TELOMERE PROTECTION AND MAINTENANCE IN

Arabidopsis thaliana

A Dissertation

by

XIANGYU SONG

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 2010

Major Subject: Biochemistry

TELOMERE PROTECTION AND MAINTENANCE IN

Arabidopsis thaliana

A Dissertation

by

XIANGYU SONG

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,	Alan E. Pepper
	Donald W. Pettigrew
	Michael Polymenis
Head of Department,	Gregory D. Reinhart

May 2010

Major Subject: Biochemistry

ABSTRACT

Telomere Protection and Maintenance in *Arabidopsis thaliana*.

(May 2010)

Xiangyu Song, B.S., Nankai University, China

Chair of Advisory Committee: Dr. Dorothy E. Shippen

Telomeres are the physical ends of linear chromosomes in eukaryotes. Telomeres not only protect chromosome ends from being recognized as double-strand breaks but also maintain the chromosome terminal sequences. These processes involve a number of telomere-related proteins. A major challenge in the field is to elucidate the full constitution of telomere-associated proteins and to understand how different protein complexes are regulated at chromosome termini.

Here, I report the identification and characterization of STN1 (Suppressor of cdc thirteen, 1), CTC1 (Conserved Telomere maintenance Component 1) and TEN1 (Telomeric pathways in association with Stn1, 1) in *Arabidopsis*. CTC1/STN1/TEN1 (CST) forms a trimeric complex that specifically associates with telomeres. Loss of any component of the CST induces catastrophic telomere loss, disrupted telomere end architecture, and massive chromosome end-to-end fusions. Thus, CST plays an essential role in chromosome end protection. I also show that CST function at telomeres is independent of a previously characterized capping complex KU70/KU80, and that ATR is responsible for a checkpoint response in plants lacking CTC1/STN1.

Additionally, I present data showing that *Arabidopsis* POT1a (Protection Of Telomere 1, a) has evolved as a telomerase recruitment factor. Unlike POT1 in other

eukaryotes which binds and protects ss telomeric DNA, AtPOT1a interacts with telomerase RNA (TER). Based on an evolutionary analysis, we found that the POT1a lineage is under positive selection in the Brassicaceae family in which *Arabidopsis* belongs. Mutations of two positive selection sites significantly reduce POT1a's activity *in vivo*. These data suggest POT1a is under pressure to evolve from a telomeric DNA binding protein to a TER binding protein. I also discovered that POT1a interacts with the novel telomere capping protein CTC1 *in vitro* and *in vivo*. Thus, I hypothesize that POT1a acts as a telomerase recruitment factor linking this enzyme to the chromosome termini via interacting with TER and CTC1. Finally, I dissected the functional domains of POT1a and demonstrated that both the N-terminus and the C-terminus of POT1a are required for its function *in vivo*.

In summary, my work has uncovered several new and essential telomere-associated proteins that provide new insight into mechanisms of chromosome end protection and maintenance.

DEDICATION

To my parents and my husband for their endless love and support,

there is no chance that I could have done this without them.

ACKNOWLEDGEMENTS

First of all, I would like to thank my advisor, Dr. Dorothy Shippen. Dorothy is one of the most energetic and enthusiastic scientists I've ever met. She keeps inspiring me with new ideas and encouraging me to aim high throughout my graduate studies. She also teaches me to think critically and creatively. Without her endless support, I could not have been to this step.

I am also grateful to my committee members, Dr. Alan Pepper, Dr. Michael Polymenis and Dr. Donald Pettigrew for their guidance and support. I deeply appreciate the help of previous committee members, Dr. David Giedroc and Dr. Paul Fitzpatrick. I want to extend my gratitude to Dr. Geoffrey Kapler and Dr. Thomas McKnight for their helpful and critical advice throughout the course of my research. I also want to thank Dr. Xiuren Zhang for his advice and encouragement.

Special thanks go to Dr. Eugene Shakirov, who was my mentor when I rotated in the lab, and he has inspired me in many ways. We collaborated on the POT1 project, which was such a journey of fun. I would also thank Dr. Yulia Surovtseva, Katie Leehy and Kara Boltz, Jon Lamb and Ross Warrington. We worked closely on the characterization of CTC1/STN1/TEN1 complex. Without all their help, I could not publish my research work in such a short time. I am grateful to all other Shippen lab members as well. I can not ask for more from lab members like Cathy, Andrew, Matt, Michelle, Kalpana, Dr. Mark Beilstein and Dr. Jung Ro Lee. It is a wonderful experience for me to work with these guys.

I am grateful to the faculty and staff in the Department of Biochemistry and Biophysics for making my time at Texas A&M University a great experience. Many

thanks to Pat, Tillie, Juanita and Sherry; their assistance has made my life much easier. I am also in debt to Guogang Dong, who helped me to settle down in College Station. All the selfless help is precious and priceless for a foreigner like me, who had just come to United States for the first time.

Finally, I would like to thank Dr. Zhen Li for reading and making corrections to this dissertation.

TABLE OF CONTENTS

CHAPTER Page

LIST OF FIGURES

Page

Page

Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

Studies of telomeres started back in the 1930s, when Barbara McClintock first discovered the importance of “natural ends” of the chromosomes in maize (McClintock, 1931). She reported the cytology of anaphase bridges caused by chromosomes with broken ends, as well as a model of breakage-fusion-bridge cycle to explain such phenomena (McClintock, 1939). Around the same time, Hermann Muller also observed that the ends of fly chromosomes are critical for genome stability (Muller, 1938). Muller named these “natural ends” telomeres (Muller, 1938).

While McClintock and Muller’s works dramatically influenced our understanding of telomeres, modern telomere studies did not explode until 1970s. In 1978, the telomeric DNA sequence was first identified by Elizabeth Blackburn (Blackburn & Gall, 1978), followed by a series of groundbreaking discoveries, including the identification and characterization of telomerase (Greider & Blackburn, 1985), telomere-associated proteins (Berman et al, 1986; Gottschling & Zakian, 1986), and telomere end architecture (Griffith et al, 1999; Klobutcher et al, 1981). Importantly, Elizabeth Blackburn, Carol Greider and Jack Szostak were awarded the 2009 Novel Prize in Physiology or Medicine for their significant contribution to telomere and telomerase studies, which substantially inspired new cancer therapies and deepened views on cell aging. All these studies, and perhaps many more to come, allow us to better visualize

_____

This dissertation follows the style of *European Molecular Biology Organization Journal*.

the ends of chromosomes and their functions in numerous cell activities in eukaryotes.

Telomeres and telomeric DNA

The first, and perhaps the biggest breakthrough of modern telomere research was the identification of telomeric DNA sequence by Elizabeth Blackburn (Blackburn & Gall, 1978). Blackburn found that chromosomes in the ciliated protozoan *Tetrahymena thermophila* end in tandem repeats of GGGGTT. Szostak and Blackburn further introduced into budding yeast a linear plasmid that carries *Tetrahymena* telomere sequence at the ends (Szostak & Blackburn, 1982). Strikingly, the linear plasmid was stable in yeast. Sequencing of the ends indicated that a new type of telomeric sequence was added, corresponding to the budding yeast-type of irregular telomeric sequence (formula TG_{1-3}) (Szostak & Blackburn, 1982).

Subsequently, it was shown that most other eukaryotes carry similar tandemly repeated G-rich sequence at telomeres. For instance, telomeric DNA is composed of TTAGGG repeats in vertebrates (Moyzis RK, 1988) and TTTAGGG repeats in most plant species (Richards & Ausubel, 1988; Zellinger & Riha, 2007). The tandem repeats of TTTAGGG are widely spread from single cell green algae *Chlorella vulgaris*, to the model dicot plant *Arabidopsis*, and to monocot crops such as maize (Zellinger & Riha, 2007). The notable exceptions are one of the alga species *Chlamydomonas reinhardtii*, which contains TTTTAGGG telomeric repeats (Petracek et al, 1990), and plants from the Asparagale order, where chromosome termini consist of vertebrate-type repeats of TTAGGG (Sykorova et al, 2006), as well as some species from Alium such as onion, where telomeric DNA contains unknown sequences (Pich et al, 1996).

These G-rich repeated telomeric sequences are maintained by telomerase and play important roles for chromosome end protection and genome stability (see below).

While the G-rich telomeric sequence is conserved in most eukaryotes, exception exists in *Drosophila* where telomeres are composed of retrotransposable elements (Biessmann et al, 1992). Sporadic addition of new transposon at the chromosome ends fully compensates the terminal loss caused by incomplete replication at telomeres (Mason & Biessmann, 1995). And as such, *Drosophila* telomere maintenance and protection involves a different set of telomere-related proteins, which will not be further discussed in this thesis.

G-overhangs

When telomeric DNA was first discovered in *Oxytricha* and *Euplotes*, sequencing of both telomeric strands indicated that telomeres end with a 3' extrusion, which is approximately 12-16 nt long in different ciliate species (Klobutcher et al, 1981; Pluta et al, 1982). This 3' ss G-rich strand is termed the G-overhang (Fig 1-1A). G-overhangs are likely present in all eukaryotes carrying G-rich telomeric repeats, including ciliates (see above), budding yeast (~ 12-14 nt) (Larrivee et al, 2004), humans (~ 250 nt) (Makarov et al, 1997) and *Arabidopsis* (~ 20 nt) (Riha et al, 2000).

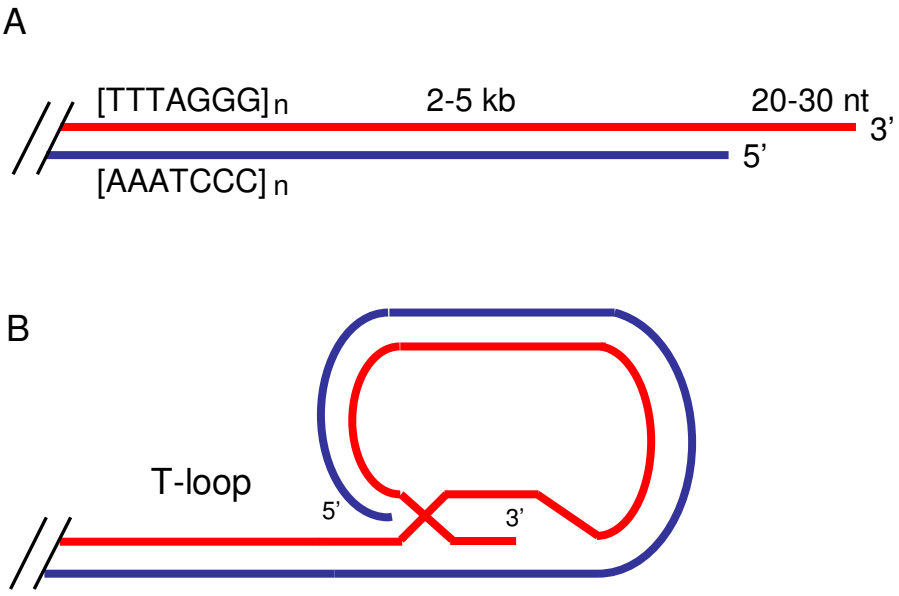


Fig 1-1. Telomere structure. (**A**) Schematic of a linear model of telomeres in *Arabidopsis* showing the 3’ G-overhang. The G-strand of the telomere is shown in red and the C-strand is shown in blue. (**B**) The t-loop structure. The 3’ G-overhang folds back and invades into the duplex telomeric DNA.

One hypothesis is that G-overhang is a product of telomerase once it extends at the G-rich strand. However, Jacob and his colleagues reported the presence of G-overhangs in a telomerase mutant, indicating telomerase is dispensable for generating G-overhangs (Jacob et al, 2003). It is now believed that G-overhangs are produced by nucleolytic resection at the C-strand rather than telomerase action (Bonetti et al, 2009; Chai et al, 2006; Jacob et al, 2003).

Abnormally extended G-overhangs are associated with impaired cell viability and senescence in yeast and human cells (Grandin et al, 2001; Grandin et al, 1997; Li et al, 2003; Nugent et al, 1996; Stewart et al, 2003). To protect the ends of chromosomes, specific G-overhang binding proteins are recruited to the telomere ends. Such proteins include POT1 (Protection Of Telomere 1) in humans and fission yeast, and Cdc13 (Cell division cycle 13) in budding yeast (see below). The G-overhangs and G-overhang associated proteins are crucial to protect telomeres from end-to-end fusions and to facilitate telomerase action (Zhu et al, 2003).

Interestingly, it has been recently reported that *C. elegans* possesses not only G-overhangs, but also C-overhangs (Raices et al, 2008). Moreover, the G- and C-overhangs appear to associate with distinct telomere binding proteins, CeOB1 and CeOB2, respectively (Raices et al, 2008). So far, C-overhangs have not been reported in any other organisms.

T-loops

First shown by Jack Griffith using electron microscopy, mammalian telomeres appear to form a lariat structure *in vitro* (Griffith et al, 1999). This structure is named the telomeric loop or t-loop (Fig 1-1B). In the same study, Griffith et al reported that t-loop

does not form in the absence of G-overhangs, indicating G-overhangs are crucial for a t-loop formation. T-loops can be observed in either a protein-free state when cross-linked with psoralen prior to DNA purification (Griffth et al, 1999), or a native state when telomeric chromatin was isolated without cross-linking (Nikitina & Woodcock, 2004). These studies indicate the t-loop structure is probably formed by invasion of the 3' G-overhang into the duplex region of the telomere tracts (Griffith et al, 1999) (Fig 1-1B). T-loops have also been detected in ciliates, plants, *C. elegans*, *K. lactis*, arguing that they are a conserved feature of eukaryote telomeres (Cesare et al, 2008; Cesare et al, 2003; Murti & Prescott, 1999; Raices et al, 2008).

While the exact function of the t-loop *in vivo* is still not clear, it is proposed to help sequester the 3' G-overhang and protect telomeres from nucleolytic attack, chromosome end-to-end fusions or other deleterious events (Palm & de Lange, 2008). It is also not known if t-loop persists throughout the cell cycle. Indeed, it seems likely that the t-loop is resolved during S-phase to allow access to telomerase and other components of the telomere replication machinery (LeBel & Wellinger, 2005). Budding yeast telomeres do not form a t-loop, perhaps because the telomeric sequences are more degenerate than other eukaryotes. Nonetheless, a fold-back structure is observed at budding yeast telomeres (de Bruin et al, 2001), reinforcing the importance of maintaining a higher order structure at chromosomal termini.

The end replication problem

In 1972, James Watson first described the “end-replication problem”: all known DNA polymerases require a polynucleotide primer, which will be removed once synthesis has been primed, leaving a 5’-terminal gap in one of the daughter strands of a linear DNA molecule (Watson, 1972) (Fig 1-2). Without solving this problem, cells carrying linear chromosomes would gradually lose terminal DNA from generation to generation, eventually leading to cell death.

In the same paper, Watson found that the ends of linear DNA molecule in T7 phage are fused to form concatenmers, and thus circumvent the end replication problem (Watson, 1972). Subsequently, Bateman and others hypothesized that telomeres may end with a hairpin structure (Bateman, 1975; Cavalier-Smith, 1974). In this model, telomeric DNA sequence is palindromic, so that the 3’ ss region can fold back and anneal to itself. In fact, this is the case in vaccinia virus, where the chromosome termini mainly consist of A/T residues which are incompletely base-paired, ending with a fold-back structure (Baroudy et al, 1982). Because of this finding, the hairpin structure model of the telomere was quite popular in the 1970s and 1980s. However, later research revealed that telomeric DNA is not palindromic and that the end replication problem is solved in most eukaryotes by a novel enzyme— telomerase (see below).

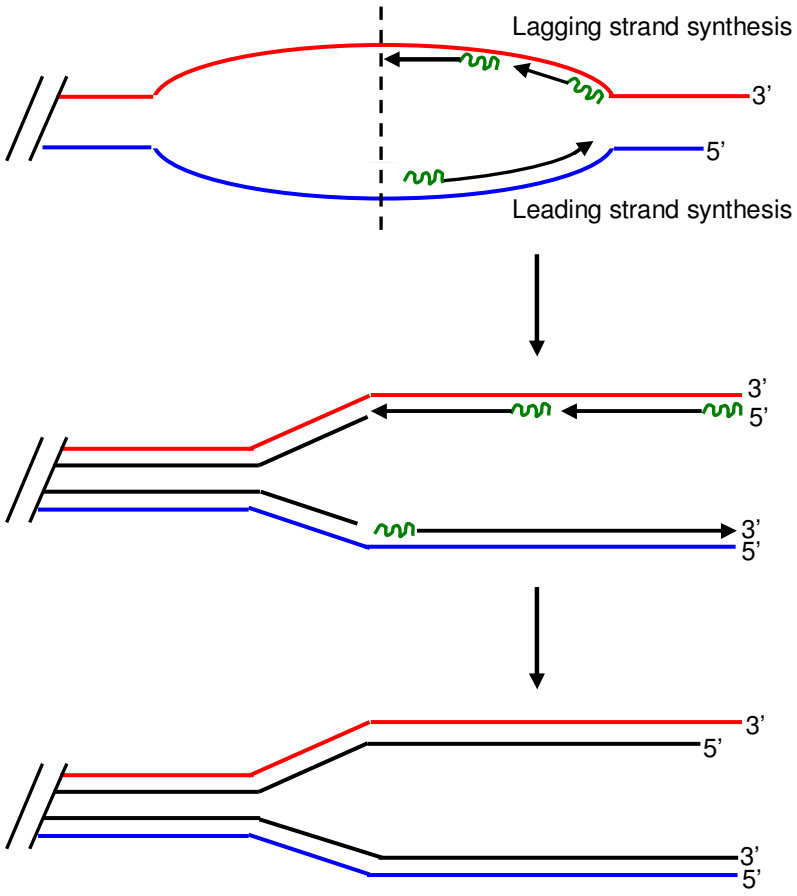


Fig 1-2. The end replication problem. For the lagging-strand DNA synthesis, polymerase α–primase complex initiates the synthesis of Okazaki fragments by laying down a short RNA primer. The RNA primers are shown as green wavy lines. The lagging-strand synthesis also requires polymerase δ and other enzymes (see text). The products of semi-conservative DNA replication are shown in black. A 5’ terminal gap is left after the RNA primer is removed, resulting in an incomplete replication at the chromosome ends.

Telomere replication

Most of the telomere tract is replicated by conventional semiconservative replication (Chakhparonian & Wellinger, 2003). Replication is initiated from an origin in the subtelomeric region, with one replication fork moving toward the chromosome end. DNA polymerase ε is involved in the leading strand synthesis (Pursell et al, 2007), while replication of the lagging strand is carried out by other polymerase complexes. In particular, polymerase α–primase complex initiates the synthesis of Okazaki fragments by laying down a short RNA primer (10–16 nt) (Fig 1-2). Polymerase α extends the primer synthesizing a short stretch of DNA before polymerase δ comes in and generates the full Okazaki fragment. The RNA primer is removed, and the gap is filled by polymerase δ (Garg & Burgers, 2005). Finally, the remaining nick ligated by DNA ligase.

At the very end of telomeres, the C-strand telomeric DNA is still synthesized by polymerase α-primase in concert with other conventional replication machinery. In addition, a special enzyme telomerase is required to fully replicate the G-strand telomeric DNA. Replication of the G- and C-strand telomeric DNA is coordinated by coupling telomerase and polymerase α-primase actions at the chromosome end (see below).

Telomerase

Telomerase is a ribonucleoprotein (RNP) reverse transcriptase that adds telomere repeats onto chromosome ends. Telomerase activity was first identified by Carol Greider and Elizabeth Blackburn in *Tetrahymena thermophila* (Greider & Blackburn, 1985), and later in other eukaryotes as well. The activity was eliminated when treated with either protease or RNase, but not DNase, indicating that telomerase is a ribonucleoprotein complex (Greider & Blackburn, 1985; Greider & Blackburn, 1989).

Telomerase is composed of two core components: the catalytic subunit, telomerase reverse transcriptase (TERT), and an RNA template subunit, telomerase RNA (TER). A biochemical purification of telomerase from Euplotes (Lingner & Cech, 1996) and a genetic screen from budding yeast (Lendvay et al, 1996) led to the identification of TERT, which contains conserved reverse transcriptase motifs (Lingner et al, 1997b). TERT was subsequently found in other eukaryotes, including fission yeast, human, *Tetrahymena*, and *Arabidopsis* (Bryan et al, 1998; Fitzgerald et al, 1999; Harrington et al, 1997; Nakamura et al, 1997). In contrast to TERT, the sequence of TER is not well conserved, although this molecule can assume a conserved secondary structure (Chen et al, 2000; Romero & Blackburn, 1981). So far, TER has been identified as a single-copy gene in ciliates, budding yeast, fission yeast, and vertebrates (Chen et al, 2000; Feng et al, 1995; Leonardi et al, 2008; Romero & Blackburn, 1981; Singer & Gottschling, 1994). Recent studies in our lab uncovered two TERs in *Arabidopsis thaliana*, that is, TER1 and TER2 (C. Cifuentes-Rojas and D.E. Shippen, unpublished data). TER contains a non-paired ss template region, which is complementary to the G-strand telomeric DNA (Chen et al, 2000; Feng et al, 1995; Greider & Blackburn, 1989; Leonardi et al, 2008; Romero & Blackburn, 1981; Shippen-

Lentz & Blackburn, 1990). When the G-overhang is bound by telomerase (Fig 1-3A), each nucleotide in the repeat is sequentially added, one at a time, by copying the template of TER (Fig 1-3B, C). After the last nucleotide in the template is copied, a translocation step is required for telomerase to further elongate the substrate (Greider & Blackburn, 1985) (Fig 1-3D).

Telomerase plays a crucial role in telomere replication: telomerase mutants in budding yeast display an “ever shorter telomere” (*est*) defect as well as decreased growth rate and cell viability (Lendvay et al, 1996; Lundblad & Szostak, 1989; Singer & Gottschling, 1994). Inactivation of telomerase in human cells is also deleterious, leading to telomere shortening, as well as senescence and cell death (Feng et al, 1995; Mitchell et al, 1999). Mice lacking telomerase can survive up to six generations, displaying progressive shortening of telomeres (~ 2 to 7 kb per generation). Late generations of mouse mutants exhibit defective reproductive organs and cells with severe genome instability, including aneuploidy and end-to-end fusions (Blasco et al, 1997; Lee et al, 1998). Similarly, telomeres in *Arabidopsis tert* mutants are shortened progressively and the mutants can live for up to ten generations (Riha etal, 2001) (see below).

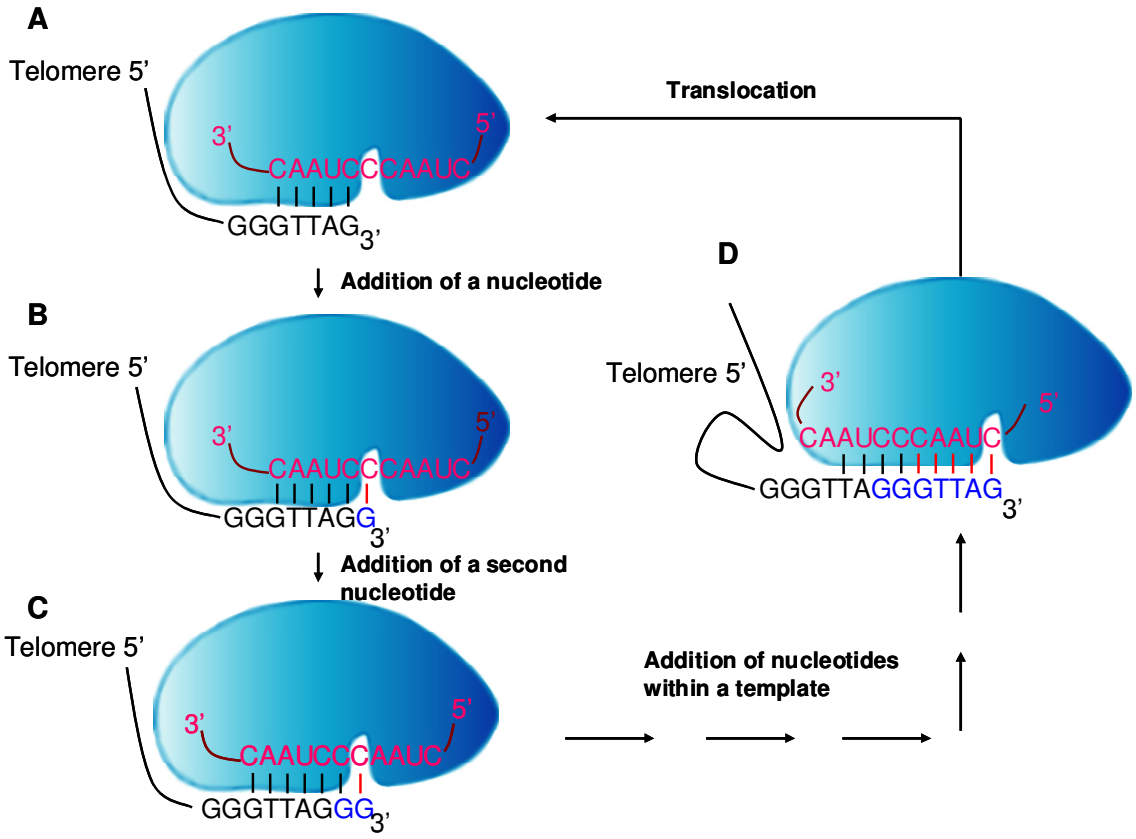


Fig 1-3. Telomere elongation by telomerase. Schematic diagram of telomerase action at the 3' end of a telomere. The catalytic subunit of telomerase TERT and the telomerase RNA TER (red) are shown. (**A**) The 3' end of a telomere is base-paired with the template region of TER. (**B**)-(**C**) Nucleotides (blue) are added to the 3' end of a telomere, one at a time, at the active site of telomerase. Once a nucleotide is added, the template moves one position, allowing the next residue fallen into the active site of the enzyme. (**D**) When the end of the template is reached, telomerase translocates and aligns with the newly synthesized 3' terminus, leading to processive addition of telomeric repeats at chromosome ends.

Telomerase expression is restricted to proliferative tissues, including fetal, new born and adult testes and ovaries. In contrast, telomerase activity is undetectable in most somatic cells (Wright et al, 1996), where telomeres become shortened each time when cells divide. Once telomeres reach a critical length, cells will stop dividing leading to senescence or cell death (Harley et al, 1990). Therefore, telomerase repression and telomere shortening is linked to cell aging. On the other hand, almost all cancer cells maintain telomere length, and strikingly, approximately 90% of them show up-regulated telomerase activity (Kim et al, 1994). Thus, telomerase has become an attractive target for cancer therapeutics (Harley, 2008). Several anti-telomerase drugs are in clinical trials (Harley, 2008).

Even in proliferative cells, telomerase is expressed at an extremely low level. Studies show that there are only five to six molecules of telomerase per cell (Cohen et al, 2007). Moreover, telomerase does not extend every telomere each cell cycle. Hemann et al first reported that telomerase preferentially extends short telomeres in mammalian cells (Hemann et al, 2001). Similar phenomena were reported in plants and in yeasts (Shakirov & Shippen, 2004; Teixeira et al, 2004). It is proposed that telomeres switch between telomerase-extendable (short telomere) and non-extendable states (long telomere), allowing establishment of telomere length homeostasis (Teixeira et al, 2004).

Telomerase-associated components

While the telomerase catalytic core consists of TERT and TER (Cohen et al, 2007), other components are associated with telomerase *in vivo*. In budding yeast, Est1 and Est3 are required for proper telomerase action *in vivo* (Lendvay et al, 1996; Lundblad & Szostak, 1989), although neither protein affects *in vitro* telomerase activity (Cohn & Blackburn, 1995; Lingner et al, 1997a).

Est1 associates with telomerase by binding TER (Seto et al, 2002). Its primary role is to recruit telomerase to the chromosome end, through an interaction with the G-overhang binding protein Cdc13 (Pennock et al, 2001) (see below). Supporting this model, fusion of the DNA binding domain of Cdc13 to Est2 (TERT) fully rescues the telomere replication defect in *est1* mutants (Pennock et al, 2001). Moreover, while *cdc13-2* (E252K) mutants and *est1-60* (K444E) mutants both show an *est* phenotype, telomere length maintenance can be fully restored in *cdc13-2 est1-60* double mutants, indicating a specific interaction site between Cdc13 and Est1 (Pennock et al, 2001).

Est3 is a stable component of the telomerase holoenzyme (Hughesa et al, 2000). While the function of Est3 is unknown, a recent study showed that Est3 harbors an oligonucleotide/oligosaccharide-binding fold (OB-fold) similar to that of the human Shelterin component, TPP1 (Yu et al, 2008), which is also involved in telomerase regulation (see below). Therefore, Est3 may be conserved in different eukaryotes.

The human TER is associated with Dyskerin, which functions in rRNA maturation, and is required for telomerase RNP biogenesis and enzyme function *in vivo* (Mitchell et al, 1999). Missense mutations in Dyskerin result in the genetic disorder Dyskeratosis Congenita, a complex syndrome characterized by abnormal skin pigmentation, bone marrow failure, telomerase enzyme deficiency, and progressive telomere shortening

(Heiss et al, 1998; Mitchell et al, 1999). Altogether, telomerase and its associated factors are crucial to maintain the terminal chromosomal sequences.

Telomere-associated proteins

Telomere-associated proteins can either bind ds or ss telomeric DNA or associate with the chromosome termini via protein-protein interactions. Well-characterized ds telomeric DNA binding proteins include Rap1 in budding yeast (Conrad et al, 1990), Taz1 in fission yeast (Cooper et al, 1997), and TRF1 and TRF2 in mammals (Broccoli et al, 1997). These proteins recognize and bind ds telomeric DNA through Myb-like motif(s). The Myb-like telomere binding motif shows similarity to the third repeat of human c-Myb, but displays higher specificity towards telomeric DNA than common Myb substrates (Bilaud et al, 1996).

The ss telomeric binding proteins have also been extensively studied, including Telomere End Binding Protein (TEBP) from a ciliate *Oxytricha nova* (Gottschling & Zakian, 1986; Price & Cech, 1987), Cdc13 from budding yeast (Garvik et al, 1995; Lin & Zakian, 1996), and Protection Of Telomeres 1 (Pot1) from fission yeast and vertebrates (Baumann & Cech, 2001). Each of these proteins binds G-strand overhangs, and is characterized by one or more OB-folds. The OB-fold is a common protein domain which consists of a five beta stranded barrel structure. OB-fold containing proteins are often involved in the recognition of single-stranded nucleic acids, including rRNA (e.g. ribosomal proteins), tRNA (e.g. class IIb tRNA synthetase), ss DNA (e.g. ssDNA-binding protein, SSB; Replication Protein A, or RPA), and telomere G-overhang (e.g. Cdc13, POT1) (Theobald et al, 2003). Interestingly, the recently characterized OB-fold containing proteins in *C. elegans*, CeOB1 and CeOB2, show specificity to G-strand and

C-strand telomeric DNA, respectively (Raices et al, 2008). Besides ds and ss telomere binding proteins, there are also other telomere-associated proteins that are recruited to telomeres via protein-protein interactions. Such proteins include Rif1 and Rif2 in budding yeast, and RAP1/TIN2/TPP1 in vertebrates (see below).

The major functions of these telomere-associated proteins are two fold. First, they regulate access of telomerase and other enzymes at telomeres and therefore control telomere length. Second, they play essential roles in chromosome end protection, which is often referred to as “telomere capping”. In the following section, the composition and function of telomere-associated proteins in budding yeast and vertebrates are described in detail.

The telomere capping function of Cdc13/Stn1/Ten1 in budding yeast

In budding yeast, the telomeric protein Cdc13 plays a multifunctional role at telomeres (Fig 1-4, middle). Cdc13 binds the G-overhang through a single OB-fold (Lin & Zakian, 1996; Nugent et al, 1996). Cdc13 serves as the platform to deliver two different complexes to telomeres: Stn1 (Suppressor of cdc thirteen, 1) and Ten1 (Telomeric pathways in association with Stn1, 1) to protect chromosome ends, as well as Est1/telomerase to maintain telomere length (Nugent et al, 1996; Pennock et al, 2001) (Fig 1-5).

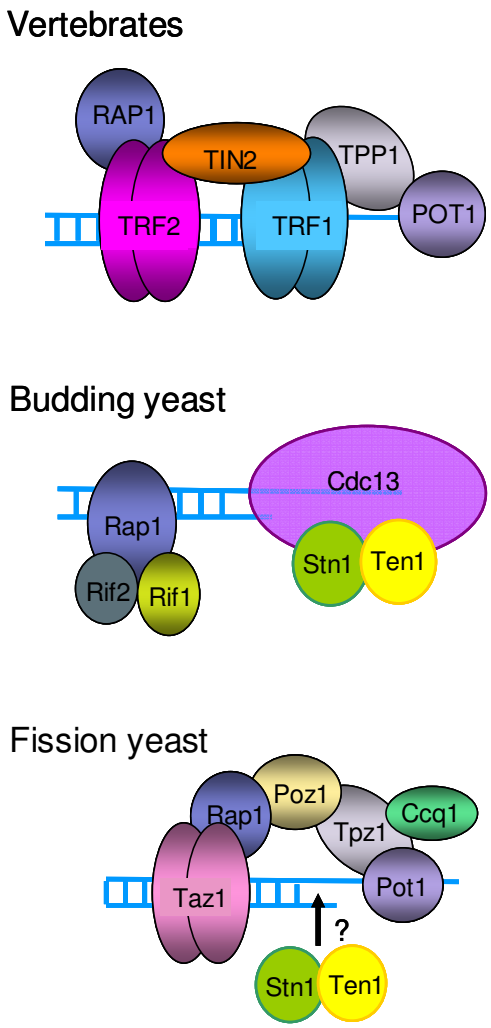


Fig 1-4. Telomere-associated proteins in vertebrates, budding yeast and fission yeast. Top, vertebrate telomeres are protected by the six-member Shelterin complex including TRF1, TRF2, RAP1, TIN2, TPP1 and POT1. Middle, budding yeast telomeres are capped by the Cdc13/Stn1/Ten1 (CST) complex. Rap1 binds to ds telomeric DNA and associates with Rif1 and Rif2 at telomeres. Bottom, telomere proteins in fission yeast consist of the Shelterin components, including Taz1 (TRF1/2 ortholog), Rap1, Poz1, Ccq1, Tpz1 (TPP1 homolog) and Pot1, and the CST components Stn1 and Ten1.

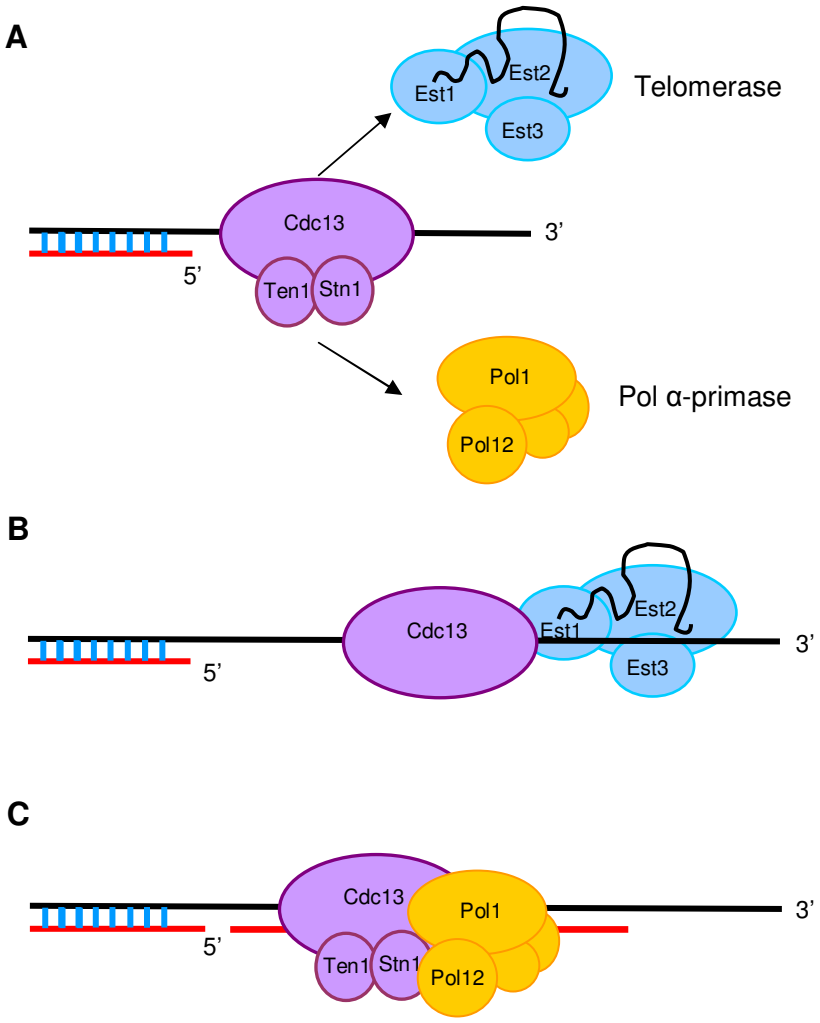


Fig 1-5. Budding yeast CST coordinates the actions of telomerase and polymerase α-primase complex to replicate telomeres. (**A**) A diagram of CST interaction with telomerase and polymerase α- primase complex. CST components are labeled in purple; telomerase components are shown in blue; and polymerase α- primase complex is shown in yellow. (**B**) By interacting with Est1, Cdc13 recruits the telomerase RNP to the chromosome end to extend the G-strand of telomeric DNA. (**C**) Cdc13 and Stn1 also contact with polymerase α- primase complex to facilitate replication of the telomeric C-strand.

Like Cdc13, Stn1 and Ten1 are also OB-fold containing proteins. It is proposed that Cdc13/Stn1/Ten1 (CST) forms a heterotrimeric complex similar to RPA (Gao et al, 2007). CST is essential for cell viability and chromosome end protection. Loss of function in either Cdc13 (*cdc13-1*) or Stn1 or Ten1 exposes the C-strand to extensive resection, which results in extremely long ss G-strand telomeric DNa and a Rad9 mediated cycle arrest at G2/M (Grandin et al, 2001; Grandin et al, 1997; Nugent et al, 1996). Interestingly, overexpression of Stn1 is sufficient to rescue the lethality of *cdc13-1* mutants (Grandin et al, 1997). In addition, *cdc13-1* lethality can also be rescued when Stn1 is ectopically delivered to telomeres by fusing Stn1 to the DNA binding domain of Cdc13 (Pennock et al, 2001). Therefore, Cdc13's primary role in end protection appears to be to deliver Stn1 to telomeres.

CST components have been found in budding yeast and its closely-related species. Only recently have Stn1 and Ten1 been identified in fission yeast (Martin et al, 2007). In Chapter II, I report that the identification and characterization of a STN1 ortholog in *Arabidopsis*. Furthermore, a novel telomere capping protein CTC1 (Chapter III) and a TEN1 homolog (Chapter IV) were identified and characterized in *Arabidopsis*. CTC1 exhibits many properties similar to budding yeast Cdc13. Excitingly, human CTC1/STN1/TEN1 complex was independently identified by the Ishikawa group (Miyake et al, 2009). Thus, CST complex is more conserved than previously expected.

Cdc13 recruits telomerase to telomere ends and plays a dual role in telomere length regulation

Cdc13 also plays an important role in telomere length regulation. On one hand, Cdc13 positively promotes telomere replication. A separation-of-function mutant of Cdc13 (*cdc13-2*) shows an *est* defect. Cdc13 interacts with Est1, through which Cdc13 recruits telomerase to telomeres (Pennock et al, 2001) (Fig 1-5B). On the other hand, Cdc13 is also a negative regulator of telomere length. A separation-of-function mutant of Cdc13 (*cdc13-5*) displays extensively elongated telomeres (Chandra et al, 2001), yet the mechanism of this negative regulation is unknown.

The interaction of CST with polymerase α- primase complex

Besides recruiting telomerase to replicate the G-strand telomeric DNA, budding yeast Cdc13 also interacts with the catalytic subunit of polymerase α (Pol 1) for C-strand telomere synthesis (Qi & Zakian, 2000). The CST-polymerase α interaction is strengthened by Stn1, which associates with the regulatory subunit of polymerase α (Pol 12) (Grossi et al, 2004) (Fig 1-5C). Point mutations in either CST or Polymerase α that affect the interaction lead to elevated ss G-strand signal as well as a modest increase of telomere length (Grossi et al, 2004). In summary, CST delivers both telomerase and polymerase α to telomeres, and hence coordinates the G- and C-strand telomere synthesis (Fig 1-5).

Other telomere-associated proteins in budding yeast

In budding yeast, ds telomeric DNA is bound by Rap1 (repressor/activator protein 1) (Conrad et al, 1990) (Fig 1-4, Middle). Rap1 was originally implicated in transcription regulation, acting as a transcriptional repressor or activator depending on the element it

binds to (Shore & Nasmyth, 1987). Rap1 is essential and a null mutant is lethal. Overexpression of a dominant negative allele of Rap1 leads to elongated telomeres, indicating Rap1 is a negative regulator of telomere length (Conrad et al, 1990). Further studies indicate that Rap1 recruits two other proteins, Rif1 and Rif2 (Rap1-interacting factors 1 and 2) to promote telomere length regulation (Hardy et al, 1992; Wotton & Shore, 1997) (Fig 1-4, Middle). Loss of Rif1 or Rif2 alone results in moderate telomere elongation, while *rif1 rif2* double mutants exhibit dramatically elongated telomeres (Wotton & Shore, 1997). Thus, both Rif1 and Rif2, together with Rap1, contribute to telomere length control in budding yeast.

Shelterin complex in mammals

Mammalian telomeres are protected by a six-member complex, called Shelterin (Palm & de Lange, 2008). Shelterin recognizes and associates with telomeres through two ds telomere binding proteins, Telomeric Repeat binding Factor 1 and 2 (TRF1 and TRF2), as well as the ss telomere binding protein, POT1 (Liu et al, 2004a; Ye et al, 2004b) (Fig 1-4, top). TRF1 and TRF2 recruit the other three Shelterin components to telomeres: the TRF2- and TRF1-Interacting Nuclear protein 2 (TIN2), Repressor/Activator Protein 1 (RAP1), and TPP1 (also known as ACD, TINT1, PTOP, or PIP1) which interacts with POT1 as well (Liu et al, 2004a; Ye et al, 2004b) (Fig 1-4, top). Recent studies suggest that some Shelterin components are conserved in fission yeast, including Taz1 (a TRF1/TRF2 homolog), Rap1, Pot1 and Tpz1 (a TPP1 homolog) (Baumann & Cech, 2001; Chikashige & Hiraoka, 2001; Ferreira & Cooper, 2001; Kanoh & Ishikawa, 2001; Miyoshi et al, 2008) (Fig 1-4, bottom).

Both TRF1 and TRF2 possess Myb domains to bind ds telomeric DNA. They associate with telomeres as homodimers or oligomers, (Broccoli et al, 1997). TRF1 and TRF2 do not interact with each other directly. TRF1 acts as a negative regulator of telomere length. Inhibition of TRF1 leads to telomere elongation, whereas overexpression of TRF1 results in telomere shortening (van Steensel & de Lange, 1997). Like TRF1, TRF2 also negatively regulates telomere length (Smogorzewska et al, 2000). It is proposed that TRF1 and TRF2 "measure" telomere length *in vivo*, through interactions with TIN2, TPP1 and POT1 (Loayza & de Lange, 2003). In addition, TRF2 is crucial for telomere protection. Expression of a dominant negative allele of TRF2 in human cells, or conditional deletion of TRF2 in mouse embryo fibroblasts results in genome-wide chromosome end fusions, a strong DNA damage response and p53-dependent senescence (Celli & de Lange, 2005; Denchi & de Lange, 2007; Karlseder et al, 1999; van Steensel et al, 1998).

POT1 was originally identified through its sequence similarity to the α subunit of the TEBPα/β telomeric binding complex in *Oxytricha nova* (Baumann & Cech, 2001). Like TEBPα, POT1 contains two OB-folds in the N-terminus that allow it to recognize the ss G-strand telomeric sequence (Baumann & Cech, 2001; Lei et al, 2004). POT1 is crucial for telomere length homeostasis *in vivo*, serving as a terminal transducer for telomere length control (Loayza & de Lange, 2003). POT1 is also required for chromosome end protection. Knockdown of hPOT1 results in a reduced G-overhang signal, as well as a DNA damage response and a modest level of chromosome end fusions (Veldman et al, 2004; Yang et al, 2005). While humans and fission yeast have only one *POT1* gene, mouse harbors *two* POT1 genes, mPOT1a and mPOT1b. Both mPOT1a and mPOT1b are required to prevent DNA damage responses at telomeres as

well as cell senescence, although minor functional differences do exist between the two paralogs (Hockemeyer et al, 2006; Wu et al, 2006).

Recent studies indicate the presence of a third OB-fold in the C-terminus of POT1 (Theobald & Wuttke, 2004), which is involved in TPP1 interactions. TPP1 brings POT1 to telomeres by contacting TIN2 in the Shelterin complex (Liu et al, 2004b; Ye et al, 2004b). TPP1 contains an OB-fold that shows considerable similarity to that of the TEBPβ subunit in *O. nova* (Wang et al, 2007; Xin et al, 2007). Altogether, POT1/TPP1 appears to be a conserved heterodimeric complex that resembles TEBPα/β telomeric binding complex. Like POT1, TPP1 negatively regulates telomere length (Liu et al, 2004b; Ye et al, 2004b). On the other hand, POT1/TPP1 acts as a telomerase processivity factor *in vitro* (Wang et al, 2007). Thus, POT1/TPP1 can switch from a telomerase inhibitor to a stimulator of telomerase activity and processivity during telomere extension (Wang et al, 2007).

TIN2 is a central bridging protein in the Shelterin complex. It contacts TRF1, TRF2 and TPP1 (Kim et al, 1999; Liu et al, 2004a; Ye et al, 2004b). *In vivo* studies suggest that TIN2 helps to stabilize TRF2 on telomeres (Ye et al, 2004a). Human RAP1 is a homolog of yeast Rap1 protein. Unlike yeast Rap1, mammalian RAP1 does not directly bind to telomeric DNA (Li et al, 2000). Rather, its association with telomeres is through TRF2 interaction (Li et al, 2000). RAP1 participates in the regulation of telomere length and heterogeneity (Li & de Lange, 2003; Li et al, 2000).

In summary, the Shelterin components participate in telomere protection in vetebrates. The recent discovery of CTC1/STN1/TEN1 in *Arabidopsis* (Chapters II-IV) and in humans indicates that the telomere-associated protein network may be more complicated than previously thought.

Telomeric repeat-containing RNA (TERRA)

The telomeric DNA and telomere-associated proteins have been extensively studied during the last three decades. In addition to these components, recent data suggest that a large noncoding RNA, called telomeric repeat-containing RNA (TERRA), is expressed and specifically localized at telomeres in yeasts and mammals (Azzalin et al, 2007; Luke & Lingner, 2009; Luke et al, 2008; Schoeftner & Blasco, 2008). Earlier studies show evidence of transcriptions at telomeres in trypanosomes (Rudenko & Van der Ploeg, 1989) and in birds (Solovei et al, 1994), indicating TERRA may be conserved in different eukaryotes.

Human TERRA contains repeats of UUAGGG, and is primarily transcribed by RNA Polymerase II in a unidirectional way (Luke et al, 2008; Schoeftner & Blasco, 2008). TERRA is extremely heterogeneous, ranging from 100 bp to ~ 9 kb based on Northern blot analysis (Azzalin et al, 2007). The 3' end of TERRA is polyadenylated (Azzalin et al, 2007; Luke et al, 2008), while it remains unclear how its 5' end is modified. The association of TERRA with telomeres is mediated in two ways: 1) TERRA can hybridize with telomeric DNA (Luke et al, 2008); 2) TERRA are associated with Shelterin components TRF1 and TRF2 (Deng et al, 2009).

TERRA RNA also contacts subunits of the origin recognition complex (ORC), heterochromatin marks H3K9me3 and heterochromatin protein 1 (HP1), as well as DNA damage pathway components (Deng et al, 2009). Knockdown of TERRA causes increased telomere dysfunction-induced foci (TIF), aberrant telomere defects at metaphase, and a loss of H3K9me3 and ORC at telomeres (Deng et al, 2009). TERRA also blocks telomerase activity *in vitro*, suggesting that TERRA may be involved in telomerase regulation at chromosome ends (Schoeftner & Blasco, 2008). Thus, TERRA

is a novel component of telomere structure, which plays a critical role in telomere maintenance and heterochromatin formation.

Telomere length homeostasis

Telomere length varies dramatically among different eukaryotes. Telomeres are only about 350 bp long in budding yeast (Teixeira et al, 2004). In contrast, they can reach to 5-15 kb in human (Lansdorp et al, 1996), and about 20-150 kb in mice (Zijlmans et al, 1997). Telomere length also varies in different plants. For instance, *Arabidopsis* harbors telomeres about 2-8 kb (Richards & Ausubel, 1988; Shakirov & Shippen, 2004), whereas telomeres in tobacco can be as long as 60-160 kb (Fajkus et al, 1995). Despite the telomere length difference, it is essential for all eukaryotes to establish and maintain telomere length homeostasis. Neither short, nor long telomeres are favorable. Critically shortened telomeres initiate a DNA damage checkpoint response, which then mediates senescence and cell death. Aberrantly elongated telomeres can also be detrimental. A *K. lactis* mutant that carries extremely long telomeres encounters reduced cell viability (McEachern & Blackburn, 1995).

It is proposed that telomere length homeostasis is established through a series of telomere shortening and lengthening events. Several pathways are involved in telomere shortening. First, in the absence of telomerase, the end replication problem accounts for progressively shortened telomeres. Second, nucleolytic attack also contributes to telomere erosion (Verdun & Karlseder, 2007). Third, telomere binding proteins, including TRF1 and TRF2 and others, act as negative regulators for telomere length (Smogorzewska et al, 2000; van Steensel & de Lange, 1997). Finally, homologous recombination at a t-loop can generate a shortened telomere as well as extrachromosomal telomeric circles (ECTC), leading to telomere rapid deletion (TRD) (Fig 1-6).

TRD stochastically shortens otherwise long telomeres back to normal range (Bucholc et al, 2001; Li & Lustig, 1996). It is proposed that TRD occurs when branch migration happens at the t-loop in human cells, followed by holiday junction resolution and cleavage (Lustig, 2003) (Fig 1-6). In budding yeast, where telomeres form a fold-back structure instead of a t-loop, it is proposed that a transient t-loop structure forms prior to homologous recombination and TRD (Lustig, 2003). TRD requires Rad52 and Mre11/Rad50/Xrs2 in yeast for homologous recombination (Bucholc et al, 2001). While originally discovered in yeast (Bucholc et al, 2001; Li & Lustig, 1996), TRD is also observed in humans and plants (Wang et al, 2004; Watson & Shippen, 2007; Zellinger et al, 2007). Remarkably, ECTC can be detected in wild-type human and plant cells (Wang et al, 2004; Zellinger et al, 2007), arguing that homologous recombination is an integral part of telomere length regulation.

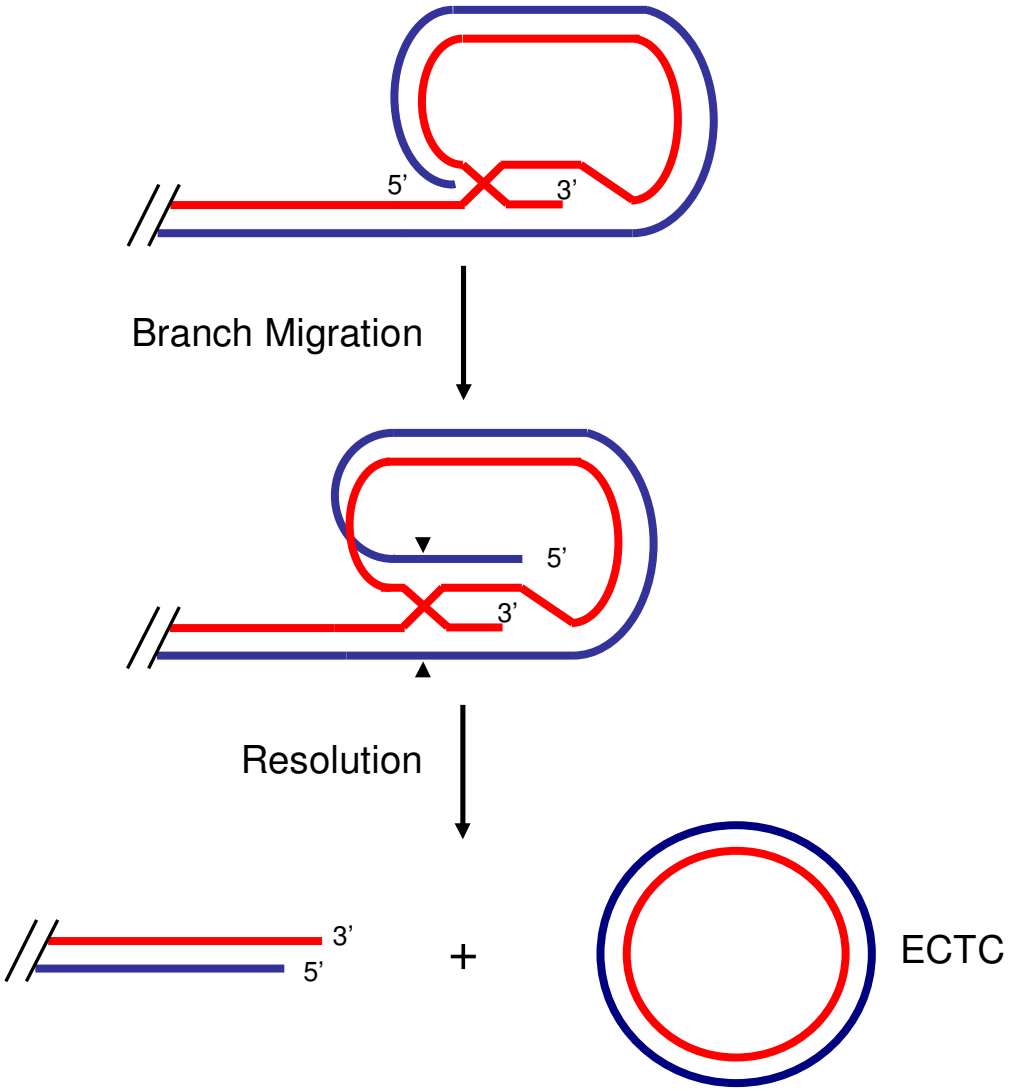


Fig 1-6. Telomere Rapid Deletion (TRD). When branch migration occurs at a t-loop, the resulting intermediate structure resembles a Holliday junction intermediate which will then be resolved (at sites indicated by the arrows). Cleavage leads to generation of a shortened telomere and ECTC.

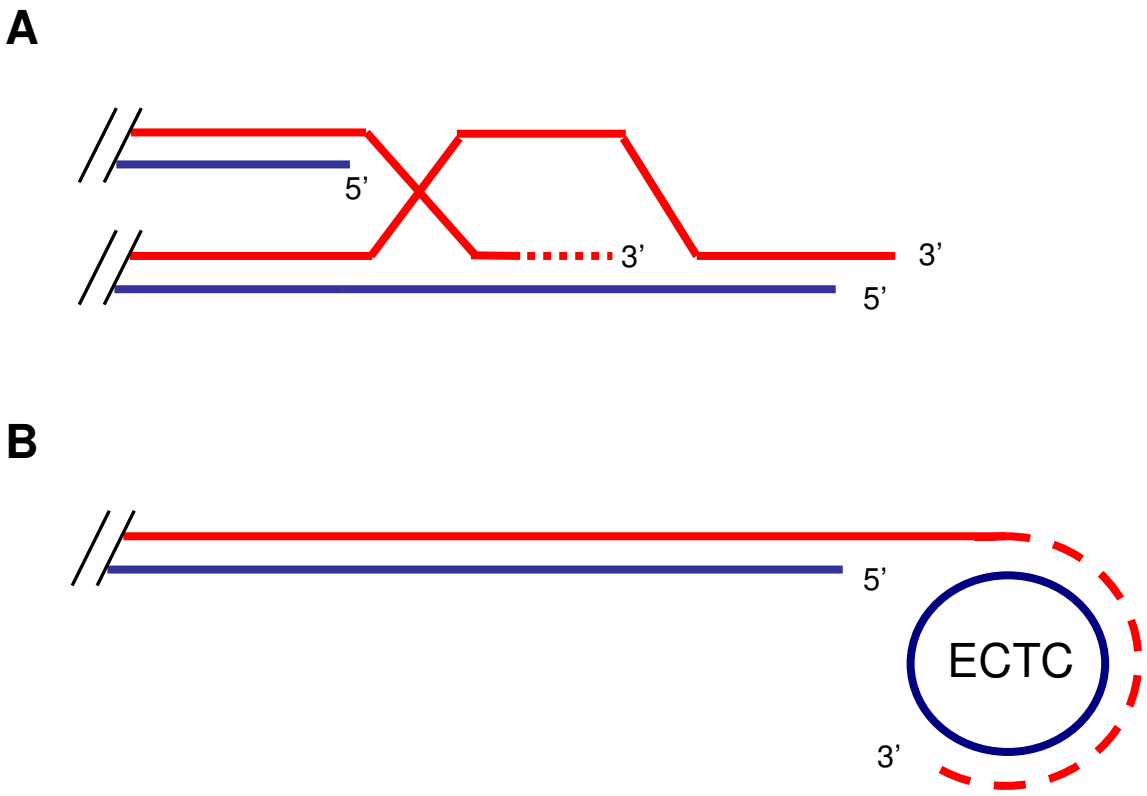


Fig 1-7. Strategies for Alternative Telomere Lengthening (ALT). (**A**) In *Saccharomyces cerevisiae* telomerase mutant survivors, telomeric sequences can be maintained in a manner where one telomere uses another as a template for extension. (**B**) Schematics of rolling-circle amplification to extend telomeres. The 3' end of a telomere is extended using an ECTC as a template.

Telomeres are primarily maintained by telomerase. In its absence, Alternative Telomere Lengthening (ALT) contributes to telomere elongation. The first example of ALT was reported in budding yeast telomerase mutants (*est1* mutants), which lose telomeric DNA gradually through generations. As expected, the majority of telomerase mutants eventually die. However, survivors were discovered that escaped the lethal consequence of telomerase defects (Lundblad & Blackburn, 1993). There are two classes of survivors, termed type I and type II. Type I survivors carry short telomeric DNA tracts, but their subtelomeric repeats (called Y' elements) are extensively amplified (Lundblad & Blackburn, 1993). Type II survivors maintain long and heterogeneous telomeric repeats where one telomere uses another telomere as a template for extension (Teng & Zakian, 1999) (Fig 1-7A). Both types of survivors are found to be dependent on Rad52, a key component of homologous recombination (Lundblad & Blackburn, 1993). Therefore, a homologous recombination-based mechanism is proposed for the elongation of telomeres in the absence of telomerase (Lundblad, 2002).

Although the majority of cancer cells upregulate telomerase, about 10% human cancer cells engage ALT to maintain telomere length (Muntoni & Reddel, 2005; Verdun & Karlseder, 2007). While the exact mechanism of ALT remains unknown, human ALT cells contain abundant ECTC (Cesare & Griffith, 2004; Wang et al, 2004). ALT also depends on the presence of MRE11/RAD50/NBS1 (Muntoni & Reddel, 2005). Therefore, homologous recombination at telomeres facilitates ALT in human cells. Consistently, ECTC have been found in *K. lactis* telomerase mutant, where telomeres are extremely elongated (Underwood et al, 2004). It is proposed that ECTC can serve as a template that allows rolling-circle amplification of telomeres (de Lange, 2004) (Fig

1-7B). To sum up, cells can utilize a combination of different telomere shortening and elongating pathways to maintain telomere length homeostasis.

Telomeres and DNA damage repair machinery

An essential role of telomeres is to distinguish themselves from double-strand breaks and to prevent action of a DNA damage response. Ironically, many DNA damage repair proteins are localized to telomeres, including MRN or MRX complex (MRE11/RAD50/NBS1 in humans or Mre11/Rad50/Xrs2 in yeast), protein signaling kinases of ATM (ataxia-telangiectasia mutated) and ATR (ATM and Rad3-related) in humans and their counterparts in budding yeast (Tel1 and Mec1), as well as KU70/KU80 heterodimer, a key player of non-homologous end joining pathway (NHEJ). Accumulating evidence reveals that these DNA damage repair proteins play critical roles in not only dysfunctional telomeres processing, but also in normal telomere length regulation and chromosome end protection.

Loss of Mre11 or Rad50 or Mec1 in budding yeast results in gradual telomere shortening (Nugent et al, 1998; Takata et al, 2004). Epistasis analysis confirms that Mre11/Rad50 regulate telomere length through the telomerase pathway (Nugent et al, 1998). Further studies indicate that MRX localizes to telomeres during late S-phase and recruits Mec1, which in turn contributes to active telomerase assembly at telomeres (Takata et al, 2004; Takata et al, 2005). Tel1, on the other hand, protects telomeres from catastrophic telomere loss and end-to-end fusions (Chan & Blackburn, 2003). Similar to budding yeast, human MRN and ATM also localize to telomeres, and prevent aberrant telomere shortening or fusion events (Verdun & Karlseder, 2007). Shelterin components regulate the functions of ATM and ATR at telomeres. The robust DNA

damage signal initiated by TRF2 loss can be rescued by simultaneous deletion of ATM, but not ATR. In contrast, TIF generated by POT1 deficiency activates ATR, but not ATM (Denchi & de Lange, 2007). Therefore, TRF2 represses ATM, presumably through a direct interaction between TRF2 and ATM, while POT1 prevents ATR pathway at telomeres.

KU70/KU80 is implicated in NHEJ repair pathway, which binds and stabilizes the ends of double-strand breaks (Boulton & Jackson, 1996a; Boulton & Jackson, 1996b). Deletion of KU in budding and fission yeast leads to shortened telomere tracts (Baumann & Cech, 2000; Gravel et al, 1998). Therefore, KU70/KU80 positively regulates telomere length in yeasts. Consistently, budding yeast KU70/KU80 interacts with TER and is involved in telomerase recruitment (Fisher et al, 2004). In contrast, telomere tracts are grossly elongated in *Arabidopsis ku70* mutants (Riha et al, 2002). Thus, KU acts as a positive regulator of telomere length in yeast and a negative regulator in *Arabidopsis*. However, in both settings, G-overhangs are dramatically extended in a telomerase-independent manner (Gravel et al, 1998; Riha & Shippen, 2003), indicating that KU is required to maintain proper telomere architecture. Altogether, these data indicate that DNA damage repair proteins are actively involved in the regulation of telomere length and protection of chromosome end architecture. The roles of ATM, ATR and KU70 in plants lacking CST components will be investigated in Chapter V.

Arabidopsis as a model to study telomere biology

Arabidopsis thaliana is a model plant with a small genome that has already been sequenced (~ 125 Mb). This multicellular eukaryote has a growth period of less than 6 weeks and is genetically tractable. *Arabidopsis* can be easily manipulated for crosses and *Agrobacterium*-based transformations. Moreover, a large collection of T-DNA insertion lines, activation tagging lines and EMS-mutagenized lines are available, making *Arabidopsis* a powerful model organism for genetic studies.

With respect to telomere biology, *Arabidopsis* shows many advantageous features. As mentioned above, telomere sequence and structure are conserved among yeasts, plants and vertebrates. In addition, the compositions of telomere-associated proteins are quite similar between *Arabidopsis* and other organisms (see below). Moreover, similar to the case in vertebrates, telomerase expression is tightly regulated in *Arabidopsis*. Telomerase is active in tissues rich in dividing cells such as flowers, seedlings and cell cultures, and is suppressed in vegetative tissues including leaves and stems (Fitzgerald et al, 1999). Thus, *Arabidopsis* shares many conserved features in terms of telomere biology.

Arabidopsis also has some unique features. The telomere length of *Arabidopsis* ranges from 2 to 5 kb in many ecotypes. The small size of telomeres facilitates accurate measurement of telomere length by Terminal Restriction Fragment (TRF) analysis (Shakirov & Shippen, 2004). Furthermore, unique subtelomeric sequences on 8 out of 10 chromosome arms provide an opportunity to examine telomere length on individual chromosome arms and to characterize the architecture of telomere fusion junctions (Heacock et al, 2004). Finally, *Arabidopsis* is remarkably tolerant to telomere dysfunction. The telomerase mutant, *tert*, can survive up to ten generations with the last

five generations displaying increasing levels of anaphase bridges and genome instability (Riha et al, 2001). Furthermore, mutations in many telomere-related genes that are lethal in mammals are viable in *Arabidopsis*. Such genes include *ATM* (Vespa et al, 2005), *ATR* (Culligan et al, 2004), *MRE11* (Bundock & Hooykaas, 2002) and *RAD50* (Gallego & White, 2001). Altogether, *Arabidopsis* has become a great model organism for telomere biology studies. Knowledge gained from *Arabidopsis* telomeres contributes to elucidate the composition of telomere-associated factors, and the mechanisms governing telomerase regulation, telomere length control and chromosome end protection.

Arabidopsis **telomerase**

The telomerase reverse transcriptase, AtTERT, was identified based on its sequence similarity to human TERT (Fitzgerald et al, 1999). Similar to humans, where TERT is only expressed in highly proliferative tissues such as stem cells and germ line cells (Wright et al, 1996), *Arabidopsis* telomerase activity is restricted to flowers, root-tips, seedlings and undifferentiated callus tissue (Fitzgerald et al, 1999). In plants lacking TERT, telomeres shorten steadily by approximately 200-500 bp per generation (Riha et al, 2001). *Arabidopsis tert* mutants can survive up to ten generations. They are morphologically indistinguishable from WT plants through the fifth generation (G5). From G6 onward, the size of leaves is substantially reduced, many of which are asymmetric and lobed. Defective anthers and reduced pollen grains begin to occur from G7 onwards. The mutants finally reach a terminal generation and arrest at a vegetative state in G9/G10 (Riha et al, 2001). The seed yields of *tert* mutants decline progressively from G6/G7. The germination rate of seeds reduces to 90% in G7, and the number

drops to 15% in G9. Anaphase bridges, which are hall marks of genome instability, occur in G5 or G6 and worsen in successive generations (Riha et al, 2001). In summary, plants lacking *TERT* display progressive telomere shortening, and only in late generations of *tert* mutants do we observe severe developmental defects and genome instability.

Unexpectedly, two TERs (TER1 and TER2) are present in *Arabidopsis* (Cifuentes-Rosias, C. and D.E. Shippen, unpublished work). Both TERs can serve as template for TERT *in vitro*, the RNAs have different functions *in vivo*. TER1 acts as the major template for telomerase, whereas TER2 negatively regulates telomerase activity. Besides TERT and TERs, *Arabidopsis* telomerase also contains Dyskerin, which contributes to maximal telomerase activity *in vivo* (Kannan et al, 2008). In addition, AtPOT1a physically associates with the telomerase RNP and is required for telomerase action *in vivo*. POT1a enriches at telomeres only during S-phase, when telomerase is thought to act at telomeres (Surovtseva et al, 2007). Like *tert* mutants, plants lacking *POT1a* also display progressive telomere shortening. Moreover, *pot1a* mutants show variable but significantly reduced telomerase activity *in vitro*. Recent data reveal that POT1a interacts with TER1 and physically associates with telomerase RNP (Surovtseva et al, 2007; Cifuentes-Rosias, C. and D.E. Shippen, unpublished work). Altogether, current data suggest that AtPOT1a stabilizes the telomerase complex and possibly regulates telomerase recruitment to telomeres.

Arabidopsis telomere-associated proteins

Because the sequence and structure of telomeres in plants are similar to those in vertebrates, it is assumed that Shelterin components are present at plant telomeres. Indeed, *Arabidopsis* harbors at least six Myb-bearing proteins that bind ds telomeric DNA *in vitro* in a manner similar to vertebrate TRF1 and TRF2 (Karamysheva et al, 2004). Furthermore, recent data suggest that at least one of these TRF-like proteins, AtTBP1, acts as a negative regulator of telomere length (Hwang & Cho, 2007). Similarly, rice mutants lacking RTBP1 display gradual telomere lengthening and exhibit telomere fusions in G2 (Hong et al, 2007).

Additionally, *Arabidopsis* encodes three OB-fold bearing POT1-like proteins POT1a, POT1b and POT1c (Shakirov et al, 2005; Surovtseva et al, 2007; Nelson, A.L.D. and D.E. Shippen, unpublished data). *Arabidopsis* POT1a and POT1b, like their homologs in humans and fission yeast, harboring two N-terminal OB-folds (OB1 and OB2) followed by a C-terminal extension. In contrast, POT1c encodes a small protein corresponding to a single OB-fold, which shows characteristics of both OB1 and OB2 of POT1a. Unlike the rodent POT1 paralogs that are 75% similar to each other, AtPOT1a and AtPOT1b are highly divergent, which share only 49% protein sequence similarity. Surprisingly, none of the *Arabidopsis* POT1 paralogs binds ss telomeric DNA *in vitro*, suggesting unusual roles of *Arabidopsis* POT1 proteins *in vivo*.

While the function of AtPOT1b and AtPOT1c is still under investigation, over-expression of a dominant negative allele of AtPOT1b or depletion of AtPOT1c lead to a telomere uncapping phenotype similar to a *pot1* deficiency in yeast and mammals (Shakirov et al, 2005; Nelson, A.L.D. and D.E. Shippen unpublished work). In contrast, AtPOT1a is dispensable for chromosome end protection and instead is required for

telomerase function (Surovtseva et al, 2007). AtPOT1a has evolved a special function as a positive regulator of telomere length and appears to be a novel component of the telomerase RNP complex. These characteristics of POT1a will be further discussed in Chapters VI and VII. Currently, orthologs for TIN2, RAP1 and TPP1 cannot be discerned in any plant genome, possibly due to the rapid evolution of these telomere genes.

Gene duplication

As mentioned above, *Arabidopsis* encodes at least six TRF-like paralogs, three POT1-like proteins, and two TERs. Although duplication of telomere components is not common, gene duplication in general is prevalent. Several ancient whole-genome polyploidization events are documented in plants as well as in yeasts and animals (Van de Peer et al, 2009). As a result, many organisms are currently polyploid, or have a polyploid ancestry. Although *Arabidopsis* is a diploid, more than 66% of the genome is duplicated (Paterson et al, 2000).

The fate of duplicated genes include non-functionalization, neo-functionalization, and sub-functionalization (Prince & Pickett, 2002). The most common outcome of gene duplication is non-functionalization, where deleterious mutations accumulate in one gene of the pair, resulting in formation of a pseudogene or even locus deletion (Lynch & Conery, 2000; Moore & Purugganan, 2003; Walsh, 1995). A less frequent fate is neo-functionalization. In this case, one duplicate is exposed to distinct selective constraints, which shapes the gene to confer an adaptive advantage (Innan & Kondrashov, 2009). Another outcome is sub-functionalization, where each of the duplicated genes retains only a subset of the ancestral gene function (Force et al, 1999).

Positive selection

The rapid expansion of genomic sequences and the development of bioinformatic tools have allowed us to detect positive selection at the molecular level. Positive selection, or Darwinian selection, is a key mechanism of evolution. Positive selection occurs when a certain phenotype is favored, leading to the increase of prevalence of advantageous alleles in a population. Well-documented positive selection events include surface antigens of parasites or viruses (Endo et al, 1996), olfactory (Gilad et al, 2000) and fertilization genes in mammals (Swanson et al, 2003), and genes involved in pathogen resistance in plants (Cavatorta et al, 2008).

For protein coding genes, genetic codon substitution can result in a change (nonsynonymous) or no change (synonyomous) of the encoded amino acid. For purifying selection, the rate of nonsynonymous substitution is lower than the rate of synonymous substitution. This implies that the nonsynonymous substitution is deleterious for the function of encoded protein. When the rate of nonsynonymous substitution equals the rate of synonymous substitution, neutral evolution occurs and indicates the absence of selective constraints on sites of interest. In the case of positive selection, the rate of nonsynonymous substitution is higher than that of synonymous substitution. Positive selection indicates that nonsynonymous substitution confers a selective advantage and increases the frequency in the population (Delport et al, 2009). Therefore, the rate ($\omega = dN/dS$) of non-synonymous substitution (dN) to synonymous substitution (dS) is widely used to determine whether an amino acid is under purifying selection ($0 < \omega < 1$), neutral evolution ($\omega = 1$), or positive selection ($\omega > 1$) (Delport et al, 2009). In Chapter VI, the *POT1a* gene lineage from the Brassicaceae family was examined for evidence of positive selection. The data suggest several amino acids in

the first OB-fold of *POT1a* genes are under positive selection. These residues are essential for POT1a function *in vivo*.

Dissertation overview

This dissertation is composed of two major parts. In the first part (Chapters II-V), I present the identification and characterization of a CST telomere capping complex in *Arabidopsis thaliana*. In the second part (Chapters VI and VII), I switch gears and explore how a Shelterin component AtPOT1a has evolved to function as a telomerase recruitment factor instead of a telomere capping protein.

Chapter II presents the identification and characterization of STN1 in *Arabidopsis*. This is the first time that a STN1 ortholog has been reported in a multicellular organism. I show that AtSTN1 encodes a single OB-fold, and localizes to telomeres *in vivo*. Loss of STN1 results in catastrophic loss of telomeric and subtelomeric DNA, increased G-overhang signal, elevated telomere recombination, and massive chromosome end-to-end fusions. These findings reveal that STN1 is essential for chromosome end protection in *Arabidopsis*.

In Chapter III, a novel telomere capping component, Conserved Telomere maintenance Component 1 (CTC1), was uncovered in *Arabidopsis* and humans. CTC1 is predicted to harbor multiple OB-folds. Plants lacking *CTC1* displayed similar, if not identical, telomere defects as *stn1* mutants. Using a genetic approach, I show that *CTC1* protects chromosome ends in the same genetic pathway as *STN1*. Furthermore, CTC1 physically associates with STN1 in an *in-vitro* co-IP assay. It is proposed that CTC1 forms a complex with STN1, and together they guard the integrity of chromosome ends in *Arabidopsis*.

In Chapter IV, a TEN1 homolog was identified in the *Arabidopsis* genome using the recently identified human TEN1 protein as a query. AtTEN1 harbors a single OB-fold and interacts with STN1 *in vitro*. Mutants with reduced expression of *TEN1* showed modestly deregulated telomere length and genome instability. Taken together, my data suggest that TEN1, STN1 and CTC1 form an essential trimeric telomere capping complex in *Arabidopsis*.

In Chapter V, a genetic approach is employed to examine the interactions of *Arabidopsis* CTC1 and STN1 with telomerase, KU70 and the DNA damage response kinases ATM and ATR. Plants doubly deficient of CTC1/STN1 and a telomerase component exhibit severe developmental defects, massive genome instability, and even shorter telomeres. Thus, telomerase action is required to stabilize telomere tracts devoid of the CST complex. Furthermore, our data indicate that maintenance of ss G-overhang in *Arabidopsis* is facilitated by at least two different pathways: one requiring CTC1 and STN1, and a second involving KU. Finally, it is demonstrated that a dramatic increase in genome instability in plants lacking CTC1/STN1 and ATR, but not ATM. This finding indicates that the CST complex protects telomeres from eliciting an ATR-dependent checkpoint response, and further that ATR plays an additional role in maintaining *Arabidopsis* telomeres.

In Chapter VI, I explore the evolution of *Arabidopsis* POT1 proteins. The data suggest that *POT1* gene duplication is rare in plants. Only two instances of independent *POT1* gene duplication were detected: one in the dicot Brassicaceae family where *Arabidopsis* belongs, and the other in the monocot Panicoideae subfamily of grasses. Phylogenetic analysis uncovered that POT1a lineage in the Brasscicaceae family is undergone positive selection. Mutating two of the positive selection sites back to

ancestral amino acids dramatically reduced AtPOT1a function in a complementation assay. These data suggest that positive selection fuels the evolution of POT1a from a telomeric DNA binding protein to a TER-associated component and/or a telomerase recruitment factor.

In Chapter VII, evidence of an interaction between CTC1 and POT1a/telomerase is presetned. I hypothesize that POT1a has evolved to recruit telomerase to chromosome ends through interaction with CTC1 and TER1. Two strategies were used to dissect functional domains in POT1a. First, we performed site-directed mutagenesis and examined the mutant POT1a function by complementation analysis. Second, we screened a collection of ethylmethanesulfonate (EMS)-mutagenized mutants for novel *pot1a* alleles. The data indicate that both the N-terminus and the C-terminus are critical for POT1a function *in vivo*.

In Chapter VIII, conclusions and future directions of the Ph.D research regarding CST and POT1a function in *Arabidopsis* are presented.

CHAPTER II

STN1 PROTECTS CHROMOSOME ENDS IN *Arabidopsis thaliana*[*]

Summary

Telomeres shield the natural ends of chromosomes from nucleolytic attack, recognition as double-strand breaks, and inappropriate processing by DNA repair machinery. The trimeric Stn1/Ten1/Cdc13 complex is critical for chromosome end protection in *Saccharomyces cerevisiae*, while vertebrate telomeres are protected by Shelterin, a complex of six proteins that does not include STN1 or TEN1. Recent studies demonstrate that Stn1 and Ten1 orthologs in *Schizosaccharomyces pombe* contribute to telomere integrity in a complex that is distinct from the Shelterin components, Pot1 and Tpp1. Thus, chromosome end protection may be mediated by distinct subcomplexes of telomere proteins. Here we report the identification of a *STN1* gene in *Arabidopsis* that is essential for chromosome end protection. AtSTN1 encodes an 18 kDa protein bearing a single oligonucleotide/oligosaccharide binding fold (OB-fold) that localizes to telomeres *in vivo*. Plants null for AtSTN1 display an immediate onset of growth and developmental defects and reduced fertility. These outward phenotypes are accompanied by catastrophic loss of telomeric and subtelomeric DNA, high levels of end-to-end chromosome fusions, increased G-overhang signals and elevated telomere recombination. Thus, AtSTN1 is a crucial component of the protective telomere cap in *Arabidopsis*, and likely in other multicellular eukaryotes.

---

*Reprinted with permission from "Stn1 protects chromosome ends in *Arabidopsis thaliana*" by X. Song, K. Leehy, R. T. Warrington, J. C. Lamb, Y. V. Surovtseva, and D. E. Shippen. 2008. *Proc. Natl. Acad. Sci. U.S.A.* 105 (50):19815-19820.

Introduction

Telomeres distinguish the natural ends of chromosomes from double-strand breaks by virtue of their unusual architecture and protein composition. Vertebrate telomeres are bound by a core complex of six proteins, termed Shelterin, which regulates the length of the telomeric DNA tract, suppresses the activation of a DNA damage response at the terminus, and protects the ends from inappropriate recombination, nuclease attack and end-to-end fusion (de Lange, 2005; Palm & de Lange, 2008). Shelterin is composed of two double-strand telomere binding proteins, TRF1 and TRF2, a single-strand telomere binding protein, POT1, and three bridging proteins TIN2, RAP1 and TPP1 (de Lange, 2005; Palm & de Lange, 2008). TRF2 and the OB-fold containing protein POT1 are critical for chromosome end protection (Hockemeyer et al, 2006; van Steensel et al, 1998; Wu et al, 2006; Yang et al, 2005). Studies in *S. pombe* confirm the presence of several Shelterin homologs, including Taz1 (a TRF1/TRF2 homolog), Rap1, Pot1 and Tpz1 (a TPP1 homolog) (Baumann & Cech, 2001; Chikashige & Hiraoka, 2001; Ferreira & Cooper, 2001; Kanoh & Ishikawa, 2001; Miyoshi et al, 2008).

In contrast, budding yeast telomeres are protected by a trimeric complex of three OB-fold proteins, Stn1/Ten1/Cdc13 (Lundblad, 2006; Lustig, 2001; Pennock et al, 2001). Recent studies demonstrate that Stn1 and Ten1 orthologs in *S. pombe* also contribute to telomere capping (Martin et al, 2007). Notably, SpStn1 and SpTen1 interact with each other, but thus far evidence is lacking for a physical interaction between these proteins and SpPot1 (Martin et al, 2007). Furthermore, Tpz1, but not Stn1/Ten1, was identified by mass spectrometry of Pot1-associated proteins in *S. pombe* (Miyoshi et al,

2008), indicating that in *S. pombe* chromosome ends are protected by two distinct telomere protein subcomplexes.

To note, several candidate orthologs of the SpStn1 protein can be found in the genomes of multicellular eukaryotes, including humans, by position-specific iterative BLAST (PSI-BLAST) (Gao et al, 2007; Martin et al, 2007). Here we use a genetic approach to demonstrate that the *STN1* gene in the flowering plant *Arabidopsis* is essential for chromosome end protection. In striking contrast to plants lacking telomerase, which display a progressive but gradual loss of telomeric DNA that ultimately leads to end-to-end chromosome fusions and worsening growth and developmental defects beginning in the sixth generation (G6) (Riha et al, 2001), telomeres are immediately and catastrophically compromised in *Arabidopsis* mutants null for STN1. Telomeric as well as subtelomeric DNA is extensively eroded and mutants exhibit increased G-overhang signals, elevated telomere recombination and massive telomere fusion, resulting in severe growth defects and sterility. These findings not only indicate that AtSTN1 is required for telomere capping in *Arabidopsis*, but further suggest that additional key components of the telomere complex remain to be elucidated in metazoa.

Materials and methods

Plant materials and plasmids

The *stn1* mutants were obtained from the *Arabidopsis* Biological Resource Center (ABRC). The T-DNA insertion lines, *stn1-1* (CS023504) and *stn1-2* (CS846727), were genotyped by PCR using primers 5'-ATGGATCGATCCCTCCAAAG-3' and 5'-TTGAATACGAACACGATAACAAC-3'. Plants were grown according to the conditions

described (Surovtseva et al, 2007). Siliques from wild-type and *stn1-1* mutants were dissected ~ 10 days after fertilization and photographed using a Zeiss Axiocam digital camera coupled to a Zeiss microscope. A transgenic construct of STN1 was prepared by inserting a C-terminal YFP tag using an Ala $(Gly)_5$ Ala linker sequence. Tagged STN1 was cloned into a Gateway entry vector pENTR (Invitrogen) and then subcloned into a binary vector pB7WG2 (Invitrogen) according to manufacturer instructions. The resultant binary vector was used to transform plants as described (Surovtseva et al, 2007).

RT-PCR

Total RNA was extracted from plant tissues using an RNA purification kit (Fisher Scientifics). Reverse transcription was performed using Superscript III reverse transcriptase (Invitrogen) per manufacturer instructions. PCR of *STN1* cDNA was performed using the above primers, with the following program: 95 ℃ 3 min; 25 cycles of 94 ℃ 20 sec, 55 ℃ 30 sec, 72 ℃ 1 min 30 sec; 72 ℃ 7 min.

Cytology, immunofluorescence and FISH

To monitor anaphase bridge formation, cells were prepared from pistils, stained with DAPI Vectashield (Vector Laboratories), and then analyzed with an epifluorescence microscope (Zeiss) as described (Riha et al, 2001). Anaphase bridges were scored as a percentage of total anaphase cells. For combined immunolocalization and FISH, second generation transformants (T2) expressing a C-terminal YFP tagged version of STN1 were grown to seedlings (~ 7-day old) and fixed with 4% formaldehyde for 30 min on ice. Root nuclei from the seedlings were extracted and dried onto

polylysine coated slides, and immunolocalization was performed as described (Onodera et al, 2005). A rabbit anti-GFP antibody (Abcam) was used as the primary antibody and a FITC-conjugated donkey anti-rabbit antibody (Jackson ImmunoResearch) was used as the secondary antibody. After immunolocalization, the nuclei were postfixed with 4% formaldehyde and 0.1% glutaraldehyde for 30 min prior to FISH. Nuclei were washed with 1 × PBS, passed through an ethanol series (70%, 80%, 90%, 100%) at -20 ℃ and then dried. Digoxygenin-dUTP labeled telomere probe was prepared as described (Armstrong et al, 2001). FISH was performed as described (Kato et al, 2004). Detection of digoxygenin labeled probes was with a rhodamine conjugated anti-digoxygenin antibody (Roche). Nuclei were counterstained with DAPI Vectashield and analyzed with an epifluorescence microscope (Zeiss).

TRF, PETRA and telomere fusion PCR

DNA from individual whole plants was extracted as described (Cocciolone & Cone, 1993). TRF analysis was performed using 50 µg of DNA digested with *Tru1*I (Fermentas) and hybridized with a ^{32}P 5' end–labeled $(T_3AG_3)_4$ oligonucleotide probe (Fitzgerald et al, 1999). The average length of bulk telomeres was determined by Telometric 1.2 (Grant et al, 2001); the range of telomere length was obtained using ImageQuant software. Subtelomeric TRF analysis was performed using 100 µg of DNA digested with *Spe*I and *Pvu*II (New England Biolabs) and hybridized with a 5R probe (Shakirov & Shippen, 2004). Telomere fusion PCR and PETRA were performed as described (Heacock et al, 2004).

In-gel hybridization and telomeric circle amplification (TCA)

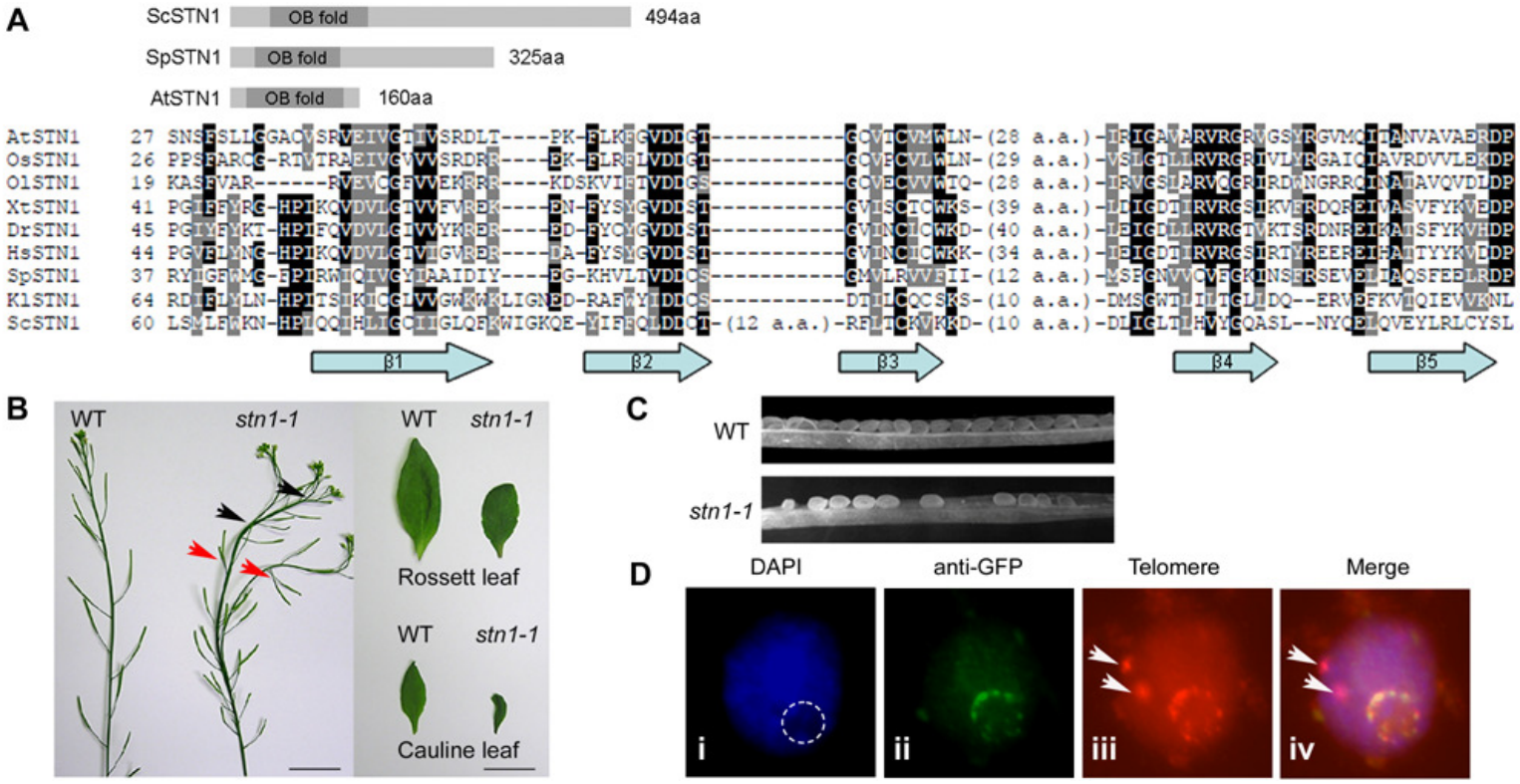
In-gel hybridization was performed as described (Heacock et al, 2007). The relative amount of single-strand G-overhang was calculated by quantifying the hybridization signal obtained from the native gel and then normalizing this value with the loading control of either interstitial telomere signal from the denaturing blot or ethidium bromide staining of the agarose gel. The single-strand G-overhang signal obtained from wild-type DNA was set to one and each sample was normalized to this value. Exonuclease treatment was performed by incubating DNA samples with T4 DNA polymerase (New England Biolabs) prior to in-gel hybridization at 12 ℃ for 30 min. The telomeric circle amplification (TCA) was performed as described (Zellinger et al, 2007).

Results

Identification of AtSTN1

To search for a STN1 protein in the plant kingdom, PSI-BLAST was employed using the protein sequence of SpStn1 as the query. In the second iteration, a previously uncharacterized protein, NP_563781, from *Arabidopsis thaliana* was uncovered with an E-value of $2e^{-06}$, well above the program threshold (0.005). The corresponding single-copy gene, At1g07130, was designated *AtSTN1*. A combination of EST database searches and 3' RACE was used to verify the boundaries of the *AtSTN1* coding region. AtSTN1 lacks introns and is predicted to encode a small protein of 160 aa that can assume a single OB-fold (Fig 2-1A).

Fig 2-1. Identification of AtSTN1 and severe morphological defects in *STN1* deficient plants. (**A**) Top, Diagram showing the OB-fold domain structure of STN1 homologs from *S. cerevisiae* (Sc), *S. pombe* (Sp) and *A. thaliana* (At). Bottom, Alignment of putative STN1 orthologs from plants and other organisms generated by Macvector and Boxshade software. The secondary structure was predicted by PSIPRED (McGuffin et al, 2000). At, *Arabidopsis thaliana* (NP_563781); Dr, *Danio rerio* (NP_956683); Hs, *Homo sapiens* (NP_079204); KL, *Kluyveromyces lactis*, (XP_452728); Ol, *Ostreococcus lucimarinus* (green algae, XP_001417183); Os, *Oryza sativa* (Rice, NP_001050181); Sc, *Saccharomyces cerevisiae* (CAA98902); Sp, *Schizosaccharomyces pombe* (XP_001713126); Xt, *Xenopus tropicalis* (NP_001004908). (**B**) Morphological defects in *stn1-1* mutants. Stems (left panel), rosette leaves (top right) and cauline leaves (bottom right) are shown for wild-type plants (WT) and *stn1-1* mutants. Fused stems (black arrows) and altered phyllotaxy (red arrows) are indicated. Bars, 1 cm. (**C**) Aborted seed development in *stn1-1* mutants. Siliques from wild-type plants and *stn1-1* mutants were visualized by microscopy. (**D**) STN1 colocalizes with telomeres. Isolated nuclei from STN1-YFP transformants were stained with DAPI (i), STN1-YFP was detected with an anti-GFP antibody (ii), and the telomeres were labeled by FISH with a telomere probe (iii) (see Materials and Methods for details). Panels (i) to (iii) were superimposed to produce panel (iv). Arrows in (iii) and (iv) indicate internal stretches of telomere signals as described in (Armstrong et al, 2001). *AtSTN1* mRNA is expressed in all plant tissues examined (Fig 2-2), unlike the mRNA for TERT, the catalytic subunit of telomerase, which accumulates only in highly proliferative organs (Fitzgerald et al, 1999).


A
ScSTN1
OB fold
494aa
SpSTN1
OB fold
325aa
AtSTN1
OB fold
160aa
AtSTN1 27 SNSESLLGGACVSRVEIVGTVSRDLT----PK-FLKGGVDDGT----------GCVTCVMWLN-(28 a.a.)-IRIGAVARVRGRVGSVRGVMCIDANVAVAERDP
OsSTN1 26 PPSEARCG-RTVTRAEIVGVVSRDRR----EK-ELRGIVDDGT----------GCVFCVLWLN-(29 a.a.)-VSLGTLLRVRGRIVLYRGAIQIAVRDVVLEKDP
OlSTN1 19 KASEVAR------RVEVCGFVVEKRR----KDSKVISTVDDGS----------GCVECVVWTQ-(28 a.a.)-IRVGSIARVQGRIRDWNGRRCINAIAVQVDLDP
XtSTN1 41 PGLEFYRG-HPIKQVDVLGTVVFVREK----EN-FYSMGVDDST----------GVISCTCWKS-(39 a.a.)-LEIGDTIRVRGSIKVERDQREIVASVFYKVEDP
DrSTN1 45 PGIYFYKT-HPIFQVDVLGIVVYKRER----ED-FYCMGVDDST----------GVINCICWKD-(40 a.a.)-LEIGDLLRVRGTVKTSRDNREIRAISFYKVHDP
HsSTN1 44 PGVELYNG-HPIKQVDVLGTVIGVRER----DA-FYSMGVDDST----------GVINCICWKK-(34 a.a.)-IEIGDTIRVRGSIRTYREEREIRAITYYKVLDP
SpSTN1 37 RYHGFWMG-EPIRWIQIVGYIAAIDIY----EG-KHVLIVDDCS----------GMVLRVVEII-(12 a.a.)-MSEGNVVCVFGKINSFRSEVELIAQSFEELRDP
KlSTN1 64 RDIEFYLN-HPITSIKICGIVVGWKWKLIGNED-RAFYYIDDCS----------DILLCQCSKS-(10 a.a.)-DMSGWTLILTGLLDQ--ERVPFKVLQIEVVKNL
ScSTN1 60 LSVLFWKN-HPIQQIHLIGCIIGLQFKWIGKQE-YIFHCLDDCT-(12 a.a.)-RFLTCKVKKD-(10 a.a.)-DLIGLTLHVYGQASL--NYQELQVEYLRLCYSL
β1
β2
β3
β4
β5
B
WT
stn1-1
WT
stn1-1
Rossett leaf
WT
stn1-1
Cauline leaf
C
WT
stn1-1
D
DAPI
anti-GFP
Telomere
Merge
i
ii
iii
iv

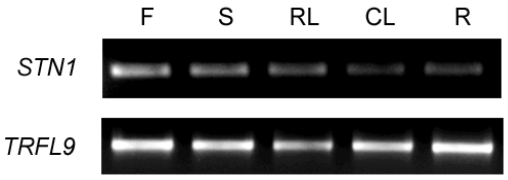


Fig 2-2. Ubiquitous gene expression of *STN1* in *Arabidopsis*. RT-PCR of *STN1* mRNA was performed in the indicated *Arabidopsis* tissues. CL, cauline leaf; F, flower; R, root; RL, rosette leaf; S, stem. RT-PCR of *TRFL9* is shown as a loading control.

Database searches revealed potential STN1 homologs from other sequenced plant genomes including rice and single-celled green algae (Fig 2-1A). As expected (Gao et al, 2007; Martin et al, 2007), putative STN1 homologs were also uncovered in a wide variety of other eukaryotes, including fishes, amphibians, birds, rodents and primates (Fig 2-1A and data not shown). In contrast to STN1 orthologs from yeast, the plant STN1 proteins lack a C-terminal extension (Fig 2-1A).

Protein sequence alignment indicated that AtSTN1 displays limited sequence similarity to SpStn1 (Fig 2-1A), but this similarity is statistically significant within the predicted OB-fold domain. Positions 7-143 of AtSTN1 align to positions 16-136 of SpStn1 with 23% identity/ 40% similarity. Secondary structure prediction by PSIPRED (McGuffin et al, 2000) indicated that residues within four of the five essential beta strands of the OB-fold (β1, β2, β3 and β4) in AtSTN1 share significant similarity to that of functionally verified STN1 protein from yeasts as well as the putative STN1 proteins from other multicellular eukaryotes (Fig 2-1A). In β5, sequence conservation is reduced in the Stn1 sequences from multicellular eukaryotes relative to their counterparts in yeasts. PFAM analysis confirmed that both AtSTN1 and SpStn1 proteins contain a "tRNA_anti" OB-fold nucleic acid binding domain, arguing that the OB-fold domain of the two proteins belongs to the same family. Results of PFAM analysis can be retrieved for AtSTN1 (http://pfam.sanger.ac.uk/protein?entry=Q9LMK5) and for SpStn1 (http://pfam.sanger.ac.uk/protein?entry=Q0E7J7).

Severe morphological defects in Arabidopsis stn1 *mutants*

We examined the *in vivo* function of *AtSTN1* by studying two T-DNA insertion lines, designated *stn1-1* and *stn1-2*, which were obtained from the *Arabidopsis* Biological

Resource Center. RT-PCR analysis of homozygous mutants confirmed that full-length *AtSTN1* mRNA was disrupted in both lines (Fig 2-3). Both mutant lines displayed a fasciated phenotype with severe morphological abnormalities in G1, although the severity of the defects varied somewhat in different individuals. In nearly all mutants, apical dominance was completely abolished, leading to multiple inflorescence bolts that were often fused (Fig 2-1B, black arrows). In addition, floral phyllotaxy was perturbed and siliques developed at irregular positions on the inflorescence bolt (Fig 2-1B, red arrows). Similar to what has been observed in late generation (G8-G9) *tert* mutants (Riha et al, 2001), leaf size was substantially reduced in *stn1* mutants, likely reflecting defects in cell proliferation (Fig 2-1B, right). *stn1* mutants produced numerous undeveloped ovules (Fig 2-1C) and the germination rate declined dramatically through successive generations. Only 17% (n=144) of the seeds from G1 mutants germinated to produce G2 plants. G2 progeny (G3) arrested early in vegetative development without producing a germline (data not shown). Many of these phenotypes are reminiscent of late generation *tert* mutants (Riha et al, 2001).

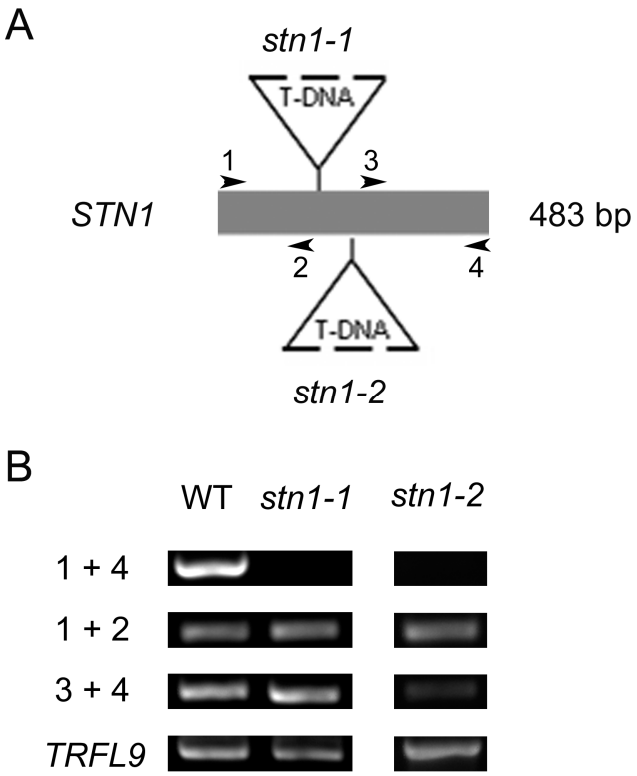


Fig 2-3. Identification of two *stn1* mutant alleles. Top, The relative positions of *stn1-1* and *stn1-2* T-DNA insertions are shown. Bottom, RT-PCR analysis of *STN1* mRNA in wild-type, *stn1-1 and stn1-2* mutants is shown. Primer positions are denoted by arrows. RT-PCR with primers flanking the T-DNA insertion suggests the full length mRNA of *STN1* was disrupted in both *stn1-1* and *stn1-2* mutants. RNA transcripts can be detected both upstream and downstream of the T-DNA insertion site. The upstream transcripts in *stn1-1* and *stn1-2* mutants encode small polypeptides (64 aa and 77 aa, respectively) and are likely to be non-functional. Additional analysis revealed that the downstream transcripts are likely derived from a cryptic promoter in the T-DNA construct and contained part of the T-DNA and an in-frame stop codon prior to the exon (data not shown). RT-PCR of *TRFL9* was used as a loading control.

AtSTN1 localizes to Arabidopsis *telomeres*

To monitor the subcellular localization of AtSTN1, we generated a *stn1-1* line expressing a C-terminal YFP tagged version of *AtSTN1* under the control of the CaMV 35S promoter. The transgene fully complemented the telomere defects in *stn1-1* mutants (see below). In root tip meristems, distinct spots of YFP signal formed a ring around the periphery of the nucleolus (data not shown). The arrangement of *Arabidopsis* telomeres at the nucleolar periphery has previously been noted in meiotic interphase (Armstrong et al, 2001). Fluorescence in situ hybridization (FISH) with a telomere probe also produced signals at the nucleolar periphery in somatic cells from roots and immature pistils (e.g., Fig 2-1D, panel iii). Immunolocalization using an anti-GFP antibody (Fig 2-1D, panel ii) combined with telomere FISH on the same nuclei produced co-localizing signals (Fig 2-1D, panel iv). This localization was specific to terminal telomeric DNA sequences as the STN1-YFP signal did not overlap with internal stretches of telomeric DNA sequence on chromosome 1 (Armstrong et al, 2001) (shown by the arrows in Fig 2-1D, panel iv). We conclude that AtSTN1 colocalizes with telomeres in *Arabidopsis*.

Extensive telomere erosion in plants lacking AtSTN1

In *S. pombe*, the absence of Stn1 leads to an immediate and profound loss of terminal DNA sequences (Martin et al, 2007). To determine if AtSTN1 protects chromosome ends in *Arabidopsis*, Terminal Restriction Fragment (TRF) analysis was performed to examine bulk telomere length. In both *stn1-1* and *stn1-2* mutants, telomere tracts appeared as a broad, heterogeneous smear (Fig 2-4A). Although the average length of bulk telomeres was only slightly shorter than in wild-type siblings (2.4

kb versus 2.7 kb, respectively), the shortest telomere tracts in *stn1-1* mutants were significantly shorter than in wild-type, trailing down to ~ 600 bp (1.4 kb shorter than the shortest wild-type telomeres) (Fig 2-4A). In contrast, telomeres in *tert* mutants decline much more gradually, reaching 600 bp in G6 or G7 (Riha et al, 2001).

Next we monitored telomere length dynamics on individual chromosome arms using subtelomeric TRF and Primer Extension Telomere Repeat Amplification (PETRA). For subtelomeric TRF, we used a probe corresponding to the right arm of chromosome 5 (5R) (Fig 2-4B). For PETRA, the left arms of chromosomes 1 and 3 (1L and 3L) were assessed (Fig 2-4C). Consistent with conventional TRF analysis, both assays revealed dramatic telomere erosion in plants lacking *AtSTN1*. Moreover, individual telomere tracts in *stn1* mutants spanned a broader size range than those in wild-type (Fig 2-4B and C). By contrast, telomere tracts on homologous chromosomes in *tert* mutants are even more homogenous in size than in wild-type, typically forming a single sharp band that spans 100-200 bp on an agarose gel (Heacock et al, 2004). We confirmed that the telomere defect in *stn1-1* mutants was due to the T-DNA insertion in the *AtSTN1* gene by complementation. Bulk telomere analysis (data not shown) and PETRA demonstrated that the profile of telomere tracts in *stn1-1* plants expressing an *AtSTN1* transgene was restored to wild-type (Fig 2-4D).

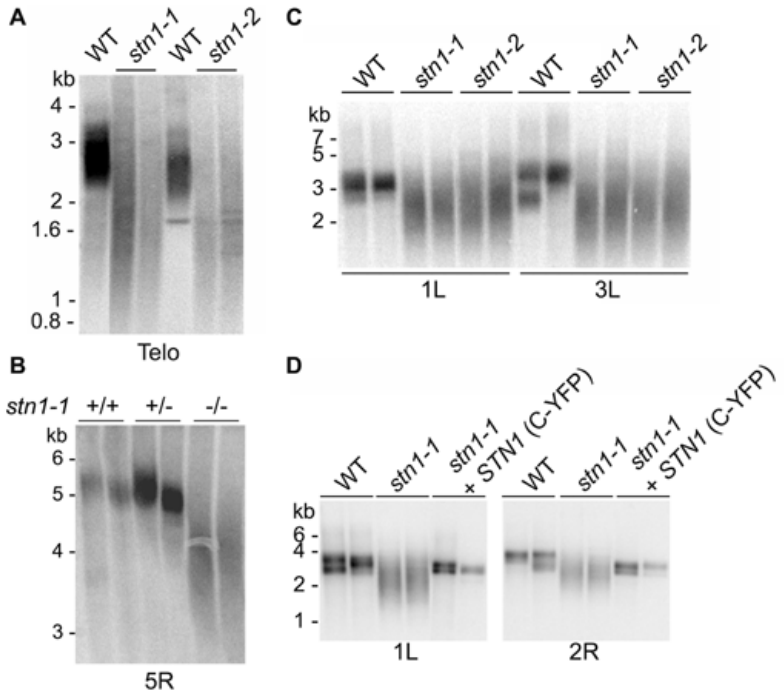


Fig 2-4. Extensive telomere erosion in *stn1* mutants. (**A**) TRF analysis of wild-type, *stn1-1* and *stn1-2* mutants. For each genotype, data from two individual sibling plants are shown. The blot was hybridized with a radiolabelled G-rich telomeric probe. Molecular weight makers are indicated. (**B**) Subtelomeric TRF analysis of wild-type, heterozygous and homozygous *stn1-1* mutants. The blot was hybridized with a probe specific for the right arm of chromosome 5 (5R). (**C**) PETRA analysis of wild-type, *stn1-1* and *stn1-2* mutants. The blot was hybridized with a telomeric probe. Telomere length on the left arm of chromosomes 1 and 3 (1L and 3L) was measured. (**D**) PETRA analysis of *stn1* mutants expressing a C-terminal YFP tagged wild-type *STN1* transgene. Telomere length was examined on the chromosome arms indicated.

Finally, we asked whether telomerase activity was diminished in *stn1* mutants using a real-time Telomere Repeat Amplification Protocol (Herbert et al, 2006; Kannan et al, 2008). *In vitro* telomerase activity levels in *stn1-1* mutants were approximately the same as in wild-type plants (Fig 2-5). Thus, the loss of telomeric DNA observed in *stn1* mutants is not due to telomerase enzyme deficiency, but we cannot rule out the possibility that telomerase access to the telomere is impeded in the absence of AtSTN1.

AtSTN1 is required to prevent telomere fusions

Extensive loss of telomeric DNA can trigger end-to-end chromosome fusions. To determine whether telomeres in *stn1* mutants engage in end-joining reactions, we monitored the frequency of anaphase bridges in the pistils of these plants. As expected, no bridged chromosomes were observed in wild-type plants (Fig 2-6A, Table 2-1). However, up to 29% of the anaphases in *stn1-1* mutants showed evidence of fused chromosomes (Fig 2-6B-D, Table 2-1). This degree of genome instability is not observed in *tert* mutants until G8 or G9 (Riha et al, 2001). The immediate and catastrophic onset of genome instability in *stn1* mutants reinforces the conclusion that AtSTN1 plays a critical role in chromosome end protection in *Arabidopsis*.

To further characterize the architecture of chromosome fusion junctions in *stn1* mutants, we employed telomere fusion PCR using primers directed at unique subtelomeric sequences on different chromosome arms (Heacock et al, 2004). Abundant telomere fusion PCR products were generated with G1 *stn1-1* DNA, which appeared as an intense, heterogeneous smear (Fig 2-6E).

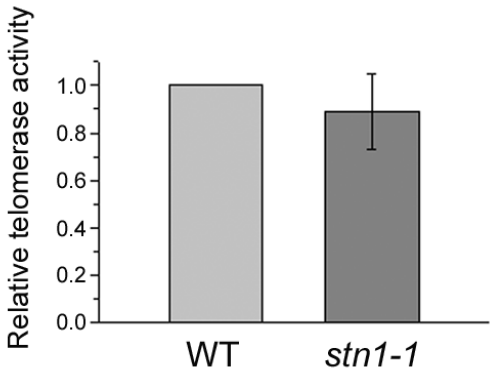
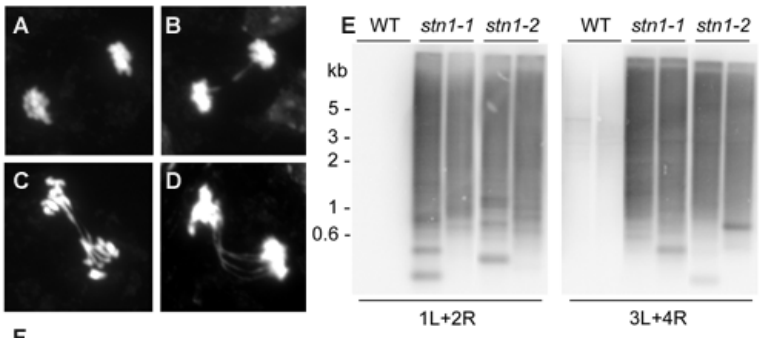


Fig 2-5. *In vitro* telomerase activity levels are approximately the same in *stn1* mutants as in wild-type plants. Real-time telomere repeat amplification protocol was performed with protein extracts from wild-type (n=3) and *stn1-1* mutants (n=3). The telomerase activity obtained from wild-type extracts was set to one and each sample was normalized to this value. Data are represented as mean ± SEM.



F

Type of Fusion	*stn1-1* (n=32) G1	*tert* (n=39) G9[1]	*ku70 tert* (n=38) G4[1]
Telomere-Telomere	0%	11%	43%
Telomere-Subtelomere	21%	78%	51%
Subtelomere-Subtelomere	79%	11%	5%
Features of Fusion Junctions			
Deletion of subtelomeric sequence (avg)	870 bp	260 bp	220 bp
Telomeric repeat retained at the fusion junction (avg)	30 bp	270 bp	270 bp

[1] Data reported in [22].

Fig 2-6. STN1 is required to prevent telomere fusions. (**A**-**D**) Cytology of mitotic chromosomes in wild-type (**A**) and *stn1-1* mutants (**B**-**D**) is shown. DAPI-stained chromosome spreads were prepared from pistils. Examples of *stn1-1* anaphases with one (**B**), two (**C**) or four (**D**) bridges are shown. (**E**) Telomere fusion PCR products obtained from wild-type, *stn1-1* and *stn1-2* mutants were hybridized using a telomeric probe. Primer pairs used to amplify specific subtelomeric regions are indicated. (**F**) Summary of DNA sequence analysis of cloned telomere fusion junctions in *stn1-1* (G1) mutants. Data for *tert* (G9) and *tert ku70* (G4) were taken from a previous study (Heacock et al, 2004).

Table 2-1. Analysis of anaphase bridges in *stn1* mutants.

	No. of pistils analyzed	total anaphase cells analyzed	anaphase cells with fusions	percentage anaphase cells with fusions
WT	4	203	0	0%
stn1-1 #54	5	241	41	17%
stn1-1 #55	3	222	54	24%
stn1-1 #70	4	229	66	29%

DAPI-stained chromosome spreads were prepared from pistils. Anaphase cells from three individual *stn1-1* mutants were analyzed. Anaphase bridges were scored as a percentage of total anaphase cells.

This observation is consistent with our previous studies showing that telomere fusion is initiated when telomeres shorten below 1kb (Heacock et al, 2004). Sequence analysis of cloned PCR products showed that the majority (79%) of end-joining events in *stn1-1* mutants involved subtelomere-to-subtelomere fusion (Fig 2-6F). In contrast, chromosome fusion junctions primarily reflect telomere-to-subtelomere joining in late generation *tert* mutants (78%), and telomere-to-telomere (43%) or telomere-to-subtelomere (51%) fusions in *ku70 tert* mutants (Fig 2-6F) (Heacock et al, 2004). Notably, the average deletion of subtelomeric DNA was four-fold greater in *stn1-1* mutants (~ 870 bp) (Fig 2-6F) than in *tert* (G9, ~ 260 bp) or *ku70 tert* mutants (G4, ~ 220 bp) (Heacock et al, 2004). Because bulk telomere length is much shorter in *tert* (G9) and in *ku70 tert* (G4) mutants where an equivalent level of genome instability is observed, our G1 *stn1-1* results indicate that at least a subset of telomeres in these mutants suffer extensive nucleolytic attack prior to being recruited into end-to-end chromosome fusions.

AtSTN1 is required to maintain proper telomere architecture and to block formation of extra-chromosomal telomeric circles

Mutations in Stn1, Ten1 or Cdc13 in *S. cerevisiae* (Grandin et al, 2001; Grandin et al, 1997; Nugent et al, 1996) and Stn1 in *K. lactis* (Iyer et al, 2005) lead to gross elongation of the G-overhang. These data are interpreted to mean that the Stn1 complex protects the telomeric C-strand from degradation. In-gel hybridization was used to determine if AtSTN1 contributes to the maintenance of telomere end structure in *Arabidopsis*. Relative to wild-type, the G-overhang signal was increased by approximately four-fold in *stn1-1* mutants (Fig 2-7A, left panel and Fig 2-7B).

Exonuclease treatment indicated that the hybridization signal detected in the native gel correlated with terminal G-overhangs (Fig 2-7A, right panel). Thus, AtSTN1 is required to maintain the proper architecture of the chromosome terminus.

The frequency of telomere recombination is dramatically increased in *K. lactis stn1* mutants (Iyer et al, 2005). To determine whether this is also true in plants lacking AtSTN1, we looked for evidence of Telomere Rapid Deletion (TRD). TRD results in large, stochastic deletions of telomere tracts and is thought to occur when the t-loop on the chromosome terminus undergoes branch migration, giving rise to a Holliday junction intermediate that is subsequently resolved to produce a truncated telomere and an extrachromosomal telomeric circle (Lustig, 2003). We monitored TRD using telomeric circle amplification (TCA), which detects the telomeric circle by-products of TRD (Zellinger et al, 2007). In this procedure, phi29 polymerase is used to amplify telomeric DNA circles into extremely long ssDNA, which is distinguished from endogenous linear telomere fragments based on its slower migration on a denaturing agarose gel. As expected, telomeric circles were enriched in our *ku70* mutant control reaction (Zellinger et al, 2007) (Fig 2-7C). A similar high molecular weight product was generated in *stn1-1* mutants, but not in the wild-type control. We conclude that STN1 suppresses telomere recombination in *Arabidopsis*.

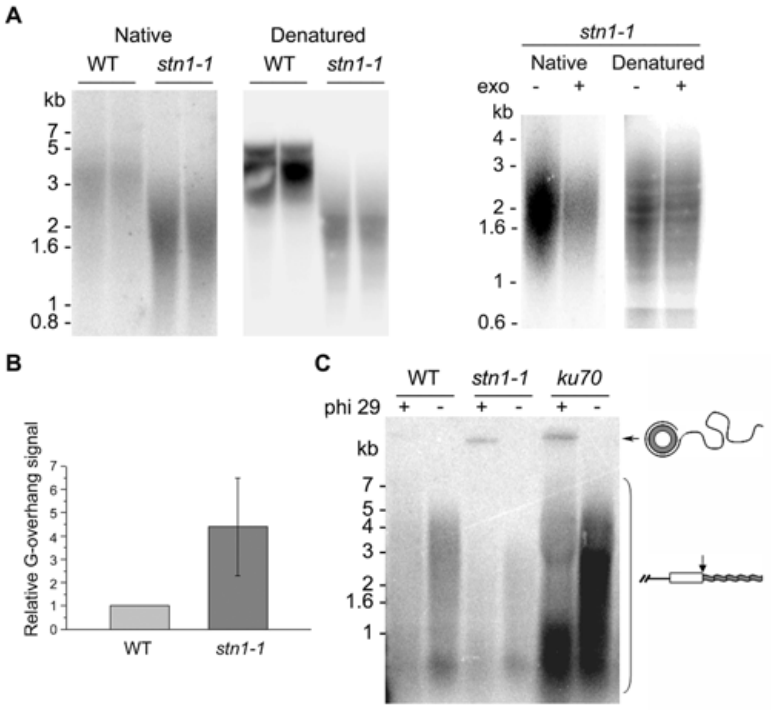


Fig 2-7. Loss of *STN1* leads to increased G-overhang signals and increased telomeric circle formation. (**A**) In-gel hybridization analysis of DNA isolated from wild-type and *stn1-1* mutants using a C-strand telomeric probe under native and denaturing conditions (left panel). The hybridization signal in the native gel was strongly reduced by exonuclease treatment, demonstrating that the signal was dependent on G-overhangs (right panel). (**B**) Quantification of the G-overhang signal. The relative G-overhang signal was determined from five independent experiments as described in Materials and Methods. Data are represented as mean ± SEM. (**C**) Telomeric circle amplification (TCA) was performed with wild-type, *stn1-1* and *ku70* mutant DNA in the presence or absence of phi 29 polymerase. DNA from *ku70* mutants served as a positive control. The hybridization signal for linear telomere tracts is indicated by the bracket.

Taken together, our data indicate that AtSTN1 is an essential component of the protective telomere cap in *Arabidopsis* that prevents nucleolytic attack, end-to-end chromosome fusions and telomere recombination.

Discussion

Although Barbara McClintock described the protective "capping" function of maize telomeres nearly 70 years ago (McClintock, 1939), we still know relatively little about why natural chromosome ends are recalcitrant to nuclease attack and end-joining reactions, especially in multicellular eukaryotes. In part, our understanding has been stymied by the rapid evolution of the telomere protein complex. Here we provide evidence that STN1 is conserved in metazoa and plays an essential role in chromosome end protection.

AtSTN1 was identified in the second iteration of PSI-BLAST as a protein bearing sequence similarity to the OB-fold domain of *S. pombe* Stn1. Subsequent analysis revealed putative STN1 orthologs in a variety of plants and vertebrates (this study; [Gao et al, 2007; Martin et al, 2007]). Structure-based alignment shows significant sequence similarity within four of the five essential beta strands of the core of the OB-fold. While the overall similarity among the Stn1 orthologs is not high, minimal sequence similarity among telomere proteins from different taxa is not without precedent. For example, Pot1 from *S. pombe* shows only 19% identity/ 40% similarity to the TEBP α subunit in ciliates, and yet the two proteins are functional and structural homologs (Baumann & Cech, 2001; Lei et al, 2003).

One notable distinction between the STN1 proteins from plants and yeasts is the absence of a C-terminal extension in the former. Recent studies indicate that the N- and

C-terminus of ScStn1 encode independent and separable functions at the telomere (Petreaca et al, 2007; Puglisi et al, 2008). The N-terminal OB-fold of ScStn1 is required for cell viability and mutation of this domain leads to an increase of single-strand DNA at the chromosome terminus (Puglisi et al, 2008), arguing that the N-terminal OB-fold is essential for chromosome end protection. In contrast, the C-terminal domain of ScStn1 is required for telomere length control and plays no detectable role in telomere capping (Puglisi et al, 2008). Like *Arabidopsis stn1* mutants, a null mutation in the *S. pombe* Stn1 leads to severe telomere deprotection phenotype, suggesting the major role of Stn1 in *S. pombe* and *Arabidopsis* may be in chromosome end protection. Notably, *S. pombe* Stn1 protein is significantly truncated relative to *S. cerevisiae* Stn1 (325 aa versus 494 aa), consistent with rapid evolution of the C-terminal domain. We hypothesize that the C-terminal domain of STN1 is not crucial for its telomere capping function in plants and hence was lost in the 1.5 billion years since plants and yeasts shared a common ancestor.

The strongest evidence that AtSTN1 is a functional homolog of the yeast Stn1 proteins is based on the genetic data. Plants lacking STN1 display phenotypes that strongly parallel the *S. pombe stn1* null mutants (Martin et al, 2007). In both cases, *stn1* mutants exhibit an immediate and profound telomere deprotection phenotype. In *Arabidopsis* mutants, both telomeric and subtelomeric tracts are subjected to extensive nuclease attack. Telomeric C-strands are particularly vulnerable to digestion, creating extended G-overhangs. As a likely consequence, *stn1* mutants exhibit increased intrachromosomal telomere recombination as evidenced by an accumulation of telomere circles. TRD may further fuel the erosion of terminal DNA sequences in this setting. The degraded telomeres engage in end-joining reactions, triggering genome-

wide instability and the cell proliferation arrest typical of plants experiencing severe telomere dysfunction (Riha et al, 2001). Thus, STN1 is a crucial component of the telomere complex in *Arabidopsis* that is essential for chromosome end protection.

Shelterin homologs have not been clearly defined in plants. *Arabidopsis* harbors at least six myb-related proteins that bind double-strand telomeric DNA *in vitro* in a manner similar to vertebrate TRF1 and TRF2 (Karamysheva et al, 2004), as well as three putative POT1 paralogs. Although the functions of AtPOT1b and AtPOT1c are still under investigation (Shakirov et al, 2005) (A.D.L. Nelson and D.E. Shippen, unpublished work), AtPOT1a is a physical component of the telomerase RNP that is required for telomerase action *in vivo* (Surovtseva et al, 2007). Strikingly, homologs for RAP1, TPP1 and TIN2 cannot be discerned in the *Arabidopsis* genome with the current search algorithms, underscoring the conclusion that telomere proteins are evolving at a rapid pace.

Besides STN1, the only other plant protein directly implicated in chromosome end protection is from rice. Like mammalian TRF2, rice telomere binding protein 1 (RTBP1) bears a myb-like DNA binding domain (Hong et al, 2007). However, in contrast to TRF2-depleted mammalian telomeres, which activate a strong DNA damage response and massive end-to-end chromosome fusions (Denchi & de Lange, 2007; Karlseder et al, 1999; van Steensel et al, 1998), plants lacking RTBP1 display very gradual telomere lengthening over successive plant generations and only in G2 do telomere fusions become evident (Hong et al, 2007). This mild phenotype may reflect functional redundancy of myb-bearing telomere proteins in plants (Shippen, 2006). *STN1*, by contrast, is a single-copy gene in all of the sequenced plant genomes we surveyed.

The conserved function of Stn1 in yeasts (Grandin et al, 1997; Martin et al, 2007; Pennock et al, 2001) and STN1 in flowering plants (this study), and the existence of putative homologs in primates, rodents, amphibians, birds and fishes (Gao et al, 2007; Martin et al, 2007) argues that this family of proteins may contribute to chromosome end protection in a broad range of eukaryotes. Notably, STN1 was not identified as a component of the Shelterin complex (Liu et al, 2004b; O'Connor et al, 2004; Ye et al, 2004a) in mammals. It is conceivable that STN1 interacts only transiently with telomeres, e.g. during a specific period of the cell cycle. Alternatively, STN1 may be part of an end protection complex distinct from Shelterin. In support of this hypothesis, a TPP1 homolog, Tpz1, but not Stn1/Ten1, was recently identified by mass spectrometry of Pot1-associated proteins in *S. pombe* (Miyoshi et al, 2008). Interestingly, SpPot1 does not interact with Stn1/Ten1 in a yeast two-hybrid assay (Martin et al, 2007), implying that *S. pombe* telomeres are composed of two distinct capping complexes, one bearing Pot1 and Tpz1 (from Shelterin) and a second containing Stn1 and Ten1. Given that mammalian Shelterin contains orthologs only from the former complex, POT1/TPP1, and that STN1 is a key component of the telomere cap in plants, the data suggest that higher eukaryotic telomeres are protected by a network of telomere protein subcomplexes, the full constituency of which is yet to be elucidated.

CHAPTER III

CONSERVED TELOMERE MAINTENANCE COMPONENT 1 INTERACTS WITH STN1 AND MAINTAINS CHROMOSOME ENDS IN HIGHER EUKARYOTES*

Summary

Orthologs of the yeast telomere protein Stn1 are present in plants, but other components of the Cdc13/Stn1/Ten1 (CST) complex have only been found in fungi. Here we report the identification of conserved telomere maintenance component 1 (CTC1) in plants and vertebrates. CTC1 encodes a novel ~ 140 kDa telomere-associated protein predicted to contain multiple OB-fold domains. *Arabidopsis* mutants null for *CTC1* display a severe telomere deprotection phenotype accompanied by a rapid onset of developmental defects and sterility. Telomeric and subtelomeric tracts are dramatically eroded, and chromosome ends exhibit increased G-overhangs, recombination, and end-to-end fusions. AtCTC1 both physically and genetically interacts with AtSTN1. Depletion of human CTC1 by RNAi triggers a DNA damage response, chromatin bridges, increased G-overhangs and sporadic telomere loss. These data indicate that CTC1 participates in telomere maintenance in diverse species and that a CST-like complex is required for telomere integrity in multicellular organisms.

---

*Reprinted with permission from "Conserved Telomere Maintenance Component 1 interacts with STN1 and maintains chromosome ends in higher eukaryotes" by Y. V. Surovtseva, D. Churikov, K. A. Boltz, X. Song, J. C. Lamb, R. T. Warrington, K. Leehy, M. Heacock, C. M. Price, and D. E. Shippen. 2009. *Mol. Cell* 36 (2): 207-218.

Introduction

The terminus of a linear chromosome must be distinguished from a double-strand (ds) break to avoid deleterious nucleolytic attack and recruitment into DNA repair reactions. Telomeres prevent such actions by forming a protective cap on the chromosome end. This cap consists of an elaborate, higher-order, DNA architecture and a suite of telomere-specific proteins. The formation of a t-loop of telomeric DNA is thought to play an important role in sequestering the terminal single-strand (ss) G-overhang from harmful activities (de Lange, 2004; Wei & Price, 2003), while ds and ss telomeric DNA binding proteins coat the chromosome terminus to further distinguish it from a ds break (Palm & de Lange, 2008).

In *Saccharomyces cerevisiae*, telomeres are bound by a trimeric protein complex, termed CST, composed of Cdc13, Stn1 and Ten1 (Gao et al, 2007; Lundblad, 2006). The three proteins interact to form an RPA-like complex with specificity for ss telomeric DNA. Cdc13 and Stn1 harbor at least one oligonucleotide-oligosaccharide binding (OB) fold, which in the case of Cdc13 is exploited to bind to the G-overhang (Guo et al, 2007; Mitton-Fry et al, 2002). Stn1 and Ten1 associate with the overhang primarily via interactions with Cdc13. The CST complex plays a key role in telomere length regulation (Bianchi & Shore, 2008). Cdc13 recruits the telomerase RNP via a direct interaction with the Est1 component of telomerase (Bianchi et al, 2004; Chandra et al, 2001), while Stn1 is thought to inhibit telomerase action by competing with Est1 for Cdc13 binding (Li et al, 2009; Puglisi et al, 2008). In addition, Cdc13 and Stn1 contribute to coupling of G- and C-strand synthesis through interactions with DNA polymerase α (Grossi et al, 2004; Qi & Zakian, 2000).

The CST complex is also essential for chromosome end protection. Mutations in any one of the CST components result in degradation of the C-strand, accumulation of ss G-rich telomeric DNA and late S/G2 cell-cycle arrest (Garvik et al, 1995; Grandin et al, 2001; Grandin et al, 1997). Telomere protection appears to be facilitated primarily by Stn1 and Ten1, and overexpression of Stn1 or Ten1 can rescue the lethality of Cdc13 depletion (Grandin et al, 2001; Petreaca et al, 2007; Puglisi et al, 2008). Finally, Cdc13 and Stn1 also inhibit telomere recombination (Iyer et al, 2005; Petreaca et al, 2006; Zubko & Lydall, 2006).

Mammalian telomeres are bound by Shelterin, a six-member complex that, unlike CST, binds both ss and ds telomeric DNA (Palm & de Lange, 2008). The Shelterin proteins TRF1 and TRF2 coat ds telomeric DNA, while POT1 binds the ss G-overhang. The TRF1/TRF2-interacting protein TIN2 and the POT1-interacting protein TPP1 associate with each other, providing a bridge between the duplex and ss regions of telomeric DNA. RAP1 associates with telomeres via interaction with TRF2. The majority of Shelterin components are implicated in telomere capping, although TRF2 and POT1 appear to play pivotal roles in this process. TRF2 associates with telomeric DNA via a myb-like DNA binding domain. Loss of telomere-bound TRF2 results in immediate degradation of the G-overhang and end-to-end chromosome fusions (Celli & de Lange, 2005), while certain dominant negative alleles cause rapid telomere shortening with extrusion of extra-chromosomal telomeric circles (ECTCs) via homologous recombination (Wang et al, 2004).

Like components of the CST complex, POT1 and its partner TPP1 harbor OB-folds. POT1 binds directly to the overhang through two adjacent OB-folds, thus sequestering the DNA 3’ terminus and reducing access to telomerase (Lei et al, 2004;

Lei et al, 2005). TPP1 does not bind DNA directly, but dimerization with POT1 increases the DNA-binding affinity of POT1 by ~10 fold (Wang et al, 2007). Knockdown of human POT1 by RNAi causes a fairly mild phenotype characterized by impaired proliferation, an increase in chromosome fusions, decreased G-overhang signals and an increase in telomere length (Hockemeyer et al, 2005; Veldman et al, 2004; Yang et al, 2005; Ye et al, 2004a). Disruption of the *POT1* gene leads to more dire consequences (Churikov et al, 2006; Hockemeyer et al, 2006; Wu et al, 2006) including activation of a strong ATR-mediated DNA damage checkpoint, G-overhang elongation, rapid telomere growth, elevated telomere recombination and ultimately cell death (Churikov & Price, 2008; Denchi & de Lange, 2007; Guo et al, 2007).

Telomere protein composition may be more conserved than previously surmised (Linger & Price, 2009). At least one Shelterin component, Rap1, is present in *S. cerevisiae*, although unlike vertebrate RAP1, ScRap1p directly binds ds telomeric DNA through two myb-like DNA binding domains and contributes to telomere length regulation and telomere silencing (Lundblad, 2006). Likewise, fission yeast contain several Shelterin orthologs including Taz1, an ortholog of mammalian TRF1/TRF2 proteins (Cooper et al, 1997), and Pot1 (Baumann & Cech, 2001). Furthermore, recent purification of SpPot1-associated proteins identified Tpz1, a presumed ortholog of vertebrate TPP1 (Miyoshi et al, 2008). Like TPP1, Tpz1 contains an OB-fold, and physical association of SpPot1 and Tpz1 is required for chromosome end protection (Miyoshi et al, 2008; Xin et al, 2007). The Pot1-Tpz1 complex recruits two additional proteins, Ccq1 and Poz1. Poz1 serves as a bridge linking the Pot1-Tpz1 complex to the ds telomere proteins Rap1 and Taz1 in a manner similar to the Shelterin component

TIN2 (Miyoshi et al, 2008). Altogether, these findings argue that the core components of the Shelterin complex are evolutionary conserved.

Emerging data indicate that components of the CST complex are also widespread. Although Cdc13 orthologs have yet to be uncovered outside of *S. cerevisiae*, a Stn1/Ten1 capping complex was recently described for *S. pombe* (Martin et al, 2007). Both proteins localize to telomeres and are essential for chromosome end protection from exonucleases and telomere fusions. Notably, no direct physical association between Stn1/Ten1 and Pot1 has been observed (Martin et al, 2007) and mass spectrometry of SpPot1-associated factors failed to identify Stn1 or Ten1 (Miyoshi et al, 2008). These findings suggest that CST and Shelterin components may constitute distinct telomere complexes.

Plants also appear to harbor both Shelterin and CST components. Several Myb-containing TRF-like proteins bind telomeric dsDNA *in vitro* (Zellinger & Riha, 2007) and in rice genetic data implicate one of these, RTBP1, in chromosome end protection (Hong et al, 2007). *Arabidopsis* encodes three OB-fold bearing POT1-like proteins (Shakirov et al, 2005; Surovtseva et al, 2007; A.D.L. Nelson and D.E. Shippen, unpubished work). Interestingly, while over-expression of a dominant negative allele of AtPOT1b or depletion of AtPOT1c lead to a telomere uncapping phenotype similar to a *pot1* deficiency in yeast and mammals (Shakirov et al, 2005) (A. Nelson, Y. Surovtseva and D. Shippen, unpublished data), AtPOT1a is dispensable for chromosome end protection and instead is required for telomerase function (Surovtseva et al, 2007). Currently, orthologs for TIN2, RAP1 and TPP1 cannot be discerned in any plant genome.

Recently, a distant homolog of the CST component STN1 was uncovered in *Arabidopsis* (Song et al, 2008). AtSTN1 bears a single OB-fold and localizes to telomeres *in vitro*. Deletion of AtSTN1 results in the immediate onset of growth defects and sterility, coupled with extensive exonucleolytic degradation of chromosome ends, increased telomere recombination, and massive end-to-end chromosome fusions (Song et al, 2008).

Here we report the identification of a novel telomere protein, termed CTC1 (conserved telomere maintenance component 1), that physically and genetically interacts with AtSTN1. We show that AtCTC1 localizes to telomeres *in vitro* and, as for AtSTN1, that loss of AtCTC1 triggers rapid telomere deprotection resulting in gross developmental and morphological defects, abrupt telomere loss, telomere recombination, and genome instability. Although not as severe as an *Arabidopsis ctc1* null mutant, the consequences of CTC1 knockdown in human cells include a DNA damage response, formation of chromatin bridges, increased G-overhang signals and loss of telomeric DNA from some chromosome ends. Altogether, these data argue that CTC1 is a component of a CST-like complex in multicellular organisms that is needed for telomere integrity. Notably, we have found that mammalian CTC1 and STN1 correspond to the two subunits of alpha accessory factor (AAF), a protein complex previously shown to stimulate mammalian DNA pol α-primase (Casteel et al, 2009; Goulian & Heard, 1990). Thus, the CST-like complex from plants and mammals may resemble the *S. cerevisiae* CST by providing a link between telomeric G- and C-strand synthesis.

Materials and methods

Mutant lines and CTC1 localization

The *ctc1-1* line was identified in the TILLING collection (Till et al, 2003). *ctc1-2* and *ctc1-3* lines were found in the SALK database (stock lines SALK_114032 and SALK_083165, respectively). Genotyping is described in supplemental methods. The *stn1-1* line was previously described (Song et al, 2008). A genetic cross was performed between plants heterozygous for *stn1-1* and for *ctc1-1*. For localization studies, a genomic copy of *CTC1* was cloned into the pB7WGC2 Gateway vector (Karimi et al, 2005). The resulting N-terminal CFP fusion was transformed into wild type *Arabidopsis* (Surovtseva et al, 2007).

Map-based cloning

Map-based cloning was performed essentially as described (Lukowitz et al, 2000). Briefly, a mutant line (*Columbia* ecotype) was out-crossed to wild type *Arabidopsis Landsberg erecta* 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 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.

siRNA-mediated knockdown of HsCTC1

HeLa, MCF7 or 293T cells were subject to two rounds of transfection 24 hrs apart using Lipofectamine2000, Oligofectamine CaPO4. The final concentration of siRNA duplex (see supplemental methods for sequences) was 50 mM (Ambion) or 100 nM (EZBiolab) for each transfection. The efficiency of CTC1 knockdown was assessed using quantitative real-time RT-PCR with SYBR Green. Regions corresponding to CTC1 and GAPDH mRNAs were amplified for each RNA sample. The GAPDH mRNA level was used as an endogenous control to normalize the level of CTC1 mRNA for each RNA sample. The normalized values were plotted relative to the mock-transfected control that was set to 100%. All reactions were performed in duplicate.

Genotyping of Arabidopsis *mutant lines, DNA and RNA extraction, and RT-PCR*

To genotype the *ctc1-1* line, a genomic region flanking the *ctc1-1* point mutation was amplified with CTC1_M2 fwd (5'-GTAATGCCCATCTCAAGTTTTG) and CTC1_M2_rev (5'-CAGCACACGCATAGCACTATG) primers and sequenced with the CTC1_M2 rev primer. Genotyping of the *ctc1-2* and *ctc1-3* lines was performed with T-DNA and gene-specific primers.

DNA was extracted from plants as previously described (Cocciolone & Cone, 1993). RNA samples were prepared using Plant RNA Purification Reagent (Invitrogen) and reverse transcription was performed using 2 µg of RNA, as described (Shakirov et al, 2005). AtCTC1 cDNA was amplified in the PCR reaction with primers CTC1_start_fwd (5'-ATGGAGAACACCACAATTCTCAC) and CTC1_stop_rev (5'-TCAGCTATTTAGCAAACCTTGGAG). To evaluate expression of the region flanking the T-DNA insertion in the *ctc1-2* allele, primers 5'-GTCACGCTTTTGAGAGGTCTG and

CTC1_M2_rev were used. For the *ctc1-3* allele, primers CTC1_M2_fwd and 5'-CACTTGAGGAACTTATCCTCTG were used.

Protein expression and co-immunoprecipitation

For *in vitro* studies, full-length CTC1 cDNA or its truncated versions were cloned into pET28a and pCITE4a vectors (Novagen) and expressed using rabbit reticulocyte lysate according to manufacturer's instructions (Promega). For *in vitro* co-immunoprecipitation, pET28a (T7-tag fusion) and pCITE4a (untagged) constructs were expressed in rabbit reticulocyte lysate in absence or presence of ^{35}S-Methionine (PerkinElmer), respectively. Coimmunoprecipitation was conducted as described (Karamysheva et al, 2004).

TRF analysis, PETRA, telomere fusion PCR, quantitative TRAP, and in-gel hybridization

TRF analysis of *Arabidopsis* telomeres 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 & Shippen, 2004). For PETRA (Heacock et al, 2004), 2 µg of DNA was used. An adapter primer was hybridized to the G-overhang and extended with ExTaq polymerase (Takara), followed by a specific chromosome arm amplification with unique subtelomeric primers as described in (Heacock et al, 2004).

Telomere fusion PCR was performed as previously described (Heacock et al, 2004). PCR products were purified, cloned into pDrive vector (Quiagen), and sequenced. Quantitative TRAP assay was performed as described (Kannan et al, 2008).

G-overhangs were analyzed by in-gel hybridization as previously described for *Arabidopsis* and human telomeres (Churikov & Price, 2008; Song et al, 2008). Genomic DNA was separated in native agarose gels, dried gels were then hybridized with ^{32}P 5' end-labeled telomeric C-strand probe $(C_3TA_3)_4$ for plant DNA and $(TA_2C_3)_4$ for human DNA). For quantification of *Arabidopsis* G-overhang signal, the hybridization signal from the native gel was normalized with the signal from the ethidium bromide-stained gel. The G-overhang signal obtained from mutant samples was compared to wild type signal, which was set to one. To quantify the G-overhang signal from human telomeres, the native gel was denatured and reprobed with the C-strand oligonucleotide. The signal from the denatured gel was used to normalize for gel loading.

Telomeric circle assays

For TCA and bubble trapping, DNA was digested with Alu1. TCA was performed using 50 µg of DNA as described (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°C, mixed, and loaded on 0.6% agarose gel. The gel was run at 20 V at 4°C for 16 hrs. DNA was then transferred to the nylon membrane and hybridized with a G-rich telomeric probe.

Cytology, immunofluorescence and FISH

For cytological analysis of *Arabidopsis* chromosomes, spreads were prepared from pistils as described (Riha et al, 2001). Chromosomes were stained with DAPI (4',6'-diamidino-2-phenylindole) and analyzed with epifluorescence microscope (Zeiss). Immunolocalization and FISH were performed on CFP-CTC1 7-days old seedlings as

discussed (Song et al, 2008). The BACs used were those described in (Vespa et al, 2007).

Human cells were fixed and stained for indirect immunofluorescence as described (Churikov & Price, 2008; Churikov et al, 2006) using monoclonal or polyclonal antibody to γ-H2AX, Ser139 and monoclonal to TRF2. Interphase bridges were visualized with DAPI. Colocalization of γ-H2AX and TRF2 foci was monitored using a colocalization plug-in written for Image J by Pierre Bourdoncle (Institut Jacques Monod, Service Imagerie, Paris). Two foci were considered colocalized if their respective intensities were higher than the set threshold of their channels, and if their intensity ratio was higher than the set value. Metaphase spreads were prepared and telomere FISH performed as described (Churikov & Price, 2008; Churikov et al, 2006; Lansdorp et al, 1996). FISH signals were scored using Image J using the Cell counter plug-in.

Results

Identification of CTC1

In an effort to identify mutations in *AtPOT1c*, we examined lines within a TILLING collection of EMS-mutagenized *Arabidopsis* plants. A mutant was uncovered that showed a profound telomere uncapping phenotype (described below). However, this phenotype did not segregate with nucleotide changes in *AtPOT1c* and therefore map-based cloning was employed to identify the lesion responsible for the phenotype. A single-nucleotide transition (G to A) was found in At4g09680, which co-segregated with telomere uncapping. At4g09680 lies on chromosome 4, while *AtPOT1c* resides on chromosome 2. At4g09680 was designated *CTC1* (conserved telomere maintenance component 1) and the point mutant was termed *ctc1-1*. *CTC1* is a single copy gene and

sequence analysis of *CTC1* cDNA from wild type plants revealed a large ORF with 16 exons that encodes a novel 142 kDa protein (Fig 3-1A). RT-PCR demonstrated that *CTC1* is widely expressed in both vegetative and reproductive organs (Fig 3-2A). Further analysis of the CTC1 protein sequence is discussed below.

CTC1 associates with telomeres in vitro

To determine whether CTC1 associates with telomeres *in vitro,* an N-terminal CFP-tagged version of CTC1 protein was expressed in transgenic *Arabidopsis* and immunolocalization experiments were performed on different tissues. Nuclear CFP signal was detected in plants expressing CFP-CTC1, but not in untransformed controls (Fig 3-1B, Fig 3-2B and data not shown). Telomere distribution was analyzed by fluorescence in-situ hybridization (FISH) using a telomere probe. In *Arabidopsis,* telomeres lie at the nucleolar periphery (Armstrong et al, 2001; Song et al, 2008) and, as expected, telomeric FISH signals were positioned in this location. Similarly, CFP-CTC1 was distributed in a punctate pattern surrounding the nucleolus. A merge of these images showed that much of the CFP-CTC1 co-localized with *Arabidopsis* telomeres (Fig 3-1B and Fig 3-2B). CTC1 association with telomeres was quantitated in flowers and seedlings, which contain cycling cells. On average, 51% (n = 38, SD = ±26%) of the telomere signals overlapped with CFP-CTC1. To determine if CTC1 colocalization with telomeres was retained in noncycling cells, we examined the apical half of rosette leaves that were at least 2 weeks old and arrested in G1 (Donnelly et al, 1999). In these cells, 44.1% (n = 28, standard deviation = ±24.5%) of the telomeres displayed an overlapping signal with CFP-CTC1. These data argue that CTC1 associates with telomeres throughout the cell cycle.

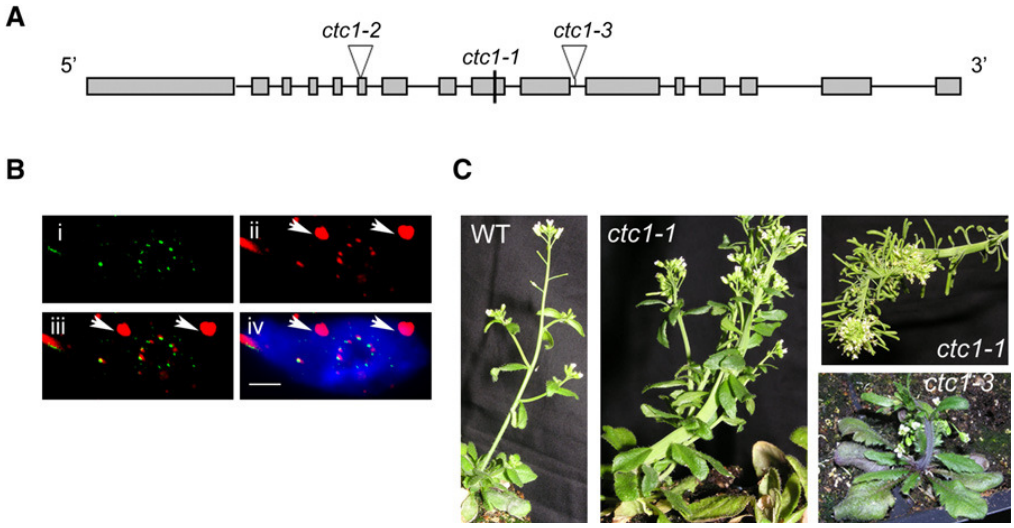


Fig 3-1. Identification of CTC1 in *Arabidopsis thaliana*. (**A**) Schematic of the AtCTC1 gene locus. Rectangles represent exons; horizontal black lines are introns. The positions of the point mutation (*ctc1-1*) and T-DNA insertions (*ctc1-2* and *ctc1-3*) are shown. (**B**) Colocalization of AtCTC1 and telomeres at the nucleolus periphery of leaf nuclei from seedlings. (i) CFP-AtCTC1 localization detected with anti-GFP antibody; (ii) telomere FISH using probe made from DIG-UTP-labeled T_3AG_3-C_3TA_3; (iii) CFP-AtCTC1-telomere merge; (iv) image from (iii) is combined with DAPI-stained nucleus. The nucleolus appears as a ring where DAPI staining is excluded, arrows in (i)–(iv) indicate internal stretches of telomeric DNA sequence (Armstrong et al, 2001). Scale bar, 2.5 mm. (**C**) Morphological defects in *ctc1* mutants. Left panel, wild-type; middle and right panels, first generation *ctc1-1* and *ctc1-3* mutants of similar age. Fasciated stems and fused organs in *ctc1* mutants are shown. The severity of morphological defects varies among *ctc1* mutants.

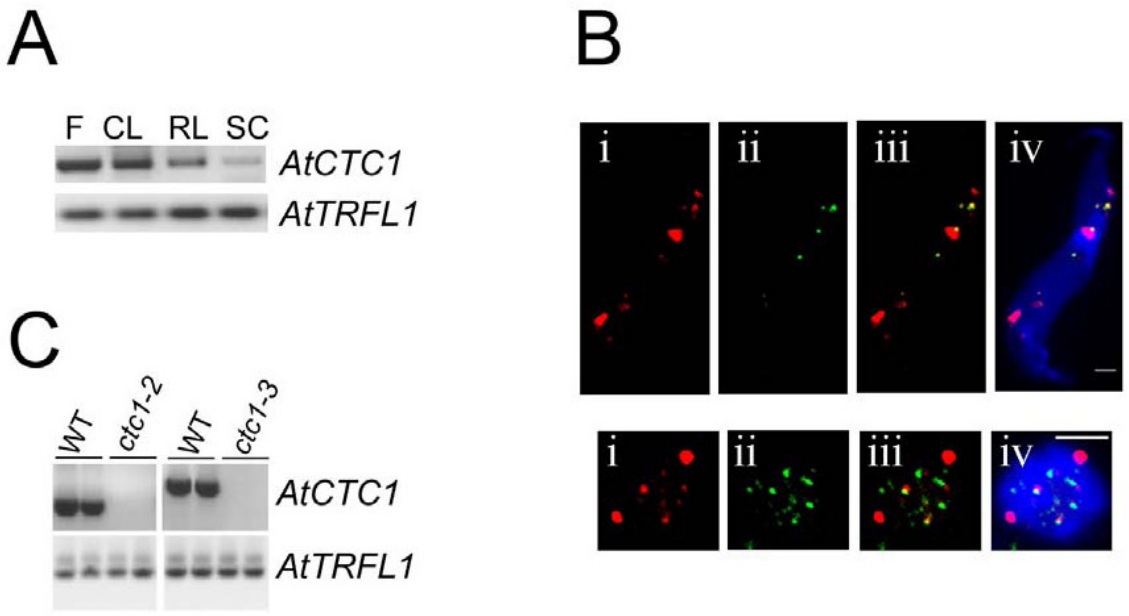


Fig 3-2. *AtCTC1* gene expression in wild type and in T-DNA insertion mutants. (**A**) RT-PCR analysis of the *AtCTC1* gene expression in different plant tissues. F, flowers; CL, cauline leaves; RL, rosette leaves; SC, suspension culture. (**B**) Co-localization of AtCTC1 and telomeres. Seedling root nucleus and flower nucleus are shown in top and bottom panels, respectively. (i) CFP-AtCTC1 localization detected with anti-GFP antibody; (ii) telomere FISH; (iii) CFP-AtCTC1 – telomere merge; (iv) image from panel (iii) is combined with DAPI stained nucleus. Scale bar = 2.5 μm. (**C**) RT-PCR analysis of AtCTC1 gene expression in *ctc1-2* and *ctc1-3* mutants. Primers flanking the insertion were used in both cases. TRFL1, a constitutively expressed gene (Karamysheva et al, 2004), was used as a loading control.

Severe growth defects and sterility in first-generation ctc1 mutants

We next examined the impact of CTC1 inactivation on plant morphology. Sequence analysis of *CTC1* cDNA from *ctc1-1* mutants revealed that the G(1935)A point mutation resulted in a nonsense codon within the ninth exon (Fig 3-1A). Two additional *CTC1* alleles, *ctc1-2* and *ctc1-3*, bearing T-DNA insertions in the sixth exon or tenth intron, respectively, were identified within the SALK database (Fig 3-1A). RT-PCR analysis showed that no *CTC1* full length mRNA was produced in either *ctc1-2* or *ctc1-3*, indicating that these lines are null alleles of *AtCTC1* (Fig 3-2C).

All three *ctc1* mutants displayed a rapid onset of severe morphological defects in the first generation (Fig 3-1C), confirming that *CTC1* lesions are responsible for telomere uncapping. The large majority of *ctc1* plants had grossly distorted floral phyllotaxy with an irregular branching pattern and fasciated (thick and broad) main and lateral stems and siliques (Fig 3-1C). Although most mutants produced an influorescence bolt, this structure was highly variable in size, ranging from very short to wild type (Fig 3-1C, compare middle and bottom right panels). Flowers and siliques were often fused, and seed yield was typically reduced to ~ 10% of wild type. The germination efficiency of the few seeds that could be recovered was extremely low, making propagation to the next generation almost impossible.

Telomere shortening and increased length heterogeneity in ctc1 *mutants*

Terminal restriction fragment (TRF) analysis was performed to examine bulk telomere length in *ctc1* plants derived from a single self-pollinated heterozygous parent. In contrast to the telomeres of their wild type and heterozygous siblings, which spanned 2-5 kb in length (Fig 3-3A, lanes 1 to 4), telomeres in homozygous *ctc1-1* mutants were severely deregulated (Fig 3-3A, lanes 5 and 6). The longest *ctc1-1* telomeres were in the wild type range, but a new population of shorter telomeres emerged, the shortest of which trailed to 0.5 kb. Homozygous *ctc1-2* and *ctc1-3* mutants showed a similar aberrant telomere length phenotype (Fig 3-4A).

We investigated how individual telomeres were affected by CTC1 loss using subtelomeric TRF analysis with probes directed at specific chromosome termini. As expected (Shakirov & Shippen, 2004), sharp bands were produced from wild type telomeres (Figs 3-3B and 3-4B). In contrast, telomeres in *ctc1* mutants gave rise to a broad heterogeneous hybridization signal spanning 1.5 kb (Figs 3-3B and 3-4B). Primer extension telomere repeat amplification (PETRA) also generated broad smears in *ctc1* mutants, confirming that the length of individual telomere tracts was grossly deregulated (Fig 3-3C). Telomere shortening and increased heterogeneity at individual telomere tracts in *ctc1* mutants is not due to a reduction in telomerase activity. Quantitative Telomere Repeat Amplification (Q-TRAP) revealed no significant difference in the *in vitro* telomerase activity levels in *ctc1* mutants relative to wild type (Fig 3-5).

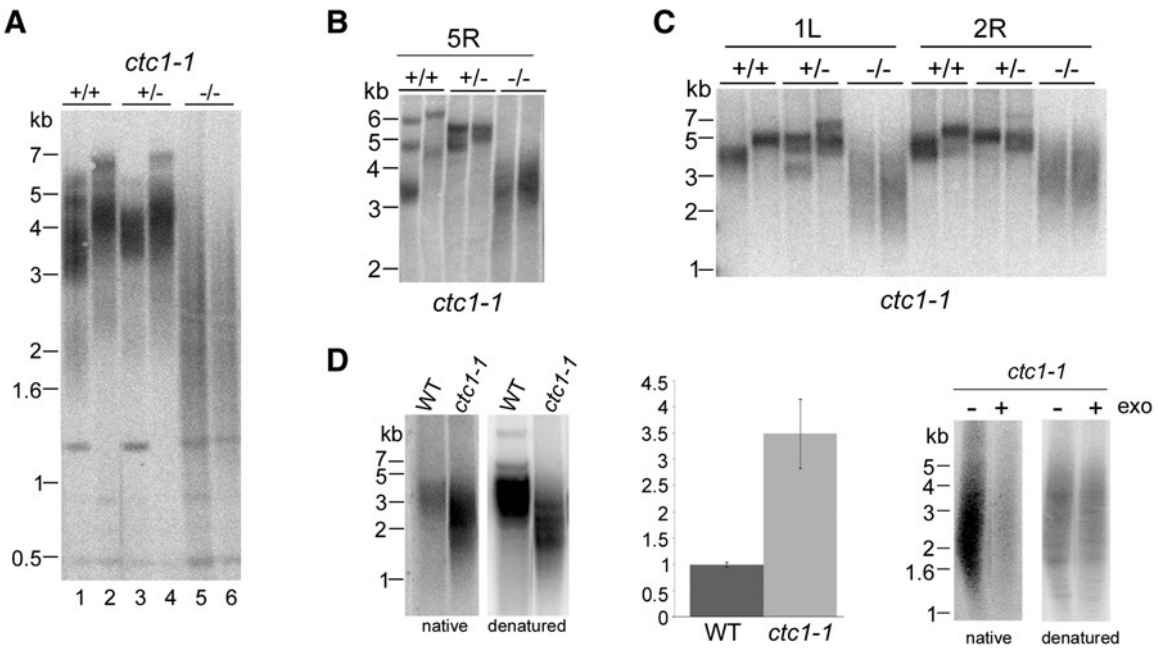


Fig 3-3. Telomere length deregulation and increased G-overhangs in *AtCTC1* mutants. (**A**) TRF analysis of *ctc1-1*. Results are shown for progeny segregated from a parent heterozygous for *ctc1*. (**B**) Subtelomeric TRF analysis of DNA from *ctc1-1* mutant. DNA blots were hybridized with a probe corresponding to subtelomeric regions on the right arm of chromosome 5 (5R). (**C**) PETRA analysis of DNA from *ctc1-1* mutants. Results for the 1L and 2R telomeres are shown. (**D**) In-gel hybridization of $(C_3TA_3)_4$ probe to telomeric restriction fragments under native and denaturing conditions (left). Quantification of *ctc1-1* signal relative to wild-type is shown in the middle panel. Data are the average of eight independent experiments ± SD (p = 1.3E-5 Student's t test). Right panel, in-gel hybridization of *ctc1-1* DNA in the absence (-) or presence (+) of 3'-5' exonuclease (T4 DNA polymerase). In (**A**) and (**C**), blots were hybridized with a radiolabeled telomeric DNA probe $(T_3AG_3)_4$. Molecular weight markers are indicated.

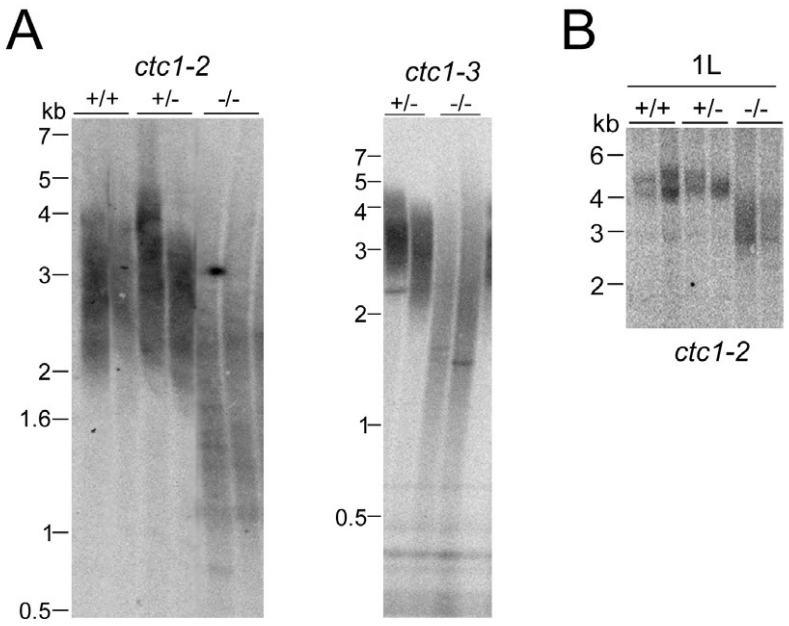


Fig 3-4. Telomere length deregulation in *AtCTC1* deficient mutants. (**A**) TRF analysis of *ctc1-2 and ctc1-3* mutants. Results are shown for progeny segregated from a parent heterozygous for *ctc1*. DNA blots were hybridized with a radiolabeled telomeric $T_3AG_3)_4$ probe. (**B**) Subtelomeric TRF analysis of DNA from *ctc1-2* mutant. Blots were hybridized with a probe corresponding to subtelomeric region on the left arm of chromosome 1 (1L). In both panels, molecular weight markers are indicated.

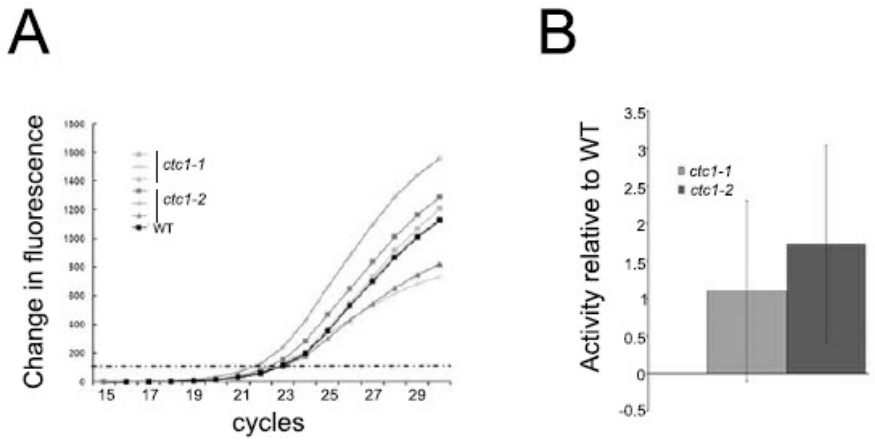
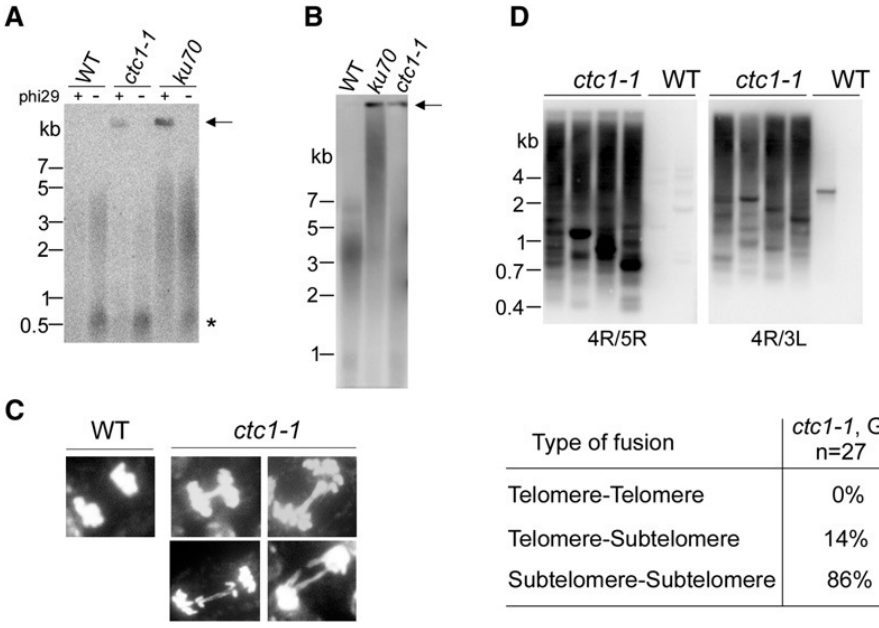


Fig 3-5. Results of real time TRAP on *ctc1-1* and *ctc1-2* mutants. Left panel shows raw data. Dashed line represents the threshold change in fluorescence. Right panel shows quantification of the telomerase activity levels in *ctc1* mutants relative to wild type.

Increased G-overhang signals and telomere recombination in ctc1 *mutants*

Next we studied the G-overhang status in *ctc1* mutants using non-denaturing in-gel hybridization. Strikingly, the G-overhang signal was ~three times greater in *ctc1* mutants relative to wild type (3.5 ± 0.7) (Fig 3-3D). A similar increase in G-overhang signal is observed in *Arabidopsis stn1* mutants (Song et al, 2008). Exonuclease treatment reduced the G-overhang signal by approximately 95%, indicating that the majority of ss telomeric DNA is associated with the chromosome terminus (Fig 3-3D, right panel).

To investigate whether telomeres in *ctc1* mutants are subjected to increased recombination, we used t-circle amplification (TCA) (Zellinger et al, 2007) to look for evidence of ECTCs, a by-product of t-loop resolution. 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 previously shown to accumulate ECTCs (Zellinger et al, 2007). A high molecular weight DNA band was detected in both *ku70* and *ctc1* DNA samples, but not in wild type (Fig 3-6A). To verify the presence of ECTCs in *ctc1* mutants, we employed the bubble trapping technique (Mesner et al, 2006), which relies on the ability of linear DNA fragments to enter the gel, while circular DNA cannot. A telomeric signal was detected in the well with DNA from *ctc1* and *ku70* mutants, but not with wild type (Fig 3-6B). These data confirm that ECTCs accumulate in the *ctc1* background and argue that loss of CTC1 results in elevated rates of homologous recombination at telomeres. Altogether, these results indicate that the architecture of the chromosome terminus is perturbed in the absence of CTC1.



Type of fusion	*ctc1-1*, G1 n=27	*tert*, G9* n=39
Telomere-Telomere	0%	11%
Telomere-Subtelomere	14%	78%
Subtelomere-Subtelomere	86%	11%

Fig 3-6. *ctc1-1* mutants display elevated telomere recombination and end-to-end fusions. (**A**) TCA with *ctc1-1* DNA. Reactions were performed in the presence or absence of phi29 polymerase. *ku70* DNA was used as a positive control. (**B**) Bubble-trapping results for *ctc1-1* and *ku70* mutants. All blots were hybridized with a radiolabeled telomeric probe. In (**A**) and (**B**), the probe hybridized to both circular and linear telomeric DNA products. Arrows mark TCA product/circles, smears correspond to TRFs, and the asterisk indicates an interstitial telomeric repeat signal. (**C**) Cytogenetic analysis of *ctc1-1* mutants showing DAPI-stained chromosome spreads with anaphase figures. (**D**) Telomere fusion PCR analysis of *ctc1-1* mutants. Primers were specific for 4R and 5R (left) or 4R and 3L (right). The table shows types of fusion junctions found after sequencing PCR products.

End-to-end chromosome fusions in ctc1 *mutants*

In *Arabidopsis,* telomeres shorter than 1 kb are prone to end-to-end chromosome fusions (Heacock et al, 2007). Since a substantial fraction of *ctc1* telomeres dropped below this critical threshold, we looked for evidence of mitotic abnormalities. Anaphase bridges were scored in four individual *ctc1-1* mutants and in their wild type siblings. As expected, there was no evidence of genome instability in wild type plants, but in all four *ctc1-1* mutants a high fraction of mitotic cells (up to 39%) exhibited anaphase bridges (Fig 3-6C and Table 3-1). Many anaphases contained multiple bridged chromosomes as well as instances of unequal chromosome segregation (Fig 3-6C). FISH using a mixture of probes from nine subtelomeric regions produced signals in 20/23 anaphase bridges, indicating that the bridges represent end-to-end fusions (Table 3-2). FISH probes from eight chromosome ends were individually applied to chromosome preparations from a single *ctc1-1* flower cluster. Signals from each probe were observed in anaphase bridges suggesting that all chromosome arms participated in chromosome fusions (Table 3-1).

Table 3-1. Frequency of anaphase bridges in *ctc1-1* mutants.

Genotype	# of analyzed pistils	# of anaphases		% anaphase bridges
		with bridges	total scored	
ctc1-1 #1	4	50	127	39
ctc1-1 #2	6	95	395	24
ctc1-1 #3	3	80	278	29
ctc1-1 #4	4	54	190	28
WT	4	1	140	0
tert, G6				6*
tert, G9				~40*

* Data reported in (Heacock et al, 2004).

Table 3-2. FISH labeling to identify chromosome ends present in anaphase bridges from *ctc1-1* mutants.

Chromosome Arms	Probe (BAC)	Bridges with Signal	Bridges Observed
All but 4R	9 BAC mix	20[a]	23
1L	F6F3	6	21
1R	F516	5	22
2R	F11L15	3	10
3R	F16M2	5	29
4R	F6N15	6	32
5L	F7J8	7	29
5R	K919	1	8
4R, 2L	25S rDNA	1	7

[a] All but four signals were doublet. Cases in which the signal was a doublet are counted as one signal.

Telomere fusion PCR confirmed end-to-end chromosome fusion. Abundant telomere fusion products were generated from *ctc1-1* homozygous plants, but not from heterozygous or wild type siblings (Fig 3-6D and data not shown). Sequence analysis of 27 cloned fusion junctions failed to detect joining events involving direct fusion of telomere repeats. Instead, telomere-subtelomere fusions (14%) and subtelomere-subtelomere fusions (86%) were recovered (Fig 3-6D), which were characterized by extensive loss of subtelomere sequences (792 bp average loss). In contrast, in G9 *tert* mutants, telomere-subtelomere fusions are the most prevalent (78%), and the average loss of subtelomeric DNA sequences is only 290 bp (Heacock et al, 2004). Thus, chromosome ends are subjected to dramatic DNA loss prior to fusion in *ctc1* mutants.

CTC1 and STN1 act in the same genetic pathway for chromosome end protection

Since the rapid telomere-uncapping phenotype associated with loss of AtCTC1 is remarkably similar to AtSTN1 deficiency (Song et al, 2008), we asked whether the two proteins act in the same genetic pathway for chromosome end protection. Plants heterozygous for *ctc1-1* were crossed to *stn1-1* heterozygotes and F1 progeny were self-pollinated to generate homozygous *ctc1-1 stn1-1* mutants, and their *ctc1-1* and *stn1-1* single mutant siblings. Double *ctc1 stn1* mutants were viable, and the severity of morphological defects was similar to the single mutants (Fig 3-7A).

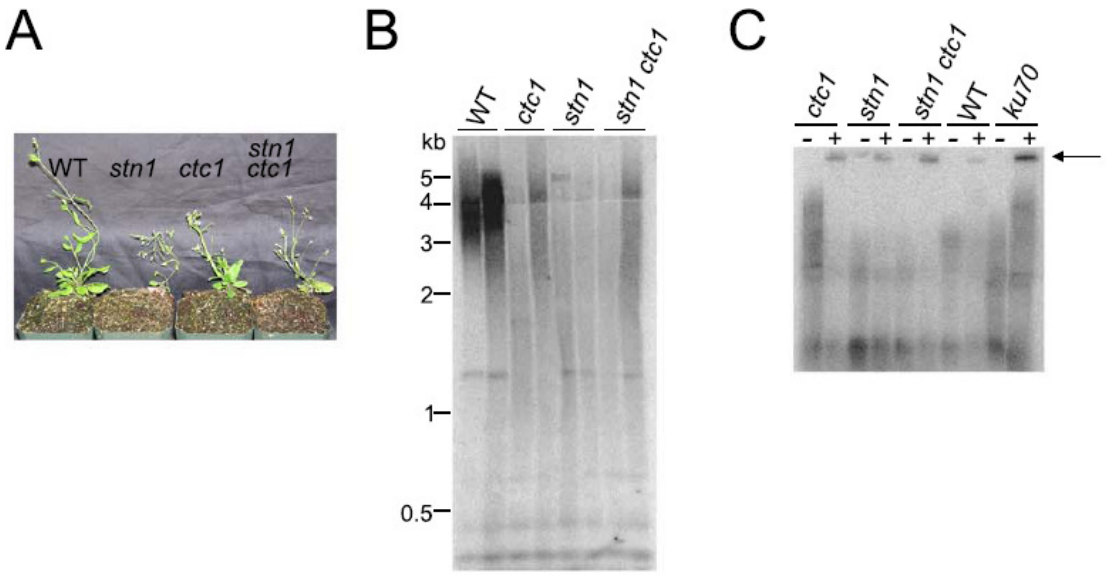


Fig 3-7. Morphological and telomere phenotypes in *ctc1-1 stn1-1* double mutants. (**A**) Morphological and developmental defects in *ctc1-1 stn1-1* double mutants and their *ctc1-1* and *stn1-1* siblings. (**B**) TRF analysis of *ctc1-1 stn1-1*, *ctc1-1* and *stn1-1* siblings. (**C**) T-circle amplification of DNA extracted from *ctc1-1 stn1-1*, *ctc1-1* and *stn1-1* siblings. All panels show progeny of a single parent heterozygous for both *ctc1-1* and *stn1-1.*

TRF analysis and PETRA revealed the same heterogeneous, shortened telomere profile in double mutants as in the *ctc1* or *stn1* single mutants (Fig 3-8A and Fig 3-7B). Similarly, G-overhang signal intensity and the level of ECTC were comparable, implying that double *ctc1-1 stn1-1* mutants did not undergo additional telomeric DNA depletion or increased telomere recombination (Fig 3-8B; Fig 3-7C). Finally, the frequency of anaphase bridges was similar in double mutants and their *ctc1* and *stn1* siblings (Table 3-3). Altogether these findings indicate that *AtCTC1* and *AtSTN1* act in the same pathway for chromosome end protection.

We looked for evidence of a physical association between AtCTC1 and AtSTN1 proteins. Full length AtSTN1 and truncation fragments of AtCTC1 were expressed in rabbit reticulocyte lysate as T7-tagged proteins or radiolabeled with ^{35}S methionine. Immunoprecipitation experiments showed no interaction between AtSTN1 and fragments A-CTC1 or D-CTC1. However, AtSTN1 bound the B-CTC1 and C-CTC1 fragments in reciprocal immunoprecipitation asays (Fig 3-8C). These data indicate that AtSTN1 and AtCTC1 directly interact *in vitro* and hence may also associate with each other *in vitro*.

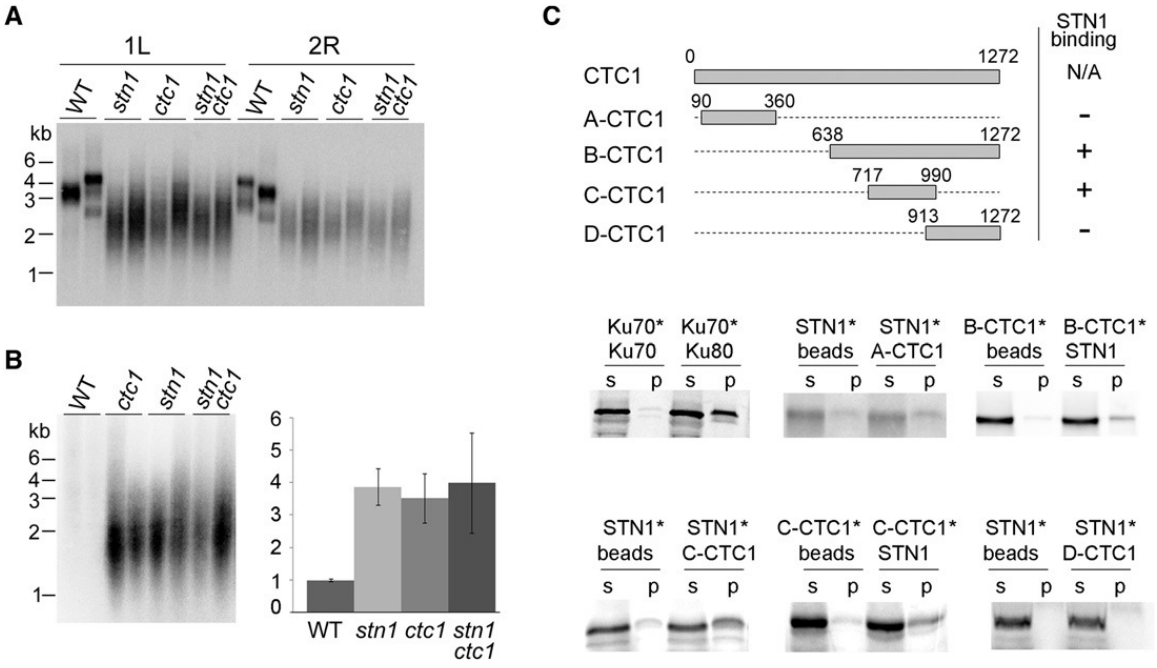


Fig 3-8. *AtCTC1* and *AtSTN1* function in the same genetic pathway for chromosome end protection and physically interact *in vitro*. (**A**) PETRA analysis of telomere length with DNA from *ctc1-1 stn1-1* double mutants, and their *ctc1-1*, *stn1-1*, and wild-type siblings. (**B**) G-overhang analysis using in-gel hybridization. Native gel and quantification results (the average of six independent experiments ± SD) are shown. p≤0.005 for all mutant samples compared to wild-type, and p≥0.4 for mutant samples compared to each other. In (**A**) and (**B**), all progeny were segregated from a double heterozygous *ctc1-1 stn1-1* parent. Blots were hybridized with a radiolabeled telomeric DNA probe. (**C**) Top, schematic of the full-length AtCTC1 protein and its truncation derivatives. AtCTC1 fragments that bind AtSTN1 are indicated. Bottom, coimmunoprecipitation experiments conducted with recombinant full-length AtSTN1 and truncated AtCTC1 fragments A-D. Asterisks indicate ^{35}S-methionine-labeled protein; the unlabeled protein was T7 tagged. S, supernatant; P, pellet. KU70-KU80 interaction was the positive control.

Table 3-3. Frequency of anaphase bridges in *ctc-1- stn1-1* double mutants and their wild type, *ctc1-1* and *stn1-1* siblings.

Genotype	# of analyzed pistils	# of anaphases		% anaphase bridges
		with bridges	total scored	
WT	1	2	207	1
ctc1-1	1	39	184	21
	2	74	273	27
stn1-1	1	42	213	20
	2	30	202	15
ctc1-1 stn1-1	1	28	234	12
	2	51	287	18

Genome instability in human cells depleted of CTC1

TBLASTN and EST database searches revealed CTC1 homologs in a wide range of plant species, while searches using PSI-BLAST and HHpred uncovered putative CTC1 orthologs in many vertebrates. Although the putative plant and animal orthologs exhibited considerable sequence divergence, a global profile-profile alignment indicated that the secondary structures had similarity throughout the length of the protein. Further analysis indicated that the C-terminal domain of human and *Arabidopsis* CTC1 shows homology to OB-fold regions from RPA orthologs, while the N-terminal domain may contain an OB-fold that is distantly related to OB2 from POT1 (Fig 3-9A, Fig 3-10).

Interestingly, the mammalian ortholog of CTC1 is identical to one subunit of Alpha Accessory Factor (AAF-132) while the second subunit of AAF (AAF-44, also known as OBFC1) corresponds to the mammalian ortholog of Stn1 (Casteel et al, 2009b; Martin et al, 2007). AAF is a heterodimeric protein that was originally identified as a factor that stimulates Pol α-primase. It was subsequently shown to enhance Pol α-primase association with ssDNA allowing the enzyme to prime and extend DNA in a reiterative fashion without falling off the DNA template (Goulian & Heard, 1990). Genes encoding the two subunits of AAF were identified recently and AAF-44 was predicted to contain OB-folds resembling those from RPA32 (Casteel et al, 2009).

Fig 3-9. Depletion of human CTC1 causes genomic instability and sudden telomere loss. (**A**) Alignment of potential OB folds in *Arabidopsis* and human CTC1 with OB-fold domains from POT1 and RPA. Dotted lines connect homologous domains. Dark shading, OB-fold homologies predicted by multiple approaches; light shading, homologies predicted by a single method. Dark rectangle within OB5 of HsCTC1 indicates putative Zn finger, which is present in human and archaeal RPAs. MjRPA, archeal RPA from *Methanococcus jannaschii*; HsRPA1, human RPA70. (**B**) Knockdown of CTC1 mRNA in HeLa cells at indicated times after the second transfection. Values are the mean of five independent experiments ± SEM. The percent knockdown is relative to the mock transfection, which was set at 100%. NC, nonsilencing control; Mock, transfection reagent alone. (**C** and **D**) Chromatin bridges and γH2AX staining after CTC1 knockdown in HeLa cells. (**C**) DAPI staining (blue) shows bridges between interphase cells, γH2AX (red) shows DNA damage foci. (**D**) Frequency of chromatin bridges. Mean of three independent experiments ± SEM, asterisks indicate significance levels (*$p < 0.05$; **$p < 0.01$) from one-tailed Student's t test. (**E** and **F**) Telomere FISH showing signal-free ends 48 hr after CTC1 knockdown in HeLa cells. (**E**) Representative metaphase spreads hybridized with Cy3-OO-$(TTAGGG)_3$ PNA probe. The top panels show magnified view of selected chromosomes. (**F**) Percent of chromosome ends that lack a telomeric DNA signal after treatment with nonsilencing control or CTC1 siRNA. Asterisks indicate significance of the increase in signal-free ends; significance levels are depicted as in (**D**).

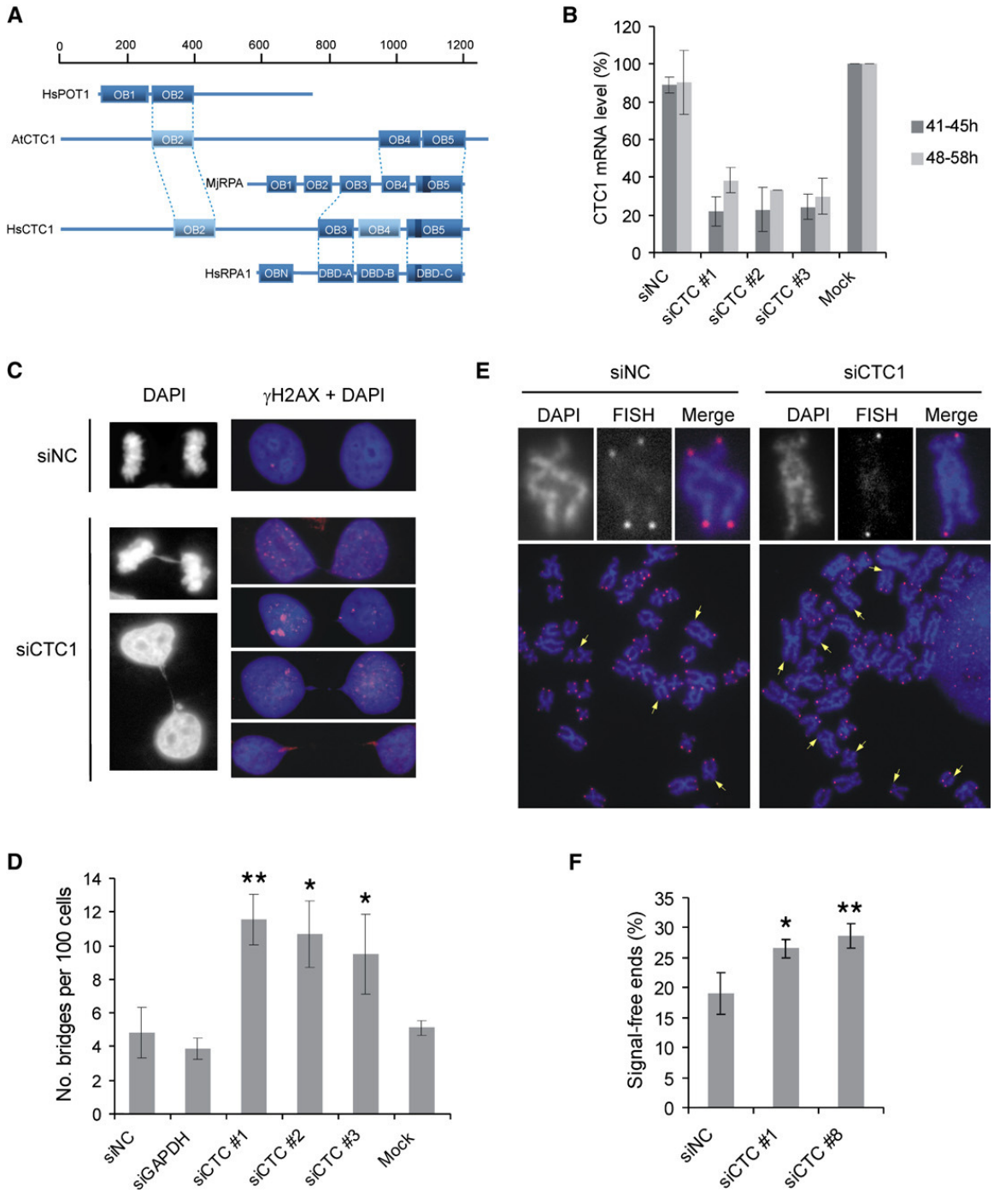
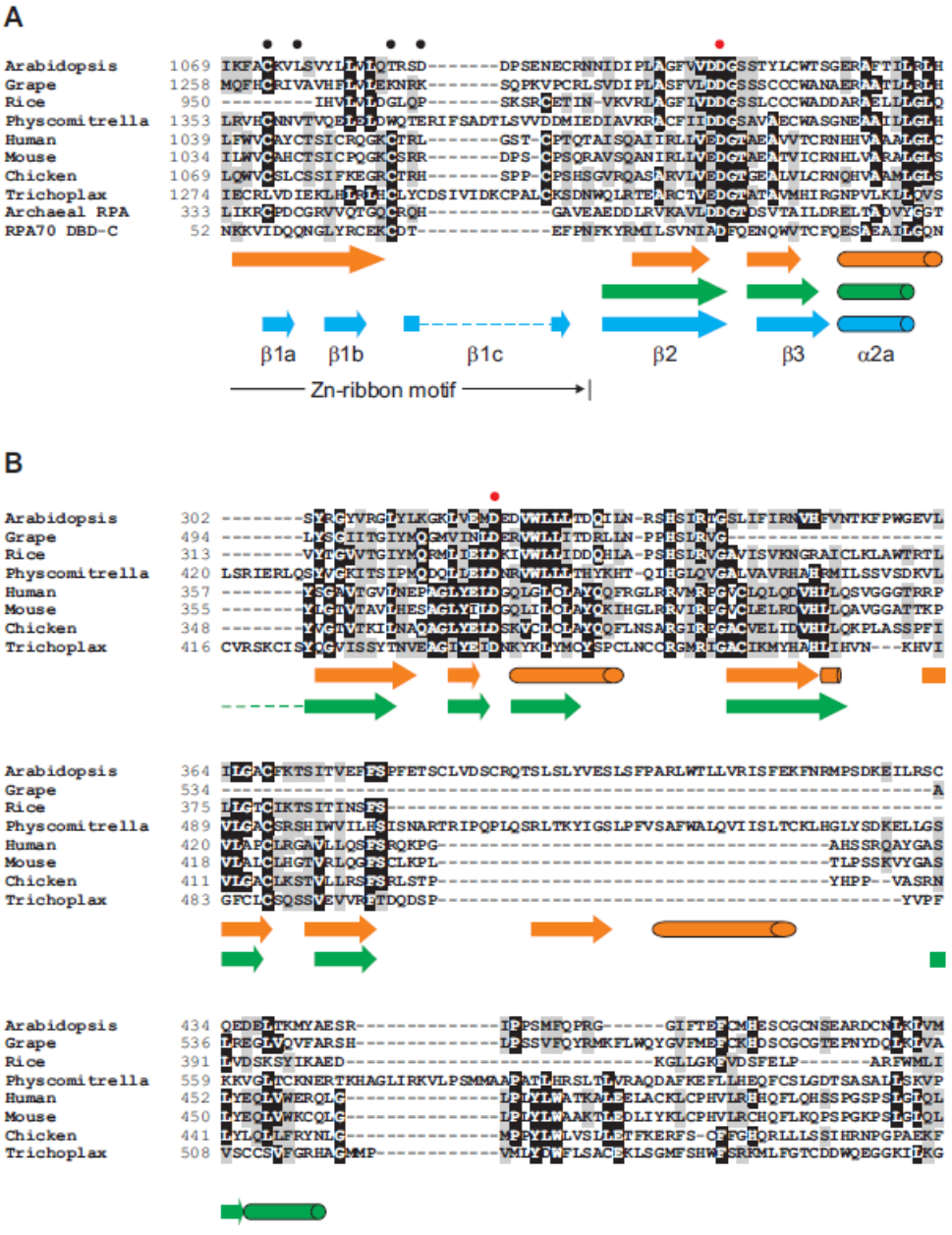

A
0
200
400
600
800
1000
1200
HsPOT1
OB1
OB2
AtCTC1
OB2
OB4
OB5
MjRPA
OB1
OB2
OB3
OB4
OB5
HsCTC1
OB2
OB3
OB4
OB5
HsRPA1
OBN
DBD-A
DBD-B
DBD-C
B
CTC1 mRNA level (%)
120
100
80
60
40
20
0
41-45h
48-58h
siNC
siCTC #1
siCTC #2
siCTC #3
Mock
C
DAPI
γH2AX + DAPI
siNC
siCTC1
E
siNC
siCTC1
DAPI
FISH
Merge
DAPI
FISH
Merge
D
No. bridges per 100 cells
14
12
10
8
6
4
2
0
**
*
*
siNC
siGAPDH
siCTC #1
siCTC #2
siCTC #3
Mock
F
Signal-free ends (%)
35
30
25
20
15
10
5
0
*
**
siNC
siCTC #1
siCTC #8

Fig 3-10. Sequence alignments showing conservation between CTC1 homologs. (**A**) Sequence alignment of the C-terminal region of CTC1 homologs with the homologous region of archeal (*Methanocaldococcus jannaschii*) RPA and human RPA70 (DBDC). Secondary structure elements were taken from the crystal structure for HsRPA (shown in blue) and were predicted with PSIPRED for AtCTC1 (orange) and HsCTC1 (green). Arrows and cylinders represent β-sheets and α-helices. Red dot indicates aspartic acid that is conserved in the second β-sheet of OB-folds. Black dots indicate conserved residues in the $CX_2CX_8CX_2H$ Zn finger motif present in archael RPAs and chicken CTC1. (**B**) Alignment of the N-terminal region that is best conserved between CTC1 homologs. Secondary structure predictions (Orange; AtCTC1, green, HsCTC1) suggest the presence of an OB fold that is distantly related to POT1 OB2.


A
Arabidopsis 1069
Grape 1258
Rice 950
Physcomitrella 1353
Human 1039
Mouse 1034
Chicken 1069
Trichoplax 1274
Archaeal RPA 333
RPA70 DBD-C 52
β1a
β1b
β1c
β2
β3
α2a
Zn-ribbon motif
B
Arabidopsis 302
Grape 494
Rice 313
Physcomitrella 420
Human 357
Mouse 355
Chicken 348
Trichoplax 416
Arabidopsis 364
Grape 534
Rice 375
Physcomitrella 489
Human 420
Mouse 418
Chicken 411
Trichoplax 483
Arabidopsis 434
Grape 536
Rice 391
Physcomitrella 559
Human 452
Mouse 450
Chicken 441
Trichoplax 508

To investigate whether the human CTC1 protein is important for telomere integrity, we examined the effect of knocking down CTC1 expression in human cells. HeLa and MCF7 cells were subject to two rounds of transfection with individual siRNAs and the level of CTC1 transcript was analyzed by quantitative real-time RT-PCR. Out of eight siRNAs tested, six routinely gave a 60-80% knockdown (Figs 3-9B and 3-11C, data not shown). The effect of CTC1 knockdown was monitored after the cells had recovered from the dual transfection.

FACS analysis of DNA content revealed that CTC1 knockdown affected cell cycle progression. MCF7 cultures showed an accumulation of cells in G1 and a decrease in the S/G2 fraction (Fig 3-12A). Microscopy of DAPI stained cells revealed that CTC1 knockdown perturbed chromosome segregation. For HeLa cells, we observed an ~ 2-fold increase in the frequency with which interphase cells remained connected by chromatin bridges (Fig 3-9C, 3-9D and 3-12B). Although the incidence of chromatin bridges was lower in MCF7 cells, there was an increase in the number of cells with micronuclei (Fig 3-12C). These micronuclei probably reflect anaphase or interphase bridges that were later resolved (Hoffelder et al, 2004). We were unable to determine whether CTC1 knockdown causes an increase in anaphase bridges as the frequency of mitotic cells was too low. However, the cut-like phenotype with interphase bridges is similar to what was observed after POT1 knockdown in HeLa cells (Veldman et al, 2004), suggesting that like *Arabidopsis* CTC1, human CTC1 is needed to prevent chromosome fusions.

Fig 3-11. Deregulation of the G-strand overhang after CTC1 knockdown in MCF7 cells. (**A**) In-gel hybridization of $(CCCTAA)_4$ probe to telomeric restriction fragments under native (upper panel) or denaturing (lower panel) conditions. +ExoI, DNA samples were treated with Exonuclease I prior to restriction digestion. (**B** and **C**) Quantification of G-strand signal (**B**) or CTC1 mRNA depletion (**C**) for experiment shown in (**A**). Change in G-strand signal or CTC1 mRNA level is shown relative to the mock transfection. (**D**) Mean change in G-strand signal after CTC1 knockdown. Data are from three independent experiments ± SEM; p values are from one-tailed Student's t test. (**E**) Mean change in CTC1 mRNA level for experiments shown in (**D**). See caption to Fig 3-9B for details.

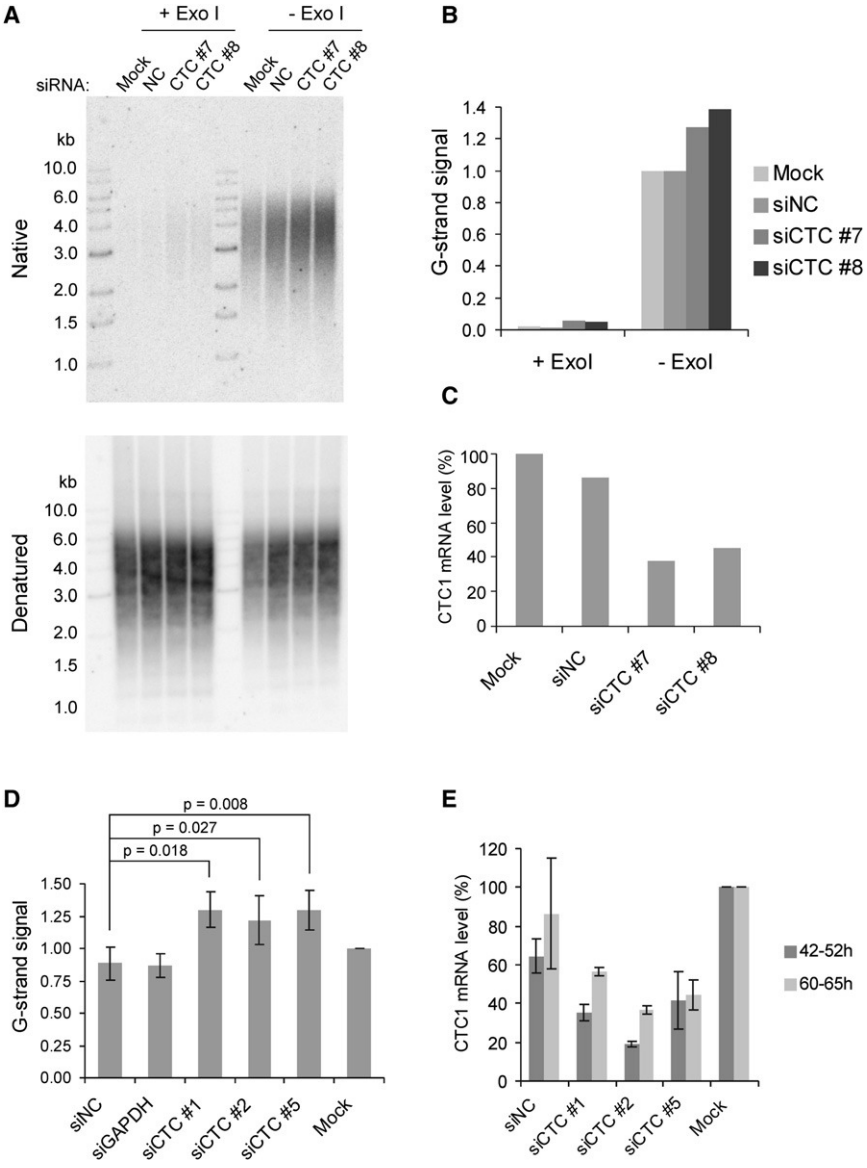
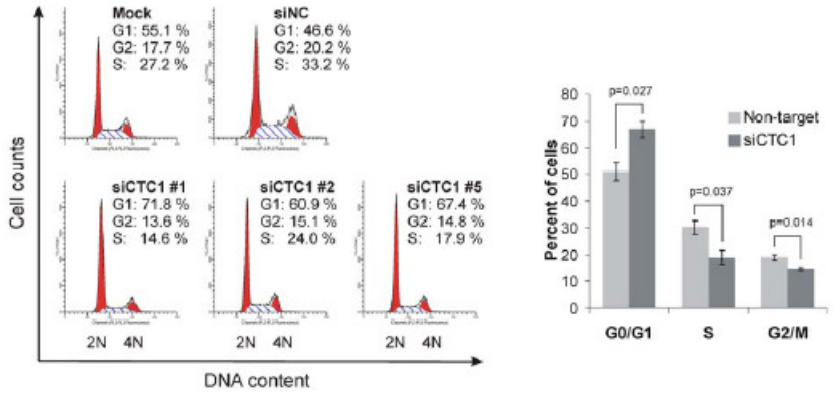

A
+ Exo I
- Exo I
siRNA:
Mock
NC
CTC #7
CTC #8
Mock
NC
CTC #7
CTC #8
kb
10.0
6.0
4.0
3.0
2.0
1.5
1.0
Native
Denatured
B
G-strand signal
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Mock
siNC
siCTC #7
siCTC #8
+ ExoI
- ExoI
C
CTC1 mRNA level (%)
100
80
60
40
20
0
Mock
siNC
siCTC #7
siCTC #8
D
p = 0.008
p = 0.027
p = 0.018
G-strand signal
1.50
1.25
1.00
0.75
0.50
0.25
0.00
siNC
siGAPDH
siCTC #1
siCTC #2
siCTC #5
Mock
E
CTC1 mRNA level (%)
120
100
80
60
40
20
0
42-52h
60-65h
siNC
siCTC #1
siCTC #2
siCTC #5
Mock

Fig 3-12. Effects of CTC1 knockdown in human cells. (**A**) FACS analysis showing accumulation of MCF7 cells in G1 at 64 hrs after treatment with CTC1 siRNA. NC, non-silencing control RNA; Mock, transfection reagent alone. The percent of cells in each phase of the cell cycle was determined using ModFit LT (Verity Software). The graph on the right shows the mean percentage of cells at each stage ± SEM. (**B**) Mean number (± SEM) of interphase chromatin bridges in HeLa cells after treatment with CTC1 siRNA. GAPDH, siRNA to GAPDH. (**C**) Micronuclei in MCF7 cells 60 hrs after treatment with CTC1 siRNA. Nuclei are stained with DAPI (blue) and antibody to γH2AX (red). (**D**) Data from 3 separate telomere FISH experiments showing the number of chromosome ends with or without FISH signals and percent of residual CTC1 mRNA.

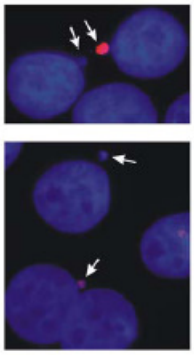
A



B

siRNA	Mean no. of bridges	SEM	Total cells analyzed
siNC	4.80	1.49	232
siGAPDH	3.87	0.63	245
siCTC #1	11.56	1.52	339
siCTC #2	10.67	2.01	222
siCTC #3	9.50	2.34	246
Mock	5.11	0.44	185

C



D

	siRNA	CTC1 mRNA	Total ends	FISH signals	Signal-free	Fraction	SEM	t-test
HeLa exp.1	siNC	100%	1151	932	219	0.190	0.035	
	siCTC #1	32%	1084	796	288	0.266	0.016	p=0.033
	siCTC #7	25%	1400	1000	400	0.286	0.021	p=0.005
HeLa exp.2	siNC	100%	1736	1368	368	0.212	0.017	
	siCTC #1	49%	1794	1313	481	0.268	0.016	p=0.184
	siCTC #7	33%	1857	1291	566	0.305	0.010	p=0.001
293T	siNC	100%	1115	967	148	0.133	0.019	
	siCTC1 #7	50%	1033	850	183	0.177	0.025	p=0.147
	siCTC1 #8	51%	1122	889	233	0.208	0.011	p=0.003

To determine whether the defects in chromosome segregation led to a DNA damage response, we looked for the appearance of γH2AX foci. Treatment with CTC1 siRNA caused an increase in foci in both HeLa and MCF7 cells. These foci were fewer in number and larger than the foci observed after UV irradiation. Moreover, they persisted for the duration of the knockdown whereas UV-induced foci were resolved after a few hours (data not shown). We looked for co-localization of γH2AX and TRF2 staining but this was not readily apparent (data not shown) suggesting that either the DNA damage was not telomeric or that disruption of CTC1 results in complete loss of the telomeric tract from a subset of telomeres. Overall our results indicate that loss of human CTC1 causes a DNA damage response and genome instability.

Depletion of human CTC1 alters G-overhang structure and results in the accumulation of signal-free ends

To determine whether CTC1 knockdown has a direct effect on telomere structure, we used non-denaturing in-gel hybridization to examine the status of the G-overhang. CTC1 depletion caused a modest but consistent increase in ss G-strand DNA in both HeLa and MCF7 cells (Fig 3-11 and data not shown). In MCF7 cells, the G-strand signal increased by 33%-41% relative to the non-silencing control siRNA (Fig 3-11). This increase was statistically significant. Treatment with Exo1 removed essentially all the G-strand signal from the control DNAs, but a small amount remained in the samples from CTC1 depleted cells (Fig 3-11A). Thus, removal of CTC1 causes an increase in G-overhang length and may also result in internal regions of ss G-strand DNA.

Given the failure of the γH2AX foci to co-localize with TRF2 after CTC1 knockdown, we analyzed metaphase spreads to determine whether depletion of CTC1

lead to sporadic telomere loss. Metaphase spreads were prepared from siRNA-treated HeLa and 293T cells and hybridized with Cy3-labeled $(TTAGGG)_3$ PNA probe. Subsequent analysis of individual chromosomes revealed an increase in signal free ends (Fig 3-9E and 3-9F). This increase was statistically significant in four out of six trials, with the greatest frequency of signal free ends correlating with the deepest CTC1 knockdown (Fig 3-12D). We therefore conclude that like *Arabidopsis* CTC1, human CTC1 is required to maintain telomere integrity.

Discussion

Although overall telomere architecture and the general mechanism of telomere replication are well conserved, telomere protein sequence and composition have evolved rapidly (Bianchi & Shore, 2008; Linger & Price, 2009). The resulting divergence has complicated telomere protein identification so it is still unclear whether the full complement of dedicated telomere proteins is known for any organism. It is also unclear whether additional telomere-specific factors are required to address the unique problems associated with replicating the DNA terminus. In this study we employed a genetic approach to uncover CTC1, a new telomere protein that is required for genome integrity in multicellular eukaryotes. The *CTC1* gene is predicted to encode a large protein (142 kDa in *Arabidopsis* and 134.5 kDa in humans) that has orthologs dispersed widely throughout the plant and animal kingdoms. Both *Arabidopsis* and human CTC1 interact with STN1, an ortholog of *S. cerevisiae* Stn1 that was recently found at *Arabidopsis* and human telomeres (this study; [Casteel et al, 2009; Dejardin & Kingston, 2009; Song et al, 2008]). Moreover, we discovered that the mammalian CTC1/STN1 complex corresponds to the recently identified DNA polymerase AAF, previously shown

to stimulate Pol α-primase (Casteel et al, 2009). Thus, CTC1 appears to be a novel protein that is required for telomere end protection and/or telomere replication.

In *Arabidopsis*, the phenotype of a *ctc1* null mutant reflects rapid and catastrophic deprotection of all chromosome ends. Telomere tracts are grossly deregulated in both length and terminal architecture and are subjected to increased recombination and extensive loss of both telomeric and subtelomeric sequences prior to end-to-end fusion. The dramatic effect of CTC1 depletion contrasts with the gradual loss of telomeric DNA in *tert* mutants and the correspondingly later onset of developmental defects (Fitzgerald et al, 1999; Riha et al, 2001). It is striking that plants null for CTC1 are viable, because in other model organisms, loss of telomere-capping proteins activates an ATM or ATR-mediated DNA damage checkpoint and is a lethal event (e.g. loss of *CDC13*, *STN1* or *TEN1* in budding yeast, *STN1*, *TEN1* or *POT1* in fission yeast, and *TRF2* or *POT1* in vertebrates (Churikov & Price, 2008; Grandin et al, 1997; Palm & de Lange, 2008). The extraordinary tolerance of plants to telomere uncapping may reflect a difference in pathways used to monitor genome integrity (Gutierrez, 2005), the partial duplication of the *Arabidopsis* genome, which permits some degree of aneuploidy. In addition, developmental plasticity may mitigate the consequences of genome instability by allowing healthy cells to assume the function of their more severely compromised neighbors.

Depletion of the human CTC1 mRNA revealed a more modest, but significant role for this protein in chromosome end protection. Several cell lines exhibited hallmarks of genome instability such as chromatin bridges, micronuclei and γH2AX staining. Moreover, telomere architecture was perturbed with cells showing an increase in G-overhang signal and sporadic telomere loss. The milder phenotypes associated with

HsCTC1 depletion relative to *Arabidopsis* may reflect the partial knockdown. Plants that are heterozygous for *CTC1* show no deleterious phenotypes, thus only low levels of protein may be needed for telomere protection. This is the case for vertebrate POT1 as the knockdown causes a less severe phenotype than the full gene knockout (Churikov et al, 2006). It is also possible that the function of HsCTC1 only partially overlaps that of AtCTC1. In *Arabidopsis*, POT1 variants seem to be telomerase subunits rather than stable components of the telomere (C. Cifuentes-Rojas, K. Kannan, J. Levy, A.D.L. Nelson, L. Tseng and D.E. Shippen, unpublished data) (Surovtseva et al, 2007). Thus, plant CTC1 may have evolved to function both in telomere end protection and telomere replication. In contrast, mammalian CTC1 may function only in telomere replication.

How CTC1 promotes telomere integrity in multicellular eukaryotes is unknown, but important clues come from recent studies of AAF (HsCTC1/STN1) (Casteel et al, 2009). AAF-44 (HsSTN1) contains an OB-fold that is required for AAF to bind ssDNA and stimulate Pol α-primase activity. Thus, as in the budding yeast Cdc13/Stn1/Ten1 (CST) complex, the mammalian CTC1/STN1 complex binds ssDNA and provides a link to the lagging strand replication machinery. This connection also appears to be conserved in plants, as AtCTC1 physically interacts with both AtSTN1 (this study) and the DNA pol α catalytic subunit (X. Song and D.E. Shippen, unpublished data). These findings raise the possibility that plant and mammalian CTC1 and STN1 are part of a CST complex that, like budding yeast CST, functions in telomere capping and/or coordination of G- and C-strand synthesis during telomere replication. If CTC1 functions in a CST-like complex, we would expect multicellular eukaryotes to possess a Ten1-like protein. Indeed, a putative TEN1 ortholog has been identified in humans (Miyake et al, 2009) and *Arabidopsis* (X. Song, K. Leehy and D.E. Shippen, unpublished data). Like its

counterpart in budding yeast, the *Arabidopsis* TEN1 protein exhibits strong affinity for AtSTN1 *in vitro*.

The observation that both *S. cerevisiae* CST and mammalian CTC1/STN1 (AAF) modulate DNA pol α-primase is particularly striking. In yeast, both Cdc13 and Stn1 interact with Pol α subunits and are proposed to couple telomeric G- and C-strand synthesis (Grossi et al, 2004; Puglisi et al, 2008; Qi & Zakian, 2000). This coupling prevents accumulation of long G-strand overhangs following G-strand extension by telomerase or C-strand resection by nuclease. Previous studies of mammalian CTC1/STN1 (AAF) only explored Pol α-primase stimulation in *vitro* and did not investigate in *vitro* telomeric function or interactions with telomeric DNA (Casteel et al, 2009; Goulian & Heard, 1990). Thus, this work did not indicate whether CTC1/STN1 promotes general DNA replication or telomere replication. Our results reveal a clear role for CTC1/STN1 in telomere maintenance. However, we cannot rule out additional non-telomeric functions. Indeed, the non-telomeric γH2AX staining after CTC1 knockdown suggests a role in DNA replication or repair. One possibility is that mammalian CST acts as a specialized replication/repair factor that is needed to reinitiate DNA synthesis by DNA Pol α if a replication block causes uncoupling of polymerase and helicase activity at the replication fork (Heller & Marians, 2006; Yao & O'Donnell, 2009). Such a function might explain the residual exonuclease-resistant G-strand signal after CTC1 depletion.

Many of the telomere defects observed after CTC1 depletion could be explained by defects in lagging strand replication either at the chromosome terminus or within the telomeric tract. For example, failure to fill in the C-strand following telomerase action or C-strand resection would lead to long G-overhangs. Damage to the G-strand might, in turn, result in telomere loss and/or telomere fusions. Likewise, failure to reinitiate

lagging-strand synthesis after replication fork stalling could lead to loss of large stretches of telomeric DNA and signal-free ends.

Given the role played by the *S. cerevisiae* CST complex, one attractive model for CTC1/STN1 function is that it serves to recruit Pol α-primase to the telomeric G-strand after telomerase action and/or C-strand processing. Pol α appears to be recruited to replication forks by Mcm10, which may in turn interact with the Cdc45/Mcm2-7/GINS replicative helicase (Warren et al, 2008). However, since the G-strand overhang cannot support a conventional replication fork, telomeres appear to require a specialized mechanism to recruit Pol α-primase for C-strand fill in. Further studies will be needed to test this model for CTC1/Stn1 function. Additional work will also be required to determine the extent to which the telomeric function of CTC1/STN1 stems from its role in telomere replication versus a more passive function in G-overhang protection. Perhaps the balance between these activities will differ between organisms. For example, the *Arabidopsis* and *S. cerevisiae* complexes may function in both capacities, while the mammalian complex is specialized for telomere replication.

CHAPTER IV

Arabidopsis TEN1 ASSOCIATES WITH STN1 AND PROTECTS CHROMOSOMAL TERMINI

Summary

In budding yeast, Cdc13/Stn1/Ten1 or CST complex is known to protect chromosome termini. In fission yeast, although a Cdc13-like component is missing, Stn1 and Ten1 homologs are reported to play critical roles in telomere capping. In *Arabidopsis*, STN1 and a novel CST component CTC1 have been recently identified to associate with each other, and to protect chromosome ends from catastrophic degradation and massive end-to-end fusions. Moreover, human STN1 is independently discovered to be associated with hCTC1 as well as a putative TEN1 homolog. It thus appears that CST (CTC1/STN1/TEN1) components are conserved in plants and mammals. Within the CST complex, it is currently unknown whether TEN1 is involved in telomere protection in multicellular eukaryotes. Here we present the identification and characterization of a TEN1 homolog in *Arabidopsis*. AtTEN1 encodes a single OB-fold, which physically associates with AtSTN1 *in vitro*. Plants with a reduced expression level of *AtTEN1* displayed a moderate level of telomere uncapping defects, including telomere length deregulation as well as chromosome end-to-end fusions. Altogether, our data reveal that TEN1 contributes to chromosome end protection in the flowering plant *Arabidopsis*, reinforcing the conservation of CST components among eukaryotes.

Introduction

Eukaryote genome stability relies on telomeric DNA, as well as telomere-associated proteins. In vertebrates, a telomere protein complex called Shelterin is well known for maintaining telomere length and protecting chromosome ends (Palm & de Lange, 2008). Shelterin is composed of six members, including ds telomere binding proteins TRF1 and TRF2, ss telomere binding protein POT1, and three bridging proteins TIN2, RAP1 and TPP1 (Liu et al, 2004a; Palm & de Lange, 2008; Ye et al, 2004b). Studies from fission yeast support that Shelterin-like components are conserved and play essential roles at telomeres. Fission yeast Shelterin includes ds telomere binding protein Taz1 (a TRF homolog), ss telomere binding protein Pot1 as well as Rap1, Tpz1 (a TPP1 homolog), Ccq1 and Poz1 (bridging proteins like TIN2) (Baumann & Cech, 2001; Ferreira & Cooper, 2001; Kanoh & Ishikawa, 2001; Miyoshi et al, 2008).

In contrast to vertebrate and fission yeast, telomeres in budding yeast are protected by a trimeric CST complex, namely Cdc13/Stn1/Ten1. Mutations in any of the CST components result in degradation of the C-strand telomeres, accumulation of single-strand G-rich telomeric DNA, and late S/G2 cell cycle arrest (Garvik et al, 1995; Grandin et al, 2001; Grandin et al, 1997). Cdc13, Stn1 and Ten1 proteins are predicted to harbor one or more OB-folds (Theobald & Wuttke, 2004). Cdc13 interacts with both Stn1 and Ten1, and is strongly associated with telomeric G-overhang (Lin & Zakian, 1996). Besides its role in telomere capping, Cdc13 is also involved in coordination of telomeric G- and C-strand replication by interacting with telomerase RNP and DNA polymerase α (Chandra et al, 2001; Evans & Lundblad, 1999; Pennock et al, 2001).

It is still poorly understood how Stn1 and Ten1 protect telomere ends. ScStn1/Ten1 contacts telomeric G-overhang *in vitro*, but with much lower affinity comparing to Cdc13 (Gao et al, 2007). It is demonstrated that Stn1 and Ten1 share structural similarities to RPA proteins (Gao et al, 2007b; Gelinas et al, 2009). Therefore, CST may form an RPA-like complex, with specificity to ss G-rich telomeric DNA. Recent studies indicate that the role of Stn1 in chromosome end protection can be separated from Cdc13. While Cdc13 is essential for telomere capping, overexpression of Stn1 can bypass the requirement of Cdc13 (Grandin et al, 1997). Moreover, a mutant Stn1 protein, which lacks the Cdc13 interaction domain, could rescue cell viability presumably through a Cdc13-independent mechanism (Petreaca et al, 2007). Mechanism of Ten1 is even less known. Qian et al show that Ten1 stimulates Cdc13 interaction with G-strand telomeric DNA *in vitro* (Qian et al, 2009). In addition, recent studies suggest that Ten1 plays a Cdc13-independent role at telomeres. Several *ten1* mutants display increased exposure of ss G-strand telomeres, even though Cdc13 is correctly localized to telomeres (Xu et al, 2009).

Emerging data suggest that CST components are more conserved in different eukaryotes. Martin et al first showed that SpStn1 and SpTen1 associate with each other and are critical to protect the telomere ends (Martin et al, 2007). Loss of either Stn1 or Ten1 in fission yeast results in catastrophic telomere loss and cell senescence. STN1 homologs have also been identified and characterized in *Arabidopsis* and humans. *Arabidopsis* STN1, like its counterpart in budding and fission yeast, plays an essential role in chromosome end protection (Song et al, 2008). AtSTN1 localizes to telomeres *in vivo*. Plants lacking *STN1* suffer dramatic telomere loss, increased G-overhang signal, massive end-to-end fusions, elevated telomere recombination and gross growth and

developmental defects. In humans, Miyake et al reported that a significant portion of hSTN1 proteins is localized to telomeres, and this localization is consistent throughout the cell cycle (Miyake et al, 2009). Human cell lines with a reduced expression of STN1 exhibit a modest increase in G-overhang signals, but no obvious defects in bulk telomere length (Miyake et al, 2009).

In addition to STN1, recent studies from several labs have identified a novel telomere-associated protein in *Arabidopsis* and mammals, named Conserved Telomere maintenance Component 1 (CTC1) (Miyake et al, 2009; Surovtseva et al, 2009). *Arabidopsis* CTC1 genetically and physically associates with STN1. Plants deficient of *CTC1* display similar telomere defects as *stn1* mutants (Surovtseva et al, 2009). In contrast to the case in *Arabidopsis*, knockdown of *CTC1* in different human cell lines only causes moderate defects, such as sporadic telomere loss on some chromosomes, a modest increase in G-overhang as well as chromatin bridges (Surovtseva et al, 2009). The milder phenotypes associated with HsCTC1 depletion may reflect a partial knockdown. Human CTC1 was independently discovered by Ishikawa's lab through pull-down of hSTN1 followed by mass-spectrometry analysis (Miyake et al, 2009). CTC1/STN1 was also found to correspond to mammalian Alpha Accessory Factor (AAF), which stimulates Pol α activity *in vitro* (Casteel et al, 2009). Thus, CTC1 appears to be a novel component that interacts with STN1 in multicellular eukaryotes. In the hSTN1 pull-down experiment, Miyake et al also identified a putative TEN1 homolog (Miyake et al, 2009). While the exact role of TEN1 in multicellular eukaryotes is still unknown, hTEN1 encodes a small single OB-fold which associates with STN1. Additionally, human STN1/TEN1 facilitates CTC1 to contact ss DNA *in vitro* (Miyake et al, 2009).

Here we show the identification and primary characterization of TEN1 in *Arabidopsis thaliana*. AtTEN1 was identified using PSI-BLAST with hTEN1 as a query. Like its human counterparts, AtTEN1 encodes a small OB-fold containing protein. Interestingly, AtTEN1 associates with AtSTN1 and AtCTC1 *in vitro*. Preliminary data showed that plants with a reduced level of *AtTEN1* expression displayed more heterogeneous bulk telomere tracts, as well as telomere end-to-end fusions. Therefore, AtTEN1 appears to play a conserved role in telomere capping in *Arabidopsis*. Altogether, these findings indicate that CST components play a conserved role at telomeres in *Arabidopsis* and possibly in other multicellular eukaryotes as well.

Materials and methods

Plant materials, plasmid construction, mutants and RT-PCR

The T-DNA insertion line, *ten1-1* (CS839995), was obtained from the *Arabidopsis* Biological Resource Center. The *ten1-1* mutant was genotyped by PCR using primers 5'- CACCCAAAACTGTCATCATTGCTTCA -3' and 5'- GCCATGGCGGCGGTGCAGTTTTTGTAGTTCCAACAAAG -3'. Plants were grown according to the conditions described (Surovtseva et al, 2007). For semi-quantitative RT-PCR, total RNA was extracted from plant tissues using an RNA purification kit (Fisher Scientifics). Reverse transcription was performed using Superscript III reverse transcriptase (Invitrogen) per manufacturer instructions. PCR of *TEN1* cDNA was performed using the above primers, with the following program: 95 ℃ 3 min; 23 cycles of 94 ℃ 20 sec, 55 ℃ 30 sec, 72 ℃ 1 min 30 sec; 72 ℃ 7 min.

In vitro *co-immunoprecipitation*

The experiment was performed as previously described (Karamysheva et al, 2004). For each reaction analyzed, one protein of interest is constructed without T7-tag, while the other protein is tagged with T7. ^{35}S methionine-labeled non-tagged proteins or T7-tagged nonradiolabeled proteins were synthesized in a TNT-coupled rabbit reticulocyte lysate translation system according to manufacturer instruction (Promega). Translation of nonradiolabled proteins was verified in the presence of ^{35}S methionine in a small aliquot from the same master mix. T7-tagged unlabled proteins and untagged radiolabeled proteins were combined and subjected to immunoprecipitation using anti-T7 antibody-conjugated agarose beads (Novagen). Precipitate and supernatant fractions were then analyzed by SDS-PAGE and autoradiography.

TRF and telomere fusion PCR

DNA from individual whole plants was extracted as described (Cocciolone & Cone, 1993). TRF analysis was performed using 50 µg of DNA digested with *Tru1*I (Fermentas) and hybridized with a ^{32}P 5' end–labeled $(T_3AG_3)_4$ oligonucleotide probe (Fitzgerald et al, 1999). The average length of bulk telomeres was determined by Telometric 1.2 (Grant et al, 2001); the range of telomere length was obtained using ImageQuant software. Subtelomeric TRF analysis was performed using 100 µg of DNA digested with *Spe*I and *Pvu*II (New England Biolabs) and hybridized with a 5R probe (Shakirov & Shippen, 2004). Telomere fusion PCR was performed as described (Heacock et al, 2004).

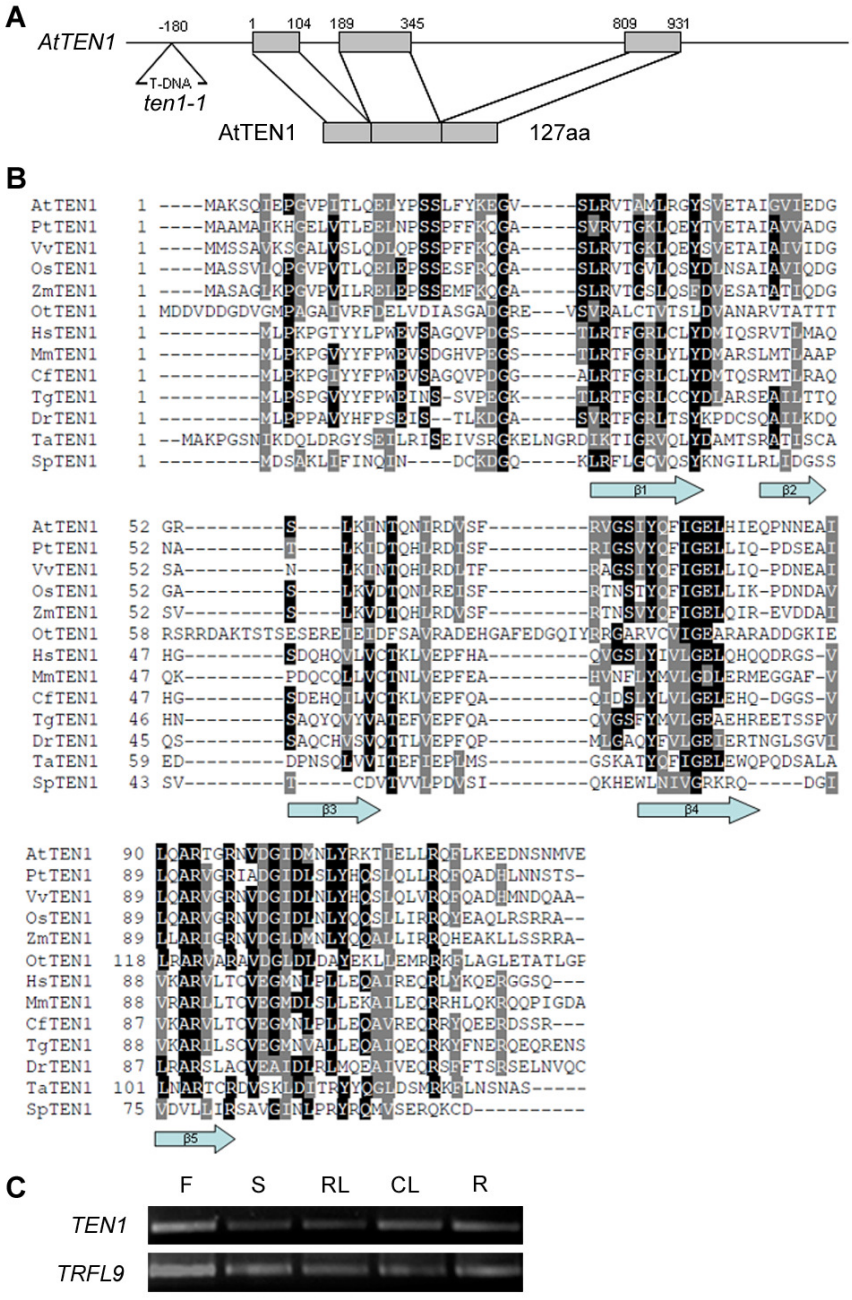
Results

Identification of TEN1 in Arabidopsis

To search for a TEN1 homolog in plants, a PSI-BLAST was employed using the protein sequence of hTEN1 as a query. A previously uncharacterized *Arabidopsis* protein, NP_176022.2, was identified in the second iteration with an E-value of $2e^{-07}$, which was designated as AtTEN1. *AtTEN1* (At1g56260) is a single-copy gene. It consists of three exons (Fig 4-1A) and is predicted to encode a small protein of 127aa (~ 14kD). Database searches revealed potential TEN1 homologs from other plant genomes including poplar, grape, rice, maize and single-celled green algae (Fig 4-1B). Putative TEN1 homologs were also uncovered in a wide variety of other eukaryotes, including fishes, birds, rodents, primates and Trichoplax, a primitive multicellular animal.

AtTEN1 protein assumes a single OB-fold, which shows 23% identity/ 48% similarity to its human homolog. Secondary structure prediction by PSIPRED (McGuffin et al, 2000) indicated that residues within five essential beta strands of the OB-fold (β1-β5) in AtTEN1 share significant similarity to that of SpTEN1 as well as the putative TEN1 proteins from other multicellular eukaryotes (Fig 4-1B). Like *STN1* and *CTC1*, *TEN1* is widely expressed and transcripts could be detected in flowers, stems, rosette leaves, cauline leaves, and roots (Fig 4-1C).

Fig 4-1. Identification of *TEN1* in *Arabidopsis*. (**A**) Schematic of *AtTEN1* gene structure. The T-DNA insertion in *ten1-1* is illustrated. (**B**) Alignment of TEN1 proteins from different eukaryotes. At, *Arabidopsis thaliana*; Pt, *Populus trichocarpa* (poplar); Vv, *Vitis vinifera* (grape); Os, *Oryza sativa* (rice); Zm, *Zea mays* (maize); Ot, *Ostreococcus tauri* (green algae); Hs, *Homo sapiens*; Mm, *Mus musculus*; Cf, *Canis familiaris* (dog); Tg, *Taeniopygia guttata* (zebra finch); Dr, *Danio rerio*; Ta, *Trichoplax adhaerens*; Sp, *Schizosaccharomyces pombe*. The positions of beta-strands of the OB-fold are indicated below the alignment. (**C**) *AtTEN1* is expressed in different plant tissues. RT-PCR of *AtTEN1* coding region is shown. F, flower; S, Stem; RL, rosette leaf; CL, cauline leaf; R, root. *TRFL9* is shown as a loading control.


A
AtTEN1
-180
1
104
189
345
809
931
T-DNA
ten1-1
AtTEN1
127aa
B
AtTEN1 1 ----MAKSQIEPGVPITLQELYPSSLFYKEGV-----SLRVTAMLRGYSVETAIGVIEDG
PtTEN1 1 ----MAAMAIKHGELVTLEEINPSSPFFKQGA-----SVRVTGKLQEYTVETAIAVVADG
VvTEN1 1 ----MMSSAVKSGALVSLQDLQPSSPFFKQGA-----SLRVTGKLQEYSVETAIAIVIDG
OsTEN1 1 ----MASSVLQPGVEVTLQELEPSSESFRQGA-----SLRVTGVLQSYDLNSAIAVIQDG
ZmTEN1 1 ----MASAGLKPGVPVILRELEPSSEMFKQGA-----SLRVTGSLQSFDVESATATIQDG
OtTEN1 1 MDDVDDGDVGMPAGAIVRFDELVDIASGADGRE---VSVRALCTVTSLDVANARVTATTT
HsTEN1 1 ----------MLPKPGTYYLPWEVSAGQVPDGS-----TLRTFGRLCLYDMIQSRVTIMAQ
MmTEN1 1 ----------MLPKPGVYYFPWEVSDGHVPEGS-----TLRTFGRLYLYDMARSLMTIAAP
CfTEN1 1 ----------MLPKPGIYYFPWEVSAGQVPDGG-----ALRTFGRLCLYDMTQSRMTIRAQ
TgTEN1 1 ----------MLPSPGVYYFPWEINS-SVPEGK-----TLRTFGRLCCYDLARSEAILTTQ
DrTEN1 1 ----------MLPPPAVYHFPSEIS--TLKDGA-----SVRTFGRLTSYKPDCSQAILKDQ
TaTEN1 1 --MAKPGSNIKDQLDRGYSEILRISEIVSRGKELNGRDIKTIGRVQLYDAMTSRATISCA
SpTEN1 1 ----------MDSAKLTFINQIN----DCKDGQ-----KLRFLGCVQSYKNGILRLHDGSS
β1
β2
AtTEN1 52 GR---------S----LKINTQNIRDVSF---------RVGSIYQFIGELHIEQPNNEAI
PtTEN1 52 NA---------T----LKIDTQHLRDISF---------RIGSVYQFIGELLIQ-PDSEAI
VvTEN1 52 SA---------N----LKINTQHLRDLTF---------RAGSIYQFIGELLIQ-PDNEAI
OsTEN1 52 GA---------S----LKVDTQNLRDISF---------RTNSIYQFIGELLIK-PDNDAV
ZmTEN1 52 SV---------S----LKVDTQHLRDVSF---------RTNSVYQFIGELQIR-EVDDAI
OtTEN1 58 RSRRDAKTSTSESERETEIDFSAVRADEHGAFEDGQIYRRGARVCVIGEARARADDGKIE
HsTEN1 47 HG---------SDQHQVLVCTKLVEPFHA---------QVGSLYIVLGELQHQQDRGS-V
MmTEN1 47 QK---------PDQCQLLVCTNLVEPFEA---------HVNFLYMVLGDLERMEGGAF-V
CfTEN1 47 HG---------SDEHQILVCTKLVEPFQA---------QIDSLYLVLGELEHQ-DGGS-V
TgTEN1 46 HN---------SAQYQVYVATEFVEPFQA---------QVGSFYMVLGEAEHREETSSPV
DrTEN1 45 QS---------SAQCHVSVQTLVEPFQP---------MLGAQYFVLGEIERTNGLSGVI
TaTEN1 59 ED---------DPNSQLVVITEFIEPLMS---------GSKATYQFIGELEWQPQDSALA
SpTEN1 43 SV---------T-----CDVIVVLPDVSI---------QKHEWLNIVGRKRQ-----DGI
β3
β4
AtTEN1 90 LQARTGRNVDGIDMNLYRKTIELLRQFLKEEDNSNMVE
PtTEN1 89 LQARVGRIADGIDLSLYHQSLQLLRQFQADHLNNSTS-
VvTEN1 89 LQARVGRNVDGIDLNLYHQSLQLVRQFQADHMNDQAA-
OsTEN1 89 LQARVGRNVDGIDLNLYQQSLLIIRQYEAQLRSRRA--
ZmTEN1 89 LLARIGRNVDGLDMNLYQQALLIIRQHEAKLLSSRRA-
OtTEN1 118 LRARVARAVDGLDLDAYEKLLEMRRKFLAGLETATLGF
HsTEN1 88 VKARVLTQVEGMNLPLLEQAIREQRLYKQERGGSQ---
MmTEN1 88 VRARLLTQVEGMDLSLLEKAILEQRRHLQKRQQPIGDA
CfTEN1 87 VKARVLTQVEGMNLPLLEQAVREQRRYQEERDSSR---
TgTEN1 88 VKARILSQVEGMNVALLEQAIQEQRKYFNERQEQRENS
DrTEN1 87 LRARSLAQVEAIDLRLMQEAIVEQRSFFTSRSELNVQC
TaTEN1 101 LNARTQRDVSKLDLTRYYQGLDSMRKFLNSNAS-----
SpTEN1 75 VDVLLIRSAVGINLPRYRQMVSERQKCD----------
β5
C
F
S
RL
CL
R
TEN1
TRFL9

TEN1 interacts with STN1 and CTC1 in vitro

One conserved feature of TEN1 is its ability to interact with STN1 (Grandin et al, 2001; Martin et al, 2007; Miyake et al, 2009). We thus investigated the interaction between AtTEN1 and AtSTN1 using an *in vitro* co-immunoprecipitation (co-IP) approach. In this assay, recombinant proteins were expressed *in vitro* in rabbit reticulocyte lysate. One protein is labeled with ^{35}S-Methionine, while the other is not radiolabeled but contains a T7 tag. Immunoprecipitation of T7-tagged AtTEN1 pulled down radiolabled AtSTN1 protein, indicating a direct interaction between these two proteins (Fig 4-2A, middle). The STN1/TEN1 interaction was confirmed by a reciprocal experiment (Fig 4-2A, right). In addition, a yeast two-hybrid assay provided further evidence for an interaction between TEN1 and STN1 (J.R. Lee and D.E. Shippen, unpublished work).

We next asked whether AtTEN1 interacts with AtCTC1. While an *in vitro* co-IP assay failed to detect any interaction between TEN1 and different truncation domains of CTC1 (Fig 4-3B), preliminary data showed that TEN1 associates with a full length recombinant CTC1 protein expressed and purified from *E. coli* (J.R. Lee and D.E. Shippen unpublished work). Thus, AtTEN1 interacts with STN1 and CTC1 *in vitro*. Therefore, *Arabidopsis* CTC1/STN1/TEN1, similar to their counterparts in budding yeast and humans, are likely to form a trimeric complex and function together in plant cells.

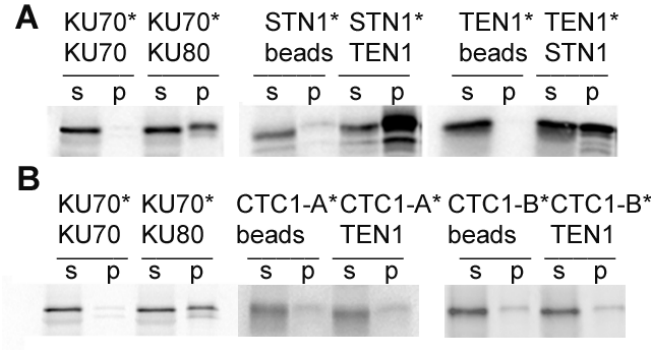


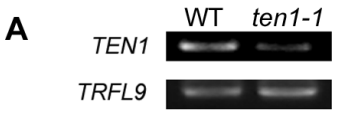
Fig 4-2. *In vitro* co-immunoprecipitation data of TEN1 with STN1 and different pieces of CTC1. (**A**) TEN1 interacts with STN1. (**B**) *In vitro* co-immunoprecipitation failed to detect an interaction between TEN1 and CTC1. CTC1-A corresponds to an N-terminal region of the protein, whereas CTC1-B represents the C-terminal of the protein (Surovtseva et al, 2009). ^{35}S Methionine labeled protein is indicated with an asterisk; The T7-tagged non-radiolabled protein is shown underneath. s, supernatant; p, pellet. KU70/KU80 is shown as a positive control; KU70/KU70 is shown as a negative control.

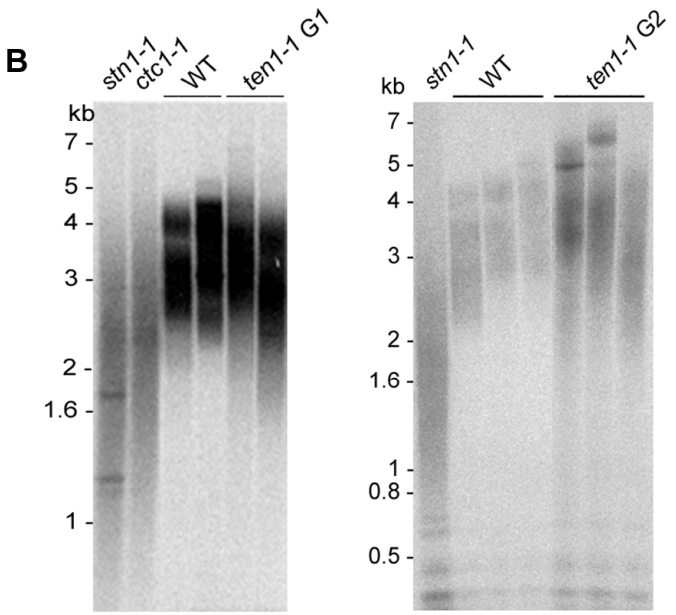
Consequences of TEN1 *depletion*

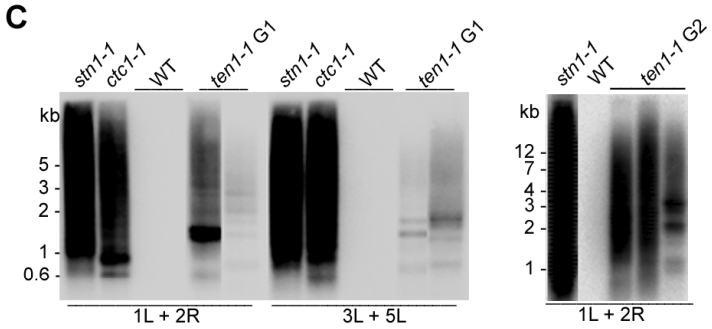
The *ten1-1* mutant harbors a T-DNA inserted in the 5' UTR of *AtTEN1* gene (~ 180bp upstream of the start codon, Fig 4-1A). Semi-quantitative RT-PCR revealed that the expression of *TEN1* mRNA in first generation (G1) plants homozygous for the *ten1-1* insertion was reduced by approximately 50% relative to wild type (Fig 4-3A). These data indicate that although *TEN1* levels are significantly reduced, *ten1-1* mutants are not null for TEN1.

The *ten1-1* mutants are morphologically indistinguishable from wild type plants. To investigate whether *TEN1* depletion affected telomere maintenance, Terminal Restriction Fragment (TRF) analysis was performed to monitor bulk telomere length. Relative to wild type where telomeres range from 2.0 kb to 5.0 kb, telomere tracts were more heterogeneous in G1 *ten1-1* mutants (1.5 kb- 5.6 kb) (Fig 4-3B, left). Such heterogeneity was exacerbated in G2 mutants, where the longest telomeres were approximately 1 kb longer than in wild type and reached ~ 6.0 kb. These data indicate that TEN1 is involved in telomere length regulation. To note, the shortest telomere tracts in G2 *ten1-1* mutants remained the same as in G1 mutants (~ 1.5 kb). It is likely that telomeres below 1.5 kb were detected as DNA damages and subject to chromosome end-to-end fusions.

Fig 4-3. Reduced expression of *TEN1* results in modest telomere defects. (**A**) *AtTEN1* expression is reduced in *ten1-1* mutants by about 50% of WT. RT-PCR of *TEN1* is shown. *TRFL9* is used as a loading control. (**B**) TRF analysis on *ten1-1* G1 and G2 mutants. WT, *stn1-1*, *ctc1-1* mutant were shown for comparison. Two *ten1-1* G1 individuals and three G2 mutants were examined. (**C**) Fusion PCR analysis on *ten1-1* mutants. The *stn1-1* and *ctc1-1* mutant samples were shown as positive controls. Two *ten1-1* individuals and three G2 individuals were examined. The primer pairs employed are indicated. (**D**) Summary of sequence analysis of cloned telomere fusion products in *ten1-1* (G2) mutants. Data for *stn1* (G1) and *ctc1* (G1) were taken from previous studies (Song et al, 2008; Surovtseva et al, 2009).







D

Type of Fusion	*ten1-1* (n=15) G2	*stn1* (n=32) G1[1]	*ctc1* (n=27) G1[2]
Telomere-Telomere	13%	0%	0%
Telomere-Subtelomere	87%	21%	14%
Subtelomere-Subtelomere	0%	79%	86%
Features of Fusion Junctions			
Deletion of subtelomeric sequence (avg)	130 bp	870 bp	N.A.
Telomeric repeat retained at the fusion junction (avg)	280 bp	30 bp	N.A.

[1] Data reported in Song et al 2008.
[2] Data reported in Surovtseva et al 2009.

To answer whether TEN1 is involved in chromosome end protection, telomeric fusion PCR was employed with *ten1-1* samples. A significant amount of fusion PCR products were detected in *ten1-1* mutants, but not their WT siblings (Fig 4-3C). Again, telomere fusions were more abundant in G2 mutants (Fig 4-3C). Sequence analysis of cloned fusion PCR products showed that the majority (87%) of end-joining events in *ten1-1* G2 mutants involved subtelomere-to-telomere fusion (Fig 4-3D). In a few cases, the PCR products corresponded to telomere-to-telomere fusions. In contrast, chromosome fusion junctions primarily reflect subtelomere-to-subtelomere joining in *stn1* mutants (79%) and *ctc1* mutants (86%) (Fig 4-3D) (Song et al, 2008; Surovtseva et al, 2009). The average deletion of subtelomeric DNA in *ten1-1* fusion PCR products was about 130 bp, much less compared to *stn1-1* samples (~ 870 bp) (Fig 4-3D). As we discussed earlier, this moderate but significant telomere deprotection phenotypes may reflect a partial knockdown of the *TEN1* gene. Overall, our data suggest that *Arabidopsis* TEN1, like STN1 and CTC1, is also involved in telomere length regulation and chromosome end protection.

Discussion

The recent discoveries of CTC1/STN1/TEN1 in plants and humans have greatly changed our view on chromosome end maintenance and protection. Data from these studies not only indicate that CST complex is widely conserved from yeasts to multicellular organisms, but also raise interesting questions such as how CST functions at telomeres in addition to the previously identified Shelterin components.

Arabidopsis STN1 and CTC1 have been previously shown to act in the same pathway and protect telomeres from aberrant degradation, end-to-end fusions and

telomere recombination (Song et al, 2008; Surovtseva et al, 2009). Here we presented the identification of a *TEN1* homolog in *Arabidopsis*, which encodes a small OB-fold containing protein. *Arabidopsis* TEN1 protein exhibits strong affinity to STN1; TEN1 also interacts with CTC1 *in vitro*. Mutants with a reduction of *TEN1* expression display a telomere deprotection phenotype, suggesting that TEN1 plays an important role in chromosome end protection in *Arabidopsis*. Thus, our data suggest that CTC/STN1/TEN1 complex, like Cdc13/Stn1/Ten1 in budding yeast, is involved in telomere protection in plants.

The telomere uncapping defects observed in *ctc1*, *stn1* or *ten1* mutants can be explained by different models. One is that CST directly protects against exonuclease attack at the telomere termini. Alternatively or additionally, depletion of CST may cause a failure to recruit pol α-primase to replicate the telomeric C-strand, leading to increased G-overhangs and genome instability. Supporting the latter assumption, mouse CTC1/STN1 corresponds to AAF (Casteel et al, 2009). AAF was previously shown to stimulate Pol α-primase activity *in vitro* by allowing the enzyme to prime and extend DNA in a reiterative fashion (Goulian & Heard, 1990). Interestingly, Cdc13 and Stn1 in budding yeast interact with the catalytic and regulatory subunit of Pol α, respectively (Grossi et al, 2004; Qi & Zakian, 2000). Such interactions are thought to coordinate the G- and C-strand telomere replication and to prevent accumulation of long G-overhangs. Consistently, we found that CTC1 physically associates with the catalytic subunit of Pol α *in vitro* (X. Song, K.A. Boltz and D.E. Shippen, unpublished work). Therefore, CTC1/STN1/TEN1 in multicellular eukaryotes, like budding yeast CST, may also be implicated in coordination of the telomeric G- and C-strand synthesis. Further studies are needed to solve the puzzle of CST in multicellular eukaryotes.

Unlike budding yeast where Cdc13/Stn1/Ten1 is sufficient to protect telomere end, telomere integrity in multicellular eukaryotes also relies on Shelterin components. For example, a crucial Shelterin component, POT1, binds ss G-overhang to protect telomere integrity and to suppress DNA damage response at telomeres (Denchi & de Lange, 2007; Palm & de Lange, 2008). Miyake et al revealed synergistic defects of double knockdown of STN1 and POT1 in human cell lines, indicating that CST and Shelterin components play redundant roles to protect telomeres (Miyake et al, 2009). So far, no direct interaction has been reported between CST and POT1 (Martin et al, 2007; Miyake et al, 2009; Miyoshi et al, 2008). Instead, a recent study suggests that STN1 interacts with TPP1 in humans (Wan et al, 2009). TPP1 is another Shelterin component that interacts with both POT1 and TIN2 (Liu et al, 2004a; Ye et al, 2004a). The STN1/TPP1 interaction suggests a cross talk between CST and Shelterin complexes in multicellular eukaryotes. Besides Shelterin, CST may interact with telomere-related proteins, including components involved in telomere DNA replication and DNA damage repair. Future studies are needed to address how chromosome ends are fully protected and replicated at the presence of both Shelterin complex and CST components in multicellular eukaryotes.

CHAPTER V

PLANTS LACKING CTC1 OR STN1 REQUIRE TELOMERASE AND ATR FOR GENOME STABILITY AND VIABILITY

Summary

In budding yeast, the trimeric complex of Cdc13/Stn1/Ten1 (ScCST) protects chromosome ends from eliciting a DNA damage response and recruits telomerase to maintain telomere length. Recent studies reveal that the CTC1/STN1/TEN1 (CST) complex is conserved in plants and humans, adding to the six-member Shelterin complex in multicellular organisms another dynamic layer of essential telomere-associated proteins. *Arabidopsis* CTC1 and STN1 localize to telomeres *in vivo* and are essential for chromosome end protection. Here we employ a genetic approach to examine the interactions of *Arabidopsis* CTC1 and STN1 with telomerase, KU70 and the DNA damage response kinases, ATM and ATR. Plants doubly deficient of a CST component and the telomerase RNP subunit TERT or POT1a exhibit massive genome instability, and harbor telomere tracts markedly shorter than in either single mutant. Thus, telomerase action is required to stabilize telomere tracts devoid of the CST complex. Our data further indicate that maintenance of single-strand (ss) G-overhang is facilitated by at least two different pathways: one requiring CTC1 and STN1, and a second involving KU. Finally, we demonstrate a dramatic increase in genome instabilityin plants lacking STN1/CTC1 and ATR, but not ATM. This finding indicates that the CST complex protects telomeres from eliciting an ATR-dependent checkpoint response.

Introduction

The essential functions of telomeres are to promote complete replication of the chromosome terminus and to distinguish the natural ends of chromosomes from double-strand (ds) breaks. The protection function of telomeres limits accessibility of the terminus to nucleolytic attack, inappropriate recombination and activation of a DNA damage response. Chromosome end protection is facilitated by a suite of telomere proteins, which include the Cdc13/Stn1/Ten1 (CST) complex in budding yeast and the six-member Shelterin complex in vertebrates.

Cdc13, Stn1 and Ten1 all contain oligonucleotide-oligosaccahride binding folds (OB-folds). The OB-fold of Cdc13 binds to ss G-overhang and therefore delivers Stn1/Ten1 to telomeres. It has been proposed that CST forms a trimer in a manner similar to Replication protein A (RPA) (Gao et al, 2007). CST is essential for cell viability. Loss of any component of CST results in extensive elongation of ss G-strand telomeric DNA, and activation of the Rad9-mediated DNA damage checkpoint leading to G2/M arrest (Grandin et al, 2001; Grandin et al, 1997; Nugent et al, 1996). In addition to chromosome end protection, CST promotes telomeric DNA replication. Telomerase-mediated synthesis of the G-rich telomeric strand is facilitated via a physical interaction between Cdc13 and the Est1 component of the telomerase RNP, while replication of the telomeric C-rich strand is stimulated by CST interactions with DNA polymerase alpha (Evans & Lundblad, 1999; Grossi et al, 2004; Qi & Zakian, 2000).

In vertebrates, Shelterin associates with telomeres through two ds telomere binding proteins, TRF1 and TRF2, a ss telomere binding protein, POT1, and three bridging factors TIN2, RAP1 and TPP1 (Palm & de Lange, 2008). TRF2 and POT1 play pivotal roles in blocking telomeric DNA degradation and end-to-end chromosome

fusions (Celli & de Lange, 2005; Hockemeyer et al, 2006; Palm & de Lange, 2008; van Steensel et al, 1998; Wu et al, 2006). TRF1 and POT1 are implicated in telomere length regulation (Loayza & de Lange, 2003; van Steensel & de Lange, 1997), but whether Shelterin is involved in coordinating telomeric G-strand and C-strand replication is unknown.

Emerging data suggest that telomeres in fission yeast, plants and mammals are protected by both CST and Shelterin proteins. Fission yeast lacking SpStn1 or SpTen1 suffer catastrophic loss of telomeric DNA and chromosome end-to-end fusions (Martin et al, 2007). Characterization of *stn1* mutants in plants and humans also revealed defects in telomere stability. *Arabidopsis stn1* mutants lose both telomeric and subtelomeric DNA and exhibit increased G-overhang signals, high level of chromosome end-to-end fusions, and elevated telomere recombination (Song et al, 2008). Like *Arabidopsis* STN1 (Song et al, 2008), human STN1 localizes to a substantial fraction of telomeres throughout the cell cycle (Miyake et al, 2009). Knock-down of human STN1 results in an increase of ss G-overhang signals, but no change in bulk telomeres. Notably, a synergistic increase in telomere dysfunction-induced foci (TIF) positive cells is observed in human cells lacking both STN1 and the Shelterin component POT1, arguing that CST and Shelterin function in separate pathways to promote telomere integrity (Miyake et al, 2009).

STN1 associates with two additional proteins in *Arabidopsis* and humans, namely the TEN1 ortholog and a novel protein termed Conserved Telomere maintenance Component 1 (CTC1) (Miyake et al, 2009; Surovtseva et al, 2009) Human TEN1 and CTC1 were identified through pull-down of hSTN1 followed by mass spectrometry analysis. BLAST search using hTEN1 as a query revealed that TEN1 is conserved in

plants (X. Song, K. Leehy and D.E. Shippen, unpublished work). AtTEN1 exhibits strong affinity to AtSTN1 *in vitro*. Plants with reduced expression levels of *TEN1* display deregulated telomeres and end-to-end chromosome fusions. In *Arabidopsis*, the novel telomere capping protein CTC1 was discovered through forward genetics. *Arabidopsis* CTC1 genetically and physically interacts with STN1. Plants lacking *CTC1* display severe telomere maintenance defects, increased recombination, and massive end-to-end chromosome fusions (Surovtseva et al, 2009). Similarly, reducd expression of CTC1 in human cells leads to an increase in chromosomes lacking detectable telomeric DNA, increased G-overhang signals, and aberrant chromatin bridges (Surovtseva et al, 2009). Altogether, these data indicate that CTC1, STN1 and TEN1 are essential components of the telomere capping complex in multicellular eukaryotes.

Remarkably, plants null for *STN1* or *CTC1* are viable despite catastrophic telomere loss and rampant genome instability. In other model organisms, mutation of telomere-capping proteins activates an ATM- or ATR-mediated DNA damage checkpoint and is a lethal event (e.g., loss of CDC13, STN1, or TEN1 in budding yeast; STN1, TEN1, or POT1 in fission yeast; and TRF2 or POT1 in vertebrates [Baumann & Cech, 2001; Churikov & Price, 2008; Garvik et al, 1995; Grandin et al, 2001; Grandin et al, 1997; Martin et al, 2007; Palm & de Lange, 2008]). In this study, we exploit the unusual tolerance of *Arabidopsis* toward telomere dysfunction, and investigate the genetic interaction of CST components with telomerase, KU and the DNA damage response kinases ATM and ATR. First, it is demonstrate that telomerase activity is required to stabilize telomere tracts devoid of the CST complex. Second, we show that KU and the CST complex mediate two distinct pathways for G-overhang maintenance. Finally, a dramatic increase in the rate of telomere shortening and genome instability

was observed in plants lacking a CST component and ATR, but not ATM. These results indicate that the CST complex prevents telomeres from eliciting an ATR-dependent checkpoint response, and that ATR plays a more direct role in promoting telomere maintenance. Altogether, these findings underscore the essential telomere capping function of the CST complex and provide new insights into how it promotes genome integrity.

Materials and methods

Plants and material

Plants were grown in chambers with 16 hr photoperiod per day at 22°C. Mutants of *stn1-1*, *ctc1-1*, *tert*, *pot1a* and *ku70* were used for crosses and genotyped as previously described (Song et al, 2008; Surovtseva et al, 2009; Surovtseva et al, 2007).

TRF and PETRA

DNA from whole plants was extracted as described (Cocciolone & Cone, 1993). TRF analysis was performed using 50 µg of DNA digested with *Tru1*I (Fermentas) and hybridized with a ^{32}P 5' end–labeled $(T_3AG_3)_4$ oligonucleotide probe (Fitzgerald et al, 1999). The blots were developed using a Pharos FX Plus Molecular Imager (Bio-Rad); and data were analyzed with Quantity One software (Bio-Rad). Primer extension telomere repeat amplification (PETRA) was performed as described (Heacock et al, 2004; Watson & Shippen, 2007). 2 µg of DNA was used per reaction for telomere extension, followed by PCR amplification. PETRA PCR products were separated on an agarose gel and subjected to Southern blotting using the same telomeric probe mentioned above.

In-gel hybridization

In-gel hybridization was performed as described (Heacock et al, 2007; Song et al, 2008). A ^{32}P 5' end-labeled telomeric C-strand probe $(C_3TA_3)_3$ was used for hybridization. The relative amount of G-overhang signal was quantified as the hybridization signal from the native gel, normalized by an interstitial telomere signal obtained from the same gel under a denaturing condition. The G-overhang signal obtained from wild-type DNA was set to one, and each sample was normalized to this value.

Telomeric circle amplification (TCA)

TCA was performed as described (Zellinger et al, 2007). 75 µg of DNA was digested with *Alu*I (Fermentas) per reaction. Phi 29 polymerase (NEB) was used to amplify from telomeric circles into long ss DNA, which was then separated on a denaturing gel (0.8% agarose, 50 mM NaOH, and 1 mM EDTA, pH 8) at 30 volts for 16 hrs, followed by Southern blotting with a ^{32}P 5' end–labeled $(T_3AG_3)_4$ oligonucleotide probe.

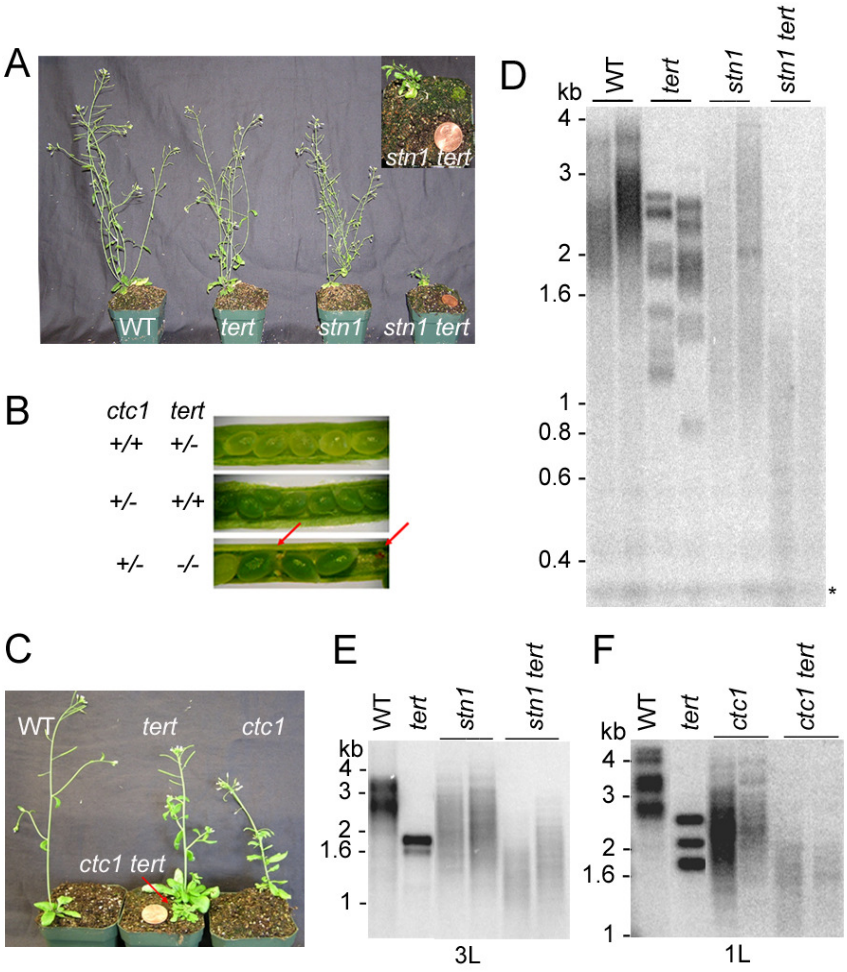
Results

TERT and CST components act synergistically to maintain telomere tracts

In budding yeast, Cdc13 recruits telomerase to telomere ends, whereas Stn1/Ten1 inhibits such process (Chandra et al, 2001; Evans & Lundblad, 1999; Pennock et al, 2001). In order to investigate CST interaction with telomerase in a multicellular eukaryote, we generated plants that are doubly deficient of a CST component and the catalytic subunit of telomerase, TERT. It has been previously shown that a null mutation in *AtTERT* results in slow, but progressive loss of telomeric DNA, and eventually leads to severe genome instability and developmental arrest in the 8^{th} or 9^{th} generation (G8-G9) of the mutant (Riha et al, 2001). In contrast, plants lacking *STN1* or *CTC1* display a much earlier onset of telomere dysfunction with mutants reaching the terminal phenotype by G2 (Song et al, 2008; Surovtseva et al, 2009).

In the progeny of a *stn1 tert* double heterozyote, *tert* mutants were morphologically indistinguishable from wild type plants in G1 as previously observed (Fig 5-1A) (Riha et al, 2001), while *stn1* siblings showed a fasciated phenotype with abnormal fused stems and flowers, as well as irregular phyllotaxy (Fig 5-1A) (Song et al, 2008). Strikingly, *stn1 tert* double mutants were significantly smaller in size than either single mutant (Fig 5-1A). It appears that double deficiency of telomerase and STN1 caused an immediate arrest of plant growth after rosette leave development, leading to plant termini in a single generation (Fig 5-1A).

Fig 5-1. TERT is required to rescue critically shortened telomeres in *stn1* and *ctc1* mutants. (**A**) Severe morphological and developmental defects in *stn1 tert* double mutants compared to *tert* or *stn1* single mutants. An enlarged picture of *stn1 tert* is shown on the right top corner. (**B**) Abandoned embryos in $tert^{-/-}$ $ctc1^{+/-}$ siliques, indicated by red arrows. (**C**) Severe morphological and developmental defects in *ctc1 tert* double mutant survivors compared to *tert* or *ctc1* single mutants. (**D**) TRF analysis of WT, *tert*, *stn1*, and *stn1 tert* mutants. For each genotype, data from two individual sibling plants are shown. The blot was hybridized with a telomeric probe. Molecular weight markers are indicated. (**E**) PETRA analysis of WT, *tert*, *stn1*, and *stn1 tert* mutants. The blot was hybridized with a telomeric probe. Telomere length on the left arm of chromosome 3 (3L) was measured. (**F**) PETRA analysis of WT, *tert*, *ctc1*, and *ctc1 tert* mutants. Telomere length on the left arm of chromosome 1 (1L) was measured.


A
WT
tert
stn1
stn1 tert
B
ctc1
tert
+/+
+/-
+/-
+/+
+/-
-/-
C
WT
tert
ctc1
ctc1 tert
D
kb
WT
tert
stn1
stn1 tert
4
3
2
1.6
1
0.8
0.6
0.4
*
E
WT
tert
stn1
stn1 tert
kb
4
3
2
1.6
1
3L
F
WT
tert
ctc1
ctc1 tert
kb
4
3
2
1.6
1
1L

Likewise, we followed segregants from a *ctc1 tert* double heterozygous parent. Interestingly, *ctc1 tert* mutants were rarely recovered, indicating that a simultaneous loss of *CTC1* and *TERT* is extremely detrimental. Consistent with this conclusion, there was an increase in the number of aborted embryos in the siliques from self-pollinated *ctc1* $^{+/-}$ *tert* $^{-/-}$ plants (20.6%) compared to those in *ctc1* $^{+/+}$ *tert* $^{-/-}$ siblings (2.2%) (Fig 5-1B). Genotyping of the progenies of a *ctc1* $^{+/-}$ *tert* $^{-/-}$ parent uncovered a few *ctc1 tert* double mutants. The infrequent *ctc1 tert* survivors, like *stn1 tert* mutants, arrested at a miniature vegetative state and failed to produce a germline (Fig 5-1C).

Terminal Restriction Fragment (TRF) analysis was performed to monitor bulk telomere length in plants segregated from a *stn1 tert* double heterozygote (Fig 5-1D). The average telomere length was 2.8 kb for wild type plants and 2.3 kb in s*tn1* mutants. *stn1 tert* mutants, however, had much shorter telomere tracts with an average length of only 1.3 kb. The shortest telomeres in these plants were ~ 0.4 kb, while in s*tn1* siblings the shortest telomeres remained above 0.8 kb. To assess telomere length dynamics on individual chromosome arms in *stn1 tert* double mutants, we employed Primer Extension Telomere Repeat Amplification (PETRA) (Fig 5-1E). As expected, telomeres in *tert* mutants were shortened by ~ 250 bp in G1 relative to wild type and were represented by a sharp banding profile. In contrast, telomeres in *stn1* mutants appeared as a broad heterogeneous smear with the shortest telomere tracts trailing down to approximately 0.8 kb. *stn1 tert* mutants displayed a similar heterogeneous telomere profile, but there was a dramatic reduction in the longest telomere tracts (Fig 5-1E). As a result, the average telomere length on the left arm of chromosome 3 (3L) was 2.2 kb in *stn1* and only 1.3 kb in *stn1 tert* mutants. Telomere length on the right arm of chromosome 2 (2R) was also examined and displayed a similar pattern (data not

shown). Based on the morphological and telomere length defects, it is expected that the genome integrity should be severely impaired in *stn1 tert* double mutants. Indeed, we observed abundant telomeric fusion PCR products using DNA sample from the double mutants (data not shown). Unfortunately, the lack of frequently dividing tissues and anaphase cells in *stn1 tert* mutants prohibited us from further quantitative analysis to compare their genome instability levels. Nonetheless, our data indicate that telomerase is required to extend critically shortened telomere tracts in plants lacking STN1.

For the rare *ctc1 tert* survivors, we were unable to perform TRF analysis due to the limited amount of plant material. Instead, PETRA was employed to analyze the telomere length in *ctc1 tert* mutants. As for *stn1 tert* mutants, telomeres in *ctc1 tert* survivors were also heterogeneous with a preferential loss of the longest telomere tracts (Fig 5-1F). Altogether, these results revealed that telomerase and CST components are synergistically required for telomere maintenance and plant viability. In addition, our data suggest that telomerase can access and extend telomeres in plants lacking STN1 or CTC1.

Telomerase activity is required to rescue plants lacking CST

TERT may have additional roles besides telomere replication. Studies in mammalian cells indicate that TERT promotes cell survival and helps to resist different stresses (Ahmed et al, 2008; Armstrong et al, 2005; Kondo et al, 1998; Sarin et al, 2005; Sharma et al, 2003; Zhu et al, 2000). In particular, expression of TERT in human cells extends the cellular lifespan, although it does not prevent telomeres from progressive erosion (Zhu et al, 1999). Thus, at least some part of TERT function is uncoupled from telomerase activity or telomere elongation. To investigate whether TERT or telomerase

enzyme activity is required to rescue plants lacking the CST complex, we examined the fate of *stn1 pot1a* double mutants. *AtPOT1a* acts in the same genetic pathway as *TERT* in telomere maintenance, serving as an accessory component of the telomerase RNP. Telomerase activity levels are reduced by approximately 13-fold in extracts prepared from *pot1a* null mutants, and telomeres in *pot1a* mutants shorten at the same rate as in *tert* mutants (Surovtseva et al, 2007) (A.D.L. Nelson and D.E. Shippen, unpublished work). Recent data reveal that AtPOT1a interacts with TER1, an RNA subunit of telomerase (C. Cifuentes-Rojas and D.E. Shippen, unpublished work). It is proposed that POT1a is involved in telomerase assembly, stabilization and/or recruitment to telomeres .

Similar to *stn1 tert* mutants, *stn1 pot1a* mutants were completely sterile and arrested early in vegetative state (Fig 5-2A). Moreover, PETRA revealed telomere tracts that were shorter and more heterogeneous in *stn1 pot1a* mutants than in *stn1* single mutants. For example, the 2R telomere had an average length of 2.3 kb in *stn1* mutants segregated from the cross, but was only 1.6 kb in *stn1 pot1a* plants. Likewise, the 5L telomere was 1.6 kb on average in *stn1 pot1a* mutants compared to 2.2 kb in *stn1* (Fig 5-2B). We conclude that telomerase activity is required to rescue telomere defects in *stn1* mutants, and the CST complex acts synergistically with telomerase to maintain telomere tracts.

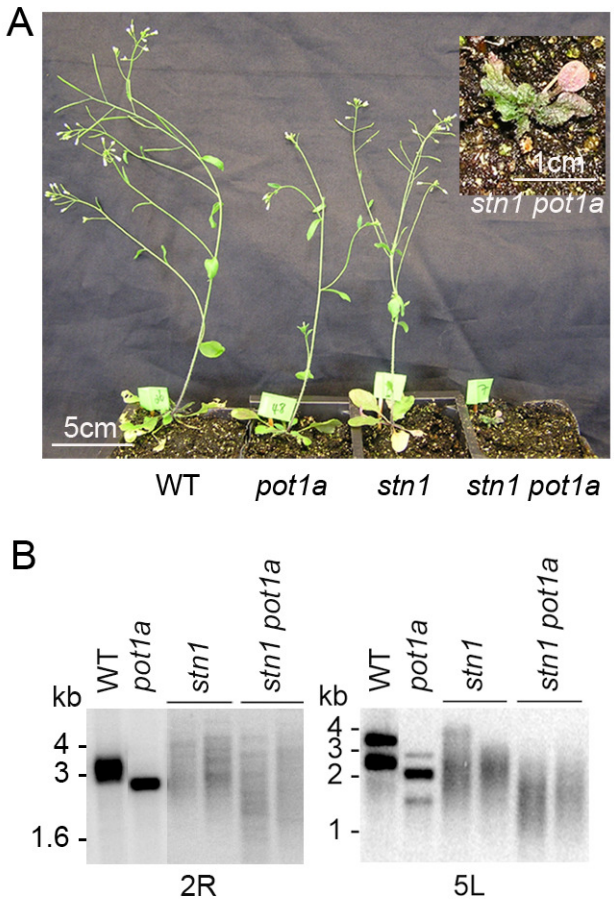


Fig 5-2. Telomerase activity is needed to compensate telomere loss in plants deficient of STN1. (**A**) Severe morphological defects of *stn1 pot1a* double mutants compared to *pot1a* or *stn1* single mutants. An enlarged *stn1 pot1a* picture is shown on the right top corner. (**B**) PETRA analysis of WT, *pot1a*, *stn1*, and *stn1 pot1a* mutants. Telomere length was examined on the chromosome arms indicated.

STN1 and CTC1 act independently from KU to maintain the G-overhang

Extended G-overhangs are associated with impaired cell viability and senescence in yeast and human cells (Grandin et al, 2001; Grandin et al, 1997; Li et al, 2003; Nugent et al, 1996; Stewart et al, 2003). CST components are critical to protect the G-overhang. Loss of Cdc13/Stn1/Ten1 in budding yeast (Garvik et al, 1995; Grandin et al, 2001; Grandin et al, 1997) or CTC1/STN1 in *Arabidopsis* and humans (Miyake et al, 2009; Song et al, 2008; Surovtseva et al, 2009) leads to a significant increase in G-overhang signal. In addition to CST, KU70/80 is implicated in G-overhang maintenance. While deletion of KU in budding and fission yeast leads to shortened telomeres (Baumann & Cech, 2000; Gravel et al, 1998). telomere tracts are grossly elongated in *Arabidopsis ku70* mutants (Riha et al, 2002). Thus, KU acts as a positive regulator of telomere length in yeast and a negative regulator in *Arabidopsis*. However, in both settings, G-overhangs are dramatically extended in a telomerase-independent manner (Gravel et al, 1998; Riha & Shippen, 2003), indicating that KU is required to maintain proper telomere architecture.

To investigate whether CST functions in the same pathway as KU70/KU80 for G-overhang maintenance, we examined telomeres in *stn1 ku70* and *ctc1 ku70* double mutants. *STN1* and *KU70* reside on chromosome 1 within ~ 10 CM to each other, which makes it difficult to obtain *stn1 ku70* mutants from a double heterozygous parent. Therefore, we segregated from a *stn1* $^{+/-}$ *ku70* $^{-/-}$ line (*ku70* G1) to obtain plants lacking both *STN1* and *KU70* in the next generation (*stn1* G1 and *ku70* G2).

As expected, *ku70* mutants did not show any morphological defects (Fig 5-3A) (Riha & Shippen, 2003; Riha et al, 2002). In contrast, *stn1 ku70* (*stn1* G1 and *ku70* G2) plants were arrested early in vegetative growth without producing any seeds (Fig 5-3A). TRF analysis revealed that telomeres in *stn1 ku70* mutants were extremely heterogeneous in length, ranging from ~ 1.5 kb to above 12 kb. Moreover, the telomere length in *stn1 ku70* mutants was dramatically shorter than their *ku70* G2 siblings or *ku70* G1 parent (Fig 5-3B).

In the case of plants deficient in *KU* and *CTC1*, *ctc1 ku70* double mutants can be segregated in a single generation. Like *stn1 ku70* mutants, *ctc1 ku70* plants were also arrested in G1 in a vegetative state without developing any reproductive tissue (Fig 5-4A). TRF analysis showed that telomeres in *ctc1 ku70* mutants were as heterogeneous as, if not more than, those in *ctc1* single mutants (Fig 5-4B). PETRA confirmed the heterogeneity of telomeres on individual chromosome arms in *ctc1 ku70* mutants (Fig 5-4C). While the results are being verified, it appears that some of the longest telomere tracts in *ctc1 ku70* double mutants were longer than in *ctc1* mutants, whereas the shortest telomeres tracts remained the same in both *ctc1* and *ctc1 ku70* mutants (Fig 5-4B & C).

To monitor G-overhangs in these mutants, in gel-hybridization was employed. *stn1 ku70* mutants showed an additive increase in G-overhang signals (6.0±1.7) relative to either single mutant (*stn1* 3.45±0.9, *ku70* 1.6±0.5) (Fig 5-5). Taken together, these data indicate that G-overhangs are maintained by at least two different pathways in *Arabidopsis*: one requiring CTC1/STN1, and the other involving KU70/KU80.

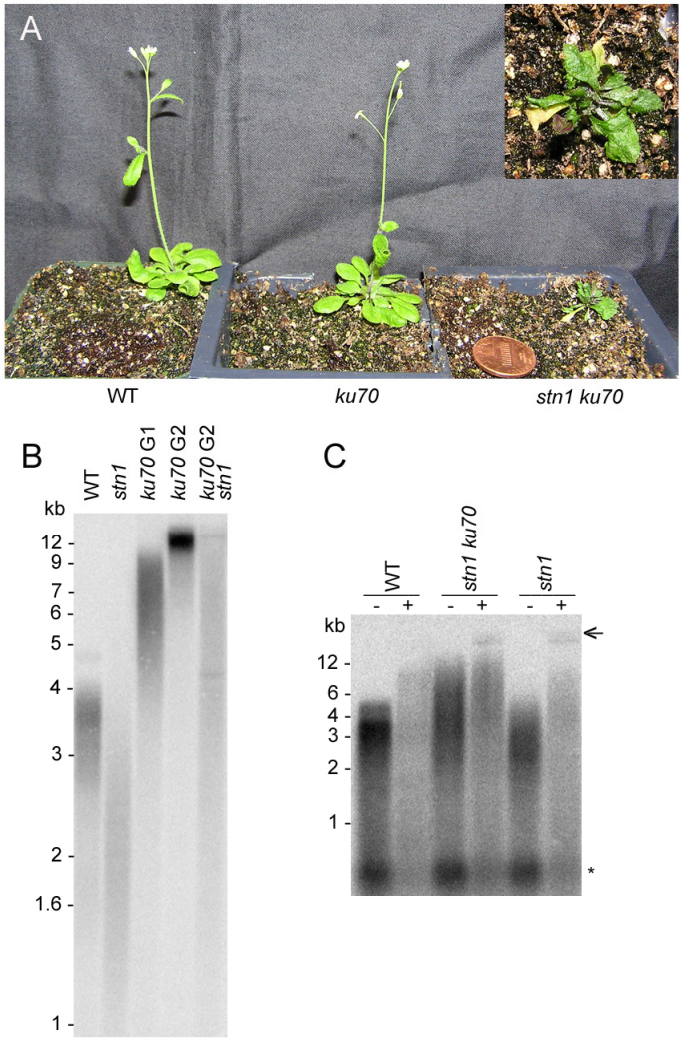


Fig 5-3. Telomere rapid deletion in *stn1 ku70* double mutants. (**A**) Severe morphological and developmental defects in *stn1 ku70* mutants. An enlarged *stn1 ku70* picture is shown on the right top corner. (**B**) TRF analysis of *stn1 ku70* mutant (*ku70* G2) comparing to WT, *stn1*, *ku70 G1* as well as *ku70* G2 mutant. (**C**) TCA of *stn1 ku70* mutant. WT was shown as a negative control. *stn1* sample was used as a positive control. Arrow indicates product from telomeric circles. Asterisk for internal telomere signal as a loading control.

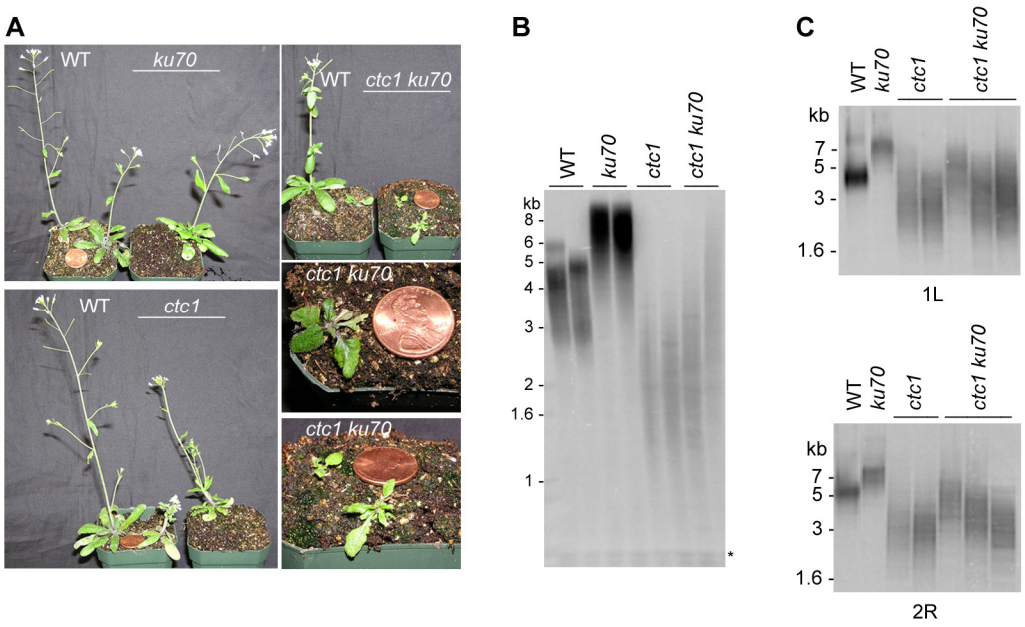


Fig 5-4. KU acts in a different pathway from CTC1 to protect proper G-overhangs. (**A**) Severe morphological and developmental defects in *ctc1 ku70* double mutants compared to *ku70* or *ctc1* single mutants. (**B**) TRF analysis of WT, *ku70*, *ctc1*, and *ctc1 ku70* mutants. For each genotype, data from two individual plants are shown. (**C**) PETRA analysis of WT, *ku70*, *ctc1*, and *ctc1 ku70* mutants. Telomere length on the left arm of chromosome 1 (1L) and the right arm of chromosome 2 (2R) was measured.

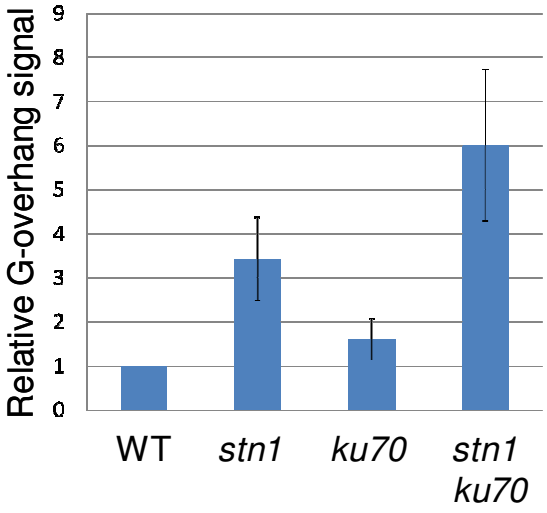


Fig 5-5. *Arabidopsis* STN1 and KU independently protect G-overhang. G-overhang signal of DNA samples from WT, *stn1*, *ku70* and *stn1 ku70* mutants were analyzed using in-gel hybridization. A ^{32}P 5' end-labeled telomeric C-strand probe $(C_3TA_3)_3$ was used for hybridization under native condition. The G-overhang signal obtained from wild-type DNA was set to one, and each sample was normalized to this value.

CST and KU are both required to inhibit telomere recombination

Plants lacking CST and KU share another common telomere defect: elevated levels of extra-chromosomal telomeric circles (ECTC) (Song et al, 2008; Surovtseva et al, 2009; Zellinger et al, 2007), a hallmark of telomeric DNA recombination. Using telomeric circle assay (Zellinger et al, 2007), ECTCs were detected in *stn1 ku70* mutants (Fig 5-3C). Although not quantitative, this result suggests that KU and the CST complex are both required to protect chromosome ends from aberrant recombination, and that telomere recombination can occur in a KU/CST-independent manner in *Arabidopsis*.

Loss of ATM does not disturb telomere defects in plants lacking STN1 or CTC1

The rapid lethality of telomere uncapping in yeast and vertebrates is triggered by the activation of DNA damage checkpoints (Churikov & Price, 2008; Garvik et al, 1995; Palm & de Lange, 2008). ATM is activated in response to TRF2 depletion, while ATR is triggered by mutation of POT1 in mammals or the CST complex in budding yeast (Denchi & de Lange, 2007; Enomoto et al, 2002; Garvik et al, 1995). The massive end-to-end chromosome fusions that accumulate in *ctc1* and *stn1* mutants are consistent with activation of a strong DNA damage response. Notably, plants with a null mutation in *ATM* or *ATR* or even both genes are viable (Culligan et al, 2004; Vespa et al, 2005). Therefore, we examined the role of ATM and ATR in activating a DNA damage response in plants lacking *STN1* or *CTC1*.

Plants doubly deficient in *ATM* and *CTC1* or *STN1* were viable, and displayed fasciated morphological defects similar to *ctc1* or *stn1* siblings (Fig 5-6A and data not shown). TRF analysis revealed no perturbation in telomere maintenance in *atm* mutants as previously examined (Fig 5-6B & C) (Vespa et al, 2005). In comparison, telomeres in *stn1 atm* and *ctc1 atm* double mutants resembled those in *stn1* and *ctc1* plants, and were characterized by highly heterogeneous tracts that trailed down to less than 1 kb in length (Fig 5-6B & C). Furthermore, telomere fusion PCR generated abundant products in plants lacking a CST component and ATM, but not in single *atm* mutants (Fig 5-7A & B). Because telomere fusion PCR amplifies only a subset of the possible end-joining reactions, we also performed cytogenetic analysis on mitotically dividing cells to obtain a more accurate estimate of the chromosome fusions in these mutants. As shown in Fig 5-7E, the same number of anaphase bridges was detected in *stn1 atm* (20%) and *stn1* mutants (21%). To summarize, disruption of ATM does not further affect telomere length or end protection defects in plants lacking CTC1/STN1. Thus, ATM is either not required, or not sufficient to suppress the DNA damage checkpoint in a CST-deficient plant.

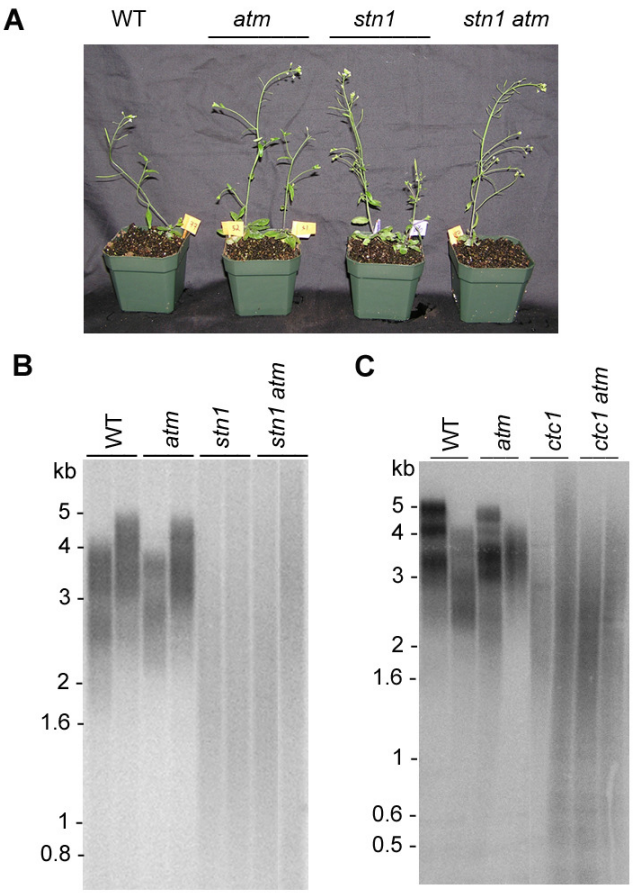
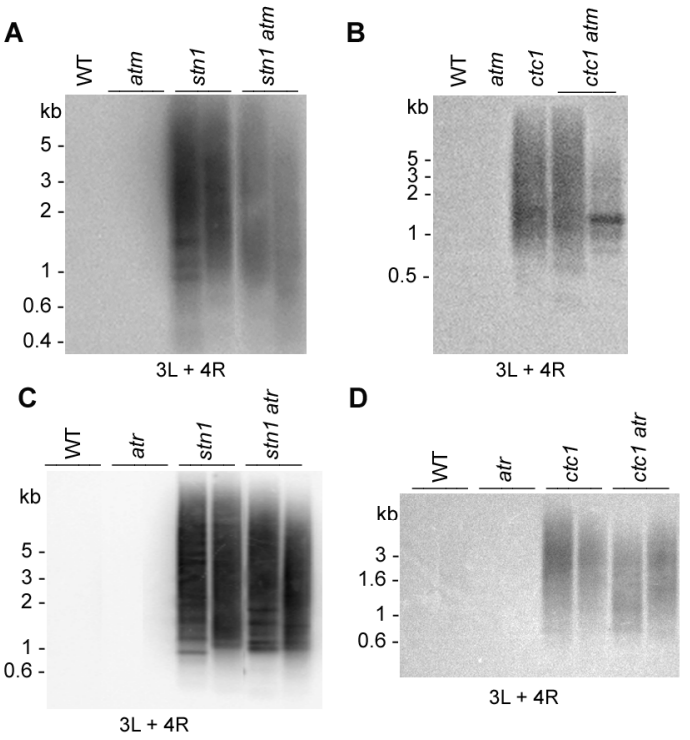


Fig 5-6. Loss of ATM does not affect telomere length in plants lacking *STN1* and *CTC1*. (**A**) Plants doubly deficient of *STN1* and *ATM* display morphological defects similar to *stn1* and *atm* single mutants. *stn1 atm* mutants show fasciated stems and irregular phyllotaxy similar to *stn1* mutants (Song et al, 2008), as well as partial sterility as *atm* mutants (Garcia et al, 2003). (**B**-**C**) TRF analysis showed similar telomere length profile in *stn1 atm* or *ctc1 atm* mutants as in *stn1* or *ctc1* mutants.



E

Genotype	Total anaphase cells analyzed	Anaphase cells with bridges	Percentage anaphase cells with bridges
WT	69	0	0%
atm	202	0	0%
atr	454	6	1%
stn1	369	76	21%
stn1 atm	137	27	20%
stn1 atr	501	288	57%

Fig 5-7. ATR, but not ATM regulates a DNA damage checkpoint in plants lacking *STN1* and *CTC1*. Abundant telomeric fusions were detected using fusion PCR analysis in *stn1 atm* or *ctc1 atm* mutants (**A**-**B**) as well as in *stn1 atr* or *ctc1 atr* mutants (**C**-**D**). (**E**) A synergistic increase of anaphase cells carrying bridges in plants lacking *STN1* and *ATR*. DAPI-stained chromosome spreads were prepared from pistils. Anaphase cells from different genotypes were analyzed. Anaphase bridges were scored as a percentage of total anaphase cells analyzed.

ATR facilitates telomere maintenance independent of STN1 *or* CTC1

A different result was obtained from plants lacking ATR and the CST complex. Unlike *stn1* or *ctc1* mutants, *stn1 atr* and *ctc1 atr* mutants were rarely found to display fascination of stems or other organs (Fig 5-8A & B). Instead, the double mutants were morphologically more like WT or *atr* single mutants. The less severe morphological phenotype in plants lacking both ATR and CST components suggest that ATR inhibits proliferation of cells with damaged telomeres in CST mutants. Consistent with this hypothesis, ATR is involved in activating a G2 cell-phase checkpoint in *Arabidopsis* (Culligan et al, 2004).

The telomere length in these mutants was examined. As previously shown (Vespa et al, 2005), disruption of *ATR* did not alter bulk telomere length measured by TRF analysis (Fig 5-8C & D). The average telomere length was about 3.2 kb in *atr* mutants similar to that in WT (3.4 kb). Strikingly, loss of *ATR* accelerated telomere shortening in plants deficient in *STN1* or *CTC1* (Fig 5-8C & D). The average length was significantly reduced in *stn1 atr* mutants (1.6 kb) comparing to that in *stn1* single mutants (2.1 kb). The longest telomeres were preferentially lost in double mutants (4.3 kb) compared to *stn1* (5.0 kb), *atr* (5.0 kb) or WT (5.2 kb) plants. Therefore, telomere tracts in *stn1 atr* mutants were similar to those in plants lacking telomerase and a CST component (Figs 5-1 and 5-2).

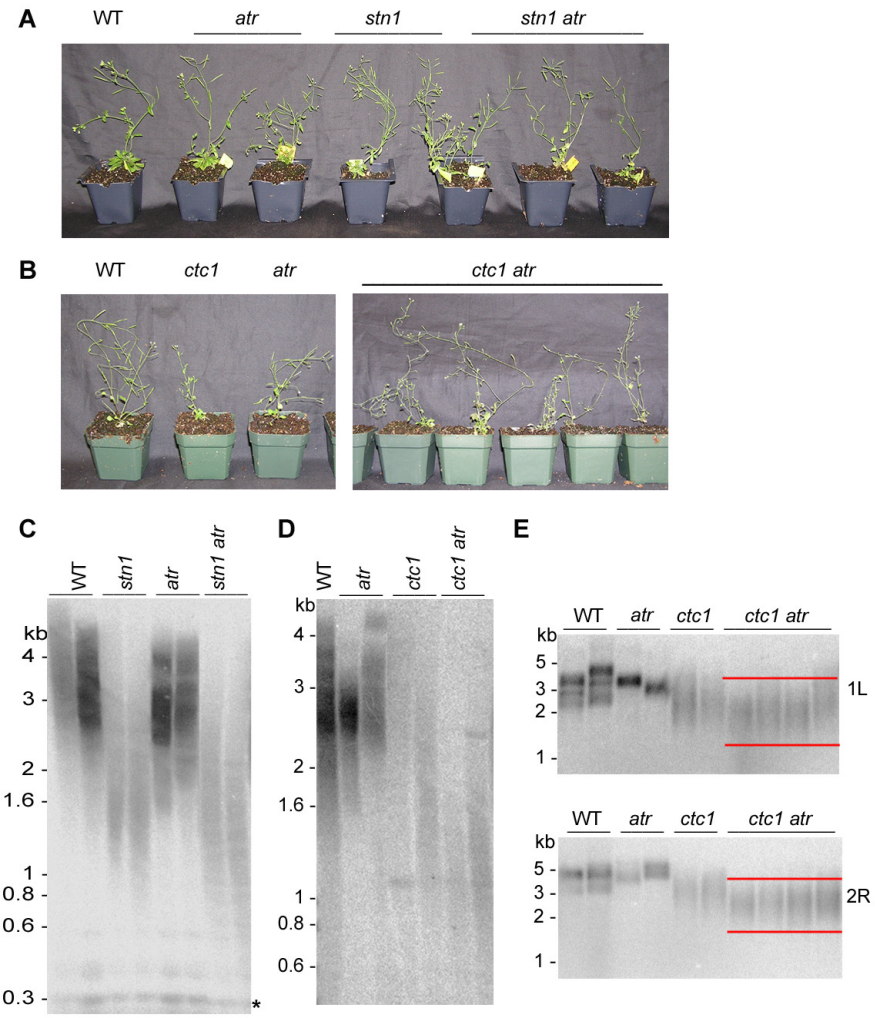


Fig 5-8. ATR and CST components protect telomere length synergistically. (**A**-**B**) *stn1 atr* and *ctc1 atr* mutants were morphologically more similar to *atr* single mutant than either *stn1* or *ctc1* mutants. TRF (**C**-**D**) and PETRA analyses (**E**-**F**) were performed on *stn1 atr* and *ctc1 atr* mutants along with their WT, *stn1*/*ctc1* and *atr* siblings. Even shorter telomeres were observed in *stn1 atr* and *ctc1 atr* mutants comparing to *stn1* and *ctc1* single mutants. The asterisk indicates an interstitial telomeric single as a loading control. For PETRA, individual chromosome arms examined were indicated.

Moreover, TRF analysis revealed a reproducible decrease in the overall hybridization signal for *stn1 atr* mutants (~ 45% of WT) relative to *stn1* single mutants (~ 60% of WT) or *atr* mutants (~ 96% of WT) (Fig 5-8C), which is consistent with a greater loss of bulk telomeric DNA. Like *stn1 atr* mutants, *ctc1 atr* mutants showed a similar pattern of shortened bulk telomere tracts (Fig 5-8D). We also performed PETRA on the double mutants on different individual chromosome ends (Fig 5-8 E & F). Consistent with the bulk telomere analysis, even shorter telomeres were observed in *stn1 atr* and *ctc1 atr* mutants compared to *stn1*, *ctc1* or *atr* single mutants. For example, the average telomere length on 1L in *ctc1 atr* mutants were about 0.2 kb shorter than *stn1* siblings (2.5 kb VS 2.7 kb); while on 2R telomeres, there was a 0.4 kb difference between *stn1 atr* (2.5 kb) and *stn1* mutants (2.9 kb). Similar PETRA data were obtained from *stn1 atr* mutants. Taken together, these data indicate that ATR and the CST complex act synergistically to maintain the length of telomere tracts.

As expected, no telomere fusion PCR products were generated with DNA from wild type or *atr* single mutants (Fig 5-7C & D) (Vespa et al, 2005). In contrast, abundant products were obtained in reactions with *stn1 atr* and *ctc1 atr* mutants similar to *stn1* or *ctc1* single mutants (Fig 5-7C & D). Strikingly, cytogenetic analysis revealed a three-fold increase in the number of anaphase cells harboring bridged chromosomes in *stn1 atr* mutants (57%) relative to *stn1* (21%) *or stn1 atm* (22%) (Fig 5-7E). Thus, the fusion of CST-deficient telomeres appears to be significantly more prevalent in the absence of ATR. Overall, these results indicate that ATR plays a more direct, CST-independent role in telomere replication. In addition, the more abundant anaphase bridges in *stn1 atr* mutants suggest that ATR is responsible for activating a checkpoint of DNA damages in

stn1 mutants. Further studies are underway to elucidate whether ATM plays any redundant role as ATR for such checkpoint responses.

Discussion

Eukaryote genome stability relies on intact chromosome ends. Each component of the Cdc13/Stn1/Ten1 complex in budding yeast is essential for cell viability, and hence analysis of the CST complex has relied on conditional alleles. In contrast, *Arabidopsis ctc1* or *stn1* mutants are viable for up to two generations, although they suffer extensive loss of telomeric/subtelomeric DNA and harbor massive chromosome end-to-end fusions.

Here we exploit the extraordinary tolerance of *Arabidopsis* toward telomere dysfunction to further investigate the function and interactions of the CST complex in a multicellular eukaryote. We found that simultaneous loss of CST and telomerase is extremely detrimental. Consistent with the severe growth and developmental defects in these double plants, bulk telomeres are much shorter than in single mutants, and a substantial population of telomeres fall below the 1 kb threshold, which was previously shown to represent the minimal functional length in *Arabidopsis* (Heacock et al, 2007). Since telomerase preferentially acts at the shortest telomeres in the population (Shakirov & Shippen, 2004), we hypothesize that telomerase temporarily rescues the lethality of plants lacking the CST complex by extending critically shortened telomeres to delay catastrophic genome instability.

Telomere tracts are significantly longer in c*tc1* and *stn1* mutants that express telomerase, suggesting that the CST complex is not required for telomerase action on chromosome ends. It is possible that CST does not play a role in telomerase

recruitment in multicellular eukaryotes. Alternatively, this function may be redundant. In budding yeast, there are two pathways for telomerase recruitment. One is facilitated by interactions between Cdc13 and the Est1 component of telomerase. This interaction is required to maintain a high level of telomere-bound Est2 (TERT) in late S/G2 phase. The other recruitment pathway is promoted by the interaction of KU with TLC1 (telomerase RNA), which leads to the association of Est2 with telomeres in G1 (Chan et al, 2008). In *Arabidopsis*, telomerase recruitment is unlikely to be mediated by KU, as this complex negatively regulates telomere length. Therefore, we propose a CST/KU-independent pathway in *Arabidopsis* to recruit telomerase for telomere synthesis. Supporting this conclusion, loss of *CTC1* and *KU* simultaneously did not show synergistically shortening of telomeres, which were observed in *ctc1 tert* survivors (Fig 5-1F and 5-4B).

Both CST and KU contribute to maintaining 3' G-overhang (Riha & Shippen, 2003; Song et al, 2008; Surovtseva et al, 2009). Our data show that mutants of *ctc1 ku70* and *stn1 ku70* suffered even more severe morphological defects than any single mutants, displaying a vegetative arrest without producing any germline. Moreover, the double mutants displayed additive G-overhang signals than *ctc1* or *ku70* single mutants. Thus, these data reinforce the importance of regulating G-overhangs in multicellular eukaryotes and indicate that CTC1/STN1 protects telomere end architecture in a different pathway from KU70/KU80.

It is currently unclear how CST deficient telomeres are monitored and repaired in a multicellular organism. In this study, we examined the role of ATM and ATR in plants lacking CTC1/STN1. No enhanced telomere length or protection defects were observed in *ctc1 atm* and *stn1 atm* mutants comparing to *ctc1* and *stn1* mutants, indicating that

ATM is not required or not sufficient to activate a DNA damage response in CST deficient mutants. In contrast, we found that lack of *ATR* in plants deficient of *CTC1* or *STN1* leads to accelerated telomere shortening, suggesting that the role of ATR in telomere maintenance is independent of CST. Supporting this notion, telomeres in *atr tert* mutants experience a faster shortening comparing to *tert* mutants (Vespa et al, 2005). Shortened telomeres are found in budding yeast deficient of Mec1 (ATR homolog) (Ritchie et al, 1999). Mec1 also promotes resolving replication forks (Cha & Kleckner, 2002). Loss of ATR in *stn1* or *ctc1* mutants also showed a synergistical increase in genome instability, with up to 57% of anaphase cells carrying anaphase bridges. Amazingly, these *stn1 atr* mutants display less severe morphological defect and are able to produce seeds for the next generation. These data not only suggest that the CST complex prevents telomeres from activating an ATR-dependent checkpoint response, but also that ATR may play a more direct role in promoting telomere maintenance independently of the CST complex. Taken together, the extreme tolerance of *Arabidopsis* toward genome instability allows us to study genes that are critical for telomere maintenance and protection. Further studies are needed to elucidate how CST guards chromosomal termini in the presence of other telomere-related components.

CHAPTER VI

POSITIVE SELECTION AND NEO-FUNCTIONALIZATION SHAPE THE MOLECULAR EVOLUTION OF *POT1* GENES IN PLANTS

Summary

The flowering plant *Arabidopsis thaliana* encodes two divergent Protection Of Telomeres 1 (POT1) proteins termed AtPOT1a and AtPOT1b. Unlike yeast and human POT1 proteins which interact with telomeric DNA and protect telomere integrity, AtPOT1a and AtPOT1b are associated with telomerase RNA (TER). AtPOT1a interacts with TER1 and is required for telomerase enzyme activity *in vivo*. In comparison, AtPOT1b associates with TER2, and participates in telomerase regulation and G-overhang maintenance. To explore the nature and origin of *AtPOT1* gene duplication, we analyzed *POT1* genes from 30 representatives of the plant kingdom. Only two examples of independent *POT1* gene duplication were uncovered. One occurred ~ 25-30 mya in the Panicoideae subfamily of grasses, and the other at least 40 mya in the Brassicaceae family, which includes *Arabidopsis thaliana*. Computer modeling of the first oligonucleotide/oligosaccharide binding fold (OB1), which in yeast and human contacts telomeric DNA, revealed no significant variation in the overall folding and location of functionally conserved residues in plant POT1 proteins. Several positively selected sites were detected in AtPOT1a, the two with the highest posterior probability values map to OB1. Mutating these two sites back to ancestral amino acids dramatically reduced AtPOT1a function in a genetic complementation assay. These results indicate that AtPOT1a is under positive selection. Moreover, a *POT1a* gene from *Brassica oleracea*, which is closely related to *Arabidopsis*, only partially rescues AtPOT1a

deficiency. Our data support a neo-functionalization model for *AtPOT1a* gene function, and reveal an extraordinarily rapid evolution of POT1 proteins in Brassicaceae.

Introduction

With over 350,000 species, approximately 250,000 species of which belong to the Division of Angiosperms, green plants provide unique opportunities to study both the antiquity of conserved eukaryotic biology and how general phenomena of eukaryotic biology have diversified, such as telomeres. Telomeres are an ancient hallmark of most eukaryotic chromosomes and are essential for genome stability and long-term proliferation capacity. The TG-rich sequence of telomeric DNA repeats is well conserved across eukaryotes (TTAGGG in vertebrates and TTTAGGG in plants), but the composition of telomere-associated proteins varies significantly among distant organisms.

Vertebrate telomeres are protected by a six-member complex called Shelterin (Palm & de Lange, 2008). Shelterin recognizes and associates with telomeres through two ds telomere binding proteins, TRF1 and TRF2, and the ss telomere binding protein, POT1 (Liu et al, 2004a; Ye et al, 2004b). In addition, three other Shelterin components are recruited to telomeres by TRF1/TRF2: TIN2, RAP1 and TPP1. TPP1 contacts with POT1 (Liu et al, 2004a; Ye et al, 2004b), and is hypothesized to function in telomerase regulation (Wang et al, 2007b; Xin et al, 2007). Recent studies suggest that Shelterin-like components are conserved in fission yeast, including Taz1 (a TRF1/TRF2 homolog), Rap1, Pot1 and Tpz1 (a TPP1 homolog) (Baumann & Cech, 2001; Chikashige & Hiraoka, 2001; Ferreira & Cooper, 2001; Kanoh & Ishikawa, 2001; Miyoshi et al, 2008). Shelterin components are evolving rapidly. For example, fission yeast Tpz1 shares only

~ 20% similarity to human TPP1 protein (Miyoshi et al, 2008). Moreover, Poz1 (Pot1-associated in *Schizosaccharomyces pombe*) bridges Taz1/Rap1 and Tpz1/Pot1 subcomplexes in fission yeast, and thus appears to be a functional homolog of TIN2. Yet, Poz1 does not show any sequence similarity to TIN2 protein (Miyoshi et al, 2008).

Only a few Shelterin components can be discerned in plant genomes. Although homologs of RAP1, TIN2 and TPP1 have not been identified in plants, six TRF-like proteins are identified in *Arabidopsis* which specifically bind ds telomeric DNA *in vitro* (Karamysheva et al, 2004). The ss telomeric DNA binding protein POT1, a key component of the Shelterin complex, is remarkably conserved across eukaryotes. Although most eukaryotes harbor a single *POT1* gene (Baumann and Cech, 2001), two POT1 paralogs have been reported in mouse (Hockemeyer et al, 2006; Wu et al, 2006) and some ciliates (Jacob et al, 2007; Wang et al, 1992). The two mouse POT1 proteins share 72% similarity and are partially redundant. Both are required to prevent DNA damage responses at telomeres as well as cell senescence (He et al, 2006; Hockemeyer et al, 2006; Wu et al, 2006). With a similar architecture of two N-terminal O-folds and a C-terminal extension, *Arabidopsis* encodes two full-length POT1 homologs, POT1a and POT1b. There is a third truncated protein termed POT1c that harbors a single OB-fold. POT1a and POT1b are highly divergent, displaying only 49% protein sequence similarity. POT1b associates with TER2 and negatively regulates telomerase (C. Cifuentes-Rojas and Shippen D.E., unpublished work). Overexpression of a dominant negative allele of POT1b or POT1c results in genome instability, implicating POT1b and POT1c in chromosome end protection (Shakirov et al, 2005). In contrast, *pot1a* null mutants display an ever-shorter telomere phenotype, but the chromosome ends, at least initially, remain fully protected (Surovtseva et al, 2007).

POT1a interacts with TER1 (C. Cifuentes-Rojas and Shippen D.E., unpublished work) and appears to be a novel accessory subunit of the telomerase RNP complex (Surovtseva et al, 2007). Genetic studies reveal that POT1a acts in the same genetic pathway as telomerase (Surovtseva et al, 2007).

The unusual activity of *Arabidopsis* POT1 proteins prompted us to study *POT1* genes in differnt plant species. To explore the origin of *AtPOT1* gene duplication, we analyzed *POT1* genes in 30 representatives of the plant kingdom. Our data suggest that *POT1* gene duplication event is rare in plants. Only two examples of independent *POT1* gene duplication were uncovered: one in the Panicoideae subfamily of grasses, and the other in the Brassicaceae family, which includes *Arabidopsis thaliana*. Using a combination of computational structural modeling, phylogenetic and positive selection analysis, and a genetic complementation approach, we found that the POT1a lineage in Brassicaceae has undergone a rapid evolution. Although the overall structure of OB1 in POT1a remains similar to its yeast or human homolog, mutating two of the positive selection sites in OB1 back to ancestral amino acids dramatically reduced AtPOT1a function *in vivo*. These results indicate that Brassicaceae POT1a is under positive selection. Moreover, introduction of a *POT1a* gene from *Brassica oleracea* only partially rescues AtPOT1a deficiency, indicating an extraordinarily rapid evolution of POT1a proteins among closely-related species in the Brassicaceae family. Taken together, the accessibility of a large collection of green plant species allows us to study the rapid evolution of *POT1* genes. Our studies indicate that neo-functionalization shapes POT1a function to act as an unusual telomerase accessory factor in *Arabidopsis*.

Materials and methods

Database searches and cDNA cloning

BLAST searches of POT1 sequences in different plant genomes were performed using the blastp or tblastn options available at the corresponding genome portals (http://asgpb.mhpcc.hawaii.edu/tools/tools.php, http://genome.jgi-psf.org/euk_cur1.html, http://www.appliedgenomics.org/blast) with *Arabidopsis* POT1 proteins as a query. For lower plants, database searches with POT1 sequences from more closely related species, such as *Physcomitrella*, improved BLAST results. BLAST searches with human or *S. pombe* POT1 proteins as queries were also attempted, but did not improve the outcome. Notably, we failed to detect POT1 orthologues in *Chlamydomonas* and *Volvox*, two closely related green algae species of the Volvocales family. We suspect that *POT1* genes in these species diverged beyond the power of bioinformatic algorithms used in this study. Alternatively, these species may engage other proteins to perform POT1 functions, such as the Gbp1 protein in *Chlamydomonas* (Petracek et al, 1994).

EST sequences from the database were used to help deduce and clone individual cDNA sequences. In cases where only partial EST sequences were available, 5'- and 3'-RACE (Ambion) were employed to amplify the full-length POT1 coding regions. cDNAs were synthesized and cloned as described (Shakirov et al, 2005).

Phylogenetic analyses and analysis of positive selection

Nucleotide sequences were aligned using Revtrans 1.4 (Wernersson & Pedersen, 2003). The aligned sequences were then manually adjusted using MacClade Vers. 4.08 (Maddison & Maddison, 2005). Phylogenetic trees were constructed using PhyML version 2.4.4 and PAUP (Guindon & Gascuel, 2003). Phylogenetic analyses of amino acid sequences were performed using PhyML with the WAG + I + Gamma model of molecular evolution. Maximum likelihood phylogenetic analysis of nucleotide data from eudicots was performed using the GTR + I + Gamma model of molecular evolution. The GTR + I + Gamma model of molecular evolution is the best fitting model of molecular evolution as determined by MrAIC version 1.4.2 (Nylander, 2004). 500 nonparametric bootstrap replicates on both amino acid sequence data and nucleotide data were performed using PhyML. Maximum Parsimony analysis was performed using PAUP*4.0 (Swofford & Sullivan, 2003) on the eudicot POT1 nucleotide sequences. 1000 nonparametric bootstrap replicates were performed with tree bisection reconnection (TBR) branch swapping and 10 random taxon addition replicates per bootstrap.

Using PAML 4.0 (Yang, 1997), the branch-site model A test was implemented with the foreground branch represented by either the POT1a or the POT1b lineage. Background branches consist of sites $0 < \omega < 1$ or $\omega = 1$ (Zhang et al, 2005). Foreground lineages contain sites $0 < \omega < 1$ or $\omega = 1$ and sites that have come under selection $\omega > 1$ (Zhang et al, 2005). Those sites with $\omega > 1$ may be derived from sites in the background lineages with $0 < \omega < 1$ (purifying selection) or $\omega = 1$ (neutrally evolving sites) (Zhang et al, 2005).

Structure prediction of plant POT1 proteins

The secondary structure of each plant POT1 protein sequences was determined using the secondary structure prediction server PSIPRED v2.6 (McGuffin et al, 2000). Generation of theoretical structural models for the OB1 domains present in each plant POT1 protein was accomplished using threading techniques. Best structural templates for each plant POT1 protein was selected based on the best alignment score generated from a CLUSTAL alignment (Larkin et al, 2007), which was found to be the N-terminal OB1 domain present in *Oxytricha nova*. Optimum sequence alignments were done using PROMALS (Pei & Grishin, 2007), which couples primary sequence homology with secondary structure prediction to align two sequences that are poorly conserved in primary sequence. PROMALS aligned sequences were submitted to the SWISS-MODEL threading server (www.swissmodel.expasy.org) (Arnold et al, 2006). Generated models were then subjected to a GROMOS96 energy minimization to adjust bond lengths, angles and geometries using Swiss-PDB Viewer (Guex & Peitsch, 1997).

Plant growth and transformation procedures

Arabidopsis seeds were cold-treated overnight at 4℃, and then placed in an environmental growth chamber and grown under a 16-hr light photoperiod at 23℃. *pot1a-1*, *ku70*, and *pot1a-1*$^{-/-}$*ku70*$^{+/-}$ mutants were described previously (Riha et al, 2002; Surovtseva et al 2007). For complementation experiments, POT1 cDNAs were cloned into the pCBKO5 binary vector carrying the bar gene as a selectable marker (Riha et al, 2002) under the control of the AtPOT1a native promoter (a 1.5 kb region immediately upstream of the start codon) or, for over-expression, under the control of the CaMV 35S promoter. Complementation constructs were introduced into the *Agrobacterium*

tumefaciens GV3101 strain, which was used to transform *pot1a-1*$^{-/-}$*ku70*$^{+/-}$ plants by a modified *in planta* method (Bechtold & Pelletier, 1998). T1 primary transformants were selected on 0.5 Murashige and Skoog basal medium supplemented with 2mg/L of phosphinothricine (BASTA) (Crescent Chemical, Islandia, New York) and genotyped by PCR to identify *pot1a-1*$^{-/-}$*ku70*$^{-/-}$ plants expressing the transgene. PCR genotyping was used to identify their siblings without the transgene.

Telomere length analysis and quantification

DNA from individual whole plants was extracted as described (Cocciolone & Cone, 1993). Terminal Restriction Fragment (TRF) analysis was performed with DNA digested with Tru1*I* (Fermentas, Hanover, MD) restriction enzyme. ^{32}P 5' end–labeled $(T_3AG_3)_4$ oligonucleotide was used as a probe (Fitzgerald et al, 1999). Radioactive signals were scanned by a Storm PhosphorImager (Molecular Dynamics, Sunnyvale, CA), and the data were analyzed by IMAGEQUANT software (Molecular Dynamics).

The average telomere length (L) was measured using Telometric-1.2 program (Grant et al, 2001). The average telomere lengths of untransformed *pot1a ku70* mutants, transformants expressing wild-type AtPOT1a and other POT1 constructs were designated as L_0, L_1, and L_x, respectively. We set the complementation level of wild-type AtPOT1a transformants (positive control) as one, and that of untransformed *pot1a ku70* mutants (negative control) as zero. The complementation efficiency (E) of each POT1 construct was calculated as: $E= (L_x- L_0) / (L_1-L_0) \times 100\%$. At least three individual transformants for each construct were analyzed for statistical support.

Nuclei isolation, antibodies and western blotting

1.5 g of one week old seedlings was harvested for each sample and ground into fine powder in liquid nitrogen. Nuclei were extracted by nuclei isolation buffer, or NIB (5mM EDTA, 50mM Tris-HCl pH 8.0, 10mM KCl, 250mM sucrose, 1.5mM $MgCl_2$, 0.3% Triton X-100, 1mM PMSF, 5mM beta mercaptoethanol, 1mM spermine, 1mM spermidine and protease inhibitors) for 10 min on ice. After filtering through a layer of miracloth, nuclei were centrifuged at 3000g for 30 min at 4ºC. The pellets were washed with Triton buffer (250mM sucrose, 10mM Tris-HCl pH8.0, 10mM $MgCl_2$, 1% Triton X-100, 5mM beta mercaptoethanol, 1mM spermine, 1mM spermidine and protease inhibitors) on ice for 10 min. After centrifugation at 2000g for 1 min, 4000g for 1 min, and 8000g for 2 min, the pellets were resuspended in 1.5M sucrose in NIB and loaded onto equal volume of the same sucrose NIB cushion buffer. After centrifugation at 14,000g for 30 min, the pellet of nuclei were resuspended in 50 to 100 µl of TMG buffer (10mM Tris-HCl pH8.0, 1mM $MgCl_2$, 1mM DTT, 15% glycerol, 0.5mM PMSF and protease inhibitors). About 5 µg of protein of each nuclei sample was loaded onto a 15% SDS polyarylamide gel for electrophoresis. The same amount of protein was loaded onto another gel and stained with Coomassie blue as a loading control. Gels for western blotting were transferred to PVDF membranes (GE Healthcare) and blocked with 5% non-fat dry milk in 1×TBST overnight at 4ºC. Blots were incubated with mouse monoclonal anti-Flag M2 antibody (Sigma, dilution of 1:1000 with 1% milk in 1×TBST) for 1 h and then washed for three times with 1% milk in 1×TBST before incubation with horseradish peroxidase conjugated goat anti-mouse antibody (Jackson ImmunoResearch, dilution of 1: 7000) for 45 min. The blots were washed for three times

as before and then exposed to ECL films (GE Healthcare) after incubation with ECL western blotting detection reagents (GE Healthcare).

Nucleotide sequence accession numbers

Accession numbers for AtPOT1a (AY884593) and AtPOT1b (AY884594) were reported previously (Shakirov et al., 2005). cDNAs encoding the following plant POT1 proteins were deposited into the GenBank: *Arabidopsis lyrata* POT1a (EU880293), *Arabidopsis lyrata* POT1b (EU880294), *Hordeum vulgare* POT1 (EU880295), *Lactuca sativa* POT1 (EU880296), *Populus trichocarpa* POT1 (EU880297), *Helianthus argophyllus* POT1 (EU880298), *Brassica oleracea* POT1a (EU880299), *Brassica oleracea* POT1b (EU880300), *Selaginella moellendorffii* POT1 (EU880301), *Physcomitrella patens* POT1 (EU880302), *Zea mays* POT1a (EU880303), *Zea mays* POT1b (EU880304), *Gossypium hirsutum* POT1 (EU880305), *Pinus taeda* POT1 (EU880306), *Solanum tuberosum* POT1 (EU883536), *Nicotiana tabacum* POT1 (EU883537), *Triticum aestivum* POT1 (EU883538), *Sorghum bicolor* POT1a (EU883539), *Sorghum bicolor* POT1b (EU883540), *Carica papaya* POT1 (EU887728). The following accession numbers were also used in this study: CAH67370 for *Oryza sativa* POT1, ABO96101 for *Ostreococcus lucimarinus* POT1 and CAL54099 for *Ostreococcus tauri* POT1. *Vitis vinifera* POT1 sequence is essentially CAO68206, except for the first 41 codons, which were determined computationally as atgggtggtgaggacgactatagattcatggccatagaagatgccatggcctcactcaaccaaaaagttaacatcatcgg cgttgtagtggaaatgggcatgcctaagcggtccaaaggaact.

Results

Phylogeny and duplication POT1 *genes in plants*

There are several well-documented genome duplication events during the evolution of land plants, many of which occurred at the radiation of the Angiosperms (Soltis et al, 2009). More than 66% of *Arabidopsis* genome is duplicated (Paterson et al, 2000) and this number is 59% and 66% for poplar and rice genomes, respectively (Tuskan et al, 2006; Yu et al, 2005). To explore the origin of *POT1* gene duplication in plants, we obtained POT1 sequences from 30 different organisms representing the major evolutionary lineages of the plant kingdom (Fig 6-1). A part of this collection includes taxa whose genomes have been sequenced (e.g. *Arabidopsis thaliana*, *Populus trichocarpa*, *Medicago truncatula*, *Carica papaya*, *Oryza Sativa*, *Ostreococcus lucimarinus* and *Ostreococcus tauri*). For other taxa, POT1 sequences were obtained through either EST databases or manually cloned from cDNA.

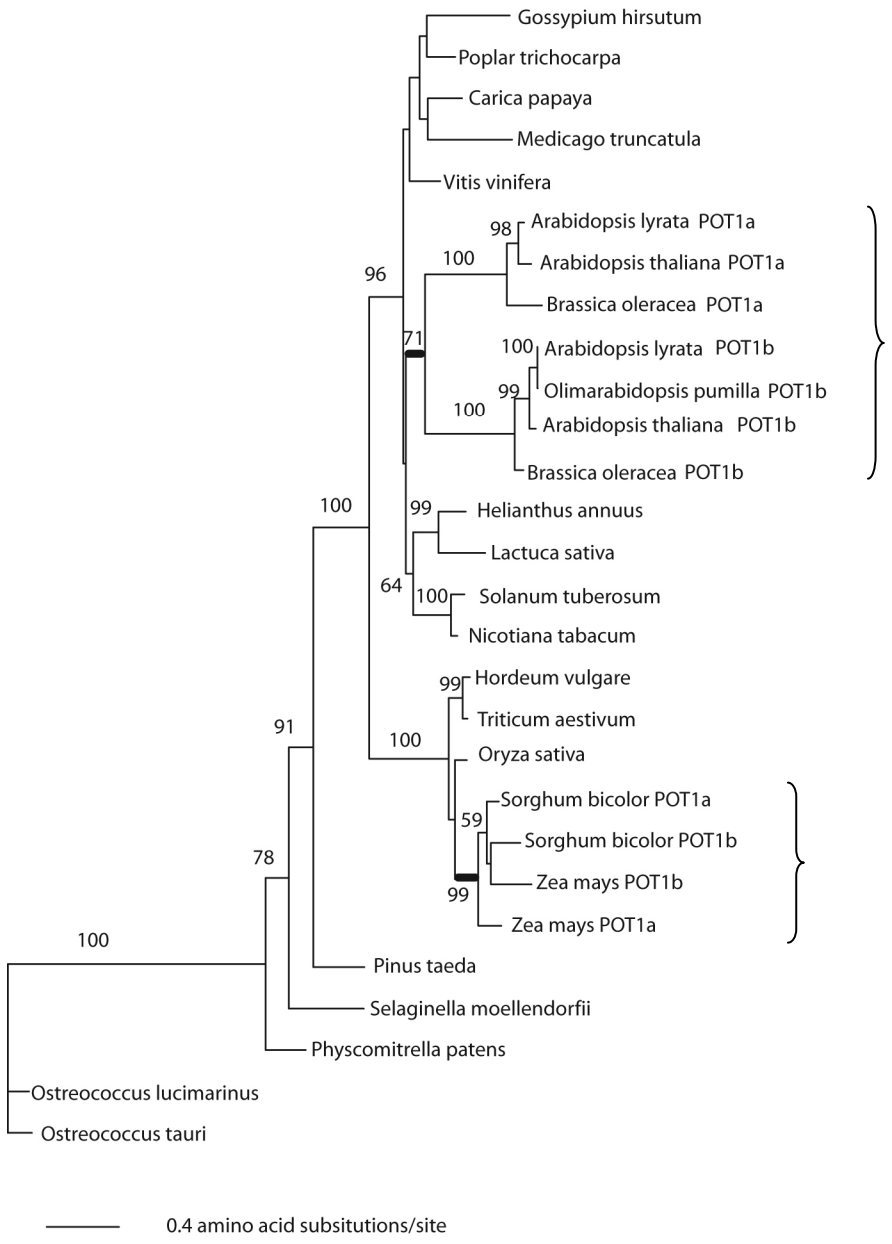


Fig 6-1. Phylogenetic tree of plant POT1 proteins. *POT1* gene sequences were obtained from 30 different taxa across the plant kingdom. Phylogenetic analyses were performed as described in Materials and methods. The two *POT1* gene duplication events were highlighted by brackets on the right.

Our bioinformatic and phylogenetic analyses suggest that duplication events of plant *POT1* genes are relatively rare. Only two *POT1* gene duplication events were detected for the taxa sampled. These events occurred independently in different lineages of the Angiosperms (Fig 6-1). The first example is a Panicoideae-specific event in grasses (maize and sorghum) that occurred less than 30 mya after the divergence of this lineage from the last common ancestor with *Oryza sativa*, *Triticum aestivum* and *Hordeum vulgare* (Paterson et al, 2004) (Fig 6-1). The overall amino acid sequence similarity between POT1 proteins within or between the different grasses is 70%-75%, similar to rodent POT1 paralogs (Hockemeyer et al, 2007). Analysis of the recently released sorghum genome indicates the presence of a third *POT1* gene (Sb06g019440), which encodes a protein that is 85% similar to SbPOT1a and likely represents a product of a more recent duplication event.

The second example of independent *POT1* gene duplication occurred in the Brassicaceae family of dicots, which includes *Arabidopsis thaliana* (Fig 6-1). *AtPOT1a* and *AtPOT1b* genes are located on chromosomes 2 and 5, which do not belong to canonical duplicate chromosomes that share large stretches of co-linearity (such as chromosomes 2 and 4) (Simillion et al, 2002). Since only a single *POT1* gene can be discerned in the other sequenced dicot genomes (e.g. *Populus trichocarpa*, *Medicago truncatula*, *Carica papaya*), we hypothesize that the *POT1* gene duplication in Brassicaceae arose in the lineage leading to *Arabidopsis* after its divergence from the last common ancestor with *Carica papaya* ~ 72 mya (Wikstrom et al, 2001).

To more precisely define the origin of *POT1* gene duplication in Brassicaceae, we cloned full-length POT1 orthologs from two other members of the same family, *Brassica oleracea* and *Arabidopsis lyrata*. In each organism, two POT1 paralogs were identified,

orthologous to AtPOT1a and AtPOT1b (Fig 6-1). Since *Arabidopsis* and *B. oleracea* diverged ~ 40 mya (Beilstein M., personal communication), the *POT1* gene duplication likely originated in the ancestor of the entire Brassicaceae family between 40 mya and 72 mya.

Conservation of POT1 sequence and structure

The presence of tandem OB-folds in POT1 proteins is a conserved feature of the POT1 family (Croy et al, 2006; Lei et al, 2004b; Theobald & Wuttke, 2004; Trujillo et al, 2005a). As expected, all of the plant POT1 proteins we characterized are predicted to encode two N-terminal OB-folds (OB1 and OB2) according to sequence analysis, secondary structure prediction, and computer generated threading models based on well-characterized POT1 structures.

The OB1 domains in plant POT1 proteins contain several structural elements conserved in the POT1 family, including a variable N-terminal region (pre-β1), an α-helical insertion present in the loop connecting β3 and β4 (L_{34}), and the presence of a C-terminal α–helix (Fig 6-2B to 2D). Besides these conserved elements, algal POT1 proteins from *O. tauri* and *O. lucimarinus* harbor large insertions in L_{23} and β5 of OB1 (Fig 6-2D). In budding yeast Cdc13p (Weinert & Hartwell, 1993), L_{23} plays an important thermodynamic role by providing an extensive array of aromatic stacking interactions that facilitate ssDNA binding (Anderson et al, 2003; Mitton-Fry et al, 2002). However, unlike Cdc13p, the majority of the amino acids in algal POT1 proteins are either polar or acidic, indicating that the nucleic acid binding interface presented by these proteins is chemically distinct from Cdc13p.

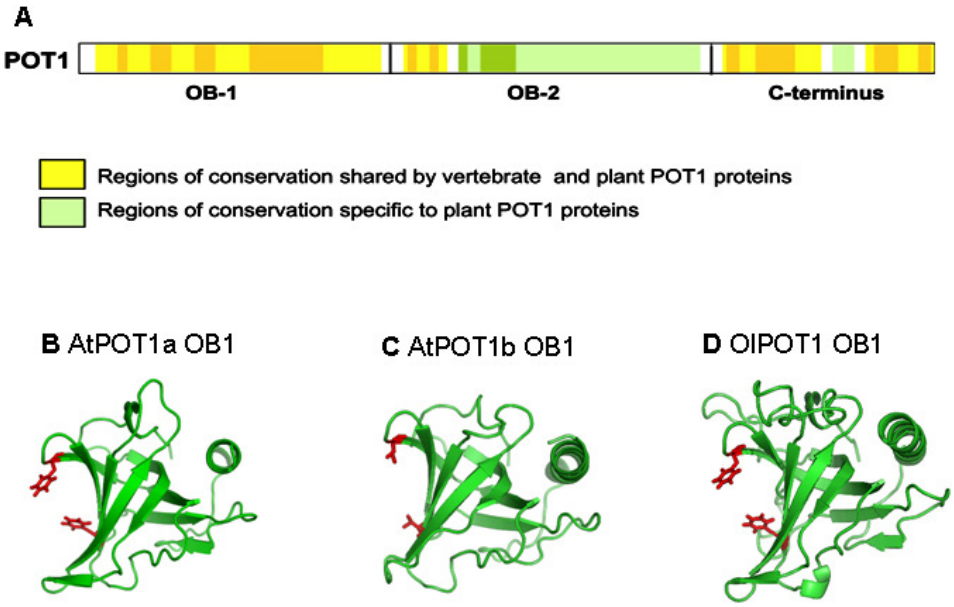


Fig 6-2. Conserved features of plant POT1 proteins. (**A**) A schematic diagram representing a generalized plant POT1 protein. Regions conserved between plant and vertebrate POT1 sequences are shown in yellow and orange (more conserved), and regions conserved only within plant POT1 proteins are shown in pale green and green (more conserved). (**B**), (**C**) and (**D**) Computer modeled structures of OB1 from AtPOT1a (**B**), AtPOT1b (**C**) and *Ostreococcus lucimarinus* OlPOT1 (**D**). In each model, two residues are shown in red, which correspond to the major DNA binding sites in human POT1 OB1.

Secondary structure analysis predicts the presence of a second OB-fold domain (OB2) directly adjacent to OB1 in all plant POT1 proteins. These putative OB2 domains are highly conserved in primary sequence among POT1 proteins within the plant kingdom, but are poorly conserved when compared with OB2 domains of non-plant POT1 proteins (Fig 6-2A). Finally, a region of approximately 160 residues at the C-termini in the plant POT1 proteins exhibit a high degree of sequence similarity to mammalian and *S. pombe* POT1 proteins. Altogether, the plant POT1 proteins identified share sequence and structural similarities to previously characterized POT1 protein family.

Evolution of Brassicaceae POT1 proteins

Given the functional differences between the duplicated POT1 paralogs in *Arabidopsis* (Shakirov et al, 2005; Surovtseva et al, 2007), we hypothesized that POT1a has undergone adaptive evolution and that the corresponding amino acid substitutions are correlated with functional diversification. To test these hypotheses, we performed an additional set of phylogenetic analyses using Maximum Likelihood (ML) and Maximum Parsimony (MP) methods on eudicot POT1 nucleotide sequences (Fig 6-3A & B). The resulting phylogenetic trees were used to test for adaptive amino acid substitutions in the POT1a lineage. Our phylogenetic results confirmed that POT1a and POT1b duplicated either prior to or during the radiation of the Brassicaceae family (Fig 6-3A & B).

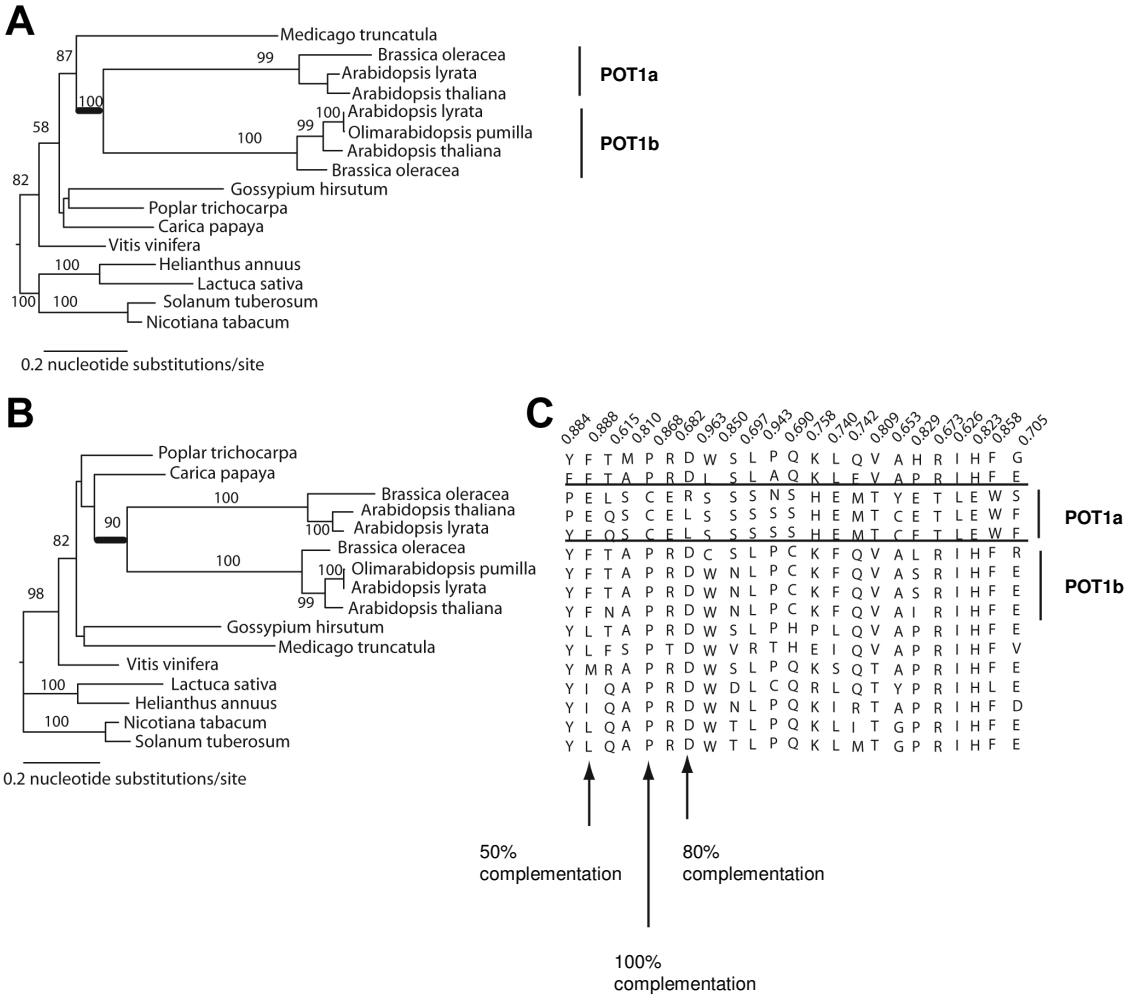


Fig 6-3. Evolutionary analysis of dicot POT1 proteins. (**A**)-(**B**) Phylogenetic trees of dicot POT1 proteins analyzed by Maximum Likelihood (ML) method (**A**) or Maximum Parsimony (MP) method (**B**). (**C**) Positive selection sites in the POT1a lineage predicted by the branch-site model A test of the PAML program. The posterior probability of each site under positive selection was calculated by Bayes Empirical Bayes (BEB) and shown on top of the amino acid. The relative complementation efficiency of E35F, C119P and L132D POT1a mutants is indicated below the corresponding sites.

Two methods of phylogenetic inference (ML and MP) of the eudicot nucleotide sequences resulted in distinguishable topologies. One major difference was the placement of the sister group to the Brassicaceae POT1 co-orthologs. The ML method resulted in the placement of the *Medicago truncatula* POT1 sequence as a sister to the clade containing the Brassicaceae POT1a and POT1b sequences (Fig 6-3A). The MP analysis resulted in trees with *Carica papaya* POT1 as a sister to the Brassicaceae POT1a and POT1b sequences (Fig 6-3B).

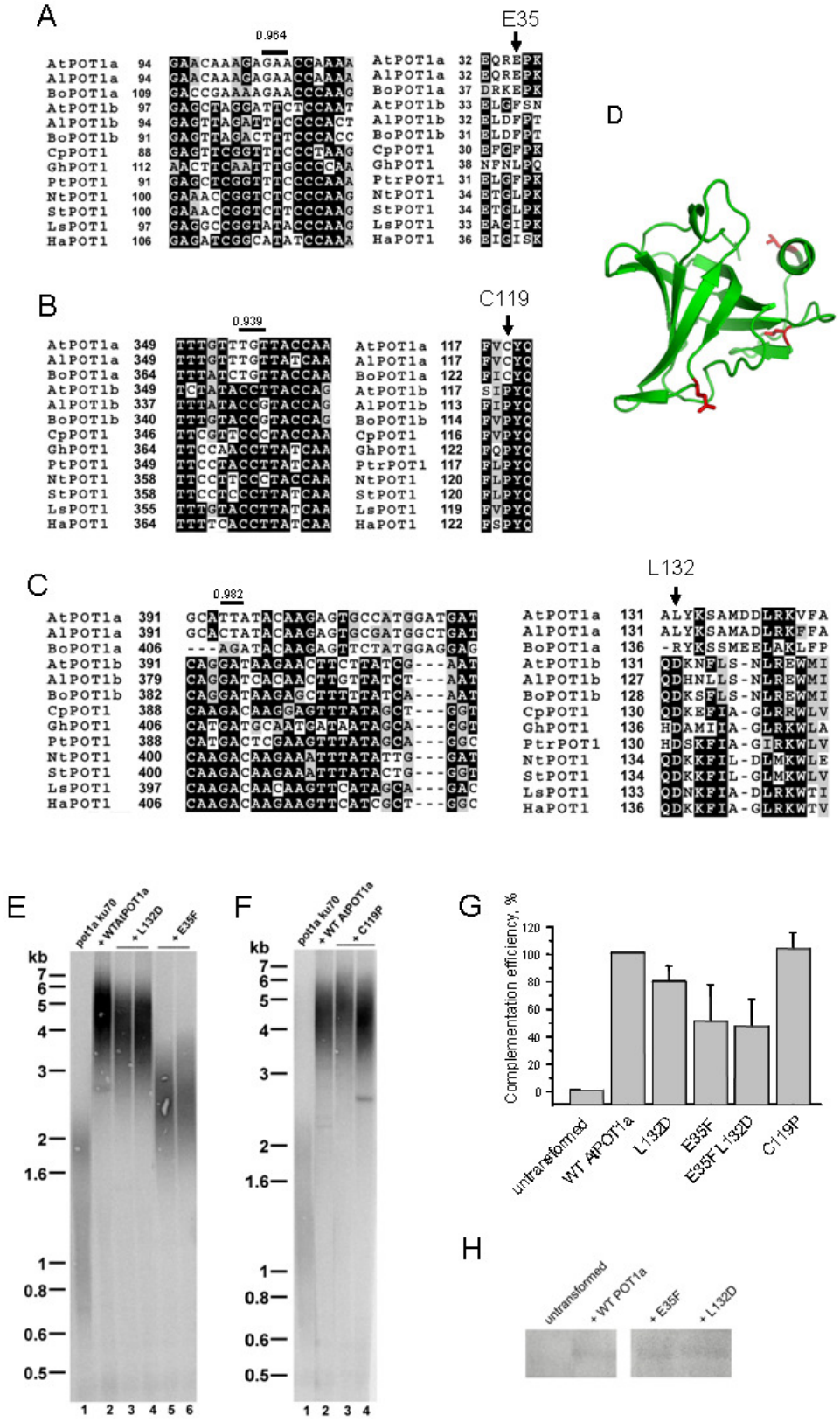
Using the ML and MP phylogenetic tree estimates, we further asked whether the amino acids in the POT1a lineage have experienced positive selection. Specifically, we examined the ratio ($\omega = dN/dS$) of non-synonymous (dN) to synonymous (dS) changes along the branches leading to POT1a. We performed the branch-site model A test (as implemented in PAML program) with the foreground branch represented by either the POT1a lineage or the POT1b lineage. Background branches consist of sites $0 < \omega < 1$ or $\omega = 1$. Foreground lineages contain sites $0 < \omega < 1$ or $\omega = 1$ and additional site classes that have come under selection $\omega > 1$. The site classes on the foreground lineage that represent $\omega > 1$ may be derived from sites in the background lineages that are in either the site class of $0 < \omega < 1$ or $\omega = 1$. This analysis resulted in a significant difference between the null and alternative models, $p = 0.00014$. Therefore, residues in the POT1a lineage are under positive selection.

The Bayes Empirical Bayes (BEB) was used to calculate the posterior probability of sites coming from the site class of ω > 1. Several putative sites were identified (Fig 6-3C), and those sites with posterior probability > 0.85 were treated as important with potentially adaptive roles in the function of POT1a. We selected three positively selected sites, E35, C119 and L132 (Fig 6-3C and Fig 6-4A, 6-4B and 6-4C) for functional tests *in vivo*. The branch-site model comparisons were not significant for the POT1b lineage ($p = 0.187$).

A genetic complementation assay for Arabidopsis POT1a.

To test the functional importance of positively selected sites *in vivo*, we developed a genetic complementation assay for AtPOT1a. Although a null mutation in AtPOT1a dramatically reduces telomerase activity *in vitro* by ~ 13 fold, only about 200 bp of telomeric DNA are lost per plant generation (Surovtseva et al. 2007).

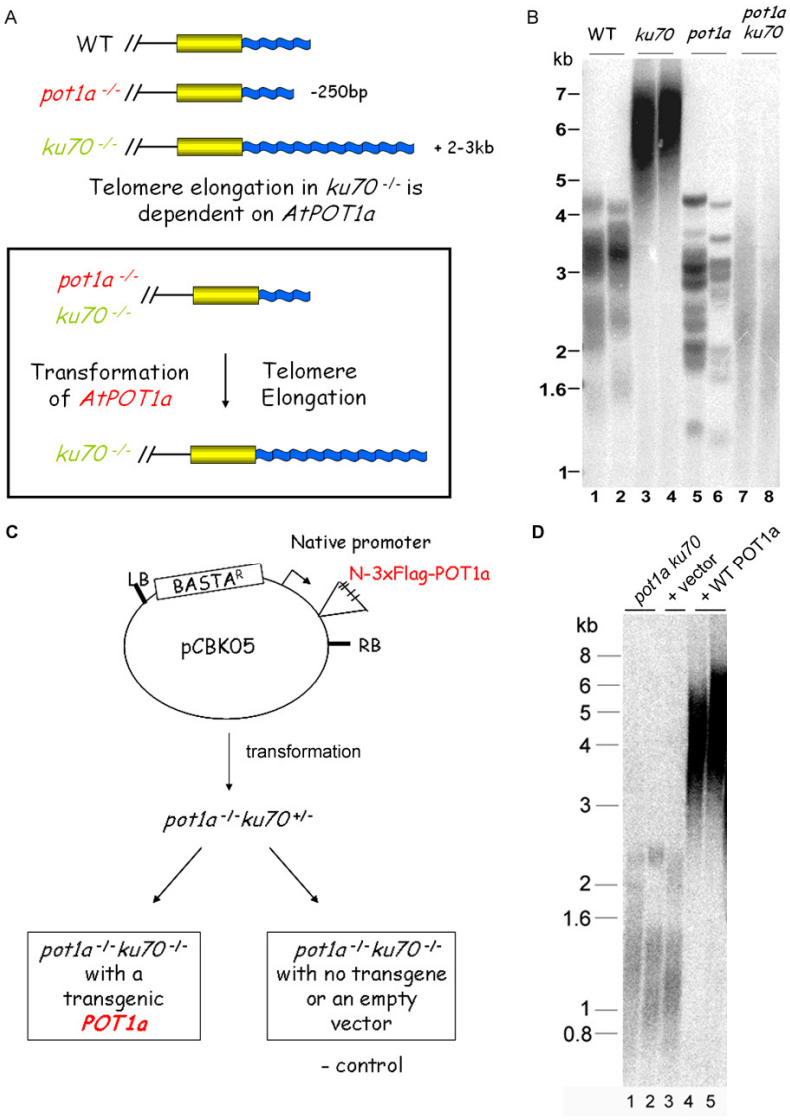
Fig 6-4. Functional analysis of positively selected sites in OB1 of AtPOT1a. (**A**)-(**C**) Nucleotide (left) and amino acid (right) sequence alignment of POT1 regions flanking positively selected E35 (**A**), C119 (**B**) and L132 (**C**) sites. Bars above the alignment indicate positively selected codons in Brassicaceae POT1a proteins, and arrows mark the location of the corresponding amino acids in the alignment. BEB-assigned posterior probability values are shown above the nucleotides. Abbreviations: At, *Arabidopsis thaliana*; Al, *Arabidopsis lyrata*; Bo, *Brassica oleracea*; Cp, *Carica papaya*; Gh, *Gossypium hirsutum*; Pt, *Populus trichocarpa*; Nt, *Nicotiana tabacum*; St, *Solanum tuberosum*; Ls, *Lactuca sativa*; Ha, *Helianthus argophyllus*. (**D**) A modeled structure of OB1 in AtPOT1a showing the location of the three positively selected sites with the highest posterior probability values (residues in red). (**E**) and (**F**) TRF analysis of transformants expressing AtPOT1a variants with positively selected sites substituted by consensus amino acids. (**E**) Lane 1, untransformed *pot1a ku70*; lanes 2-6, *pot1a ku70* mutants expressing WT AtPOT1a (lane 2), AtPOT1a L132D (lanes 3, 4), or AtPOT1a E35F (lanes 5, 6). (**F**) Lane 1, untransformed *pot1a ku70*; lanes 2-4, *pot1a ku70* mutants expressing WT AtPOT1a (lane 2) or AtPOT1a C119P (lanes 3, 4). (**G**) A histogram of complementation efficiency of AtPOT1a variants used in (**E**) and (**F**). (**H**) A western blot showing that E35F and L132D mutant proteins were expressed at a similar level as the transgenic WT POT1a, which fully complemented *pot1a* deficneicy.


A
E35
0.964
AtPOT1a 94 GAACAAAGAGAACCAAAA AtPOT1a 32 EQREPK
AlPOT1a 94 GAACAAAGAGAACCAAAA AlPOT1a 32 EQREPK
BoPOT1a 109 GACCGAAAAGAACCCAAG BoPOT1a 37 DRKEPK
AtPOT1b 97 GAGCTAGGATTCTCCAAT AtPOT1b 33 ELGFSN
AlPOT1b 94 GAGTTAGATTTCCCCACT AlPOT1b 32 ELDFPT
BoPOT1b 91 GAGTTAGACTTTCCCCACC BoPOT1b 31 ELDFPT
CpPOT1 88 GAGTTCGGTTTCCCTAAG CpPOT1 30 EFGFPK
GhPOT1 112 AACTTCAATTTGCCGCAA GhPOT1 38 NFNLPQ
PtPOT1 91 GAGCTCGGTTTCCCCAAA PtrPOT1 31 ELGFPK
NtPOT1 100 GAAACCGGTCTCCCCAAG NtPOT1 34 ETGLPK
StPOT1 100 GAAACCGGTCTTCCCAAG StPOT1 34 ETGLPK
LsPOT1 97 GAGGCCGGTATACCCAAG LsPOT1 33 EAGIPK
HaPOT1 106 GAGATCGGCATATCCAAA HaPOT1 36 EIGISK
B
C119
0.939
AtPOT1a 349 TTTGTTTGTTACCAA AtPOT1a 117 FVCYQ
AlPOT1a 349 TTTGTTTGTTATCAA AlPOT1a 117 FVCYQ
BoPOT1a 364 TTTATCTGTTACCAA BoPOT1a 122 FICYQ
AtPOT1b 349 TCTATACCTTACCAG AtPOT1b 117 SIPYQ
AlPOT1b 337 TTTATACCGTACCAG AlPOT1b 113 FIPYQ
BoPOT1b 340 TTTGTACCGTACCAG BoPOT1b 114 FVPYQ
CpPOT1 346 TTCGTTCCCTACCAA CpPOT1 116 FVPYQ
GhPOT1 364 TTCCAACCTTATCAA GhPOT1 122 FQPYQ
PtPOT1 349 TTCCTACCTTATCAA PtrPOT1 117 FLPYQ
NtPOT1 358 TTCCTTCCCTACCAA NtPOT1 120 FLPYQ
StPOT1 358 TTCCTCCCTTATCAA StPOT1 120 FLPYQ
LsPOT1 355 TTTGTACCTTATCAA LsPOT1 119 FVPYQ
HaPOT1 364 TTTTCACCTTATCAA HaPOT1 122 FSPYQ
D
C
L132
0.982
AtPOT1a 391 GCATTATACAAGAGTGCCATGGATGAT AtPOT1a 131 ALYKSAMDDLRKVFA
AlPOT1a 391 GCACTATACAAGAGTGCGATGGCTGAT AlPOT1a 131 ALYKSAMADLRKFFA
BoPOT1a 406 ---AGATACAAGAGTTCTATGGAGGAG BoPOT1a 136 -RYKSSMEELAKLFP
AtPOT1b 391 CAGGATAAGAACTTCTTATCG---AAT AtPOT1b 131 QDKNFLS-NLREWMI
AlPOT1b 379 CAGGATCACAACTTGTTATCT---AAT AlPOT1b 127 QDHNLLS-NLREWMI
BoPOT1b 382 CAGGATAAGAGCTTTTTATCA---AAT BoPOT1b 128 QDKSFLS-NLREWMI
CpPOT1 388 CAAGACAAGGAGTTTATAGCT---GGT CpPOT1 130 QDKEFIA-GLRRWLV
GhPOT1 406 CATGATGCAATGATATAGCA---GGT GhPOT1 136 HDAMIIA-GLRKWLA
PtPOT1 388 CATGACTCGAAGTTTATAGCA---GGC PtrPOT1 130 HDSKFIA-GIRKWLV
NtPOT1 400 CAAGACAAGAAATTTATATTG---GAT NtPOT1 134 QDKKFIL-DLMKWLE
StPOT1 400 CAAGACAAGAAATTTATACTG---GGT StPOT1 134 QDKKFIL-GLMKWLV
LsPOT1 397 CAAGACAACAAGTTCATAGCA---GAC LsPOT1 133 QDNKFIA-DLRKWTI
HaPOT1 406 CAAGACAAGAAGTTCATCGCT---GGC HaPOT1 136 QDKKFIA-GLRKWTV
E
pot1a ku70
+ WTAtPOT1a
+ L132D
+ E35F
kb
7
6
5
4
3
2
1.6
1
0.8
0.6
0.5
1 2 3 4 5 6
F
pot1a ku70
+ WT AtPOT1a
+ C119P
kb
7
6
5
4
3
2
1.6
1
0.8
0.6
0.5
1 2 3 4
G
Complementation efficiency, %
120
100
80
60
40
20
0
untransformed
WT AtPOT1a
L132D
E35F
E35F L132D
C119P
H
untransformed
+ WT POT1a
+ E35F
+ L132D

To exaggerate the differences in telomere length between wild-type plants and *pot1a* mutants, complementation test constructs were introduced into a *ku70* null background, where telomeres are elongated in an AtPOT1a-dependent manner (Fig 6-5A & B). KU heterodimer (KU70/KU80) is a negative regulator of telomere length in *Arabidopsis* (Riha et al, 2002). In its absence, telomeres are dramatically elongated by about two fold per plant generation (Riha & Shippen, 2003; Surovtseva et al, 2007) (Fig 6-5B). If a transgenic construct fully complements a POT1a deficiency in the *ku70* background, telomeres will become dramatically elongated, just as in *ku70* mutants (Fig 6-5A).

A transgenic POT1a construct was introduced into pot1a -/- ku70+/- plants. pot1a ku70 transformants were selected and subjected to TRF analysis. To keep the expression of POT1a transgene similar to wild type level, a putative POT1a native promoter (a 1.5 kb fragment upstream of POT1a gene) was used to drive transgene expression. In addition, an N-terminal 3×Flag tag was attached to each construct to detect exogenous POT1 protein expression (Fig 6-5C). As predicted, introduction of full-length WT POT1a into this background led to telomere elongation by 2-3 kb compared to untransformed plants or plants transformed with an empty vector (Fig 6-5D, compare lanes 4, 5 to lanes 1-3). Thus, a sensitive genetic complementation assay was successfully developed to examine POT1a activity *in vivo*.

Fig 6-5. Genetic complementation system for *POT1a*. (**A**) A complementation system for *POT1a* was set up in a *pot1a ku70* background. (**B**) TRF analysis of WT, *pot1a*, *ku70*, and *pot1a ku70* mutants (Surovtseva et al, 2007). In the absence of *POT1a*, telomeres are progressively shortened by ~ 250 bp per plant generation (lanes 5 and 6). In *ku70* mutants, telomeres are elongated by ~ 2-3 kb per plant generation (lanes 3 and 4). Telomeres remain short in *pot1a ku70* mutants (lanes 7 and 8). (**C**) POT1a transgenes were driven by its putative native promoter. Telomere length of *pot1a ku70* transformants was analyzed to calculate complementation efficiency. (**D**) TRF results of untransformed *pot1a ku70* mutants (lanes 1 and 2), transforments with an empty vector (lane 3), and mutants with a transgenic copy of WT POT1a (lanes 4 and 5).


A
WT
pot1a -/-
-250bp
ku70 -/-
+ 2-3kb
Telomere elongation in ku70 -/- is dependent on AtPOT1a
pot1a -/-
ku70 -/-
Transformation of AtPOT1a
Telomere Elongation
ku70 -/-
B
WT
ku70
pot1a
pot1a ku70
kb
7
6
5
4
3
2
1.6
1
1 2 3 4 5 6 7 8
C
Native promoter
LB
BASTAR
N-3xFlag-POT1a
pCBK05
RB
transformation
pot1a -/- ku70 +/-
pot1a -/- ku70 -/- with a transgenic POT1a
pot1a -/- ku70 -/- with no transgene or an empty vector
- control
D
pot1a ku70
+ vector
+ WT POT1a
kb
8
6
5
4
3
2
1.6
1
0.8
1 2 3 4 5

Sites of positive selection are required for AtPOT1a function in vivo

If AtPOT1a was indeed subjected to extensive evolutionary sweep, substituting positively selected amino acids with consensus residues found in other dicot POT1 proteins should decrease AtPOT1a complementation efficiency. In contrast to conventional Alanine substitution, this mutagenesis strategy will revert the POT1a sequence back to the highly invariable ancestral residues, effectively erasing millions of years of protein evolution (Chang et al, 2002). E35, L132 and C119 were mutated to Phe (E35F), Asp (L132D) and Pro (C119P), respectively.

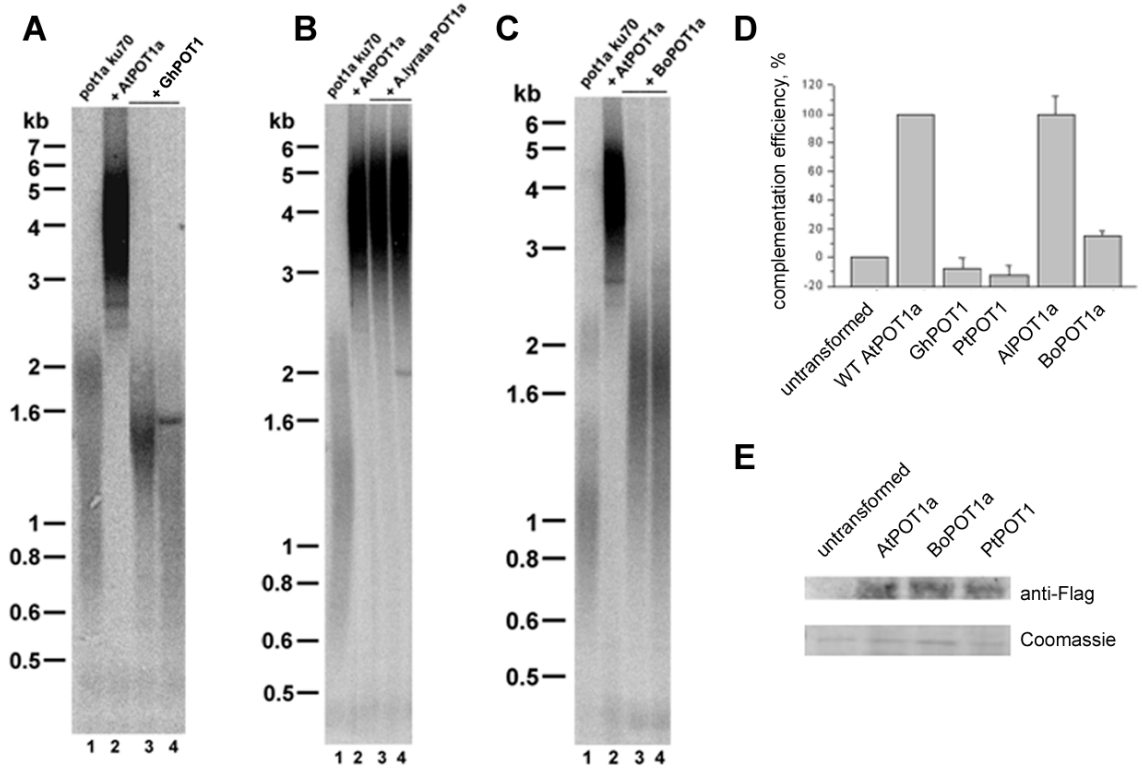
Based on our threading models, E35 is located in β-strand 1 (β1), which lies in proximity to the putative nucleic acid binding interface in AtPOT1a, while L132 is located in the C-terminal α-helical region, suggesting that it is not directly involved in ssDNA binding (Fig 6-4D). To compare different POT1 transgene activity *in vivo*, we set the complementation efficiency of WT AtPOT1a transgene as one and the untransformed plants as zero. The complementation levels of all other transgenic constructs were then compared to the WT AtPOT1a transgene. Interestingly, the level of telomere extension (or complementation efficiency) in *pot1a* null plants expressing exogenous AtPOT1a L132D was reduced by 20% (Fig 6-4E, lanes 3, 4 and Fig 6-4G) of the wild-type AtPOT1a levels. This value was reduced to 50% in transgenic plants expressing AtPOT1a E35F (Fig 6-4E, lanes 5, 6 and Fig 6-4G). The double E35F L132D mutant displayed complementation efficiency similar to the AtPOT1a E35F mutant (Fig 6-4G). These results demonstrate that both E35 and L132 contribute to AtPOT1a function *in vivo*.

C119 has the third most significant posterior probability value in OB1 (0.868). Based on our threading models, C119 is located in the loop connecting β-strand 5 (β5) to the C-terminal α-helix. The C119P mutation converts the Brassicaceae POT1a-specific Cys to a Pro, which is conserved in all other plant POT1 proteins, including those from algae (Fig 6-4B). Notably, this AtPOT1a-C119P transgene fully complemented the AtPOT1a deficiency (Fig 6-4F, lanes 3, 4 and Fig 6-4G). Thus, either residue C119 has little effect on AtPOT1a function, or both Pro and Cys can function similarly in this context. This experiment illustrates the need for functional validation of *in silico* predictions.

Neo-functionalization of POT1a

The neo-functionalization model of gene evolution predicts that if the *POT1a* genes from Brassicaceae species have acquired a novel function, a single-copy *POT1* gene from closely related non-Brassicaceae plants will fail to complement AtPOT1a deficiency. Consistent with this prediction, the single copy *POT1* genes from *Gossypium hirsutum* and *Populus trichocarpa*, which shared the last common ancestor with *Arabidopsis* 85 and 100 mya, respectively (Wikstrom et al, 2001), failed to complement AtPOT1a deficiency (Fig 6-6A, lanes 3, 4; Fig 6-6D and data not shown).

Fig 6-6. Cross-species complementation analysis of *AtPOT1a* deficiency. (**A**) TRF analysis of transgenic plants expressing *Gossypium hirsutum* (Gh) POT1. Lane 1, untransformed *pot1a ku70*; lane 2, *pot1a ku70* mutants expressing WT AtPOT1a, *pot1a ku70* mutants expressing GhPOT1 (lanes 3, 4). (**B**) and (**C**) TRF analysis of transgenic plants expressing POT1 proteins from other Brassicaceae species. (**B**) Lane 1, untransformed *pot1a ku70*; lanes 2-4, *pot1a ku70* mutants expressing WT AtPOT1a (lane 2), or *A. lyrata* POT1a (lanes 3, 4). (**C**) Lane 1, untransformed *pot1a ku70*; lanes 2-4, *pot1a ku70* mutants expressing WT AtPOT1a (lane 2) or *Brassica oleracea* (Bo) POT1a (lanes 3, 4). (**D**) A histogram of relative complementation efficiency of *POT1* transgenes from *Gossypium hirsutum*, *Populus trichocarpa*, and *POT1a* genes from *Arabidopsis lyrata* and *Brassica oleracea* comparing to WT AtPOT1a. (**E**) Top, a representative western blot with an anti-Flag antibody to show the expression of exogenous POT1 proteins. Bottom, the same amount of protein samples used in the western blot were loaded onto another gel for coomassie staining.


A
pot1a ku70
+ AtPOT1a
+ GhPOT1
kb
7
6
5
4
3
2
1.6
1
0.8
0.6
0.5
1 2 3 4
B
pot1a ku70
+ AtPOT1a
+ A.lyrata POT1a
kb
6
5
4
3
2
1.6
1
0.8
0.6
0.5
1 2 3 4
C
pot1a ku70
+ AtPOT1a
+ BoPOT1a
kb
6
5
4
3
2
1.6
1
0.8
0.6
0.5
1 2 3 4
D
complementation efficiency, %
120
100
80
60
40
20
0
-20
untransformed
WT AtPOT1a
GhPOT1
PtPOT1
AtPOT1a
BoPOT1a
E
untransformed
AtPOT1a
BoPOT1a
PtPOT1
anti-Flag
Coomassie

Most of the candidate sites under positive selection with relatively high posterior probability values in AtPOT1a are similar or identical to sites in POT1a protein from Arabidopsis lyrata, a species that shared the last common ancestor with A. thaliana only ~ 5.2 mya (Koch et al, 2000) (Fig 6-7). As expected, cross-species complementation using AlPOT1a cDNA fully rescued AtPOT1a deficiency phenotype (Fig 6-6B, lanes 3, 4, and Fig 6-6D). By comparison, Arabidopsis thaliana and Brassica oleracea diverged ~ 40 mya (Beilstein M., personal communication) and BoPOT1a protein exhibits 74% similarity to AtPOT1a overall. Strikingly, BoPOT1a displayed only ~ 15% of complementation efficiency relative to wild-type AtPOT1a (Fig 6-6C, lanes 3, 4 and Fig 6-6D), indicating that the protein function of POT1a from distant members of Brassicaceae has significantly diverged.

Taken together, our data suggest specific diversifications occur in the POT1a lineage, but not the POT1b lineage in Brassicaceae. Thus, significant functional amino acid differences have accumulated since POT1 genes are duplicated in the Brassicaceae family.

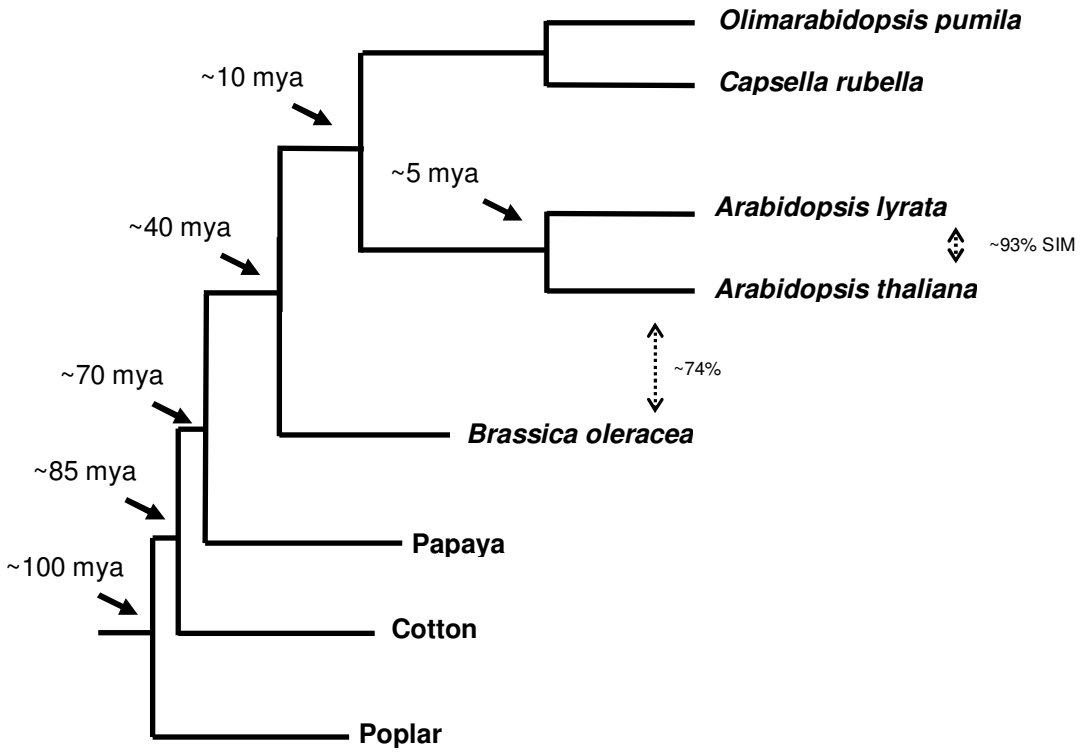


Fig 6-7. Phylogenetic relationship of Brassicaceae and other dicot species. Approximate divergence times are shown above each node. Percent overall amino acid similarity between corresponding POT1a proteins in indicated.

Disscussion

Implications of positive selection for POT1a function

Unlike POT1 proteins from yeast and mammals that are esstial for telomere protection, none of the *Arabidopsis* POT1 paralogs directly binds telomeric DNA *in vitro*. AtPOT1a interacts with TER1 and positively regulate telomere length *in vivo*. In contrast, AtPOT1b binds to TER2 and negatively regulates telomerase activity *in vivo*. Moreover, POT1b and POT1c are involved in G-overhang protection. But the mechanism is yet to be determined. It is proposed that *Arabidopsis* POT1 proteins evolve from a DNA binding protein to a telomerase RNA associated factor.

Here we employed a combination of evolutionary, structural and molecular genetic approaches to analyze representative *POT1* genes across the plant kingdom. We obtained POT1 sequences from 30 distinct plants, which evolved within the last one billion years. The identification of a large number of plant POT1 sequences provided an opportunity to search for evidence of selective pressure on the *Arabidopsis* POT1 proteins. The BEB test identified a number of positively selected sites in POT1a, which were prioritized on the basis of BEB-assigned statistical values and the availability of a high quality structural model. To investigate the functional contribution of specific residues within AtPOT1a *in vivo*, we developed a genetic complementation system that provides a quantitative read-out of POT1a activity. Notably, many of our candidate positive selection sites, including the highly significant E35 and L132, localize to OB1, the most highly conserved region in all eukaryotic POT1 proteins. Structural modeling suggests that none of these sites is likely to play an important role in direct nucleic acid binding, although E35 lies in the vicinity of DNA-binding pocket. We hypothesize that

E35 and L132 could be involved in protein-protein interactions and higher order complex assembly, or play roles in protein stability.

POT1 *gene duplication is rare and is accompanied by rapid functional diversification.*

Gene duplication is a major source of evolutionary advances (Gilbert et al, 1997). However, despite multiple ancient whole-genome polyploidization events in vertebrates, yeasts and plants (Cui et al, 2006; Dehal & Boore, 2005; Kellis et al, 2004), *POT1* gene remains single-copy in most eukaryotic genomes. Among the 30 representatives of the green plant lineage we surveyed, only two instances of independent *POT1* gene duplication were detected, one in the dicot Brassicaceae family and one in the monocot Panicoideae subfamily of grasses. The most common outcome of gene duplication is the accumulation of deleterious mutations in one member of the pair, followed by subsequent silencing and eventual gene loss (Lynch & Conery, 2000; Moore & Purugganan, 2003; Walsh, 1995). Given the usually larger number of average gene family members in plants, including the well-documented increase in genes encoding putative double-strand telomere binding proteins in *Arabidopsis* (Karamysheva et al, 2004), our finding argues that duplication of POT1 may be deleterious, unless the duplicated genes can be preserved through beneficial neo-functionalization events.

In the neo-functionalization model for the retention of duplicated genes, one gene copy acquires a novel, beneficial function, while the other copy retains most of the ancestral gene functions (Lynch & Conery, 2000). The current data are consistent with this model for the evolution of *POT1* genes in *Arabidopsis*. First, overexpression of a dominant-negative of AtPOT1b or depletion of AtPOT1c results in telomere deprotection phenotypes similar to yeast and vertebrate *pot1* mutants (Shakirov et al, 2005). In

addition, POT1b interacts with TER2 and negatively regulates telomerase. In contrast, AtPOT1a protein has evolved novel interactions with the telomerase RNP and is required for enzyme function *in vivo* (Surovtseva et al, 2007). Second, the data presented here reveal an evolutionary sweep that targeted several regions in AtPOT1a, diverging them from consensus amino acids present in all other plant POT1 proteins. In most cases, the corresponding residues in AtPOT1b remained unchanged and closely resemble the consensus sequence. Third, the inability of single-copy *POT1* genes from close relatives of *Arabidopsis* (*Gossypium hirsutum* and *Populus trichocarpa*) to complement AtPOT1a deficiency strongly argues for neo-functionalization. One explanation for this failure is that POT1 proteins from non-Brassicaceae species cannot interact with the binding partner(s) of AtPOT1a. For example, GhPOT1 or PtPOT1 may not be able to contact AtTER1 and thus fail to complement AtPOT1a deficiency. A similar model has been proposed for the mouse POT1 paralogs (Hockemeyer et al, 2007). Although expression of a single-copy human POT1 fails to complement mouse POT1a or POT1b deficiencies, co-expression of human POT1 with its interacting partner TPP1 rescues the phenotypes associated with mouse POT1a mutation. Interestingly, mouse POT1b deficiency is not rescued with these same constructs, suggesting that mouse POT1b evolved a novel function or interacts with a different mouse protein (Hockemeyer et al, 2007). The latter would be consistent with neo-functionalization model of evolution not for just one gene, but for the entire mouse POT1b gene network. Thus, recent acquisition of novel functions for POT1 may be a common theme for plants as well as animals.

The two most distantly related Brassicaceae species analyzed here, *Arabidopsis thaliana* and *Brassica oleracea*, shared the last common ancestor only 40 mya

(Beilstein M., personal communication). Thus, the limited complementation of AtPOT1a deficiency by the BoPOT1a protein suggests that Brassicaceae POT1a proteins are under strong selective pressure and are continuing to diverge rapidly. Recent data also suggest that *Arabidopsis* TERT may also be experiencing an elevated mutation rate (Shakirov et al, 2008). Thus, evolutionary pressure may be simultaneously acting on several components of the telomerase enzyme and the protective telomere cap. Given its genetic tractability and the presence of sequence homologues for many important proteins involved in human telomere biology, *Arabidopsis* may prove to be an excellent model to study evolution of the entire telomere maintenance network.

CHAPTER VII

DISSECTION OF POT1a FUNCTION IN *Arabidopsis*

Summary

Protection Of Telomeres 1 (POT1) is a single copy gene in yeasts and humans that plays essential roles in chromosome end protection. In contrast, the flowering plant *Arabidopsis thaliana* encodes three highly divergent POT1 proteins, termed AtPOT1a, AtPOT1b and AtPOT1c. Previous studies indicate that AtPOT1b and AtPOT1c contribute to chromosome end protection, while AtPOT11a functions in the telomerase pathway for telomere maintenance. AtPOT1a binds the telomerase RNA subunit TER1, and is required for telomerase enzyme activity *in vitro* and *in vivo*. Here we show that POT1a binds to a telomere capping protein CTC1 *in vitro* and *in vivo*. This result suggests that POT1a is a telomerase recruitment factor that delivers telomerase to CTC1 *in vivo*. To further investigate the function of POT1a, a genetic complementation assay was used to dissect POT1a functional domains. We found that a Phe residue (F65) in the first OB-fold of AtPOT1a, corresponding to a conserved amino acid that contacts telomeric DNA in human POT1, is critical for *Arabidopsis* telomere length regulation *in vivo*. In addition, deletion of the last five amino acids from the extreme C-terminus, or substitution of Thr 463 into Ala significantly reduced POT1a function. To identify more *pot1a* alleles, we screened a collection of ethylmethanesulfonate (EMS)-mutagenized *Arabidopsis* plants. This screen uncovered a novel *pot1a* allele, *pot1a-3*, which contains D385N mutation in the protein C-terminus. *pot1a-3* mutants exhibit progressively shorter telomeres through plant generations, similar to a *pot1a* null. Preliminary results revealed that mutations of F65A and D385N reduce POT1a affinity

for TER1, but do not decrease CTC1 interaction. Altogether, our data suggest that both the N-terminal OB-folds and the C-terminus of POT1a are required for its function in *Arabidopsis*.

Introduction

In most eukaryotes, telomeres terminate in a 3' extrusion of ss G-rich DNA, called the G-overhang. The presence of the G-overhang and G-overhang binding proteins help maintain the architecture of linear chromosome ends and protect them from being recognized as double strand breaks (Palm & de Lange, 2008). In budding yeast, Cdc13 shows strong affinity to ss G-strand telomeric DNA via an oligonucleotide/oligosaccharide binding fold (OB-fold) (Lin & Zakian, 1996). Cdc13 interacts with Stn1 and Ten1 to form a trimeric complex of CST, which plays an essential role in telomere capping (Garvik et al, 1995). Recently, a CST complex was identified in plants and vertebrates, which contains STN1, TEN1 and a novel protein, Conserved Telomere maintenance Component 1 (CTC1) (Miyake et al, 2009; Surovtseva et al, 2009). CTC1 has structural and functional similarities to Cdc13. It directly associates with ss telomeric DNA *in vitro*, localizes to telomeres in *vivo*, and together with STN1 and TEN1, it protects the integrity of chromosome ends.

In addition to CTC1, POT1 is another well-studied ss telomere binding protein in fission yeast and mammals. POT1, like Cdc13/CTC1, is an OB-fold containing protein. It belongs to a six-member telomere protein complex called Shelterin, whose full composition includes two ds telomere binding proteins TRF1 and TRF2, as well as RAP1, TIN2 and TPP1, which bridge the ds telomere binding priteins to POT1 (Palm &

de Lange, 2008). TPP1 binds POT1, while TIN2 contacts both TPP1 and TRF2 to hold all Shelterin components together.

In fission yeast and humans, *POT1* is a single-copy gene. Deletion of *POT1* in fission yeast is lethal and leads to catastrophic telomere loss (Baumann & Cech, 2001). In humans, POT1 is involved in chromosome end protection as well as telomere length regulation (Colgin et al, 2003; Loayza & de Lange, 2003; Yang et al, 2005). On one hand, POT1 transduces a regulative signal from TRF1 to inhibit telomere elongation (Loayza & de Lange, 2003). On the other hand, TPP1 together with POT1 associate with the telomerase RNP *in vivo* and increase telomerase processivity *in vitro* (Wang et al, 2007; Xin et al, 2007). It is proposed that the POT1-TPP1 complex plays a dual role in telomerase regulation: first by limiting telomerase access to telomeres; and then increasing telomerase processivity once the telomeric DNA is bound by telomerase (Xin et al, 2007).

Several organisms encode more than one *POT1* genes, including ciliates, worms, mice and *Arabidopsis* (He et al, 2006; Hockemeyer et al, 2006; Jacob et al, 2007; Raices et al, 2008; Shakirov et al, 2005; Wu et al, 2006). These POT1 paralogs make distinct contributions to telomere biology. In *Tetrahymena*, there are two POT1 homologs. While TtPOT1b's function is still unknown, TtPOT1a is a negative regulator of telomere length and prevents activation of a cell-cycle checkpoint (Jacob et al, 2007). In *C.elegans*, CeOB1 and CeOB2 are found to show structural similarities to the second and first OB-fold of human POT1, respectively (Raices et al, 2008). While CeOB1 binds ss G-strand telomeric DNA, CeOB2 binds ss C-strand telomeric DNA. CeOB1 negatively regulates telomere length and is required for proper G-overhang structure. In contrast, depletion of CeOB2 causes telomere length heterogeneity (Raices et al, 2008).

Finally, in mouse, mPOT1a and mPOT1b share 72% sequence similarity and are partially redundant (Hockemeyer et al, 2006). Although *pot1b* null mice are viable, knockout of *POT1a* results in embryonic lethality. mPOT1a represses the DNA damage response at chromosome ends, whereas mPOT1b is implicate in regulation of ss G-overhang (Hockemeyer et al, 2006). Both mPOT1a and mPOT1b are implicated in repression of non-homologous end joining and homologous recombination at telomeres (He et al, 2006; Wu et al, 2006). Taken together, these findings point to a rapid evolution of POT1 proteins.

Arabidopsis harbors three *POT1* genes, *AtPOT1a*, *AtPOT1b* and *AtPOT1c*. AtPOT1a and AtPOT1b, like their counterparts in vertebrates and fission yeast, encode two N-terminal OB-folds (OB1 and OB2) followed by a C-terminal extension (Shakirov et al, 2005). AtPOT1a and AtPOT1b are highly divergent, and share only 49% protein sequence similarity to each other. AtPOT1c encodes a small protein corresponding to a single OB-fold, and shows similarities to both OB1 and OB2 of AtPOT1a. While the functions of AtPOT1b and AtPOT1c are still not clear, both of them are required for proper G-overhang regulation (A.D.L. Nelson and D.E. Shippen, unpublished work). Overexpression of a dominant negative allele of POT1b or depletion of POT1c results in genome instability, implicating POT1b and POT1c in chromosome end protection (Shakirov et al, 2005; A.D.L. Nelson and D.E. Shippen, unpublished work). In contrast, *pot1a* null mutants display progressive telomere shortening at the same rate as telomerase mutants, but the chromosome ends, at least initially, remain fully protected (Surovtseva et al, 2007). AtPOT1a acts in the same genetic pathway as telomerase. Strikingly, none of the *Arabidopsis* POT1 proteins bind telomeric DNA. Biochemical data indicate that AtPOT1a binds TER1 and associates with telomerase activity (C.

Cifuentes-Rojas and D.E. Shippen, unpublished work). Thus, AtPOT1a has evolved to bind TER1 instead of ss G-rich telomeric DNA (Shakirov et al, 2009).

In this study, we report that POT1a interacts with the telomere capping protein CTC1 *in vitro* and *in vivo*. The data suggest that POT1a is a recruitment factor for telomerase that brings the enzyme to chromosome ends via interaction with TER1 and CTC1. To define functional domains in POT1a required for these interactions, we used two strategies. First, we performed site-directed mutagenesis and examined mutant POT1a activity by a genetic complementation assay. Second, we screened a collection of ethylmethanesulfonate (EMS)-mutagenized TILLING (Targeting Induced Local Lesions IN Genomes) for novel *pot1a* alleles. Analysis of these mutants revealed that residues in OB1 as well as the C-terminus are critical for POT1a function.

Materials and methods

Plant materials and plasmid construction

The *pot1a-3* mutant (CS95038) was obtained from *Arabidopsis* Biological Resource Center. Genotyping was performed by PCR with 5'-TTGGGACACATTTCATTCTGGTGT -3' and 5'-TCATTAATGAAGTAGTCTAGTACCAAAG -3', followed by sequencing with both primers. Plants were grown according to the conditions described (Surovtseva et al, 2007). For complementation, wild type or mutant POT1a cDNA was cloned into a binary vector pCBKO5 under the control of the putative native promoter of POT1a (~ 1.5kb upstream from the start codon of POT1a gene). Plant transformation was performed as described (Surovtseva et al, 2007).

RT-PCR

Total RNA was extracted from plant tissues using an RNA purification kit (Fisher Scientifics). Reverse transcription was performed using Superscript III reverse transcriptase (Invitrogen) per the manufacturer instructions. PCR of *POT1a* cDNA was performed using primers 5'- ATGGCGAAGAAGAGAGAGAG -3' and 5'- TTAATGAAGTAGTCTAGTACCAAAG -3', with the following program: 95 ℃ 3 min; 25 cycles of 94 ℃ 20 sec, 55 ℃ 30 sec, 72 ℃ 2 min 30 sec; 72 ℃ 7 min.

In vitro *co-immunoprecipitation, immunoprecipitation and western blot analysis*

In vitro co-immunoprecipitation was performed as previously described (Karamysheva et al, 2004). For each reaction, one protein was constructed with a T7-tag, while the other protein was ^{35}S methionine-labeled. The two proteins were synthesized in a TNT-coupled rabbit reticulocyte lysate translation system according to manufacturer instructions (Promega). Translation of unlabled proteins was verified by the presence of ^{35}S methionine in a small aliquot from the same master mix. T7-tagged unlabled proteins and untagged radiolabeled proteins were combined and subjected to immunoprecipitation using anti-T7 antibody-conjugated agarose beads (Novagen). Pellet and supernatant fractions were then analyzed by SDS-PAGE and autoradiography.

For plant extracts immunoprecipitation, ~ 0.5 g of WT or CTC1-CFP seedlings (Surovtseva et al, 2009) were grown and harvested. The protein was extracted using W buffer (50 mM Tris·acetate pH 7.5, 5 mM $MgCl_2$, 100 mM potassium glutamate, 20 mM EGTA, 1.0 mM DTT, 0.6 mM vanadyl ribonucleoside complex, 1.5% polyvinylpyrrolidone, 10% glycerol) as described (Fitzgerald et al, 1996). After pre-clear with protein-A beads for 1hr, the supernatant of the protein extract was incubated with 10ul POT1a peptide antibody and pre-blocked protein-A beads for 3 hr. The beads were washed three times with W300 buffer (20mM Tris·acetate pH 7.5, 10% Glycerol, 1mM EDTA, 5mM $MgCl_2$, 200mM NaCl, 300mM potassium glutamate, 1% NP40, 0.5mM sodium deoxycholate, 1mM DTT), and once with TMG buffer (10mM Tris·acetate pH 8.0, 1mM $MgCl_2$, 10% Glycerol, 1mM DTT). The precipitates were resuspended in TMG buffer and subjected to western blot analysis or the telomere repeat amplification protocol (TRAP). For CTC1-CFP immnoprecipitation, anti-GFP antibody conjugated beads (Abcam) were used following a similar immnoprecipitation protocol. Western blot was performed as described (Surovtseva et al, 2007), using a peptide antibody raised against AtPOT1a (Surovtseva et al, 2007), an anti-Flag antibody (Sigma), or an anti-GFP antibody (Abcam) as indicated.

TRF analysis, PETRA, TRAP and Q-TRAP

Genomic DNA was extracted using a cetyltrimethylammonium bromide (CTAB)-based method as described (Cocciolone & Cone, 1993). TRF analysis was performed using 50 µg of DNA digested with *Tru1*I (Fermentas) overnight at 65℃ (Fitzgerald et al, 1999). The digested DNA samples were subjected to a Southern blot using a telomeric probe.

For POT1a complementation studies, the average telomere length (L) was measured using Telometric-1.2 program (Grant et al, 2001). The average telomere lengths of untransformed *pot1a ku70* mutants, transformants expressing wild-type *AtPOT1a* and other *POT1* constructs were designated as L_0, L_1, and L_x, respectively. We set the complementation level of wild-type *AtPOT1a* transformants (positive control) as one, and that of untransformed *pot1a ku70* mutants (negative control) as zero. The complementation efficiency (*E*) of each *POT1* construct was calculated as: $E = (L_x - L_0) / (L_1 - L_0) \times 100\%$. At least three individual transformants for each construct were analyzed for statistical support.

Primer extension telomere repeat amplification (PETRA) analysis was carried out with 2 µg DNA using 2U *Ex-Taq* polymerase (Takara) per reaction and performed as previously described (Heacock et al, 2004; Watson & Shippen, 2007). The blots were hybridized with a ^{32}P 5' end–labeled $(T_3AG_3)_4$ oligonucleotide probe. Telomerase activity was examined by TRAP (Fitzgerald et al, 1996) or a real-time telomere repeat amplification protocol (qTRAP) (Herbert et al, 2006; Kannan et al, 2008).

Results

POT1a interacts with CTC1 in vitro *and* in vivo

CTC1 has been recently identified as a novel telomere-associated component in plants and vertebrates (Miyake et al, 2009; Surovtseva et al, 2009) (see Chapter III). CTC1, like budding yeast Cdc13, directly associates with ss telomeric DNA *in vitro*, and forms a complex with STN1 and TEN1 to protect the chromosome ends *in vivo* (Miyake et al, 2009; Surovtseva et al, 2009) (J.R. Lee and D.E. Shippen, unpublished work). While budding yeast Cdc13 is involved in telomerase recruitment, it is currently unknown whether CTC1 has a similar function. To explore the possibility, we

investigated whether *Arabidopsis* CTC1 interacts with POT1a, a novel telomerase accessory factor that binds TER1. Co-immunoprecipitation was performed with recombinant POT1a and different domains of CTC1 *in vitro* (Fig 7-1A, top). POT1a was associated with CTC1-B, CTC1-C, CTC1-D proteins, but not CTC1-A. CTC1-B corresponds to the C-terminus of the protein, which covers regions of both CTC1-C and CTC1-D (Fig 7-1A, bottom). These results were confirmed by a reciprocal binding experiment (Fig 7-1A, top), indicating that POT1a directly associates with CTC1 *in vitro*.

Next, we asked if POT1a interacts with CTC1 in plant extracts. Protein samples from wild type or a transgenic line with a CFP-tagged CTC1 (Surovtseva et al, 2009) were incubated with anti-GFP antibody conjugated beads. As expected, western blot analysis showed that CTC1-CFP protein was precipitated by these beads (Fig 7-1B, top). In addition, POT1a was detected in the pull-down fraction of CTC1-CFP sample, but not wild type extracts. This result indicates that POT1a is associated with CTC1 *in vivo* (Fig 7-1B, middle). Since POT1a binds TER1, we asked whether CTC1 is associated with this RNA. RT-PCR showed that TER1 was precipitated with CTC1-CFP (Fig 7-1C), indicating that CTC1 interacts with TER1 *in vivo*. Finally, we asked whether CTC1 associates with enzymatically active telomerase by Telomere Repeat Amplification Protocol (TRAP). TRAP assay revealed that telomerase was pulled down in the CTC1-CFP, but not wild type sample (Fig 7-1D). These findings indicate that CTC1 interacts with the telomerase RNP complex containing POT1a. It remains to be determined whether the CTC1-telomerase interaction is dependent on POT1a.

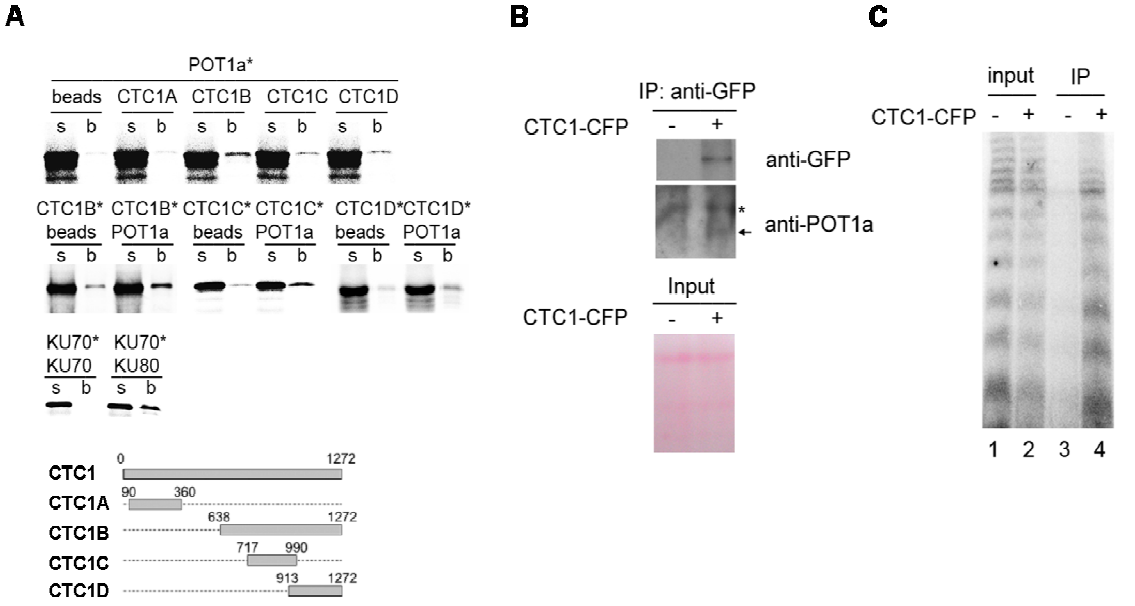


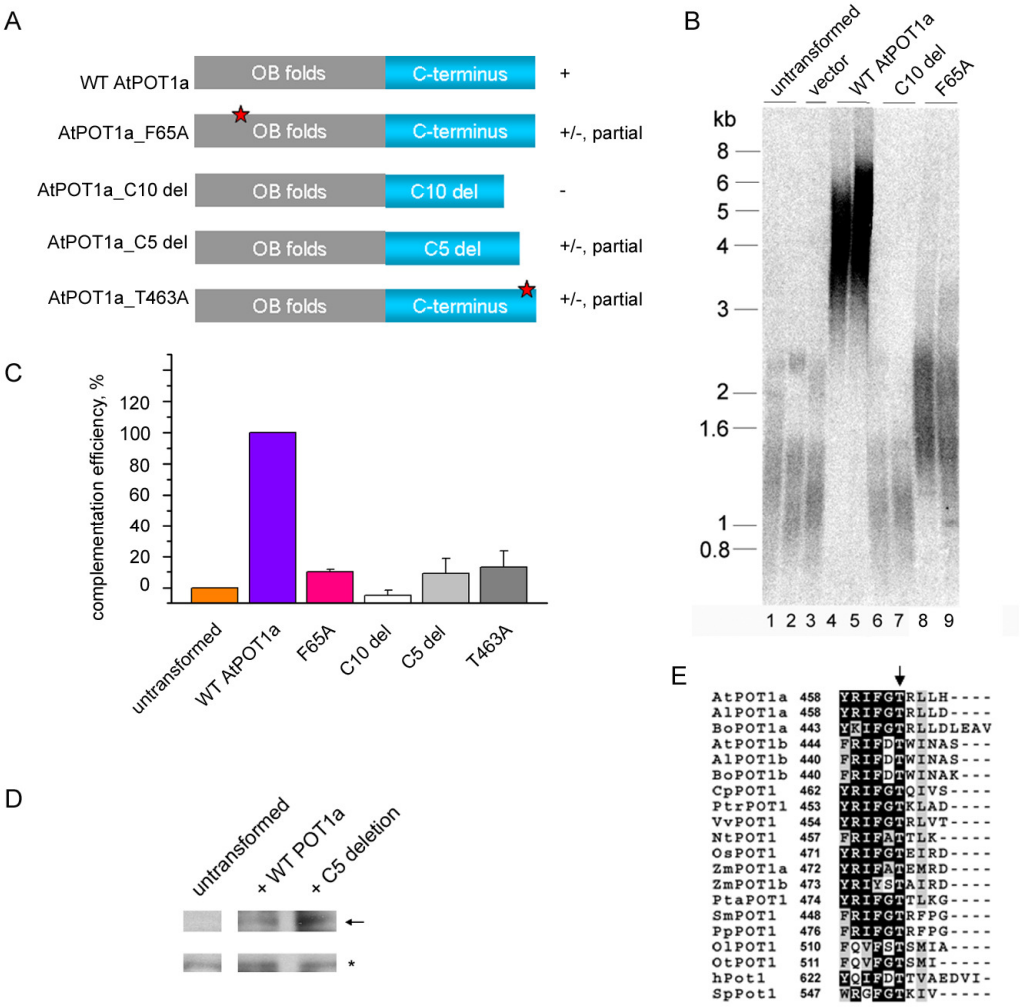
Fig 7-1. POT1a interacts with CTC1 *in vitro* and *in vivo*. (**A**) POT1a interacts with CTC1 *in vitro*. Top, *in vitro* co-immunoprecipitation results with POT1a and different domains of CTC1 are shown. ^{35}S Methionine labeled protein is indicated with an asterisk; The T7-tagged non-radiolabled protein is shown underneath. s, supernatant; b, bound. KU70/KU80 serves as a positive control; KU70/KU70 as a negative control. Bottom, diagram of different CTC1 domains. (**B**) CTC1 is associated with POT1a *in vivo*. Protein extracts from wild type or transformants with a CFP-tagged CTC1 were immunoprecipitated with anti-GFP conjugated beads, and then subject to a western blot. The western blot was performed with either an anti-GFP antibody to detect CTC1 (top) or an anti-POT1a peptide antibody (middle). Inputs were run on a separate gel and stained with Ponceau S as a loading control (bottom). (**C**) CTC1 is associated with TER1. RT-PCR results of *TER1*, *U6* and *ACTIN* from the input and immunoprecipitated samples. (**D**) CTC1 is associated with active telomerase. The input and immunoprecipitated protein samples were examined by TRAP assay.

A Phe residue in OB-fold 1 is critical for AtPOT1a function in vivo

The interaction of POT1a with TER1 and CTC1 promoted us to dissect the functional domains in the protein. Previous studies showed that the OB-folds of POT1a are critical for interacting with TER1 *in vitro* (C. Cifuentes-Rojas and D.E. Shippen, unpublished work). Computer modeling of AtPOT1a OB1 onto the crystal structure of hPOT1 failed to detect any significant structural variation that could account for AtPOT1a's preference for TER1 as opposed to ss telomeric DNA (Croy, J. and Wuttke, D., University of Colorado, see Chapter VI). In fact, a Phenyalanine residue, which plays a critical role in telomeric DNA interaction in human POT1 (Lei et al, 2004b; Wu et al, 2006a), is conserved in *Arabidopsis* POT1a, that is, F65.

To examine POT1a functional domains, a genetic complementation approach was used (see Chapter VI). Introduction of a construct bearing an F65A mutation in POT1a only partially rescued the shortened telomeres in the complementation assay (Fig 7-2B, lanes 10 and 11). The complementation efficiency of F65A was about 10% compared to wild type POT1a (Fig 7-2C). Since the corresponding Phe in human POT1 is critical for binding ss telomeric oligo, we asked whether F65 in AtPOT1a contributes to TER1 interaction. Preliminary results showed that recombinant POT1a-F65A protein has reduced TER1 binding activity (Fig 7-3), suggesting that the Phe residue responsible for contacting nucleic acids in hPOT1 is conserved in AtPOT1a and is critical for the protein function.

Fig 7-2. AtPOT1a function relies on Phe 65 in OB1 and the last ten amino acids at the extreme C-terminus. (**A**) Diagram of different POT1a transgenic constructs. Red stars show position of mutations. The relative complementation efficiency is shown on the right for each construct. (**B**) TRF results are shown for untransformed *pot1a ku70* mutants, transformed with an empty vector, or constructs expressing wild type AtPOT1a, POT1a C10 del and POT1a F65A. (**C**) Histogram of complementation efficiency levels with different POT1a constructs. The complementation efficiency of each construct was calculated as described in Materials and Methods. Telomere elongation in untransformed plants was set as 0%, and that in wild type is set as 100%. Error bars show standard deviation. (**D**) Western blot results confirming that the mutant protein C5 del does not disturb protein stability. Arrow points to a band specific for POT1a protein; the asterisk indicates a nonspecific band as a loading control. (**E**) Alignment of the extreme C-terminus of POT1 proteins from different plants, human and fission yeast. Arrow indicates Thr 463 in AtPOT1a.


A
WT AtPOT1a
OB folds
C-terminus
+
AtPOT1a_F65A
OB folds
C-terminus
+/-, partial
AtPOT1a_C10 del
OB folds
C10 del
-
AtPOT1a_C5 del
OB folds
C5 del
+/-, partial
AtPOT1a_T463A
OB folds
C-terminus
+/-, partial
B
untransformed
vector
WT AtPOT1a
C10 del
F65A
kb
8
6
5
4
3
2
1.6
1
0.8
1 2 3 4 5 6 7 8 9
C
complementation efficiency, %
120
100
80
60
40
20
0
untransformed
WT AtPOT1a
F65A
C10 del
C5 del
T463A
D
untransformed
+ WT POT1a
+ C5 deletion
*
E
AtPOT1a 458 YRIFGTRLLH----
AlPOT1a 458 YRIFGTRLLD----
BoPOT1a 443 YKIFGTRLLDLEAV
AtPOT1b 444 FRIFDTWINAS---
AlPOT1b 440 FRIFDTWINAS---
BoPOT1b 440 FRIFDTWINAK---
CpPOT1 462 YRIFGTQIVS----
PtrPOT1 453 YRIFGTKLAD----
VvPOT1 454 YRIFGTRLVT----
NtPOT1 457 FRIFATTLK-----
OsPOT1 471 YRIFGTEIRD----
ZmPOT1a 472 YRIFATEMRD----
ZmPOT1b 473 YRIYSTAIRD----
PtaPOT1 474 YRIFGTTLKG----
SmPOT1 448 FRIFGTRFPG----
PpPOT1 476 FRIFGTRFPG----
OlPOT1 510 FQVFSTSMIA----
OtPOT1 511 FQVFGTSMI-----
hPot1 622 YQIFDTTVAEDVI-
SpPot1 547 WRGFGTKIV-----

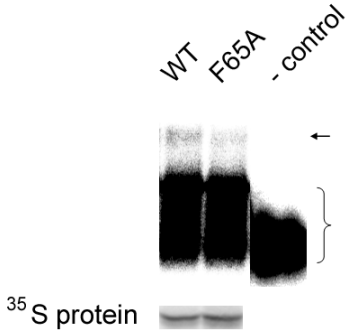


Fig 7-3. Mutation of F65A in POT1a disrupts the interaction with TER1 RNA *in vitro*. Gel shift results with radiolabeled TER1 RNA and wild type, F65A mutant POT1a. A no protein control (- control) is also shown. Arrow points to a gel-shifted band. Bracket shows the free TER1 probe. TER1 RNA was transcribed in the presence of ^{32}P CTP *in vitro*. Recombinant protein was translated in rabbit reticulocyte lysate. In a separate reaction, recombinant protein was expressed in the presence of ^{35}S Methionine and run on a SDS-PAGE gel to monitor protein expression levels.

The C-terminus of POT1a is required for its function

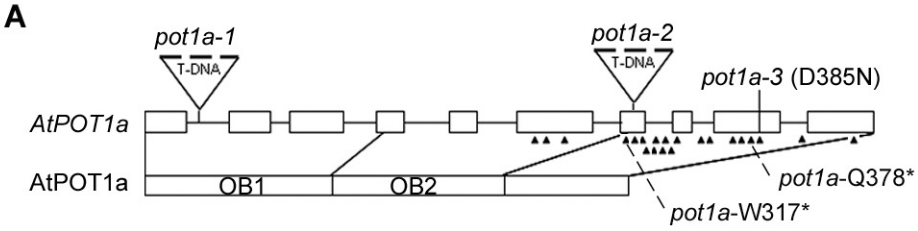
In mammals, POT1's recruitment to telomeres relies on its association with TPP1. The TPP1 interaction domain has been mapped to a short conserved motif in extreme C-terminus of hPOT1 (Liu et al, 2004b). Although no sequence homolog of TPP1 can be detected in the *Arabidopsis* genome, we found that at least one of the TPP1-interacting motifs is conserved in AtPOT1a (Fig 7-2E), corresponding to the last ten amino acids. Deletion of this motif (C10 del) completely abolished AtPOT1a activity *in vivo* (Fig 7-2B, lanes 8 and 9). Furthermore, a construct with only five amino acids deleted at the C-terminus (C5 del), rescued the telomere length phenotype to only 9% the level of wild type POT1a, indicating that these five amino acids are critical for POT1a function (Fig 7-2C). We confirmed that the mutant protein was expressed at a similar level as wild type POT1a protein (Fig 7-2D).

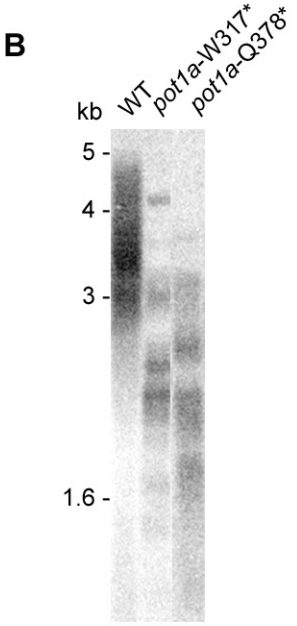
There is a highly conserved Thr residue within the last five amino acids of AtPOT1a (Fig 7-2E). Mutation of this residue to Ala (T463A) significantly reduced POT1a function *in vivo*, resulting in only ~ 10% complementation relative to wild type POT1a (Fig 7-2C). Thus, Thr 463 plays an important role for POT1a *in vivo*. Because Thr is a potential target of protein kinases, one interesting possibility is that T463 modulates AtPOT1a function through its phosphorylation status. Taken together, AtPOT1a's role in telomere length regulation is modulated by residues in the extreme C-terminus. Our data also suggest the presence of a TPP1-like protein in *Arabidopsis*, which contacts the AtPOT1a C-terminus.

A TILLING screen to identify novel pot1a *alleles*

In an effort to identify novel *pot1a* alleles, we screened a collection of EMS-mutagenized plants from Seattle *Arabidopsis* TILLING Project. The initial trial was targeted for mutations in the 5' end of the *POT1a* genomic sequence, which corresponds to the N-terminal OB-folds. Unfortunately, all the mutations recovered localized to *POT1c* locus and not *POT1a* (data not shown). POT1c displays 92% identify with the 5' terminus of POT1a and so this result is not too surprising. A second genomic region (~ 1.2 kb) was chosen for further screening, covering the 6^{th}-10^{th} exons of POT1a (Fig 7-4A). This screen yielded several new alleles of *pot1a* (Fig 7-4A and Table 7-1). 14 lines carried mutations in POT1a exons (groups I to III), while the others had mutations in the introns (group IV). Group I mutants contain nonsense mutations that result in premature stop codons; group II mutants carry missense mutations; and group III mutants contain synonymous mutations.

Fig 7-4. A screen of TILLING mutants for novel *pot1a* alleles. (**A**) Schematic of *pot1a* TILLING mutants. The structure of the *AtPOT1a* gene is shown as exons (rectangle boxes) and introns (lines), with the corresponding protein domains indicated beneath. Positions of T-DNA insertions are indicated in the previously identified *pot1a-1* and *pot1a-2* alleles (Surovtseva et al, 2007). Black triangles denote point mutations uncovered in TILLING mutants (also see Table 7-1). Dashed lines indicate positions of nonsense mutations in *pot1a-W317** and *pot1a-Q378**. A vertical line corresponds to *pot1a-3*, a mutant that contains a substitution of Asp 385 to Asn (D385N). (**B**) TRF analysis of *pot1a-W317** and *pot1a-Q378** mutants. The blot was hybridized with a radiolabeled G-rich telomeric probe. Molecular weight markers are indicated. (**C**) Alignment of POT1 proteins from different species. At, *Arabidopsis thaliana*; Bo, *Brassica oleracea*; Hs, *Homo sapiens*; Mm, *Mus musculus*; Gg, *Gallus gallus*; Sp, *Schizosaccharomyces pombe*.


A
pot1a-1
T-DNA
pot1a-2
T-DNA
pot1a-3 (D385N)
AtPOT1a
AtPOT1a
OB1
OB2
pot1a-Q378*
pot1a-W317*


B
kb
WT
pot1a-W317*
pot1a-Q378*
5 -
4 -
3 -
1.6 -

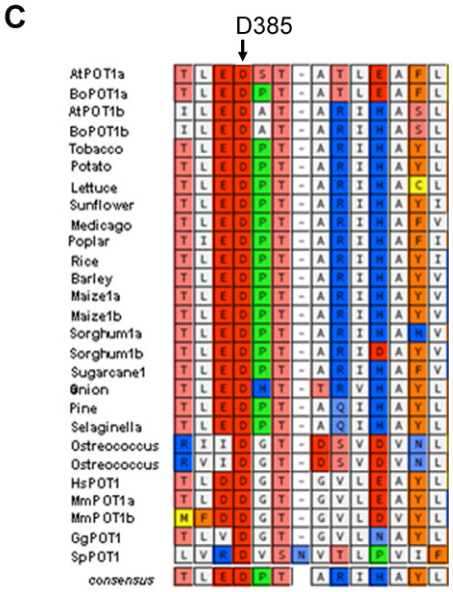

C
D385
AtPOT1a
BoPOT1a
AtPOT1b
BoPOT1b
Tobacco
Potato
Lettuce
Sunflower
Medicago
Poplar
Rice
Barley
Maize1a
Maize1b
Sorghum1a
Sorghum1b
Sugarcane1
Onion
Pine
Selaginella
Ostreococcus
Ostreococcus
HsPOT1
MmPOT1a
MmPOT1b
GgPOT1
SpPOT1
consensus

Table 7-1. AtPOT1a TILLING mutant lines.

Group	Nucleotide change	Effect	ABRC stock
I	G1660A	W317*	CS92009
			CS93904
	C2029T	Q378*	CS85995
II	C1317T	P238S	CS93939
	C1429T	S275F	CS93849
	C1676T	P323S	CS91331
	G1680A	G324E	CS92100
	G1996A	D367N	CS87091
			CS91835
	G2050A	D385N	CS95038
III	C1305T	L234=	CS95221
	C1815T	A333=	CS90504
	G2013A	E372=	CS95381
	C2308T	L443=	CS92004
IV	G1689A	intron	CS91901
	C1698T	intron	CS92020
	G1721A	intron	CS95000
	C1757T	intron	CS95536
	C1772T	intron	CS93714
	C1781T	intron	CS88421
	C1885T	intron	CS86438
	G1923A	intron	CS94929
	C2155T	intron	CS90788
			CS93786

Note: for G1660A, G1996A, and C2155T mutants, two stock lines were available at the stock center.

We first examined bulk telomere length in group I mutants (*pot1a*-W317* and *pot1a*-Q378*), which carry a nonsense mutation at Trp 317 and Gln 378, respectively. Using TRF analysis, we found that telomeres were shorter in both mutants than in wild type and displayed a distinct sharp banding pattern (Fig 7-4B). Similar results were observed with *pot1a* null mutants that harbor a T-DNA insertion (Surovtseva et al, 2007). Thus, the two nonsense mutations result in loss-of-function alleles. As expected, all mutants from group III (synonymous mutations) and group IV (mutations in introns), and most mutants from group II (missense mutations) did not show any telomere length defects (data not shown).

Identification of a pot1a-3 *mutant*

One missense mutant from group II harbors a G to A mutation in the 7th exon of *POT1a* gene, which leads to a substitution of Asp 385 to Asn (D385N) (Fig 7-4A, Table 7-1). This *Arabidopsis* TILLING mutant line was designated *pot1a-3*. Charged amino acids like Asp contribute to electrostatic forces between molecules (Voet et al, 2006). Interestingly, Asp 385 is located in a conserved region in all plant POT1 proteins examined, including both POT1a and POT1b from *Arabidopsis* and cauliflower and the single copy POT1 proteins from poplar, rice and single cell alga (Fig 7-4C). In the case of human and yeast POT1, there may be a conserved Asp residue in the same region (Fig 7-4C).

pot1a-3 mutants, like all *Arabidopsis pot1a* mutants identified, did not show any morphological defects (data not shown). RT-PCR revealed that the mRNA expression of *POT1a* was not disturbed in *pot1a-3* mutants (Fig 7-5A). We cloned and sequenced full-length *POT1a* cDNA in *pot1a-3* mutants, and confirmed that D385N is the only mutation in the corresponding protein sequence. Immunoprecipitation followed by western blotting indicated that the mutant protein POT1a was expressed at a level similar to wild type POT1a (Fig 7-5B).

Progressive telomere shortening in pot1a-3 *mutant*

Using TRF analysis, we found that the telomeres in *pot1a-3* mutants were much shorter than those in wild type plants and displayed a homogeneous banding pattern (Fig 7-5C), similar to those in *pot1a* null mutants. Notably, *pot1a-3* mutants were already homozygous when we obtained them from the stock center. The telomere length of *pot1a-3* mutants was similar to that of fourth generation (G4) *pot1a-1* mutants, suggesting the *pot1a-3* line has been propagated for several generations at the stock center. Bulk telomeres of *pot1a-3* mutants were shorter in the next generation (Fig 7-5C). To measure the rate of telomere loss in *pot1a-3* mutants from parent to progeny, primer extension telomere repeat amplification (PETRA) was employed. Using this assay, a decline of approximately 250bp was observed in *pot1a-3* progeny on the two chromosome ends examined (e.g. the right arm of chromosome 2 and the left arm of chromosome 3) (2R and 3L) (Fig 7-5D). A similar rate of telomere shortening was reported for *pot1a* null mutants (Surovtseva et al, 2007). Taken together, these data indicate that amino acid Asp 385 is essential for POT1a function *in vivo*.

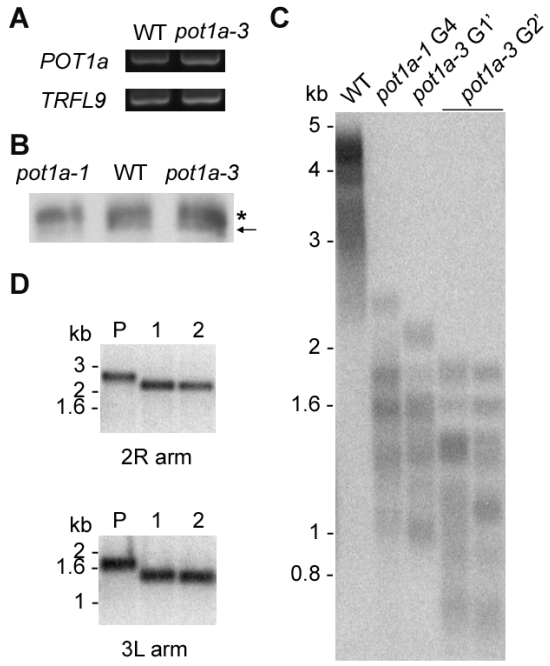


Fig 7-5. Identification and characterization of *pot1-3* allele. (**A**) RT-PCR of *POT1a* coding region in wild type and *pot1a-3* mutants. *TRFL9* is shown as a loading control. (**B**) Similar levels of POT1a protein expression were detected in *pot1a-3* mutants and wild type plants by immunoprecipitation and western blot with a POT1a peptide antibody (Surovtseva et al, 2007). *pot1a-1* null mutant is shown as a negative control. Asterisk shows a nonspecific band for loading control. Arrow indicates a band specific for POT1a protein. (**C**) TRF analysis of a first generation *pot1a-3* mutant obtained from the stock center (G1') and its progeny (G2'). Results for wild type and a G4 *pot1a-1* mutant are shown for comparison. (**D**) PETRA results for *pot1a-3* mutants with a parent (P) and two progeny (1 and 2). The blot was hybridized with a telomeric probe. Telomere length on 2R and 3L was measured.

The telomere defect in pot1a-3 *mutants is linked to D385N mutation*

As TILLING plants are known to harbor multiple mutations in a single mutant line (Greene et al, 2003), we asked whether the telomere defect in *pot1a-3* is linked to D385N mutation in POT1a. The *pot1a-3* mutant was backcrossed to wild type plants, and the resulting heterozygote was self-pollinated and segregated. As expected, TRF analysis of wild type and heterozygous *pot1a-3* mutants showed normal telomere length profile (Fig 7-6A). Only plants homozygous for POT1a-D385N displayed telomere shortening defects (n=9, Fig 7-6A and data not shown). This result suggests that the D385N mutation is linked to telomere defects in *pot1a-3* mutants. Consistent with this finding, introduction of a transgenic copy of wild-type POT1a complemented the telomere length defect in *pot1a-3* mutants (Fig 7-6B).

The D385N mutation in POT1a disturbs in vitro *telomerase activity and affects POT1a interaction with TER1* in vitro

One intriguing defect of *pot1a* null mutants is that the protein extracts exhibit significantly reduced telomerase activity (Surovtseva et al, 2007). Real-time telomere repeat amplification protocol (Q-TRAP) revealed that telomerase activity was reduced by 90% in *pot1a* protein samples relative to wild type (Surovtseva et al, 2007) (A.D.L. Nelson and D.E. Shippen, unpublished data). To ask whether the D385N mutation in POT1a affects telomerase activity, we performed Q-TRAP on *pot1a-3* mutant samples. A similar level of decreased telomerase activity was observed in *pot1a-3* mutants (~ 11% of wild type samples), comparable to that in *pot1a-2* null mutants (Fig 7-7A). Thus, Asp 385 in POT1a is required for maintaining telomerase activity.

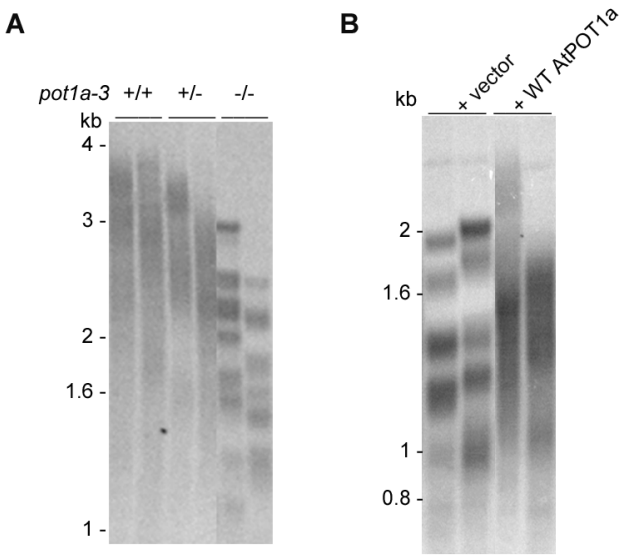


Fig 7-6. The progressively shortening telomere defect in *pot1a-3* is linked to mutation of D385N. (**A**) TRF analysis of segregants from a self-pollinated heterozygous *pot1a-3* parent. Two individuals each of wild type, *pot1a-3* heterozygous and homozygous mutants are shown. (**B**) Introduction of a transgenic copy of wild type POT1a complemented the telomere length defects in *pot1a-3* mutants. Two individuals each are shown for *pot1a-3* mutants transformed with an empty vector and a wild type *POT1a* transgene.

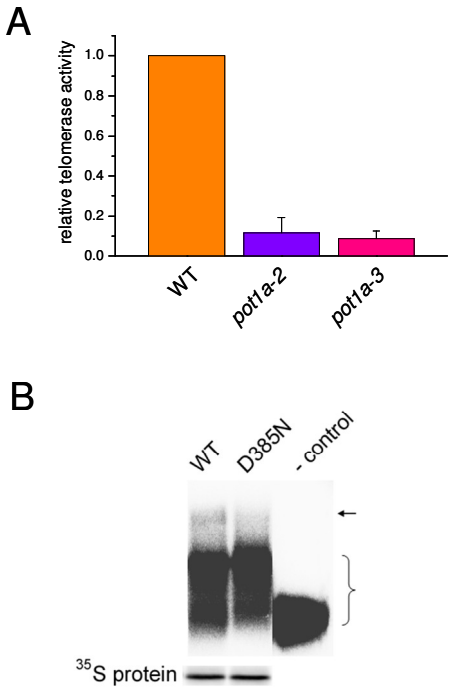


Fig 7-7. Mutation of D385N in POT1a disturbs telomerase activity and affects POT1a interaction with TER1 *in vitro*. (**A**) Q-TRAP was performed with protein extracts from wild type (WT), *pot1a-2* and *pot1a-3* mutants (n=3). The telomerase activity obtained from WT extracts was set to one and each sample was normalized to this value. Data are represented as mean ± SEM. (**B**) D385N mutation in POT1a disrupts its interaction with TER1 RNA. Top, gel shift analysis of TER1 RNA with POT1a wild type or D385N mutant protein. The TER1 RNA was transcribed at the presence of ^{32}P CTP *in vitro*. Arrow indicates a gel-shifted band. Bracket denotes the free TER1 probe. Recombinant protein was translated in rabbit reticulocyte lysate and then subjected to gel shift with TER1 RNA. Bottom, in a separate reaction, recombinant protein was expressed in the presence of ^{35}S Methionine and run on a SDS-PAGE gel as a loading control.

While the OB-fold 1 of POT1a is sufficient for TER1 interaction *in vitro* (C. Cifuentes-Rojas and D.E. Shippen, unpublished work), the D385N mutation in the C-terminus of POT1a protein reduced TER1 binding activity (Fig 7-7B). Although this result is preliminary, it suggests that the C-terminal domain of POT1a can modulate TER1 interaction. Consistent with this observation, substitution of the corresponding Asp to Asn in Asparagus POT1 disrupts its interaction with telomeric oligo nucleotides (E.V. Shakirov and D.E. Shippen, unpublished work).

Discussion

In contrast to fission yeast and humans, where there is only a single *POT1* gene, *Arabidopsis* harbors three POT1 paralogs. The duplication of *POT1* genes in *Arabidopsis* and other plant species was discussed in Chapter VI. It is believed that positive selection shapes the function of POT1a in the Brassicaceae family. POT1a has evolved a novel function in telomere maintenance. It contacts TER1 and is required for telomerase activity *in vitro* and *in vivo*. The interaction of POT1 proteins with TER has not been reported in other species, although hPOT1-TPP1 complex enhances telomerase processivity *in vitro* (Lei et al, 2005; Wang et al, 2007).

In this study, I show evidence for a novel AtPOT1a function associated with the telomere capping protein CTC1 *in vitro* and *in vivo*. I propose that POT1a recruits telomerase to telomeres by interacting with both TER1 and CTC1. In an analogous situation, the telomerase recruitment in budding yeast relies on an interaction between Est1 and Cdc13 (Pennock et al, 2001). There are several similarities between POT1a and Est1. First, both POT1a and Est1 bind to telomerase RNA (Seto et al, 2002) (C. Cifuentes-Rojas and D.E. Shippen, unpublished work). Second, both are positive

regulators of telomere length (Lundblad & Szostak, 1989; Surovtseva et al, 2007). Third, similar to Est1 which associates with Cdc13 (Pennock et al, 2001), preliminary data indicate an interaction between POT1a and CTC1. We do notice some differences between Est1 and POT1a. For example, unlike POT1a, Est1 is dispensable for telomerase activity (Lingner et al, 1997a). In addition, Est1 exhibits a weak interaction with ss telomeric DNA (Virta-Pearlman et al, 1996), which has not been identified for POT1a. The interaction between Est1 and ss telomeric DNA is thought to contribute to telomerase recruitment to the chromosome ends. Future studies of POT1a should be focused on when and how it is interacted with CTC1 *in vivo*, and whether this interaction is required and sufficient to deliver telomerase to telomere ends.

To dissect the functional domains of POT1a, two different strategies were employed. First, we developed a sensitive complementation assay to monitor POT1a activity *in vivo*. Our data indicate that a Phe residue (F65) in the OB1, which corresponds to a residue protein that contacts telomeric DNA in human POT1, is important for AtPOT1a function. Moreover, the proper function of POT1a requires the last five amino acids in the extreme C-terminus, especially a Thr residue (T463). Preliminary data also indicate that Asp 385 and Phe 65 contribute to the POT1a-TER1 interaction. Second, a collection of EMS-mutagenized *Arabidopsis* TIILING plants were screened to identify novel *pot1a* mutant alleles. Through this screen, an Asp residue (Asp 385) was identified to be conserved in all plant POT1 proteins and critical for POT1a function *in vivo*. Altogether, these data indicate that both the N-terminus and the C-terminus of POT1a promote its association with telomerase and positive regulation of telomere length.

Notably, a significant reduction in telomerase activity was detected in the TILLING mutant *pot1a-3*, which carries a D385N mutation in POT1a. A similar level of reduced telomerase activity was observed in *pot1a* null mutants (Surovtseva et al, 2007) (A.D.L. Nelson and D.E. Shippen, unpublished work). As the D385N mutation disturbs POT1a-TER1 binding *in vitro*, it is reasonable to suspect that POT1a regulate telomerase and telomere length through binding to TER1. One explanation is that POT1a is necessary for telomerase enzyme stability. Consistent with this idea, a *Tetrahymena* protein p65 was reported to play critical roles in telomerase RNA accumulation (Witkin & Collins, 2004). POT1a may also act to increase the telomerase access to telomeres, or enhance its processivity as has been shown for hPOT1 (Lei et al, 2005; Wang et al, 2007). Experiments are currently underway to examine the interaction of different POT1a mutants with TER1, which may help to elucidate how POT1a is involved in telomerase regulation. Overall, our findings of POT1a interaction with TER1 and CTC1 provide new insight into the mechanisms of telomerase recruitment and telomere maintenance.

CHAPTER VIII

CONCLUSIONS AND FUTURE DIRECTIONS

Telomeres play pivotal roles in preserving genome stability in eukaryotes. One of the biggest challenges in the field is to characterize the full compositioin of telomere-associated proteins and to elucidate the relationship between different telomere components. Another fundamental aspect of telomere biology is to explore how telomeres are maintained by telomerase and the replication machinery, and how this process is regulated.

Originally identified in ciliates, telomeres and telomerase are widely conserved in yeasts, plants, and animals. Studies from numerous labs revealed that different eukaryotes share similar telomeric DNA sequences, telomere end structure, telomerase components and telomere-associated proteins. In particular, the model plant *Arabidopsis thaliana* is a fascinating system for telomere biology studies because of its short lifespan, the sequenced genome, easy access to mutant stocks, genetic tractability and most importantly, the extreme tolerance to genome instability.

In this dissertation, I present the identification and characterization of a previously unknown telomere capping/maintenance complex CTC1/STN1/TEN1 in *Arabidopsis* (Chapters II to V). Moreover, I report how a Shelterin-like component AtPOT1a has evolved a novel function as a telomerase accessory component (Chapters VI and VII). In this chapter, conclusions from these studies and speculations of future directions are described.

CTC1/STN1/TEN1 complex plays a crucial role in telomere integrity in *Arabidopsis*

Previous studies showed that telomeres in budding yeast are protected by a trimeric complex of Cdc13/Stn1/Ten1 (ScCST), whereas a six-member Shelterin complex protects the chromosome ends in vertebrates and fission yeast. Our data show that the integrity of telomeres in *Arabidopsis* and humans also relies on a protein complex of CTC1/STN1/TEN1 (AtCST or hCST), which localizes at telomeres consistently during the cell cycle and plays a critical role in chromosome end maintenance and protection (Miyake et al, 2009; Surovtseva et al, 2009).

In Chapter II, I report the discovery that loss of *STN1* in *Arabidopsis* induces catastrophic telomere loss, elevated G-overhang signal, frequent chromosome end-to-end fusions, and aberrant recombination at telomeres. Furthermore, STN1 specifically localizes at telomeres *in vivo*. These data indicate the presence of a second telomere-capping complex containing STN1, which is distinct from the previously characterized Shelterin complex. In Chapter III, a novel telomere capping protein CTC1 was identified in *Arabidopsis*. *Arabidopsis ctc1* mutants display the same telomere defects as *stn1* mutants. Moreover, we discovered that CTC1 physically interacts with STN1, and

together they protect the chromosome ends in the same genetic pathway. In Chapter IV, I present the identification of a TEN1 ortholog in *Arabidopsis*. AtTEN1 interacts with STN1 and CTC1. Reduced expression of AtTEN1 results in deregulation of telomere length as well as chromosome end-to-end fusions. Taken together, *Arabidopsis* CTC1/STN1/TEN, like ScCST, forms a trimeric complex and plays an essential role in telomere end protection. To note, *Arabidopsis stn1* and *ctc1* null mutants are viable for at least two plant generations, despite the exhibition of catastrophic telomere loss and massive chromosome fusions (Chapters II & III). In contrast, loss of Stn1 causes senescence and cell death in budding yeast. Thus, *Arabidopsis* is a great model organism to study telomere-related gene functions.

While we were characterizing *Arabidopsis* CST, human CTC1/STN1/TEN1 complex was independently discovered by Fuyuki and colleagues (Miyake et al, 2009). In comparison to *Arabidopsis* CST, where loss-of-function alleles are available, study of human CST relies on gene knockdown. Reduced expression of CTC1 results in sporadic telomere loss on some chromosome ends, increased G-overhang signals, as well as chromatin bridges (Surovtseva et al, 2009). Knockdown of STN1 also results in increased G-overhang signals (Miyake et al, 2009). Moreover, double knockdown of STN1 and POT1 leads to synergistically elevated TIF cells (Miyake et al, 2009). Thus, mammalian CST and Shelterin components play redundant roles to maintain telomere end architecture as well as to protect chromosome termini.

CTC1/STN1/TEN1 appears to be a higher-eukaryote counterpart of the budding yeast Cdc13/Stn1/Ten1 complex. Like budding yeast CST, CTC1, STN1 and TEN1 in plants and mammals contain OB-folds that show structural similarity to those of RPA proteins (Miyake et al, 2009; Song et al, 2009; Surovtseva et al, 2009). While Cdc13

specifically binds to ss G-rich telomeric DNA (Lin & Zakian, 1996; Nugent et al, 1996), human CTC1 contacts ss DNA *in vitro* in a sequence nonspecific manner (Miyake et al, 2009). This result suggests that hCTC1 is recruited to telomeres by other telomere-associated components. Nonetheless, the CTC1-ss DNA association indicates that CTC1/STN1/TEN1 is involved in ss telomere maintenance in multicellular eukaryotes. Last but not least, functional studies revealed that CTC1, STN1 and TEN1 are critical for G-overhang maintenance and telomere capping in *Arabidopsis* and in humans. Interestingly, mammalian CTC1/STN1 has been recently found to correspond to Polymerase Alpha Accessory Factor (AAF), which stimulates Polymerase alpha activity *in vitro* (Casteel et al, 2009; Goulian & Heard, 1990). Moreover, preliminary data revealed that *Arabidopsis* CTC1 interacts with the catalytic subunit of Pol alpha *in vitro* (X. Song, K.A. Boltz and D.E. Shippen, unpublished work). Likewise, budding yeast Cdc13 and Stn1 are also involved in Polymerase alpha interaction (Grossi et al, 2004; Qi & Zakian, 2000), which is interpreted to coordinate telomere replication. In summary, CTC1/STN1/TEN1 in multicellular eukaryotes shares structural, biochemical and functional similarities with budding yeast Cdc13/Stn1/Ten1. CST complex is thus much more conserved than previously thought and provides another set of players that, in addition to the existing Shelterin components, protect telomere integrity in multicellular eukaryotes.

Testing a working hypothesis for CST function in Arabidopsis

In the current working model, we propose that AtCTC1 binds ss G-overhang and recruits STN1/TEN1 to protect telomere integrity *in vivo* (Fig 8-1). Supporting this hypothesis, Dr. Jung Ro Lee in our lab revealed that AtCTC1 directly interacts with ss

DNA *in vitro*, but it is not yet known if this interaction is specific for G-strand telomeric DNA. In comparison, STN1, TEN1 alone or co-expressed STN1/TEN1 failed to show DNA binding (J.R. Lee and D.E. Shippen, unpublished work). To test if CTC1 interacts with telomeric DNA *in vivo*, chromatin immunoprecipitation (ChIP) should be performed using an anti-CTC1 antibody. This antibody is currently under production. Alternatively, we can utilize a transgenic line with a tagged-CTC1 for telomere ChIP analysis. One important goal would be to follow CTC1-telomere association throughout the cell cycle. This will allow us to study the dynamics of CTC1 action at telomeres. It is expected that CTC1 is associated with telomeres throughout the cell cycle, but the association may peak at the S-phase, when the t-loop is supposed to unfold and the chromosome ends are extended and processed.

To uncover the role of CTC1 in STN1/TEN1 recruitment, the association of STN1/TEN1 with telomeres in a *ctc1* mutant could be examined using immunofluorescence of STN1/TEN1 in combination with Telomere FISH. If STN1/TEN1 is recruited to telomeres by CTC1, it is expected that loss of CTC1 will disrupt STN1/TEN1 telomere localization. Alternatively, we can test whether the major role of CTC1 is to recruit STN1 by fusing STN1 to the DNA binding domain of CTC1. If the chimera rescues the *ctc1* null phenotype, this would support our model. Similar experiments have been done in budding yeast, where Stn1 is fused to the DNA binding domain of Cdc13 (Pennock et al, 2001) and the fusion protein rescues the lethality of *cdc13* mutant.

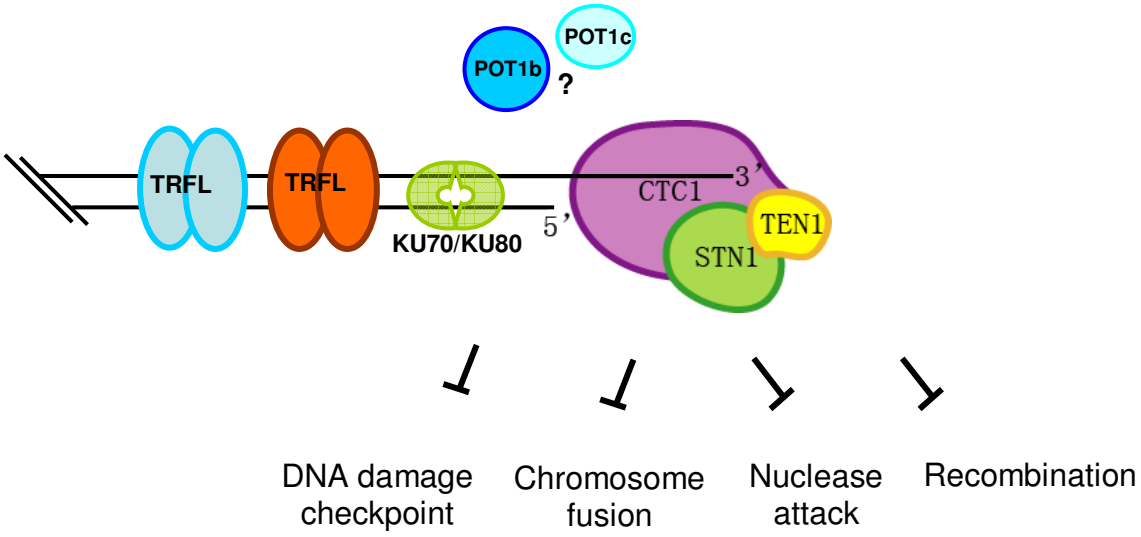


Fig 8-1. A model of how *Arabidopsis* telomeres are protected by CST and other telomere-associated components. Several protein complexes have been reported to protect telomere integrity in *Arabidopsis*, including CST, Shelterin (TRFLs and POT1b/POT1c) and KU70/KU80. While there appears to be no direct interaction between POT1b/POT1c and telomeric DNA, CTC1 binds ss DNA together with STN1/TEN1. In addition, TRFL proteins and KU70/KU80 bind to ds telomeric DNA and are required for chromosome end protection (see text for details).

It will be more challenging to elucidate the mechanism of STN1/TEN1 capping function. One hypothesis is that STN1/TEN1 stabilizes CTC1 and/or promotes CTC1 function when forming a trimeric complex. To test this hypothesis, we need to investigate whether *Arabidopsis* STN1/TEN1 boosts CTC1-telomeric DNA interactions. This seems to be true in both ScCST and hCST (Miyake et al, 2009; Qian et al, 2009). In addition, it remains to be determined whether STN1/TEN1 affects CTC1 protein stability, its localization to telomeres, or its interaction with Pol alpha *in vivo*. These activities of CTC1 can be examined in a *stn1* or *ten1* mutant background.

Another possibility is that STN1/TEN1 has a CTC1-independent protection function at telomeres. Supporting this idea, overexpression of Stn1 and Ten1 rescues the telomere uncapping defect of *cdc13-1* mutants (Grandin et al, 1997; Petreaca et al, 2006). In addition, a conditional mutant of Ten1, which does not affect Cdc13's association with telomeres, still causes accumulation of ss G-strand telomeric DNA (Xu et al, 2009). These data revealed a Cdc13-independent role of Stn1 and Ten1 in budding yeast. In *Arabidopsis*, overexpression of AtSTN1 rescues the telomere deprotection phenotype in a *ctc1-1* mutant (Y.V. Surovtseva and D.E. Shippen, unpublished work). While additional controls are needed to draw firm conclusion from these experiments, it is possible that AtSTN1 has a CTC1-independent telomere capping function. As described later, the identification of new STN1-interaction partners may help elucidate how STN1 executes the protection function.

Different telomere capping complexes in Arabidopsis

Besides CST, there are several other telomere-related factors that are implicated in telomere capping in plants (Fig 8-1). First, *Arabidopsis* harbors at least six Myb-

bearing proteins (TRFLs) that specifically bind ds telomeric DNA *in vitro* (Karamysheva et al, 2004). Among these, AtTBP1 acts as a negative regulator of telomere length (Hwang & Cho, 2007). Rice mutants lacking RTBP1 display gradual telomere lengthening and exhibit telomere fusions in G2, consistent with a role in telomere capping (Hong et al, 2007). Second, AtPOT1b and AtPOT1c are involved in chromosome end protection. Over-expression of a dominant negative allele of POT1b or depletion of POT1c leads to a telomere uncapping phenotype similar to a *pot1* deficiency in yeast and mammals (Shakirov et al, 2005; A.D.L. Nelson and D.E. Shippen, unpublished work). Both proteins are required to maintain G-overhang structure. Surprisingly, neither POT1b nor POT1c binds ss telomeric DNA *in vitro*. Therefore, it is unclear how POT1b and POT1c contribute to telomere protection (Fig 8-1). Third, telomere integrity in *Arabidopsis* also requires KU70/KU80. KU70/KU80 binds to ds telomeric DNA and acts as a negative regulator of telomere length in *Arabidopsis* (Gallego et al, 2003; Riha & Shippen, 2003). Moreover, KU is implicated in maintaining proper G-overhang architecture and prohibiting degradation of the C-rich telomeric strand (Riha & Shippen, 2003). Loss of KU results in increased G-overhang signals similar to *ctc1* or *stn1* mutants, although to a more moderate level. In summary, CST, Shelterin (TRFLs and POT1b/POT1c) and KU all contribute to telomere protection in *Arabidopsis*.

In Chapter V, I show that CST protects G-overhang structure in a different genetic pathway from KU70/KU80. In addition, *in vitro* co-IP failed to detect any interaction between CST components and KU. It is possible that KU70/KU80 sits at the ds/ss telomere junction and inhibits exonuclease activity on the C-strand telomeric DNA, whereas CST binds ss G-overhang and recruits Pol alpha-primase complex for C-strand telomere synthesis (see below). This model predicts that a mutation of Pol alpha in a CST mutant background will not further increase G-overhang signal. In contrast, a Pol alpha/KU double mutant would display more severe G-overhang defects.

Another fundamental question to ask is why CST and Shelterin are both required to maintain telomere integrity. The synergistic defects of depletion of human STN1 and POT1 indicate that CST and Shelterin protect telomeres independently. Is this true in *Arabidopsis*? To answer this question, *Arabidopsis stn1* mutants can be crossed with plants lacking TBP1 and POT1b. In the case of POT1c, knockdown of POT1c will be introduced in a *ctc1* or *stn1* mutant since there is no *pot1c* null mutant available. If *Arabidopsis* CST and Shelterin independently protect telomeres, it is expected that even more severe telomere defects will be observed in CST/Shelterin double mutants. As both CST and Shelterin are involved in G-overhang protection and telomere maintenance, one would expect some interactions between these two. Indeed, an interaction between CTC1 and POT1b/POT1c was detected by an *in vitro* co-IP experiment (Y.V. Surovtseva, X. Song, A.D.L. Nelson and D.E. Shippen, unpublished work). Additionally, human STN1 associates with TPP1 *in vivo* (Wan et al, 2009). In *Arabidopsis*, the TPP1 homolog has not been identified. Dr. Jung Ro Lee in the lab is taking a yeast-two hybrid assay to screen putative interaction partners of CTC1, STN1 and TEN1. We hope to identify AtTPP1 and other CST partners in this assay. I would

also recommend performing a pull-down of CST followed by mass spectrometry analysis. For this experiment, synchronization of the plant cells prior to pull-down may help to enrich transiently associated protein partners. In particular, we are interested in CST partners during S-phase, when there may be a conformational change that allows telomerase access. Further questions can be pursued as we dig deeper, such as whether the interaction between CST and POT1b/POT1c/TPP1 (or other factors) is required for CST's action at telomeres. To sum up, future efforts are needed to elucidate how CST protects chromosome ends in the presence of other telomere-associated proteins.

The role of CST in telomere replication

In budding yeast, Cdc13 recruits telomerase to telomeres through interaction with Est1 (Qi & Zakian, 2000). Cdc13 also interacts with the catalytic subunit of Pol alpha (Pol1), while Stn1 is associated with the regulatory subunit of Pol alpha (Pol12) (Grossi et al, 2004; Qi & Zakian, 2000). Such interactions indicate that ScCST coordinates the action of telomerase and Pol alpha at telomeres, and couples the G- and C-strand telomere replication. In multicellular organisms, the regulation of telomerase and other telomere replication machinery is not clear. Human TPP1/POT1 appears to stimulate telomerase processivity *in vitro* (Wang et al, 2007). In addition, recent findings revealed that human CTC1/STN1 corresponds to Pol alpha accessory factor (Casteel et al, 2009), which facilitates Pol alpha activity *in vitro* by allowing the enzyme to prime and extend DNA in a reiterative fashion (Goulian & Heard, 1990). Thus, CST may recruit Pol alpha for telomere replication in vertebrates.

In *Arabidopsis*, direct interaction between STN1 and Pol alpha has yet to be determined, however, preliminary data indicate that AtCTC1 contacts the catalytic subunit of Pol alpha *in vitro* (X. Song, K.A. Boltz and D.E. Shippen, unpublished work). It is proposed that CTC1 is involved in recruiting Pol alpha to replicate the C-strand telomeres (Fig 8-2). To further elucidate the genetic interaction between CST and Pol alpha in *Arabidopsis*, a cross of *stn1* to a Pol alpha point mutant was generated and will be analyzed for telomere length, G-overhang structure analyses (K.A. Boltz, X. Song and D.E. Shippen, unpublished work). It is predicted that double mutants of *stn1* and *pol alpha* would not elicit synergistic defects at telomeres. To follow up on these, an *in vitro* Polymerase alpha enzyme activity assay could be developed (Goulian et al, 1990), in which protein extracts from different CST mutants and wild type plants can be tested for Pol alpha activity *in vitro*. It is expected that the loss of CST may impede polymerase alpha function on some types of substrates, including repetive telomeric DNA, stalled replicaction forks and so on. Finally, efforts should be made to obtain a separation-of-function allele of CTC1 or Pol alpha, which has an impaired CTC1-Pol alpha binding site. Analysis of such mutant may help us to dissect the function of CTC1, and to determine whether the telomere capping function of CST is related to the recruitment of Pol alpha to chromosome ends.

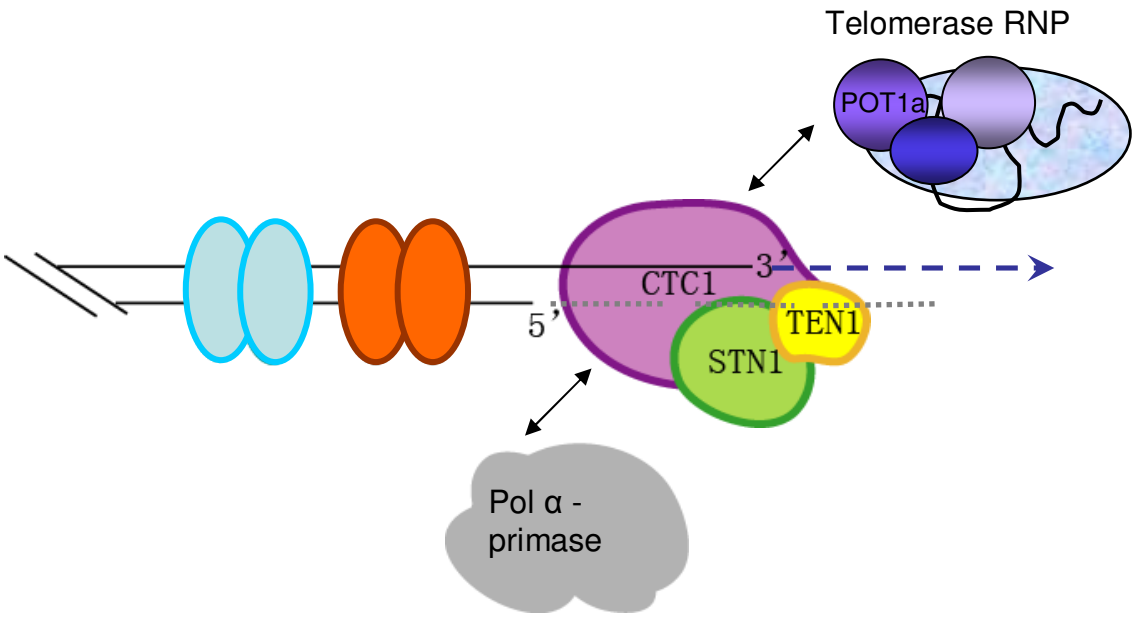


Fig 8-2. A proposed model of *Arabidopsis* CTC1 in telomere replication. CTC1 shows a direct interaction with the catalytic subunit of Pol alpha *in vitro*, suggesting CST may be involved in regulating the access of conventional replication machinery to telomere ends. In addition, CTC1 associates with the telomerase accessory subunit, POT1a *in vitro* and *in vivo*. It is hypothesized that CTC1 recruits both Pol alpha and telomerase and co-ordinates the G- and C-strand telomere synthesis in *Arabidopsis*.

Budding yeast Cdc13 also recruits telomerase to telomeres. This function is executed through an interaction between Cdc13 and Est1, a component of telomerase complex (Qi & Zakian, 2000). In Chapter VII, I show that CTC1 interacts with POT1a *in vitro* and *in vivo*, and that a tagged CTC1 protein is associated with active telomerase in plant extracts. These data, although preliminary, suggest that CTC1 is involved in telomerase recruitment to telomeres through POT1a interaction in *Arabidopsis* (Fig 8-2). As discussed below, *Arabidopsis* POT1a shares several characteristics with Est1. Future experiments should aim to confirm CTC1-POT1a-telomerase interaction (see below). Katie Leehy in the lab is working to identify new *ctc1* mutants in a TILLING EMS mutant collection. If our model is correct, a separation-of-function mutation of CTC1 that disrupts interaction with POT1a should result in an *est* defect. It is noticed though POT1a, STN1 and Pol alpha all interact with the C-terminal region of CTC1 (Fig 3-8, Fig 7-1 and data not shown). This result indicates these three proteins compete for a similar binding site on CTC1. If this is true, we may need to obtain separation-of-function alleles of POT1a, STN1 and Pol alpha to dissect CTC1 function.

A possible non-telomeric role for CST

Recent data from our lab revealed that the N-terminal OB-folds of AtCTC1 are sufficient for ss G-rich telomeric DNA interaction *in vitro* (J.R. Lee and D.E. Shippen, unpublished work). It remains unclear whether CTC1 can contact ss C-rich telomeric DNA, ss non-telomeric DNA, ds telomeric DNA or RNA molecules (e.g. TERRA RNA, or telomerase RNA). Once binding specificity is determined, further experiments should include, but not limit to determining the minimal binding site, the contribution of all nucleotides involved in the interaction, and the preference of CTC1 toward 3' or 5'

sequence. These biochemical experiments will help to reveal the function and dynamics of CTC1 at telomeres.

Notably, only about 20% of CTC1 or STN1 protein localizes to telomeres throughout the cell cycle in *Arabidopsis* and in humans (Miyake et al, 2009; Surovtseva et al, 2009). Where does the other 80% of CTC1/STN1 protein go? Does CST bind to non-telomeric sequence *in vivo*? If it does, what is the function of CST at these loci? The non-telomeric role of CST may be as important as its telomere function. This speculation is consistent with the observation that budding yeast Stn1 is enriched at non-telomeric sequences (Nugent, C.I., personal communication). One hypothesis is that CST assists the action of Polymerase alpha at genomic regions that are difficult to replicate (Wellinger, 2009). Such regions include telomeres, as well as satellite DNA, LINE-L1, LTR in human genome. The replication of these regions finishes at late S-phase or even G2-phase (Gilson & Geli, 2007; Hansen et al, 2010). One way to investigate non-telomeric substrates of CST is to perform a ChIPseq (ChIP-sequencing) assay (Johnson et al, 2007). In this experiment, protein-DNA complexes are crossed-linked prior to immunoprecipitation with an antibody against CTC1 (or STN1/TEN1), and the precipitated DNA will be purified and sequenced. If non-telomeric sequences are associated with CST, we should further determine whether there is a replication defect at these loci in CST deficient mutants.

ATR is involved in a checkpoint response in plants lacking CTC1 or STN1

To preserve genome integrity, eukaryotes have developed a sophisticated surveillance system to monitor and repair DNA damage. Damaged DNA activates the ATM and ATR kinases, leading to a signaling cascade that causes cell cycle arest until

the DNA damage is repaired. This ensures cells to repair DNA damage prior to entering mitosis. In particular, ATM responds to ds breaks (Garcia et al, 2003), whereas ATR is activated by replication stress or ss DNA damage (Culligan et al, 2004).

In Chapter V, we reported that plants doubly deficient of STN1 and ATR exhibited more anaphase bridges than *stn1* single mutants. It is conceivable that the abundant ss G-strand telomeric DNA in *stn1* mutants triggers ATR, which activates a cell cycle checkpoint response. Without ATR, the damaged cells will escape cell cycle arrest and end up with higher genome instability. To test this hypothesis, Kara Boltz in the lab is trying to determine if checkpoint-related genes are upregulated in *stn1* mutants. An alternative strategy is to examine the nuclear DNA content by flow cytometry. If our model is right, more G2 arrested cells will be found in *stn1* mutants than *stn1 atr* mutants. Another way to look at this result is that STN1 (or CST) in normal cells inhibits ATR activation at telomeres. How does this inhibition occur? It is of interest to investigate if CST physically interact with ATR or ATRIP, an ATR-associated protein. These studies will help us to better understand the role of CST and DNA damage response machinery at telomeres.

Taken together, CST complex in multicellular eukaryotes not only adds another layer to protect chromosome ends, but also provides a link between replication machinery and telomeres. Further exploration of CST functions would help to elucidate the mechanisms governing eukaryote genome stability, cell proliferation and cancer formation.

***Arabidopsis* POT1a is a novel telomerase regulator**

POT1 is a single copy gene that plays essential roles in chromosome end protection in yeast and humans (Baumann & Cech, 2001). In comparison, the flowering plant *Arabidopsis thaliana* encodes three highly divergent POT1 proteins, termed POT1a, POT1b and POT1c. While the exact role of POT1b and POT1c is still not clear, both proteins are implicated in G-overhang maintenance and chromosome end protection (Shakirov et al, 2005; A.D.L. Nelson and D.E. Shippen, unpublished work). POT1b is also involved in telomerase regulation possibly through its interaction with TER2. In contrast, AtPOT1a has evolved a novel function to bind TER1 (C. Cifuentes- and D.E. Shippen, unpublished work), and is required for telomerase enzyme activity *in vitro* and *in vivo* (Surovtseva et al, 2007).

To explore the nature and origin of *POT1* gene duplication in plants, *POT1* genes from 30 representatives of the plant kingdom were analyzed in Chapter VI. We found that *POT1* gene duplication is rare in plants. Only two examples of independent *POT1* gene duplication were uncovered: one occurred in the Panicoideae subfamily of grasses, and the other in the Brassicaceae family that includes *Arabidopsis*. Computer modeling of OB1 revealed no obvious structural variation between plant POT1 proteins and POT1 proteins from *S.pombe* and humans. However, several positively selected sites were detected in AtPOT1a. Two of these sites are located in the OB1, E35 and L132, which exhibit statistical significance for positive selection. Mutating of E35 and L132 back to ancestral amino acids dramatically reduced AtPOT1a function in a genetic complementation assay. Structural modeling suggests that none of these sites is likely to play an important role in direct nucleic acid binding. Preliminary results showed that mutation of E35 and L132 barely affect TER1 binding (data not shown). We hypothesize

that E35 and L132 is involved in protein-protein interactions and possibly telomerase complex assembly. Moreover, the single copy *POT1* genes from cotton or poplar failed to complement *Atpot1a* mutants. In addition, POT1a from *B. oleracae*, which is closely related to *Arabidopsis*, only partially complements AtPOT1a deficiciency. These data support a neo-functionalization model for *AtPOT1a* gene function, and reveal that Brassicaceae POT1a proteins are under strong selective pressure.

POT1a interacts with CTC1 *in vitro* and *in vivo* (Chapter VII). Therefore, I hypothesize that POT1 is a telomerase recruitment factor that links the telomerase enzyme to chromosome ends through interacting with CTC1 and TER1 (Fig 8-2). To further dissect the functional domains in POT1a, two different strategies were undertaken in Chapter VII. First, a collection of EMS-mutagenized *Arabidopsis* TIILING plants were screened to identify novel *pot1a* mutant alleles. Through this screen, I found that Asp 385, which is conserved in POT1 proteins, is critical for POT1a function *in vivo*. Second, we used site-directed mutagenesis and a complementation strategy to monitor POT1a activity *in vivo*. The data suggest that a Phe 65 in OB1, which corresponds to an amino acid that contacts telomeric DNA in human POT1 protein (Lei et al, 2004), is important for POT1a function. Moreover, the proper function of POT1a requires the last five residues at the extreme C-terminus, especially Thr 463. Preliminary data indicate that Asp 385 and Phe 65 contribute to the POT1a-TER1 interaction. Taken together, both the N-terminus and the C-terminus of POT1a are required for its function in telomerase association and telomere length regulation.

The hypothesis that POT1a is a functional homolog of budding yeast Est1

Arabidopsis POT1a shares many functional similarities with Est1 in budding yeast. Like Est1 (Seto et al, 2002), POT1a directly associates with telomerase RNA (C. Cifuentes and D.E. Shippen, unpublished work). In addition, loss of POT1a leads to an *est* phenotype similar to *est1* mutants in budding yeast. Moreover, similar to Est1, which associates with Cdc13 (Pennock et al, 2001), preliminary data indicate an interaction between POT1a and CTC1 *in vitro* and *in vivo* (Chapter VII). Finally, our collaborator Sue Armstrong found that POT1a is not stably expressed during the cell cycle (Armstrong, S., personal communication). Likewise, Est1 is degraded in G1 by proteasome (Osterhage et al, 2006). Taken together, it is conceivable that POT1a is a functional homolog of Est1. Two methods are proposed here to test the Est1-homolog hypothesis for AtPOT1a. First, we need to identify separation-of-function alleles of POT1a or CTC1. If the model is correct, mutations that disrupt CTC1-POT1a interaction should result in shortened telomere tracts. Second, it will be interesting to investigate whether the requirement of POT1a can be bypassed by fusing CTC1 directly with TERT. Similar experiments were performed in budding yeast (Pennock et al, 2001). Once the interaction of CTC1-POT1a is confirmed, it is interesting to follow whether the interaction is changed during cell-cycle. Such studies will help to reveal how telomere replication is coordinated with cell cycle progression in multicellular eukaryotes.

There are some differences between Est1 and POT1a functions. For example, Est1 exhibits a weak interaction with ss telomeric DNA (Virta-Pearlman et al, 1996), which has not been reported for POT1a. In addition, Est1 is dispensable for telomerase activity (Lingner et al, 1997a), while POT1a is required for full enzyme activity *in vitro* (Surovtseva et al, 2007). This result suggests that POT1a has an additional activity to

stabilize or stimulate telomerase. Catherine Cifuentes-Rojas in the lab is working on developing a direct telomerase assay using recombinant TERT and TER1. Once this assay is set up, it will be intriguing to test whether adding POT1a will boost the telomerase activity and processivity *in vitro*. Finally, it is important to note that the proposed model is not perfect. In Chapter V, we showed that the rare survivors of *ctc1 tert* mutants carry shorter telomere tracts than *ctc1* single mutants. Although this result does not disprove the role of CTC1 in telomerase regulation, it suggests that there is a CTC1-independent telomerase recruitment pathway in *Arabidopsis thaliana*.

What other proteins interact with POT1a?

In fission yeast and mammals, POT1 interacts with TPP1 to protect chromosome ends and to regulate telomere length. POT1 and TPP1 correspond to the alpha and beta subunit of TEBP protein in *Oxytrica*, respectively (Baumann & Cech, 2001; Xin et al, 2007). Although three POT1-like proteins are present in the *Arabidopsis* genome, the TPP1 homolog has not been discovered. Interestingly, one of the TPP1-interacting motifs of hPOT1 (Liu et al, 2004b) is conserved in the extreme C-terminus of AtPOT1a. Deletion of this motif completely abolished POT1a activity *in vivo* (Chapter VII). This result suggests that a TPP1 functional homolog may exist in the *Arabidopsis* genome, whose activity is required for complete POT1a function. A yeast-two hybrid assay could be used to identify TPP1 homolog and other POT1a partners. Once TPP1 (or other POT1a interaction partner) is identified, its role in telomere length regulation and chromosome end protection can be investigated along with its interaction with other known telomere-associated proteins.

The limitation of current studies on POT1a

With our genetic complementation strategy, we have a handful of *pot1a* mutants that show significantly reduced protein activity *in vivo*, including positive selection sites mutants E35F and L132D, the TILLING mutant D385N, point mutants of F65A and T463A, and the truncation mutants lacking either five or ten amino acids at the C-terminus. In addition, we have cloned several POT1a genes in the Brassicaceae family, including BoPOT1a and AlPOT1a. Although BoPOT1a shares ~ 70% protein sequence similarity to AtPOT1a, it barely complements AtPOT1a deficiency. It remains largely speculative about how these mutations or sequence differences (in the case of BoPOT1a) affect POT1a function in *Arabidopsis*. Preliminary results showed that D385N and F65A POT1a mutants disrupt the protein interaction with TER1. We should optimize this assay and test all POT1a mutants available. Furthermore, although TERT and TER are sufficient to reconstitute telomerase activity *in vitro* (C. Cifuentes-Rojas and D.E. Shippen, unpublished work), we should test whether the addition of wild type or mutant POT1a proteins affect telomerase activity. Finally, the interactions need to be examined between POT1a mutants and CTC1. Taken together, further studies on POT1a and POT1a-associated proteins will help to understand how telomerase is regulated in *Arabidopsis*.

Conclusions

In summary, I have used a combination of genetic, biochemical, molecular and evolutionary tools to study *Arabidopsis* CTC1/STN1/TEN1 and POT1a. My work has uncovered several new telomere-related proteins with essential functions in chromosome end protection and telomere replication. These studies not only improve our understanding on plant telomere composition, but also provide new insights into fundamental aspects of telomere protection and maintenance in multicellular eukaryotes.

REFERENCES

Ahmed S, Passos JF, Birket MJ, Beckmann T, Brings S, Peters H, Birch-Machin MA, von Zglinicki T, Saretzki G (2008) Telomerase does not counteract telomere shortening but protects mitochondrial function under oxidative stress. *J Cell Sci* **121:** 1046-1053

Anderson EM, Halsey WA, Wuttke DS (2003) Site-directed mutagenesis reveals the thermodynamic requirements for single-stranded DNA recognition by the telomere-binding protein Cdc13. *Biochemistry* **42:** 3751-3758

Armstrong L, Saretzki G, Peters H, Wappler I, Evans J, Hole N, von Zglinicki T, Lako M (2005) Overexpression of telomerase confers growth advantage, stress resistance, and enhanced differentiation of ESCs toward the hematopoietic lineage. *Stem Cells* **23:** 516-529

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

Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22:** 195-201

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

Baroudy BM, Venkatesan S, Moss B (1982) Incompletely base-paired flip-flop terminal loops link the two DNA strands of the vaccinia virus genome into one uninterrupted polynucleotide chain. *Cell* **28:** 315-324

Bateman A (1975) Letter: simplification of palindromic telomere theory. *Nature* **253:** 379-380

Baumann P, Cech TR (2000) Protection of telomeres by the Ku protein in fission yeast. *Mol Biol Cell* **11:** 3265-3275

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. *Mol Cell Biol* **22:** 8079-8087

Bechtold N, Pelletier G (1998) In planta Agrobacterium-mediated transformation of adult *Arabidopsis thaliana* plants by vacuum infiltration. *Methods Mol Biol* **82:** 259-266

Berman J, Tachibana CY, Tye BK (1986) Identification of a telomere-binding activity from yeast. *Proc Natl Acad Sci U S A* **83:** 3713-3717

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

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

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

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

Blasco MA, Lee H-W, Hande MP, Samper E, Lansdorp PM, DePinho RA, Greider CW (1997) Telomere shortening and tumor formation by mouse cells lacking telomerase RNA. *Cell* **91:** 25-34

Bonetti D, Martina M, Clerici M, Lucchini G, Longhese MP (2009) Multiple pathways regulate 3' overhang generation at *S. cerevisiae* telomeres. *Mol Cell* **35:** 70-81

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

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

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

Bryan TM, Sperger JM, Chapman KB, Cech TR (1998) Telomerase reverse transcriptase genes identified in *Tetrahymena thermophila* and *Oxytricha trifallax*. *Proc Natl Acad Sci U S A* **95:** 8479-84

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

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

Casteel DE, Zhuang S, Zheng Y, Perrino FW, Boss GR, Goulian M, Pilz RB (2009) A DNA polymerase-alpha/primase cofactor with homology to replication protein A-32 regulates DNA replication in mammalian cells. *J Biol Chem* **284:** 5807-5818

Cavalier-Smith T (1974) Palindromic base sequences and replication of eukaryuote chromosome ends. *Nature* **250:** 467-470

Cavatorta J, Savage A, Yeam I, Gray S, Jahn M (2008) Positive Darwinian selection at single amino acid sites conferring plant virus resistance. *J Mol Evol* **67**(5)**:** 551-559

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, Groff-Vindman C, Compton SA, McEachern MJ, Griffith JD (2008) Telomere loops and homologous recombination-dependent telomeric circles in a *Kluyveromyces lactis* telomere mutant strain. *Mol Cell Biol* **28:** 20-29

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

Cha RS, Kleckner N (2002) ATR homolog Mec1 promotes fork progression, thus averting breaks in replication slow zones. *Science* **297:** 602-606

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

Chakhparonian M, Wellinger RJ (2003) Telomere maintenance and DNA replication: how closely are these two connected? *Trends Genet* **19:** 439-446

Chan A, Boule J-B, Zakian VA (2008) Two pathways recruit telomerase to *Saccharomyces cerevisiae* telomeres. *PLoS Genet* **4:** e1000236

Chan SWL, Blackburn EH (2003) Telomerase and ATM/Tel1p protect telomeres from nonhomologous end joining. *Mol Cell* **11:** 1379-1387

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

Chang BSW, Jonsson K, Kazmi MA, Donoghue MJ, Sakmar TP (2002) Recreating a functional ancestral archosaur visual pigment. *Mol Biol Evol* **19:** 1483-1489

Chen J-L, Blasco MA, Greider CW (2000) Secondary structure of vertebrate telomerase RNA. *Cell* **100:** 503-514

Chikashige Y, Hiraoka Y (2001) Telomere binding of the Rap1 protein is required for meiosis in fission yeast. *Curr Biol* **11:** 1618-1623

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

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

Cohen SB, Graham ME, Lovrecz GO, Bache N, Robinson PJ, Reddel RR (2007) Protein composition of catalytically active human telomerase from immortal cells. *Science* **315:** 1850-1853

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

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

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

Croy JE, Podell ER, Wuttke DS (2006) A new model for *Schizosaccharomyces pombe* telomere recognition: the telomeric single-stranded DNA-binding activity of Pot11-389. *J Mol Biol* **361:** 80-93

Croy JE, Wuttke DS (2006) Themes in ssDNA recognition by telomere-end protection proteins. *Trends Biochem Sci* **31:** 516-525

Cui L, Wall PK, Leebens-Mack JH, Lindsay BG, Soltis DE, Doyle JJ, Soltis PS, Carlson JE, Arumuganathan K, Barakat A, Albert VA, Ma H, dePamphilis CW (2006) Widespread genome duplications throughout the history of flowering plants. *Genome Res* **16:** 738-749

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

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 (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

Dehal P, Boore JL (2005) Two rounds of whole genome duplication in the ancestral vertebrate. *PLoS Biol* **3:** e314

Delport W, Scheffler K, Seoighe C (2009) Models of coding sequence evolution. *Brief Bioinform* **10:** 97-109

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

Deng Z, Norseen J, Wiedmer A, Riethman H, Lieberman PM (2009) TERRA RNA binding to TRF2 facilitates heterochromatin formation and ORC recruitment at telomeres. *Mol Cell* **35:** 403-413

Donnelly PM, Bonetta D, Tsukaya H, Dengler RE, Dengler NG (1999) Cell cycling and cell enlargement in developing leaves of *Arabidopsis*. *Dev Biol* **215:** 407-419

Endo T, Ikeo K, Gojobori T (1996) Large-scale search for genes on which positive selection may operate. *Mol Biol Evol* **13:** 685-690

Enomoto S, Glowczewski L, Berman J (2002) MEC3, MEC1, and DDC2 are essential components of a telomere checkpoint pathway required for cell cycle arrest during senescence in *Saccharomyces cerevisiae*. *Mol Biol Cell* **13:** 2626-2638

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

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

Fajkus J, Sykorova E, Leitch AR (2005) Telomeres in evolution and evolution of telomeres. *Chromosome Res* **13:** 469-479

Feng J, Funk WD, Wang S-S, Weinrich SL, Avilion AA, Chiu C-P, Adams RR, Chang E, Allsopp RC, Yu J, Le S, West MD, Harley CB, Andrews WH, Greider CW, Villeponteau B (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

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

Fitzgerald MS, McKnight TD, Shippen DE (1996) Characterization and developmental patterns of telomerase expression in plants. *Proc Natl Acad Sci USA* **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 USA* **96:** 14813-14818

Fitzgerald MS, Shakirov EV, Hood EE, McKnight TD, Shippen DE (2001) Different modes of *de novo* telomere formation by plant telomerases. *Plant J* **26:** 77-87.

Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J (1999) Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* **151:** 1531-1545

Gallego ME, White CI (2001) RAD50 function is essential for telomere maintenance in *Arabidopsis*. *Proc Natl Acad Sci U S A* **98:** 1711-1716

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 V, Bruchet H, Camescasse D, Granier F, Bouchez D, Tissier A (2003) AtATM is essential for meiosis and the somatic response to DNA damage in plants. *Plant Cell* **15:** 119-132

Garg P, Burgers PMJ (2005) DNA polymerases that propagate the eukaryotic DNA replication fork. *Crit Rev Biochem Mol Biol* **40:** 115-128

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

Gilad Y, Segre D, Skorecki K, Nachman MW, Lancet D, Sharon D (2000) Dichotomy of single-nucleotide polymorphism haplotypes in olfactory receptor genes and pseudogenes. *Nat Genet* **26:** 221-224

Gilbert W, de Souza SJ, Long M (1997) Origin of genes. *Proc Natl Acad Sci U S A* **94:** 7698-7703

Gilson E, Geli V (2007) How telomeres are replicated. *Nat Rev Mol Cell Biol* **8:** 825-838

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

Goulian M, Heard C (1990) The mechanism of action of an accessory protein for DNA polymerase alpha/primase. *J Biol Chem* **265:** 13231-13239

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

Grant JD, Broccoli D, Muquit M, Manion FJ, Tisdall J, Ochs MF (2001) Telometric: a tool providing simplified, reproducible measurements of telomeric DNA from constant field agarose gels. *BioTechniques* **31:** 1314-1318

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

Greene EA, Codomo CA, Taylor NE, Henikoff JG, Till BJ, Reynolds SH, Enns LC, Burtner C, Johnson JE, Odden AR, Comai L, Henikoff S (2003) Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in *Arabidopsis*. *Genetics* **164:** 731-740

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

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

Grossi S, Puglisi A, Dmitriev PV, Lopes M, Shore D (2004) Pol12, the B subunit of DNA polymerase a, functions in both telomere capping and length regulation. *Genes Dev* **18:** 992-1006

Guex N, Peitsch M (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18:** 2714-2723

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* **52:** 696-704

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

Gutierrez C (2005) Coupling cell proliferation and developement in plants. *Nat Cell Biol* **7:** 535-541

Hansen RS, Thomas S, Sandstrom R, Canfield TK, Thurman RE, Weaver M, Dorschner MO, Gartler SM, Stamatoyannopoulos JA (2010) Sequencing newly replicated DNA reveals widespread plasticity in human replication timing. *Proc Natl Acad Sci U S A* **107:** 139-144

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 (2008) Telomerase and cancer therapeutics. *Nat Rev Cancer* **8:** 167-179

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

Harrington L, Zhou W, McPhail T, Oulton R, Yeung DS, Mar V, Bass MB, Robinson MO (1997) Human telomerase contains evolutionarily conserved catalytic and structural subunits. *Genes Dev* **11:** 3109-15

He H, Multani AS, Cosme-Blanco W, Tahara H, Sen Pathak JM, 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

Heiss NS, Knight SW, Vulliamy TJ, Klauck SM, Wiemann S, Mason PJ, Poustka A, Dokal I (1998) X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. *Nat Genet* **19:** 32-38

Heller RC, Marians KJ (2006) Replication fork reactivation downstream of a blocked nascent leading strand. *Nature* **439:** 557-562

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

Herbert BS, Hochreiter AE, Wright WE, Shay JW (2006) Nonradioactive detection of telomerase activity using the telomeric repeat amplification protocol. *Nat Protoc* **1:** 1583-1590

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

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

Hoffelder DR, Luo L, Burke NA, Watkins SC, Gollin SM, Saunders WS (2004) Resolution of anaphase bridges in cancer cells. *Chromosoma* **112:** 389-397

Hong JP, Byun MY, Koo DH, An K, Bang JW, Chung IK, An G, Kim WT (2007) Suppression of *RICE TELOMERE BINDING PROTEIN1* 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

Hughesa TR, Evans SK, Weilbaecherc RG, Lundblad V (2000) The Est3 protein is a subunit of yeast telomerase. *Curr Biol* **10:** 809-812

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

Innan H, Kondrashov F (2009) The evolution of gene duplications: classifying and distinguishing between models. *Nat Rev Genet* **11:** 97-108

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 (2007) *Tetrahymena* POT1a regulates telomere length and prevents activation of a cell cycle checkpoint. *Mol Cell Biol* **27:** 1592-1601

Johnson DS, Mortazavi A, Myers RM, Wold B (2007) Genome-wide mapping of *in vivo* protein-DNA interactions. *Science* **316:** 1497-1502

Kannan K, Nelson ADL, 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

Karimi M, De Meyer B, Hilson P (2005) Modular cloning in plant cells. *Trends Plant Sci* **10:** 103-105

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

Kato A, Lamb JC, Birchler JA (2004) Chromosome painting using repetitive DNA sequences as probes for somatic chromosome identification in maize. *Proc Natl Acad Sci USA* **101:** 13554-13559

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

Kellis M, Birren BW, Lander ES (2004) Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. *Nature* **428:** 617-624

Kellogg EA (1998) Relationships of cereal crops and other grasses. *Proc Natl Acad Sci USA* **95:** 2005-2010

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

Kim N, Piatyszek M, Prowse K, Harley C, West M, Ho P, Coviello G, Wright W, Weinrich S, Shay J (1994) Specific association of human telomerase activity with immortal cells and cancer. *Science* **266:** 2011-2015

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

Klobutcher LA, M T Swanton, P Donini, 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

Koch MA, Haubold B, Mitchell-Olds T (2000) Comparative Evolutionary Analysis of Chalcone Synthase and Alcohol Dehydrogenase Loci in *Arabidopsis*, *Arabis*, and Related Genera (Brassicaceae). *Mol Biol Evol* **17:** 1483-1498

Kondo Y, Kondo S, Tanaka Y, Haqqi T, Barna BP, Cowell JK (1998) Inhibition of telomerase increases the susceptibility of human malignant glioblastoma cells to cisplatin-induced apoptosis. *Oncogene* **16:** 2243-2248

Kuchar M, Fajkus J (2004) Interactions of putative telomere-binding proteins in *Arabidopsis thaliana*: identification of functional TRF2 homolog in plants. *FEBS Lett* **578:** 311-315

Lansdorp P, Verwoerd N, van de Rijke F, Dragowska V, Little M, Dirks R, Raap A, Tanke H (1996) Heterogeneity in telomere length of human chromosomes. *Hum Mol Genet* **5:** 685-691

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23:** 2947-2948

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

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

Lee H-W, Blasco MA, Gottlieb GJ, Horner JW, Greider CW, DePinho RA (1998) Essential role of mouse telomerase in highly proliferative organs. *Nature* **392:** 569-574

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

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

Leonardi J, Box JA, Bunch JT, Baumann P (2008) TER1, the RNA subunit of fission yeast telomerase. *Nat Struct Mol Biol* **15:** 26-33

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

Li S, Makovets S, Matsuguchi T, Blethrow JD, Shokat KM, Blackburn EH (2009) Cdk1-dependent phosphorylation of Cdc13 coordinates telomere elongation during cell-cycle progression. *Cell* **136:** 50-61

Lin J-J, Zakian VA (1996) The *Saccharomyces* CDC 13 protein is a single-strand TG_{1-3} telomeric DNA-binding protein *in vitro* that affects telomere behavior *in vivo*. *Proc Natl Acad Sci USA* **93:** 13760-13765

Linger BR, Price CM (2009) Conservation of telomere protein complexes: shuffling through evolution. *Crit Rev Biochem Mol Biol* **44:** 434-446

Lingner J, Cech TR (1996) Purification of telomerase from *Euplotes aediculatus*: requirement of a primer 3' overhang. *Proc Natl Acad Sci U S A* **93:** 10712–10717

Lingner J, Cech TR, Hughes TR, Lundblad V (1997a) Three ever shorter telomere (EST) genes are dispensable for *in vitro* yeast telomerase activity. *Proc Natl Acad Sci U S A* **94:** 11190-11195

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

Liu D, O'Connor MS, Qin J, Songyang Z (2004a) 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 (2004b) 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

Luke B, Lingner J (2009) TERRA: telomeric repeat-containing RNA. *EMBO J* **28:** 2503-2510

Luke B, Panza A, Redon S, Iglesias N, Li Z, Lingner J (2008) The Rat1p 52 to 32 exonuclease degrades telomeric repeat-containing RNA and promotes telomere elongation in *Saccharomyces cerevisiae*. **32:** 465-477

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 (2006) Budding yeast telomeres. In *Telomeres, Second Edition,* de Lange T, Lundblad V and Blackburn EH (ed) pp 345-386. Cold Spring Harbor, New York, USA, Cold Spring Harbor Laboratory Press

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

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

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

Lynch M, Conery JS (2000) The evolutionary fate and consequences of duplicate genes. *Science* **290:** 1151-1155

Maddison D, Maddison W (2005) MacClade: analysis of phylogeny and character evolution. Sunderland, Massachusetts, USA, Sinauer Associates

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

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

Mason JM, Biessmann H (1995) The unusual telomeres of *Drosophila*. *Trends in Genetics* **11:** 58-62

McClintock B (1931) Cytological observations of deficiencies involving known genes, translocations and an inversion in *Zea mays*. *Mo Agric exp Res Stn Res Bull* **163:** 4-30

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

McCormick-Graham M, Haynes WJ, Romero DP (1997) Variable telomeric repeat synthesis in *Paramecium tetraurelia* is consistent with misincorporation by telomerase. *EMBO J* **16:** 3233-3242

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

McGuffin LJ, Bryson K, Jones DT (2000) The PSIPRED protein structure prediction server. *Bioinformatics* **16:** 404-405

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

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

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

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

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

Miyoshi T, Kanoh J, Saito M, Ishikawa F (2008) Fission yeast Pot1-Tpp1 protects telomeres and regulates telomere length. *Science* **320:** 1341-1344

Mizuno H, Wu J, Katayose Y, Kanamori H, Sasaki T, Matsumoto T (2008) Chromosome-specific distribution of nucleotide substitutions in telomeric repeats of rice (*Oryza sativa* L.). *Mol Biol Evol* **25:** 62-68

Moore RC, Purugganan MD (2003) The early stages of duplicate gene evolution. *Proc Natl Acad Sci U S A* **100:** 15682-15687

Moyzis RK BJ, 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. *The Collecting Net* **8:** 182-195

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

Murti KG, Prescott DM (1999) Telomeres of polytene chromosomes in a ciliated protozoan terminate in duplex DNA loops. *Proc Natl Acad Sci U S A* **96:** 14436-14439

Nakamura TM, Morin GB, Chapman KB, Weinrich SL, Andrews WH, Lingner J, Harley CB, Cech TR (1997) Telomerase catalytic subunit homologs from fission yeast and human. *Science* **277:** 955-9

Nikitina T, Woodcock CL (2004) Closed chromatin loops at the ends of chromosomes. *J Cell Biol* **166:** 161-165

Nugent CI, Bosco G, Ross LO, Evans SK, Salinger AP, Moore JK, Haber JE, Lundblad V (1998) Telomere maintenance is dependent on activities required for end repair of double-strand breaks. *Curr Biol* **8:** 657-662

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

Nylander J (2004) MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

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

Onodera Y, Haag JR, Ream T, Nunes PC, Pontes O, Pikaard CS (2005) Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* **120:** 613-622

Osterhage JL, Talley JM, Friedman KL (2006) Proteasome-dependent degradation of Est1p regulates the cell cycle-restricted assembly of telomerase in *Saccharomyces cerevisiae*. *Nat Struct Mol Biol* **13:** 720-728

Palm W, de Lange T (2008) How shelterin protects mammalian telomeres. *Annu Rev Genet* **42:** 301-314

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

Paterson AH, Bowers JE, Burow MD, Draye X, Elsik CG, Jiang C-X, Katsar CS, Lan T-H, Lin Y-R, Ming R, Wright RJ (2000) Comparative genomics of plant chromosomes. *Plant Cell* **12:** 1523-1540

Paterson AH, Bowers JE, Chapman BA (2004) Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. *Proc Natl Acad Sci U S A* **101:** 9903-9908

Pei J, Grishin NV (2007) PROMALS: towards accurate multiple sequence alignments of distantly related proteins. *Bioinformatics* **23:** 802-808

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

Petracek ME, Konkel LM, Kable ML, Berman J (1994) A Chlamydomonas protein that binds single-stranded G-strand telomere DNA. *EMBO J* **13:** 3648-3658

Petracek ME, Lefebvre PA, Silflow CD, Berman J (1990) Chlamydomonas telomere sequences are A+T-rich but contain three consecutive G-C base pairs. *Proc Natl Acad Sci U S A* **87:** 8222-8226

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

Petreaca RC, Chiu HC, Nugent CI (2007) The role of Stn1p in *Saccharomyces cerevisiae* telomere capping can be separated from its interaction with Cdc13p. *Genetics* **177:** 1459-1474

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

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

Prince VE, Pickett FB (2002) Splitting pairs: the diverging fates of duplicated genes. *Nat Rev Genet* **3:** 827-837

Puglisi A, Bianchi A, Lemmens L, Damay P, Shore D (2008) Distinct roles for yeast Stn1 in telomere capping and telomerase inhibition. *EMBO J* **27:** 2328–2339

Pursell ZF, Isoz I, Lundstrom E-B, Johansson E, Kunkel TA (2007) Yeast DNA polymerase ε participates in leading-strand DNA replication. *Science* **317:** 127-130

Qi H, Zakian VA (2000) The *Saccharomyces* telomere-binding protein Cdc13p interacts with both the catalytic subunit of DNA polymerase α and the telomerase-associated Est1 protein. *Genes Dev* **14:** 1777-1788

Qian W, Wang J, Jin N-N, Fu X-H, Lin Y-C, Lin J-J, Zhou J-Q (2009) Ten1p promotes the telomeric DNA-binding activity of Cdc13p: implication for its function in telomere length regulation. *Cell Res* **19:** 849-863

Raices M, Verdun RE, Compton SA, Haggblom CI, Griffith JD, Dillin A, Karlseder J (2008) *C. elegans* telomeres contain G-strand and C-strand overhangs that are bound by distinct proteins. *Cell* **132:** 745-757

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

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 (2003) Ku is required for telomeric C-rich strand maintenance but not for end-to-end chromosome fusions in *Arabidopsis*. *Proc Natl Acad Sci USA* **100:** 611-615

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:** 2819-1826

Ritchie KB, Mallory JC, Petes TD (1999) Interactions of TLC1 (which encodes the RNA subunit of telomerase), TEL1, and MEC1 in regulating telomere length in the yeast *Saccharomyces cerevisiae*. *Mol Cell Biol* **19:** 6065-6075

Romero DP, Blackburn EH (1981) A conserved secondary structure for telomerase RNA. *Cell* **67:** 343-353

Rossignol P, Collier S, Bush M, Shaw P, Doonan JH (2007) *Arabidopsis* POT1A interacts with TERT-V(I8), an N-terminal splicing variant of telomerase. *J Cell Sci* **120:** 3678-3687

Rudenko G, Van der Ploeg LH (1989) Transcription of telomere repeats in protozoa. *EMBO J* **8:** 2633–2638.

Sarin KY, Cheung P, Gilison D, Lee E, Tennen RI, Wang E, Artandi MK, Oro AE, Artandi SE (2005) Conditional telomerase induction causes proliferation of hair follicle stem cells. *Nature* **436:** 1048-1052

Schoeftner S, Blasco MA (2008) Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. *Nat Cell Biol* **10:** 228-236

Schranz ME, Mitchell-Olds T (2006) Independent ancient polyploidy events in the sister families Brassicaceae and Cleomaceae. *Plant Cell* **18:** 1152-1165

Seto AG, Livengood AJ, Tzfati Y, Blackburn EH, Cech TR (2002) A bulged stem tethers Est1p to telomerase RNA in budding yeast. *Genes Dev* **16:** 2800-2812

Shakirov EV, McKnight TD, Shippen DE (2009) POT1-independent single-strand telomeric DNA binding activities in Brassicaceae. *Plant J* **58:** 1004-1015

Shakirov EV, Salzberg SL, Alam M, Shippen DE (2008) Analysis of *Carica papaya* telomeres and telomere-associated proteins: insights into the evolution of telomere maintenance in Brassicales. *Trop Plant Biol* **1:** 202-215

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

Sharma GG, Gupta A, Wang H, Scherthan H, Dhar S, Gandhi V, Iliakis G, Shay JW, Young CSH, Pandita TK (2003) hTERT associates with human telomeres and enhances genomic stability and DNA repair. *Oncogene* **22:** 131-146

Shippen DE (2006) Plant telomeres. In *Telomeres, Second Edition,* de Lange T, Lundblad V and Blackburn EH (ed) pp 525-550. Cold Spring Harbor, New York, USA, Cold Spring Harbor Laboratory Press

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

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

Simillion C, Vandepoele K, Van Montagu MCE, Zabeau M, Van de Peer Y (2002) The hidden duplication past of *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A* **99:** 13627-13632

Singer MS, Gottschling DE (1994) TLC1: template RNA component of *Saccharomyces cerevisiae* telomerase. *Science* **266:** 404-9

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

Solovei I, Gaginskaya ER, Macgregor HC (1994) The arrangement and transcription of telomere DNA sequences at the ends of lampbrush chromosomes of birds. *Chromosome Res* **2:** 460-470

Soltis DE, Albert VA, Leebens-Mack J, Bell CD, Paterson AH, Zheng C, Sankoff D, dePamphilis CW, Wall PK, Soltis PS (2009) Polyploidy and angiosperm diversification. *Am J Bot* **96:** 336-348

Song X, Leehy K, Warrington RT, Lamb JC, Surovtseva YV, Shippen DE (2008) STN1 protects chromosome ends in *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A* **105:** 19815-19820

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, Churikov D, Boltz KA, Song X, Lamb JC, Warrington R, Leehy K, Heacock M, Price CM, Shippen DE (2009) Conserved telomere maintenance component 1 interacts with STN1 and maintains chromosome ends in higher eukaryotes. *Mol Cell* **36:** 207-218

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

Swanson WJ, Nielsen R, Yang Q (2003) Pervasive adaptive evolution in mammalian fertilization proteins. *Mol Biol Evol* **20:** 18-20

Swofford DL, Sullivan J (2003) Plylogeny inference based on parsimony and other methods using PAUP. In *The phylogenetic handbook: a practical approach to DNA and protein phylogeny*, Salemi M and Vandamme A (ed) pp160-206. Cambridge, UK, Cambridge University Press

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

Sykorova E, Lim KY, Chase MW, Knapp S, Leitch IJ, Leitch AR, Fajkus J (2003) The absence of *Arabidopsis*-type telomeres in *Cestrum* and closely related genera *Vestia* and *Sessea* (Solanaceae): first evidence from eudicots. *Plant J* **34:** 283-291

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

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 rriggers hierarchical assembly of telomere replication proteins in *Saccharomyces cerevisiae*. *Mol Cell* **17:** 573-583

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

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

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, Schultz SC (2003) Nucleotide shuffling and ssDNA recognition in *Oxytricha nova* telomere end-binding protein complexes. *EMBO J* **22:** 4314-4324

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

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

Underwood DH, Carroll C, McEachern MJ (2004) Genetic dissection of the *Kluyveromyces lactis* telomere and evidence for telomere capping defects in TER1 mutants with long telomeres. *Eukaryotic Cell* **3:** 369-384

Van de Peer Y, Maere S, Meyer A (2009) The evolutionary significance of ancient genome duplications. *Nat Rev Genet* **10:** 725-732

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

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, 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

Vespa L, Warrington RT, Mokros P, Siroky J, Shippen DE (2007) ATM regulates the length of individual telomere tracts in *Arabidopsis*. *Proc Natl Acad Sci U S A* **104:** 18145-18150

Virta-Pearlman V, Morris DK, Lundblad V (1996) Est1 has the properties of a single-stranded telomere end-binding protein. *Genes Dev* **10:** 3094-3104

Voet D, Voet JG, Pratt C (2006) *Fundamentals of Biochemistry: Life at the Molecular Level, Second Edition.* Hoboken, New Jersey, USA, John Wiley & Sons Inc,

Walsh JB (1995) How often do duplicated genes evolve new functions? *Genetics* **139:** 421-428

Wan M, Qin J, Songyang Z, Liu D (2009) OB-fold containing protein 1 (OBFC1), a human homologue of yeast Stn1, associates with TPP1 and is implicated in telomere length regulation. *J Biol Chem* **284:** 26725-26731

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

Warren EM, Vaithiyalingam S, Haworth J, Greer B, Bielinsky AK, Chazin WJ, Eichman BF (2008) Structural basis for DNA binding by replication initiator Mcm10. *Structure* **16:** 1892-1901

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

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

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

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

Weinert TA, Hartwell LH (1993) Cell cycle arrest of *cdc* mutants and specificity of the RAD9 checkpoint. *Genetics* **134:** 63-80

Wellinger RJ (2009) The CST complex and telomere maintenance: the exception becomes the rule. *Mol Cell* **36:** 168-169

Wernersson R, Pedersen AG (2003) RevTrans: multiple alignment of coding DNA from aligned amino acid sequences. *Nucl Acids Res* **31:** 3537-3539

Wikstrom N, Savolainen V, Chase MW (2001) Evolution of the angiosperms: calibrating the family tree. *Proc R Soc Lond B Biol Sci* **268:** 2211-2220

Witkin KL, Collins K (2004) Holoenzyme proteins required for the physiological assembly and activity of telomerase. *Genes Dev* **18:** 1107-1118

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 W, Piatyszek M, Rainey W, Byrd W, Shay J (1996) Telomerase activity in human germline and embryonic tissues and cells. *Genes Dev* **18:** 173-179

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

Xin H, Liu D, Songyang Z (2008) The telosome/shelterin complex and its functions. *Genome biol* **9:** 232

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-β and interacts with POT1 to recruit telomerase. *Nature* **445:** 559-562

Xu L, Petreaca RC, Gasparyan HJ, Vu S, Nugent CI (2009) TEN1 is essential for CDC13-mediated telomere capping. *Genetics* **183:** 793-810

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

Yang Z (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci* **13:** 555-556

Yao NY, O'Donnell M (2009) Replisome structure and conformational dynamics underlie fork progression past obstacles. *Curr Opin Cell Biol* **21:** 336-43

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-S, 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

Yoon HS, Hackett JD, Ciniglia C, Pinto G, Bhattacharya D (2004) A molecular timeline for the origin of photosynthetic eukaryotes. *Mol Biol Evol* **21:** 809-818

Yu EY, Wang F, Lei M, Lue NF (2008) A proposed OB-fold with a protein-interaction surface in *Candida albicans* telomerase protein Est3. *Nat Struct Mol Biol* **15:** 985-989.

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

Zhang J, Nielsen R, Yang Z (2005) Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol Biol Evol* **22:** 2472-2479

Zhu H, Fu W, Mattson MP (2000) The catalytic subunit of telomerase protects neurons against amyloid beta-peptide-induced apoptosis. *J Neurochem* **75:** 117-124

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 X-D, Niedernhofer L, Kuster B, Mann M, Hoeijmakers JHJ, 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

Zijlmans JMJM, Martens UM, Poon SSS, Raap AK, Tanke HJ, Ward RK, Lansdorp PM (1997) Telomeres in the mouse have large inter-chromosomal variations in the number of T_2AG_3 repeats. *Proc Natl Acad Sci U S A* **94:** 7423-7428

APPENDIX

POT1 PROTEINS IN GREEN ALGAE AND LAND PLANTS: DNA BINDING PROPERTIES AND EVIDENCE OF CO-EVOLUTION WITH TELOMERIC DNA*

Summary

Telomeric DNA terminates with a single-stranded 3' G-overhang that in vertebrates and fission yeast is bound by POT1 (Protection Of Telomeres). However, no *in vitro* telomeric DNA binding is associated with *Arabidopsis* POT1 paralogs. To further investigate POT1-DNA interaction in plants, we cloned *POT1* genes from 11 plant species representing major branches of plant kingdom. Telomeric DNA binding was associated with POT1 proteins from the green alga *Ostreococcus lucimarinus* and two flowering plants, maize and asparagus. Site-directed mutagenesis revealed several residues critical for telomeric DNA recognition in vertebrates are functionally conserved in plant POT1 proteins. However, the plant proteins varied in their minimal DNA binding sites and nucleotide recognition properties. Green alga POT1 exhibited a strong preference for the canonical plant telomere repeat sequence TTTAGGG with no detectable binding to hexanucleotide telomere repeat TTAGGG found in vertebrates and some plants, including Asparagus. In contrast, POT1 proteins from maize and Asparagus bound TTAGGG repeats with only slightly reduced affinity relative to the TTTAGGG sequence. We conclude that the nucleic acid binding site in plant POT1 proteins is evolving rapidly, and that the recent acquisition of TTAGGG telomere

* Reprint with permission from "POT1 proteins in green algae and land plants: DNA-binding properties and evidence of co-evolution with telomeric DNA" E. V. Shakirov, X. Song, J. A. Joseph and D. E. Shippen 2009..*Nucleic Acids Res.* 37(22): 7455–7467.

repeats in Asparagus appears to have co-evolved with changes in POT1 DNA sequence recognition.

Introduction

Telomeres are the ancient nucleoprotein structures that define the physical ends of eukaryotic chromosomes, protecting them from deleterious activities such as recombination and nucleolytic attack, and providing a means to replenish telomeric DNA lost during replication (Gilson & Geli, 2007). Defects in telomere structure or length maintenance result in cell proliferation and genome maintenance abnormalities, senescence or apoptosis (Xin et al, 2008). Telomere structure and composition is conserved across different eukaryotic lineages. Telomeric DNA typically consists of tandem arrays of short G-rich repeats that can reach thousands of nucleotides in length. The extreme 3'-ends of the chromosomes terminate in a single-stranded protrusion termed the G-overhang. Several evolutionarily conserved proteins bind directly to the double-stranded region of the telomeric DNA or to the single-strand G-overhang to form the first layer of telomere-associated protein factors. Together with bridging proteins, these DNA binding factors constitute a telomere-specific protein complex termed shelterin (Palm & de Lange, 2008; Xin et al, 2008).

The green plant lineage represents a monophyletic group of photosynthetic organisms that evolved near the base of eukaryotic life and shared the last common ancestor with fungi and animals approximately 1.5 billion years ago (bya) (Yoon et al, 2004). Despite such long divergence time, many aspects of telomere biology are well conserved between plants and animals. The telomere repeat sequence in the vast majority of plants is TTTAGGG, one nucleotide longer than the six-base sequence

TTAGGG found in vertebrates (McKnight & Shippen, 2004). Interestingly, a few outliers exist in the plant kingdom that harbor atypical or unknown telomere sequence. For example, while many green algae exhibit the canonical plant TTTAGGG repeats (Higashiyama et al, 1995)(E. Shakirov and D. Shippen, unpublished), telomeres in the model fresh water alga *Chlamydomonas reinhardtii* are composed of the eight nucleotide repeat TTTTAGGG (Petracek et al, 1990). Onions (*Allium cepa*) and related species lack simple telomere repeats and instead appear to harbor terminally located satellite DNA (Pich et al, 1996). Most intriguing is the situation in two phylogenetically unrelated groups of flowering plants (angiosperms) in which canonical TTTAGGG repeats have been replaced by the vertebrate-type hexanucleotide TTAGGG sequence. One of these plant groups is comprised of a handful of obscure Solanaceae (tomato family) species native to South America (Sykorova et al, 2003). The second group consists of a large number of families in the Asparagales order, and represents one of the most successful lineages of extant flowering plants, with 22,000-25,000 currently known species, or nearly 10% of all angiosperms, including Irises, Hyacinths, Agaves and Amaryllis (Fajkus et al, 2005). The switch in telomeric DNA sequence from TTTAGGG to TTAGGG is thought to have occurred approximately 90 million years ago (mya) (Wikstrom et al, 2001) and likely corresponds to a single nucleotide deletion in the template region of the telomerase RNA subunit. This sequence change may have posed a challenge for the plant shelterin complex to maintain chromosome end protection, although the successful diversification of Asparagales argues that these plants accommodated the mutation in telomeric DNA sequence in a short evolutionary time frame. Such compensatory changes likely involved co-evolution of telomere

binding proteins to allow binding to the new telomere repeat sequence, but the molecular mechanisms of these changes are currently unknown.

In vertebrates and fission yeast, POT1 (Protection Of Telomeres) binds single-strand G-rich telomeric DNA with high affinity and plays a pivotal role in mediating telomere length regulation as well as chromosome end protection and cell viability (Baumann & Cech, 2001; Palm & de Lange, 2008; Xin et al, 2008). POT1 proteins are defined by the presence of two N-terminal oligosaccharide/oligonucleotide binding folds (OB-folds), which are responsible for the specific interaction with single-stranded telomeric DNA. OB1 and OB2 contacts with telomeric DNA are primarily mediated by aromatic amino acids, which form stacking interactions with DNA nucleotides (reviewed in (Croy & Wuttke, 2006). Most POT1 proteins studied to date exhibit a minimum binding site (MBS) of 10-12 nucleotides, roughly corresponding to two telomeric repeats, though the most preferred repeat permutation in each case appears to be species-specific (Croy et al, 2006; Lei et al, 2004; Wei & Price, 2004). Interestingly, only a subset of MBS nucleotides is specifically recognized by POT1 proteins, with nucleotides crucial for protein binding scattered throughout the MBS. For many POT1 proteins, the most 3'-terminal MBS nucleotide is buried deep inside the OB-folds (Horvath et al, 1998; Lei et al, 2004), suggesting a mechanism for how POT1 can protect the G-overhang from nucleases or telomerase action. In addition, while some POT1 proteins clearly prefer 3' terminal telomeric repeats (Baumann & Cech, 2001) , others can associate with telomeric sequences in the middle or on the 5'-terminus of oligonucleotide substrate (Loayza et al, 2004; Wei & Price, 2004), suggesting that *in vivo* they can also bind to the displaced G-rich strand in the context of the T-loop. While specific interaction with telomeric DNA *in vivo* is an essential feature of all POT1 proteins studied to date,

human POT1 is delivered to the telomere via protein-protein interactions with its binding partner, TPP1/ACD. The interaction with TPP1 is achieved through a structurally undefined C-terminal domain (Liu et al, 2004b; Ye et al, 2004b). The POT1-TPP1 heterodimer is required for proper shelterin assembly and to regulate telomerase access and processivity (Kelleher et al, 2005; Liu et al, 2004b; Loayza & de Lange, 2003; Wang et al, 2007).

POT1-like proteins have also been identified in plants (Baumann et al, 2002; Kuchar & Fajkus, 2004; Shakirov et al, 2009; Shakirov et al, 2008; Tani & Murata, 2005). Similar to the situation in most other eukaryotes, only a single *POT1* gene has been detected in most plants surveyed (Shakirov et al., in preparation). However, *Arabidopsis* is an exception as it encodes three highly divergent POT1-like proteins (Shakirov et al, 2005; Nelson et al., in preparation). All three *Arabidopsis* POT1 proteins are involved in telomere biology, but their functions differ. AtPOT1a is a positive regulator of telomerase activity that physically interacts with the telomerase RNP (Rossignol et al, 2007; Surovtseva et al, 2007), while AtPOT1b and AtPOT1c negatively regulate telomerase activity and participate in chromosome end protection (Shakirov et al, 2005; Nelson et al., in preparation). Strikingly, although AtPOT1 proteins have an architecture similar to yeast and vertebrate POT1 with two N-terminal OB-folds and a C-terminal domain, no *in vitro* telomeric DNA binding has been demonstrated for POT1 proteins from *Arabidopsis* or two other closely related plants (Shakirov et al, 2009; Surovtseva et al, 2007). Thus, it is unclear whether telomeric DNA binding is a conserved function of POT1 proteins from the plant kingdom.

Plant systematics has witnessed a remarkable influx of new data revealing evolutionary relationship of various lineages of the green plants. We took advantage of

this detailed phylogenetic map to clone *POT1* genes from eleven representative members of major plant evolutionary branches. Here we report the initial characterization of the DNA binding activities of three plant POT1 proteins and provide evidence for significant biochemical differences in POT1 proteins across the plant kingdom. We also demonstrate that POT1 proteins from angiosperms have strong affinity for both TTAGGG and TTTAGGG telomeric repeats, providing a possible explanation for how Asparagales adapted to the recent switch in its telomeric DNA repeat sequence.

Materials and methods

Asparagus telomere length analysis

DNA from Asparagus (*Asparagus officinalis*) shoots was extracted as described (Cocciolone & Cone, 1993). Terminal restriction fragment (TRF) analysis was performed as described (Fitzgerald et al, 1999) with DNA digested with either *Tru1*I or *Alu*I restriction enzymes (Fermentas, Hanover, MD). ^{32}P 5' end–labeled $(T_3AG_3)_4$ and $(T_2AG_3)_4$TTAG oligonucleotides were used as heptanucleotide and hexanucleotide probes, respectively. Radioactive signals were scanned by a Pharos FX Plus Molecular Imager (Bio-Rad laboratories, Hercules, CA), and the data were analyzed by Quantity One v.4.6.5 software (Bio-Rad).

In vitro translation and EMSA assays

Expression of plant POT1 proteins in rabbit reticulocyte lysate (RRL) was performed as described (Baumann et al, 2002). EMSA assays were conducted as described (Shakirov et al, 2009) with slight modifications. Briefly, each reaction (15 µl

total volume) contained equal amounts (4 µl) of RRL-translated plant POT1 protein, 0.5 pmole of the corresponding ^{32}P-labeled telomeric oligonucleotide, 3 µl of 5X DNA binding buffer (100 mM Tris-HCl pH 8.0, 250 mM NaCl, 50 mM MgCl, 5mM EDTA, 5 mM DTT, 25% glycerol), and 1 µl each of non-specific RNA, single-stranded and double-stranded DNA competitors as described in (Baumann et al, 2002). Reactions were incubated at RT for 15 min. For competition assays, 2.5 pmole of cold competitor oligonucleotide was added and the reactions were incubated for an additional 15 min. The complexes were separated on 5% polyacrylamide gel (acrylamide: bisacrylamide 29:1) for 2 h at 150 volts in 0.8X TBE at RT, dried and exposed to PhosphorImager screens. Screens were scanned by a Pharos FX Plus Molecular Imager and signal intensity was quantified by Quantity One v.4.6.5 software. Each EMSA result was reproduced several times, but due to variations in protein expression levels in RRL, only one representative gel and the corresponding quantification result are shown for each experiment.

Site-directed mutagenesis

Mutagenesis reactions were performed with *Pfu* Turbo DNA polymerase (Stratagene) according to the manufacturer's instructions using the following PCR conditions: 94°C 5 min; 18 cycles of 94°C 30 sec, 55°C 1 min, 68°C 40 min; followed by 10 min at 68°C. After *Dpn*I (Promega) treatment at 37°C for 4 hr, the reaction product was transformed into TOP10F' competent cells (Invitrogen). Plasmids were purified and mutations verified by sequencing.

Nucleotide sequence accession numbers

cDNAs encoding the following plant POT1 proteins were deposited into the GenBank: HvPOT1 (EU880295), PtrPOT1 (EU880297), HaPOT1 (EU880298), SmPOT1 (EU880301), ZmPOT1a (EU880303), ZmPOT1b (EU880304), GhPOT1 (EU880305), PtaPOT1 (EU880306), StPOT1 (EU883536), AoPOT1 (FJ516399). The nucleotide sequence of cloned OlPOT1 cDNA was found to correspond to the previously deposited sequence ABO96101.

Results

Telomeric DNA binding by plant POT1 proteins.

As part of a larger study on the molecular evolution and functional divergence of POT1 proteins in plants, we cloned *POT1* cDNAs from eleven plant species representing major branches on the plant evolutionary tree. These include POT1 sequences from the green alga *Ostreococcus lucimarinus,* a spikemoss (*Selaginella moellendorffii*), a pine, three monocotyledonous (barley, maize and Asparagus) and four dicotyledonous (potato, sunflower, poplar and cotton) flowering plants (Fig A-1).

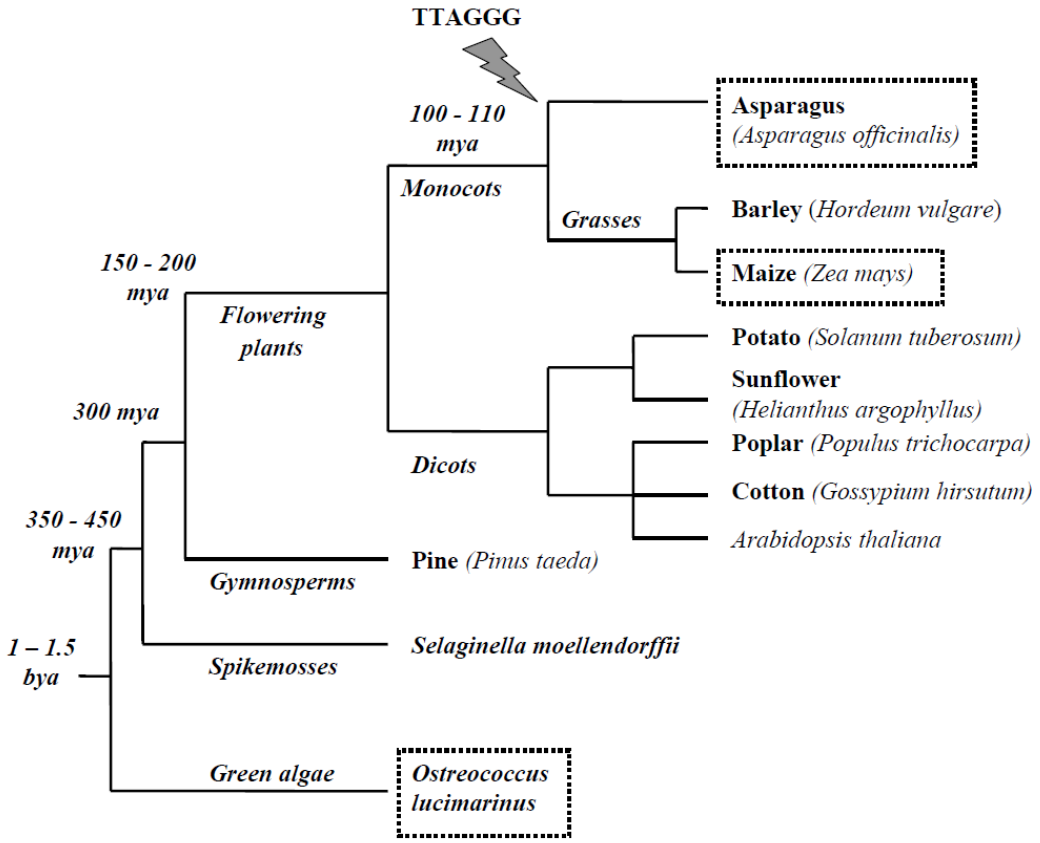


Fig A-1. A schematic evolutionary tree of plant species with POT1 proteins analyzed in this study. The common and binomial names of species used in this study are indicated. Major plant taxons are italicized and shown below each node, and the approximate time of their first appearance is indicated. Branch lengths do not represent degree of divergence. The lightning bolt points to the Asparagales lineage which experienced a mutation that changed its telomere repeat to the hexanucleotide TTAGGG sequence. Boxed species indicate plants with POT1 proteins that exhibit telomeric DNA binding in vitro. *Arabidopsis thaliana* included for comparative reasons. Divergence times are inferred from (Wikstrom et al, 2001).

Among the plant species analyzed in our study, three genomes have been sequenced: *Ostreococcus lucimarinus, Selaginella moellendorffii* and *Populus trichocarpa* (poplar). Ostreococcus *POT1* is a single-exon gene. *POT1* genes in both Selaginella and poplar, as well as the two previously characterized *Arabidopsis thaliana POT1* genes, harbor 10 exons with conserved intron positions (data not shown). This evolutionarily conserved gene structure supports the conclusion that the plant *POT1* genes are indeed orthologs. Among all plant species analyzed in this study, only maize appears to encode more than one POT1 protein. The two maize *POT1* genes were most likely retained after whole-genome duplication in the ancestor of maize (Paterson et al, 2004). *ZmPOT1a* and *ZmPOT1b* encode 54.5 and 54.6 kDa proteins, respectively, with 75% overall amino acid similarity to each other. Like the POT1 proteins from the Brassicaceae plants (Shakirov et al, 2009), all eleven of the new POT1 proteins display significant sequence conservation and are predicted to harbor two N-terminal DNA binding OB-folds with secondary structures similar to the human and fission yeast POT1 proteins (J. Croy and D. Wuttke, personal communication).

To assess the DNA binding properties of the new POT1 proteins, we attempted to obtain soluble recombinant proteins using standard expression protocols in *E. coli*. However, as was previously shown for POT1 proteins from *Arabidopsis* and related species (Shakirov et al, 2009; Surovtseva et al, 2007), we were unable to generate a sufficient amount of soluble protein from any clone for DNA binding studies. Therefore, we turned to an *in vitro* rabbit reticulocyte lysate (RRL) expression system, previously shown to produce soluble vertebrate and plant POT1 proteins (Baumann et al, 2002; He et al, 2006; Shakirov et al, 2009). While such approach will not allow us to define DNA binding constants, we could perform qualitative binding experiments that in previous studies with the mammalian POT1 proteins have provided important comparative insights into POT1 interaction with telomeric DNA (Baumann et al, 2002; He et al, 2006). RRL-expressed plant POT proteins were obtained in a soluble form (Fig A-2) and were subjected to electrophoretic mobility shift assays (EMSA) (Fig A-3).

For the yeast and vertebrate POT1 proteins, two telomeric repeats are sufficient for *in vitro* binding (Croy & Wuttke, 2006). Therefore, we used a cocktail of ^{32}P 5'-labeled oligonucleotides corresponding to two repeats of the seven possible permutations of the plant telomere repeat (2PLANT cocktail probe) for EMSA. Under standard gel-shift conditions (He et al, 2006), stable telomeric DNA binding was observed for two full-length plant POT1 proteins: OlPOT1 from the green alga *Ostreococcus lucimarinus* (Fig A-3, lane 2) and AoPOT1 from *Asparagus officinalis* (garden asparagus) (see below). In addition, a band with intensity slightly above background was observed in the well for the maize (*Zea mays*) ZmPOT1b protein (Fig A-3, lane 7). The intensity of this band increased when a specific oligonucleotide was used instead of the oligonucleotide cocktail (data not shown).

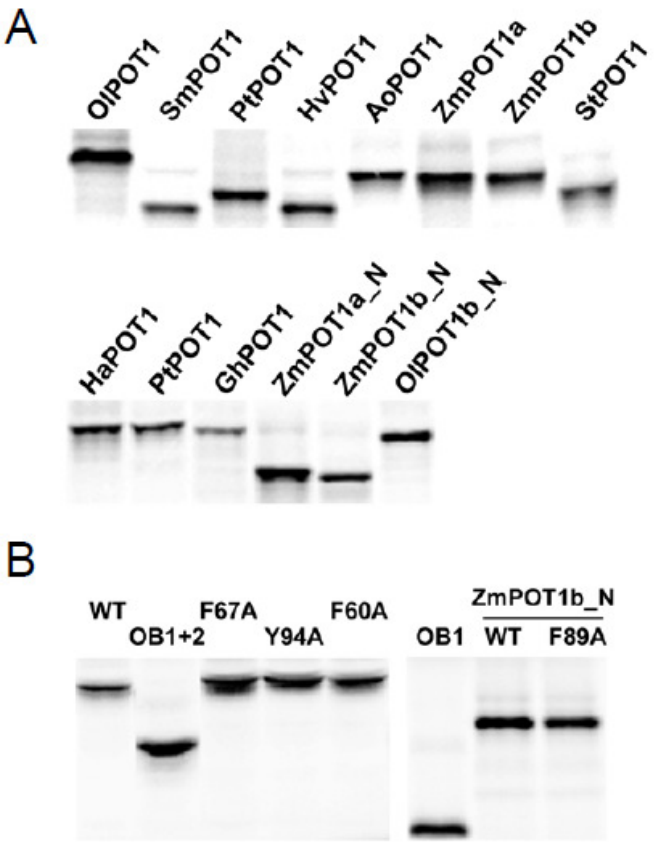


Fig A-2. RRL-expressed plant POT1 proteins are soluble. Reactions containing *in vitro* translated plant POT1 proteins (^{35}S) were spun for 10 min at 14,000 g and the soluble fractions were analyzed by SDS-PAGE. (**A**) RRL-expressed full-length and OB-fold truncated plant POT1 proteins. (**B**) RRL-expressed AoPOT1 and ZmPOT1b_N truncation and point mutants.

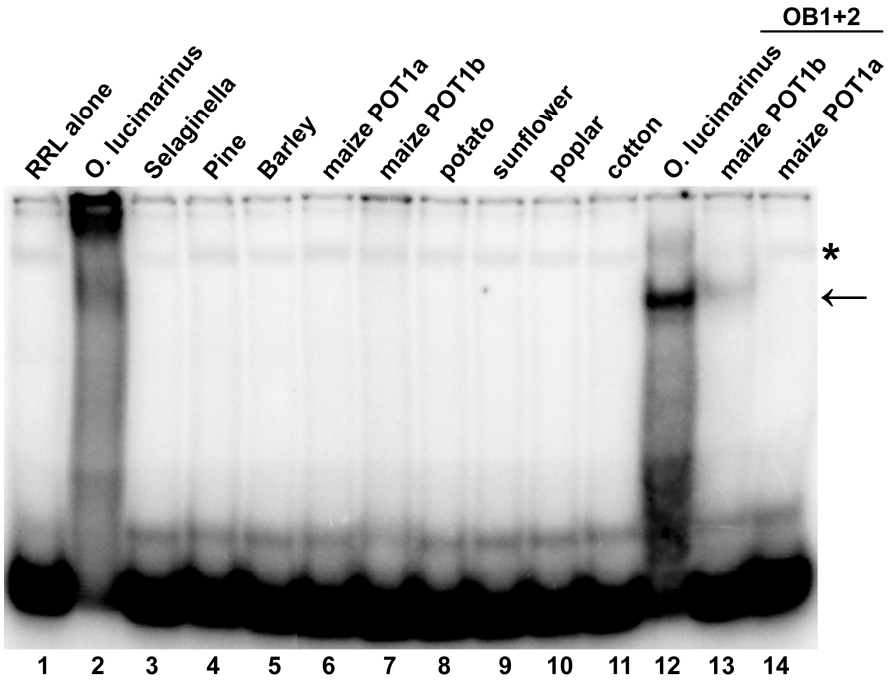


Fig A-3. Analysis of DNA binding capacity of recombinant plant POT1 proteins. EMSA was performed with a cocktail of seven ^{32}P-labeled 2PLANT oligonucleotides. POT1 proteins from the corresponding plant species are shown above each lane. Binding assays were performed with either full-length POT1 (lanes 2-11) or with truncated proteins bearing only two N-terminal OB-fold domains (lanes 12-14). Asterisk designates a non-specific band often present in the negative RRL-only control (lane 1). Protein-DNA complexes specific to maize and *O. lucimarinus* POT1 proteins are indicated by an arrow.

Telomeric DNA binding by mammalian and yeast POT1 proteins is enhanced with constructs containing only the N-terminal OB-folds (Baumann & Cech, 2001; He et al, 2006; Wu et al 2006), possibly due to modulation of DNA binding by the protein C-terminus. The OB-fold domains of all eleven plant POT1 proteins were expressed in RRL (Fig A-2 and data not shown) and tested in EMSA. Deletion of the C-terminus improved the binding of OlPOT1_N (amino acids 1-363) to the plant telomeric DNA cocktail, resulting in the formation of a single well-defined protein-DNA complex (Fig A-3, lane 12). In addition, a truncated version of maize POT1b, ZmPOT1b_N (amino acids 1-326) also formed single weak protein-DNA complex (Fig A-3, lane 13). However, the second maize POT1 protein, ZmPOT1a, failed to bind telomeric DNA as either a full-length protein or an OB-fold truncation (Fig A-3, lanes 6 and 14). With the exception of AoPOT1_N (see below), we detected no telomeric DNA binding by any other OB-fold truncated plant POT1 protein (data not shown). We also asked whether plant POT1 proteins could bind an oligonucleotide corresponding to the C-rich strand of telomeric DNA, as is the case for one of the POT1-like proteins from *C. elegans* (Raices et al, 2008). No C-strand binding was observed for any of the plant POT1 proteins (data not shown). Nevertheless, our data imply that the ability of POT1 to bind G-rich telomeric DNA has not been completely lost throughout plant kingdom.

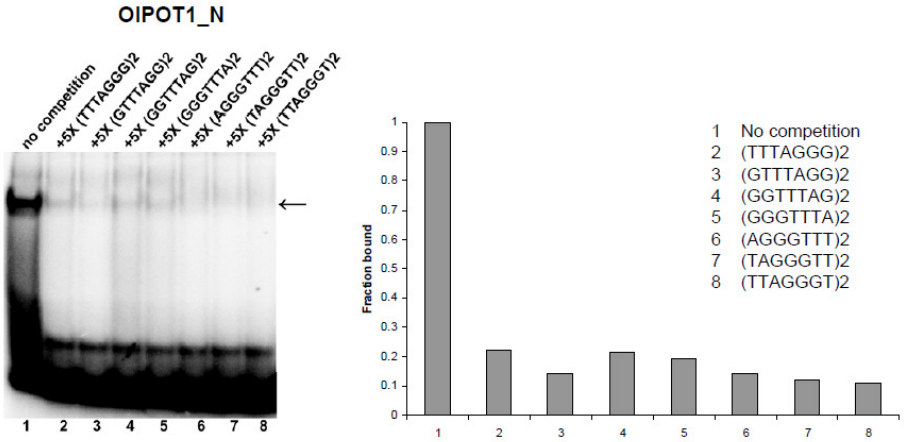
DNA binding properties of OlPOT1_N from green alga

Ostreococcus lucimarinus is a species of Prasinophytes, a clade of green algae that belongs to the oldest diverging (over 1 bya) branch of the photosynthetic eukaryotic lineage, and is a sister clade to all land plants (Yoon et al, 2004). Consequently, analysis of the DNA binding characteristics of OlPOT1_N may provide insight into the mechanisms of telomeric DNA recognition by the ancestral plant POT1 protein and how these properties have evolved in land plants.

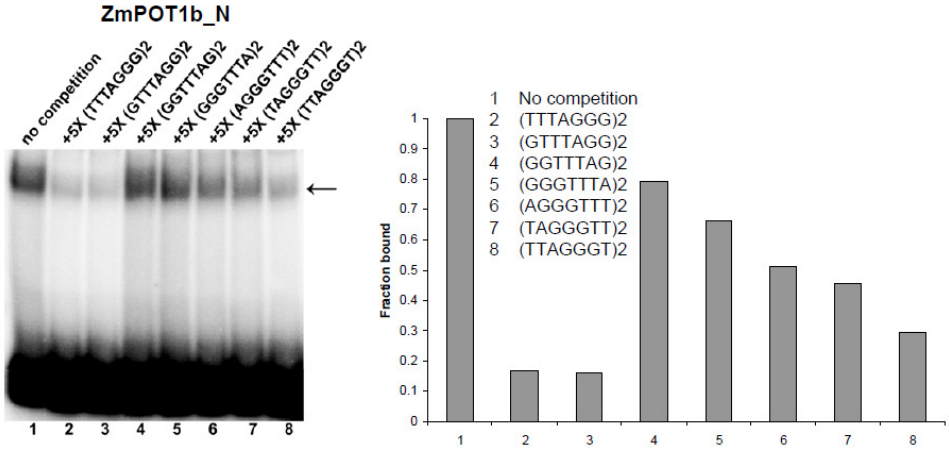
To determine which of the seven permutations of the plant telomeric repeat OlPOT1_N binds best, we performed competition experiments with ^{32}P-labeled 2PLANT cocktail probe and a 5-fold excess of individual cold 2PLANT oligonucleotides. A representative gel and corresponding quantification are shown in Fig A-4A. The intensity of the shifted band (fraction bound) in the absence of competitors was measured and set as 1.0, and the remaining signal intensity after the addition of competitors was expressed as a fraction of 1. We found that all individual 2PLANT oligonucleotides competed efficiently for binding (Fig A-4A).

Fig A-4. Identification of the preferred telomere repeat permutation for plant POT1 proteins. Competition assays for OlPOT1_N (**A**) and ZmPOT1b_N (**B**) bound to a radioactively labeled cocktail containing equal amounts of seven oligonucleotides representing all possible permutations of two plant telomere repeats in the absence (lane 1) or the presence of 5X excess of cold competitors representing individual permutations (lanes 2-8). Representative EMSA scans are shown in left panels. For each individual scan, the signal intensity (fraction of protein bound) is plotted on the right with binding in the absence of competitors set at 1.0. (**C**) Competition assay for AoPOT1 bound to the radioactively labeled 2HEXA cocktail containing equal amounts of six oligonucleotides representing all possible permutations of two hexanucleotide telomere repeats in the absence (lane 1) or the presence of 5X excess of cold competitors representing individual permutations (lanes 2-7). Quantification as in (**A**) and (**B**). Arrow denotes protein-DNA complex.

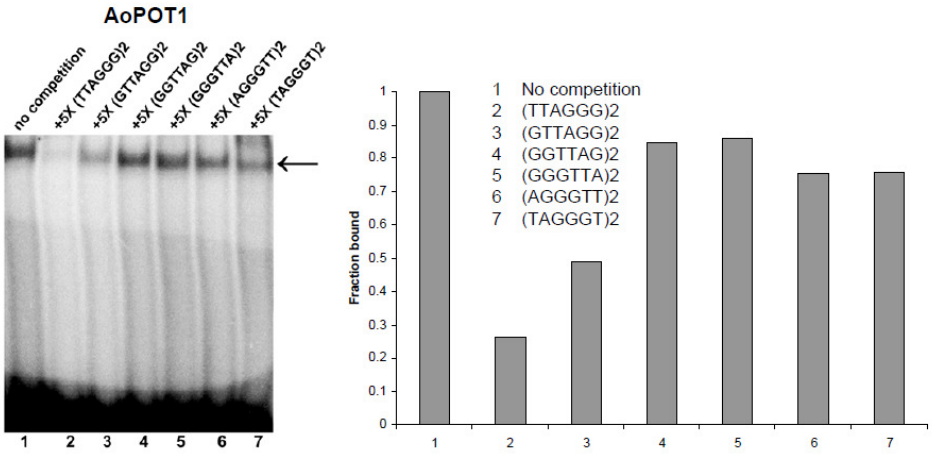
A


OlPOT1_N
no competition
+5X (TTTAGGG)2
+5X (GTTTAGG)2
+5X (GGTTTAG)2
+5X (GGGTTTA)2
+5X (AGGGTTT)2
+5X (TAGGGTT)2
+5X (TTAGGGT)2
1 2 3 4 5 6 7 8
Fraction bound
1 No competition
2 (TTTAGGG)2
3 (GTTTAGG)2
4 (GGTTTAG)2
5 (GGGTTTA)2
6 (AGGGTTT)2
7 (TAGGGTT)2
8 (TTAGGGT)2

B


ZmPOT1b_N
no competition
+5X (TTTAGGG)2
+5X (GTTTAGG)2
+5X (GGTTTAG)2
+5X (GGGTTTA)2
+5X (AGGGTTT)2
+5X (TAGGGTT)2
+5X (TTAGGGT)2
1 2 3 4 5 6 7 8
Fraction bound
1 No competition
2 (TTTAGGG)2
3 (GTTTAGG)2
4 (GGTTTAG)2
5 (GGGTTTA)2
6 (AGGGTTT)2
7 (TAGGGTT)2
8 (TTAGGGT)2

C


AoPOT1
no competition
+5X (TTAGGG)2
+5X (GTTAGG)2
+5X (GGTTAG)2
+5X (GGGTTA)2
+5X (AGGGTT)2
+5X (TAGGGT)2
1 2 3 4 5 6 7
Fraction bound
1 No competition
2 (TTAGGG)2
3 (GTTAGG)2
4 (GGTTAG)2
5 (GGGTTA)2
6 (AGGGTT)2
7 (TAGGGT)2

Next we determined the minimum DNA sequence required for OlPOT1_N binding. EMSA was performed using TAGGGTTTAGGGTT and a series of single nucleotide truncations from either the 5' or 3' end of this substrate (Fig A-5A). Deletion of the first two nucleotides from the 5' end (oligonucleotide 12) did not decrease binding. However, removal of three nucleotides decreased binding by over 30% (oligonucleotide 11') and deletion of four abolished nearly all binding (oligonucleotide 10''). In both cases, a smear trailing down to free probe was observed, suggesting partial dissociation of the complex during electrophoresis. Only one nucleotide could be removed from the 3'end without detectable loss of DNA binding (oligonucleotide 11). Deletion of two nucleotides (oligonucleotide 10) decreased binding to only 15% of the full-length oligonucleotide. Therefore, the minimum tight-binding sequence (core MBS) for OlPOT1_N appears to be GGTTTAGGGT (oligonucleotide 10'). Addition of one G residue at the oligonucleotide 5' end improved binding almost two-fold (oligonucleotide 11). Thus, as little as 10 nucleotides are necessary for OlPOT1_N binding and may comprise its MBS, while a 11-nt GGGTTTAGGGT oligonucleotide represents the best tight-binding substrate.

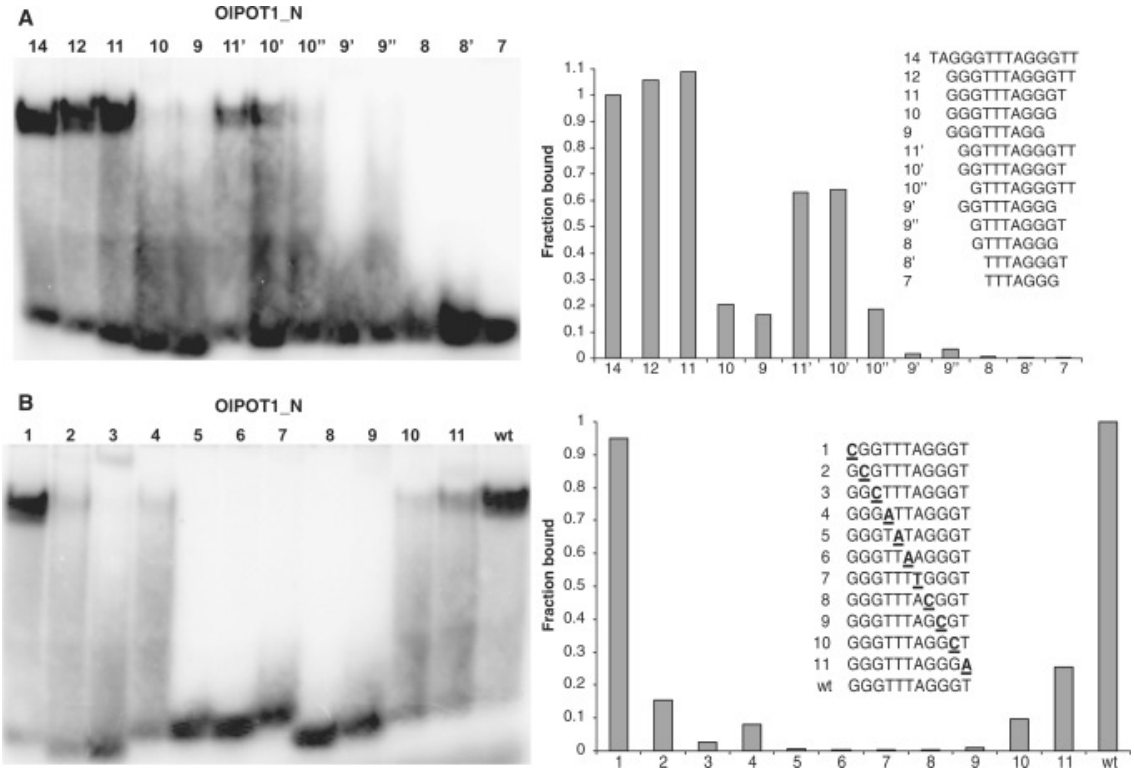


Fig A-5. Characterization of DNA binding activity of recombinant POT1_N protein from *Ostreococcus lucimarinus*. (**A**) EMSA identifies the minimum binding site of OlPOT1_N. Equal amounts of OlPOT1_N were incubated with the indicated oligonucleotides and the protein-DNA complexes were separated by native PAGE. (**B**) Identification of nucleotides recognized by OlPOT1_N in GGGTTTAGGGT. Numbers indicate nucleotide positions that were substituted with complementary nucleotides (bold and underlined). Representative EMSA scans are shown in left panels. For each scan, the signal intensity (fraction of protein bound) is plotted on the right with binding to TAGGGTTTAGGGTT (**A**) or GGGTTTAGGGT (wt) (**B**) set at 1.0.

Studies with non-plant POT1 proteins reveal that only a subset of nucleotides within the MBS are specifically recognized (Croy & Wuttke, 2006). To identify nucleotides in the OlPOT1_N MBS critical for protein interaction, a series of oligonucleotides with single complementary nucleotide substitutions were tested for OlPOT1_N binding (Fig A-5B). Unexpectedly, we discovered that all core MBS nucleotides are crucial for OlPOT1_N binding (Fig A-5B, lanes 2-11). The only nucleotide change that did not affect OlPOT1_N binding was the most 5'-terminal G (Fig A-5B, lane 1), which is not a part of the core MBS. A nucleotide in this position may be required for improved protein-DNA complex stability, but may not contribute to specific interactions with OlPOT1_N protein. Overall, we conclude that OlPOT1_N requires at least 10 telomeric nucleotides for efficient binding. However, unlike the vertebrate and yeast POT1 proteins, all nucleotides in the core MBS appear to make specific and crucial contacts with OlPOT1_N.

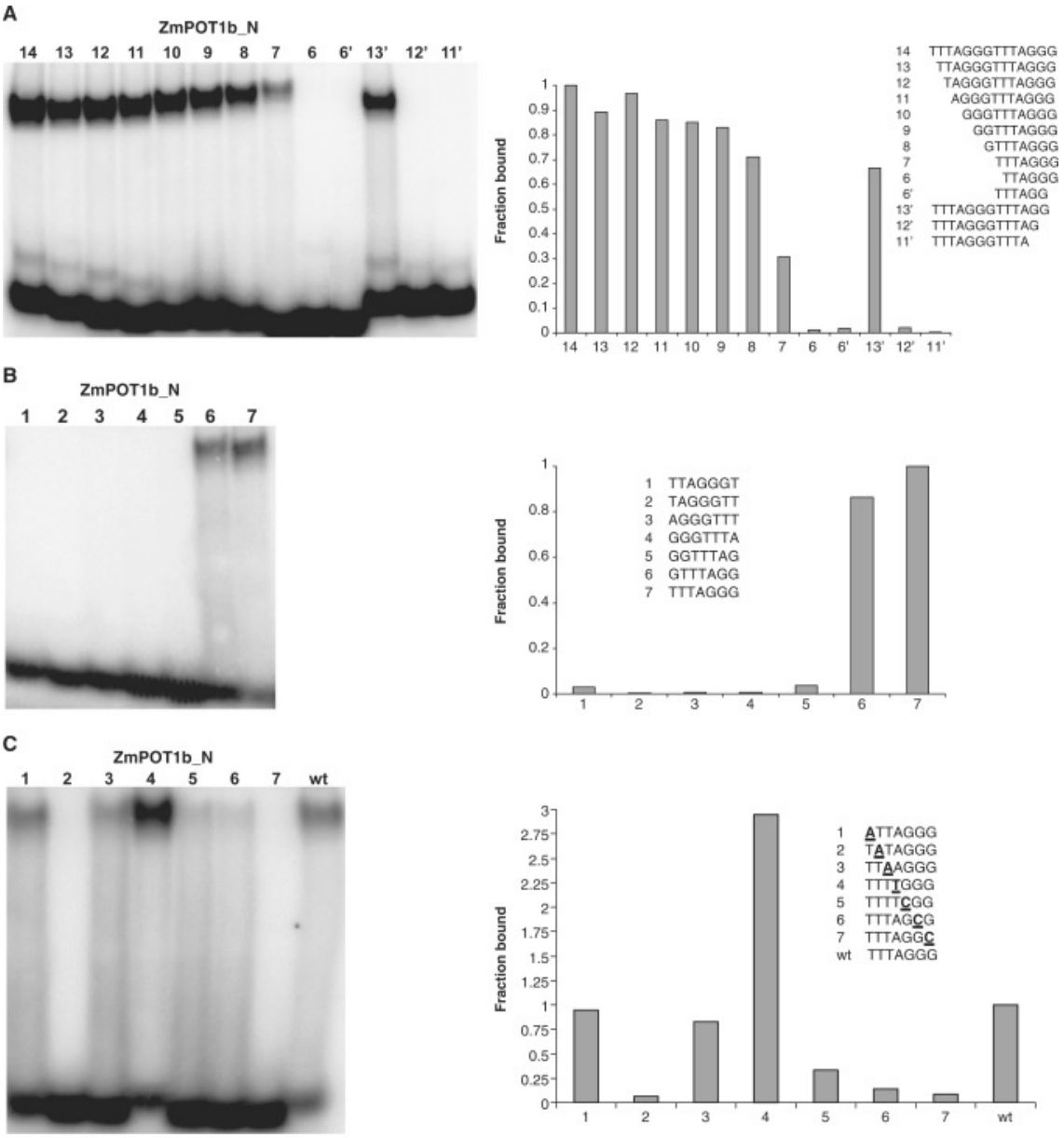
DNA binding properties of POT1b_N from maize

Maize is an angiosperm species that harbors canonical plant TTTAGGG telomere repeats (Burr et al, 1992). We tested ZmPOT1b_N binding to the seven permutations of the plant telomere repeat sequence using competition assays (Fig A-4B). Although incubation with several 2PLANT competitors leads to slightly decreased signal intensity (Fig A-4B, lanes 6-8), only the addition of $(TTTAGGG)_2$ and $(GTTTAGG)_2$ resulted in significant competition (Fig A-4B, lanes 2 and 3). These data indicate that ZmPOT1b_N recognizes two different permutations of the plant telomere repeat.

Next we determined the minimum number of nucleotides required for ZmPOT1b_N interaction with telomeric DNA (Fig A-6A). In contrast to the situation with OlPOT1_N, removal of as many as 7 nucleotides from the 5'-end of the oligonucleotide (a full plant TTTAGGG repeat) did not abolish binding (Fig A-6A, oligonucleotide 7). On the other hand, only a single nucleotide could be removed from the 3'-end of $(TTTAGGG)_2$ oligonucleotide (Fig A-6A, compare oligonucleotides 13' and 12'). Since two different probes containing 6 nucleotides each failed to bind ZmPOT1b_N (Fig A-6A, lanes 6 and 6'), these data indicate that the MBS necessary for efficient ZmPOT1b_N binding consists of seven nucleotides, one full plant repeat.

To further evaluate the MBS of ZmPOT1b_N, we tested ZmPOT1b_N binding to oligonucleotides representing the seven permutations of the plant telomere repeat (1PLANT). Similar to the permutation analysis of 14-nt 2PLANT probes (Fig A-4B), ZmPOT1b_N stably associated with two 7-nt 1PLANT probes, TTTAGGG and GTTTAGG (Fig A-6B, lanes 6 and 7). Among the two, TTTAGGG appeared to be a slightly better substrate, suggesting that this sequence represents the preferred MBS for ZmPOT1b_N.

Fig A-6. Characterization of ZmPOT1b_N interaction with telomeric DNA. (**A**) Identification of the MBS for ZmPOT1b_N. Equal amounts of ZmPOT1b_N were incubated with the indicated oligonucleotides and protein-DNA complexes were separated by native PAGE. (**B**) Analysis of ZmPOT1b_N binding to different permutations of a single plant telomere repeat. Numbers indicate seven possible repeat permutations. (**C**) Analysis of nucleotides specifically recognized by ZmPOT1b_N in TTTAGGG. Numbers indicate nucleotide positions that were substituted with complementary nucleotides (bold and underlined). For all panels, EMSA scans are shown on the left, and radioactive signal intensity is plotted on the right with ZmPOT1b_N binding to TTTAGGGTTTAGGG (**A**) or TTTAGGG (**B**, **C**) set at 1.0.


A
ZmPOT1b_N
14 13 12 11 10 9 8 7 6 6' 13' 12' 11'
Fraction bound
14 TTTAGGGTTTAGGG
13 TTAGGGTTTAGGG
12 TAGGGTTTAGGG
11 AGGGTTTAGGG
10 GGGTTTAGGG
9 GGTTTAGGG
8 GTTTAGGG
7 TTTAGGG
6 TTAGGG
6' TTTAGG
13' TTTAGGGTTTAGG
12' TTTAGGGTTTAG
11' TTTAGGGTTTA
B
ZmPOT1b_N
1 2 3 4 5 6 7
Fraction bound
1 TTAGGGT
2 TAGGGTT
3 AGGGTTT
4 GGGTTTA
5 GGTTTAG
6 GTTTAGG
7 TTTAGGG
C
ZmPOT1b_N
1 2 3 4 5 6 7 wt
Fraction bound
1 ATTAGGG
2 TATAGGG
3 TTAAGGG
4 TTTTGGG
5 TTTTCGG
6 TTTAGCG
7 TTTAGGC
wt TTTAGGG

We tested the relative importance of each MBS nucleotide for efficient ZmPOT1b_N binding using complementary nucleotide substitutions (Fig A-6C). ZmPOT1b_N binding was significantly reduced or abolished with substitutions in positions 2 (TATAGGG), 5 (TTTACGG), 6 (TTTAGCG) and 7 (TTTAGGC) (Fig A-6C, lanes 2, 5-7). Nucleotide changes in position 1 (ATTAGGG) and position 3 (TTAAGGG) did not lead to a substantial decrease in ZmPOT1b_N binding, while a change in position 4 (TTTTGGG) actually improved binding three-fold over the wild-type (Fig A-6C, compare lane 4 with wt). These data suggest that unlike the situation with OlPOT1_N, not all nucleotides in ZmPOT1b_N MBS make crucial contributions to binding.

Interestingly, only three out of seven nucleotide positions (1, 4 and 5) differ in the two acceptable ZmPOT1b_N permutations, $\underline{T}_1T_2T_3\underline{A}_4\underline{G}_5G_6G_7$ and $\underline{G}_1T_2T_3\underline{T}_4\underline{A}_5G_6G_7$. Since changes in positions 1 and 4 do not lead to decreased protein binding, the increased ZmPOT1b_N binding to TTTAGGG versus GTTTAGG may reflect a difference at position 5, with G being preferred over A. Overall, although TTTAGGG is the natural telomere repeat sequence in both *O. lucimarinus* and maize, the MBS and the relative importance of individual nucleotides within it vary dramatically between OlPOT1_N and ZmPOT1b_N.

DNA binding properties of Asparagus POT1 protein

The plant lineages leading to maize and Asparagus diverged only 100-110 mya (Wikstrom et al, 2001), a relatively recent event in the evolutionary history of land plants. To study Asparagus POT1, we first confirmed that *Asparagus officinalis* possesses hexanucleotide TTAGGG repeats using the TRF assay. As expected, a consensus plant telomere probe $(T_3AG_3)_4$ hybridized to the control *Arabidopsis* telomeric DNA (Fig

A-7A, lanes 1 and 3), but not to Asparagus DNA (lanes 2 and 4), while a TTAGGG-specific telomere probe recognized Asparagus DNA, but not *Arabidopsis* (Fig A-7A, lanes 5-8). Strikingly, Asparagus telomere tracts appear to be at least an order of magnitude longer than in *Arabidopsis* (Fig A-7A, compare lanes 1 and 3 with 6 and 8).

To evaluate Asparagus POT1 binding to the hexanucleotide telomere repeat sequences, we examined the affinity of full-length AoPOT1 for an oligonucleotide cocktail containing equal amounts of all six possible permutations of two TTAGGG telomere repeats (2HEXA cocktail probe). A single shifted band was formed (Fig A-4C, lane 1). To determine which permutation(s) of the human telomere repeat were recognized, competition assays were performed with a 5-fold excess of individual cold 2HEXA oligonucleotides (Fig A-4C, lanes 2-7). Among all 2HEXA oligonucleotides, $(TTAGGG)_2$ was the most efficient competitor (Fig A-4C, lane 2). The specificity of AoPOT1 for $(TTAGGG)_2$ was confirmed in a direct binding assay. AoPOT1 efficiently bound $(TTAGGG)_2$, while the mouse mPOT1a_N protein was unable to bind this repeat permutation (Supplementary Fig A-8B, compare lanes 1 and 3). Thus, although the telomere repeat sequence in Asparagus and vertebrates is the same, AoPOT1 prefers the permutation terminating in TTAGGG, while mPOT1a_N and other vertebrate POT1 proteins prefer the permutation terminating in GGTTAG (He et al, 2006; Lei et al, 2004; Loayza et al, 2004; Wei & Price, 2004).

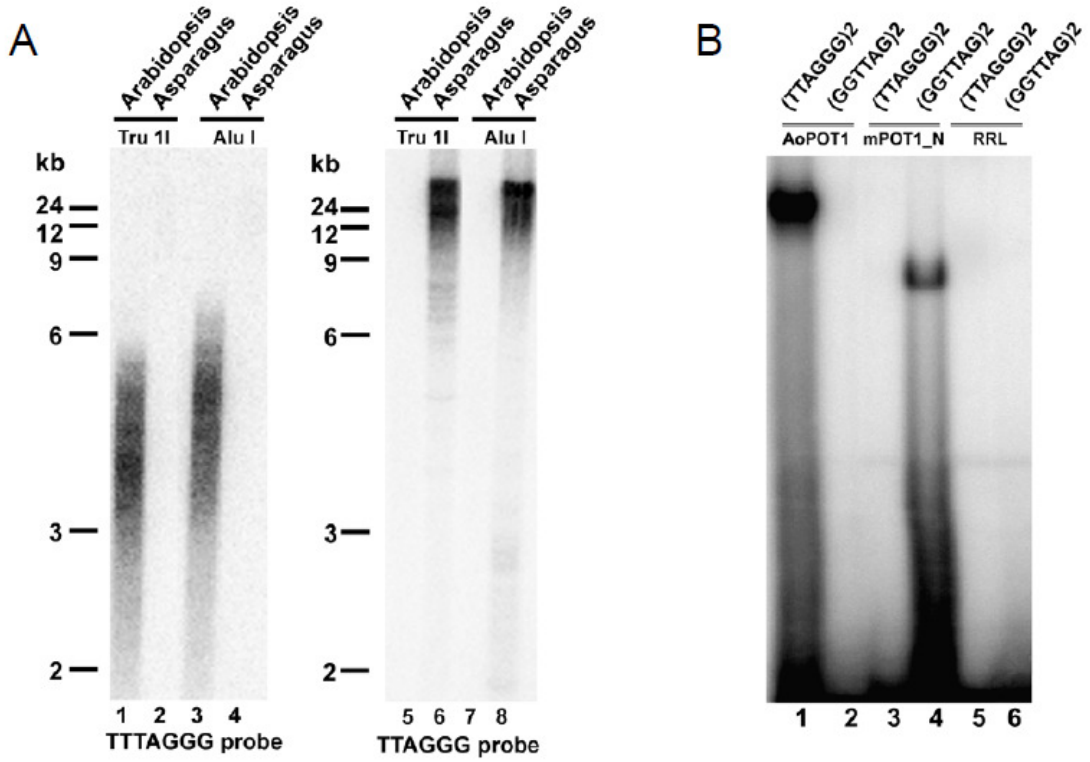


Fig A-7. Analysis of telomere repeat composition and POT1 binding in Asparagus. (**A**) Terminal restriction fragment analysis (TRF) of Asparagus telomeres. Asparagus and *Arabidopsis* genomic DNA was digested with Tru1I or AluI. Digested DNA was hybridized first with a plant telomere-specific $(TTTAGGG)_4$ probe (lanes 1-4), then stripped and re-hybridized again with an Asparagales and vertebrate telomere-specific $(TTAGG)_4TTAG$ probe under the same conditions (lanes 5-8). Molecular weight markers are shown on the left of each panel. (**B**) EMSA results for AoPOT1 and mPOT1a_N. AoPOT1 and mPOT1a_N were incubated with $(TTAGGG)_2$ (lanes 1 and 3, respectively) and $(GGTTAG)_2$ (lanes 2 and 4, respectively). RRL alone used as a negative control for both DNA oligonucleotides (lanes 5 and 6).

AoPOT1 binding was assessed with a series of single-nucleotide truncations of the $(TTAGGG)_2$ oligonucleotide to define the MBS. Removal of even a single nucleotide from the 3'-end of the oligonucleotide was not tolerated (Fig A-8A, oligonucleotides 11'-7'). On the other hand, removal of as many as 4 nucleotides from the 5'-end did not completely abolish binding, though it led to a substantial reduction in signal intensity (Fig A-8A, oligonucleotide 8). While these data suggest that the MBS is the 8-nt GGTTAGGG, we noticed that AoPOT1 binding to a 9-nt oligonucleotide GGGTTAGGG is improved two-fold (oligonucleotide 9), suggesting that GGGTTAGGG represents the best tight-binding substrate.

We next examined the relative contribution of each MBS nucleotide to AoPOT1 binding using a series of complementary nucleotide substitutions in GGGTTAGGG (Fig A-8B). Although mutations in the first three 5'-terminal positions decreased signal intensity to only 34-45% of the original oligonucleotide (Fig A-8B, oligonucleotides 1-3), mutations in all other positions abolished binding almost completely (Fig A-8B, oligonucleotides 4-9). Thus, nucleotides 4-9 (a complete TTAGGG telomere repeat) make the most important contributions to AoPOT1 binding, while guanines 1-3 may be important for the stability of AoPOT1-DNA interaction. This situation is different from OlPOT1_N, where all the nucleotides in the MBS are important for binding, and ZmPOT1b_N, where several nucleotide positions in the MBS exhibit more relaxed specificity.

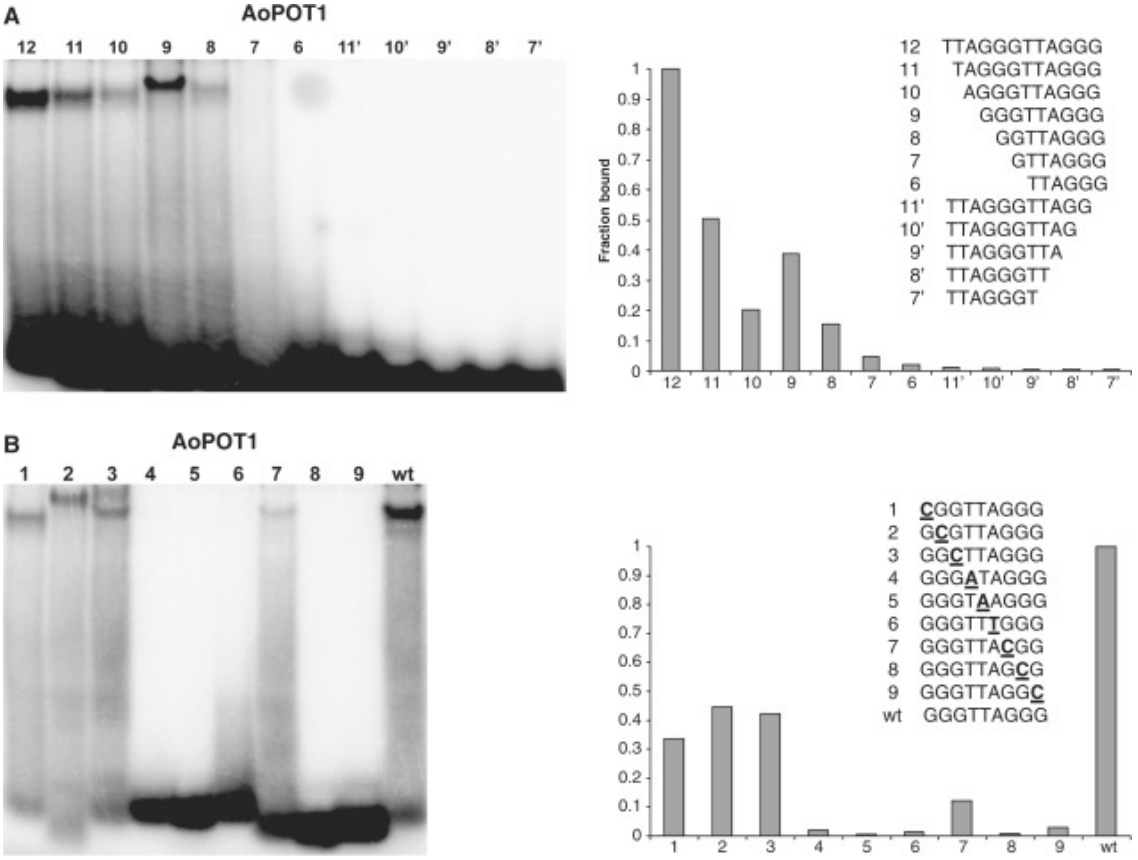


Fig A-8. Analysis of Asparagus POT1 interaction with telomeric DNA. (**A**) Identification of the MBS for AoPOT1. Equal amounts of AoPOT1 were incubated with the indicated oligonucleotides and protein-DNA complexes were separated by native PAGE. (**B**) Analysis of nucleotides specifically recognized by AoPOT1 in GGGTTAGGG. Numbers indicate nucleotide positions that were substituted with complementary nucleotides (bold and underlined). For all panels, EMSA scans are shown on the left, and radioactive signal intensity is plotted on the right with AoPOT1 binding to TTAGGGTTAGGG (**A**) or GGGTTAGGG (**B**) set at 1.0.

POT1 proteins from Asparagus and maize bind both hexa- and heptanucleotide telomere repeats.

Our TRF and EMSA data indicate that AoPOT1 may bind TTAGGG repeats *in vivo*, however it was unclear whether this protein retained the ability to interact with the ancestral TTTAGGG sequence. We found that AoPOT1 binding to two heptanucleotide repeats is two-fold better than to two hexanucleotide repeats (Fig A-9, lanes 1 and 2), suggesting that the ancestral TTTAGGG sequence is still a preferred substrate for AoPOT1. We next tested whether the ability to bind both types of telomere repeats is a conserved feature of POT1 proteins. As with AoPOT1, ZmPOT1b_N bound both types of repeats (Fig A-9, lanes 7 and 8), displaying 2.5-fold better binding to $(TTTAGGG)_2$ than to $(TTAGGG)_2$. In striking contrast, we could not detect binding by *O. lucimarinus* POT1_N to any permutation of the hexanucleotide repeat in a direct binding assay (Fig A-9, lanes 4 and 10). Similarly, mouse POT1a_N failed to bind any permutation of the plant telomere repeat (Fig A-9, lanes 6 and 11). Thus, recognition of both hexanucleotide and heptanucleotide telomere repeats is an evolutionarily conserved feature of POT1 proteins from maize and Asparagus, which dates back to at least 100-110 mya. The inability of OlPOT1_N to bind hexanucleotide repeats suggests that the algal POT1 protein either lost the ability to bind such repeats after the divergence of land plants and green algae, or this property evolved independently in angiosperms.

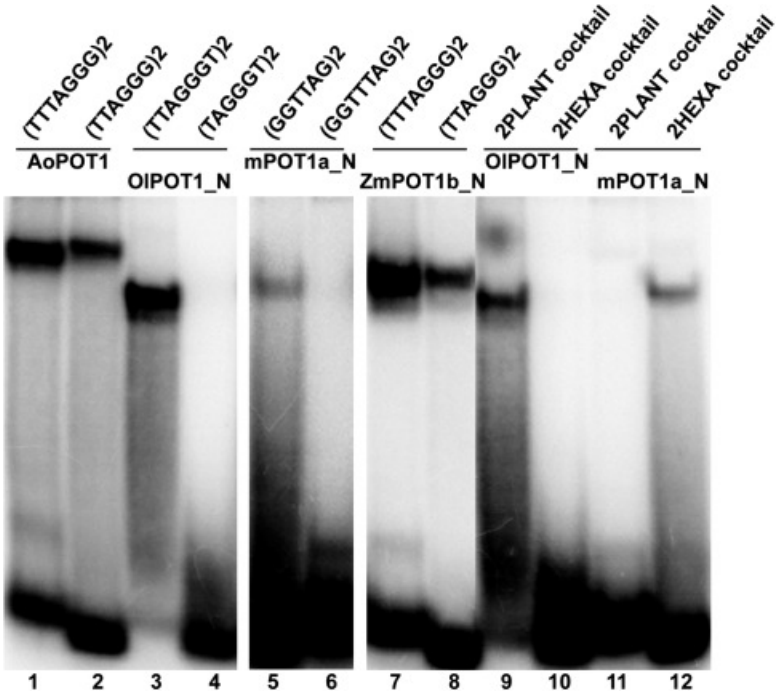


Fig A-9. AoPOT1 and ZmPOT1b_N bind both hepta- and hexanucleotide telomere repeats. EMSA results are shown for AoPOT1, OlPOT1_N, mPOT1a_N and ZmPOT1b_N. POT1 proteins were incubated with the indicated oligonucleotides consisting of two full hepta- or hexanucleotide telomere repeats, or with an oligonucleotide cocktail containing all possible permutations of the heptanucleotide or hexanucleotide telomere repeat.

We noted that the first six 3'-terminal positions in the MBS of POT1 proteins from angiosperms (GGGTTAGGG for AoPOT1 and TTTAGGG for ZmPOT1b_N) are identical and constitute one full hexanucleotide repeat. The remaining 5'-terminal nucleotides in each MBS do not appear to make crucial contacts with the corresponding POT1 proteins (Fig A-8B, lanes 1-3 and A-6C, lane 1). Thus, the ability of AoPOT1 and ZmPOT1b_N to bind both hexa- and hepta-nucleotide telomere repeats can be explained if both proteins fail to discriminate between different nucleotides in the 5'-terminal positions of their respective MBS. In support of this model, we found that AoPOT1 and ZmPOT1b_N can bind to all variations of GGNTTAGGG and NTTAGGG, respectively (Fig A-10).

End-binding specificity of plant POT1 proteins

Vertebrate and yeast POT1 proteins differ in their preference for telomeric repeats at the 3'-end of the DNA substrate (Croy & Wuttke, 2006). Therefore, we asked if plant POT1 proteins also exhibit such preference by performing competition assays with oligonucleotides containing two telomere repeats at the 5' or 3'-end or in the middle of a longer DNA oligonucleotide (Fig A-11).

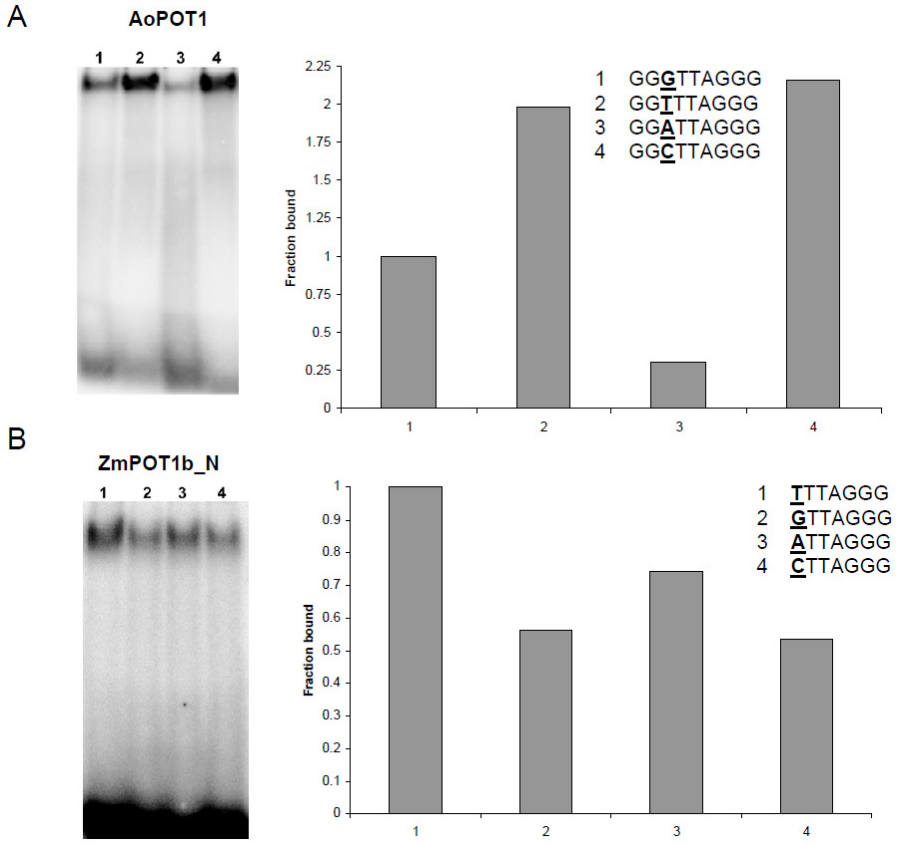
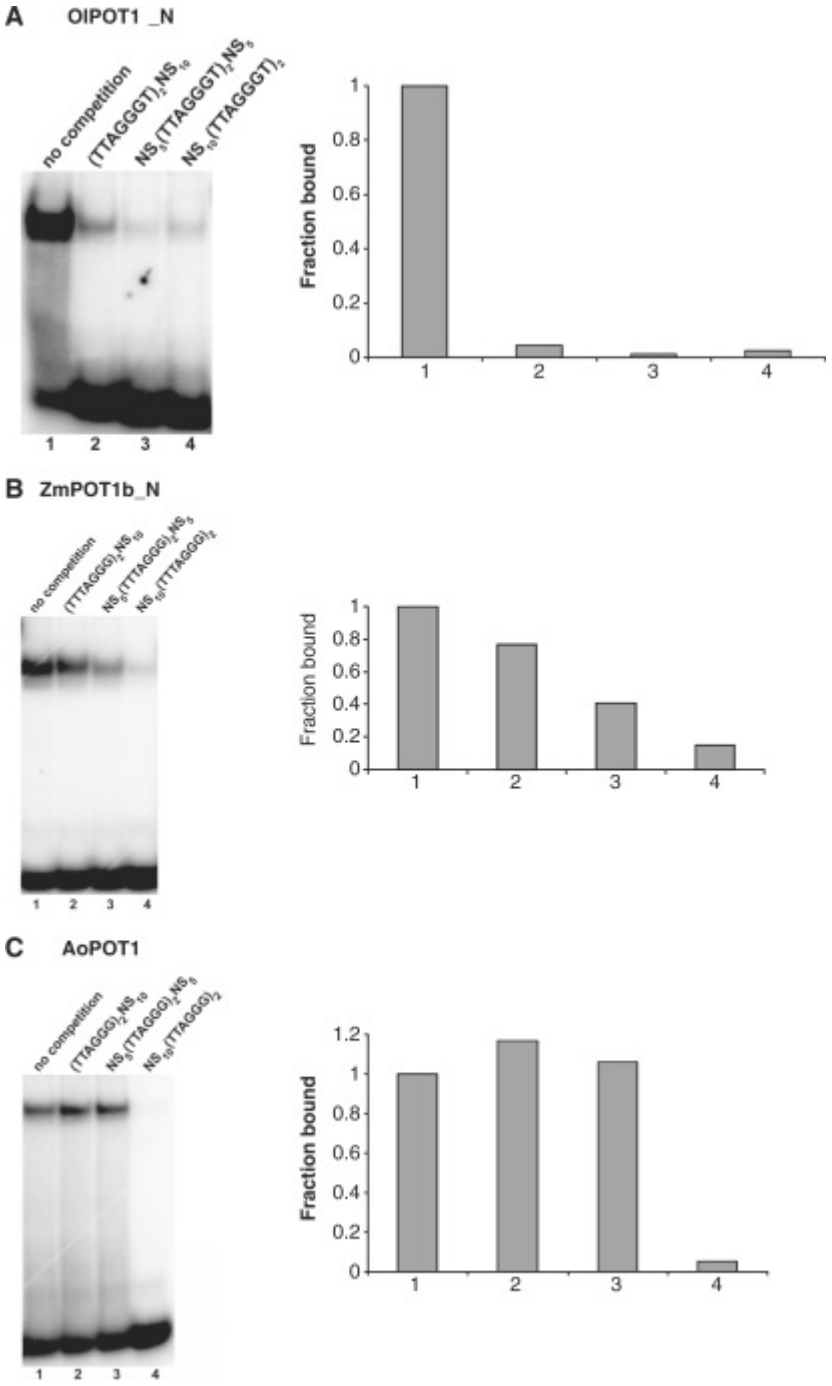


Fig A-10. AoPOT1 ZmPOT1b_N accommodate guanine nucleotides instead of thymidine in the context of hexanucleotide telomere repeats. (**A**) AoPOT1 was incubated with the MBS sequence GGGTTAGGG (lane 1) or three oligonucleotides in which the seventh 3'-terminal G was substituted for T (lane 2), A (lane 3) or C (lane 4). (**B**) ZmPOT1b_N was incubated with either its MBS TTTAGGG (lane 1), or three oligonucleotides in which the seventh 3'-terminal T was substituted for C (lane 2), A (lane 3) or G (lane 4). Representative EMSA scans are shown in left panels. For each individual scan, the signal intensity is plotted on the right with AoPOT1 binding to GGGTTAGGG (**A**) or ZmPOT1b_N binding to TTTAGGG (wt) (**B**) set at 1.0.

Fig A-11. End-binding preference of plant POT1 proteins. (**A**) EMSA results are shown for reactions with OlPOT1_N using radioactively labeled $(TTAGGGT)_2$ oligonucleotide as a probe in the absence (lane 1) or presence of 5X excess cold competitors containing the same telomeric sequence located either 5'-terminally (lane 2) or 3'-terminally (lane 4) to the 10-nt non-telomeric sequence NS10 (CTCTACCAAA), or flanked by 5-nt non-telomeric NS5 sequences (CTCTA and CCAAA) on both ends (lane 3). The fraction of complex bound to the labeled oligonucleotide is plotted on the right with binding in the absence of competitor set at 1.0. (**B**) and (**C**) Competition assays for ZmPOT1b_N bound to radioactively labeled $(TTTAGGG)_2$ oligonucleotide (**B**) and AoPOT1 bound to radioactively labeled $(TTAGGG)_2$ oligonucleotide (**C**). Lane designation and quantification as in (**A**).

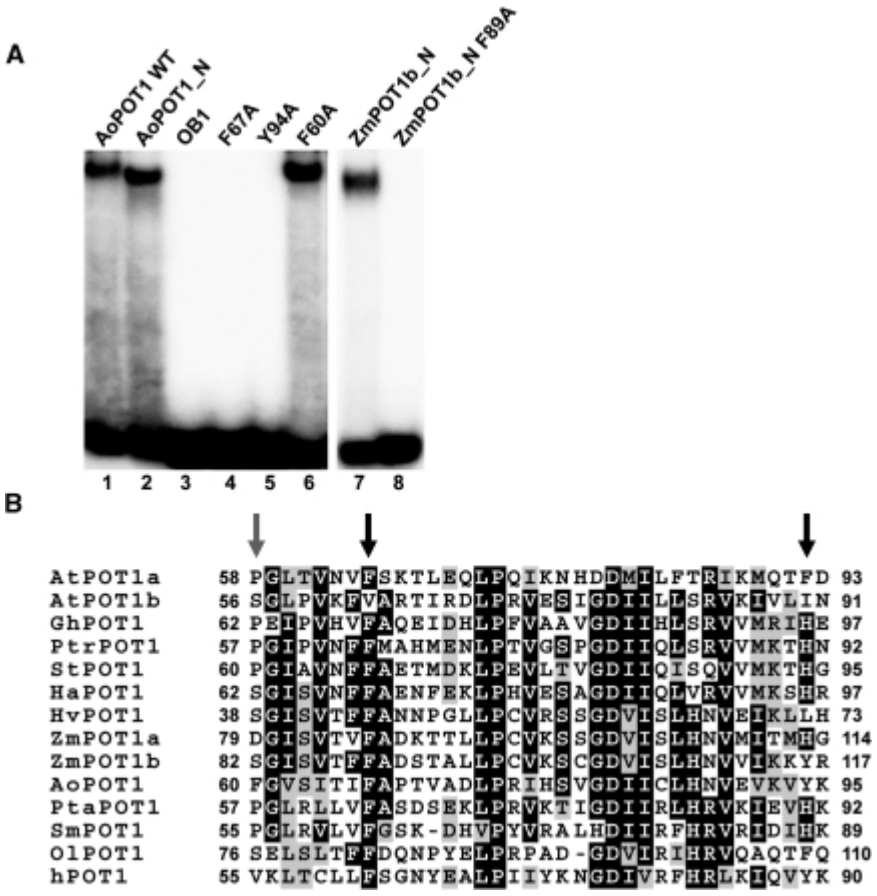

A OlPOT1 _N
no competition
(TTAGGGT)2NS10
NS5(TTAGGGT)2NS5
NS10(TTAGGGT)2
1 2 3 4
Fraction bound
1
0.8
0.6
0.4
0.2
0
1 2 3 4
B ZmPOT1b_N
no competition
(TTAGGG)2NS10
NS5(TTAGGG)2NS5
NS10(TTAGGG)2
1 2 3 4
Fraction bound
1
0.8
0.6
0.4
0.2
0
1 2 3 4
C AoPOT1
no competition
(TTAGGG)2NS10
NS5(TTAGGG)2NS5
NS10(TTAGGG)2
1 2 3 4
Fraction bound
1.2
1
0.8
0.6
0.4
0.2
0
1 2 3 4

Interestingly, the plant POT1 proteins behaved differently in these competition assays. Complex formation of OlPOT1_N with labeled $(TTAGGGT)_2$ was significantly reduced or abolished with all competitors, suggesting that this protein binds telomeric repeats regardless of their position in the substrate (Fig A-11A). On the other hand, increasing competition was observed for ZmPOT1b_N with oligonucleotides carrying the telomere repeats on the 5' end, middle and on the 3' end (Fig A-11B), respectively. Finally, AoPOT1 binding to $(TTAGGG)_2$ could only be competed with a substrate harboring two hexanucleotide telomere repeats on the 3'-end of the oligonucleotide (Fig A-11C). We conclude that preference for the position of telomeric repeats on the DNA substrate is species-specific and not evolutionarily conserved.

Mutational analysis of plant POT1 proteins

Structurally characterized POT1 proteins from humans and S. pombe share a number of conserved primary sequence and secondary structure elements, which are crucial for specific interaction with telomeric DNA (Lei et al, 2003; Lei et al, 2004). The availability of Asparagus POT1 provides an opportunity to analyze the importance of these amino acids and protein regions for TTAGGG repeat recognition in the context of a full-length plant POT1 protein. As expected from studies in yeast and humans (Baumann & Cech, 2001; He et al 2006; Wu et al, 2006), the C-terminal region of AoPOT1 (amino acids 322-504) was dispensable for DNA binding (Fig A-12A, lane 2).

Fig A-12. Mutational analysis of plant POT1 proteins. (**A**) EMSA assays of AoPOT1 and ZmPOT1b_N truncation and point mutants. DNA binding reactions were performed with wild-type AoPOT1 (lane 1), AoPOT1 truncation constructs (lanes 2 and 3) and point mutants (lanes 4 - 6) as well as with wild-type ZmPOT1b_N (lane 7) and its point mutant F89A (lane 8). The labeled probes are GGGTTAGGG for AoPOT1 and TTTAGGG for ZmPOT1b_N. (**B**) Partial alignment of plant POT1 proteins with human POT1. An OB1 region with a high degree of inter-kingdom amino acid similarity is shown. Black arrows indicate the positions of two catalytically important human POT1 residues and the corresponding aromatic amino acids F67 and Y94 in AoPOT1. Grey arrow designates the location of F60 in AoPOT1, which has no effect on telomeric DNA binding. Numbers indicate amino acid positions relative to the start codon. Abbreviations: At, *Arabidopsis thaliana*; Gh, *Gossypium hirsutum*; Ptr, *Populus trichocarpa*; St, *Solanum tuberosum*; Ha, *Helianthus argophyllus*; Hv, *Hordeum vulgare*; Zm, *Zea mays*; Ao, *Asparagus officinalis*; Pta, *Pinus taeda*; Sm, *Selaginella moellendorffii*; Ol, *Ostreococcus lucimarinus*. Alignment was generated with MEGA 3 software (Kumar et al, 2004) and visualized in the BOXSHADE format.


A
AoPOT1 WT
AoPOT1_N
OB1
F67A
Y94A
F60A
ZmPOT1b_N
ZmPOT1b_N F89A
1 2 3 4 5 6 7 8
B
AtPOT1a 58 PGLTVNVFSKTLEQLPQIKNHDDMILFTRIKMQTFD 93
AtPOT1b 56 SGLPVKFVARTIRDLPRVESIGDIILLSRVKIVLIN 91
GhPOT1 62 PEIPVHVFAQEIDHLPFVAAVGDIIHLSRVVMRIHE 97
PtrPOT1 57 PGIPVNFFMAHMENLPTVGSPGDIIQLSRVVMKTHN 92
StPOT1 60 PGIAVNFFAETMDKLPEVLTVGDIIQISQVVMKTHG 95
HaPOT1 62 SGISVNFFAENFEKLPHVESAGDIIQLVRVVMKSHR 97
HvPOT1 38 SGISVTFFANNPGLLPCVRSSGDVISLHNVEIKLLH 73
ZmPOT1a 79 DGISVTVFADKTTLLPCVKSSGDVISLHNVMITMHG 114
ZmPOT1b 82 SGISVTFFADSTALLPCVKSCGDVISLHNVVIKKYR 117
AoPOT1 60 FGVSITIFAPTVADLPRIHSVGDIICLHNVEVKVYK 95
PtaPOT1 57 PGLRLLVFASDSEKLPRVKTIGDIIRLHRVKIEVHK 92
SmPOT1 55 PGLRVLVFGSK-DHVPYVRALHDIIRFHRVRIDIHK 89
OlPOT1 76 SELSLTFFDQNPYELPRPAD-GDVIRIHRVQAQTFQ 110
hPOT1 55 VKLTCLLFSGNYEALPIIYKNGDIVRFHRLKIQVYK 90

Moreover, full-length AoPOT1 and AoPOT1_N have similar binding properties, although the latter showed slightly reduced binding to the 8-nt GGTTAGGG substrate (Fig A-13). Further truncation of AoPOT1 to eliminate the second OB-fold (amino acids 1 - 167) completely abolished DNA binding activity (Fig A-12A, lane 3). Thus, two N-terminal OB-folds are sufficient for telomeric DNA binding by AoPOT1.

The crystal structure of human POT1 protein identified conserved F62 and Y89 residues in OB1, which are critical for POT1 interaction with telomeric DNA (Fig A-12B) (Lei et al, 2004). Most plant POT1 proteins have a nearly invariant phenylalanine in the first position (F67 in AoPOT1) and a large amino acid with a bulky side chain (mostly Y, H or F) in the second position (Y94 in AoPOT1). Consistent with previous reports for mammalian POT1 proteins (He et al, 2006; Lei et al, 2004), alanine substitutions of the corresponding Asparagus POT1 amino acids F67 and Y94 completely abolished DNA binding (Fig A-12A, lanes 4, 5). A similar result was obtained in ZmPOT1b_N with a F89A mutation (corresponding to F62 in human POT1) (Fig A-12A, compare lanes 7 and 8). As a control, a F60A mutation in AoPOT1, which affects a non-conserved amino acid, had no effect on AoPOT1 binding (Fig A-12A, lane 6). Altogether, these results argue that evolutionarily conserved aromatic amino acids in OB1 are important for telomeric DNA binding across kingdoms and imply that the overall architecture of OB1 in plant POT1 proteins is similar to that of its mammalian and yeast counterparts.

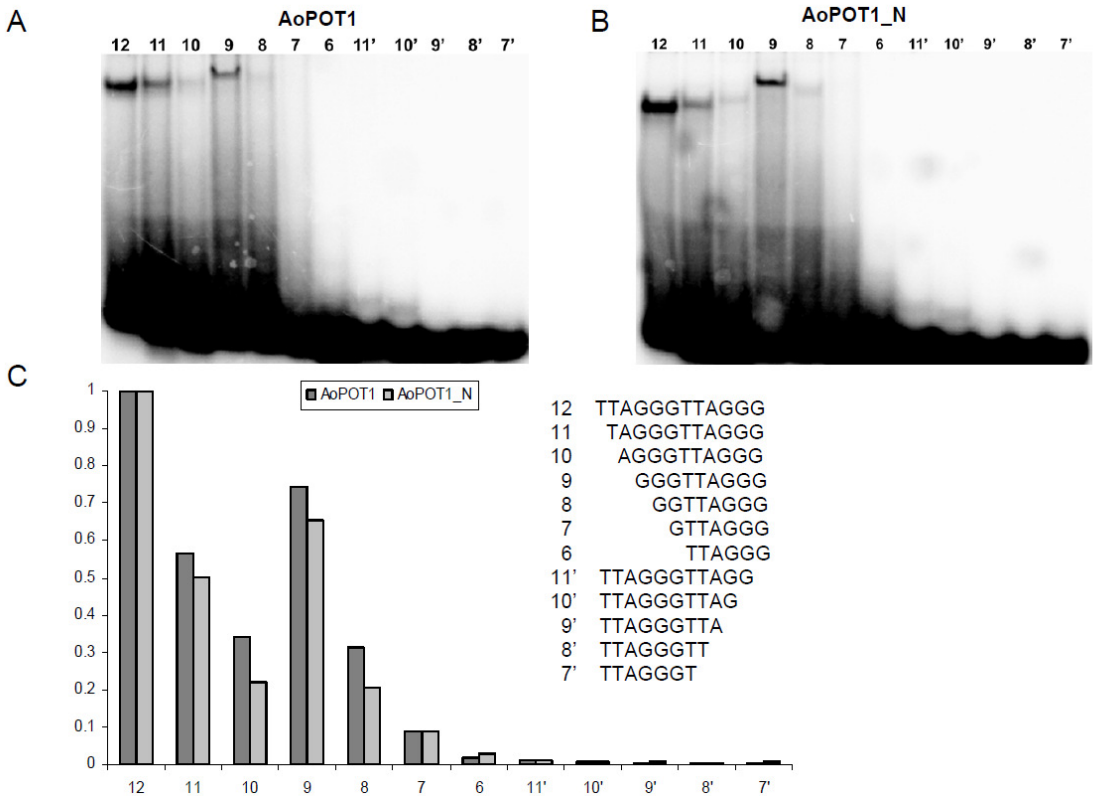


Fig A-13. Comparison of DNA binding properties of AoPOT1 and AoPOT1_N. EMSA results for AoPOT1 and AoPOT1_N. Equal amounts of either AoPOT1 (**A**) or AoPOT1_N (**B**) were incubated with the indicated oligonucleotides and protein-DNA complexes were separated by native PAGE. (**C**) The fraction of AoPOT1 and AoPOT1_N bound to various telomeric substrates in (**A**) and (**B**) was plotted with binding to the TTAGGGTTAGGG oligonucleotide containing two full hexanucleotide telomere repeats set at 1.0.

Discussion

The POT1 protein family represents an evolutionarily conserved group of telomeric DNA-binding factors with essential functions in chromosome end protection and telomere length regulation. Although the major POT1 functions appear to be conserved in most branches of eukaryotic life, previous data in plants indicated that *Arabidopsis* POT1 proteins evolved unusual functions in regulating telomerase, a property not dependent on physical contact with telomeric DNA (Surovtseva et al, 2007). This study addresses this phenomenon further and provides a reconciling view that in some plant species POT1 may indeed have functions similar to those described for non-plant POT1 proteins.

DNA binding properties of plant POT1 proteins

A well-established phylogenetic hierarchy for the major groups within the green plant lineage allows us an opportunity to examine changes in telomere-related genes in an evolutionary context. Here we compare the DNA binding properties of three plant POT1 proteins: OlPOT1_N, from the earliest branching lineage analyzed in our study, and AoPOT1 and ZmPOT1b_N from angiosperms. Our biochemical analysis reveals fundamental differences in the nucleic acid binding activity of POT1 proteins across the plant kingdom (Table A-1).

Table A-1. Summary of MBS sequences from POT1 proteins.

Species	POT1 protein	MBS length	MBS sequence [a]	Refs
H. sapiens	HsPOT1	10	T**TAGG**GTTA**G**	(Lei et al, 2004)
S. pombe	SpPot1pN	6	**GGTTA**C	(Lei et al, 2002)
S. pombe	SpPot1$^{1-389}$	12	G**GTT**ACG**GTT**AC	(Croy et al, 2006)
C. elegans	CeOB1	12	TTAGGCTTAGGC [b]	(Raices et al, 2008)
C. elegans	CeOB2	6	GCCTAA [b]	(Raices et al, 2008)
G. gallus	cPOT1	12	GGTTAGGGTTAG [b]	(Wei & Price, 2004)
O. lucimarinus	OlPOT1_N	10	**GGTTTAGGGT**	This study [c]
Z. mays	ZmPOT1b_N	7	T**T**TA**GGG**	This study [c]
A. officinalis	AoPOT1	9	GGG**TTAGGG**	This study [c]

[a] Specifically recognized nucleotides are in bold

[b] Specifically recognized nucleotides not determined

[c] Binding measured in qualitative assays

Several of the DNA binding properties of OlPOT1 are reminiscent of non-plant POT1 proteins and contrast sharply with AoPOT1 and ZmPOT1b. Specifically, the minimum tight-binding sequence of OlPOT1_N consists of ten nucleotides, a number similar to MBS of POT1 proteins from yeast and animals (Croy & Wuttke, 2006). In contrast, POT1 proteins from angiosperms require fewer nucleotides for efficient binding, with Asparagus POT1 (8-9 nucleotides) being on the lower end of the spectrum and ZmPOT1b_N displaying a short MBS of only 7 nucleotides. Interestingly, two polypeptides harboring a single OB-fold, *S. pombe* Pot1pN and *C. elegans* CeOB2, recognize a shorter, 6-nucleotide MBS (Lei et al, 2002; Raices et al, 2008). Thus, ZmPOT1b_N has the smallest MBS among all currently characterized POT1 proteins bearing at least two OB-folds.

Second, plant POT1 proteins show significant variation in the way they interact with cognate DNA. For other POT1 proteins, only a subset of MBS nucleotides is specifically recognized and makes important contributions to binding (Croy & Wuttke, 2006). A similar situation is observed for ZmPOT1b_N, where four out of seven nucleotides in the MBS are required for binding. Asparagus POT1 follows the same trend, with six of the nine MBS nucleotides needed for binding, although all these crucial AoPOT1 MBS nucleotides are localized at the 3'-end of the oligonucleotide. In contrast, all 10 MBS nucleotides are required for OlPOT1 binding, a phenomenon not previously observed for other POT1 proteins. These data suggest that the mechanism responsible for specific recognition of single-stranded telomeric DNA may significantly differ between POT1 proteins from green algae and land plants. Overall, we conclude that the telomeric DNA binding properties of plant POT1 proteins are evolving rapidly.

Evidence for co-evolution of telomeric DNA and POT1 proteins in plants

The biochemical similarities and differences between ZmPOT1b_N and AoPOT1 may provide clues to the apparent co-evolution of Asparagus POT1 and the telomere repeat sequence in this species. ZmPOT1b_N requires only 7 nts (one full TTTAGGG repeat) for efficient binding, while AoPOT1 requires a very similar, but longer nine-nucleotide substrate GGGTTAGGG. Since the three 5'-terminal Gs do not appear to make significant contributions to AoPOT1 binding, our data suggest that these additional nucleotides in Asparagus POT1 MBS may stabilize this protein's interaction with shorter human-type telomere repeats. Notably, the remaining six nucleotides in AoPOT1 MBS, TTAGGG, are identical to the 3'-end nucleotides present in ZmPOT1b_N MBS. Complementary substitutions in most of these nucleotides completely abolish protein binding, suggesting that TTAGGG sequence is crucial for specific interaction with POT1 proteins. These data may also explain the requirement of AoPOT1 and, to a lesser extent, ZmPOT1b_N for the presence of telomeric repeats on the 3'-end of DNA oligonucleotides. Such 3'-end positioning is likely necessary to improve or stabilize Asparagus POT1 interaction with the G-overhang, which, in turn, may result in better regulation of G-overhang length or interaction with telomerase.

What is the molecular basis for the stable association of POT1 proteins from angiosperms with hexanucleotide telomere repeats? Since both ZmPOT1b_N and AoPOT1 are capable of specifically binding to TTAGGG repeats, this biochemical feature must have evolved in the common ancestor of maize and Asparagus. We note that AoPOT1 and ZmPOT1b have a similar tolerance to nucleotide substitutions in certain MBS positions. Specifically, two different nucleotides, T (the extra T nucleotide present only in the plant TTTAGGG repeat, but not in the hexanucleotide TTAGGG

sequence) and G (which replaces this T in the context of AoPOT1 MBS GGGTTAGGG) are equally tolerated by ZmPOT1b and AoPOT1 in the seventh MBS position (counting from the 3'-end of each oligonucleotide). This property may have originally evolved as a response to the known ability of many plant telomerases to naturally generate mutant telomere repeats containing one less or one more T (Fitzgerald et al, 2001; Mizuno et al, 2008; Shakirov et al, 2008). These so-called T-slippage events represent the most commonly detected type of telomerase error in plants *in vitro* and *in vivo*. The ability to tolerate T-slippage could have potentially allowed the ancestral POT1 protein to remain bound to the mutant human-type TTAGGG sequences. Another possibility is that decreased affinity to the cognate telomere repeats may help to dislodge POT1 from the G-overhang by other DNA-binding proteins, such as RPA or the CST complex during telomere replication (Gao et al, 2007).

Recognition of non-cognate telomerfe repeats is not unique to higher plant POT1 proteins. *Oxytricha nova* telomere end-binding protein (OnTEBP) stably associates with non-cognate telomeric sequences by facilitating significant conformational changes in DNA oligonucleotides via a phenomenon termed nucleotide shuffling, during which DNA sequence register shifts and entire nucleotides are excluded from the protein-DNA complex (Theobald & Schultz, 2003). Chicken and human POT1 proteins are also capable of interacting with non-cognate DNA oligonucleotides in competition assays (Baumann & Cech, 2001; Wei & Price, 2004). Similarly, in direct EMSA assays, *Schizosaccharomyces pombe* POT1 specifically binds DNA sequences resembling telomere repeats present in *T. thermophyla, O. nova* and even *S. cerevisiae* (Trujillo et al, 2005). Taken together, these observations suggest that relaxed DNA sequence specificity may be a common characteristic of POT1 proteins. This property could be

especially beneficial in organisms such as Paramecium, where telomerase synthesizes an unusually high number of mutant telomere repeats (McCormick-Graham et al, 1997). Likewise, in Asparagus plant lineage relaxed POT1 telomeric DNA sequence specificity would be beneficial in an evolutionary context and may have contributed to the survival of this entire plant order.

Evolutionary changes in plant POT1 functions

Interestingly, we failed to detect *in vitro* telomeric DNA binding for eight POT1 proteins from the evolutionarily diverse group of plant organisms analyzed in this study and for six previously characterized POT1 proteins from the Brassicaceae family of plants, which includes *Arabidopsis* (Shakirov et al, 2009). We can not rule out the possibility that some RRL-expressed plant POT1 proteins lack proper post-translational modifications or other requisites for efficient binding to telomeric DNA *in vitro*. However, we note that RRL-expressed POT1 proteins from yeast (Baumann & Cech, 2001) , mammals (Wu et al, 2006), and three different plants (this study) can efficiently bind telomeric DNA under the same conditions. Moreover, several lines of evidence suggest that telomeric DNA binding may not be the major *in vivo* function of POT1 proteins in *Arabidopsis* and, perhaps, in other plants. In striking contrast to yeast and mammalian POT1 proteins, *Arabidopsis* POT1a acts as a positive regulator of telomerase activity and is only enriched at the telomeres in S-phase when telomerase is thought to act (Surovtseva et al, 2007). Moreover, *Arabidopsis* POT1b appears to be a negative regulator of telomerase activity (E. Shakirov, A. Nelson and D. Shippen, in preparation). Recent data indicate that AtPOT1a and AtPOT1b associate directly with *Arabidopsis* telomerase RNA in regions outside the telomere template domain (C. Cifuentes-Rojas

et al., in preparation). While RNA binding is associated with other OB-fold containing proteins, e. g. translation factors (reviewed in Theobald et al, 2003), POT1 proteins have not been previously reported to bind RNA. Thus, the interaction of *Arabidopsis* POT1a and POT1b with telomerase RNA appears to represent a major evolutionary shift in plant POT1 functions from DNA to RNA binding.

Arabidopsis and maize are currently the only plants known to harbor more than one *POT1* orthologue. In both cases, *POT1* genes were likely duplicated around 30 mya, when the lineages leading to *Arabidopsis* and maize experienced independent whole-genome duplication events (Kellogg, 1998; Paterson et al, 2004; Schranz & Mitchell-Olds, 2006). Despite a similar evolutionary timeframe, the fate of these duplicated *POT1* genes appears to be distinct. First, like the two mouse POT1 proteins, which evolved partially non-overlapping functions (Hockemeyer et al, 2006; Palm et al, 2009; Wu et al, 2006), the maize POT1 paralogs share ~ 75% amino acid similarity. In contrast, *Arabidopsis* POT1a and POT1b display much lower sequence conservation, retaining only ~ 50% amino acid similarity overall. Second, while both *Arabidopsis* POT1 proteins bind telomerase RNA instead of telomeric DNA, only one of the maize POT1 proteins, ZmPOT1a, lost the ability to bind telomeric DNA, raising the interesting possibility that ZmPOT1a evolved to bind the maize telomerase RNA. Although further analysis of the telomere complex in maize will be required to test this model, the comparative biochemical analysis of plant POT1 proteins described here underscores the remarkably rapid evolution of the OB-fold nucleic acid binding interface.

VITA

Name:	Xiangyu Song
Address:	Texas A&M University
	Department of Biochemistry and Biophysics
	College Station, TX, 77843-2128
Email Address:	xys@neo.tamu.edu
Education:	Ph.D., Biochemistry, Texas A&M University, 2010
	B.S., Biotechnology, Nankai University, China, 2004