I. Identification and phylogenetic analysis of plant TERT genes, II. Identification of Ribulosebisphosphate carboxylase/oxygenase activase in wheat
by Ruschelle Ann Love

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Plant Sciences
Montana State University
© Copyright by Ruschelle Ann Love (2000)

Abstract:
Telomerase is a ribonucleoprotein complex that adds telomeres to the ends of eukaryotic chromosomes. Two subunits are associated with telomerase, the catalytic subunit, commonly referred to as TERT, and the RNA subunit. Telomerase has been extensively studied in mammalian and yeast systems, but little is known about the roles of telomeres and telomerase in plants or their effect on plant growth and development. This thesis reports the identification of two plant TERT genes, Arabidopsis ihaliana and Oryza sativa (rice). From the multiple sequence alignments it was concluded that, as with previously identified TERTs, plant TERTs contain the conserved reverse transcriptase motifs 1, 2 and A-E as well as the TERT specific motif T. In addition, the alignment of all known TERTs showed seven additional conserved motifs, five upstream of motif T (TEL 1-5) and two downstream of the RT motif E (F and G). Phylogenetic analysis revealed that the TERT catalytic subunit in plants is conserved throughout evolution, and, that plant TERT proteins resemble those of higher eukaryotes (human, mouse and hamster) more closely than to those of lower eukaryotes (yeast or ciliates).

Ribulose bis-phosphate carboxylase activase (Rubisco activase) is an enzyme that catalyzes the activation of Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco). Rubisco activase has been identified in a number of plant species, including barley, rice, maize, Arabidopsis, tobacco, tomato and spinach, but only within the monocot barley have two distinct Rubisco activase genes been identified. All other plants have one gene associated with Rubisco activase. In this thesis, we set out to identify the Rubisco activase gene(s) in an additional monocot, wheat. Screening of a cDNA library resulted in identification of two wheat Rubisco activase genes, WRcaA and WRcaB. The 432 amino acid WRcaB polypeptide was found to be 96% identical to the barley BRcaB polypeptide and the 363 amino acid WRcaA polypeptide was found to be 95% identical to the barley BRcaA polypeptide. Multiple sequence alignment also showed WRcaA and WRcaB contain the same structural features found through studies on other plant species.

I. IDENTIFICATION AND PHYLOGENETIC ANALYSIS OF PLANT TERT GENES

II. IDENTIFICATION OF RIBULOSEBISPHOSPHATE CARBOXYLASE/OXYGENASE ACTIVASE IN WHEAT

by

Ruschelle Ann Love

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Plant Sciences

MONTANA STATE UNIVERSITY
Bozeman, Montana

July, 2000

N378
L9425

APPROVAL

of a thesis submitted by

Ruschelle Ann Love

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style and consistency, and is ready for submission to the College of Graduate Studies.

Dr. Gary A. Strobel _____ (Signature) 7/5/00 Date

Approved for the Department of Plant Sciences

Dr. Norm F. Weeden _____ (Signature) 6/30/00 Date

Approved for the College of Graduate Studies

Dr. Bruce M. McLeod _____ (Signature) 7-7-00 Date

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Montana State University, I agree that the library shall make it available to borrowers under rules of the library.

If I have indicated my intention to copyright this thesis by including a copyright notice, copying is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for permission for extended quotation from or reproduction of this thesis in whole or in parts may be granted only by the copyright holder.

Signature Russell A. Love

Date July 6, 2000

TABLE OF CONTENTS

TABLE OF CONTENTS - Continued

LIST OF TABLES

Page

Table

LIST OF FIGURES

Page

Figure

LIST OF FIGURES - Continued

Page

Figure

ABSTRACT

Telomerase is a ribonucleoprotein complex that adds telomeres to the ends of eukaryotic chromosomes. Two subunits are associated with telomerase, the catalytic subunit, commonly referred to as TERT, and the RNA subunit. Telomerase has been extensively studied in mammalian and yeast systems, but little is known about the roles of telomeres and telomerase in plants or their effect on plant growth and development. This thesis reports the identification of two plant TERT genes, *Arabidopsis thaliana* and *Oryza sativa* (rice). From the multiple sequence alignments it was concluded that, as with previously identified TERTs, plant TERTs contain the conserved reverse transcriptase motifs 1, 2 and A-E as well as the TERT specific motif T. In addition, the alignment of all known TERTs showed seven additional conserved motifs, five upstream of motif T (TEL1-5) and two downstream of the RT motif E (F and G). Phylogenetic analysis revealed that the TERT catalytic subunit in plants is conserved throughout evolution, and, that plant TERT proteins resemble those of higher eukaryotes (human, mouse and hamster) more closely than to those of lower eukaryotes (yeast or ciliates).

Ribulose bis-phosphate carboxylase activase (Rubisco activase) is an enzyme that catalyzes the activation of Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco). Rubisco activase has been identified in a number of plant species, including barley, rice, maize, *Arabidopsis*, tobacco, tomato and spinach, but only within the monocot barley have two distinct Rubisco activase genes been identified. All other plants have one gene associated with Rubisco activase. In this thesis, we set out to identify the Rubisco activase gene(s) in an additional monocot, wheat. Screening of a cDNA library resulted in identification of two wheat Rubisco activase genes, WRcaA and WRcaB. The 432 amino acid WRcaB polypeptide was found to be 96% identical to the barley BRcaB polypeptide and the 363 amino acid WRcaA polypeptide was found to be 95% identical to the barley BRcaA polypeptide. Multiple sequence alignment also showed WRcaA and WRcaB contain the same structural features found through studies on other plant species.

IDENTIFICATION AND PHYLOGENETIC ANALYSIS OF PLANT TELOMERASE GENES

Introduction

Telomeres, or the specialized structures found at the ends of eukaryotic chromosomes, were first characterized by Muller in 1927 (Blackburn, 1995). In *Drosophila*, Muller produced genetic mutations via X-ray irradiation and closely studied the breakage and rejoining of chromosomes. Analysis of the mutations led to the finding that breakage and rejoining (chromosomal rearrangements) never occurred at the terminal end of chromosomes, thus leading to the concept of a "cap" at the end of the chromosome. These discoveries were supported by maize chromosome breakage studies done by Barbara McClintock in 1941 (Blackburn, 1995). McClintock used a breakage-fusion-bridge cycle to study newly formed chromosomes ends. The breakage-fusion-bridge cycle consisted of the movement of two centromeres of a dicentric chromosome to opposite poles of the mitotic spindle, thus forming a bridge between poles. The breaking of the bridge, during anaphase or telophase, produced two monocentric chromosomes with newly formed ends. Throughout these studies, McClintock was able to draw several conclusions concerning broken chromosomes, including the idea that newly formed chromosome ends were "sticky" and would fuse with other newly formed chromosome ends (normal, unbroken, chromosomes were stable and would not fuse to other ends) (Blackburn, 1995).

These initial studies, on chromosome rearrangement, were followed by cytological

studies of chromosome ends, showing many cases in which the ends of chromosomes behaved in special ways, and molecular studies of telomeres (Blackburn, 1995). The molecular analysis of telomeres began with the discovery that the most distal portion of the *Tetrahymena thermophila* chromosomes consisted of tandem arrays of repeated sequence (Blackburn, 1995). Since then, progress has been made in the characterization of telomere structure and function. It was soon discovered, by looking at various eukaryotic organisms, that telomeres, regardless of species, were structurally and functionally similar. All telomeres consist of repeated sequences with a G-rich strand oriented in the 5' to 3' direction towards the chromosome terminus (Blackburn, 1995). Each terminus contains a 3' overhang of two repeat units (Regad, 1994; Richards, 1991; Zentgraf, 1995) which enables telomerase, an RNA-dependent DNA polymerase, to synthesize additional G-rich repeats (Blackburn, 1995). The telomeric DNA usually consists of tandem repeats of 5-8 bp sequence elements. Some common sequence motifs include: TTAGGG (vertebrates); TTTAGGG (angiosperms); TTAGGG (fungi); and $T_{2-4}GGGG$ (ciliates) (Blackburn, 1995). Other organisms have telomeric repeats that are much longer and more highly variable. *Saccharomyces cerevisiae*, for example, has an imperfect telomeric repeat, $T(G)_{2-3}(TG)_{1-6}$ (Zakian, 1996), and *Candida albicans* has a unique 23 bp repeat (Blackburn, 1995). The length of the telomeric tandem arrays for the above species has been shown to vary, ranging from 5-20kb in humans, 100-150kb mice, 200-250bp in *Arabidopsis*, to very short stretches of 18-20bp in *Oxytricha* and *Euplotes* (Zentgraf, 1994 and Richards, 1988).

An additional interesting feature of telomere structure is the presence of bound proteins associated with them. Proteins associated with telomeres have been isolated and

characterized for different systems and are associated with different parts of the telomeric tract, such as the single-stranded 3' overhang or the double-stranded telomeric repeat tracts (Zentgraf, 1995, Shippen, 1998 and Zentgraf, 2000). The first group of telomere-associated proteins interact with the telomeric single-stranded 3' overhang in ciliates. They are assumed to function as molecular chaperones, capping the G-quartets formed by the single-stranded 3' overhang (the G-rich overhanging strands of telomeres that dimerize to form stable complexes in solution via the formation of G-quartet, or cyclic, tetramers). The second group of telomere-binding proteins, double-stranded telomere-binding proteins, has been characterized in human and yeast and are classified as a family of Myb-related (protooncogene) telomeric DNA-binding proteins. Rap1p, one example of a double-stranded telomere-binding protein, is involved in telomeric position effects as well as in telomere length regulation. The telomeric position effect refers to the ability of telomeres to repress the transcription of genes in their vicinity (Stavenhagen, 1998). Plant telomere-binding proteins have also recently been discovered (Zentgraf, 1995). In contrast to human/yeast, they appear to form complexes with both double-stranded and single-stranded telomeric DNA effectively. Thus, the plant telomere-associated proteins appear to fit simultaneously into both categories described above, suggesting that they may have multiple functions.

The functional roles of telomeres have been well established. It is known that telomeres perform at least three basic functions. First, telomeres may maintain genome integrity by forming protective caps at the ends of chromosomes that prevent chromosome end to end fusion, recombination, or exonucleolytic degradation (McEachern, 1996 and Gall, 1995). Second, they may serve a function in nuclear organization, as they seem to be

associated with the nuclear matrix (Zentgraf, 1995). Third, telomeres are needed for the complete replication of chromosomal termini as a solution to the "end replication problem" (discussed below). Because RNA primers are required to initiate DNA strand replication, conventional DNA polymerases cannot completely replicate eukaryotic chromosomes containing 3' terminal extensions (Lingner, 1997). Thus, the enzyme telomerase, a ribonucleoprotein, is needed for replication of chromosomal termini.

Telomerase, first purified from *Tetrahymena* (Ligner, 1997), is an RNA-dependent DNA polymerase containing an RNA and a protein subunit. The RNA subunit provides the template for addition of short sequence repeats to the 3' end of the telomere and was first identified in *Tetrahymena* by Greider and Blackburn (Greider, 1989). To date, a variety of telomerase RNA subunits have been identified from different organisms, including ciliates, yeasts and mammals (reviewed in Nugent, 1998). The telomerase catalytic protein component, also first identified and sequenced in *Tetrahymena*, has now been identified in ciliated protozoans, mammals, yeasts and the plant-*Arabidopsis thaliana*. The protein component serves as the catalyst for telomere repeat synthesis and contains sequence similarity to reverse transcriptase, including the presence of seven reverse transcriptase (RT) motifs: 1,2, and A-E (Bryan, 1999). In addition to containing the RT sequence motifs, a telomerase-specific motif, T, is also present in all telomerases identified to date, but its function has not yet been elucidated. Identification of the eight (T, 1, 2, and A-E) telomerase reverse transcriptase motifs led to the naming of the catalytic subunit as "TERT" (Telomerase Reverse Transcriptase) by Nakamura et al., 1997.

Genes in the RT family encode a wide variety of elements including telomerases,

retroviruses, long terminal repeat (LTR) retrotransposons and group II introns to name a few (Nakamura, 1998). And although they are a diverse group, they all contain RT motifs. The seven canonical RT sequence elements have been found in the catalytic subunit of telomerase, which is not surprising, as telomerase polymerizes DNA using an RNA template, itself the definition of an RT (Nakamura, 1998). Telomerase has been found in evolutionarily diverse organisms, and in all cases contain an RNA template and a conserved TERT subunit, suggesting that telomere maintenance by telomerase is an ancient mechanism.

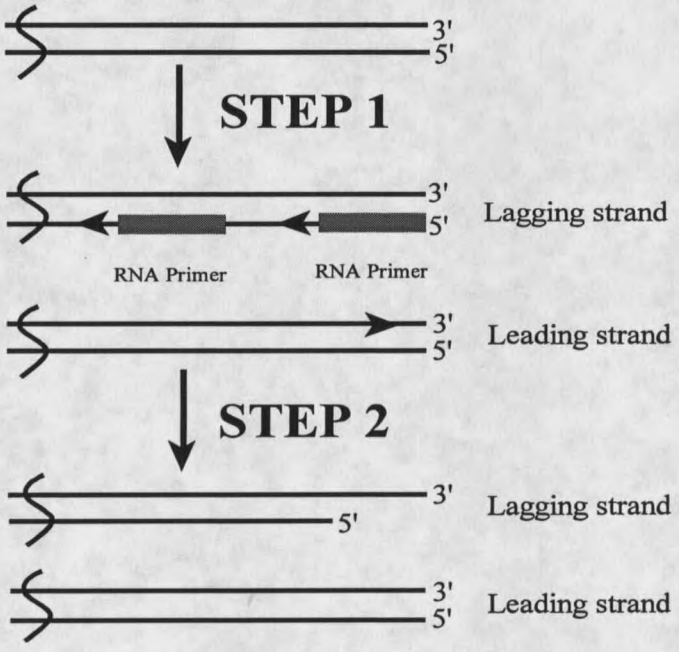

Figure 1-1. End Replication Problem. STEP 1: Denaturation of the template occurs first. The leading strand synthesis results in complete replication of the 5' template sequence. Both lagging and leading strand synthesis are initiated by RNA primers. STEP 2: RNA primers are removed from the lagging strand, resulting in an incompletely replicated end. A 3' overhang at one end of each "new" DNA strand (chromosome) is thus created. Only one end of the chromosome is illustrated. The left-hand curved line indicates the rest of the chromosome.

Chromosome replication poses a problem for eukaryotic cells. Replication by DNA polymerase does not work for replication of chromosome ends, as the 3'-end of chromosomes are not completely copied via this mechanism. Thus, each successive round of conventional DNA replication results in a loss of telomeric sequence as DNA polymerase fails to complete synthesis of the lagging strand (Figure 1-1). To prevent continuous loss of sequence, telomerase has evolved to add sequence onto 3' chromosomal termini. The model of telomere repeat addition by telomerase (Figure 1-2) was proposed by Greider and Blackburn (Greider, 1989) and has since been supported by additional studies. Telomerase binds to the 3' overhang of the chromosome end, with the RNA template base pairing with the most terminal telomeric repeat. The telomere is then elongated by the addition of nucleotides to the 3' terminus. Once the end of the templating domain is reached, a translocation step is necessary to reposition the DNA back to the beginning of the RNA template for another round of nucleotide synthesis (Blackburn, 1995 and Shippen, 1998).

Thus, telomere maintenance is dependent on telomerase and therefore telomerase often determines the future of the cell. In healthy cells, telomerase activity appears to be tightly regulated and maintains proper telomere length. However, in cancer cells or aging cells, telomeres are abnormally long or short. In human somatic cells, telomeres are found to shorten with each successive cell division. Furthermore, in both mammalian and fungal systems (Nakamura, 1997 and Ligner, 1997), severe telomere degradation is implicated in cellular senescence and aging. On the other hand, human immortalized cell lines and cancer cell lines are often found to have higher levels of telomerase, suggesting that unchecked

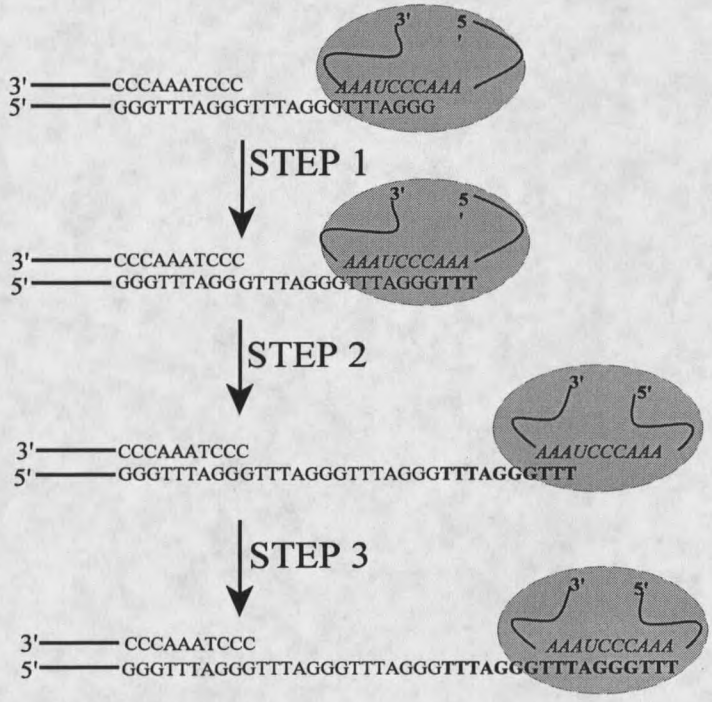

Figure 1-2. Telomerase Elongation Model. STEP 1: The TTTAGGG repeat of the telomere is base paired with the RNA template sequence of telomerase. Elongation occurs as the RNA is copied to the end of the template region. STEP 2: Translocation repositions the telomere sequence and exposes additional template sequence. STEP 3: Another round of template copying produces additional TTTAGGG repeats. New bases are seen in bold and letters in italics indicate RNA template.

addition of telomeres by telomerase may result in immortality of cell lines (Zakian, 1997).

Despite increasing interest in telomeres and telomerase over the past decade, studies of plant telomeres and telomerase have been limited. This is somewhat surprising as the telomere concept was developed in maize (Blackburn, 1995) and the first higher eukaryotic telomere was cloned from *Arabidopsis thaliana* (Richards, 1988).

Current plant telomere and telomerase research includes focus on the developmental

control of telomere length and of telomerase expression in plants. Plant telomeres, as stated above, were first identified in the plant *Arabidopsis thaliana* by Richards and Ausubel (Richards, 1988). The plant telomere sequence appears to be TTTAGGG for many identified plant species, including *Arabidopsis*, barley, tomato, maize, wheat, and *Melandrium album* (white campion) (Shippen, 1998, Riha, 1998). Interestingly, plants of the onion family apparently lack the TTTAGGG telomeric repeats, as neither *in situ* hybridization nor Southern blotting have been successful in detection of this telomeric sequence (Shippen, 1998). Although the telomere repeat itself is conserved within most plants, telomere length varies to a large degree between species as well as between different tissues of the same species. For example, *Arabidopsis* telomeres average 2-4 kb (Richards, 1988) whereas tobacco telomeres extend up to 130 kb (Shippen, 1998). It has also been determined that telomere length varies between tissues of the same plant. For example, plant telomeres are shortened in embryos and inflorescences during differentiation in barley. Moreover, the telomere shortening is believed to be due to an absence of telomerase, as seen in the similar process in humans (Kilian, 1995). The telomeres in barley are especially long during undifferentiated growth in suspension cultures, likely the result of higher levels of telomerase activity (Kilian, 1995).

Telomerase expression is also believed to be developmentally regulated. Telomerase expression was first reported in tobacco using a biochemical telomerase (oligo extension) assay method of detection (Fajkus, 1996). More recently, the PCR-based Telomere Repeat Amplification Protocol (TRAP) assay has been successfully used to detect telomerase activity in a variety of plants including: barley, *Arabidopsis thaliana*, cauliflower, tobacco, soybean

and *M. album* (McKnight, 1997; Fajkus, 1998; Shippen, 1998; Riha, 1998). Utilizing such methods, expression of plant telomerase activity in various tissues has been explored (Killian, 1998, McKnight, 1998, Riha, 1998,and Heller, 1996). Telomerase activity was easily detected in organs containing rapidly proliferating meristems, such as germinating seedlings, root tips, embryos, anthers, carpels, flowers and the floral buds. Telomerase activity, however, was either low or absent in such tissues as the shoot apex, leaf, or stem. The above observations suggest that telomere length in plants, as in human and yeasts, is maintained by a telomerase-mediated mechanism.

Despite recent progress in investigating plant telomeres and telomerase, much work must still be done before a solid understanding of the role of telomerase and its controls are achieved. Identification and study of both the TERT and RNA components of plant telomerase will be an essential step towards this goal. In this thesis, with the goal of identifying plant TERT genes, we utilized the higher plant preliminary sequence data made available by the Stanford Sequencing Center (Stanford, California) to identify putative plant *TERT* gene fragments. Using sequence information from the putative *Oryza sativa* (rice) and *Arabidopsis thaliana TERT* gene fragments, primers were made with the goal of identifying full length *TERT* genes by either PCR or by screening a cDNA library. The deduced polypeptide sequence from resulting genes were subject to phylogenetic analysis, against known TERT sequences, in an attempt to reinforce the findings that the TERT catalytic subunit is a conserved component of telomerase. Surprisingly, it was found that *Arabidopsis* TERT and *O. sativa* TERT more closely resemble the TERT proteins of higher eukaryotes than TERTs from lower eukaryotes.

Materials and Methods

Database Identification of Putative *Arabidopsis thaliana* TERT

A putative *Arabidopsis thaliana TERT* gene was found in the GenBank *Arabidopsis* (incomplete genome) database via the "TBLASTN" algorithm using a 76 amino acid fragment of the known human TERT (hTERT) amino acid sequence (motifs B, C and D) as query. Parameters were set at default and the hTERT query pulled up multiple hits including a 625 bp *Arabidopsis thaliana* genomic DNA fragment, accession number B27802. The deduced polypeptide sequence of the unidentified *Arabidopsis* gene was 41% identical to hTERT with 65% similarity. This protein had regions which showed substantial similarity to reverse transcriptase RT motifs C, D and E of hTERT, and was therefore considered a putative TERT.

Collection of *Arabidopsis thaliana* DNA

RNAse treated *Arabidopsis thaliana* DNA (Colombia ecotype) was obtained from B. Sharrock, Montana State University. The DNA was concentrated by ethanol precipitation and then resuspended in TE buffer (10mM Tris pH8, 1mM EDTA). The concentration was quantitated via the Beckman Du-50 spectrophotometer at A_{260}, and DNA was inspected by running a 1% agarose gel to ensure it was not degraded.

Identification of *Arabidopsis thaliana* TERT (*At*TERT)

Exact and degenerate primers were constructed using sequence information from the putative *Arabidopsis TERT* fragment found in GenBank and from *Saccharomyces cerevisiae,*

Schizosaccharomyces pombe, Tetrahymena thermophila, and *Oxytricha trifallax TERT* sequences. The primers used for identification of TERT from genomic *Arabidopsis* DNA were: Arab C1 (rev)-exact, 5'-AGA CAC AAA ATG TAG TCA TC-3'; Arab C2 (rev)-exact, 5'- GTC ATCA ATA AAT CTC AGT AA-3'; Arab C1 (rev)-deg., 5-GAY GAY TAY ATH TTY GTN TCN-3'; Arab C2 (rev)-deg., 5'-TYG NYG NSA NTT MTT VAT SAY-3'; B (for)-yeast combo, 5'-MRA RAA GWT GGT MTY YYT CAR GG-3'; B (rev)-yeast combo, 5- NSW NCC YTG NGG DAT NCC-3'; and T (for)-yeast combo, 5'-RWC STT TYT TYT AYD KCA CBG A-3'. IUB group codes, seen above, to identify degenerate sequences are as follows: R= A+G, Y= C+T, M= A+C, K= G+T, S= G+C, W= A+T, H= A+T+C, B= G+T+C, D= G+A+T, V= G+A+C, and N= A+G+C+T.

The first round of PCR was performed in a Perkin Elmer DNA thermal cycler 480 using 50µl reactions containing the following: 2µg *Arabidopsis* genomic DNA, 1 unit of *taq* polymerase (Promega), 0.1 mM dNTPs, 250 µM of each primer and buffer C (provided by the manufacturer) with [Mg^{++}] at 1.5mM. A 5 minute denaturation step at 94° was followed by 30 cycles of 94°C for 1 minute, 50°C for 2 minutes, and 72°C for 1 1/2 minutes. The appropriate number of cycles was determined experimentally by quantitating the amount of PCR product in a cycle-dependent manner, comparing product yield after 21, 23, 25, 27, 29, 31, 33, and 35 cycles. Primer sets used, in separate attempts to identify the TERT gene, for the first round of PCR were: Arab C1 (rev)-exact and T (for)-yeast combo, Arab C1 (rev)-exact alone, T (for)-yeast combo alone, Arab C1 (rev)-deg. and T (for)-yeast combo, and Arab C1 (rev)-deg. The approximately 1600 bp band, produced by Arab C1 (rev)-exact primer alone, was gel purified. The band was cut from the gel, purified by the pulp-spin

method (spin DNA through eppendorf tube containing paper pulp in TE buffer for 10 minutes at 14 K and collect supernatant), ethanol precipitated and resuspended in 20 μl of water. The gel purified DNA (1μl) was electrophoresed on a 1% agarose gel for quantitation and to determine purity, and was then used as template in the second round of PCR.

The second round of PCR, nested PCR, used 0.5ng of the ~1600 bp Arab C1 (rev) alone fragment as template. All other PCR conditions were kept the same as first round PCR. Nested primers sets used in the nested PCR reactions were as follows: Arab C1 (rev)-exact and T (for)-yeast combo, Arab C1 (rev)-exact alone, T (for)-yeast combo alone, Arab C2 (rev)-exact and T (for)-yeast combo, Arab C2 (rev)-exact alone, Arab C2 (rev)-deg. and T (for)-yeast combo, Arab C2 (rev)-deg. alone, Arab C2 (rev)-exact and B (for)- yeast combo, B (for)-yeast combo alone, B (rev)-yeast combo and T (for)-yeast combo, and B (rev)-yeast combo alone. From these primer sets, various PCR product sizes were separated on a 1.5% agarose gel stained with 0.5μg/ml ethidium bromide and visualized under UV_{254}. Resulting bands of the correct size, estimated from the size of other telomerases, were seen for Arab C2 (rev)-exact and T (for)-yeast combo (1002 bp), Arab C2 (rev)-exact and B (for)-yeast combo (153 bp), and B (rev)-yeast combo and T (for)-yeast combo (849 bp) (Figure 1-3). The bands were cloned into the pGem-T cloning vector using the pGem-T Vector System II (Promega), as per the manufacturers directions, and sequenced. Each PCR product (2μl) was ligated into the pGem-T vector and transformed into *E. coli* JM109 competent cells (Promega). The clones were screened for plasmids with inserts of the desired size and the plasmids containing correct sized inserts were used directly in sequencing reactions.

Sequencing of Clones

Sequencing reactions were set up for the ABI PRISM 377 as described by the ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin Elmer). For each 20μl reaction, 0.2μg of DNA, 3.2pmole primer, and 4μl of BIGDYE mix (A-Dye Terminator, C-Dye Terminator, G-Dye Terminator, T-Dye, Terminator, dITP, dATP, dCTP, dTTP, Tris-HCl (pH 9.0), $MgCl_2$ thermal stable pyrophosphatase, and AmpliTaq DNA Polymerase) were mixed and placed in the GeneAmp PCR system 9600. Thermal cycling was as follows: 96°C for 15 seconds, 50°C for 10 seconds, and 60°C for 4 minutes. This cycle was repeated 25 times and PCR products were precipitated with 75% isopropanol and pellets were washed with 70% ethanol. The resulting pellet was dried and sequenced by the Montana State University center for sequencing.

The clones were sequenced using universal primers SP6 (5'-ATT TAG GTG ACA CTA TAG-3') and T7 (5'-TAA TAC GAC TCA CTA TAG GG-3'). Internal primers were constructed to verify the sequence in both directions. The primers were exact sequences taken from SP6 and T7 sequencing results and included: ArabNB1 (rev) 5'- TAT CCA AAC CAT TAG ATC-3', ArabNC1 (rev) 5'-CAA AGT CCT CTC GAG ATG-3', ArabNT1 (for) 5'-TGG GAA AGA TTA ATA AGC-3', and ArabNT2 (for) 5'-GAA CCA GAT GTT CTT GG-3'.

Database Identification of Putative *Oryza sativa* (Rice) TERT

A segment of DNA containing a potential *Oryza sativa* TERT was identified by searching the *Arabidopsis thaliana* database (Stanford University Home Page) via the

"TBLASTN" algorithm. The search utilized a 28 amino acid segment of the *Arabidopsis thaliana* TERT protein sequence in the region identified as the "C motif" as query against the higher plant sequence database with the expect parameter set at 100. The second match, with a match score of 74, was accession number AQ510589 from the *Oryza sativa* sequencing project at Clemson University. AQ510589 was a 531 base pair genomic fragment. Sequence alignment of this deduced polypeptide sequence to the TERT protein sequence of *Arabidopsis thaliana* identified multiple regions of sequence similarity, including the canonical reverse transcriptase motifs C, D, and E.

Collection of *Oryza sativa* (Rice) DNA

Clemson University provided the *Oryza sativa* Rice BAC clone, in the pBeloBACII vector, (catalog # nbxb0095N19f) as a stab culture in *E. Coli* strain DH10B. One liter of culture was grown in LB (Lennox L Broth Base) at 37°C with overnight shaking. Cells were harvested and DNA extracted via the Wizard Plus Maxipreps DNA purification system (Promega) using the vacuum filtration protocol, according to the manufacturers directions. Recovered DNA was resuspended in 0.75ml TE buffer and the final yield was 20μg of DNA. DNA yield was determined by running a 1% agarose gel with 0.5μg/ml ethidium bromide and visualizing DNA concentration by ultraviolet light. The fragment was sequenced using universal SP6 and T7 primers to ensure the correct sequence. Cultures were prepared for long term storage by: 1) directly storing recovered DNA from the above purification at -70°C and 2) preparing a glycerol stab containing culture in LB and 8% glycerol and storing at -70°C.

Oryza sativa (California ecotype) genomic DNA, 1μg/μl, was obtained from M. Giroux, Montana State University. Concentration and quality of the DNA was checked on a 1% agarose gel.

Confirmation of *Oryza sativa* (Rice) TERT Fragment Sequence

The BAC genomic fragment, nbxb0095N19f (GenBank accession AQ510589), containing the Rice TERT was amplified by PCR and sequenced to verify the sequence obtained from the database. The purified DNA, from Wizard Plus Maxipreps DNA purification system (described above), was used as template in a PCR reaction utilizing two end primers, made from the database sequence. One internal primer, to amplify just the 270 bp TERT fragment, was also made. Primers used for PCR were: Rice EP-2 (for) 5'-CCT GAA TAT TTG TTA ATG AGG-3, Rice EP-3 (rev) 5'-GGA AAC ACA TCG TCC AAG TCA AT-3', and Rice ER-1 (rev) 5'-GTC ATA CCT CGT ATA ATC AGC-3'. Internal primer Rice ER-1 (rev) and primer EP-2 (for) amplified the 270 bp TERT fragment whereas the other 261 bp of the 531 bp fragment did not align with known TERT sequences. The PCR reaction, to obtain the fragment of interest from the BAC DNA, was done using 2µg BAC DNA, 1 unit of *Tfl* polymerase, 0.1mM dNTPs, 250µM primers, and a 1.5mM concentration of $MgSO_4$ in *Tfl* buffer. A 5 minute denaturation step at 94°C was followed by 40 cycles of 94°C for 1 minute, 52°C for 1 minute, and 68°C for 2 minutes. The resulting fragments were purified using the QIAquick PCR product purification protocol (Qiagen) and the products were used as template in the sequencing reactions. The sequencing was done, as described above, using Rice EP-2 (for), Rice EP-3 (rev) and Rice ER-1 (rev) as primers.