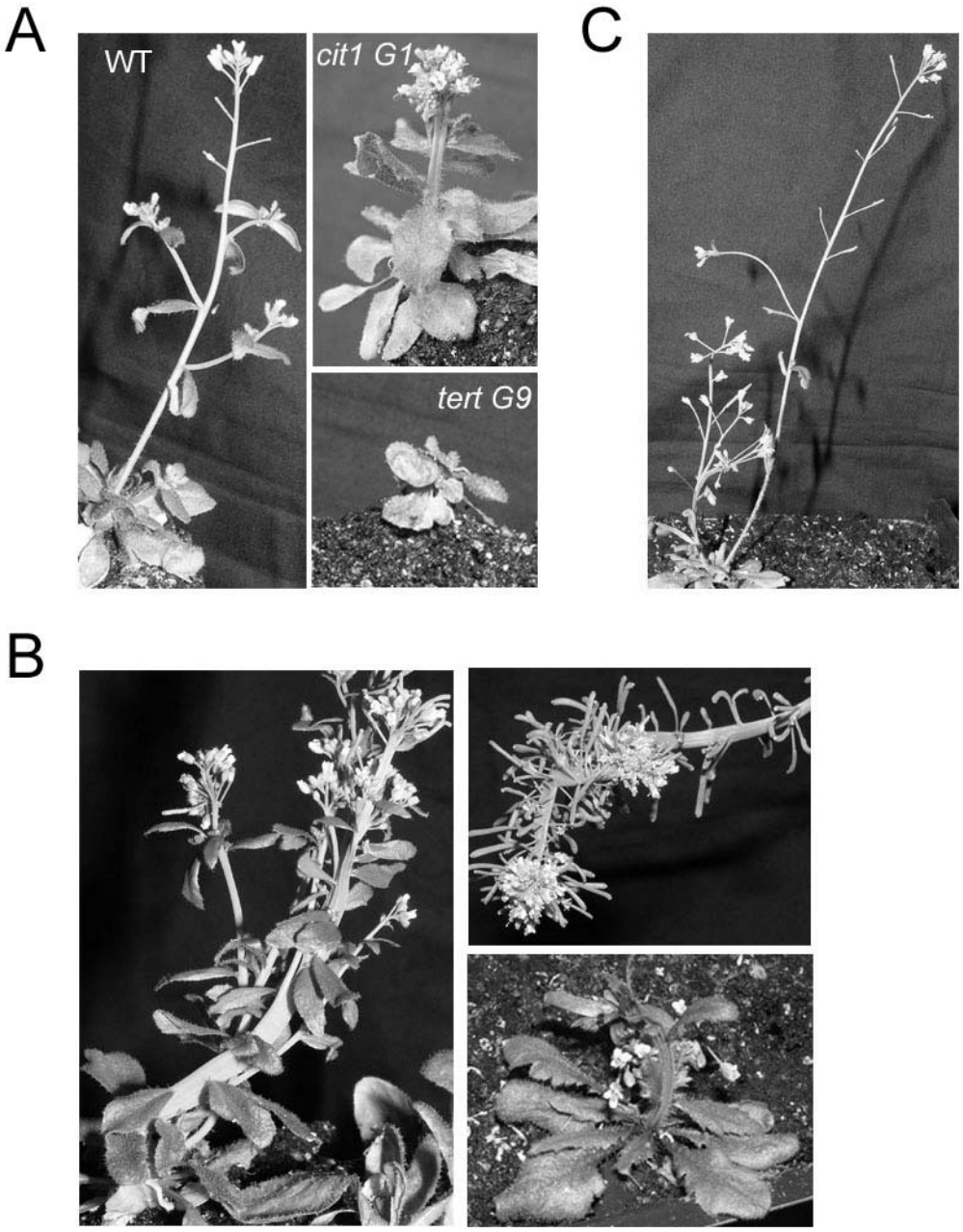


Figure 40. Morphological defects in *cit1* mutants. (A) Wild type, first generation *cit1* and ninth generation *tert* plants at 3-4 weeks of age. (B) Fasciated stems and fused organs in *cit1-1* mutants. The level of phenotypical defects varies dramatically between different *cit1-1* plants. (C) *cit1-2* mutant. The severity of defects varies between different branches of the plant.

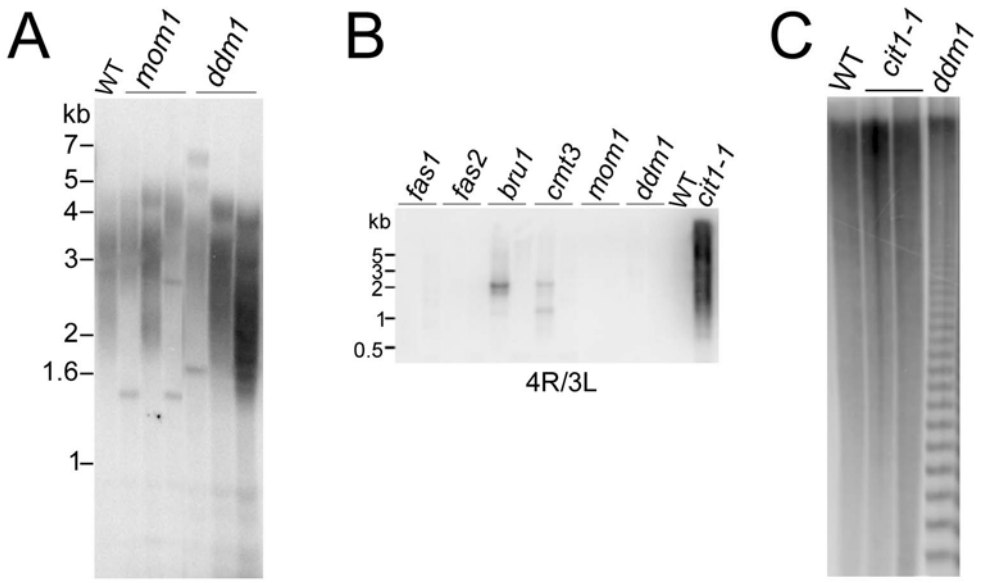


Figure 41. Telomere dysfunction does not result from global epigenetic perturbations. (A) TRF analysis of *mom1* and *ddm1* epigenetic mutants. (B) Fusion PCR analysis. No telomere fusions are observed in *Arabidopsis* mutants known to affect epigenetic regulation. (C) Analysis of centromere repeat methylation in *cit1-1* and *ddm1* mutants. The methylation profile is similar in wild type and in *cit1-1*.

similar to telomere end binding protein TEBP, the protein implicated in chromosome end protection in *O. nova* (Gottschling and Zakian, 1986; Horvath *et al*., 1998; Price and Cech, 1987). This observation suggests that CIT1 provides a capping function for *Arabidopsis* telomeres by direct interaction with telomeric DNA or a telomere-binding protein.

Discussion

A key property of telomeres is their ability to distinguish linear chromosome ends from double-strand breaks. This function is mediated by telomere-associated proteins that form a protective cap at the chromosome terminus. Although the overall organization of telomeres is highly conserved, the telomere complex is undergoing extraordinarily rapid evolution, which has resulted in extensive divergence in protein composition in different organisms. Although several essential components of the telomere complex have been defined in yeast and mammals, identification of proteins needed for chromosome end protection in plants has lagged. Key components of shelterin complex in mammals, TRF2, RAP1, TPP1 or TIN2, can not be discerned in *Arabidopsis* genome. It is possible that rapid divergence of these proteins precluded recognition by available sequence alignment tools. Alternatively, novel proteins may have evolved to provide chromosome end protection in plants.

In this study, we uncovered a novel telomere capping protein CIT1 in *Arabidopsis*. Disruption of CIT1 leads to a severe telomere deprotection phenotype reminiscent of a deficiency in the ds or ss telomere capping proteins found in other organisms. Similar to *S. cerevisiae* Cdc13 and *S. pombe* POT1 inactivation (Baumann and Cech, 2001; Garvik *et al*., 1995), telomeres in *cit1* mutants are subjected to

massive nuclease attack, resulting in both extensive C-strand resection and loss of telomeric and subtelomeric sequences. FISH and fusion PCR analysis detect a strong bias for fusions involving subtelomeric regions of DNA. Moreover, like *S. pombe* POT1 deletion, loss of CIT1 leads to profound genome instability and end-to-end chromosome fusions (Baumann and Cech, 2001). However, similar to TRF2 (Wang *et al*., 2004), AtCIT1 appears to protect telomeres from homologous recombination: loss of CIT1 results in the accumulation of ECTCs, the hallmark of increased telomere recombination. Altogether, these data demonstrate that AtCIT1 is essential for chromosome end protection in *Arabidopsis*.

Unexpectedly, we found that telomerase is required for viability of *cit1* mutants. It is possible that telomerase is able to rescue a fraction of the critically shortened telomeres in *cit1* mutants by telomere elongation. Alternatively, telomerase may play a more direct role in chromosome end protection as has been hypothesized in mammals (Blackburn, 2005; Hsu *et al*., 2007; Kim *et al*., 2003; Zhu *et al*., 1999). Thus, the combined loss of CIT1 and telomerase may completely abrogate chromosome end protection.

A key observation uncovered in this study is a fasciated phenotype displayed by *cit1* mutants. *cit1* mutants display features consistent with altered chromatin in *Arabidopsis* (Kaya *et al*., 2001; Takeda *et al*., 2004). Although we ruled out global epigenetic changes in *cit1* mutants, it is possible that telomeric chromatin is specifically altered upon CIT1 loss. This is intriguing since increasing evidence suggests the relationship between telomere chromatin state and telomere function (reviewed in Blasco, 2007). While epigenetic DNA and histone modifications influence telomere length homeostasis, the epigenetic status of telomeric chromatin is also affected by

telomere length. It is therefore conceivable that an abrupt reduction in telomere length leads to changes in telomeric chromatin in *cit1* mutants. Alternatively, CIT1 could be involved in the regulation of telomere-specific chromatin-modifying activities. We favor the first model since disruption of the *Arabidopsis* STN1 homolog leads to morphological defects extremely similar to the defects in *cit1* mutants (X. Song and D. Shippen, unpublished). Thus, the change in epigenetic state of telomeric chromatin may be a common consequence of telomere uncapping. In this regard, it is intriguing that decrease in subtelomeric DNA methylation results in increased levels of homologous recombination at mammalian telomeres (Gonzalo *et al*., 2006). It will be interesting to see whether the change in DNA methylation correlates with elevated telomeric recombination observed in *cit1* mutants.

Remarkably, the morphological defects observed in terminal germination of telomerase mutants significantly differ from CIT1 or STN1 deficiency (Riha *et al*., 2001). Therefore, the gradual telomere uncapping associated with progressive erosion of telomeric DNA in plants lacking telomerase appears to be distinct from abrupt telomere uncapping observed upon deletion of essential telomeric proteins. It has been argued that deprotected telomeres are recognized by DNA repair machinery in different ways depending on the context of telomere dysfunction. While fusion of deprotected telomeres upon mouse TRF2 loss requires ligase 4 (Celli and de Lange, 2005), dysfunctional telomeres that arise in telomerase-deficient mice can activate additional ligase 4-independent DNA repair pathways (Maser *et al*., 2007). Therefore, the mode of telomere deprotection may influence the molecular and physiological outcomes of telomere uncapping. It is conceivable, for example, that gradual telomere shortening

allows the cell to adapt to the changing status of the telomeric chromatin, while immediate telomere uncapping does not provide an opportunity for such a response.

It is unknown how CIT1 provides protection for the chromosome end. One possibility is that CIT1 indirectly affects the chromosome end protection by regulating the stability and/or function of essential telomere protein. For example, tankyrase 1 is a poly (ADP-ribose) polymerase that modifies TRF1 protein, thus decreasing its ability to bind telomeric DNA. Since TRF1 is a negative regulator of telomere length in vertebrates, over-expression of tankyrase leads to increase in telomere length (Smith and de Lange, 2000). Another more intriguing possibility is that CIT1 physically associates with telomeres and directly protects chromosome ends. The presence of an OB-fold in the CIT1 protein is consistent with its direct interaction at the telomere. In addition to three POT1-like *Arabidopsis* proteins and the STN1 protein, CIT1 appears to be yet another OB-fold containing protein required for proper telomere function. Several reports have suggested that in both fission yeast and budding yeast, telomeres are protected by multiple OB-fold containing complexes (Larrivee and Wellinger, 2006; Martin *et al*., 2007; Petreaca *et al*., 2006; Zubko and Lydall, 2006). Thus, CIT1 may function in the context of other OB-fold proteins to achieve full protection of chromosome termini.

It is also possible is that CIT1 is not a stable component of the telomere like TRF2 or POT1, but rather associates with telomeres in a cell-cycle dependent manner. It is now established that telomere structure and composition are highly dynamic and undergo significant changes during the cell cycle (Takata *et al*., 2004; Takata *et al*., 2005; Verdun *et al*., 2005; Verdun and Karlseder, 2006). During S phase, telomeres are thought to unfold their t-loops to assume an accessible state that allows extension

by telomerase and by conventional replication machinery. This more open configuration will make telomeres more vulnerable to nucleolytic processing and/or recombination. Indeed, telomeres are recognized and processed by DNA damage machinery in G2 of the cell cycle (Verdun *et al*., 2005; Verdun and Karlseder, 2006). The recruitment of DNA damage repair factors correlates with the decrease in telomere-associated POT1 (Verdun *et al*., 2005). Thus, transient association of CIT1 with telomeres in late S/early G2 could limit the DNA damage response, reduce C-strand resection, and allow telomeres to assume a protected state upon completion of telomere replication and processing. Although vertebrate TRF2 was proposed as a putative candidate for this function (Verdun *et al*., 2005), this has not yet been tested experimentally.

Whatever the precise mechanism for CIT1-mediated chromosome end protection, this protein is conserved among plants and has putative orthologs in vertebrates. Genetic and biochemical searches for telomere-associated factors have thus far failed to identify CIT1 homologs in other organisms (Liu *et al*., 2004; O'Connor *et al*., 2004; Ye *et al*., 2004b). One possibility is that CIT1 associates with the G-overhang and it was overlooked due to its sub-stoichiometric amounts. Another possibility, although not mutually exclusive, is that CIT1 association with telomeres is cell-cycle dependent (as discussed above). Although more work will be required to decipher the exact mechanism of chromosome end protection by CIT1, our data argue that we have yet to identify all of the key components of the telomere complex.

CHAPTER VII

CONCLUSIONS AND FUTURE DIRECTIONS

Since all known DNA polymerases require a primer to initiate DNA replication, loss of terminal sequences due to the removal of the primer is an inevitable outcome of the replication of linear genomes. Although multiple solutions for the end replication problem exist, most of eukaryotes have evolved the telomere as a protective cap to overcome the end replication problem (reviewed in de Lange, 2004). In most normal human cells, telomere sequences undergo progressive shortening with each cell division, which ultimately leads to telomere dysfunction, activation of DNA damage pathways, and cell death. In contrast, germ line, stem cells and most cancer cells express the telomerase enzyme which can extend telomeres, thus providing an unlimited proliferation capacity to the cell. Recent studies indicate that telomeres comprise a biological clock that influences cellular life span. As a consequence, understanding the structure of telomeres and their maintenance by telomerase has been a major focus in medical research as it may impact how we treat aging and cancer-related diseases (reviewed in Shay and Wright, 2005).

Although the composition of telomere proteins varies between different species, the overall architecture of the telomeric complex is evolutionarily conserved. Therefore, different model systems have been used to investigate telomere structure and function. While each model system has its advantages and disadvantages, knowledge obtained from these systems lays the groundwork for understanding fundamental mechanisms of telomere biology. In this dissertation, I described the characterization of telomere-

associated proteins in *Arabidopsis thaliana*. Here, I discuss the significance of these results and their importance in the telomere field.

Characterization of *Arabidopsis* ds telomeric DNA binding proteins

Identification of plant telomere-associated proteins has so far mostly relied on BLAST searches for sequence homologs of known telomere components from other species. A common feature of ds telomeric DNA binding proteins is the presence of the conserved Myb-type DNA binding domain (Bilaud *et al*., 1996). We therefore used a genomic approach to look for Myb-containing proteins in *Arabidopsis*. In Chapter II, I described the identification and biochemical characterization of twelve TRFL proteins, which contain a C-terminal Myb-domain and fall into two categories based on their sequence and biochemical properties (Karamysheva *et al*., 2004).

Despite the sequence similarity of TRFL proteins to vertebrate TRF1 and TRF2, and the ability of TRFL proteins from the family I to bind ds telomeric DNA *in vitro*, genetic analysis of single gene mutants failed to uncover a role for these proteins in telomere length regulation or in chromosome end protection (L. Vespa, R. Warrington and D. Shippen, unpublished). We suspect that the TRFL genes are redundant and at least a subset of corresponding proteins has overlapping functions. This is not surprising since the *Arabidopsis* genome has undergone several rounds of duplication, and many genes are present in more that one copy (Blanc and Wolfe, 2004). It might therefore be necessary to obtain plant lines with a deficiency in multiple TRFL proteins.

The papaya genome, which has recently been sequenced, is evolutionary close to *Brassicaceae*, the order to which *Arabidopsis* belongs. Interestingly, papaya encodes only three TRFL genes which appear to be ancestors of TBP1, TRFL2 and

TRFL9 in *Arabidopsis thaliana* (E. Shakirov and D. Shippen, unpublished). It is therefore possible that TBP1, TRFL2 and TRFL9 are the key TRF proteins in *Arabidopsis*. A mutant that lacks all three of these genes is now being analyzed to test this hypothesis (R. Warrington and D. Shippen, unpublished).

An alternative approach currently being pursued is the over-expression of dominant negative alleles of the TRFL proteins that displace the TRFL proteins from telomeres. Since the Myb and Myb-extension are conserved in all TRFL proteins, this domain was chosen for the over-expression. Preliminary data indicate that truncated proteins localize at telomeres and cause telomere shortening (R. Warrington, J. Lamb and D. Shippen, unpublished). However, further work is needed. Localization studies and telomeric FISH analysis, together with chromatin immunoprecipitation will also be required to decipher which of the TRFL proteins bind telomeres *in vitro*. Since these proteins are the putative ds telomeric DNA binding proteins, they should be present at telomeres in the amounts sufficient for these types of experiments.

Since our study of TRFL proteins was begun, a putative TRF1 homolog was identified in rice (RTBP1) and *Arabidopsis* (AtTBP1) by two other labs. Disruption of RTBP1 leads to longer telomeres in the first generation, but the length new set point is stabilized in the subsequent generations (Hong *et al*., 2007). In contrast, telomere lengthening is only observed in the third generation of *tbp1* mutants in *Arabidopsis* (Hwang and Cho, 2007). Rice TBP1 appears to contribute to chromosome end protection: mutants display end-to-end fusions and genome instability (Hong *et al*., 2007). In contrast, no signs of genome instability were observed upon TBP1 inactivation in *Arabidopsis* (Hwang and Cho, 2007), arguing that this protein is

dispensable for chromosome end protection, or that it acts in concert with another telomere protein.

In summary, despite the progress in the identification of plant ds telomeric DNA binding protein, *Arabidopsis* TRF1 and TRF2 homologs have not been yet found. Rather, it is likely that multiple proteins fulfill these roles in plants. Identification of the ds telomere binding proteins proteins and their interaction partners will be necessary for complete understanding of the structure and function of plant telomeres.

Characterization of *Arabidopsis* ss telomeric DNA binding proteins

In Chapters III, IV and V, I described the characterization of three POT1-like proteins in *Arabidopsis*, AtPOT1a, AtPOT1b and AtPOT1c. Although these proteins were identified based on their sequence similarity to POT1 protein from humans and fission yeast, our data suggest that *Arabidopsis* POT1 proteins have evolved distinct functions in telomere biology.

Role of Arabidopsis *POT1a*

Although the AtPOT1a gene was found in 2001 (Baumann and Cech, 2001), the null mutant was not discovered until 2006, blocking a thorough genetic analysis of the role of AtPOT1a protein in telomere biology for several years. Therefore, we took an alternative method to investigate the function of AtPOT1a, using a transgenic approach to produce a dominant negative allele of AtPOT1a. A portion of AtPOT1a that lacks DNA binding domain (P1ΔN) was over-expressed in wild type plants. Telomere shortening in first-generation transgenic plants implicated AtPOT1a in the regulation of telomere length homeostasis (Shakirov *et al*., 2005). We hypothesized that P1ΔN

dislodged endogenous AtPOT1a and its interacting proteins involved in telomerase action at telomeres, leading to telomere shortening. These data indicated that AtPOT1a positively regulates telomere length in *Arabidopsis* (Shakirov *et al*., 2005).

Somewhat surprisingly, in contrast to *Arabidopsis* P1a∆N, over-expression of the corresponding region of human POT1 {P1(∆OB)} results in the opposite phenotype, telomere lengthening (Kelleher *et al*., 2005; Loayza and de Lange, 2003). A current model proposes that human POT1 localizes to telomeres primarily via interactions with ds telomeric DNA binding proteins. These interactions are mediated by the C-terminus of hPOT1 (Liu *et al*., 2004; Ye *et al*., 2004b). Subsequently, hPOT1 is loaded on the G-overhang, where it binds ss telomeric DNA in an OB-fold dependent manner (Loayza and De Lange, 2003a). Therefore, over-expression of the human POT1 mutant that lacks OB-fold still leads to the accumulation of the protein at telomeres, but the negative regulation of telomerase is abolished, leading to telomere lengthening. Telomere shortening observed upon over-expression of the similar region of AtPOT1a indicated that *Arabidopsis* POT1a regulates telomere length through a different mechanism.

Screening several *Arabidopsis* knockout collections allowed us to identify two mutant lines bearing T-DNA insertions in the *AtPOT1a* gene (Surovtseva *et al*., 2007). Consistent with the role of AtPOT1a in telomere length regulation, *pot1a* null mutants displayed ever-shorter-telomere phenotype that was remarkably similar to the telomerase deficiency (Surovtseva *et al*., 2007). Therefore, we investigated the relationship between AtPOT1a and TERT. Genetic analysis of single and double *pot1a* and *tert* mutant lines indicated that AtPOT1a works in the telomerase pathway. To explore the mechanism of AtPOT1a's role in telomerase regulation, we used a biochemical approach. We first showed that telomerase activity is reduced ~13-fold in

pot1a mutants (Surovtseva *et al*., 2007; A. Nelson and D. Shippen, unpublished). We also demonstrated that AtPOT1a is a physical component of telomerase RNP using co-immunoprecipitation. Furthermore, we showed that AtPOT1a is associated with telomeres *in vivo* only during S phase, when telomerase is known to act. Thus, these findings indicate that AtPOT1a has evolved a novel function – to physically associate with the telomerase RNP and play an essential role in telomerase function *in vitro* and *in vivo* (Surovtseva *et al*., 2007).

Evolution and function of AtPOT1a is one of the major areas of interest in our lab. The genetic and biochemical analysis described in Chapters III and IV uncovered an unexpected mechanism of AtPOT1a-mediated telomere maintenance. *Arabidopsis* POT1a was identified and named based on its sequence similarity to human and *S. pombe* POT1. However, in contrast to human and *S. pombe* proteins that associate with single-strand telomeric DNA with high affinity (Baumann and Cech, 2001; Lei *et al*., 2002), AtPOT1a does not bind telomeric DNA *in vitro*. This finding is surprising since AtPOT1a encodes two OB-folds, the structural motifs commonly used for binding to ss DNA. New data from our lab suggest that AtPOT1a interacts with *Arabidopsis* TER, the telomerase RNA subunit (C. Cifuentes-Rojas and D. Shippen, unpublished). Importantly, the N-terminus of AtPOT1a containing two OB-folds is sufficient for RNA binding. Moreover, a point mutation in *AtPOT1a* that corresponds to an important DNA-binding residue in human Pot1 abolishes binding of AtPOT1a to TER. These exciting data argue that *Arabidopsis* POT1a associates with the telomerase RNP via an interaction with TER, and suggest that the OB folds in AtPOT1a have evolved to allow protein binding the telomerase RNA instead of telomere.

Xiangyu Song in our lab uncovered a plant line with a point mutation in *AtPOT1a* that gives rise to an ever-shorter-telomere phenotype, identical to the phenotype of *pot1a* null mutants. This line was designated *pot1a-3*. The mutation in *pot1a-3* resides in the C-terminus of AtPOT1a and does not affect the OB-folds (X. Song and D. Shippen, unpublished). This finding validates our studies with the dominant negative allele of AtPOT1a (Pt1ΔN) and argues that the C-terminus of AtPOT1a protein is required for its role in the positive regulation of telomere length. Remarkably, *pot1a-3* mutation does not affect TER binding (C. Cifuentes-Rojas and D. Shippen, unpublished). It is possible that this mutation abolishes interaction of AtPOT1a with another protein important for telomerase recruitment to telomeres. In *S. cerevisiae*, telomerase is recruited to telomeres via an interaction of Est1p component of telomerase RNP with Cdc13p, the POT1 ortholog in budding yeast (Evans and Lundblad, 1999). Mutations that abolish Est1p-Cdc13p interactions result in ever-shorter-telomere phenotype due to inability of telomerase to extend telomeres *in vivo* (Chandra *et al*., 2001; Pennock *et al*., 2001). Similarly, the C-terminus of AtPOT1a might mediate the interaction with an *Arabidopsis* G-overhang binding protein, allowing telomerase recruitment and telomere elongation. Intriguingly, in mammals, TPP1 protein associates with the G-overhang binding protein POT1and physically interacts with active telomerase (Xin *et al*., 2007). Similar to *Arabidopsis* POT1a, TPP1 harbors a canonical OB-fold but cannot bind telomeric DNA on its own (Wang *et al*., 2007; Xin *et al*., 2007). Moreover, like AtPOT1a, TPP1 associates with the telomerase in an OB-fold-dependent manner and is required for telomere length control (Xin *et al*., 2007). Interestingly, no TPP1 homolog has been identified in *Arabidopsis* genome. It is therefore possible that AtPOT1a fulfils the role of TPP1 in *Arabidopsis*. Altogether,

these data argue that telomerase action at the chromosome terminus is regulated by interactions of telomerase-associated proteins with the G-overhang binding proteins, and that this interaction is evolving rapidly.

According to our model, AtPOT1a mediates telomerase recruitment to telomeres via an interaction with a G-overhang binding protein. AtPOT1b is an attractive candidate for this role, as its sequence is similar to human G-overhang binding protein POT1 (15% identity, 28% similarity). Moreover, AtPOT1b harbors two OB-folds at the N-terminus which are 13% identical and 29% similar to the corresponding region of human POT1. Finally, co-imunoprecipitation end yeast-two-hybrid experiments demonstrated that AtPOT1a and AtPOT1b interact with each other. However, AtPOT1b does not bind telomeric DNA *in vitro*, and genetic analysis of the null *pot1b* mutant argues against the role of this protein in telomerase recruitment (see below).

Identification of the AtPOT1a binding partner that mediates telomerase recruitment to the G-overhang will be necessary to understand the mechanism of telomerase regulation in *Arabidopsis*. A yeast-two-hybrid screen of the *Arabidopsis* cDNA library identified two putative POT1a-interacting factors (M. Jasti, X. Song and D. Shippen, unpublished). Interestingly, one of these proteins encodes a putative RNA-binding protein. Telomere length analysis in the corresponding null mutants will be required to understand what role, if any, these proteins play in telomere biology. Additional approaches, such as immunoprecipitation with AtPOT1a antibody and mass-spectrometry can also be used to identify AtPOT1a-interacting factors.

Another unanswered question is why telomerase activity is affected by AtPOT1a *in vitro*. Telomerase activity is decreased by approximately 10-fold in *pot1* mutants compared to wild type (Surovtseva *et al*., 2007; A. Nelson and D. Shippen,

unpublished). It is possible that AtPOT1 affects the processivity of repeat addition by telomerase. A non-PCR based telomerase activity assay will be required to test this hypothesis, but such assay is currently unavailable for plants. Another possibility is that AtPOT1a is involved in telomerase RNP assembly and/or stability. Since AtPOT1a binds telomerase RNA, it is conceivable that this protein is required for TER stability. It will therefore be interesting to check the levels of TER by reverse transcription-quantitative PCR in *pot1* mutants.

Establishing of the localization of AtPOT1a throughout the cell cycle may also help us to reveal the role of this protein in telomere biology. Does it co-localize with TERT all the time? Does it bind telomeres in a cell-cycle dependent manner? Our ChIP data suggest that AtPOT1a is enriched at telomeres in S phase. However, ChIP might not be sensitive enough to detect low levels of the protein at telomeres outside S phase. Additional experiments such as immunolocalization combined with telomeric FISH might be informative.

Altogether, our data demonstrate that in contrast to *S. pombe* and human POT proteins, AtPOT1a does not bind telomeric DNA, but rather associates with the telomerase RNP. Moreover, in contrast to POT proteins from other organisms, AtPOT1a is dispensable for chromosome end protection, but it has evolved a specialized role in positive regulation of telomere length by working in the context of telomerase RNP.

Role of Arabidopsis *POT1b*

G-overhang binding proteins from yeast and vertebrates are essential for chromosome end protection. Mutations in these proteins have outcomes characteristic of uncapped

and dysfunctional telomeres reviewed in (Wei and Price, 2003). In contrast, our analysis of *Arabidopsis pot1a* mutants demonstrated that this protein is dispensable for telomere capping. We therefore hypothesized that end protection in *Arabidopsis* is mediated by POT1b. As for AtPOT1a, we initially employed a transgenic approach to examine the function of AtPOT1b, since a null mutant was not available at the time. In Chapter III, I described the characterization of a AtPOT1b dominant negative mutant line over-expressing the DNA binding domain of AtPOT1b (P2ΔC). Plants over-expressing P2ΔC displayed severe morphological defects, sterility and a high incidence of anaphase bridges (Shakirov *et al*., 2005). We therefore concluded that AtPOT1b is implicated in chromosome end protection in *Arabidopsis*.

To further characterize the role of AtPOT1b in telomere biology, a plant line bearing a T-DNA insertion in the *AtPOT1b* gene was identified. Unexpectedly, analysis of the null AtPOT1b mutant did not detect any major defects chromosome end protection. Unlike dominant negative AtPOT1b mutants, plants null for AtPOT1b did not display any telomere length change or signs of genome instability (E. Shakirov and D. Shippen, unpublished). We hypothesize that the function of AtPOT1b is redundant with another *Arabidopsis* telomere capping protein. AtPOT1a is an unlikely candidate since double *pot1a pot1b* mutants display a *pot1a*-like phenotype without any signs of telomere uncapping (E. Shakirov and D. Shippen, unpublished). It is possible that AtPOT1b exerts its capping function in cooperation with AtPOT1c. Consistent with this hypothesis, over-expression of AtPOT1c results in chromosome end deprotection (Chapter V). It is conceivable that over-expression of P2ΔC dislodges both AtPOT1b and AtPOT1c, or somehow affects interacting proteins that are involved in end protection. A more definitive understanding of the role of AtPOT1b in telomere biology

will require localization studies of this protein in wild type and P2ΔC mutant backgrounds. Despite several attempts, we failed to produce a reliable AtPOT1b antibody. Therefore, expression of the tagged version of AtPOT1b driven by its native promoter in null mutants where the expression of the endogenous protein is inactivated will be necessary to study the localization of AtPOT1b.

Complete inactivation of both *AtPOT1b* and *AtPOT1c* will be necessary to test whether these genes are redundant for chromosome end protection in *Arabidopsis*. It is possible, however, that proteins other that POT1c associate with POT1b to provide chromosome end protection. A fascinating candidate is AtCIT1 (as discussed below). Search for AtPOT1b-interacting proteins by yeast-two-hybrid or through *in vivo* pull down experiments should shed some light on AtPOT1b function.

Role of Arabidopsis *POT1c*

In Chapter V, I described identification and characterization of a third *Arabidopsis* POT1 protein, AtPOT1c. *AtPOT1c* appears to be a very recent gene duplication of *AtPOT1a* specific for *Arabidopsis thaliana*. Analysis of plants over-expressing AtPOT1c demonstrated that like AtPOT1a, AtPOT1c works in telomerase pathway for telomere length regulation. We have previously shown that AtPOT1a is essential for telomerase action *in vivo*; null pot1a mutants display phenotype indistinguishable from *tert* deficiency (Surovtseva *et al*., 2007). In Chapter V, I showed that AtPOT1c over-expression leads to deregulation of telomere length at individual chromosome arms and a two-fold decrease in telomerase activity *in vitro*.

The exact relationship between AtPOT1a and AtPOT1c in telomere length regulation remains unclear. Very recent data from our lab suggest that both AtPOT1a

and AtPOT1c bind to the newly identified *Arabidopsis* telomerase RNA, implying that OB-folds in both of these proteins have similar features allowing TER binding (C. Cifuentes-Rojas and D. Shippen, unpublished). It is therefore conceivable that AtPOT1c may have evolved a negative role in regulating telomerase activity by competing with AtPOT1a for telomerase RNA association. We hypothesize that AtPOT1c displaces AtPOT1a, a positive regulator of telomerase, from the enzyme, leading to a reduction in the overall level of telomerase activity *in vitro* and *in vivo*. Analysis of the AtPOT1c null mutant will be required to decipher the exact contribution of AtPOT1c in telomere length regulation.

We have also shown that despite its recent emergence, AtPOT1c has evolved functions distinct from its ancestor AtPOT1a. While AtPOT1a is dispensable for chromosome end protection, AtPOT1c is involved in defining G-overhang structure, likely through the control of the C-strand processing (as discussed in Chapter V). Since reduction in the G-overhang signal was observed when AtPOT1c was over-expressed, we hypothesize that AtPOT1c over-expression results in the titration of the essential protective protein away from telomeres. Interestingly, while G-overhangs are intact in single *tert* mutants, over-expression of AtPOT1c in *tert* mutants renders G-overhangs almost undetectable. It therefore appears that *Arabidopsis* possess several partially redundant end-protection complexes, and telomerase itself appears to be one of them.

One of the puzzling observations described in Chapter V is the genome stability of *Arabidopsis* despite the devastating perturbations of telomere architecture and extreme shortening of the G-overhang. It has been previously shown that in mammals, accessibility of G-overhangs can trigger end-joining reactions (Zhu *et al*., 2003). Moreover, G-overhang exposure has been implicated in senescence in human cells (Li

et al., 2003). Lack of end-to-end fusions or any signs of genome instability in AtPOT1c over-expressing mutants suggests that the effects of the G-overhang depletion may be different in plants. It is possible that loss of G-overhang leads to DNA damage response at *Arabidopsis* telomeres. Localization studies of DDR proteins will be necessary to test this hypothesis.

Future experiments for understanding of the exact role of AtPOT1c in telomere biology will require an identification of the null mutant. I have designed and cloned four different RNAi constructs targeting coding region and 5' UTR of *AtPOT1c*. Characterization of putative AtPOT1c knockdown lines is currently underway. A plant was identified where reduced levels of AtPOT1c transcript correlate with dramatic telomere shortening, likely through nuclease attack (A. Nelson and D. Shippen, unpublished). Analysis of more AtPOT1c-depleted mutants will be required to confirm the role of AtPOT1c in chromosome end protection. It is also important to focus on identification of proteins that interact with AtPOT1c. In co-immunoprecipitation *in vitro* experiments, I failed to detect direct AtPOT1c interaction with AtPOT1a, AtPOT1b, KU70, KU80, and TERT, the known *Arabidopsis* telomere components (Y. Surovtseva and D. Shippen, unpublished). Therefore, a yeast-two-hybrid screen or biochemical purification of AtPOT1c may be necessary to identify AtPOT1c interacting partners. Another unresolved question is the functional significance of altenative splicing for *AtPOT1c*. In our experiments, a genomic copy of *AtPOT1c* was transformed into plants, allowing over-expression of both AtPOT1c isoforms. It will be interesting to see if the consequences differ when only AtPOT1cS or AtPOT1cL is over-expressed. Finally, localization studies of AtPOT1c will be necessary to better understand the biology of this protein. Does it predominantly co-localize with telomerase? Is there a

telomere-bound portion of AtPOT1c? Does AtPOT1c localization change in the cell cycle? Altogether, these experiments will help understand the role of this novel unique protein at *Arabidopsis* telomeres.

Evolution of plant POT1 proteins

Arabidopsis is one of the few organisms that encode more than one POT1-like protein. Ciliated protozoa are also an exception (Wang *et al*., 1992). In *Euplotes crassus*, there are two G-overhang binding proteins that are highly divergent in sequence and in function. While the telomere protein (TP) tightly binds the 3' terminus of the G-overhang (Price, 1990) and is therefore thought to protect chromosome ends, the replication telomere protein (rTP) is only detected during DNA replication (Skopp *et al*., 1996). Mouse also encodes two POT1 proteins, POT1a and POT1b. In contrast to *Euplotes*, mouse POT1a and POT1b share 72% similarity and appear to be partially redundant for telomere length regulation and chromosome end protection (He *et al*., 2006; Hockemeyer *et al*., 2006; Wu *et al*., 2006).

In *Arabidopsis*, the extraordinary sequence divergence of POT1 proteins implies that these proteins make distinct contributions in telomere biology. This notion is confirmed by genetic data described in this dissertation. It therefore appears that many of the known functions of ss telomeric DNA binding proteins have segregated between multiple POT1 proteins in *Arabidopsis*. Interestingly, most of other plant species appear to possess only one *POT1* gene (E. Shakirov, X. Song and D. Shippen, unpublished). Since the entire *Arabidopsis* genome has undergone several rounds of polyploidization (Blanc and Wolfe, 2004), we hypothesize that *Arabidopsis* acquired multiple POT genes via gene duplication. Over time, accumulation of sequence mutations may result in the

loss of function of one of the duplicated genes, or in the separation of function (subfunctionalization) of each gene (Adams and Wendel, 2005). Alternatively, one of the duplicated genes might acquire a new function (neofunctionalization). It is possible that AtPOT1b and AtPOT1c represent a subfunctionalization, with both proteins contributing to different aspects of chromosome end protection. In contrast, both AtPOT1a and AtPOT1c appear to represent neofunctionalization, with their OB-folds evolved to bind to telomerase RNA. In this regard, it will be interesting to investigate how telomerase access to telomeres is mediated in other plant species that encode a single POT1 gene. Altogether, our data indicate that POT genes in *Arabidopsis* have evolved distinct functions in telomere biology (Figure 42). Thus, *Arabidopsis* is an important model system to study the role and the evolution of this class of telomeric proteins.

Identification and characterization of CIT1

In chapter VI, I described the identification and characterization of the *CIT1* gene. I found *CIT1* through map-based cloning of a mutation which leads to a severe telomere deprotection phenotype. I then demonstrated that *CIT1* is conserved among higher plants and has putative orthologs in vertebrates. The unique combination of telomere functions for the CIT1 protein that has not been ascribed to any known telomere-associated factors argues that CIT1 is a novel and essential telomere cap component.

Even though CIT1 homologs exist in numerous species, none of them encodes a conserved motif or has an annotated function that will provide a clue into the mechanism of CIT1 action. However, secondary structure prediction by the 3D-PSSM program revealed a putative OB-fold in CIT1 protein. To confirm this prediction, we

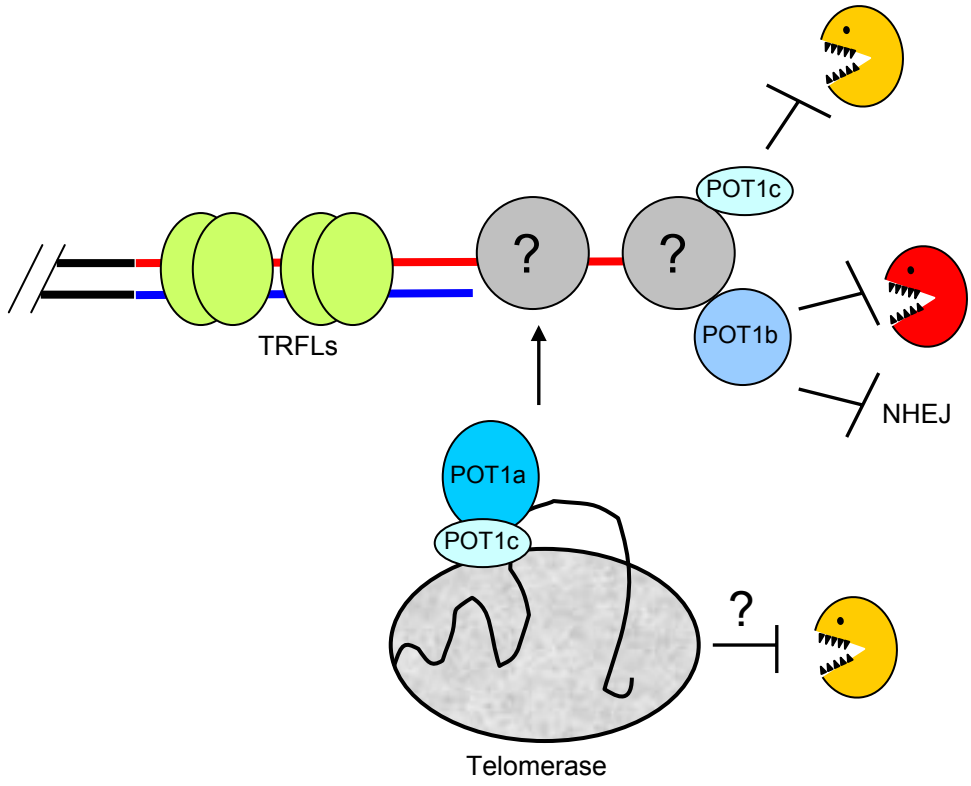


Figure 42. Model for POT1 proteins functions in *Arabidopsis*. None of the *Arabidopsis* POT1 proteins bind telomeric DNA directly. AtPOT1a is a telomerase accessory factor necessary for telomere maintenance *in vivo* and *in vitro* (Chapter IV). AtPOT1c acts in collaboration with telomerase for telomere length maintenance and in regulating telomerase activity (Chapter V). AtPOT1c is also important for G-overhang maintenance and for the control of the C-strand resection by nucleases, possibly through an interaction with telomere-binding protein (Chapter V). Both AtPOT1a and AtPOT1c interact with *Arabidopsis* TER (C. Cifuentes-Rojas and D. Shippen, unpublished). AtPOT1b is implicated in chromosome end protection from end-to-end fusions and nuclease attack (Chapter III). AtPOT1b role in chromosome end protection is likely redundant with AtPOT1c or some other telomere proteins (Chapter VII).

have begun a collaboration with Dr. Douglas Theobald, a computational biologist with an expertise in the identification of OB-fold motifs.

The presence of an OB fold in CIT1 is consistent with the possibility that CIT1 binds the G-overhang. *In vitro* binding experiments will be important to test whether CIT1 directly binds ss telomeric DNA. However, the large size of CIT1 (~142 kDa) makes its expression in *E. coli* or rabbit reticulocyte lysate technically challenging. Therefore, mapping of the CIT1 OB-fold will allow us to examine this domain for its putative interaction with telomeric DNA using *in vitro* gel shift assays with ss telomeric DNA oligo or with the oligonucleotide mimicking the junction between the ss and ds telomeric DNA. In addition, several mutants have been identified in our lab that give rise to either short or long telomeric G-overhangs (Riha and Shippen, 2003a). Localization studies of CIT1 in these mutant backgrounds will be necessary to test the possibility that CIT1 binds telomeres in a G-overhang dependent manner.

It is unclear why CIT1 protein has never been found in thorough extensive biochemical purifications of vertebrate telomere complexes (Liu *et al*., 2004; O'Connor *et al*., 2004; Ye *et al*., 2004a). One possibility is that sub-stoichiometric amounts of CIT1 precluded its identification. As discussed in Chapter VI, it is also possible that CIT1 does not localize to telomeres all the time, but rather dynamically associates telomeres in a cell-cycle dependent manner. Both of these models can be tested in plants by chromatin immunoprecipitation and immunolocalization studies of CIT1 protein. We have already transformed *cit1-1* heterozygous mutants with a construct that encodes *CIT1* cDNA fused to a Strep-tag. The same construct was transformed into *Arabidopsis* cell culture, which can be synchronized. Immunolocalization of CIT1 in

these transformants combined with telomeric FISH in synchronized cells may provide new insight into the role of AtCIT1 in telomere biology.

Model for AtCIT1 role at telomeres

While these important localization experiments are underway, our current data allow us to propose a model for CIT1 function at telomere. I favor the hypothesis that CIT1 is essential for telomere protection in late S/early G2 phases of cell cycle (Figure 43). At this time, the t-loop is unfolded and telomeres assume an open and therefore less protected state to allow replication by telomerase and by conventional replication machinery. It has recently been established that DNA damage repair proteins are recruited to the telomeres in lateS/early G2 (Verdun *et al*., 2005; Verdun and Karlseder, 2006). The current model suggests that DNA damage machinery is required for exonucleolytic processing of the C strand after telomere replication (Chai *et al*., 2006; Larrivee *et al*., 2004; Verdun and Karlseder, 2006). C-strand processing is, in turn, obligatory for the formation of the t-loop and thus the formation of the protective telomere state (Dionne and Wellinger, 1996; Jacob *et al*., 2003; Jacob *et al*., 2001; Wellinger *et al*., 1992; Wellinger *et al*., 1993). Although resection of the C strand by the DNA damage repair proteins is essential for proper telomere function, it has to be controlled. Thus, I propose that CIT1 controls the extent of C strand processing. Consistent with this role, *cit1* mutants display a gross increase in G-overhang signal, up to 10 times that of wild type. Currently, I cannot rule out the possibility that the increased G-overhang signal upon CIT1 loss is dependent on telomerase. However, since a role in C-strand protection from degradation has been ascribed to Cdc13p, Stn1p and Ten1p, the three critical OB-fold containing ss telomere proteins from

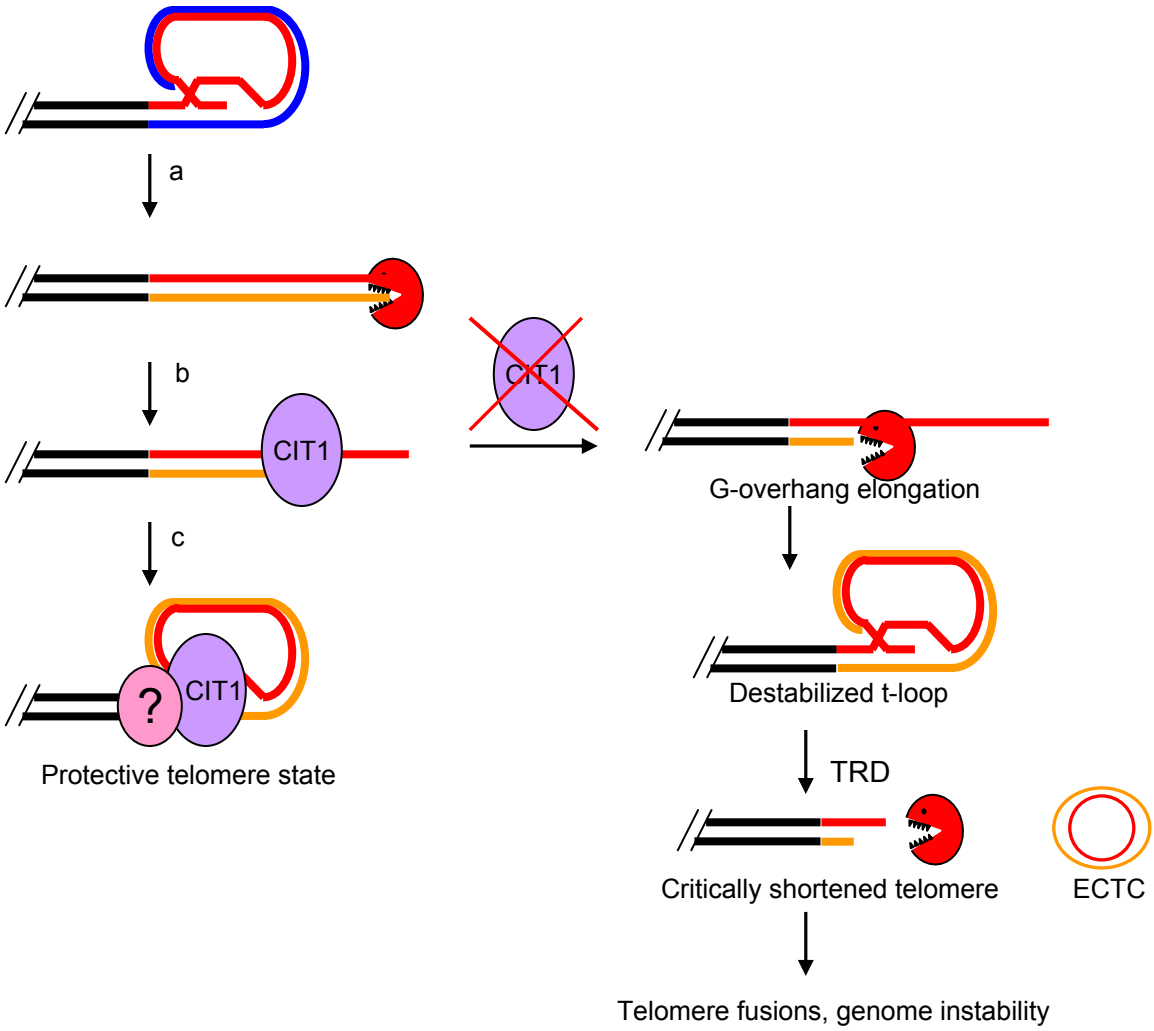


Figure 43. Model for CIT1 function. In S phase, the t-loops are unfolded to allow passage of replication fork and telomere replication by telomerase (a). Following replication, C-strand processing by nucleases leads to a formation of the G-overhang. CIT1 may control C-strand processing, preventing excessive nucleolytic degradation (b). CIT1 could also stabilize the newly formed t-loop through an unknown mechanism (c). In the absence of CIT1, uncontrolled C-strand degradation leads to an increase in the G-overhang length. Although t-loops are formed, they are unstable, resulting in t-loop cleavage through TRD. This in turn leads to accumulation of extrachromosomal telomeric circles (ECTCs) and formation of critically shortened telomeres. Deprotected telomeres are then recruited into end-to-end fusions, and genome instability is initiated.

budding yeast, I speculate that CIT1 has a similar C-strand protection function (Garvik *et al*., 1995; Grandin *et al*., 2001; Grandin *et al*., 1997; Nugent *et al*., 1996).

Since at least a subset of telomeres in *cit1* mutants possesses G-overhangs, the t-loop is apparently formed in the absence of CIT1. However, identification of extrachromosomal telomeric circles (ECTC) in the *cit1* background argues that the t-loops are unstable and telomeres undergo recombination-based cleavage of the t-loop via TRD. It is possible that TRD is increased in *cit1* mutants due to long recombinogenic G-overhangs. Alternatively, CIT1 might stabilize the t-loop via the recruitment of other essential telomere capping components. A protein such as TRF2 is a putative candidate for this role, since it binds at the junction of ss and ds telomeric DNA and plays role in the formation and stabilization of the t-loop *in vitro* (Fouche *et al*., 2006; Stansel *et al*., 2001).

Finally, I speculate that the elevated TRD in *cit1* mutants is responsible for the critically short telomeres in this background, which cannot form the t-loop and therefore cannot assume a protected state. Nucleolytic processing of such deprotected chromosome ends will lead to the loss of telomeric and subtelomeric sequences, recruitment of dysfunctional telomeres into end-to-end fusions, formation of the dicentric chromosomes, and devastating genome instability characteristic of the CIT1 deficiency (Figure 43).

A protective role for telomerase in cit1 *mutants*

One unexpected finding uncovered in my study is the embryonic lethality of double *tert cit1* mutants. Two models are proposed to explain the synthetic lethality of these mutations. One possibility is that telomerase is able to rescue at least a subset of

critically shortened telomeres in *cit1* mutants by extending them. Consistent with this model, it has been shown in plants and other systems that telomerase is preferentially recruited to the shortest telomeres in the cell (Hemann *et al*., 2001; Shakirov and Shippen, 2004; Teixeira *et al*., 2004). Analysis of CIT1 loss in the *ku70* mutant background may be required to test this hypothesis. KU is a negative regulator of telomerase in *Arabidopsis*; *ku70* mutants display extensive telomerase-dependent telomere elongation (Riha and Shippen, 2003a). It is therefore possible that KU inhibition will rescue the *cit1* phenotype due to an enhanced ability of telomerase to extend critically shortened telomeres.

Alternatively, telomerase may be more directly involved in telomere capping. A capping function for TERT has been proposed in both yeast and mammals based on the fact that the effect of inactivating telomerase on cell viability and can be uncoupled from the effect on telomere length maintenance reviewed in (Blackburn, 2005). While some TERT alleles that cannot prevent telomere shortening extend cellular lifespan (Zhu *et al*., 1999), other alleles cause cell death without changes in bulk telomere length (Kim *et al*., 2003). Therefore, the TERT subunit of telomerase may be involved in chromosome capping. This may also be true in *Arabidopsis*. If so, TERT may function independently from CIT1, and the dysfunction of both of these protective pathways may not be tolerated by the cell. It should be possible to identify a TERT allele that results in inactivation of telomerase, but which does not abolish TERT protein accumulation. Combining this mutant with CIT1 deficiency may indicate whether TERT or the telomerase can protect chromosome ends. The latter analysis will also be required to test whether the deregulation of telomere length and the lengthening of the G-overhang upon CIT1 loss are dependent on telomerase activity.

An unusual morphological phenotype in cit1 *mutants*

As described in the Chapter VI, *cit1* mutant plants display an unusual morphological phenotype termed the fasciated phenotype, which was first described in *clavala* mutants (Clark *et al*., 1993). The fasciated phenotype appears to reflect defects in a complex signaling network required for proper meristem function. Meristems are plant tissues containing stem cells (reviewed in Carles and Fletcher, 2003). While stem cell identity is maintained by multiple loci, CLAVATA promotes cell differentiation and incorporation into organ primordial (Clark, 2001). Loss of function *clavata* mutants display a fasciated phenotype similar to morphological defects in *cit1* mutants (Figure 44) (Clark *et al*., 1993; Kayes and Clark, 1998). It is therefore possible that telomere defects can result in meristem defects, or that meristem defects can give rise to telomere defects. It is, however, unclear which comes first and what the primary cause of morphological defects in such mutants is. To test whether defects in meristem maintenance cause telomere deprotection in plants, TRF analysis and telomere fusion PCR was performed on different *clavata* mutants. In all cases, telomeres fell into the wild type length range, and no evidence for end-to-end fusions was observed in *clavala* mutants (Figure 44). Thus, stem cell dysfunction does not indirectly affect telomere homeostasis.

By contrast, telomere uncapping does give rise to defects in the meristem in *Arabidopsis* (Riha *et al*., 2001). Late generation telomerase-deficient mutants display complex changes in shoot apical meristem (Riha *et al*., 2001). Two models were proposed. In the first model, genome instability associated with telomere dysfunction directly affects genes involved in meristem growth and stem cell maintenance (Laux and Mayer, 1998; Meyerowitz, 1997). Alternatively, telomere dysfunction could lead to

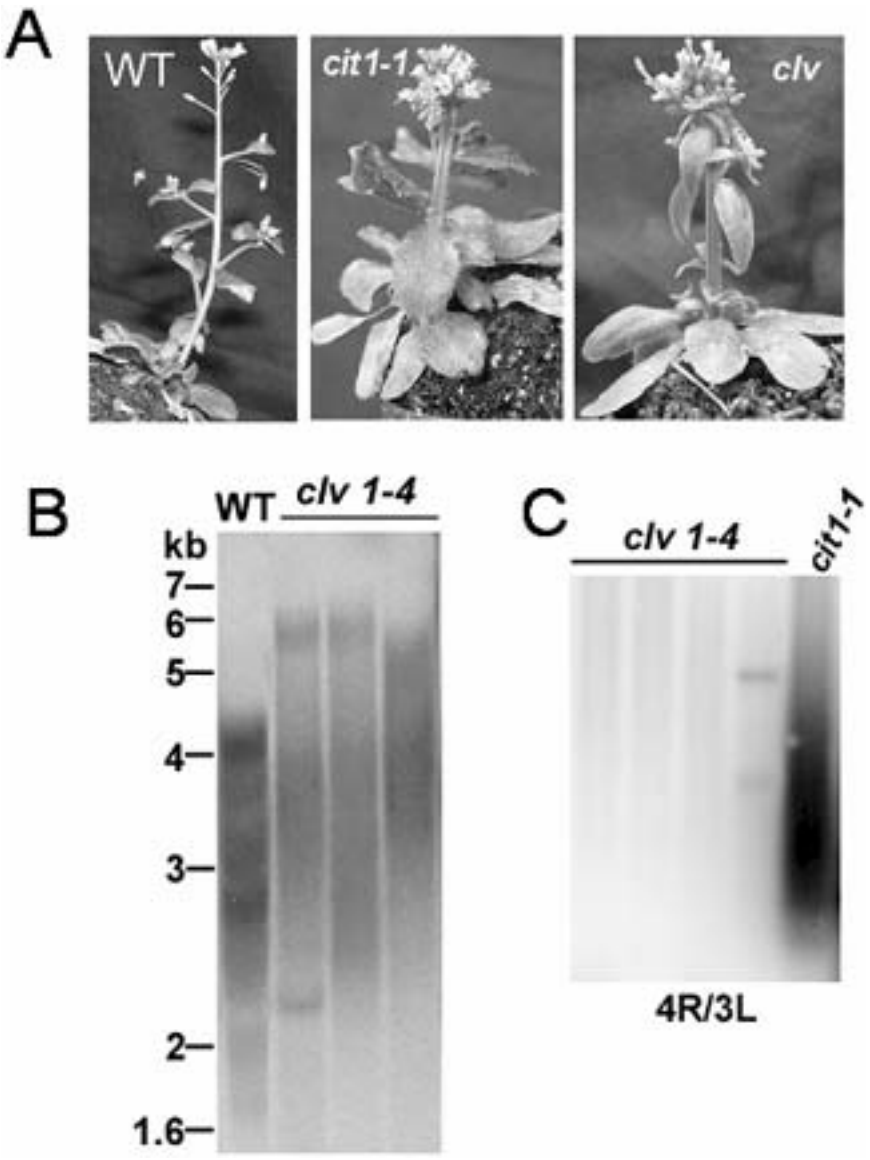


Figure 44. Stem cell dysfunction does not affect telomere length homeostasis. (A) Morphological defects in *cit1-1* and *clavata* (*clv*) plants. Both mutants display fasciation of the stem and fused floral organs. (B) TRF analysis of clavata mutants. (C) Telomere fusion PCR analysis of *clavata* mutants. Primers specific for 4R and 3L subtelomeric sequences were used. *cit1-1* DNA was used as a positive control. Blots shown in panels B and C were hybridized with a radiolabeled telomeric probe. The faint bands in panel C represent non-specific amplification of centromere regions and are sometimes observed in wild type DNA samples.

loss of intracellular communications and meristem integrity (Fletcher and Meyerowitz, 2000). Whatever the mechanism, our analysis of *cit1* mutants confirm that telomere uncapping and accompanying genome instability cause defects in the stem cell population of *Arabidopsis*.

Chromosome end protection in Arabidopsis

Although the TRF2 homolog has not been yet found in plants, several proteins that contribute to telomere capping were found in *Arabidopsis*. As discussed above, transgenic studies implicated AtPOT1b and AtPOT1c in chromosome end protection. CIT1 is another protein essential for proper telomere function. In addition, a sequence homolog of STN1 protein has recently been found in *Arabidopsis*. Like its yeast homolog, *Arabidopsis* STN1 is essential for chromosome end protection: null mutants display severe telomere deprotection phenotype that is remarkably similar to that observed upon CIT1 loss (X. Song and D. Shippen, unpublished). Notably, all of the proteins implicated in protection of *Arabidopsis* telomeres encode an OB-fold. Several questions remain unresolved. Do these proteins interact with each other and how do they contribute to chromosome end protection? Do they associate with telomeres dynamically? Do they bind G-overhang directly, or do they use their OB-fold for protein-protein interactions? Do they function in a single pathway to provide chromosome end protection? Biochemical analysis of interactions between these proteins, genetic analysis of double mutants, and molecular analysis of the cell cycle localization will be required to answer at least some of these questions.

Conclusions

Using a combination of genetic, biochemical and molecular tools, my work has resulted in the identification and characterization of essential *Arabidopsis* telomere-associated proteins. These studies have not only improved our understanding of plant telomere composition, but also provided new insights into the overall structure and evolution of the telomeric complex.

REFERENCES

Adams KL, Wendel JF (2005) Polyploidy and genome evolution in plants. *Curr Opin Plant Biol* **8**: 135-141

Armbruster BN, Linardic CM, Veldman T, Bansal NP, Downie DL, Counter CM (2004) Rescue of an hTERT mutant defective in telomere elongation by fusion with hPot1. *Mol Cell Biol* **24**: 3552-3561

Armstrong SJ, Franklin FC, Jones GH (2001) Nucleolus-associated telomere clustering and pairing precede meiotic chromosome synapsis in *Arabidopsis thaliana*. *J Cell Sci* **114**: 4207-4217

Autexier C, Greider CW (1994) Functional reconstitution of wild-type and mutant *Tetrahymena* telomerase. *Genes Dev* **8**: 563-575

Autexier C, Lue NF (2006) The structure and function of telomerase reverse transcriptase. *Annu Rev Biochem* **75**: 493-517

Baumann P. (2006) Are mouse telomeres going to pot? *Cell* **126**: 33-36

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

Baumann P, Podell E, Cech TR (2002) Human Pot1 (protection of telomeres) protein: cytolocalization, gene structure, and alternative splicing. *Molecular & Cellular Biology* **22**: 8079-8087

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

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

Biessmann H, Valgeirsdottir K, Lofsky A, Chin C, Ginther B, Levis RW, Pardue ML (1992) HeT-A, a transposable element specifically involved in "healing" broken chromosome ends in *Drosophila melanogaster*. *Mol Cell Biol* **12**: 3910-3918

Bilaud T, Brun C, Ancelin K, Koering CE, Laroche T, Gilson E (1997) Telomeric localization of TRF2, a novel human telobox protein. *Nature Genetics* **17**: 236-239

Bilaud T, Koering CE, Binet-Brasselet E, Ancelin K, Pollice A, Gasser SM, 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 EH (1991) Structure and function of telomeres. *Nature* **350**: 569-573

Blackburn EH (2001) Switching and signaling at the telomere. *Cell* **106**: 661-673

Blackburn EH (2005) Telomeres and telomerase: their mechanisms of action and the effects of altering their functions. *FEBS Lett* **579**: 859-862

Blackburn EH, Gall JG (1978) A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in *Tetrahymena*. *J Mol Biol* **120**: 33-53

Blanc G, Wolfe KH (2004) Functional divergence of duplicated genes formed by polyploidy during *Arabidopsis* evolution. *Plant Cell* **16**: 1679-1691

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

Boulton SJ, Jackson SP (1996) Identification of a *Saccharomyces cerevisiae* Ku80 homologue: roles in DNA double strand break rejoining and in telomeric maintenance. *Nucleic Acids Res* **24**: 4639-4648

Boulton SJ, Jackson SP (1998) Components of the Ku-dependent non-homologous end-joining pathway are involved in telomeric length maintenance and telomeric silencing. *EMBO Journal* **17**: 1819-1828

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

Bryan TM, Englezou A, Dalla-Pozza L., Dunham MA, Reddel RR (1997) Evidence for an alternative mechanism for maintaining telomere length in human tumors and tumor-derived cell lines. *Nat Med* **3**: 1271-1274

Bryan TM, Goodrich KJ, Cech TR (2000) A mutant of *Tetrahymena* telomerase reverse transcriptase with increased processivity. *J Biol Chem* **275**: 24199-24207

Bucholc M, Park Y, Lustig AJ (2001) Intrachromatid excision of telomeric DNA as a mechanism for telomere size control in *Saccharomyces cerevisiae*. *Mol Cell Biol* **21**: 6559-6573

Bunch JT, Bae NS, Leonardi J, Baumann P (2005) Distinct requirements for Pot1 in limiting telomere length and maintaining chromosome stability. *Mol Cell Biol* **25**: 5567-5578

Bundock P, Hooykaas P (2002) Severe developmental defects, hypersensitivity to DNA-damaging agents, and lengthened telomeres in *Arabidopsis* MRE11 mutants. *Plant Cell* **14**: 2451-2462

Burr B, Burr FA, Matz EC, Romero-Severson J (1992) Pinning down loose ends: mapping telomeres and factors affecting their length. *Plant Cell* **4**: 953-960

Cao X, Jacobsen SE (2002) Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. *Proc Natl Acad Sci U S A* **99**: 16491-16498

Carles CC, Fletcher JC (2003) Shoot apical meristem maintenance: the art of a dynamic balance. *Trends Plant Sci* **8**: 394-401

Cech TR (2004) Beginning to understand the end of the chromosome. *Cell* **116**: 273-279

Celli GB, de Lange T (2005) DNA processing is not required for ATM-mediated telomere damage response after TRF2 deletion. *Nat Cell Biol* **7**: 712-718

Cesare AJ, Griffith JD (2004) Telomeric DNA in ALT cells is characterized by free telomeric circles and heterogeneous t-loops. *Mol Cell Biol* **24**: 9948-9957

Cesare AJ, Quinney N, Willcox S, Subramanian D, Griffith JD (2003) Telomere looping in *P. sativum* (common garden pea). *Plant J* **36**: 271-279

Chai W, Sfeir AJ, Hoshiyama H, Shay JW, Wright WE (2006) The involvement of the Mre11/Rad50/Nbs1 complex in the generation of G-overhangs at human telomeres. *EMBO Rep* **7**: 225-230

Chan SW, Blackburn EH (2002) New ways not to make ends meet: telomerase, DNA damage proteins and heterochromatin. *Oncogene* **21**: 553-563

Chan SW, Chang J, Prescott J, Blackburn EH (2001) Altering telomere structure allows telomerase to act in yeast lacking ATM kinases. *Curr Biol* **11**: 1240-1250

Chandra A, Hughes TR, Nugent CI, Lundblad V (2001) Cdc13 both positively and negatively regulates telomere replication. *Genes Dev* **15**: 404-414

Chen CM, Wang CT, Ho CH (2001) A plant gene encoding a Myb-like protein that binds telomeric GGTTAG repeats *in vitro*. *J Biol Chem* **276**: 16511-16519

Chen JL, Greider CW (2004) Telomerase RNA structure and function: implications for dyskeratosis congenita. *Trends Biochem Sci* **29**: 183-192

Chong L, van Steensel B, Broccoli D, Erdjument-Bromage H, Hanish J, Tempst P, Lange T (1995) A human telomeric protein. *Science* **270**: 1663-1667

Churikov D, Price CM (2008) Pot1 and cell cycle progression cooperate in telomere length regulation. *Nat Struct Mol Biol* **15**: 79-84

Churikov D, Wei C, Price CM (2006) Vertebrate POT1 restricts G-overhang length and prevents activation of a telomeric DNA damage checkpoint but is dispensable for overhang protection. *Mol Cell Biol* **26**: 6971-6982

Clark SE (2001) Cell signalling at the shoot meristem. *Nat Rev Mol Cell Biol* **2**: 276-284

Clark SE, Running MP, Meyerowitz EM (1993) CLAVATA1, a regulator of meristem and flower development in *Arabidopsis*. *Development* **119**: 397-418

Cocciolone SM, Cone KC (1993) Pl-Bh, an anthocyanin regulatory gene of maize that leads to variegated pigmentation. *Genetics* **135**: 575-588

Colgin LM, Baran K, Baumann P, Cech TR, Reddel RR (2003) Human POT1 facilitates telomere elongation by telomerase. *Curr Biol* **13**: 942-946

Collins K (2006) The biogenesis and regulation of telomerase holoenzymes. *Nat Rev Mol Cell Biol* **7**: 484-494

Conrad MN, Wright JH, Wolf AJ, Zakian VA (1990) RAP1 protein interacts with yeast telomeres *in vivo*: overproduction alters telomere structure and decreases chromosome stability. *Cell* **63**: 739-750

Cooper JP, Nimmo ER, Allshire RC, Cech TR (1997) Regulation of telomere length and function by a Myb-domain protein in fission yeast. *Nature* **385**: 744-747

Counter CM, Avilion AA, LeFeuvre CE, Stewart NG, Greider CW, Harley CB, Bacchetti S (1992) Telomere shortening associated with chromosome instability is

arrested in immortal cells which express telomerase activity. *EMBO J* **11**: 1921-1929

Culligan K, Tissier A, Britt A (2004) ATR regulates a G2-phase cell-cycle checkpoint in *Arabidopsis thaliana*. *Plant Cell* **16**: 1091-1104

D'Amours D, Jackson SP (2002) The Mre11 complex: at the crossroads of DNA repair and checkpoint signalling. *Nat Rev Mol Cell Biol* **3**: 317-327

da Costa e Silva O, Klein L, Schmelzer E, Trezzini GF, Hahlbrock K (1993) BPF-1, a pathogen-induced DNA-binding protein involved in the plant defense response. *Plant J* **4**: 125-135

de Bruin D, Zaman Z, Liberatore RA, Ptashne M (2001) Telomere looping permits gene activation by a downstream UAS in yeast. *Nature* **409**: 109-113

de Lange T (2002) Protection of mammalian telomeres. *Oncogene* **21**: 532-540

de Lange T (2004) T-loops and the origin of telomeres. *Nat Rev Mol Cell Biol* **5**: 323-329

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

De Rycker M, Venkatesan RN, Wei C, Price CM (2003) Vertebrate tankyrase domain structure and sterile alpha motif (SAM)-mediated multimerization. *Biochem J* **372**: 87-96

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

Dionne I, Wellinger RJ (1996) Cell cycle-regulated generation of single-stranded G-rich DNA in the absence of telomerase. *Proc Natl Acad Sci U S A* **93**: 13902-13907

Dragon F, Pogacic V, Filipowicz W (2000) *In vitro* assembly of human H/ACA small nucleolar RNPs reveals unique features of U17 and telomerase RNAs. *Mol Cell Biol* **20**: 3037-3048

Etheridge KT, Banik SS, Armbruster BN, Zhu Y, Terns RM, Terns MP, Counter CM (2002) The nucleolar localization domain of the catalytic subunit of human telomerase. *J Biol Chem* **277**: 24764-24770

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

Evans SK, Lundblad V (2000) Positive and negative regulation of telomerase access to the telomere. *J Cell Sci* **113**: 3357-3364

Evans SK, Lundblad V (2002) The Est1 subunit of *Saccharomyces cerevisiae* telomerase makes multiple contributions to telomere length maintenance. *Genetics* **162**: 1101-1115

Fairall L, Chapman L, Moss H, de Lange T, Rhodes D (2001) Structure of the TRFH dimerization domain of the human telomeric proteins TRF1 and TRF2. *Molecular Cell* **8**: 351-361

Fajkus J, Kovarik A, Kralovics R, Bezdek M (1995) Organization of telomeric and subtelomeric chromatin in the higher plant *Nicotiana tabacum*. *Mol Gen Genet* **247**: 633-638

Fang G, Cech TR (1993) The beta subunit of *Oxytricha* telomere-binding protein promotes G-quartet formation by telomeric DNA. *Cell* **74**: 875-885

Feng J, Funk WD, Wang SS, Weinrich SL, Avilion AA, Chiu CP, Adams RR, Chang E, Allsopp RC, Yu J, *et al*. (1995) The RNA component of human telomerase. *Science* **269**: 1236-1241

Ferreira, MG, Cooper JP (2001) The fission yeast Taz1 protein protects chromosomes from Ku-dependent end-to-end fusions. *Mol Cell* **7**: 55-63

Ferreira MG, Miller KM, Cooper JP (2004) Indecent exposure: when telomeres become uncapped. *Mol Cell* **13**: 7-18

Fisher TS, Taggart AK, Zakian VA (2004) Cell cycle-dependent regulation of yeast telomerase by Ku. *Nat Struct Mol Biol* **11**: 1198-1205

Fisher TS, Zakian VA (2005) Ku: a multifunctional protein involved in telomere maintenance. *DNA Repair (Amst)* **4**:1215-1226

Fitzgerald MS, McKnight TD, Shippen DE (1996) Characterization and developmental patterns of telomerase expression in plants. *Proc Natl Acad Sci U S A* **93**: 14422-14427

Fitzgerald MS, Riha K, Gao F, Ren S, McKnight TD, Shippen DE (1999) Disruption of the telomerase catalytic subunit gene from *Arabidopsis* inactivates telomerase and leads to a slow loss of telomeric DNA. *Proc Natl Acad Sci U S A* **96**: 14813-14818

Fletcher JC, Meyerowitz EM (2000) Cell signaling within the shoot meristem. *Curr Opin Plant Biol* **3**: 23-30

Forsythe HL, Jarvis JL, Turner JW, Elmore LW, Holt SE (2001) Stable association of hsp90 and p23, but Not hsp70, with active human telomerase. *J Biol Chem* **276**: 15571-15574

Fouche N, Cesare AJ, Willcox S, Ozgur S, Compton SA, Griffith JD (2006) The basic domain of TRF2 directs binding to DNA junctions irrespective of the presence of TTAGGG repeats. *J Biol Chem* **281**: 37486-37495

Fransz P, De Jong JH, Lysak M, Castiglione MR, Schubert I (2002) Interphase chromosomes in *Arabidopsis* are organized as well defined chromocenters from which euchromatin loops emanate. *Proc Natl Acad Sci U S A* **99**: 14584-14589

Gallego ME, Jalut N, White CI (2003) Telomerase dependence of telomere lengthening in Ku80 mutant *Arabidopsis*. *Plant Cell* **15**: 782-789

Gallego ME, Jeanneau M, Granier F, Bouchez D, Bechtold N, White CI (2001) Disruption of the *Arabidopsis* RAD50 gene leads to plant sterility and MMS sensitivity. *Plant J* **25**: 31-41

Gao H, Cervantes RB, Mandell EK, Otero JH, Lundblad V (2007) RPA-like proteins mediate yeast telomere function. *Nat Struct Mol Biol* **14**: 208-214

Garcia-Cao M, O'Sullivan R, Peters AH, Jenuwein T, Blasco MA (2004) Epigenetic regulation of telomere length in mammalian cells by the Suv39h1 and Suv39h2 histone methyltransferases. *Nat Genet* **36**: 94-99

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

Gonzalo S, Jaco I, Fraga MF, Chen T, Li E, Esteller M, Blasco MA (2006) DNA methyltransferases control telomere length and telomere recombination in mammalian cells. *Nat Cell Biol* **8**: 416-424

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

Grandin N, Damon C, Charbonneau M (2000) Cdc13 cooperates with the yeast Ku proteins and Stn1 to regulate telomerase recruitment. *Mol Cell Biol* **20**: 8397-8408

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

Grandin N, Reed SI, Charbonneau M (1997) Stn1, a new *Saccharomyces cerevisiae* protein, is implicated in telomere size regulation in association with Cdc13. *Genes Dev* **11**: 512-527

Gravel S, Larrivee M, Labrecque P, Wellinger RJ (1998) Yeast Ku as a regulator of chromosomal DNA end structure. *Science* **280**: 741-744

Gray JT, Celander DW, Price CM, Cech TR (1991) Cloning and expression of genes for the *Oxytricha* telomere-binding protein: specific subunit interactions in the telomeric complex. *Cell* **67**: 807-814

Greider CW, Blackburn EH (1985) Identification of a specific telomere terminal transferase activity in *Tetrahymena* extracts. *Cell* **43**: 405-413

Greider CW, Blackburn EH (1987) The telomere terminal transferase of *Tetrahymena* is a ribonucleoprotein enzyme with two kinds of primer specificity. *Cell* **51**: 887-898

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

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

Grunstein M (1997) Molecular model for telomeric heterochromatin in yeast. *Curr Opin Cell Biol* **9**: 383-387

Guo X, Deng Y, Lin Y, Cosme-Blanco W, Chan S, He H, Yuan G, Brown EJ, Chang S (2007) Dysfunctional telomeres activate an ATM-ATR-dependent DNA damage response to suppress tumorigenesis. *EMBO J* **26**: 4709-4719

Hall AR, Milner J (1997) Specific p53-DNA complexes contain an mdm2-related protein. *Oncogene* **14**: 1371-1376

Hande MP, Balajee AS, Tchirkov A, Wynshaw-Boris A, Lansdorp PM (2001) Extra-chromosomal telomeric DNA in cells from Atm(-/-) mice and patients with ataxia-telangiectasia. *Hum Mol Genet* **10**: 519-528

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

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

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

Hayflick L, Moorhead PS (1961) The serial cultivation of human diploid cell strains. *Exp Cell Res* **25**: 585-621

He H, Multani AS, Cosme-Blanco W, Tahara H, Ma J, Pathak S, Deng Y, Chang S (2006) POT1b protects telomeres from end-to-end chromosomal fusions and aberrant homologous recombination. *EMBO J* **25**: 5180-5190

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

Heacock ML, Idol RA, Friesner JD, Britt AB, Shippen DE (2007) Telomere dynamics and fusion of critically shortened telomeres in plants lacking DNA ligase IV. *Nucleic Acids Res* **35**: 6490-6500

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

Henson JD, Neumann AA, Yeager TR, Reddel RR (2002) Alternative lengthening of telomeres in mammalian cells. *Oncogene* **21**: 598-610

Hicke BJ, Celander DW, MacDonald GH, Price CM, Cech TR (1990) Two versions of the gene encoding the 41-kilodalton subunit of the telomere binding protein of *Oxytricha nova. Proc Natl Acad Sci U S A* **87**: 1481-1485

Higashiyama T, Maki S, Yamada T (1995) Molecular organization of *Chlorella vulgaris* chromosome I: presence of telomeric repeats that are conserved in higher plants. *Mol Gen Genet* **246**: 29-36

Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77**: 51-59

Hockemeyer D, Daniels JP, Takai H, de Lange T (2006) Recent expansion of the telomeric complex in rodents: two distinct POT1 proteins protect mouse telomeres. *Cell* **126**: 63-77

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

Holt SE, Aisner DL, Baur J, Tesmer VM, Dy M, Ouellette M, Trager JB, Morin GB, Toft DO, Shay JW, Wright WE, White MA (1999) Functional requirement of p23 and Hsp90 in telomerase complexes. *Genes Dev* **13**: 817-826

Hong JP, Byun MY, Koo DH, An K, Bang JW, Chung IK, An G, Kim WT (2007) Suppression of RICE TELOMERE BINDING PROTEIN 1 results in severe and gradual developmental defects accompanied by genome instability in rice. *Plant Cell* **19**: 1770-1781

Horvath MP, Schweiker VL, Bevilacqua JM, Ruggles JA, Schultz SC (1998) Crystal structure of the *Oxytricha nova* telomere end binding protein complexed with single strand DNA. *Cell* **95**: 963-974

Houghtaling BR, Cuttonaro L, Chang W, Smith S (2004) A dynamic molecular link between the telomere length regulator TRF1 and the chromosome end protector TRF2. *Curr Biol* **14**: 1621-1631

Hsu M, McEachern MJ, Dandjinou AT, Tzfati Y, Orr E, Blackburn EH, Lue NF (2007) Telomerase core components protect *Candida* telomeres from aberrant overhang accumulation. *Proc Natl Acad Sci U S A* **104**: 11682-11687

Huet J, Cottrelle P, Cool M, Vignais ML, Thiele D, Marck C, Buhler JM, Sentenac A, Fromageot P (1985) A general upstream binding factor for genes of the yeast translational apparatus. *EMBO J* **4**: 3539-3547

Hwang MG, Cho MH (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 Res* **35**: 1333-1342

Hwang MG, Chung IK, Kang BG, Cho MH (2001) Sequence-specific binding property of *Arabidopsis thaliana* telomeric DNA binding protein 1 (AtTBP1). *FEBS Letters* **503**: 35-40

Iyer S, Chadha AD, McEachern MJ (2005) A mutation in the STN1 gene triggers an alternative lengthening of telomere-like runaway recombinational telomere elongation and rapid deletion in yeast. *Mol Cell Biol* **25**: 8064-8073

Jacob NK., Kirk KE, Price CM (2003) Generation of telomeric G strand overhangs involves both G and C strand cleavage. *Mol Cell* **11**: 1021-1032

Jacob NK, Lescasse R, Linger BR, Price CM (2006) *Tetrahymena* POT1a regulates telomere length and prevents activation of a cell cycle checkpoint. *Mol Cell Biol* **11**: 1021-1032

Jacob NK, Skopp R, Price CM (2001) G-overhang dynamics at *Tetrahymena* telomeres. *EMBO J* **20**: 4299-4308

Kannan K, Nelson AD, Shippen DE (2008) Dyskerin is a component of the *Arabidopsis* telomerase RNP required for telomere maintenance. *Mol Cell Biol* **28**: 2332-2341

Kanoh J, Ishikawa F (2001) spRap1 and spRif1, recruited to telomeres by Taz1, are essential for telomere function in fission yeast. *Curr Biol* **11**: 1624-1630

Karamysheva ZN, Surovtseva YV, Vespa L, Shakirov EV, Shippen DE. (2004) A C-terminal Myb extension domain defines a novel family of double-strand telomeric DNA-binding proteins in *Arabidopsis*. *J Biol Chem* **279**: 47799-47807

Karlseder J, Broccoli D, Dai Y, Hardy S, de Lange T (1999) p53- and ATM-dependent apoptosis induced by telomeres lacking TRF2. *Science* **283**: 1321-1325

Karlseder J, Hoke K, Mirzoeva OK, Bakkenist C, Kastan MB, Petrini JH, de Lange T (2004) The telomeric protein TRF2 binds the ATM kinase and can inhibit the ATM-dependent DNA damage response. *PLoS Biol* **2**: E240

Kaya H, Shibahara KI, Taoka KI, Iwabuchi M, Stillman B, Araki T (2001) FASCIATA genes for chromatin assembly factor-1 in *Arabidopsis* maintain the cellular organization of apical meristems. *Cell* **104**: 131-142

Kayes JM, Clark SE (1998) CLAVATA2, a regulator of meristem and organ development in *Arabidopsis*. *Development* **125**: 3843-3851

Keegan CE, Hutz JE, Else T, Adamska M, Shah SP, Kent AE, Howes JM, Beamer WG, Hammer GD (2005) Urogenital and caudal dysgenesis in adrenocortical dysplasia (acd) mice is caused by a splicing mutation in a novel telomeric regulator. *Hum Mol Genet* **14**: 113-123

Kelleher C, Kurth I, Lingner J (2005) Human Protection of Telomeres 1 (POT1) is a negative regulator of telomerase activity *in vitro*. *Mol Cell Biol* **25**: 808-818

Kelley LA, MacCallum RM, Sternberg MJ (2000) Enhanced genome annotation using structural profiles in the program 3D-PSSM. *J Mol Biol* **299**: 499-520

Kilian A, Stiff C, Kleinhofs A (1995) Barley telomeres shorten during differentiation but grow in callus culture. *Proc Natl Acad Sci U S A* **92**: 9555-9559

Kim M, Xu, L, Blackburn EH (2003) Catalytically active human telomerase mutants with allele-specific biological properties. *Exp Cell Res* **288**: 277-287

Kim SH, Beausejour C, Davalos AR, Kaminker P, Heo SJ, Campisi J (2004) TIN2 mediates functions of TRF2 at human telomeres. *J Biol Chem* **279**: 43799-43804

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

Klobutcher LA, Swanton MT, Donini P, Prescott DM (1981) All gene-sized DNA molecules in four species of hypotrichs have the same terminal sequence and an unusual 3' terminus. *Proc Natl Acad Sci U S A* **78**: 3015-3019

Kobryn K, Chaconas G (2001) The circle is broken: telomere resolution in linear replicons. *Curr Opin Microbiol* **4**: 558-564

Konig P, Fairall L, Rhodes D (1998) Sequence-specific DNA recognition by the myb-like domain of the human telomere binding protein TRF1: a model for the protein-DNA complex. *Nucleic Acids Res* **26**: 1731-1740

Konig P, Giraldo R, Chapman L, Rhodes D (1996) The crystal structure of the DNA-binding domain of yeast RAP1 in complex with telomeric DNA. *Cell*, **85**: 125-136

Kranz HD, Denekamp M, Greco R, Jin H, Leyva A, Meissner RC, Petroni K, Urzainqui A, Bevan M, Martin C, Smeekens S, Tonelli C, Paz-Ares J, Weisshaar B (1998) Towards functional characterisation of the members of the R2R3-MYB gene family from *Arabidopsis thaliana*. *Plant J* **16**: 263-276

Krutilina RI, Oei S, Buchlow G, Yau PM, Zalensky AO, Zalenskaya IA, Bradbury EM, Tomilin NV (2001) A negative regulator of telomere-length protein trf1 is associated with interstitial (TTAGGG)n blocks in immortal *Chinese hamster* ovary cells. *Biochem Biophys Res Commun* **280**: 471-475

Kyrion G, Boakye KA, Lustig AJ (1992) C-terminal truncation of RAP1 results in the deregulation of telomere size, stability, and function in *Saccharomyces cerevisiae*. *Mol Cell Biol* **12**: 5159-5173

Larrivee M, LeBel C, Wellinger RJ (2004) The generation of proper constitutive G-tails on yeast telomeres is dependent on the MRX complex. *Genes Dev* **18**: 1391-1396

Larrivee M, Wellinger RJ (2006) Telomerase- and capping-independent yeast survivors with alternate telomere states. *Nat Cell Biol* **8**: 741-747

Laux T, Mayer KF (1998) Cell fate regulation in the shoot meristem. *Semin Cell Dev Biol* **9**: 195-200

Lebel C, Wellinger RJ (2005) Telomeres: what's new at your end? *J Cell Sci* **118**: 2785-2788

Lei M, Baumann P, Cech TR (2002) Cooperative binding of single-stranded telomeric DNA by the Pot1 protein of *Schizosaccharomyces pombe*. *Biochemistry* **41**: 14560-14568

Lei M, Podell ER, Baumann P, Cech TR (2003) DNA self-recognition in the structure of Pot1 bound to telomeric single-stranded DNA. *Nature* **426**: 198-203

Lei M, Podell ER, Cech TR (2004) Structure of human POT1 bound to telomeric single-stranded DNA provides a model for chromosome end-protection. *Nat Struct Mol Biol* **11**: 1223-1229

Lei M, Zaug AJ, Podell ER, Cech TR (2005) Switching human telomerase on and off with hPOT1 protein *in vitro*. *J Biol Chem* **280**: 20449-20456

Leibfried A, To JP, Busch W, Stehling S, Kehle A, Demar M, Kieber JJ, Lohmann JU (2005) WUSCHEL controls meristem function by direct regulation of cytokinin-inducible response regulators. *Nature* **438**: 1172-1175

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

Levy MZ, Allsopp RC, Futcher AB, Greider CW, Harley CB (1992) Telomere end-replication problem and cell aging. *J Mol Biol* **225**: 951-960

Li B, de Lange T (2003) Rap1 affects the length and heterogeneity of human telomeres. *Mol Biol Cell* **14**: 5060-5068

Li B, Lustig AJ (1996) A novel mechanism for telomere size control in *Saccharomyces cerevisiae*. *Genes Dev* **10**: 1310-1326

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

Li GZ, Eller MS, Firoozabadi R, Gilchrest BA (2003) Evidence that exposure of the telomere 3' overhang sequence induces senescence. *Proc Natl Acad Sci U S A* **100**: 527-531

Lingner J, Hughes TR, Shevchenko A, Mann M, Lundblad V, Cech TR (1997) Reverse transcriptase motifs in the catalytic subunit of telomerase. *Science* **276**: 561-567

Liu D, O'Connor MS, Qin J, Songyang Z (2004) Telosome, a mammalian telomere-associated complex formed by multiple telomeric proteins. *J Biol Chem* **279**: 51338-51342

Liu D, Safari A, O'Connor MS, Chan DW, Laegeler A, Qin J, Songyang Z (2004) PTOP interacts with POT1 and regulates its localization to telomeres. *Nat Cell Biol* **6**: 673-680

Loayza D, de Lange T (2003) POT1 as a terminal transducer of TRF1 telomere length control. *Nature* **423**: 1013-1018

Loayza D, Parsons H, Donigian J, Hoke K, de Lange T (2004) DNA binding features of human POT1: A nonamer 5'-TAGGGTTAG-3' minimal binding site, sequence specificity, and internal binding to multimeric sites. *J Biol Chem* **279**: 13241-13248

Lugert T, Werr W (1994) A novel DNA-binding domain in the Shrunken initiator-binding protein (IBP1). *Plant Mol Biol* **25**: 493-506

Lukowiak AA, Narayanan A, Li ZH, Terns RM, Terns MP. (2001) The snoRNA domain of vertebrate telomerase RNA functions to localize the RNA within the nucleus. *RNA* **7**: 1833-1844

Lukowitz W, Gillmor CS, Scheible WR (2000) Positional cloning in *Arabidopsis.* Why it feels good to have a genome initiative working for you. *Plant Physiol* **123**: 795-805

Lundblad V (2002) Telomere maintenance without telomerase. *Oncogene* **21**: 522-531

Lundblad V, Blackburn EH (1993) An alternative pathway for yeast telomere maintenance rescues est1- senescence. *Cell* **73**: 347-360

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

Lustig AJ (2001) Cdc13 subcomplexes regulate multiple telomere functions. *Nat Struct Biol* **8**: 297-299

Lustig AJ (2003) Clues to catastrophic telomere loss in mammals from yeast telomere rapid deletion. *Nat Rev Genet* **4**: 916-923

Lustig AJ, Kurtz S, Shore D (1990) Involvement of the silencer and UAS binding protein RAP1 in regulation of telomere length. *Science* **250**: 549-553

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

Makarov VL, Lejnine S, Bedoyan J, Langmore JP (1993) Nucleosomal organization of telomere-specific chromatin in rat. *Cell* **73**: 775-787

Marcand S, Gilson E, Shore D (1997) A protein-counting mechanism for telomere length regulation in yeast. *Science* **275**: 986-990

Marchler-Bauer A, Bryant SH (2004) CD-Search: protein domain annotations on the fly. *Nucleic Acids Res* **32**: W327-331

Marian CO, Bordoli SJ, Goltz M, Santarella RA, Jackson LP, Danilevskaya O, Beckstette M, Meeley R, Bass HW (2003) The maize Single myb histone 1 gene, Smh1, belongs to a novel gene family and encodes a protein that binds telomere DNA repeats *in vitro*. *Plant Physiol* **133**: 1336-1350

Martin V, Du LL, Rozenzhak S, Russell P (2007) Protection of telomeres by a conserved Stn1-Ten1 complex. *Proc Natl Acad Sci U S A* **104**: 14038-14043

Maser RS, Wong KK, Sahin E, Xia H, Naylor M, Hedberg HM, Artandi SE, DePinho RA (2007) DNA-dependent protein kinase catalytic subunit is not required for dysfunctional telomere fusion and checkpoint response in the telomerase-deficient mouse. *Mol Cell Biol* **27**: 2253-2265

McClintock B (1939) The behavior in successive nuclear divisions of a chromosome broken at meiosis. *Proc Natl Acad Sci U S A* **25**: 405-416

McClintock B. (1941) The stability of broken ends of chromosomes in *Zea Mays*. *Genetics* **26**: 234-282

McEachern MJ, Blackburn EH (1995) Runaway telomere elongation caused by telomerase RNA gene mutations. *Nature* **376**: 403-409

McKnight TD, Riha K, Shippen DE (2002) Telomeres, telomerase, and stability of the plant genome. *Plant Mol Biol* **48**: 331-337

McKnight TD, Shippen DE (2004) Plant telomere biology. *Plant Cell* **16**: 794-803

Menges M, Murray JA (2002) Synchronous *Arabidopsis* suspension cultures for analysis of cell-cycle gene activity. *Plant J* **30**: 203-212

Mesner LD, Crawford EL, Hamlin JL (2006) Isolating apparently pure libraries of replication origins from complex genomes. *Mol Cell* **21**: 719-726

Meyerowitz EM (1997) Genetic control of cell division patterns in developing plants. *Cell* **88**: 299-308

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

Miller KM, Ferreira MG, Cooper JP (2005) Taz1, Rap1 and Rif1 act both interdependently and independently to maintain telomeres. *EMBO J* **24**: 3128-3135

Mitchell JR, Cheng J, Collins K (1999) A box H/ACA small nucleolar RNA-like domain at the human telomerase RNA 3' end. *Mol Cell Biol* **19**: 567-576

Mitton-Fry RM, Anderson EM, Hughes TR, Lundblad V, Wuttke DS (2002) Conserved structure for single-stranded telomeric DNA recognition. *Science* **296**: 145-147

Moyzis RK, Buckingham JM, Cram LS, Dani M, Deaven LL, Jones MD, Meyne J, Ratliff RL, Wu JR (1988) A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes. *Proc Natl Acad Sci U S A* **85**: 6622-6626

Muller HJ (1938) The remaking of chromosomes. *Collecting Net*, **8**: 182-198

Munoz-Jordan JL, Cross GA, de Lange T, Griffith JD (2001) t-loops at trypanosome telomeres. *EMBO J* **20**: 579-588

Muntoni A, Reddel RR (2005) The first molecular details of ALT in human tumor cells. *Hum Mol Genet* **14**: R191-196

Nishikawa T, Nagadoi A, Yoshimura S, Aimoto S, Nishimura Y (1998) Solution structure of the DNA-binding domain of human telomeric protein, hTRF1. *Structure* **6**: 1057-1065

Nishikawa T, Okamura H, Nagadoi A, Koig P, Rhodes D, Nishimura Y (2001) Structure of the DNA-binding domain of human telomeric protein, TRF1 and its interaction with telomeric DNA. *Nucleic Acids Res Suppl* 273-274

Nugent CI, Hughes TR, Lue NF, Lundblad V (1996) Cdc13p: a single-strand telomeric DNA-binding protein with a dual role in yeast telomere maintenance. *Science* **274**: 249-252

O'Connor MS, Safari A, Liu D, Qin J, Songyang Z (2004) The human Rap1 protein complex and modulation of telomere length. *J Biol Chem* **279**: 28585-28591

Ogino H, Nakabayashi K, Suzuki M, Takahashi E, Fujii M, Suzuki T, Ayusawa D (1998) Release of telomeric DNA from chromosomes in immortal human cells lacking telomerase activity. *Biochem Biophys Res Commun* **248**: 223-227

Olovnikov AM (1971) [Principle of marginotomy in template synthesis of polynucleotides]. *Dokl Akad Nauk SSSR* **201**: 1496-1499

Opresko PL, Mason PA, Podell ER, Lei M, Hickson ID, Cech TR, Bohr VA (2005) POT1 stimulates RecQ helicases WRN and BLM to unwind telomeric DNA substrates. *J Biol Chem* **280**: 32069-32080

Peersen OB, Ruggles JA, Schultz SC (2002) Dimeric structure of the *Oxytricha nova* telomere end-binding protein alpha-subunit bound to ssDNA. *Nat Struct Biol* **9**: 182-187

Pennock E, Buckley K, Lundblad V (2001) Cdc13 delivers separate complexes to the telomere for end protection and replication. *Cell* **104**: 387-396

Petreaca RC, Chiu HC, Eckelhoefer HA, Chuang C, Xu L, Nugent CI (2006) Chromosome end protection plasticity revealed by Stn1p and Ten1p bypass of Cdc13p. *Nat Cell Biol* **8**: 748-755

Pich U, Fuchs J, Schubert I (1996) How do *Alliaceae* stabilize their chromosome ends in the absence of TTTAGGG sequences? *Chromosome Res* **4**: 207-213

Pina B, Fernandez-Larrea J, Garcia-Reyero N, Idrissi FZ (2003) The different (sur)faces of Rap1p. *Mol Genet Genomics* **268**: 791-798

Pitt CW, Moreau E, Lunness PA, Doonan JH (2004) The pot1+ homologue in *Aspergillus nidulans* is required for ordering mitotic events. *J Cell Sci* **117**: 199-209

Polotnianka RM, Li J, Lustig AJ (1998) The yeast Ku heterodimer is essential for protection of the telomere against nucleolytic and recombinational activities. *Curr Biol* **8**: 831-834

Price CM (1990) Telomere structure in *Euplotes crassus*: characterization of DNA-protein interactions and isolation of a telomere-binding protein. *Mol Cell Biol* **10**: 3421-3431

Price CM, Cech TR (1987) Telomeric DNA-protein interactions of *Oxytricha* macronuclear DNA. *Genes Dev* **1**: 783-793

Qi, H. and Zakian, V.A. (2000) The *Saccharomyces* telomere-binding protein Cdc13p interacts with both the catalytic subunit of DNA polymerase alpha and the telomerase-associated est1 protein. *Genes Dev*, **14**, 1777-1788

Raices M, Verdun RE, Compton SA, Haggblom CI, Griffith JD, Dillin A, Karlseder J (2008) *C. elegans* telomeres contain both G-strand and C-strand overhangs that are bound by distinct telomere-binding proteins. *Mol Cell*, in press

Ranganathan V, Heine WF, Ciccone DN, Rudolph KL, Wu X, Chang S, Hai H, Ahearn IM, Livingston DM, Resnick I, Rosen F, Seemanova E, Jarolim P, DePinho RA, Weaver DT (2001) Rescue of a telomere length defect of Nijmegen breakage syndrome cells requires NBS and telomerase catalytic subunit. *Curr Biol* **11**: 962-966

Regad F, Lebas M, Lescure B (1994) Interstitial telomeric repeats within the *Arabidopsis thaliana* genome. *J Mol Biol* **239**: 163-169

Regev A, Cohen S, Cohen E, Bar-Am I, Lavi S (1998) Telomeric repeats on small polydisperse circular DNA (spcDNA) and genomic instability. *Oncogene* **17**: 3455-3461

Reichenbach P, Hoss M, Azzalin CM, Nabholz M, Bucher P, Lingner J (2003) A human homolog of yeast Est1 associates with telomerase and uncaps chromosome ends when overexpressed. *Curr Biol* **13**: 568-574

Rhodes D, Fairall L, Simonsson T, Court R, Chapman L (2002) Telomere architecture. *EMBO Reports* **3**: 1139-1145

Richards EJ, Ausubel FM (1988) Isolation of a higher eukaryotic telomere from *Arabidopsis thaliana*. *Cell* **53**: 127-136

Riha K, Heacock ML, Shippen DE (2006) The role of the nonhomologous end-joining DNA double-strand break repair pathway in telomere biology. *Annu Rev Genet* **40**: 237-277

Riha K, McKnight TD, Fajkus J, Vyskot B, Shippen DE (2000) Analysis of the G-overhang structures on plant telomeres: evidence for two distinct telomere architectures. *Plant J* **23**: 633-641

Riha K, McKnight TD, Griffing LR, Shippen DE (2001) Living with genome instability: plant responses to telomere dysfunction. *Science* **291**: 1797-1800

Riha K, Shippen DE (2003a) Ku is required for telomeric C-rich strand maintenance but not for end-to-end chromosome fusions in *Arabidopsis. Proc Natl Acad Sci U S A* **100**: 611-615

Riha K, Shippen DE (2003b) Telomere structure, function and maintenance in *Arabidopsis*. *Chromosome Res* **11**: 263-275

Riha K, Watson JM, Parkey J, Shippen DE (2002) Telomere length deregulation and enhanced sensitivity to genotoxic stress in *Arabidopsis* mutants deficient in Ku70. *EMBO J* **21**: 2819-2826

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**: 406-425

Schrumpfova P, Kuchar M, Mikova G, Skrisovska L, Kubicarova T, Fajkus J (2004) Characterization of two *Arabidopsis thaliana* myb-like proteins showing affinity to telomeric DNA sequence. *Genome* **47**: 316-324

Seimiya H, Sawada H, Muramatsu Y, Shimizu M, Ohko K, Yamane K, Tsuruo T (2000) Involvement of 14-3-3 proteins in nuclear localization of telomerase. *EMBO J* **19**: 2652-2661

Shakirov EV, Shippen DE (2004) Length regulation and dynamics of individual telomere tracts in wild-type *Arabidopsis*. *Plant Cell* **16**: 1959-1967

Shakirov EV, Surovtseva YV, Osbun N, Shippen DE (2005) The *Arabidopsis* Pot1 and Pot2 proteins function in telomere length homeostasis and chromosome end protection. *Mol Cell Biol* **25**: 7725-7733

Shampay J, Szostak JW, Blackburn EH (1984) DNA sequences of telomeres maintained in yeast. *Nature* **310**: 154-157

Shay JW, Wright WE (2005) Senescence and immortalization: role of telomeres and telomerase. *Carcinogenesis* **26**: 867-874

Shore D, 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

Silverman J, Takai H, Buonomo SB, Eisenhaber F, de Lange T (2004) Human Rif1, ortholog of a yeast telomeric protein, is regulated by ATM and 53BP1 and functions in the S-phase checkpoint. *Genes Dev* **18**: 2108-2119

Singer MS, Gottschling DE (1994) TLC1: template RNA component of *Saccharomyces cerevisiae* telomerase. *Science* **266**: 404-409

Skopp R, Wang W, Price C (1996) rTP: a candidate telomere protein that is associated with DNA replication. *Chromosoma* **105**: 82-91

Smith S, de Lange T (2000) Tankyrase promotes telomere elongation in human cells. *Curr Biol* **10**: 1299-1302

Smith S, Giriat I, Schmitt A, de Lange T (1998) Tankyrase, a poly(ADP-ribose) polymerase at human telomeres. *Science* **282**: 1484-1487

Smogorzewska A, de Lange T (2004) Regulation of telomerase by telomeric proteins. *Annu Rev Biochem* **73**: 177-208

Smogorzewska A, van Steensel B, Bianchi A, Oelmann S, Schaefer MR, Schnapp G, de Lange T (2000) Control of human telomere length by TRF1 and TRF2. *Mol Cell Biol* **20**: 1659-1668

Snow BE, Erdmann N, Cruickshank J, Goldman H, Gill RM, Robinson MO, Harrington L (2003) Functional conservation of the telomerase protein Est1p in humans. *Curr Biol* **13**: 698-704

Stansel RM, de Lange T, Griffith JD (2001) T-loop assembly i*n vitro* involves binding of TRF2 near the 3' telomeric overhang. *EMBO J* **20**: 5532-5540

Stewart SA, Ben-Porath I, Carey VJ, O'Connor BF, Hahn WC, Weinberg RA (2003) Erosion of the telomeric single-strand overhang at replicative senescence. *Nat Genet* **33**: 492-496

Surovtseva YV, Shakirov EV, Vespa L, Osbun N, Song X, Shippen DE (2007) *Arabidopsis* POT1 associates with the telomerase RNP and is required for telomere maintenance. *EMBO J* **26**: 3653-3661

Sykorova E, Leitch AR, Fajkus J (2006) *Asparagales* telomerases which synthesize the human type of telomeres. *Plant Mol Biol* **60**: 633-646

Szostak JW, Blackburn EH (1982) Cloning yeast telomeres on linear plasmid vectors. *Cell* **29**: 245-255

Taggart AK, Teng SC, Zakian VA (2002) Est1p as a cell cycle-regulated activator of telomere-bound telomerase. *Science* **297**: 1023-1026

Takai H, Smogorzewska A, de Lange T (2003) DNA damage foci at dysfunctional telomeres. *Curr Biol* **13**: 1549-1556

Takata H, Kanoh Y, Gunge N, Shirahige K, Matsuura A (2004) Reciprocal association of the budding yeast ATM-related proteins Tel1 and Mec1 with telomeres *in vivo*. *Mol Cell* **14**: 515-522

Takata H, Tanaka Y, Matsuura A (2005) Late S phase-specific recruitment of Mre11 complex triggers hierarchical assembly of telomere replication proteins in *Saccharomyces cerevisiae*. *Mol Cell* **17**: 573-583

Takeda S, Tadele Z, Hofmann I, Probst AV, Angelis KJ, Kaya H, Araki T, Mengiste T, Mittelsten Scheid O, Shibahara K, Scheel D, Paszkowski J (2004) BRU1, a novel link between responses to DNA damage and epigenetic gene silencing in *Arabidopsis*. *Genes Dev* **18**: 782-793

Tani A, Murata M (2005) Alternative splicing of Pot1 (Protection of telomere)-like genes in *Arabidopsis thaliana*. *Genes Genet Syst* **80**: 41-48

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

Teixeira MT, Forstemann K, Gasser SM, Lingner J (2002) Intracellular trafficking of yeast telomerase components. *EMBO Rep* **3**: 652-659

Teng SC, Zakian VA (1999) Telomere-telomere recombination is an efficient bypass pathway for telomere maintenance in *Saccharomyces cerevisiae*. *Mol Cell Biol* **19**: 8083-8093

Theimer CA, Feigon J (2006) Structure and function of telomerase RNA. *Curr Opin Struct Biol* **16**: 307-318

Theobald DL, Mitton-Fry RM, Wuttke DS (2003) Nucleic acid recognition by OB-fold proteins. *Annu Rev Biophys Biomol Struct* **32**: 115-133

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

Till BJ, Colbert T, Tompa R, Enns LC, Codomo CA, Johnson JE, Reynolds SH, Henikoff JG, Greene EA, Steine MN, Comai L, Henikoff S (2003) High-throughput TILLING for functional genomics. *Methods Mol Biol* **236**: 205-220

Tomaska L, Willcox S, Slezakova J, Nosek J, Griffith JD (2004) Taz1 binding to a fission yeast model telomere: formation of telomeric loops and higher order structures. *J Biol Chem* **279**: 50764-50772

Tommerup H, Dousmanis A, de Lange T (1994) Unusual chromatin in human telomeres. *Mol Cell Biol* **14**: 5777-5785

Tremousaygue D, Garnier L, Bardet C, Dabos P, Herve C, Lescure B (2003) Internal telomeric repeats and 'TCP domain' protein-binding sites co-operate to regulate gene expression in *Arabidopsis thaliana* cycling cells. *Plant J* **33**: 957-966

Tremousaygue D, Manevski A, Bardet C, Lescure N, Lescure B (1999) Plant interstitial telomere motifs participate in the control of gene expression in root meristems. *Plant J* **20**: 553-561

Trujillo KM, Bunch JT, Baumann P (2005) Extended DNA binding site in Pot1 broadens sequence specificity to allow recognition of heterogeneous fission yeast telomeres *J Biol Chem* **280**: 9119-9128

Vaillant I, Schubert I, Tourmente S, Mathieu O (2006) MOM1 mediates DNA-methylation-independent silencing of repetitive sequences in *Arabidopsis*. *EMBO Rep* **7**: 1273-1278

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

van Steensel B, Smogorzewska A, de Lange T (1998) TRF2 protects human telomeres from end-to-end fusions. *Cell* **92**: 401-413

Vaziri H, West MD, Allsopp RC, Davison TS, Wu YS., Arrowsmith CH, Poirier GG, Benchimol S (1997) ATM-dependent telomere loss in aging human diploid fibroblasts and DNA damage lead to the post-translational activation of p53 protein involving poly(ADP-ribose) polymerase. *EMBO J* **16**: 6018-6033

Veldman T, Etheridge KT, Counter CM (2004) Loss of hPot1 function leads to telomere instability and a cut-like phenotype. *Curr Biol* **14**: 2264-2270

Verdun RE, Crabbe L, Haggblom C, Karlseder J (2005) Functional human telomeres are recognized as DNA damage in G2 of the cell cycle. *Mol Cell* **20**: 551-561

Verdun RE, Karlseder J (2006) The DNA damage machinery and homologous recombination pathway act consecutively to protect human telomeres. *Cell* **127**: 709-720

Verdun RE, Karlseder J (2007) Replication and protection of telomeres. *Nature* **447**: 924-931

Vespa L, Couvillion M, Spangler E, Shippen DE (2005) ATM and ATR make distinct contributions to chromosome end protection and the maintenance of telomeric DNA in *Arabidopsis*. *Genes Dev* **19**: 2111-2115

Vongs A, Kakutani T, Martienssen RA, Richards EJ (1993) *Arabidopsis thaliana* DNA methylation mutants. *Science* **260**: 1926-1928

Wang F, Podell ER., Zaug AJ, Yang Y, Baciu P, Cech TR, Lei M (2007) The POT1-TPP1 telomere complex is a telomerase processivity factor. *Nature* **445**: 506-510

Wang RC, Smogorzewska A, de Lange T (2004) Homologous recombination generates T-loop-sized deletions at human telomeres. *Cell* **119**: 355-368

Wang W, Skopp R, Scofield M, Price C (1992) *Euplotes crassus* has genes encoding telomere-binding proteins and telomere-binding protein homologs. *Nucleic Acids Res* **20**: 6621-6629

Watson JD (1972) Origin of concatemeric T7 DNA. *Nat New Biol* **239**: 197-201

Watson JM, Bulankova P, Riha K, Shippen DE, Vyskot B (2005) Telomerase-independent cell survival in *Arabidopsis thaliana. Plant J* **43**: 662-674

Watson JM, Shippen DE (2007) Telomere rapid deletion regulates telomere length in *Arabidopsis thaliana. Mol Cell Biol* **27**: 1706-1715

Wei C, Price CM (2004) Cell cycle localization, dimerization, and binding domain architecture of the telomere protein cPot1. *Mol Cell Biol* **24**: 2091-2102

Wei C, Price M (2003) Protecting the terminus: t-loops and telomere end-binding proteins. *Cell Mol Life Sci* **60**:2283-2294

Weinrich SL, Pruzan R, Ma L, Ouellette M, Tesmer VM, Holt SE, Bodnar AG, Lichtsteiner S, Kim NW, Trager JB, Taylor RD, Carlos R, Andrews WH, Wright WE, Shay JW, Harley CB, Morin GB (1997) Reconstitution of human telomerase with the template RNA component hTR and the catalytic protein subunit hTRT. *Nat Genet* **17**: 498-502

Wellinger RJ, Ethier K, Labrecque P, Zakian VA (1996) Evidence for a new step in telomere maintenance. *Cell* **85**: 423-433

Wellinger RJ, Wolf AJ, Zakian VA (1992) Use of non-denaturing Southern hybridization and two dimensional agarose gels to detect putative intermediates in telomere replication in *Saccharomyces cerevisiae*. *Chromosoma* **102**: S150-156

Wellinger RJ, Wolf AJ, Zakian VA (1993) *Saccharomyces* telomeres acquire single-strand TG1-3 tails late in S phase. *Cell* **72**: 51-60

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

Wright WE, Piatyszek MA, Rainey WE, Byrd W, Shay JW (1996) Telomerase activity in human germline and embryonic tissues and cells. *Dev Genet* **18**: 173-179

Wright WE, Tesmer VM, Huffman KE, Levene SD, Shay JW (1997) Normal human chromosomes have long G-rich telomeric overhangs at one end. *Genes Dev* **11**: 2801-2809

Wu L, Multani AS, He H, Cosme-Blanco W, Deng Y, Deng JM, Bachilo O, Pathak S, Tahara H, Bailey SM, Deng Y, Behringer RR, Chang S (2006) Pot1 deficiency initiates DNA damage checkpoint activation and aberrant homologous recombination at telomeres. *Cell* **126**: 49-62

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

Yang Q, Zheng YL, Harris CC (2005) POT1 and TRF2 cooperate to maintain telomeric integrity. *Mol Cell Biol* **25**: 1070-1080

Yang SW, Kim DH, Lee JJ, Chun YJ, Lee JH, Kim YJ, Chung IK, Kim WT (2003) Expression of the telomeric repeat binding factor gene NgTRF1 is closely coordinated with the cell division program in tobacco BY-2 suspension culture cells. *J Biol Chem* **278**: 21395-21407

Yang SW, Kim SK, Kim WT (2004) Perturbation of NgTRF1 expression induces apoptosis-like cell death in tobacco BY-2 cells and implicates NgTRF1 in the control of telomere length and stability. *Plant Cell* **16**: 3370-3385

Ye JZ, Donigian JR, van Overbeek M, Loayza D, Luo Y, Krutchinsky AN, Chait BT, de Lange T (2004a) TIN2 binds TRF1 and TRF2 simultaneously and stabilizes the TRF2 complex on telomeres. *J Biol Chem* **279**: 47264-47271

Ye JZ, Hockemeyer D, Krutchinsky AN, Loayza D., Hooper SM, Chait BT, de Lange T (2004b) POT1-interacting protein PIP1: a telomere length regulator that recruits POT1 to the TIN2/TRF1 complex. *Genes Dev* **18**: 1649-1654

Yu EY, Kim SE, Kim JH, Ko JH, Cho MH, Chung IK (2000) Sequence-specific DNA recognition by the Myb-like domain of plant telomeric protein RTBP1. *J Biol Chem* **275**: 24208-24214

Zakian VA (1996) Structure, function, and replication of *Saccharomyces cerevisiae* telomeres. *Annual Review of Genetics* **30**: 141-172

Zaug AJ, Podell ER, Cech TR (2005) Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension *in vitro*. *Proc Natl Acad Sci U S A* **102**: 10864-10869

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

Zellinger B, Riha K (2007) Composition of plant telomeres. *Biochim Biophys Acta* **1769**: 399-409

Zentgraf U (1995) Telomere-binding proteins of *Arabidopsis thaliana*. *Plant Molecular Biology* **27**: 467-475

Zentgraf U, Hinderhofer K, Kolb D (2000) Specific association of a small protein with the telomeric DNA-protein complex during the onset of leaf senescence in *Arabidopsis thaliana*. *Plant Molecular Biology* **42**: 429-438

Zhou XZ, Lu KP (2001) The Pin2/TRF1-interacting protein PinX1 is a potent telomerase inhibitor. *Cell* **107**: 347-359

Zhu J, Wang H, Bishop JM, Blackburn EH (1999) Telomerase extends the lifespan of virus-transformed human cells without net telomere lengthening. *Proc Natl Acad Sci U S A* **96**: 3723-3728

Zhu XD, Niedernhofer L, Kuster B, Mann M, Hoeijmakers JH, de Lange T (2003) ERCC1/XPF removes the 3' overhang from uncapped telomeres and represses formation of telomeric DNA-containing double minute chromosomes. *Mol Cell* **12**: 1489-1498

Zou L, Elledge SJ (2003) Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. *Science* **300**: 1542-1548

Zubko MK, Lydall D (2006) Linear chromosome maintenance in the absence of essential telomere-capping proteins. *Nat Cell Biol* **8**: 734-740

VITA

Name:	Yulia V. Surovtseva
Address:	Department of Biochemistry and Biophysics 2128 TAMU College Station, TX, 77843-2128
Education:	Ph.D. Biochemistry, Texas A&M University, 2008 B.S., Medical Biology, Novosinibrsk State University, Russia, 2002