Structure, Function and Evolution of Filamentous Fungal Telomerase RNA

by

Xiaodong Qi

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2011 by the
Graduate Supervisory Committee:

Julian Chen, Chair
John Chaput
Giovanna Ghirlanda

ARIZONA STATE UNIVERSITY

May 2011

ABSTRACT

Telomerase ribonucleoprotein is a unique reverse transcriptase that adds telomeric DNA repeats to chromosome ends. Telomerase RNA (TER) is extremely divergent in size, sequence and has to date only been identified in vertebrate, yeast, ciliate and plant species. Herein, the identification and characterization of TERs from an evolutionarily distinct group, filamentous fungi, is presented. Based on phylogenetic analysis of 69 TER sequences and mutagenesis analysis of *in vitro* reconstituted *Neurospora* telomerase, we discovered a conserved functional core in filamentous fungal TERs sharing homologous structural features with vertebrate TERs. This core contains the template-pseudoknot and P6/P6.1 domains, essential for enzymatic activity, which retain function *in trans*. The *in vitro* reconstituted *Neurospora* telomerase is highly processive, synthesizing canonical TTAGGG repeats. Similar to *Schizosaccharomycetes pombe*, filamentous fungal TERs utilize the spliceosomal splicing machinery for 3' processing. *Neurospora* telomerase, while associating with the Est1 protein *in vivo*, does not bind homologous Ku or Sm proteins found in both budding and fission yeast telomerase holoenzyme, suggesting a unique biogenesis pathway. The development of *Neurospora* as a model organism to study telomeres and telomerase may shed light upon the evolution of the canonical TTAGGG telomeric repeat and telomerase processivity within fungal species.

DEDICATION

This thesis is dedicated to my most beloved ones:

To my parents for their tremendous supports during my entire life. Without their wisdom and guidance, I would not reach my dream.

And to my girlfriend for her unconditional supports and love.

ACKNOWLEDGMENTS

I give my sincere appreciations to all people who have offered supports through my entire graduate study.

Dr. Julian J.L. Chen, my PhD advisor and mentor, introduced me into the wonderful scientific research and offered tremendous guidance, encouragement and supports during my whole graduate life.

My graduate and oral committee members, Dr. John Chaput, Dr. Giovanna Ghirlanda, Dr. Hao Yan and Dr. Anne Jones, kindly supported me in my research through years.

I really enjoyed the graduate life with all former and current fellow graduate students in Chen Laboratory: Mingyi Xie, Chris Bley, Lena Li, Sandhya Tadepalli, Tracy Niday, Xiaowei Liu, Josh Podlevsky, Andrew Brown and Lina Franco. I sincerely appreciate Josh Podlevsky for proofreading of my thesis.

TABLE OF CONTENTS

Page

LIST OF TABLES vi

LIST OF FIGURES vii

LIST OF SYMBOLS AND ABBREVIATIONS x

CHAPTER

1 INTRODUCTION 1

1.1 Discovery of telomeres and telomerase enzyme 1

1.2 Telomerase RNA evolution and RNP biogenesis 4

1.3 Projects 9

1.4 References 12

2 IDENTIFICATION OF NEUROSPORA CRASSA TELOMERASE RNA 16

2.1 Abstract 17

2.2 Introduction 18

2.3 Materials and Methods 20

2.4 Results 31

2.5 Discussion 46

2.6 References 48

3 STRUCTURE AND EVOLUTION OF FILAMENTOUS FUNGAL TELOMERASE RNAS 51

3.1 Abstract 52

3.2 Introduction 53

CHAPTER Page

3.3 Materials and Methods 55

3.4 Results 60

3.5 Discussion 81

3.6 References 86

4 FUNCTIONAL STUDY OF N. CRASSA TELOMERASE RNA ... 92

4.1 Abstract 93

4.2 Introduction 94

4.3 Materials and Methods 96

4.4 Results 98

4.5 Discussion 112

4.6 References 113

5 CONCLUSION 115

REFERENCES 117

APPENDIX

A NEUROSPORA CRASSA STRAINS USED IN THIS STUDY 126

B SUPPLEMENTAL INFORMATION FOR CHAPTER 2 128

C SUPPLEMENTAL INFORMATION FOR CHAPTER 3 135

D SUPPLEMENTAL INFORMATION FOR CHAPTER 4 140

E IDENTIFICATION OF A PUTATIVE CAB BOX BINDING PROTEIN NOPP140 142

F CO-AUTHOR APPROVAL 176

LIST OF TABLES

Table Page

2.1 *N. crassa* telomerase RNA candidates with more than 20 reads 39

3.1 List of all filamentous fungal TER1s identified in this study 68

A.1 List of *Neurospora crassa* strains utilized in this thesis 127

S2.1 Comparison of telomeric repeats from three groups 129

S2.2 Telomere length of *N. crassa* chromosome VR 130

E.1 T7 phage titer from plaque assay 159

E.2 Blast results for each DNA band 161

LIST OF FIGURES

Figure Page

1.1 Secondary structure comparison of three major telomerase RNAs 9

2.1 Schematic of northern blot capillary transferring 26

2.2 Construction of an *N. crassa* strain expressing ncTERT-3xFLAG ... 32

2.3 *N. crassa* transformants screening 33

2.4 *N. crassa* telomerase holoenzyme purification and ncTER candidates search procedure 35

2.5 Purification of *N. crassa* telomerase holoenzyme through gel filtration chromatography 36

2.6 Telomerase TRAP activity assay after ncTERT-3xFLAG IP 37

2.7 *N. crassa ter1* gene analysis 41

2.8 Schematic of three ncTER1 forms and northern blot analysis 43

2.9 Expression ncTER1 template mutants in *N. crassa* generates telomere DNA mutations 45

3.1 Electro-transferring of RNA onto Hybond XL membrane using Owl VEP-3 electroblotting system 60

3.2 Phylogenetic tree of the phylum Ascomycota 62

3.3 Sequence alignment of filamentous fungal TER1 splicing signal 65

3.4 Phylogenetic comparison of template-pseudoknot region 71

3.5 Phylogenetic comparison of P6/6.1 73

3.6 Secondary structure of ncTER1 conserved regions 74

Figure Page

3.7 Eurotiomycetes TER1 template sequence analysis 76

3.8 Telomerase TRAP assay post anti-FLAG IP 78

3.9 Northern blot and western blot after SmD3, Ku80 and Est1 immunoprecipitation 79

3.10 Anti-TMG immunoprecipitation from *N. crassa* total RNA..... 81

4.1 Schematic of telomerase translocation 95

4.2 Human and *N. crassa* telomerase activity assay reconstituted in RRL 100

4.3 *N. crassa* telomerase activities with ncTER1 template mutants 102

4.4 Telomerase activity assay with truncated ncTER1 fragments 105

4.5 Secondary of ncTER1 template-pseudoknot and P6/6.1 107

4.6 Telomerase activity assay with ncTER1 truncated fragments 108

4.7 Telomerase activity assay with ncTER1 pseudoknot hairpin truncations 110

5.1 Comparison of four major telomerase RNAs 116

S2.1 Gel filtration analysis of RNAs for Solexa sequencing 131

S2.2 ncTER1 3' RACE 132

S2.3 ncTER1 full length sequence 133

S2.4 Southern blot of ncTER1 template mutant strains 134

S3.1 *Aspergillus nidulans* TER1 full length sequence 136

S3.2 Sequence alignment of P6/6.1 from 69 filamentous fungal TER1s 137

Figure Page

S3.3 Secondary structure of *Aspergillus nidulans* TER1 conserved regions 139

S4.1 Predicted secondary structure of *M. oryzae* pseudoknot 141

E.1 Schematic of T7 phage plaque assay 147

E.2 Secondary structure of hTER and CR7, CR7m 154

E.3 RNA stability test in T7 phage lysate 155

E.4 CR7 and CR7m RNA secondary structure test..... 157

E.5 Schematic of T7 phage display selection cycle 158

E.6 PCR analysis for phage display with human breast tumor cDNA library 160

E.7 PCR analysis for phage display with human colon tumor cDNA library 161

E.8 Schematic of Nopp140 cDNA inserts in phage libraries..... 162

E.9 Purification of Nopp140 from *E. coli* 164

E.10 Purified Nopp140-C34 protein from *E. coli* 165

E.11 Phosphorylation of Nopp140 by Casein Kinase II 166

E.12 Gel mobility shift assay of CR7 and CR7m with purified Nopp140 protein 168

E.13 UV crosslinking of Nopp140 and RNAs 169

E.14 T7 phage pull down assay 170

E.15 Nopp140 and dyskerin protein association *in vivo*..... 172

LIST OF SYMBOLS AND ABBREVIATIONS

Symbol/Abbreviation

ATP ..... adenosine triphosphate

bp ..... base pair(s)

BSA ..... bovine serum albumin

^{o}C ..... degree celcius

cDNA ..... complementary DNA

CKII ..... casein kinase II

DNA ..... deoxyribonucleic acid

dNTP ..... deoxynucleotide triphosphate

DTT ..... dithiothreitol

EDTA ..... ethylene-diamine-tetra-acetic acid

FGSC ..... fungal genetics stock center

hph ..... hygromycin B phosphotransferase

IP ..... immunoprecipitation

IPTG isopropyl-β-D-1-thiogalactopyranoside

kb ..... kilo base(s)

l.c. ..... loading control

MCS ..... multiple cloning sites

mRNA ..... messenger RNA

ncTER1 ..... *Neurospora crassa* telomerase RNA

ncTERT ..... *Neurospora crassa* telomerase reverse transcriptase

NHEJ ..... non homologous end joining

Symbol/Abbreviation

O.D. ..... optical density

P6 .....pair region 6

PAGE ..... polyacrylamide gel electrophoresis

PCR ..... polymerase chain reaction

pg ..... pico grams

PK ..... pseudoknot

PNK .....polynucleotide kinase

PVDF .....polyvinylidene fluoride

RACE .....rappid amplification of cDNA end

rDNA .....ribosome DNA

RIP ..... repeat induced point mutation

rRNA ..... ribosome RNA

RNA .....ribonucleotide

RNP ..... ribonucleoprotein

rpm ..... revolutions per minute

RRL .....rabbit reticulocyte lysate

RT .....reverse transcriptase

RT-PCR ..... reverse transcription PCR

ScaRNA .....small cajal body RNA

SDS ..... sodium dodecyl sulfate

SGS .....second-generation sequencing

snRNA .....small nuclear RNA

Symbol/Abbreviation

spTrt1 .....*Schizosaccharomycetes pombe* telomerase reverse transcriptase

TAE ..... Tris, Acetic Acid, EDTA

TBE .....Tris, Boric Acid, EDTA

TER .....telomerase RNA

TERT ..... telomerase reverse transcriptase

TLC1 .....*Saccharomycetes cerevisiae* telomerase RNA

TMG .....trimethylguanosine

tRNA ..... transfer RNA

TWJ .....three way junction

UTR .....untranslated region

UV .....ultra violet

v/v ..... volume to volume ratio

w/v ..... weight to volume ratio

Chapter 1

INTRODUCTION

1.1 Discovery of telomeres and telomerase enzyme

All life on earth evolved from a common ancestor, eventually splitting into the three major domains of life: Bacteria, Archaea and Eukarya. Bacteria and Archaea contain circular chromosomes while linear chromosomes are almost exclusively found within eukaryotic cells. Telomeres were coined by Hermann J. Muller in 1938 as a protecting element at the ends of the linear chromosomes. Muller discovered that within the fruitfly, *Drosophila*, chromosome breakages generated by X-ray irradiation would only fuse to another broken chromosome but not to the native end of the chromosomes (Muller, 1938). Meanwhile, Barbara McClintock was working with maize chromosomes, and noted a similar phenomenon, indicating that the telomeres delineate natural chromosomal ends from induced double stranded breaks for maintaining the integrity of the linear chromosomes (McClintock, 1939, 1941). After the discovery of DNA structure and DNA polymerases, the famous end-replication problem was proposed by James Watson in 1972, describing the dilemma cells face when replicating linear chromosomes. DNA polymerases are structured to synthesize nucleotide addition in a 5’ to 3’ direction from an RNA primer, necessitating a special mechanism infilling the 5’ region vacated by removal of the RNA primer (Watson, 1972). A mechanism for the telomere length maintenance was proposed and shortly afterward, the discovery of telomerase enzyme.

Telomeric DNA was first cloned by Elizabeth Blackburn and Joseph Gall from rDNA mini-chromosomes of a ciliated protozoa, *Tetrahymena thermophila* (Blackburn and Gall, 1978). The terminal sequences of rDNA molecules were found to contain 20 to 70 tandem hexa-nucleotide repeats, CCCCAA. Soon, telomeric DNA from *Tetrahymena* other chromosomes termini was also identified and found to contain the same telomere sequences as rDNA (Yao et al., 1981). Expanding on this, Blackburn and Szostak generated artificial linear chromosomes in the budding yeast, *Saccharomycetes cerevisiae*. Linear plasmids were generally unstable in yeast cells. However they determined that a linear yeast plasmid with *Tetrahymena* telomere tandem repeat sequences capping the termini were stably replicated within the yeast cells (Szostak and Blackburn, 1982). Furthermore, it was also observed that these linear yeast plasmids, capped by transgenic *Tetrahymena* telomere sequences, had the telomeric DNA maintained and extended within the yeast cells, with yeast telomere sequences added to the termini of linear yeast plasmids (Shampay et al., 1984; Szostak and Blackburn, 1982). Considering the differences in the telomere repeat sequences between *T. thermophila* and *S. cerevisiae*, it is unlikely for recombination to occur for telomere lengthening. Thus, Blackburn and Szostak proposed that an unknown enzyme is specifically responsible for telomere DNA addition. In 1985, this enzyme activity was first observed in *T. thermophila* cell lysate by Carol Greider and Elizabeth Blackburn that a synthetic telomere DNA primer was extended by 6 nucleotide repeats (Greider and Blackburn, 1985). This enzyme was originally named as telomere terminal transferase and was later renamed to telomerase.

Similar enzyme activity was later determined in human cancer cell lysate which incorporates human telomeric DNA repeats, TTAGGG, onto chromosome termini (Morin, 1989).

RNA dependent DNA polymerase (reverse transcriptase) was first discovered in 1970 by two independent research groups (Baltimore, 1970; Temin and Mizutani, 1970). At that time, no one would consider that a unique reverse transcriptase could be responsible for telomere length maintenance. However, soon after the discovery of telomerase in *Tetrahymena thermophila*, Greider examined telomerase activity under DNase or RNase treatment, concluding that telomerase activity was RNA dependent, implying the telomerase enzyme is a ribonucleoprotein complex (Greider and Blackburn, 1987). RNA sequencing from purified *Tetrahymena* telomerase enzyme resulted in the discovery of telomerase RNA (TER) which harbors a putative telomeric template sequence, CAACCCCAA, essential for telomerase activity (Greider and Blackburn, 1989). Telomerase was thus proposed to function as a reverse transcriptase, with the intrinsic RNA component serving as the template. In 1990, Shippen and Blackburn experimented with various DNA primers, complementary to the putative RNA template flanking region of a ciliate *Euplotes crassus* telomerase RNA, to provide functional evidence for the presence of the RNA template (Shippen-Lentz and Blackburn, 1990). Later in the same year, the Blackburn group provided *in vivo* support of the RNA template hypothesis (Yu et al., 1990). The over-expression of *Tetrahymena thermophila* telomerase RNA template mutants in *Tetrahymena* cells resulted in a dominant negative phenotype. The

incorporation of mutant telomeric DNA sequences encoded by the mutated template sequences resulted in altered telomere sequences as well as cellular senescence. This added significantly credence in support of telomerase functioning as a unique reverse transcriptase, with an intrinsic RNA moiety as template. Additional telomerase RNAs were identified in evolutionarily distinct groups of species including yeast and human (Feng et al., 1995; Singer and Gottschling, 1994). Within each identified telomerase RNA, a template sequence complementary to the species telomeric DNA sequences was identified, further supporting the telomerase RNA as the template for reverse transcription. Further evidence was presented with the identification of the protein component, the telomerase reverse transcriptase (TERT), from ciliate, yeast and human containing a conserved reverse transcriptase (RT) domain (Bryan et al., 1998; Lingner et al., 1997b; Meyerson et al., 1997). For their initial and substantial contribution to the discovery of how chromosomes are protected by telomeres and telomerase, Elizabeth H. Blackburn, Carol W. Greider and Jack W. Szostak were awarded the Nobel Prize in Physiology or Medicine in 2009.

1.2 Telomerase RNA evolution and RNP biogenesis

Since the discovery of *Tetrahymena* TER, tremendous efforts have been expended in the identification of TER from other evolutionarily distinct and divergent groups of species as well as determining conserved secondary structure, regions and domains necessary for ribonucleoprotein functionality. The identification of *Tetrahymena thermophila* TER brought about the opportunity to use this sequence for hybridization studies for identifying additional ciliate TER

sequences (Romero and Blackburn, 1991), due to their sequence similarity. The ciliate TER secondary structure model, derived from a phylogenetic comparative analysis, was first presented from seven tetrahymenine ciliates (Romero and Blackburn, 1991). It has been shown that although those ciliate TER sequences diverged quite rapidly, the secondary structure has remained strikingly conserved. Therefore, phylogenetic comparison is a powerful tool for determination of TER secondary structure.

The budding yeast *S. cerevisiae* telomerase RNA (TLC1) was identified in 1994 from a cDNA library screening (Singer and Gottschling, 1994). The screen would identify genes, whereby over-xpression suppresses telomeric silencing. The TLC1 gene was then identified which suppressed the telomeric silencing and shortened the telomeric lengths. The counterintuitive result of over-expression of a telomerase component inducing telomere erosion was explained as high levels of TLC1 titrates away the yeast telomerase holoenzyme essential component, Ku heterodimer (Evans et al., 1998). The length of the TLC1 RNA at approximately 1300 nt is significantly larger than the ciliate TER at approximately 150 to 200 nt, suggesting a complicated and divergent TER evolution.

The human telomerase RNA (hTER) was cloned in 1995 from the comparison of cDNA libraries derived from telomerase positive and negative cell lines. The difference between the two libraries narrowed the number of transcribed genes considerably and RNA candidates were analyzed to identify hTER. This 451nt telomerase RNA bore no sequence similarity to either ciliate or yeast TER.

Telomerase RNAs identified from three evolutionarily divergent groups of species, vertebrate, ciliate, and yeast, exhibit striking variation. Additional TER sequences from within the same lineage, therefore, are required for predicting and determining TER secondary structure. The vertebrate TER secondary structure was determined in 2000, from cloning 35 vertebrate species and phylogenic analysis of the TER sequences (Chen et al., 2000). With length greater than 1000 nt, the budding yeast TER secondary structure was predicted in 2004 (Dandjinou et al., 2004), derived from phylogenetic comparison of closely related yeast species. The structure of TER from three groups contains two conserved core structures, a template-pseudoknot domain and a three way junction structure (Brown et al., 2007; Lin et al., 2004).

Telomerase RNA template typically contains a sequence complementary to 1.5 telomeric DNA repeats. The complete repeat at the 5' end region of the RNA is reverse transcribed to generate telomeric DNA, while the remaining 3' region of the template serves to align and base-pair with the telomeric DNA. In ciliate and vertebrate, telomerase synthesizes numerous telomeric DNA repeats before complete dissociation from a telomeric primer. Upon reaching the 5' end of TER template sequence, telomerase must undergo a translocation step to re-align the telomeric primer to the alignment region of template and regenerate the template for synthesis of a second repeat. The re-alignment of the template RNA with the telomeric primer must occur without complete dissociating of the enzyme from the telomeric primer (Greider and Blackburn, 1989). This unique feature of telomerase, called repeat addition processivity, differentiates it from other reverse

transcriptases which synthesize a long DNA sequence continuously until reaching the RNA 5' end. Yeast TER typically has a long template sequence responsible for synthesizing irregular telomeric DNA sequences. Repeat addition processivity, however, is not observed in most yeast species, since only one or two telomeric DNA repeats are synthesized before complete dissociation of the enzyme from a given primer (Cohn and Blackburn, 1995).

The pseudoknot and three way junction (P6/6.1) domains are responsible for TERT binding and telomerase enzymatic activity in vertebrate (Mitchell and Collins, 2000). Similarly, pseudoknot and stem loop IV are important for both *in vitro* and *in vivo* telomerase activity in ciliate. However, the three way junction domain in yeast is not essential for either *in vitro* or *in vivo* telomerase enzymatic activity (Zappulla et al., 2005).

Beyond these two conserved domains, additional regions of the telomerase RNA, not required for *in vitro* telomerase activity, are important in telomerase RNP biogenesis. Telomerases from different groups acquired completely divergent telomerase RNP biogenesis pathways during the evolution. Ciliate telomerase RNA is an RNA polymerase III transcript (Yu et al., 1990), whereas yeast and vertebrate telomerase RNAs are RNA polymerase II transcripts (Chen et al., 2000; Hinkley et al., 1998; Seto et al., 1999). A conserved ScaRNA domain was identified at the 3' end of vertebrate TER (Jady et al., 2004), which is essential for telomerase localization and function *in vivo*, allowing TER to share the same RNP biogenesis pathway as other ScaRNAs. In budding yeast, Est1 and Ku heterodimer proteins associate with telomerase RNA, to recruit the telomerase

to the telomeres. A conserved Sm binding site was identified at the 3' end of all yeast TERs, to associate with Sm protein complexes. The underlying cause for the telomerase RNP acquiring such divergence in biogenesis during evolution remains elusive.

Telomerase RNAs recently have been identified in the fission yeast *Schizosaccharomycetes pombe* (Leonardi et al., 2008; Webb and Zakian, 2008), and the higher plant *Arabidopsis thaliana* (Cifuentes-Rojas et al., 2011) from completely independent approaches. Additional RNP biogenesis variations have been determined. Thus, the identification of telomerase RNA from novel evolutionarily distinct groups of species will provide important insights into RNA and RNP evolution. A universal approach for novel telomerase RNA identification is necessary to expand the number and diversity of telomerase RNAs for evolutionary study.

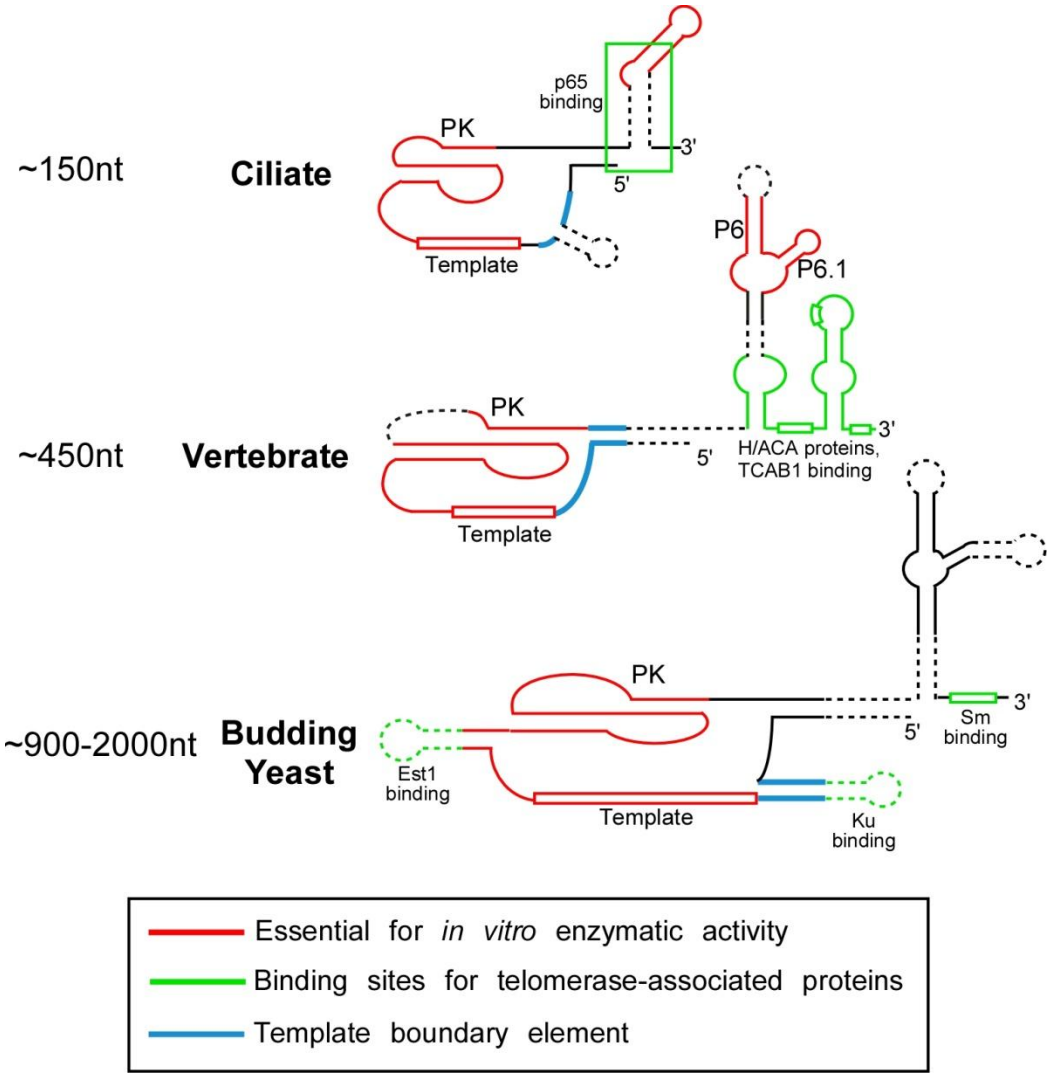


Figure 1.1 Secondary structure comparison of three major telomerase RNAs

The red boxes represent the template region, whereas dashed lines represent variable sequences within the same phylogenetic group. PK indicates the pseudoknot domain. The structures essential for *in vitro* enzymatic activity were labeled in red. TER template boundary elements were indicated with blue color. Telomerase associated proteins binding sites were highlighted with green color, with proteins indicated adjacent.

1.3 Projects

Telomeres and telomerase research has been exhaustively studied within a few yeast species. Yeast is an evolutionarily early branching group which belongs

to the phylum Ascomycota. The later branching group Pezizomycotina is the largest subphylum of Ascomycota, containing most filamentous fungi species (Spatafora et al., 2006). Unlike yeast, most filamentous fungi contain vertebrate type telomeric DNA repeat TTAGGG, including two famous model organisms: *Neurospora crassa* (Schechtman, 1987) and *Aspergillus nidulans* (Bhattacharyya and Blackburn, 1997; Connelly and Arst, 1991). However, beyond the telomeric DNA sequences, little is known about the telomeres structure or the telomerase enzyme components within filamentous fungi. Understanding telomerase regulation in filamentous fungi, another group with conserved canonical short telomeric DNA repeats, will shed lights into the telomerase RNA and RNP evolution.

Neurospora crassa is a well characterized model organism in genetics, having been studied for over 70 years (Davis, 2000). The Nobel Prize in Physiology or Medicine in 1958 was awarded to George W. Beadle and Edward L. Tatum for their famous 'one gene, one enzyme' hypothesis proposed from *N. crassa* mutagenesis study. The 38 Mb genome of *N. crassa* was sequenced in 2003, about 10,000 genes predicted, almost double the number of yeast genes. DNAs ectopically transformed of *N. crassa* typically integrate into the genome through random insertions, as non-homologous end joining (NHEJ) is the major DNA repair system in *N. crassa*. This is problematic for gene replacement strategies and has recently been overcome by disrupting either the *Ku70* or *Ku80* genes eliminating NHEJ (Ninomiya et al., 2004). Within either the Ku70 or Ku80 deficient strain, virtually all observed recombination events were homologous,

providing an efficient approach for gene knock-out and gene replacement. Furthering this system, a high throughput gene knock-out procedure has been developed in *N. crassa*, with most protein coding gene knock-out strains developed available from fungal genetics stock center (FGSC). These systems present *N. crassa* as advantageous in the study of telomerase regulation and telomeres biology.

The identification of the *N. crassa* telomerase RNA (ncTER1) using a novel biochemical and bioinformatics approach is described in Chapter 2 along with the characterization and verification of RNA. Furthermore, the 3' end processing mechanism, found to be similar to *S. pombe* TER1 is detailed. Chapter 3 and 4, is a continuation of the *N. crassa* telomerase study, with the ncTER1 secondary structure model determined from the phylogenetic comparative analysis of 69 identified filamentous fungal TER1 sequences. The secondary structure model indicates conserved domains shared between vertebrate and yeast telomerase RNAs. *N. crassa* telomerase *in vitro* reconstitution system was developed and human type telomerase activity was observed in *N. crassa* with significant and processive activity. Only *N. crassa* Est1 protein, homolog to yeast Est1, was identified and found to be associated with active *N. crassa* telomerase *in vivo*, suggesting a unique telomerase RNP biogenesis might be present in filamentous fungal telomerase RNA.

1.4 References

Baltimore, D. (1970). RNA-dependent DNA polymerase in virions of RNA tumour viruses. Nature *226*, 1209-1211.

Bhattacharyya, A., and Blackburn, E.H. (1997). Aspergillus nidulans maintains short telomeres throughout development. Nucleic Acids Res *25*, 1426-1431.

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

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

Bryan, T.M., Sperger, J.M., Chapman, K.B., and Cech, T.R. (1998). Telomerase reverse transcriptase genes identified in Tetrahymena thermophila and Oxytricha trifallax. Proc Natl Acad Sci U S A *95*, 8479-8484.

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

Cifuentes-Rojas, C., Kannan, K., Tseng, L., and Shippen, D.E. (2011). Two RNA subunits and POT1a are components of Arabidopsis telomerase. Proc Natl Acad Sci U S A *108*, 73-78.

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

Connelly, J.C., and Arst, H.N., Jr. (1991). Identification of a telomeric fragment from the right arm of chromosome III of Aspergillus nidulans. FEMS Microbiol Lett *64*, 295-297.

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

Davis, R.H. (2000). Neurospora : contributions of a model organism (Oxford ; New York, Oxford University Press).

Evans, S.K., Sistrunk, M.L., Nugent, C.I., and Lundblad, V. (1998). Telomerase, Ku, and telomeric silencing in Saccharomyces cerevisiae. Chromosoma *107*, 352-358.

Feng, J., Funk, W.D., Wang, S.S., Weinrich, S.L., Avilion, A.A., Chiu, C.P., Adams, R.R., Chang, E., Allsopp, R.C., Yu, J., *et al.* (1995). The RNA component of human telomerase. Science *269*, 1236-1241.

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

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

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

Hinkley, C.S., Blasco, M.A., Funk, W.D., Feng, J., Villeponteau, B., Greider, C.W., and Herr, W. (1998). The mouse telomerase RNA 5"-end lies just upstream of the telomerase template sequence. Nucleic Acids Res *26*, 532-536.

Jady, B.E., Bertrand, E., and Kiss, T. (2004). Human telomerase RNA and box H/ACA scaRNAs share a common Cajal body-specific localization signal. J Cell Biol *164*, 647-652.

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

Lin, J., Ly, H., Hussain, A., Abraham, M., Pearl, S., Tzfati, Y., Parslow, T.G., and Blackburn, E.H. (2004). A universal telomerase RNA core structure includes structured motifs required for binding the telomerase reverse transcriptase protein. Proc Natl Acad Sci U S A *101*, 14713-14718.

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

McClintock, B. (1939). The Behavior in Successive Nuclear Divisions of a Chromosome Broken at Meiosis. Proc Natl Acad Sci U S A *25*, 405-416.

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

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

Mitchell, J.R., and Collins, K. (2000). Human telomerase activation requires two independent interactions between telomerase RNA and telomerase reverse transcriptase. Mol Cell *6*, 361-371.

Morin, G.B. (1989). The human telomere terminal transferase enzyme is a ribonucleoprotein that synthesizes TTAGGG repeats. Cell *59*, 521-529.

Muller, H.J. (1938). The remaking of chromosomes. Collecting Net *13*, 181-198.

Ninomiya, Y., Suzuki, K., Ishii, C., and Inoue, H. (2004). Highly efficient gene replacements in Neurospora strains deficient for nonhomologous end-joining. Proc Natl Acad Sci U S A *101*, 12248-12253.

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

Schechtman, M.G. (1987). Isolation of telomere DNA from Neurospora crassa. Mol Cell Biol *7*, 3168-3177.

Seto, A.G., Zaug, A.J., Sobel, S.G., Wolin, S.L., and Cech, T.R. (1999). Saccharomyces cerevisiae telomerase is an Sm small nuclear ribonucleoprotein particle. Nature *401*, 177-180.

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

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

Singer, M.S., and Gottschling, D.E. (1994). TLC1: template RNA component of Saccharomyces cerevisiae telomerase. Science *266*, 404-409.

Spatafora, J.W., Sung, G.H., Johnson, D., Hesse, C., O'Rourke, B., Serdani, M., Spotts, R., Lutzoni, F., Hofstetter, V., Miadlikowska, J., *et al.* (2006). A five-gene phylogeny of Pezizomycotina. Mycologia *98*, 1018-1028.

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

Temin, H.M., and Mizutani, S. (1970). RNA-dependent DNA polymerase in virions of Rous sarcoma virus. Nature *226*, 1211-1213.

Watson, J.D. (1972). Origin of concatemeric T7 DNA. Nat New Biol *239*, 197-201.

Webb, C.J., and Zakian, V.A. (2008). Identification and characterization of the Schizosaccharomyces pombe TER1 telomerase RNA. Nat Struct Mol Biol *15*, 34-42.

Yao, M.C., Blackburn, E., and Gall, J. (1981). Tandemly repeated C-C-C-C-A-A hexanucleotide of Tetrahymena rDNA is present elsewhere in the genome and may be related to the alteration of the somatic genome. J Cell Biol *90*, 515-520.

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

Zappulla, D.C., Goodrich, K., and Cech, T.R. (2005). A miniature yeast telomerase RNA functions in vivo and reconstitutes activity in vitro. Nat Struct Mol Biol *12*, 1072-1077.

Chapter 2

IDENTIFICATION OF NEUROSPORA CRASSA TELOMERASE RNA

2.1 Abstract

Telomerase RNA (TER) is the most divergent RNA reported. To date, it has only been identified in ciliate, yeast, vertebrate, and plant species with striking variations in size, sequence, structure, and *in vivo* biogenesis. Telomerase RNA has yet to be identified in filamentous fungi, including the most characterized genetics model organism *Neurospora crassa*, which harbors the vertebrate type telomere repeat, TTAGGG. Herein, the *Neurospora crassa* telomerase RNA (ncTER1) is identified employing novel biochemical and bioinformatics approaches. Telomerase holoenzyme was purified from an *N. crassa* strain expressing a recombinant tagged TERT, ncTERT-3xFLAG. Solexa deep sequencing of RNA co-purified with ncTERT and transcriptome assembly produced a limited number of ncTER candidates. These candidates were then screened for the predicted template sequence. An ncTER candidate, herein named ncTER1, containing a putative 9 nt template sequence, 5'-UAACCCUAA-3', was cloned and 5'- 3'-RACE determined the mature RNA to be 2,049 nt after 3' processing. However, two additional polyadenylated ncTER1 forms were identified, a 3' intron harboring precursor and a spliced form derived from splicing the precursor. Several ncTER1 template mutants were over-expressed in *N. crassa*, adding mutant telomeric repeats to the chromosome termini and the corresponding telomere mutants were cloned. The incorporation of mutant sequences synthesized from the ncTER1 template mutants suggests ncTER1 is indeed the telomerase RNA component and reconstitutes with ncTERT to form an active telomerase enzyme *in vivo*

2.2 Introduction

Telomeres cap the eukaryotic linear chromosomal ends with tandem repetitive sequences bound to several telomere DNA binding proteins, the telosome/shelterin complex. The end-replication problem erodes the telomeres with each cell replication cycle. The telomerase enzyme, a ribonucleoprotein minimally consisting of a catalytic telomerase reverse transcriptase (TERT) and a telomerase RNA (TER) (Greider and Blackburn, 1989; Lingner et al., 1997b), maintains telomere length and counterbalances the telomere loss from eukaryotic cellular proliferation. Telomeric DNA sequences are reverse transcribed by telomerase onto chromosomal termini from an intrinsic RNA moiety (Shippen-Lentz and Blackburn, 1990; Yu et al., 1990). TERT is highly conserved among evolutionarily distinct groups of species, whereas TER is quite divergent in size, sequence and structure within even closely associated species. Over the past two decades, a multitude of endeavors attempted, numerous methodologies developed and tremendous efforts put forth for the isolation and identification of telomerase RNAs from evolutionarily diverse groups of species, including ciliates, yeasts, vertebrates, and plants (Blasco et al., 1995; Chen et al., 2000; Cifuentes-Rojas et al., 2011; Feng et al., 1995; Greider and Blackburn, 1989; Leonardi et al., 2008; Singer and Gottschling, 1994; Webb and Zakian, 2008; Xie et al., 2008). Different telomerase RNP biogenesis pathways were also identified to be shared with other known RNPs, indicating complicated telomerase RNP evolution. The identification of additional telomerase RNAs from divergent, as well as closely

associated, groups will add to the body of knowledge regarding telomerase RNA evolution and shed insights into RNP biogenesis.

Yeasts and filamentous fungi (Pezizomycotina) both belong to Ascomycota, the largest phylum in the fungi kingdom. Although telomerase and telomeres biology has been exhaustively investigated within many yeast species, including both the budding and fission yeasts, little is known regarding the telomerase composition, RNA sequence, telosome/shelterin complex or telomere maintenance in any filamentous fungi species. Unlike yeast species, most filamentous fungi were reported to maintain a vertebrate-like short 6 bp telomeric repeat sequence of TTAGGG (Supplemental Table S2.1), including the well studied and characterized model organisms: *Neurospora crassa* (Schechtman, 1987) and *Aspergillus nidulans* (Bhattacharyya and Blackburn, 1997; Connelly and Arst, 1991). A homolog of human telomere binding protein Pot1 was also identified in *N. crassa* (Baumann et al., 2002), indicating a similar telomere maintenance mechanism could reside within both human and *N. crassa*. The *N. crassa* genome, published in 2003, revealed a conservation of higher eukaryotic genes, which are absent in yeast species (Galagan et al., 2003). The advantages of this filamentous fungus can provide another system for telomerase regulation and telomeres biology study, and the characterization of this telomerase will provide important insights to aid the understanding the evolution of telomerase and telomeres regulation.

Second-generation sequencing technology (SGS, e.g. 454 and solexa) is currently broadly applied towards genome sequencing, transcriptome sequencing,

RNA modification analysis, and for this study novel RNA identification. Solexa sequencing employs high through-put analysis of short DNA or RNA fragments, generating a large array of short sequence reads, with up to giga-bases of combined nucleotide read length, from a single sequencing run (Metzker, 2010). These short sequence reads must then assemble into larger fragments for bioinformatics analysis. The solexa second-generation sequencing technology was applied here for the discovery of novel telomerase RNA genes.

We hereby present the identification of *N. crassa* telomerase RNA (ncTER1) through novel biochemical and bioinformatics approaches. Solexa deep sequencing of RNA species extracted from the purified *N. crassa* telomerase ribonucleoprotein provided ncTER candidates. The characterization of ncTER1 revealed three ncTER1 forms with the predominant mature RNA at 2,049 nt, the largest confirmed telomerase RNA to date. The spliceosome mediated intron-partially splicing, a mechanism whereby the 3' exon is discarded along with the intron was originally discovered in *S. pombe* TER1. This novel mechanism is also important for ncTER1 3' end processing. Several ncTER1 template mutants were constructed and over-expressed in *N. crassa*, adding mutant telomeric repeats to the chromosome termini, and the corresponding telomeric DNA mutations identified, confirming ncTER1 encodes *N. crassa* telomerase RNA.

2.3 Materials and methods

2.3.1 *N. crassa* growth media

Vogel's medium is the minimum medium used for *N. crassa* mycelia growth and stock culture. To make 1L of vogel's medium, 20ml 50x vogel's salt solution

(premade, see below) and 15g sucrose were added into 980ml of dH_2O, and the mixture was autoclaved for 30min at 121^oC. If solid medium was needed, 1.5% agar was added before autoclave. Other supplements such as antibiotics and L-histidine can also be added to the medium after autoclave.

50x vogel's salt solution: 1L

To 750ml of dH_2O, add ingredients with the following order:

Na_3Citrate $2H_2O$	126.7g
KH_2PO_4 (anhydrous)	250g
NH_4NO_3	100g
$MgSO_4$ (anhydrous)	4.88g
$CaCl_2$ (anhydrous)	3.77g
Biotin stock	5ml
Trace element stock	5ml

dH_2O was utilized to bring the volume to 1L. 10ml of Chloroform was added as preservative. Solution was stored in a screw-capped, 1 liter bottle at room temperature.

Biotin stock: 5mg biotin was dissolved in 100ml 50% ethanol and stored at 4^oC in a screw-capped bottle.

Trace element stock: 100ml

To 95ml of dH2O, add the ingredients with the following order:

Citric acid $1H_2O$	5g
$ZnSO_4$ 7H2O	5g
$Fe(NH_4)_2SO_4$ $6H_2O$	1g

$CuSO_4$ $5H_2O$	0.25g
$MnCl_2$ $4H_2O$	0.0586g
H_3BO_3	0.05g
Na_2MoO_4 $2H_2O$	0.05g

dH_2O was utilized to bring the volume to 100ml. 2ml of Chloroform was added as preservative. Solution was stored in a screw-capped, 1 liter bottle at room temperature.

2.3.2 *N. crassa* plating medium

1x Vogel's salt agar was prepared by taking 20ml of 50x Vogel's salt solution into 880ml of dH_2O and 15g of agar. The medium was autoclaved for 30min at 121^oC, and was added by 100ml of 10x sugar solution: 20% L-(-)-Sorbose (MP Biomedicals), 0.5% glucose, and 0.5% fructose. Solid plating medium was made by pouring the medium (~20ml) onto the 100mm sterile plates. Supplements such as antibiotics or histidine can also be included to the medium before pouring the plates.

2.3.3 *N. crassa* transformation

N. crassa transformation was carried out as previously described (Davis, 2000). *N. crassa* strain NC1 (*mat a, his3*$^-$*, mus 52*$^-$) was inoculated onto the center of a 500ml flask filled with 50ml vogel's medium. *N. crassa* cells were incubated at 30-32^oC for 2-3 days in the dark, and then room temperature (~25^oC) with constant light for 10-14 additional days. Conidia were harvested with 50ml sterile water and filtered through 1 layer of sterile cheese cloth. Conidia were transferred into a 50ml falcon tube and centrifuged at 1,500g for 3min with a swinging

bucket rotor. After removing the supernatant, conidia were resuspended and washed with 30ml sterile cold 1M sorbitol, and centrifuged at 1,500 g for 3min. Washing was repeated twice with 1M sorbitol and conidia were resuspended in 0.5-1ml of 1M sorbitol, resulting in ~$2.5x10^9$ conidia/ml. 1-2 μg of DNA (PCR DNA or linearized plasmid DNA) was mixed with 40-50μl conidia in an Eppendorf tube and incubated on ice for 4 min. The conidia and DNA suspension was then transferred into a 2mm electro-cuvette (USA scientific). After electro-pulse the cuvette at 1,500V, 600ohms, and 25μF, 850μl of ice-cold 1M sorbitol was immediately added into the cuvette, and transferred into an Eppendorf tube. 300μl of conidia were plated onto the plating medium containing 2% L-sorbose, and incubated at 30^{o}C for 2-3 days until colonies are visible.

2.3.4 *N. crassa* genomic DNA and total RNA isolation

Conidia were obtained following the same procedure as described in *N. crassa* transformation, and inoculated and grown in the liquid Vogel's medium for 16 h at 30^{o}C with gentle shaking (~180rpm). Mycelia were harvested the next day by filtration through whatman #1 filter paper in a funnel, and stored in -80^{o}C. One hundred mg of mycelia tissue was completely homogenized in liquid nitrogen with a mortar and pestle. Genomic DNA was isolated using Wizard Genomic DNA purification system (Promega) following manufacturer's instruction. The genomic DNA pellet was dissolved in 1xTE buffer (10mM Tris-HCl pH 7.4, 1mM EDTA) and stored at 4^{o}C. Total RNA was isolated from 100mg of mycelia tissue using Tri-reagent following manufacturer's instruction. The concentrations of genomic DNA and total RNA samples were determined by

A260nm measurement using Nanodrop ND-1000 spectrophotometer (Nano-Drop Technologies).

2.3.5 Southern blotting

Southern blotting analysis was performed using modified in-gel hybridization assay. Two micrograms of genomic DNA isolated from *N. crassa* was completed digested with 10u of appropriate restriction enzymes (New England Biolabs) in recommended 1x NEB buffer at 37^{o}C for 8h up to overnight. Digested genomic DNA samples were resolved in a 1% agarose gel at 36 V for 12 hours. The agarose gel was washed twice in 500ml of 1.5M NaCl/0.5M NaOH for 20min each, following with 500ml of 1.5M NaCl/0.5M Tris-HCl pH7.4 washing twice for 15min each. After the gel was washed in 1L of dH_2O for 30min, it was dried in a gel drier at 50^{o}C for 1 hour. The dried gel was then prehybridized for 1hour at 52-58^{o}C (Temperature varies upon different probes) in 20ml hybridization buffer (5x SSC, 5x Denhardt's reagent, 0.1%SDS, and 20mM sodium phosphate pH7.0). Gel was then hybridized with ^{32}P end-labeled oligonucleotide overnight (~16 h) at 52-58^{o}C. Gel was then washed three times with 3x SSC/0.1% SDS at 58^{o}C for 20min each, following by three more stringent wash with 2x SSC/0.1% SDS at 58^{o}C for 20min each. The gel was wrapped with saran wrap and exposed with phosphorimager screen (Bio-Rad, 1hour to overnight).

2.3.6 ncTER1 Northern blotting

Five micrograms of total RNA or 500pg of *in vitro* transcribed *N. crassa* telomerase RNA (mature and precursor forms) were resolved in a 1.5%

formaldehyde denaturing agarose gel (in 1x MOPS buffer and 20% volume of formaldehyde solution 37%) at 10V/cm for 3-5 hours. The gel was rinsed with 500ml of dH_2O twice with 15min each, and then washed in 500ml of 10x SSC buffer (1.5M NaCl, 0.15M of sodium citrate, pH7.0) for 10min. Three pieces of 3MM filter paper and one piece of Hybond XL membrane (GE healthcare) were cut with the same size as the gel and rinsed with 10x SSC buffer. The agarose gel, hybond XL membrane, filter paper and paper towels were assembled as indicated in Figure 2.1. Transferring was carried out at room temperature for 8-24 hours. The RNA transferred was then crosslinked to the Hybond XL membrane by XL-1500 UV crosslinker (Spectronics corporation) using optimal crosslink mode. The membrane was then prehybridized with 20 ml of Ultra-hyb hybridization buffer (Applied Biosystems) in a hybridization bottle at 65^{o}C for 30 min, and hybridized with radiolabeled riboprobe at 65^{o}C overnight. Riboprobe was prepared by *in vitro* transcription using Maxiscript kit (Applied Biosystems) following manufacturer's instruction and was purified through G-25 gel filtration mini-columns (Roche). Following hybridization overnight, the membrane was washed twice with 20ml 1x SSC and 0.2% SDS at 65^{o}C for 10min and twice with 20 ml 0.2x SSC and 0.1 % SDS at 65^{o}C for 20min. After wrapping with saran wrap, the membrane was exposed and analyzed with phosphorimager (Bio-Rad) for 1h up to overnight.

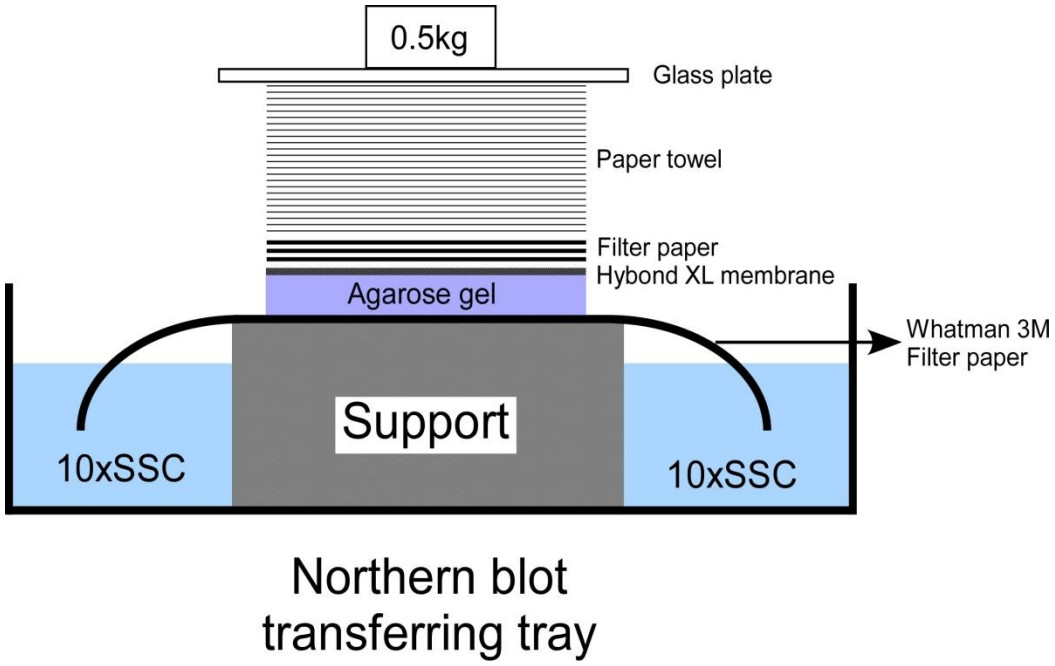


Figure 2.1 Schematic of northern blot capillary transferring.

10x SSC buffer was added to the transferring tray close to the top of the support. One piece of whatman 3MM filter paper was prepared to be wider than the agarose gel and longer enough to touch the buffer. The washed agarose gel was inverted and placed on top of the filter paper. The hybond XL membrane was placed on top of the agarose gel without any bubble formed between. Three pieces of whatman 3MM filter paper and a stack of paper towel were placed on the top and covered with a glass plate and 0.5kg of weight.

2.3.7 *N. crassa* nuclear extract isolation

Two grams of *N. crassa* mycelia was disrupted with a Bead Beater in a ice-cold buffer consisting of 1M sorbitol, 7% Ficoll (w/v), 20% glycerol (v/v), 5mM $MgCl_2$, 10mM $CaCl_2$, and 1% Triton X-100. After five 10-second pulses in the bead beater, with 10-second intervals, the cell slurry was collected and centrifuged at 1,500g for 10 minutes. The supernatant was transferred and centrifuged at 15,000g for 20 minutes. The nuclei pellet was then washed once with buffer consisting of 50mM Tris-HCl, pH 7.5, 5mM $MgCl_2$, 10mM BME, 20% glycerol and 0.6M sucrose. CHAPS lysis buffer (10mM Tris-HCl pH 7.5, 400mM

NaCl, 0.5% CHAPS, 5% glycerol, 1mM EGTA, 1mM $MgCl_2$, supplemented with 5mM BME and 1x protease inhibitor cocktail [Roche] before use) was used to lyse the nuclei and nuclear extract was collected by centrifuge at 14,000g for 10minutes. The isolated nuclear extract was immediately loaded for gel filtration chromatography or stored at -80^{o}C.

2.3.8 Size exclusion chromatography (Gel filtration)

XK26/60 column packed with Sephacryl S-500 resin (GE healthcare) was used for initial purification of *N. crassa* holoenzyme. 2ml of the nuclear extract was loaded onto the column balanced with the column buffer (10mM Tris-HCl pH7.5, 400mM NaCl, 5% glycerol). Fractions were collected with 5ml per tube and UV absorbance was monitored. Telomerase TRAP assay (see below) was carried out for each fraction.

2.3.9 Anti-FLAG antibody immunoprecipitation

Twenty microliters of anti-FLAG m2 antibody agarose beads (50% slurry, Sigma-Aldrich) was washed three times with 1xTBS buffer (10mM Tris-HCl pH7.4, 150mM NaCl). Cell lysate was directly added to the beads and incubated at 4^{o}C for 1-2 hours with gentle rotation. Beads were collected via centrifugation at 1,000g for 15s, and washed with 300ul of 1xTBS buffer three times. RNA was extracted directly from the beads with Phenol/Cloroform extraction for Northern blotting analysis, or 1x Laemmli buffer (0.125M Tris-HCl, pH6.8, 2% SDS, 10% glycerol, 5% 2-mercaptoehanol and 0.0025% bromophenol blue) was directly applied to the beads for Western Blotting analysis.

2.3.10 Telomerase TRAP assay

A modified two-tube TRAP assay was utilized to monitor the *N. crassa* telomerase activity during the holoenzyme purification. In the initial telomerase reaction, 1µl of lysate or immunoprecipiteted enzyme was incubated at 30^{o}C for 1 hour in 1x PE buffer (50mM Tris-HCl pH 8.3, 2mM DTT, 0.5mM $MgCl_2$, and 1mM spermidine) together with 0.5µM of TS primer (5'-AATCCGTCGAGCAGAGTT-3'), 50µM of dATP, dTTP and dGTP. The TS primer was then purified with phenol/chloroform extraction following with ethanol precipitation. The precipitated sample (or diluted with dH_2O) was then used in a 25µl PCR reaction mix consisting of 1X Taq PCR buffer (New England Biolabs), 0.1mM of each dNTP, 0.4µM of ^{32}P end-labeled TS primer, 0.4µM of ACX primer (5'-GCGCGGCTTACCCTTACCCTTACCCTAACC-3'), 0.4µM of NT primer (5'-ATCGCTTCTCGGCCTTTT-3'), $4x10^{-13}$ M TSNT primer (5'-CAATCCGTCGAGCAGAGTTAAAAGGCCGAGAAGCGATC-3', the concentration could be varied depending on the intensity of telomerase activity), and 1U of Taq DNA polymerase (New England Biolabs). The PCR mix was first denatured at 94^{o}C for 2 minutes, followed with 25-28 cycles of (94^{o}C 25s, 50^{o}C 25s, and 68^{o}C 60s) and a final 68^{o}C extension for 5 minutes. Five microliters of the PCR reaction was resolved in a 10% native polyacrylamide/2% glycerol gel. The gel was then dried in a gel-drier and exposed to phosphorimager screen (Bio-Rad).

2.3.11 5' and 3' Rappid Amplification of cDNA End (RACE)

5' RACE was performed using the FirstChoice RLM-RACE Kit (Applied Biosystems) according to the manufacturer's instructions. For 3' RACE, ~7.5µg

of total RNA was incubated with or without 600U of yeast Poly(A) polymerase (USB Corporation) in 20mM Tris-HCl pH 7.0, 0.6mM $MnCl_2$, 0.02mM EDTA, 0.2mM DTT, 0.1 mg/ml BSA, 10% glycerol, and 0.5 mM ATP at 37^{o}C for 15min. Total RNA was then purified by acid phenol/chloroform extraction and ethanol precipitation, and reverse transcribed with 200U of MMLV (GenScript) and oligo-dT anchor primer (5'-TTCCTCACCATAGTTGCGTCTGTACT_{25}V-3') in 1x RT buffer (GenScript), 250μM of each dNTP, 0.5U/μl of SUPERase IN (Applied Biosystems) and 0.6 M Trehalose at 48^{o}C for 40min. PCR was performed with RT products and a gene-specific primer and an anchor primer (primer sequences were listed in the figure legend). The PCR products were cloned into pZero vector and sequenced.

2.3.12 *N. crassa* telomere G-overhang capture assay (GOCA)

N. crassa telomere and subtelomere DNA sequences were obtained from *N. crassa* genome database (Broad Institute of MIT and Harvard). A Poly-C tail was first added to the G-overhang strand of telomere DNA by incubating 200ng genomic DNA with 6U of Terminal deoxynucleotide Transferase (TdT, USB Corporation) in 10ul 1x TdT buffer (20mM Tris-HCl pH7.8, 50mM KAc, 10mM $MgCl_2$, and 1mM dCTP) for 30min at 37^{o}C. After heat inactivating the TdT at 70^{o}C for 10min, the poly-C tailed genomic DNA (without purification) was directly added into 30ul of PCR mix with the final buffer condition: 67mM Tris-HCl pH8.8, 16mM $(NH_4)_2SO_4$, 5% Glycerol, 0.01% Tween-20, 0.2mM of each dNTP 1U of Taq DNA polymerase (New England Biolabs), 0.4uM of primers ASU2627 (5'-CGGGATCCG_{18}) and ASU2871 (5'-

CAATTTACGACCCCCTCATATCAGCCTCGTTTAGCCTTAGCCGGATTAG -3). The reaction was heat denatured at 95^{o}C for 2min following with 45 cycles of (94℃ 20s, 62℃ 15s, 72℃ 25s) and a final 72^{o}C extension for 2min. 1ul of 100-fold diluted PCR DNA samples were used for nested PCR template in 1X HF buffer (Finnzymes), 0.2mM of each dNTP, 1U of Phusion DNA polymerase (Finnzymes), 0.4μM of primers ASU2627 and ASU2872 (5'-CAAACCAGCCCCTAAGAACCGTAGCATACCGTATACTTACAGTCTACA CTC-3'). The reaction was heat denatured at 98^{o}C for 30s following with 30 cycles of (98℃ 12s, 65℃ 10s, 72℃ 10s) and a final 72^{o}C extension for 1min. The PCR products (~300bps) were gel purified and ligated into the positive selection vectors, and plasmid DNAs were sent for sequencing using M13F and M13R primers.

2.3.13 Construction of positive selection vectors for PCR DNA cloning

pCR4-TOPO vector (Invitrogen) was used for construction of positive selection vectors. Three DNA adaptors were created by annealing two oligonucleotides together respectively: a) 5'-AATTCGGCGATATCGAATT-3' and 5'-AATTCGATATCGCCGAATT-3'; b) 5'-AATTCGAGATATCGGAATT-3' and 5'-AATTCCGATATCTCGAATT-3'; c) 5'-AATTCCGATATCTCGAATT-3' and 5'-AATTCGAGATATCGGAATT-3'. These DNA adaptors were then ligated into *EcoR* I digested pCR4-TOPO vector and single adaptor insertions (pCR4-TOPO a, b, and c) were selected in *E.coli* strain DB3.1 (Invitrogen). The three plasmids (3.3ug each) purified from DB3.1 cells were then digested with 20U of *EcoR* V enzyme (New England Biolabs) in

1x buffer 3 at 37^{o}C for 1hour. PCR DNAs or other blunt-ended DNAs were then ligated into the purified linearized vectors (Promega Wizard SV gel and PCR purification kit) with 10U T4 DNA ligase (Invitrogen) in 1x ligation buffer at room temperature for 2-4 hours. The self ligated vector kills *E.coli* DH10B cells via expression of toxic ccdB protein, while the insertion of PCR DNA will be selected by disrupting the ccdB ORF or providing a stop codon. Three adaptor constructs (a, b and c) increase the possibility for having a stop codon within the insert when disrupting the ccdB ORF is not available.

2.4 Results

2.4.1 Construction of an *N. crassa* strain with expression of ncTERT-3xFLAG.

All *N. crassa* strains constructed and utilized in this study were listed in the Supplemental Table A.1 in the Appendix A. The predicted *N. crassa tert* gene sequence was obtained from GenBank, and *tert* transcript was analyzed via 5' and 3' RACE and the full length new *tert* sequence was deposited into the GenBank. In order to purify the *N. crassa* telomerase holoenzyme, an ncTERT-3xFLAG *N. crassa* strain was generated by gene replacement for the insertion of a 3xFLAG tag towards the 3' end of ncTERT. Overlapping PCR was carried out to amplify the whole gene replacement construct which was then integrated into *N. crassa* strain NC1 (*mat A, his-3, mus-52*) (Fig 2.2).

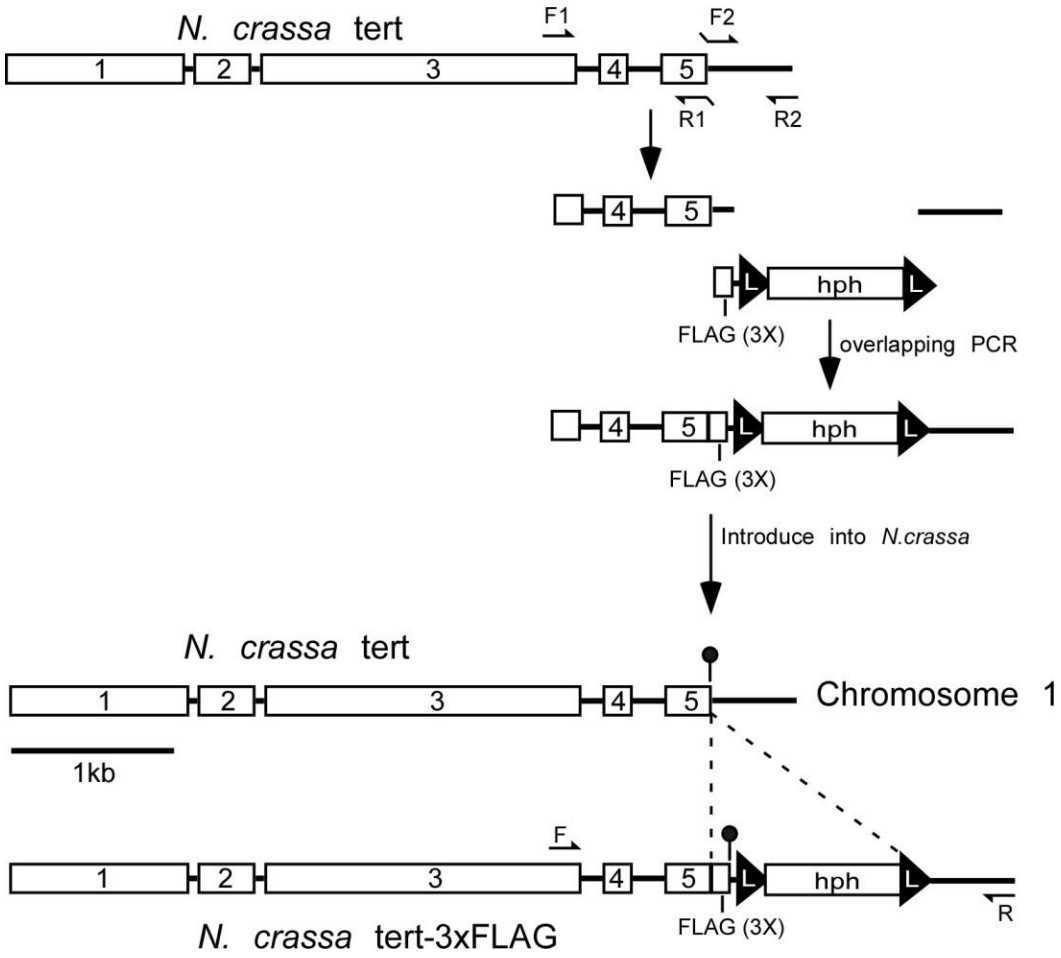
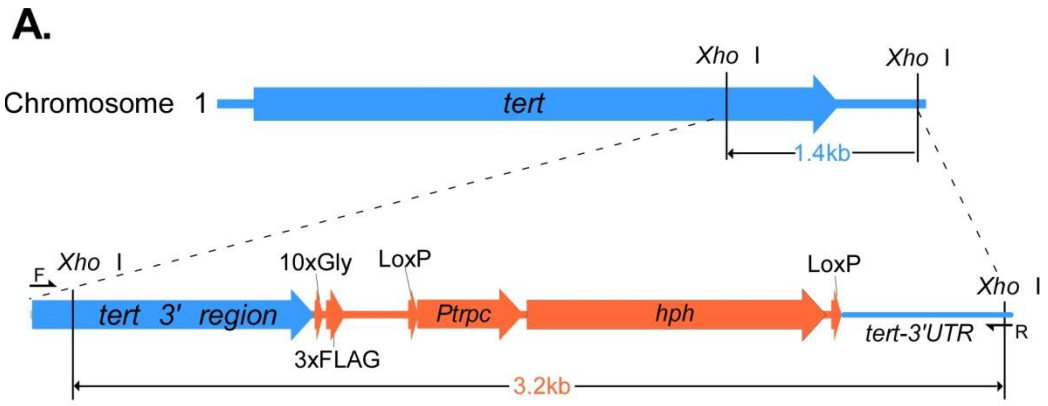


Figure 2.2 Construction of an *N. crassa* strain expressing ncTERT-3xFLAG

N. crassa tert locus locates at chromosome 1 with five exons shown. 1000bp *tert* 3'-end sequence and 500bp of *tert* 3'UTR sequence were amplified from *N. crassa* genomic DNA with primers F1+R1 and F2+R2 respectively. The 10xGly-3xFLAG-loxP-hph sequence was amplified from p3xFLAG-hph vector. Overlapping PCR was achieved using these three PCR fragments and F1 and R2 primers. The entire PCR fragment was introduced into *N. crassa* strain NC1 by electroporation as described in the materials and methods. Hygromycin was utilized for transformants selection. L: LoxP; hph: hygromycin B phosphotransferase. The black dot indicates the ncTERT native stop codon which was switched to the C-terminal end of 3xFLAG tag upon integration of PCR DNA into the genome.

Selection for hygromycin resistance yielded several transformants. The C-terminal 3xFLAG tag insertion was confirmed by southern blotting analysis. Probe for the 3' terminus of the *tert* gene together with *Xho* I endogenous *tert* loci

digested generated a 1.4 kb wild-type band, while a 3.2 kb band was detected from insertion of the 3xFLAG tag together with *hph*, hygromycin resistance gene (Fig 2.3). Of the 12 transformants examined, 3 were homokarya (#2, 6, 12) in which all nuclei contain the C-terminal 3xFLAG tagged TERT, while the remaining 9 were either heterokarya (#4-5, 7-11) or ectopic insertion (#3). A homokaryon was then crossed to the wildtype strain and an *N.crassa* strain (NC3) with *his-3*$^{+}$, *mus52*$^{+}$ *and hph*$^{+}$ genotype was selected.



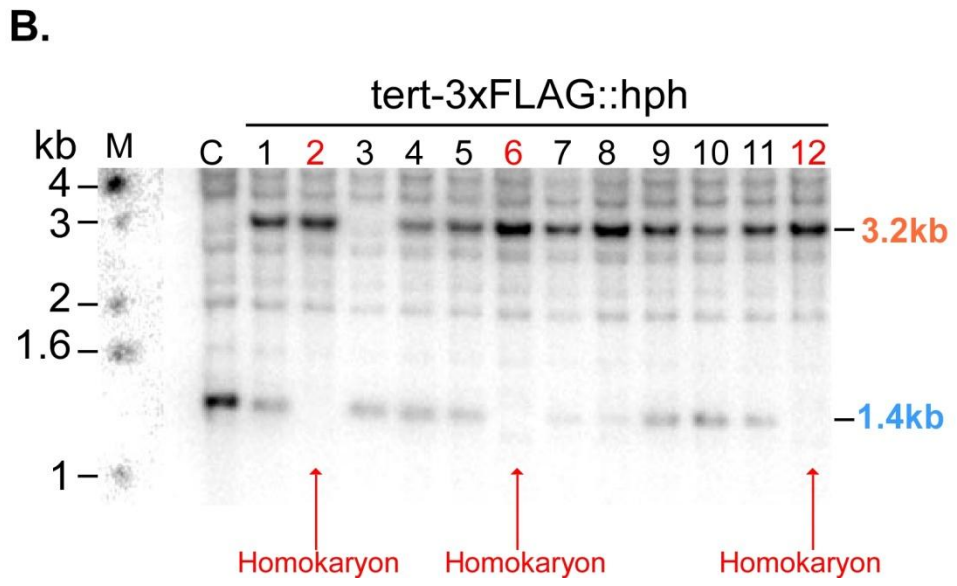


Figure 2.3 *N. crassa* transformants screening.

(A) Schematic drawing shows the *N. crassa tert* loci before and after the insertion. The green color indicates the endogenous *tert* loci, whereas the orange color shows the insertion. Digestion of the genomic DNA with *Xho* I enzyme results the 1.4kb band in endogenous *tert* loci which increases to 3.2kb upon the insertion. PtrpC: *Aspergillus nidulans* trpC promoter; hph: hygromycin B phosphotransferase. (B) Southern blot analysis of wt (indicated in lane C) and 12 transformants (labeled from 1 to 12) with radiolabled probes against *tert* 3' region. Red arrows indicate #2, 6 and 12 clones are homokarya as no endogenous band was detected. DNA fragment size marker (labeled as lane M) was labeled with sizes on the left.

2.4.2 Purification of *N. crassa* telomerase holoenzyme from nuclear extract.

After isolation of the nuclear extract from ncTERT-3xFLAG *N.crassa* strain, a two-step purification scheme was performed with TRAP assay to follow the telomerase enzymatic activity (Fig 2.4). The first step was gel filtration, size exclusion chromatography. TRAP activity assays were performed on the collected fractions (Fig 2.5). Fractions with peak activity were combined and immunoprecipited with anti-FLAG M2 antibody affinity gel (Sigma). TRAP assay was again performed to monitor the telomerase enzymatic activity bound to the affinity gel (Fig 2.6).

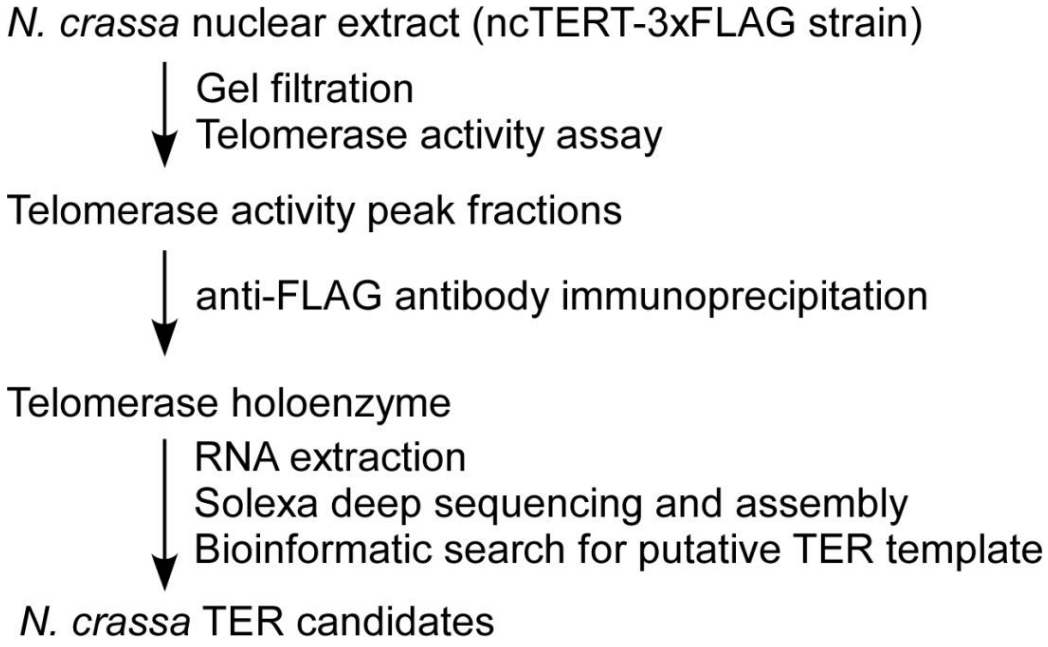


Figure 2.4 *N. crassa* telomerase holoenzyme purification and ncTER candidates search procedure

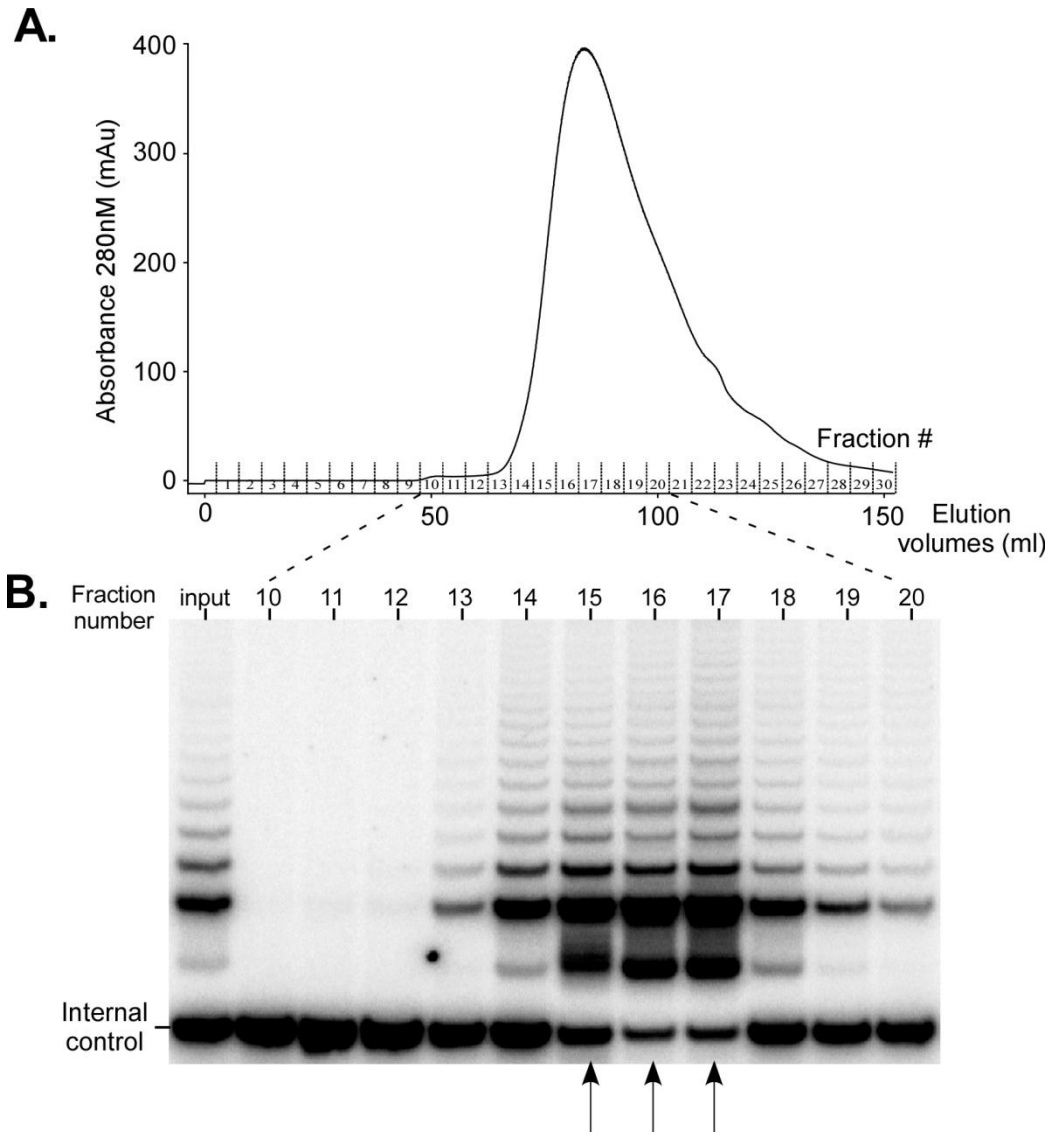


Figure 2.5 Purification of *N. crassa* telomerase holoenzyme through gel filtration chromatography.

(A) Nuclear extract isolated from 2g of mycelia (NC1 strain) was loaded onto the S-500 gel filtration chromatography. Fractions were collected with 5ml each as indicated. UV absorbance 280nm was monitored with an FPLC machine. Fractions collected from # 10 to 20 were subjected for telomerase TRAP activity assay together with the original nuclear extracts (B). The PCR internal control band is shown as labeled. Fractions with peak activities (#15, 16 and 17) were collected and combined for the next purification step.

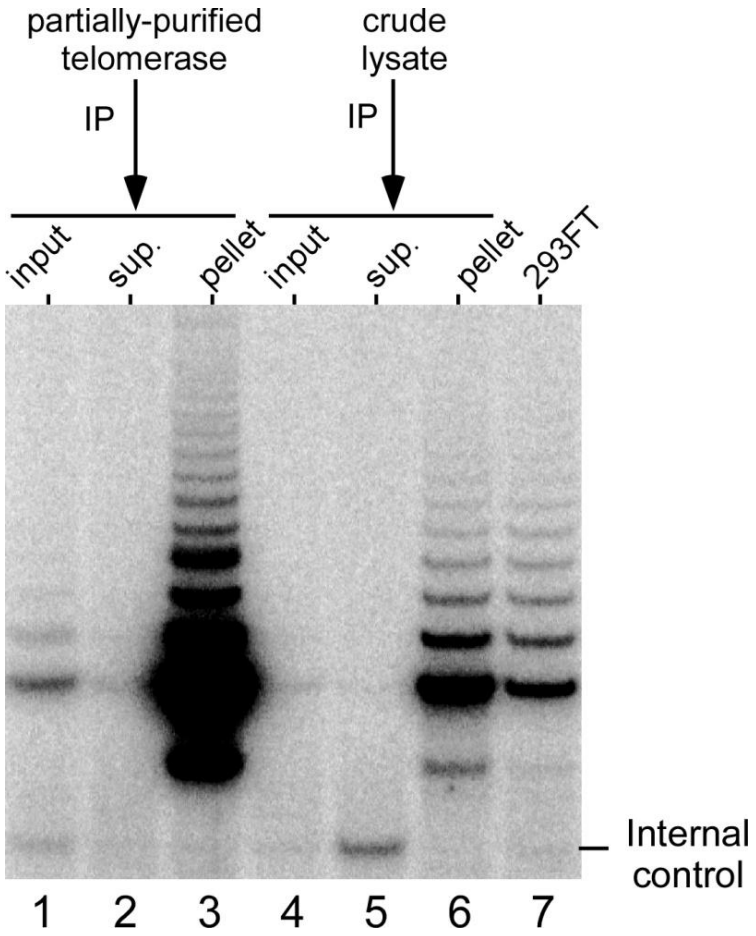


Figure 2.6 Telomerase TRAP activity assay after ncTERT-3xFLAG IP

Telomerase peak activity fractions from gel filtration (15, 16 and 17) were immunoprecipitated against ncTERT-3xFLAG by using anti-FLAG antibody conjugated beads. 1μl of input and supernatant (sup.) after IP together with 5% of the beads were subjected to TRAP activity assay. The original *N. crassa* nuclear extract before gel filtration was handled with the same procedure. 1μg of 293FT whole cell protein fraction was used as a comparison for TRAP activity. RNA was recovered from the rest 95% beads as described in the experimental procedures.

2.4.3 Solexa deep sequencing assembly and identification of *N. crassa* TER candidates

RNAs bound to the immunoprecipitation beads were recovered for Solexa deep sequencing. More than 15 million reads were generated with 50-60 bp

lengths for each read. The reads were assembled and candidates were identified as follows:

For *in silico* detection we developed a pipeline consisting of two steps: (1) Generating Candidate Set (2) Filtering Candidate Set.

Candidate Set: The reverse complement of 5'-TTAGGG-3' is expected to be part of the telomerase RNA. Candidates were set to have 5'-CCCUAA-3' with a length of 8nt on both strands. The initial candidate set consists of sequences including the template and a 500nt downstream region.

Filter Steps: Starting with 342 708 initial candidates, 90% of the *Neurospora crassa* candidates were removed by the following procedure:

(a) The template sequence occurs in known telomerase RNAs at most 1.5 times. Candidates with more than three repetitions of the template are interpreted as repeats and therefore removed.

(b) Candidates with more than 10 unknown neighbouring nucleotides (N) were removed.

(c) Sequences shorter than 100 nt were removed, since the following scoring scheme could not be calculated.

(d) Identical or highly similar sequences were identified by blastclust (NCBI toolkit) and removed. Since most genomes are assembled on scaffold or contig level sequences with multiple genomic occurrences can be removed.

(e) Potential protein sequences were removed. These sequences were identified by comparing them with blastx (Altschul et al., 1997) against known proteins from NCBI.

The short-read deep sequencing data were then used to further reduce the genome predicted candidate set. The short read library with more than 15 million sequencing reads was mapped to the genome predicted TER candidates using segemehl (Hoffmann et al., 2009) requiring a minimum accuracy rate of 85% and a maximum seed E-value of 10. Each read was permitted up to 100 matches with optimal score in the target genome, all suboptimal hits were discarded. This sensitive parameter setting ensured that almost 88% of the reads were matched to the reference genome. Ambiguous reads were corrected by hit normalization where the expression of a read was evenly distributed among all best hits. Only candidates with significant expression were considered in the downstream analysis.

The final candidates were sorted by the number of reads mapping to the candidate sequences and candidates with greater than 20 reads are presented in Table 2.1.

Table 2.1:

N. crassa telomerase RNA candidates with more than 20 reads

Supercontig	Strand	Range	# of reads	Size (nt)	Template length (nt)	RNA type
supercontig_9.8	-	92906-99105	6692000	6200	10 CCCUAACCCU	rRNA
supercontig_9.2	-	2947024-2948936	3364	1913	9 UAACCCUAA	TER1
supercontig_9.1	-	9366025-9366391	119	367	9 CCUAACCCU	Hypothetical mRNAs
supercontig_9.2	-	3793084-3793573	101	490	9 AACCCUAAC	
supercontig_9.3	+	973366-973926	63	561	8 CCUAACCC	

The candidate with the greatest number of reads was determined to be a precursor rRNA transcript (including 18s, 5.8s and 26s rRNAs), which harbors a 10 nt putative template sequence: 5'-CCCUAACCCU-3'. The second candidate had 3,364 reads, significantly greater number of reads than the remaining candidates. This candidate TER is a hypothetical mRNA sequences in GenBank, whereas the remaining candidates were not predicted from the gnome to produce an RNA transcript. With the exception of the rRNA precursor, the remaining candidates could not be identified by comparison to known and predicted protein sequences. The high number of reads and inability to associate this transcript with a known fugal gene together with the presence of a 9 nt template sequence, 5'-UAACCCUAA-3', prompted us to putatively name the second candidate as 'TER1', as a potential *N. crassa* telomerase RNA.

2.4.4 Cloning and characterization of the putative *N. crassa* telomerase RNA gene

The *N.crassa ter1* putative gene is mapped to chromosome 2 in a region flanked by several hypothetical protein coding genes (Fig 2.7 A). Southern blot analysis indicates that *ter1* is located at a single site in the genome (Fig 2.7 B), consistent with gene duplication suppression by repeat induced point mutation (RIP) in *N. crassa* (Selker, 1990).

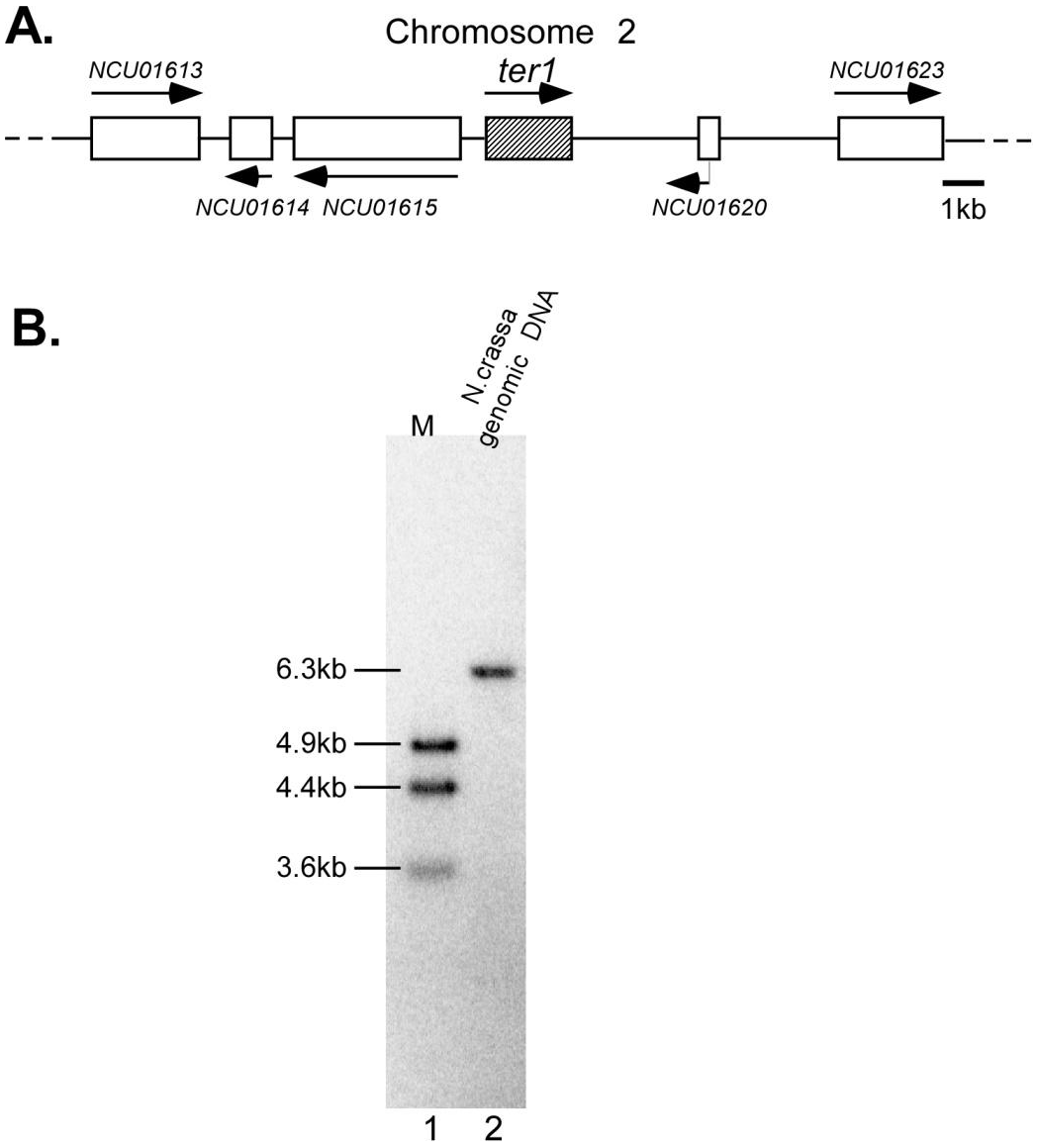


Figure 2.7 *N. crassa ter1* gene analysis.

(A) *N. crassa ter1* gene was mapped to chromosome 2 with several unknown protein coding genes as indicated. The arrows indicate the gene transcription directions. (B) Southern blot analysis of *N. crassa ter1* gene. *N. crassa* Genomic DNA was digested with *Bsp* HI (New England Biolabs) and resolved in 1% agarose gel (lane 2). Southern blot was carried out as described in materials and methods. Linearized plasmid DNAs containing *ter1* sequences were utilized as the size marker (Lane 1). Radiolabeled oligonucleotides specific for *N. crassa ter1* gene were utilized as the hybridization probes. The probe sequences are: 5'-CCAACCACAAGTCGGTCTCTGTTTTG-3', 5'-GAGAGGGCTGAAAGACT-CGATG-3' and 5'-GACCAAACTCGTTGAAAACCAATGC-3'. The sizes of the marker are listed on the left side of the gel.

The *N.crassa ter1* putative gene was verified to encode for an RNA transcript and the full length ncTER1 sequence was determined by 5' and 3' rapid amplification of cDNA end (RACE) assay. Two ncTER1 clones of 5' RACE mapped to the same position (Supplemental fig S2.3). The 3' RACE of ncTER1 performed from poly(A) polymerase treated total RNA amplified a single band. Sequencing results indicated that there was heterogeneity at the 3' end of *in vivo* expressed ncTER1, with the longest transcript terminated at the 2,049 nt. Surprisingly, the 3' RACE for ncTER1 performed from untreated total RNA revealed a larger band, with sequencing identifying two major transcripts. Of these two transcripts, one has an intron sequence spliced out and the other with intron remained (Fig 2.8 and Supplemental fig S2.2). The intron contains a non-canonical 5' splice signal, identified as 5'-AUAAGU-3', following by a putative branch point sequence 5'-GCUGAC-3' and terminating with a 3' splice signal 5'-UAG-3'. The most frequent 3' end of the non-polyadenylated ncTER1 locates immediately upstream of the 5' splice signal. This indicates that spliceosome mediates 3' end processing of ncTER1, a mechanism previously reported in *S. pombe* telomerase RNA (spTER1) (Box et al., 2008). The smallest ncTER1 resulting from the partial splicing, is the only one to be amplified by 3' RACE after *in vitro* polyadenylation, suggesting *N. crassa* employs a similar mechanism. We thus named these three RNA forms as precursor, spliced and mature ncTER1 (Fig 2.8 A). The full length ncTER1 sequence with 5' and 3' RACE mapping is presented in supplemental fig S2.3. To verify that the shortest ncTER1 transcript is the mature RNA, we carried out the northern blot analysis to specifically detect

different forms of ncTER1 (Fig 2.8 B). With the riboprobe specifically against exon 2, no signal was detected within the total RNA pool (Fig 2.8B, left). The ncTER1 from the total RNA, however, could only be detected with the riboprobe against exon 1, with the same size as the *in vitro* transcribed mature RNA (2,049nt) (Fig 2.8 B, right).

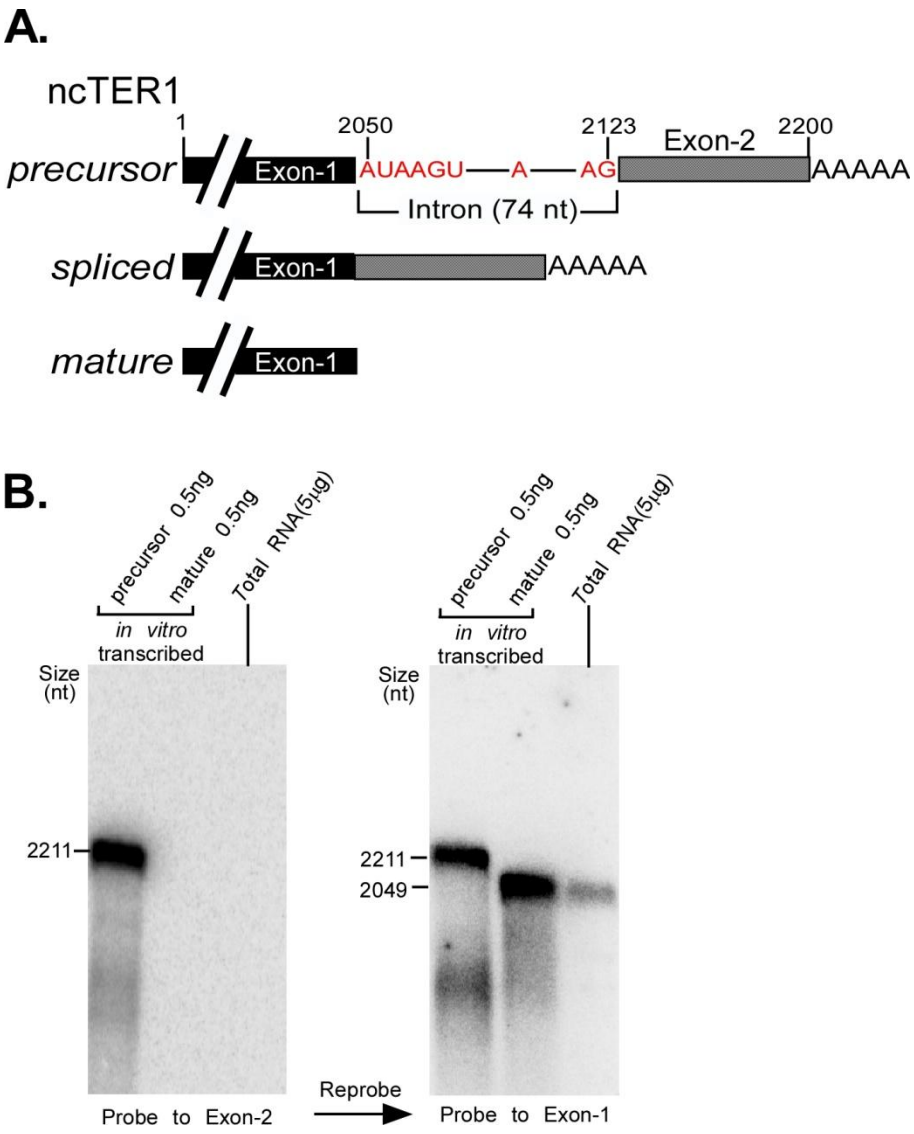


Figure 2.8 Schematic of three ncTER1 forms and northern blot analysis

Figure 2.8 Schematic of three ncTER1 forms and northern blot analysis

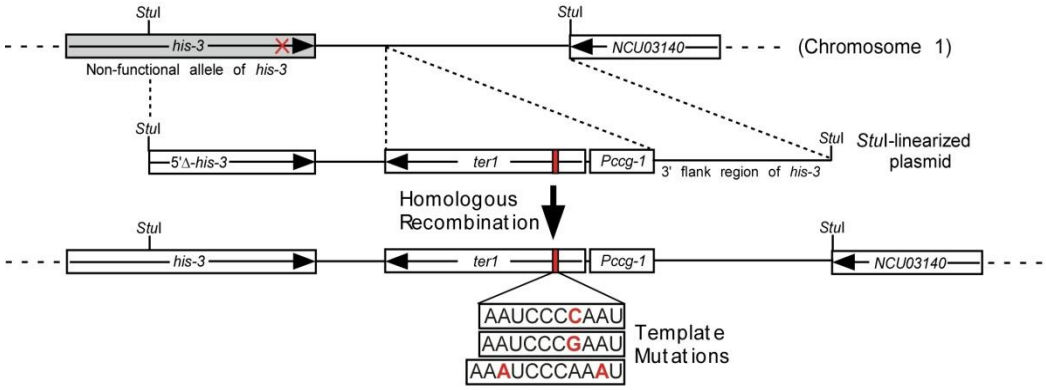
(A) Schematic of three *N. crassa* TER1 forms identified from 5' and 3' RACE. The numbers indicate the nucleotide position within the ncTER1 precursor. The 74nt intron was shown in precursor RNA with 5' splice signal (AUAAGU), branch point nucleotide A, and 3' splice signal (AG) shown in red. Five As were drawn to the 3' end of precursor and spliced RNA forms, as indicating the polyA tail. (B) Northern blotting analysis with 0.5ng *in vitro* transcribed RNAs and 5 μg of *N. crassa* total RNA. The membrane was initially probed with riboprobe against exon 2 (2124-2204nt), then reprobed against exon 1 (78-594nt) which was shown on the right side. The sizes of *in vitro* transcribed precursor RNA (2,211nt) and mature RNA (2,049nt) were shown.

2.4.5 Expressing ncTER1 template mutants in *N. crassa*

The *in vivo* expression of ncTER1 template mutants in *N. crassa* (Fig 2.9 A) via gene-directed integration would incorporate mutant sequences to the telomeres, confirming ncTER1 as the telomerase RNA component. Site-directed mutagenesis generated mutant templates: a) 5'-UAACCCCUAA-3', b) 5'-UAAACCCUAAA-3', c) 5'-UAAGCCCUAA-3' from wildtype template 5'-UAACCCUAA-3'. *N. crassa ter1* genes with mutant templates were integrated and expressed separately in the *N. crassa* strain NC1 (*matA, his-3*$^{-}$*, Δmus-52::bat*$^{+}$) under the *ccg-1* promoter, a strong promoter from *N. crassa*. Gene-directed integration to the *his-3* locus through homologous recombination was confirmed by Southern blotting (Supplemental fig S2.4). *N. crassa* strains with expression of ncTER1 template mutants were first inoculated on the center of a flask with solid Vogel's minimum media. Conidia were collected after 14 days and inoculated into liquid media for overnight growth. Mycelia were harvested and genomic DNA was isolated from each strain as described in materials and method. G-overhang capture assay (Forstemann et al., 2000) amplified telomeric

DNA repeats from chromosomes IIIL and VR from each mutant strain. About 20% of the telomere DNA sequences obtained harbored one or more corresponding mutant telomere repeats, all of which were close to the 3' end as expected (Fig 2.9 B). This confirmed our discovery that ncTER1 is the *N. crassa* telomerase RNA component.

A.



B.

Template	No. of wt clones	No. of mutant clones	Sequences of mutant clones
3'–AAUCCCCAAU–5'	7	3	---TTA(GGGTTA)$_{19}$GGGGTTA(GGGTTA)$_{7}$-3' ---TTA(GGGTTA)$_{23}$GGGGTTAGG-3' ---TTA(GGGTTA)$_{21}$GGGGTTAGG-3'
3'–AAUCCCGAAU–5'	13	2	---TTA(GGGTTA)$_{16}$GGGCTTAGGGCTTAGGGCTTAG-3' ---TTA(GGGTTA)$_{19}$GGGCTTAGGGCT-3'
3'–AAAUCCCAAAU–5'	12	2	---TTA(GGGTTA)$_{9}$GGGTTTA(GGGTTA)$_{5}$-3' ---TTA(GGGTTA)$_{12}$GGGTTTA(GGGTTA)$_{4}$GGGTTTA (GGGTTA)$_{2}$GGGTTTAGGGTTTTA(GGGTTA)$_{1}$GGGTTTAGGGTTT-3'

Figure 2.9 Expression ncTER1 template mutants in *N. crassa* generates telomere DNA mutations

(A) Ectopic ncTER1 template mutants were expressed in *his-3* loci driven by *ccg-1* promoter. NC1 strain was used which expresses a non-fuctional his-3 protein with the mutation indicated as a red cross mark. Three ncTER1 mutants were shown with template sequences as: 5'-UAACCCCUAA-3', 5'-UAAGCCCUAA-3' and 5'-UAAACCCUAAA-3' with inserted nucleotides marked as red color. *Stu* I-linearized plasmid DNA containing 5' truncated *his-3*, 3' end flanking region of *his-3*, Pccg-1 and ncTER1 template mutant was introduced into *N. crassa* NC1 strain with homologous recombination. The new *his-3* locus contains functional *his-3* gene which was utilized as selection marker. (B) Telomere PCR and sequencing result. Number of wildtype telomere and mutated telomere sequences was shown for each ncTER1 template mutant strain. Each telomere sequence with mutant clones were shown with subtelomere sequences omitted. The mutated nucleotides in telomere sequences were shown in red color, whereas wildtype sequences were indicated as blue.

2.5 Discussion

TER is strikingly divergent in length, sequence and structure among evolutionarily distinct groups of species. In this study, *N. crassa* telomerase RNA (ncTER1) was successfully identified using novel techniques and found to have a surprisingly large length. Characterization of ncTER1 indicates similar 3' end processing system to spTER1.

2.5.1 A novel and universal technique for telomerase RNA identification

N. crassa telomerase RNA was identified with novel biochemical and bioinformatical approaches. These approaches can be applied towards telomerase RNA identification from a variety of additional species and has only two criteria: a sequenced genome available and a telomerase activity assay accessible. Within a sequenced genome, the catalytic *tert* gene is easily identified from homologous blast and cloned. The inability to genetically tag TERT can be offset by employing a TERT peptide antibody for affinity purification of the telomerase

holoenzyme from the cell lysate. The active telomerase holoenzyme is followed during each purification step with the telomerase activity assay. The identification of ncTER1 poses limitations on previous methods of TER discovery, anti-TMG IP and template RT-PCR employed recently for *S. pombe* TER1 isolation (Leonardi et al., 2008; Webb and Zakian, 2008). These previous methods rely on 5'-TMG capping, a non-universal characteristic of TER. Additionally, an rRNA, the most abundant cellular RNA, may harbor a portion, or complete template sequence obscuring RT-PCR amplification. The improvement in sequencing techniques is expanding the number genomes that have, are currently, or soon to be sequenced. Our methodology of combining telomerase holoenzyme purification with screening putative genome and transcriptome template sequences has the potential to become the standard procedure for novel telomerase RNA identification, possibly altered and expanded for additional RNA species. Second generation sequencing technology utilizes PCR based amplification potentially resulting in a sequencing bias. Third-generation sequencing, such as single molecule sequencing, has the potential to overcome this potential problem (Schadt et al., 2010).

2.5.2 Evolution of ncTER1

Similar to the yeast telomerase RNA, the ncTER1 is a long telomerase RNA transcript. In *S. cerevisiae* TLC1, large sections of the RNA were not essential for telomerase function either *in vitro* or *in vivo*. The *N. crassa* telomerase RNA might also have non-essential regions. Several protein binding sites within 1.3 kb *S. cerevisiae* TLC1 have been characterized with the RNP biogenesis well studied.

The ncTER1 having a larger size compared to *S. cerevisiae* TLC1 RNA implies there are several protein binding motifs. Additional filamentous fungal species TER sequences are required for a complete structural and functional analysis.

The ncTER1 precursor is polyadenylated, suggesting that this RNA is an RNA polymerase II (Pol II) transcript, the same polymerase as vertebrate and yeast TER. However, the 3' processing mechanism for vertebrate and fungi species is different. The spliceosome cleavage mediated ncTER1 3' end formation, is quite similar to *S. pombe* TER1 and Candida TLC1, suggesting this mechanism is evolutionarily conserved minimally among the Ascomecetes lineage. Additional telomerase RNA genes from filamentous fungi as well as early branched fungal species would expand our understanding of the evolution of TER 3' end processing.

2.6 References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res *25*, 3389-3402.

Baumann, P., Podell, E., and Cech, T.R. (2002). Human Pot1 (protection of telomeres) protein: cytolocalization, gene structure, and alternative splicing. Mol Cell Biol *22*, 8079-8087.

Bhattacharyya, A., and Blackburn, E.H. (1997). Aspergillus nidulans maintains short telomeres throughout development. Nucleic Acids Res *25*, 1426-1431.

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

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

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

Cifuentes-Rojas, C., Kannan, K., Tseng, L., and Shippen, D.E. (2011). Two RNA subunits and POT1a are components of Arabidopsis telomerase. Proc Natl Acad Sci U S A *108*, 73-78.

Connelly, J.C., and Arst, H.N., Jr. (1991). Identification of a telomeric fragment from the right arm of chromosome III of Aspergillus nidulans. FEMS Microbiol Lett *64*, 295-297.

Davis, R.H. (2000). Neurospora : contributions of a model organism (Oxford ; New York, Oxford University Press).

Feng, J., Funk, W.D., Wang, S.S., Weinrich, S.L., Avilion, A.A., Chiu, C.P., Adams, R.R., Chang, E., Allsopp, R.C., Yu, J., *et al.* (1995). The RNA component of human telomerase. Science *269*, 1236-1241.

Forstemann, K., Hoss, M., and Lingner, J. (2000). Telomerase-dependent repeat divergence at the 3' ends of yeast telomeres. Nucleic Acids Res *28*, 2690-2694.

Galagan, J.E., Calvo, S.E., Borkovich, K.A., Selker, E.U., Read, N.D., Jaffe, D., FitzHugh, W., Ma, L.J., Smirnov, S., Purcell, S., *et al.* (2003). The genome sequence of the filamentous fungus Neurospora crassa. Nature *422*, 859-868.

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

Hoffmann, S., Otto, C., Kurtz, S., Sharma, C.M., Khaitovich, P., Vogel, J., Stadler, P.F., and Hackermuller, J. (2009). Fast Mapping of Short Sequences with Mismatches, Insertions and Deletions Using Index Structures. Plos Comput Biol *5*.

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

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

Metzker, M.L. (2010). Sequencing technologies - the next generation. Nat Rev Genet *11*, 31-46.

Schadt, E.E., Turner, S., and Kasarskis, A. (2010). A window into third-generation sequencing. Hum Mol Genet *19*, R227-240.

Schechtman, M.G. (1987). Isolation of telomere DNA from Neurospora crassa. Mol Cell Biol *7*, 3168-3177.

Selker, E.U. (1990). Premeiotic instability of repeated sequences in Neurospora crassa. Annu Rev Genet *24*, 579-613.

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

Singer, M.S., and Gottschling, D.E. (1994). TLC1: template RNA component of Saccharomyces cerevisiae telomerase. Science *266*, 404-409.

Webb, C.J., and Zakian, V.A. (2008). Identification and characterization of the Schizosaccharomyces pombe TER1 telomerase RNA. Nat Struct Mol Biol *15*, 34-42.

Xie, M., Mosig, A., Qi, X., Li, Y., Stadler, P.F., and Chen, J.J. (2008). Structure and function of the smallest vertebrate telomerase RNA from teleost fish. J Biol Chem *283*, 2049-2059.

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

Chapter 3

STRUCTURE AND EVOLUTION OF FILAMENTOUS FUNGAL TELOMERASE RNAS

3.1 Abstract

Neurospora crassa telomerase RNA is a surprisingly long transcript, 2,049nt for the mature RNA. Covariation and sequence divergence limits the ability to predict the secondary structure without comparative analysis of additional sequence. The ncTER1 sequence was quarried to identify 68 other putative filamentous fungal TER1 sequences. An ncTER1 secondary structure model was derived from the phylogenetic comparison of 69 filamentous fungal TER1 sequences. The filamentous fungal TER1s share several conserved structure features from both vertebrate and yeast TERs. All filamentous fungal TER1s harbor a highly conserved splicing signal for 3' end processing. Several filamentous fungal TER1 template variations were identified corresponding with their irregular telomeric DNA repeats, indicating the fast evolution of telomeric sequences within filamentous fungi. The Est1 protein, essential telomerase *in vivo* component in yeast, was also identified to be associated with active telomerase in *N. crassa*.

3.2 Introduction

Telomerase RNAs have been identified from ciliate, vertebrate, yeast, and plant species. The divergence in the TER sequence renders aligning evolutionarily distant groups of species impossible. Thus phylogenetic comparative analysis was performed for TERs from closely related species. The comparison of similar species can identify conserved sequences, regions and structures (Chen et al., 2000; Romero and Blackburn, 1991). RNA phylogenetic comparison employs covariation, whereby sequence may vary while the architecture of the RNA, base pairing, loops, and other structures are maintained between species. Therefore covariation provides significant supports for RNA helical structure and RNA overall secondary structure as well.

Telomerase RNA secondary structure models from three distantly related groups, ciliate, vertebrate and yeast, suggest two conserved and universal domains: the template-pseudoknot and the stem terminus element (Blackburn and Collins, 2010; Brown et al., 2007; Lin et al., 2004). The template-pseudoknot interacts with the TERT protein and is essential for telomerase activity (Lai et al., 2001). The single stranded template region contains a 5' template sequence that is reverse transcribed for telomere DNA synthesis and a 3' realignment sequence for the annealing to the telomeric DNA primer. Telomerase RNA template typically contains 1.5 repeat of the telomere complementary sequence, whereby the full repeat is the template and 0.5 repeat is the primer annealing region. The conserved pseudoknot is located nearby the template within the secondary structure, and is common to all telomerase RNAs so far identified. The

pseudoknot domain contains several conserved triple-helix A:U:U base pairings, conserved in all known telomerase RNAs and is essential for telomerase activity (Shefer et al., 2007; Theimer et al., 2005). The mechanism of how pseudoknot interacts with TERT and stimulates activity still remains elusive. A telomerase crystal structure would provide an explanation of the interactions, and expand our understanding of telomerase function.

The other conserved domain: the stem terminus element (STE) (Blackburn and Collins, 2010) is typically located a distance from the template-pseudoknot region. This element exists in ciliates as a stem loop structure, while a three-way junction structure is identified in yeast and vertebrate TER (in vertebrate TER, it is usually named as CR4-CR5 or P6/6.1, fig 1.1). In ciliates and vertebrates, the STE motif binds to TERT independently from the TERT and template-pseudoknot interaction (O'Connor et al., 2005). The STE-TERT interactions significantly increases telomerase activity in ciliates (O'Connor and Collins, 2006; Richards et al., 2006) and is essential for vertebrate telomerase activity (Mitchell and Collins, 2000). Recombinant TERT combined with *in vitro* transcribed template-pseudoknot and STE RNA fragments *in trans* can reconstitute telomerase enzymatic activity *in vitro*. The yeast telomerase RNA STE (three-way junction), however, is not essential for telomerase activity either *in vitro* or *in vivo* (Zappulla et al., 2005), suggesting an alternative telomerase enzymatic mechanism.

Beyond these two conserved domains, telomerase RNAs from distantly related groups of species exhibit group specific conserved regions. The ScaRNA

domain is highly conserved in vertebrate TER, which is believed to be acquired from ScaRNA family during evolution along the vertebrate lineage (Chen et al., 2000; Jady et al., 2004; Mitchell et al., 1999; Xie et al., 2008). Although the scaRNA domain is not required for telomerase activity reconstitution *in vitro*, it is crucial for telomerase RNA 3' end processing and telomerase RNP biogenesis *in vivo* (Jady et al., 2004; Mitchell et al., 1999). The budding yeast TLC1 has binding sites for Est1, the Ku heterodimer and the Sm protein complex, which are all essential for telomerase RNP biogenesis (Evans et al., 1998; Seto et al., 1999; Zhou et al., 2000). The divergence and the numerous of telomerase biogenesis mechanisms indicates a rapid evolution of the telomerase RNA and the RNP.

For the prediction and analysis of ncTER1 structural features, 68 additional filamentous fungal TER1 sequences were identified through a bioinformatic search. The phylogenetic comparative analysis of 69 filamentous fungal TER1 sequences identified conserved regions and a secondary structural model. A conserved template-pseudoknot domain and a three-way junction domain (named P6/6.1 for the similarity to the vertebrate namesake) were identified. A spliceosome-mediated 3' processing mechanism was found to be conserved among all filamentous fungal TERs sequences identified. Est1 protein, but not Ku or Sm proteins, was found to be associated with the active *N. crassa* telomerase *in vivo*. The structural analysis of filamentous fungal TERs provides important insights into telomerase RNA and RNP evolution and biogenesis.

3.3 Materials and Methods

3.3.1 Filamentous fungal TER1 sequence alignment and ncTER1 secondary structure prediction

Filamentous fungal genome sequences were obtained from various resources (Table 3.1). Blast of filamentous fungi TER1 was first carried out by using ncTER1 as the query sequence for genome sequence blast with E value setting between 1 and $1x10^{-3}$. The newly identified TER1 sequence from a new class will then be used for blast of more TER1 sequences from the same class. Sequence alignment was carried out using BioEdit with ClustalW Multiple alignment. Manually adjustment was performed according to the base pairing covariations. ncTER1 secondary structure of conserved regions was drawing according to the sequence alignment, and supported by base pairing covariations.

3.3.2 Genome walking for identification of TER1 from *Phymatotrichum omnivorum*

Phymatotrichum omnivorum (Strain name: OKAlf8) genomic DNA was obtained from Dr. Stephen Marek of Oklahoma State University. Genome walking was performed for *P.omnivorum ter1* using the Universal GenomeWalker Kit (ClonTech) following manufacturer's instruction.

3.3.3 Culturing of *Aspergillus nidulans*

Wild-type *Aspergillus nidulans* strain (FGSC A4) was obtained from fungal genetics stock center (FGSC) and grown under the condition described previously (Bhattacharyya and Blackburn, 1997). Briefly, *A. nidulans* was inoculated in solid YAG medium (5g/l yeast extract, 20g/l glucose, 15g/l agar and 1ml/l of *N. crassa* trace element solution) for 3 days at 37^{o}C. Conidia were harvested as described

for *N. crassa* and were inoculated in liquid YG medium (YAG medium but omit agar) for overnight with gentle shaking at 37^{o}C. Mycelia were harvested with the same procedure as *N. crassa* mycelia collection.

3.3.4 Culturing of *Mycosphaerella graminicola*

Mycosphaerella graminicola wildtype strain IPO323 was obtained from Dr. Gert Kema and grown in yeast glucose broth (YGB; 1% yeast extract, 3% glucose) at 18^{o}C with gentle shaking. *M.graminicola* cells were harvested by centrifugation of liquid culture at 1,500g for 10 min.

3.3.5 Plasmids construction

N. crassa SmD3, Ku80 and Est1 protein coding genes were amplified through RT-PCR using gene-specific primers, and were cloned into pCCG::C-Gly::3xFLAG vector (Honda and Selker, 2009) through multiple cloning sites. Linearized plasmids (*Stu* I or *Bsp* EI) were introduced into NC1 strain as described in materials and methods in chapter 2. Vogel minimum medium with 2 percent L-sorbose instead of sucrose was utilized for *his-3*$^{+}$ genotype colony screening. Western blot analysis was carried out for the screening of transformants expressing ectopic 3xFLAG tagged proteins.

3.3.6 RNA isolation

Total RNA isolation from *Aspergillus nidulans* and *Mycoshpaerella graminicola* was performed with the same procedure as *N. crassa* total RNA isolation described in the materials and methods of chapter 2. RNA isolation from cell extract was carried out by mixing 1ml Tri-reagent (Molecular research) per 50μl of cell lysate and following the manufacturer's instruction. To extract RNA

bound to anti-FLAG M2 affinity gel (Sigma), the gel was resuspended in 100μl RNA extraction buffer (20mM Tris-HCl pH 7.5, 10mM EDTA, and 0.5% SDS), following by acid phenol/chloroform (Applied Biosystems) extraction and ethanol precipitation. All precipitated RNA samples were dissolved in dH_2O, and concentration was measured by Nanodrop.

3.3.7 Cell lysate isolation

N. crassa whole cell lysate was obtained from ground mycelia tissue using CHAPS lysis buffer (10mM Tris-HCl pH 7.5, 400mM KCl, 1mM EDTA, 0.5% CHAPS, 10% glycerol, with 5mM β-mercaptoethanol and 1X protease inhibitor cocktail [Roche] added before use). Briefly, 500μl of ice cold CHAPS lysis buffer was added into 100mg of ground tissue and incubation was carried out for 5-10min at 4^oC with gentle rotation. Whole cell lysate was obtained by centrifugation at 16,000g for 5min, and was immediately utilized for immunoprecipitation or quick frozen in liquid nitrogen and stored at -80^oC.

3.3.8 Western blotting

Cell lysate or immunoprecipitated samples were heated at 90^oC for 10min with 1X Laemmli sample buffer (125mM Tris-HCl pH6.8, 2% SDS, 10% glycerol, 5% β-mercaptoethanol and 0.0025% bromophenol blue). Heated samples were resolved in 8% (for Ku80 and Est1 proteins) or 12% (for SmD3 protein) SDS-PAGE gels. After transferring to the PVDF membrane (Bio-rad), the membrane was blocked in 5% nonfat milk/1X TTBS (20 mM Tris-HCl, pH 7.5, 150 mM NaCl and 0.05% Tween 20) for 1 hour at room temperature with gentle rotation, followed by incubation with anti-FLAG M2 antibody (mouse monoclonal

antibody, Sigma) in 5% nonfat milk/1X TTBS for 1 hour at room temperature. The membrane was washed five times with 1X TTBS with 5 min each and blocked in 5% nonfat milk/1X TTBS for 10 min, then incubated with the HRP-conjugated goat-anti-mouse secondary antibody (Bio-Rad) in 5% nonfat milk/1X TTBS for 1 hour at room temperature. The membrane was washed four times with 1X TTBS with 5min each, and developed using the Immobilon Western Chemiluminescent HRP substrate (Millipore) following the manufacturer's instructions, and imaged using a Gel Logic440 system (Kodak).

3.3.9 Northern blotting

ncTER1 northern blotting analysis procedure was described in materials and methods of chapter 2. For *N. crassa* U2 snRNA northern blot analysis, total RNA isolated from cell lysate or immunoprecipitated samples was resolved in a 4% polyacrylamide 8M urea denaturing gel, and RNA was electro-transferred onto the Hybond-XL membrane with Owl VEP-3 large tank electroblotting system (Thermo Scientific Owl separation system) at 0.5 amp for 1 hour as shown in Figure 3.1. After transferring, RNA was UV crosslinked onto the membrane using optimal crosslink mode and hybridized with the same procedure as ncTER1 northern blot analysis.

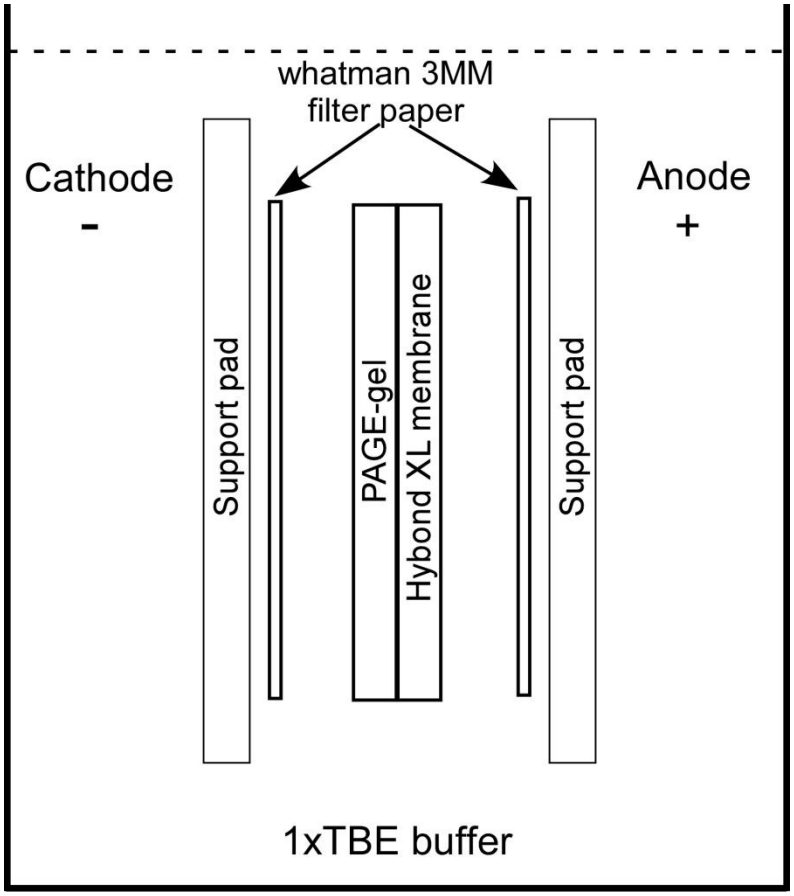


Figure 3.1 Electro-transferring of RNA onto Hybond XL membrane using Owl VEP-3 electroblotting system.

3L of 1xTBE buffer was utilized as transferring buffer. Two pieces of whatman 3MM filter paper, PAGE gel and one piece of Hybond XL membrane were assembled as indicated.

3.4 Results

3.4.1 Identification of 69 filamentous fungal TER1

The ncTER1 sequence is surprisingly long, thus it is impossible to predict its structure without a phylogenetic comparative analysis. The decreasing costs, the small genome size, and the importance of filamentous fungi resulted in a substantial number of the filamentous fungi genomes available. A bioinformatic search from several genome resources (listed in Table 3.1) found 69 putative

filamentous fungal TER sequences (including *N. crassa* TER1) within 5 classes from the phylum Ascomycota: 26 Sordariomycetes, 3 Leotiomycetes, 26 Eurotiomycetes, 12 Dothidiomycetes, and 2 Pezizomycetes (Figure 3.1). The *Phymatotrichum omnivorum* genome was sequenced for only 6-fold coverage at Oklahoma University (URL: http://www.genome.ou.edu/fungi.html) using 454 sequencing technology. The putative *P.omnivorum* TER sequence is, however, incomplete containing on a region from the template to the pseudoknot. Genome walking was performed to obtain the full length sequence for *Phymatotrichum omnivorum* TER.

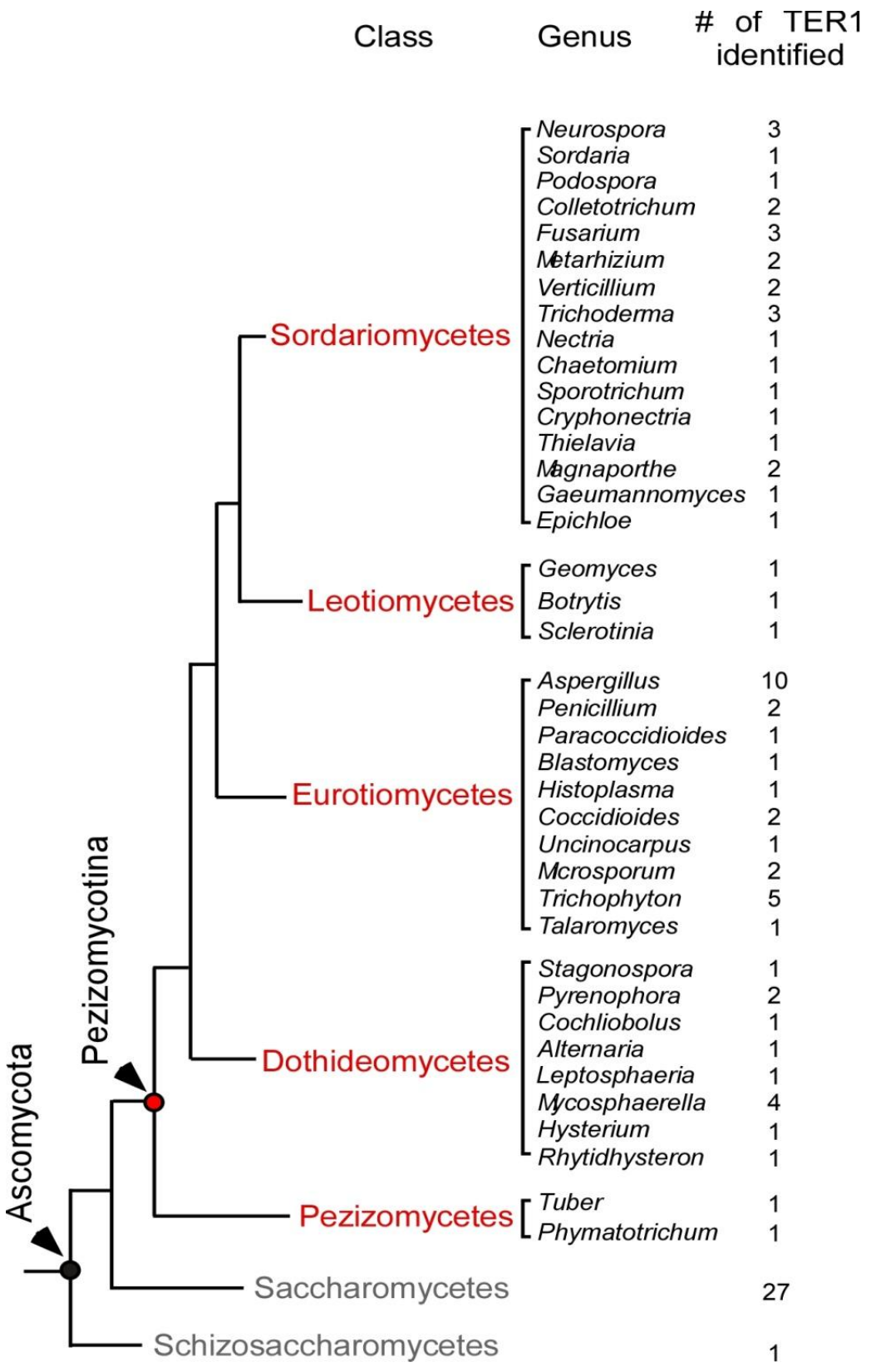

Class
Genus
# of TER1 identified
Ascomycota
Pezizomycotina
Sordariomycetes
Leotiomycetes
Eurotiomycetes
Dothideomycetes
Pezizomycetes
Saccharomycetes
Schizosaccharomycetes
Neurospora 3
Sordaria 1
Podospora 1
Colletotrichum 2
Fusarium 3
Metarhizium 2
Verticillium 2
Trichoderma 3
Nectria 1
Chaetomium 1
Sporotrichum 1
Cryphonectria 1
Thielavia 1
Magnaporthe 2
Gaeumannomyces 1
Epichloe 1
Geomyces 1
Botrytis 1
Sclerotinia 1
Aspergillus 10
Penicillium 2
Paracoccidioides 1
Blastomyces 1
Histoplasma 1
Coccidioides 2
Uncinocarpus 1
Microsporum 2
Trichophyton 5
Talaromyces 1
Stagonospora 1
Pyrenophora 2
Cochliobolus 1
Alternaria 1
Leptosphaeria 1
Mycosphaerella 4
Hysterium 1
Rhytidhysteron 1
Tuber 1
Phymatotrichum 1
27
1

Figure 3.2 Phylogenetic tree of the phylum Ascomycota.

Filamentous fungal TER1s identified in this study are from 5 classes: Sordariomycetes, Leotiomycetes, Eurotiomycetes, Dothidiomycetes, and Pezizomycetes, which are shown in red. The genus names from these five classes were also shown to each class. Number of the TER1 from each genus was listed. Telomerase RNAs were identified previously in the early branched classes, Saccharomycetes (Dandjinou et al., 2004; Gunisova et al., 2009; Hsu et al., 2007; Kachouri-Lafond et al., 2009; Seto et al., 2002; Singer and Gottschling, 1994) and Schizosaccharomycetes (Leonardi et al., 2008; Webb and Zakian, 2008), which are shown in grey. The evolution time is not drawn as proportion.

3.4.2 Filamentous fungal TER1s employ a conserved splicing signal for 3' end processing

Secondary structure modeling of filamentous fungal TERs is based on the phylogenetic comparison. The sequence alignment of Sordariomycetes and Leotiomycetes was performed independent of Eurotiomycetes, Dothidiomycetes and Pezizomycetes. The highly conserved regions within these two alignments served as anchor points for generating a comprehensive alignment for all filamentous fungal species. A conserved splicing signal was identified within all 69 filamentous fungal TERs, with the consensus sequence motif as: 5'-AUAAGU—A—YAG-3' (Fig 3.3), suggesting filamentous fungal TERs share a similar spliceosome-mediated 3' end processing mechanism as originally reported in *S. pombe* telomerase RNA (spTER1). The 5' splicing signal of spTER1 has the canonical sequence of 5'-GUAUGU-3', whereas the highly conserved 5' splicing signal in filamentous fungal TERs is a non-canonical sequence of 5'-AUAAGU-3'. As in *S. pombe*, the *N. crassa* mature TER1 terminates just upstream of the 5' splicing signal upon spliceosome-mediated cleavage. The mature form for the additional filamentous fungal TER is proposed to terminate immediately upstream

of this highly conserved 5’ splicing site. Further evidence to support this putative splicing site will come from the verification of additional filamentous fungal TERs and the 3’ ends.

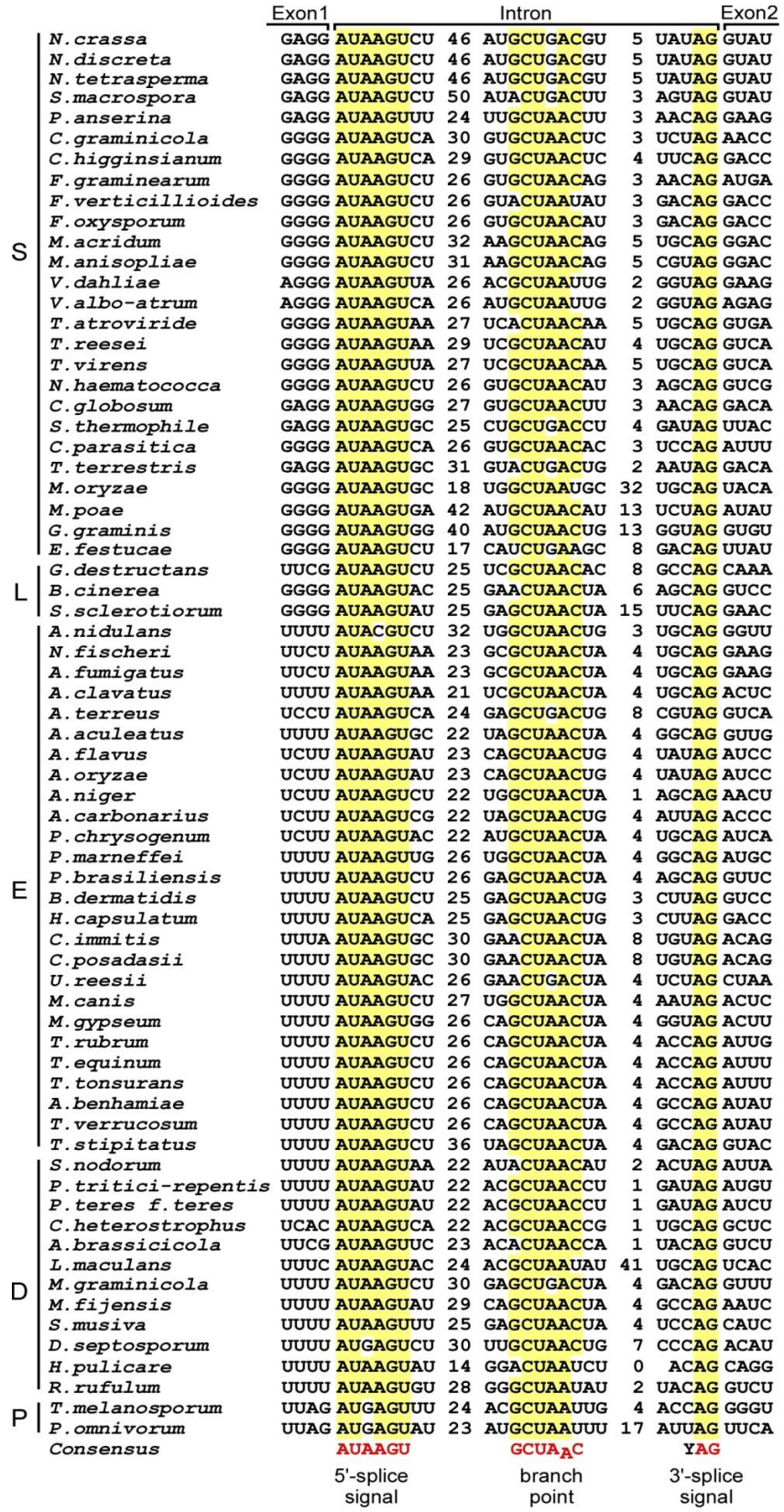

Exon1
Intron
Exon2
S
L
E
D
P
N.crassa GAGG AUAAGUCU 46 AUGCUGACGU 5 UAUAG GUAU
N.discreta GAGG AUAAGUCU 46 AUGCUGACGU 5 UAUAG GUAU
N.tetrasperma GAGG AUAAGUCU 46 AUGCUGACGU 5 UAUAG GUAU
S.macrospora GAGG AUAAGUCU 50 AUACUGACUU 3 AGUAG GUAU
P.anserina GAGG AUAAGUUU 24 UUGCUAACUU 3 AACAG GAAG
C.graminicola GGGG AUAAGUCA 30 GUGCUAACUC 3 UCUAG AACC
C.higginsianum GGGG AUAAGUCA 29 GUGCUAACUC 4 UUCAG GACC
F.graminearum GGGG AUAAGUCU 26 GUGCUAACAG 3 AACAG AUGA
F.verticillioides GGGG AUAAGUCU 26 GUACUAAUAU 3 GACAG GACC
F.oxysporum GGGG AUAAGUCU 26 GUGCUAACAU 3 GACAG GACC
M.acridum GGGG AUAAGUCU 32 AAGCUAACAG 5 UGCAG GGAC
M.anisopliae GGGG AUAAGUCU 31 AAGCUAACAG 5 CGUAG GGAC
V.dahliae AGGG AUAAGUUA 26 ACGCUAAUUG 2 GGUAG GAAG
V.albo-atrum AGGG AUAAGUCA 26 AUGCUAAUUG 2 GGUAG AGAG
T.atroviride GGGG AUAAGUAA 27 UCACUAACAA 5 UGCAG GUGA
T.reesei GGGG AUAAGUAA 29 UCGCUAACAU 4 UGCAG GUCA
T.virens GGGG AUAAGUUA 27 UCGCUAACAA 5 UGCAG GUCA
N.haematococca GGGG AUAAGUCU 26 GUGCUAACAU 3 AGCAG GUCG
C.globosum GAGG AUAAGUGG 27 GUGCUAACUU 3 AACAG GACA
S.thermophile GAGG AUAAGUGC 25 CUGCUGACCU 4 GAUAG UUAC
C.parasitica GGGG AUAAGUCA 26 GUGCUAACAC 3 UCCAG AUUU
T.terrestris GAGG AUAAGUGC 31 GUACUGACUG 2 AAUAG GACA
M.oryzae GGGG AUAAGUGC 18 UGGCUAAUGC 32 UGCAG UACA
M.poae GGGG AUAAGUGA 42 AUGCUAACAU 13 UCUAG AUAU
G.graminis GGGG AUAAGUGG 40 AUGCUAACUG 13 GGUAG GUGU
E.festucae GGGG AUAAGUCU 17 CAUCUGAAGC 8 GACAG UUAU
G.destructans UUCG AUAAGUCU 25 UCGCUAACAC 8 GCCAG CAAA
B.cinerea GGGG AUAAGUAC 25 GAACUAACUA 6 AGCAG GUCC
S.sclerotiorum GGGG AUAAGUAU 25 GAGCUAACUA 15 UUCAG GAAC
A.nidulans UUUU AUACGUCU 32 UGGCUAACUG 3 UGCAG GGUU
N.fischeri UUCU AUAAGUAA 23 GCGCUAACUA 4 UGCAG GAAG
A.fumigatus UUCU AUAAGUAA 23 GCGCUAACUA 4 UGCAG GAAG
A.clavatus UUUU AUAAGUAA 21 UCGCUAACUA 4 UGCAG ACUC
A.terreus UCCU AUAAGUCA 24 GAGCUGACUG 8 CGUAG GUCA
A.aculeatus UUUU AUAAGUGC 22 UAGCUAACUA 4 GGCAG GUUG
A.flavus UCUU AUAAGUAU 23 CAGCUAACUG 4 UAUAG AUCC
A.oryzae UCUU AUAAGUAU 23 CAGCUAACUG 4 UAUAG AUCC
A.niger UCUU AUAAGUCU 22 UGGCUAACUA 1 AGCAG AACU
A.carbonarius UCUU AUAAGUCG 22 UAGCUAACUG 4 AUUAG ACCC
P.chrysogenum UCUU AUAAGUAC 22 AUGCUAACUA 4 UGCAG AUCA
P.marneffei UUUU AUAAGUUG 26 UGGCUAACUA 4 GGCAG AUGC
P.brasiliensis UUUU AUAAGUCU 26 GAGCUAACUA 4 AGCAG GUUC
B.dermatidis UUUU AUAAGUCU 25 GAGCUAACUG 3 CUUAG GUCC
H.capsulatum UUUU AUAAGUCA 25 GAGCUAACUG 3 CUUAG GACC
C.immitis UUUA AUAAGUGC 30 GAACUAACUA 8 UGUAG ACAG
C.posadasii UUUU AUAAGUGC 30 GAACUAACUA 8 UGUAG ACAG
U.reesii UUUU AUAAGUAC 26 GAACUGACUA 4 UCUAG CUAA
M.canis UUUU AUAAGUCU 27 UGGCUAACUA 4 AAUAG ACUC
M.gypseum UUUU AUAAGUGG 26 CAGCUAACUA 4 GGUAG ACUU
T.rubrum UUUU AUAAGUCU 26 CAGCUAACUA 4 ACCAG AUUG
T.equinum UUUU AUAAGUCU 26 CAGCUAACUA 4 ACCAG AUUU
T.tonsurans UUUU AUAAGUCU 26 CAGCUAACUA 4 ACCAG AUUU
A.benhamiae UUUU AUAAGUCU 26 CAGCUAACUA 4 GCCAG AUAU
T.verrucosum UUUU AUAAGUCU 26 CAGCUAACUA 4 GCCAG AUAU
T.stipitatus UUUU AUAAGUCU 36 UAGCUAACUA 4 GACAG GUAC
S.nodorum UUUU AUAAGUAA 22 AUACUAACAU 2 ACUAG AUUA
P.tritici-repentis UUUU AUAAGUAU 22 ACGCUAACCU 1 GAUAG AUGU
P.teres f.teres UUUU AUAAGUAU 22 ACGCUAACCU 1 GAUAG AUCU
C.heterostrophus UCAC AUAAGUCA 22 ACGCUAACCG 1 UGCAG GCUC
A.brassicicola UUCG AUAAGUUC 23 ACACUAACCA 1 UACAG GUCU
L.maculans UUUC AUAAGUAC 24 ACGCUAAUAU 41 UGCAG UCAC
M.graminicola UUUU AUAAGUCU 30 GAGCUGACUA 4 GACAG GUUU
M.fijensis UUUU AUAAGUAU 29 CAGCUAACUA 4 GCCAG AAUC
S.musiva UUUU AUAAGUUU 25 GAGCUAACUA 4 UCCAG CAUC
D.septosporum UUUU AUGAGUCU 30 UUGCUAACUG 7 CCCAG ACAU
H.pulicare UUUU AUAAGUAU 14 GGACUAAUCU 0 ACAG CAGG
R.rufulum UUUU AUAAGUGU 28 GGGCUAAUAU 2 UACAG GUCU
T.melanosporum UUAG AUGAGUUU 24 ACGCUAAUUG 4 ACCAG GGGU
P.omnivorum UUAG AUGAGUAU 23 AUGCUAAUUU 17 AUUAG UUCA
Consensus AUAAGU GCUAAC YAG
5'-splice signal
branch point
3'-splice signal

Figure 3.3 Sequence alignment of filamentous fungal TER1 splicing signal.

Four nucleotides upstream and downstream the splicing signal are shown and labeled as part of the exon 1 and exon 2. Nucleotides with 80% or more of identity are shaded in yellow. The consensus sequences of the 5' splice signal, branch point and 3' splice signal are also shown under the alignment. The 3' splice signal YAG are predicted as the first YAG sequence downstream of the branch point, except TER1s from species *N. crassa*, *A. nidulans* and *M. graminicola* in which they were verified by 3' RACE. Numbers between the nucleotides indicate nucleotides that are omitted due to the high sequence variation.

3.4.3 Identification of the longest telomerase RNA

Although the 3' end of the mature filamentous fungal TER1 could be reasonably predicted by the splicing signal alignment, extensive sequence variations among filamentous fungal TER1s prevented alignment of the 5' end. Thus the 5' end of the filamentous TER1s could not be accurately or confidently predicted. To overcome this difficulty, the sequence length from the template to the 3' end was measured, as listed in Table 3.1. We identified putatively the largest TER from *Mycosphaerella graminicola* with the longest length, 2,050 nt, from the template to the predicted 3' end, a size greater than the entire mature ncTER1 (Table 3.1). 5' and 3' RACE was employed to determine the length of *Mycosphaerella graminicola* and *Aspergillus nidulans*, another well characterized fungal model organism, TER1s. As previously observed in ncTER1, three RNA forms were identified in both TERs. The *A. nidulans* full length sequence is shown in Supplemental figure S3.5, and a predicted intron from the phylogenetic comparison is present at the 3' end. The longest telomerase RNA previously identified length was the *Plasmodium falciparum* TER, at ~2.2 kb (Chakrabarti et

al., 2007) by RNase protection assay. The mature *M. graminicola* TER1 is 2,425 nt, and is currently the largest telomerase RNA yet identified (Table 3.1).

Table 3.1

List of all filamentous fungal TER1s identified in this study

Class	Species	Template to 3' end (nt)	Full length (nt)	Sources
Sordariomycetes	*Neurospora crassa*	1608	2049[a]	JGI, Broad
	Neurospora discreta	1604	2032[b]	JGI
	Neurospora tetrasperma	1629	2065[b]	JGI
	Sordaria macrospora	1644		NCBI
	Podospora anserina	1358		NCBI
	Colletotrichum graminicola	1267		Broad
	Colletotrichum higginsianum	1280		Broad
	Fusarium graminearum	1278		Broad
	Fusarium verticillioides	1297		Broad
	Fusarium oxysporum	1303		Broad
	Metarhizium acridum	1197		NCBI
	Metarhizium anisopliae	1224		NCBI
	Verticillium albo-atrum	1290		Broad
	Verticillium dahilae	1288		Broad
	Trichoderma atroviride	1330		JGI
	Trichoderma reesei	1309		JGI
	Trichoderma virens	1266		JGI
	Nectria haematococca	1310		JGI
	Chaetomium globosum	1373		JGI, Broad
	Sporotrichum thermophile	1407		JGI
	Cryphonectria parasitica	1346		JGI
	Thielavia terrestris	1372		JGI
	Magnaporthe oryzae	1436		Broad
	Magnaporthe poae	1696		Broad
	Gaeumannomyces graminis	1571		Broad
	Epichloe festucae	1328		OU
Leotiomycetes	*Geomyces destructans*	1559		Broad
	Botrytis cinerea	1547		Broad
	Sclerotina sclerotiorum	1541		Broad
Eurotiomycetes	*Aspergillus nidulans*	1315	1584[a]	JGI, Broad
	Neosartorya fischeri	1755		Broad
	Aspergillus fumigatus	1763		Broad
	Aspergillus clavatus	1752		Broad
	Aspergillus terreus	1766		Broad
	Aspergillus aculeatus	1654		JGI
	Aspergillus flavus	1732		Broad
	Aspergillus oryzae	1734		Broad
	Aspergillus niger	1728		JGI, Broad
	Aspergillus carbonarius	1713		JGI
	Penicillium chrysogenum	1762		NCBI
	Penicillium marneffei	1567		NCBI
	Paracoccidioides brasiliensis	1993		NCBI
	Blastomyces dermatidis	2001		Broad
	Histoplasma capsulatum	1977		Broad
	Coccidioides immitis	1910		Broad
	Coccidioides posadasii	1902		Broad

Table 3.1 continue:

Class	Species	Template to 3' end (nt)	Full length (nt)	Sources
Eurotiomycetes	*Uncinocarpus reesii*	1863		Broad
	Microsporum gypseum	1829		Broad
	Microsporum canis	1883		Broad
	Trichophyton rubrum	1837		Broad
	Trichophyton equinum	1841		Broad
	Trichophyton tonsurans	1841		Broad
	Arthroderma benhamiae	1835		NCBI
	Trichophyton verrucosum	1836		NCBI
	Talaromyces stipitatus	1585		NCBI
Dothidiomycetes	*Stagonospora nodorum*	1711		JGI
	Pyrenophora tritici-repentis	1709		JGI
	Pyrenophora teres f. teres	1708		NCBI
	Cochliobolus heterostrophus	1709		JGI
	Alternaria brassicicola	1705		JGI
	Leptosphaeria maculans	1721		INRA
	Mycosphaerella graminicola	2050	2425[a]	JGI
	Mycosphaerella fijiensis	1794		JGI
	Septoria musiva	2030		JGI
	Dothistroma septosporum	1887		JGI
	Hysterium pulicare	1877		JGI
	Rhytidhysteron rufulum	1860		JGI
Pezizomycetes	*Tuber melanosporum*	1500		INRA
	Phymatotrichum omnivorum	1389[c]		OU

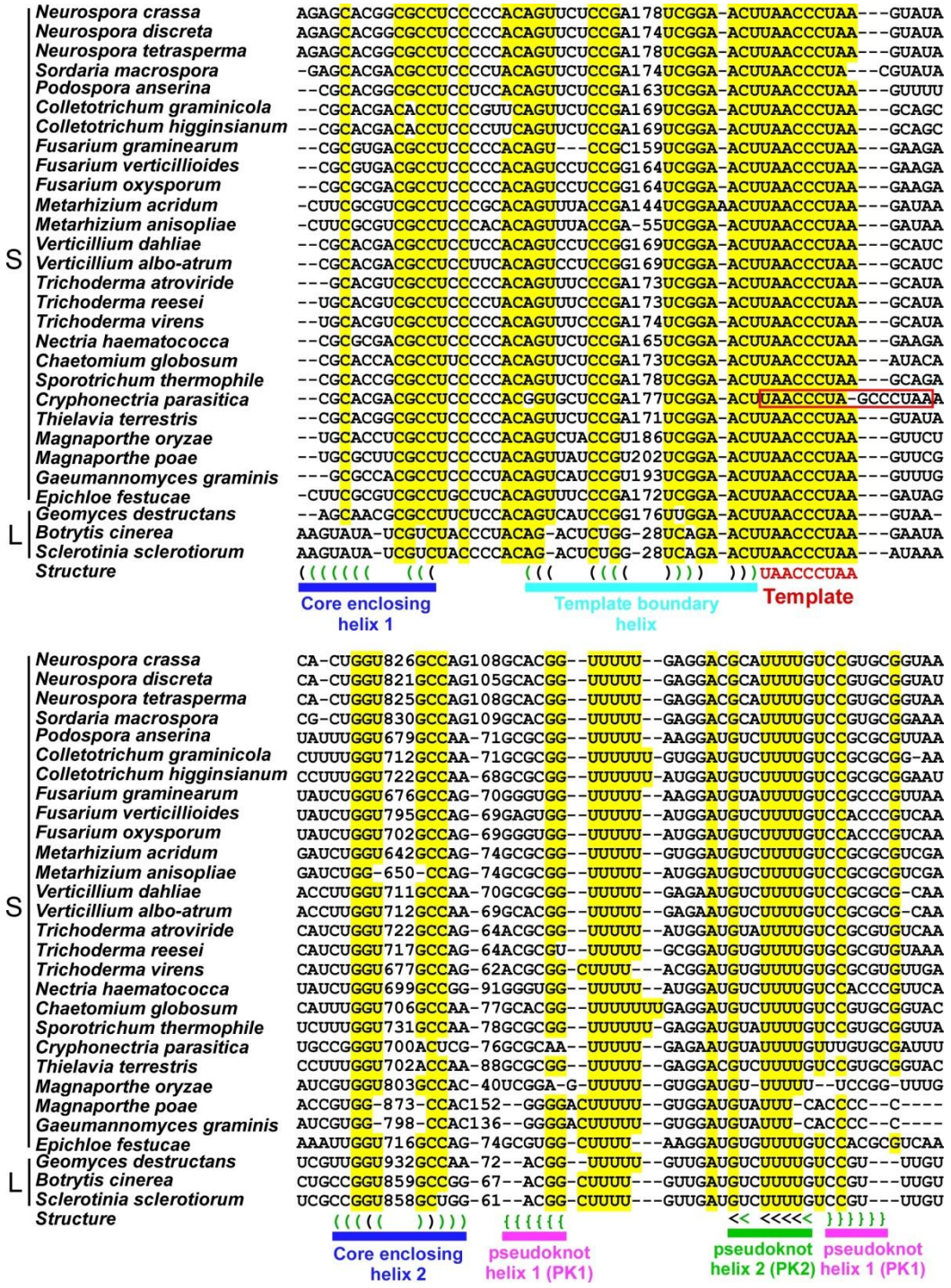
a: Full length mature TER1 confirmed by 5'/3' RACE
b: Full length mature TER1 predicted by alignment with ncTER1
c: Genome Walking performed for TER1
JGI: DOE Joint Genome Institute, URL: www.jgi.doe.gov
Broad: Broad institute of MIT and Harvard, URL: www.broadinstitute.org
NCBI: National center for Biotechnology Information, URL: www.ncbi.nlm.nih.gov
INRA: Institut national de la recherch é agronomique, URL: www.inra.fr
OU: Oklahoma University, URL: www.genome.ou.edu/fungi.html

3.4.4 Secondary structure of ncTER1 conserved domains

In addition to the splicing signal, two conserved regions: the template-pseudoknot and the P6/6.1 domains were also discovered. The template-pseudoknot domain alignment was performed for Sordariomycetes and Leotiomycetes TER sequences. The TER template within these two classes of species is highly conserved, with the consensus of 9 nt, 5'-UAACCCUAA-3', with the exception of *Cryphonectria parasitica*, initially predicted to have an 8 nt

template, 5’-UAACCCUA-3’. Further analysis, however, predicted sequences 5’-UAACCCUAGCCCUAA-3’ or 5’-CCCUAGCCCUA-3’ as template to synthesize telomere repeat: 5’-TTAGGGCTAGGG-3’ and 5’-CTAGGG-3’ respectively. Interestingly, within the genome sequencing data of *Cryphonectria parasitica*, multiple 5’-TTAGGGCTAGGG-3’ and 5’-CTAGGG-3’ repeated sequences can be found at the ends of several assembled scaffolds. Consecutive canonical 5’-TTAGGG-3’ repeats, however, were not identified within these repetitive sequences, suggesting a different template sequence within this region of the TER. Further experimentation, such as telomere cloning, must be performed for *C. parasitica* to confirm our hypothesis.

A conserved putative template boundary helix was predicted upstream of the ncTER1 template sequence, similar to yeast telomerase RNA. The prediction of a canonical pseudoknot structure containing a conserved A:U:U triple helix indicates a highly conserved motif within all fungal telomerase RNAs identified. Hairpin structures were also predicted within the pseudoknot with a highly conserved stem region towards the base. The truncation of the top variable region did not alter activity, however, complete deletion of these whole hairpins dramatically decreased activity (see chapter 4), suggesting these unique stem loops contribute to enzyme assembly.


S
Neurospora crassa
Neurospora discreta
Neurospora tetrasperma
Sordaria macrospora
Podospora anserina
Colletotrichum graminicola
Colletotrichum higginsianum
Fusarium graminearum
Fusarium verticillioides
Fusarium oxysporum
Metarhizium acridum
Metarhizium anisopliae
Verticillium dahliae
Verticillium albo-atrum
Trichoderma atroviride
Trichoderma reesei
Trichoderma virens
Nectria haematococca
Chaetomium globosum
Sporotrichum thermophile
Cryphonectria parasitica
Thielavia terrestris
Magnaporthe oryzae
Magnaporthe poae
Gaeumannomyces graminis
Epichloe festucae
L
Geomyces destructans
Botrytis cinerea
Sclerotinia sclerotiorum
Structure
Core enclosing helix 1
Template boundary helix
Template
S
Neurospora crassa
Neurospora discreta
Neurospora tetrasperma
Sordaria macrospora
Podospora anserina
Colletotrichum graminicola
Colletotrichum higginsianum
Fusarium graminearum
Fusarium verticillioides
Fusarium oxysporum
Metarhizium acridum
Metarhizium anisopliae
Verticillium dahliae
Verticillium albo-atrum
Trichoderma atroviride
Trichoderma reesei
Trichoderma virens
Nectria haematococca
Chaetomium globosum
Sporotrichum thermophile
Cryphonectria parasitica
Thielavia terrestris
Magnaporthe oryzae
Magnaporthe poae
Gaeumannomyces graminis
Epichloe festucae
L
Geomyces destructans
Botrytis cinerea
Sclerotinia sclerotiorum
Structure
Core enclosing helix 2
pseudoknot helix 1 (PK1)
pseudoknot helix 2 (PK2)
pseudoknot helix 1 (PK1)

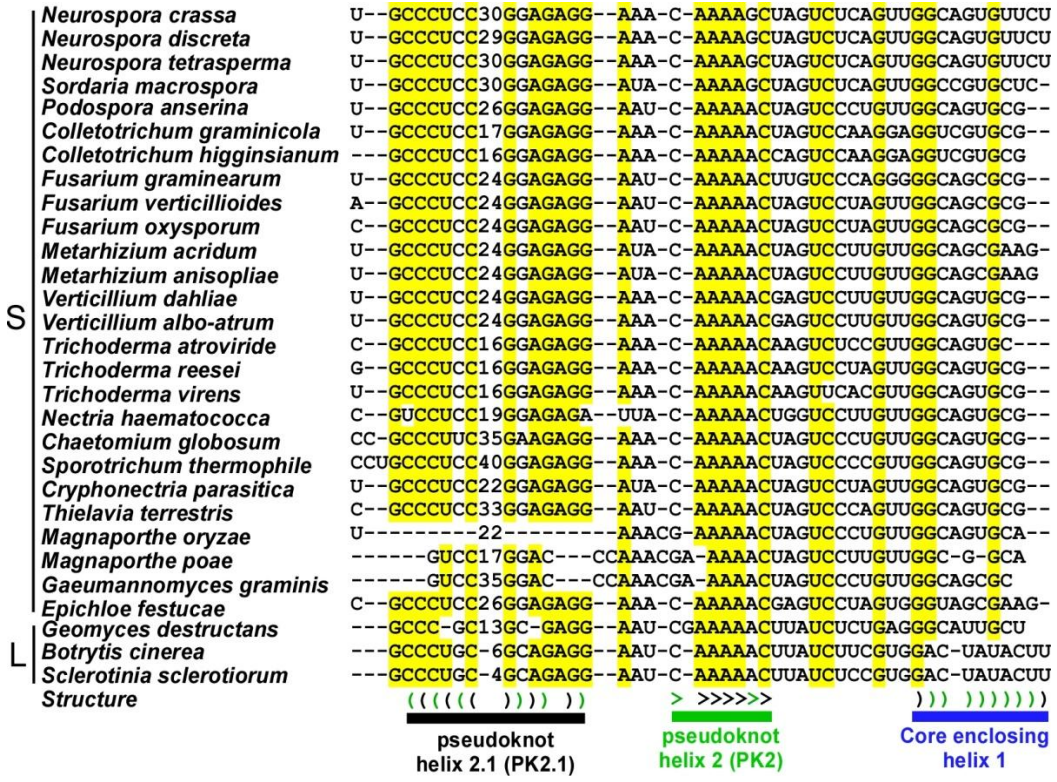


Figure 3.4 Phylogenetic comparison of template-pseudoknot region.

Alignment of template-pseudoknot region of telomerase RNAs from two classes of species: Sordariomycetes and Leotiomycetes. Nucleotides with 90% or more of identity are shaded in yellow. Numbers within the sequences indicate nucleotides that were omitted. Each structure feature shown in figure 3.6 are labeled under the sequences. The base pairings with covariations are shown in the structure with green parentheses. Putative template sequence of *Cryphonectria parasitica* is boxed. S: Sordariomycetes; L: Leotiomycetes.

Another conserved domain, P6/6.1, with similar structure to vertebrate P6/6.1 domain (CR4/CR5 domain) was predicted through the sequence alignment for all 69 filamentous fungal TERs. This region is highly conserved among all filamentous fungal species, as most of the TER1 sequences were identified by P6/6.1. The length of the filamentous fungal TER P6/6.1 is quite stringent, with most sequences around 38-39 nt, a similar length to teleost fish TER (Xie et al., 2008).

Figure 3.5 Phylogenetic comparison of P6/6.1.

Alignment of P6/6.1 region from filamentous fungal telomerase RNA sequence of 5 classes of species is shown. Similar sequences from the same genus are omitted for alignment. Nucleotides with 90% or more of identity are shaded in yellow. Structures are shown under the sequences. The base pairings with covariations are shown in the structure with green parentheses. S: Sordariomycetes; L: Leotiomycetes; E: Eurotiomycetes; D: Dothidiomycetes; P: Pezizomycetes.

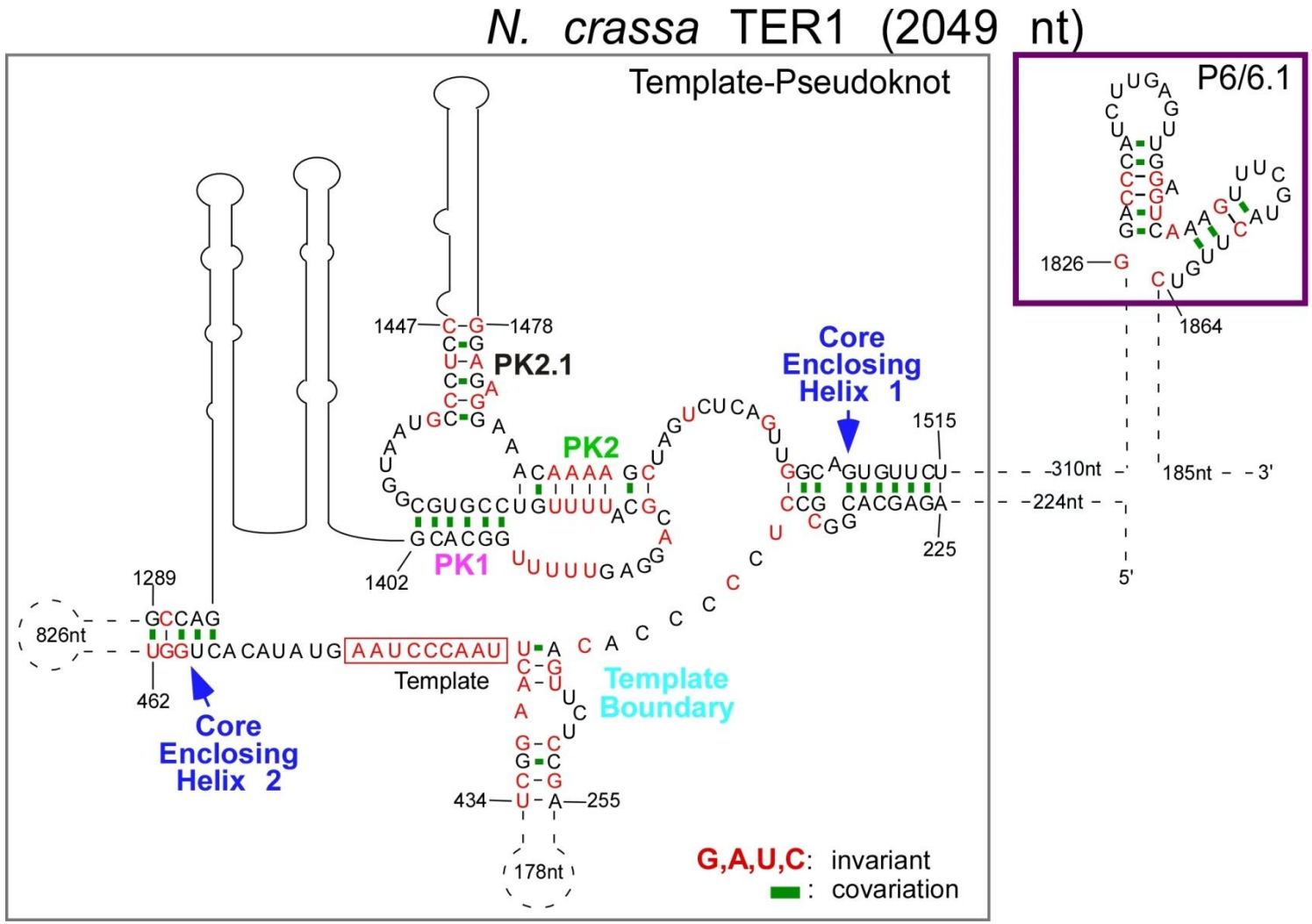

N. crassa TER1 (2049 nt)
Template-Pseudoknot
P6/6.1
PK2.1
PK2
PK1
Core Enclosing Helix 1
Core Enclosing Helix 2
Template
Template Boundary
1447
1478
1402
1289
826nt
462
1515
225
1826
1864
310nt
224nt
185nt
3'
5'
434
255
178nt
G,A,U,C: invariant
: covariation

Figure 3.6 Secondary structure of ncTER1 conserved regions

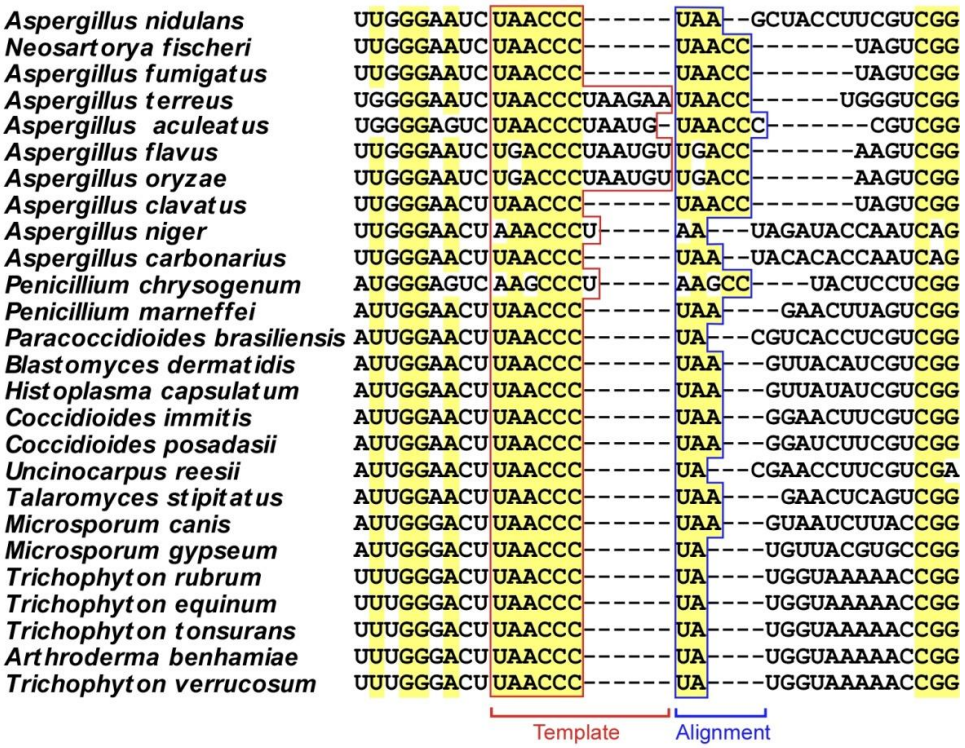
Proposed secondary structure of ncTER1 conserved regions 'Template-Pseudoknot' and 'P6/6.1' were determined by phylogenetic comparison as shown in figure 3.4 and 3.5. The 'Template-Pseudoknot' structure was determined by phylogenetic comparison of telomerase RNAs from two classes: Sordariomycetes and Leotiomycetes, while the 'P6/6.1' structure was determined from all filamentous fungal telomerase RNAs. Invariant nucleotides from the sequence alignment are shown with red color. The base pairs supported with covariations are shown as green filling boxes. The major structure features are shown: Core enclosing helix 1, template boundary, template, core enclosing helix 2, pseudoknot helix 1, 2, and 2.1 according to the sequence alignment shown in figure 3.4 and 3.5.

3.4.5 Eurotiomycetes telomeres and TER1 template sequences analysis

To determine if additional filamentous fungi contained non-canonical TER template and putative telomere sequences, as was observed in *Cryphonectria parasitica*, we aligned and compared template sequences from 69 filamentous fungal species. Surprisingly, we identified several putative non-canonical template sequences within the class of Eurotiomycetes, with the majority from the genus *Aspergillus*. The telomeric repeats were predicted from the template sequences and compared with telomeric repeats found within the genome sequencing data. The telomeric DNA of *Aspergillus oryzae* was previously reported to consist of a dodeca-nucleotide repeat: 5'-TTAGGGTCAACA-3' (Kusumoto et al., 2003). The sequence alignment predicted a template sequence of 5'-UGACCCUAAUGUUGACC-3' to synthesize a non-canonical telomere DNA repeat. The telomeric DNA repeats found within the *A. flavus* and *P. chrysogenum* genomes correlated to this putative template sequence. However, there were some variations within *A. aculeatus* and *A. niger* telomeric DNA

repeats, irregular for this template sequence. The telomeres from the strains with non-canonical, irregular telomere repeats must be cloned to verify our discovery.

A



B

Species	Predicted telomere repeat from template	Putative telomere repeat from genome
A.terreus	TTAGGG or TTAGGGTTATTC	N.A.
A.aculeatus	TTAGGG or TTAGGGTTACA	TTAGGGTTTACA
A.flavus *A.oryzae*	TTAGGGTCAACA	TTAGGGTCAACA
A.niger	TTTAGGG	T_{3-5}AGGG or T_{3-4}ATTAGGG
P.chrysogenum	TTAGGGC	TTAGGGC
C.parasitica	TTAGGG, CTAGGG or TTAGGGCTAGGG	CTAGGG or TTAGGGCTAGGG

Figure 3.7 Eurotiomycetes TER1 template sequence analysis

(A) Phylogenetic comparison of Eurotiomycetes TER1 template and flanking regions. The predicted template and alignment regions are indicated under the sequences. Nucleotides with 90% or more of identity are shaded in yellow. (B) List of filamentous fungi species with predicted and putative irregular telomere repeat sequences. *Aspergillus oryzae* and its telomere sequences are shown in red, as it was confirmed previously (Kusumoto et al., 2003). Irregular telomere repeat sequences were first predicted from the template alignment. And putative non-canonical telomere repeat sequences were obtained from genome sequencing. No telomere repeats were found for *Aspergillus terreus* in its genome database. *Cryphonectria parasitica* TER1 template and telomere repeat sequences were discussed previously in figure 3.4.

3.4.6 ncEst1 associates with active telomerase *in vivo*

Although catalytic TERT subunit and telomerase RNA are minimally required for reconstitution of activity *in vitro*, a variety of proteins are essential for telomerase stability, biogenesis, and localization to the telomeres *in vivo*. Since filamentous fungal TERs lack a conserved H/ACA domain, there is low probability for these RNAs to associate with dyskerin-like protein complex. The general structure similarity to yeast telomerase RNAs and similar 3' end processing as fission yeast TER1, prompted screening *N. crassa* telomerase for associated proteins homologous to yeast telomerase associated proteins: Sm proteins, Ku70/Ku80 heterodimer, and Est1 protein (Peterson et al., 2001; Seto et al., 1999; Zhou et al., 2000).

N. crassa SmD3, Ku70, Ku80, and Est1 gene sequences were obtained from the GenBank, and 5' and 3' RACE were performed to verify their 5' and 3' ends. Confirmed sequences identified from RACE and RT-PCR were submitted to GenBank. Ku80-3xFLAG was analyzed in this study to restore the heterodimer since the *N. crassa* strain for transformation is Ku80 deficient. C-terminal

3xFLAG tagged proteins were expressed under *ccg-1* promoter in *his-3* loci following the same procedure as previously described for the ncTER1 template mutants. TRAP activity assay was performed after each immunoprecipitation with anti-FLAG antibody beads. The Est1 immunoprecipitation had significant telomerase activity, while Ku70 and SmD3 did not. This indicates that only Est1 has stable association with the active telomerase enzyme *in vivo*.

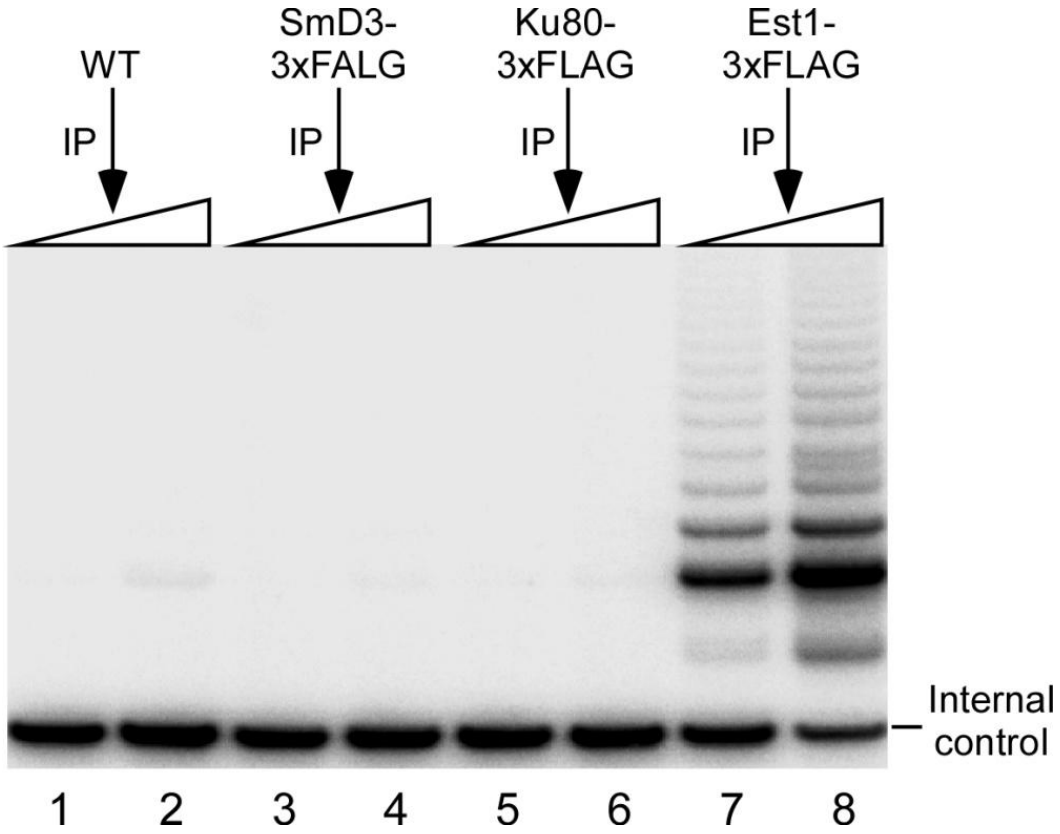


Figure 3.8 Telomerase TRAP assay post anti-FLAG IP.

After immuno-precipitation from wt, SmD3-3xFLAG, Ku80-3xFLAG and Est1-3xFLAG lysate (each was isolated from 100mg of mycelia tissue) with anti-FLAG antibody affinity gel, telomerase activity from each immuno-precipitated fraction (1:500 and 1:100 dilutions) was detected with TRAP assay as described in the materials and methods.

To further confirm Est1 is interacting specifically with the telomerase enzyme, northern blot analysis was performed. The RNA recovered from the immuno-precipitation was probed against ncTER1 and ncU2 snRNA. The Est1 protein coimmunoprecipitated ncTER1, whereas the SmD3 protein coimmunoprecipitated U2 snRNA. Thus it appears that neither Sm proteins nor Ku70/Ku80 heterodimer stably associate with ncTER1, in contrast to budding yeast telomerase RNA binding Sm and Ku and fission yeast telomerase RNA binding only Sm proteins. This result, however, supports the lack of a conserved Sm binding site within the phylogenetic comparison of filamentous fungal TER1s. A previous study revealed Ku70 or Ku80 deficient *N. crassa* strain did not show telomere length alteration (Ninomiya et al., 2004), suggesting the Ku heterodimer functions vary in budding yeast and other fungi species.

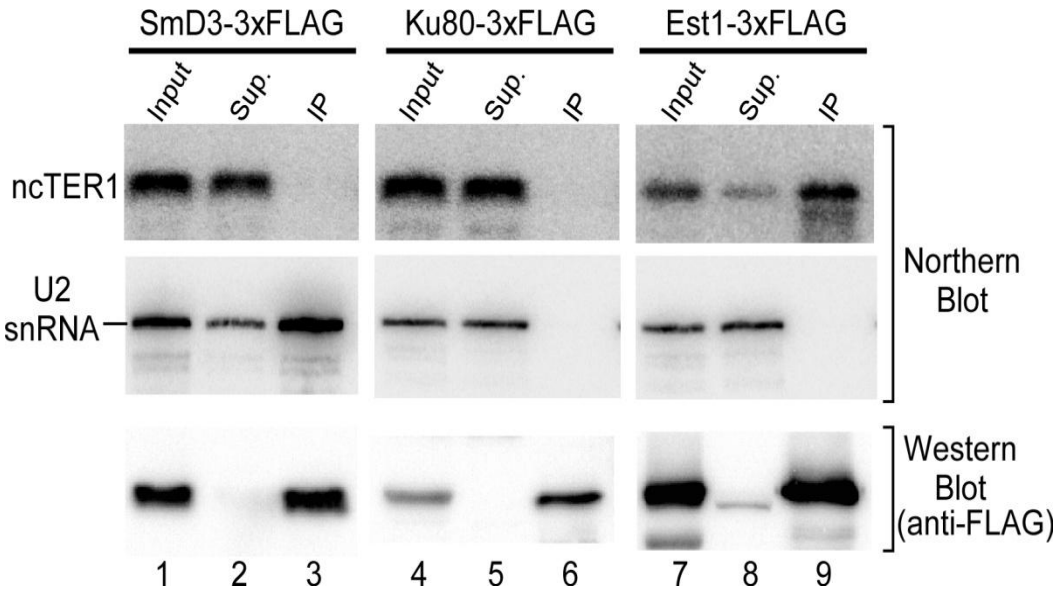


Figure 3.9 Northern blot and western blot after SmD3, Ku80 and Est1 immunoprecipitation. (See next page for figure legend)

Figure 3.9 Northern blot and western blot after SmD3, Ku80 and Est1 immunoprecipitation.

Northern blot and Western blot assay were performed after immuno-precipitation of each C-terminal 3xFLAG tagged protein as performed in Figure 3.8. ncTER1 and *N. crassa* U2 snRNA riboprobes were used for northern blot analysis. Anti-FLAG M2 antibody (Sigma) was used for western blot analysis.

3.4.5 ncTER1 is not 5' TMG capped

The immunoprecipitation of SmD3 indicated ncTER1 is not part of this small nuclear ribonucleoprotein complex (snRNP). This leads us to hypothesize ncTER1 is not 5'-2,2,7- trimethylguanosine (TMG) capped, since 5' TMG capping correlates to Sm protein binding. A 5' TMG cap was observed for both budding and fission yeast telomerase RNAs, and harbor an Sm binding site at the 3' end (Leonardi et al., 2008; Seto et al., 1999), whereas vertebrate telomerase RNAs employ a ScaRNA motif. To validate our hypothesis that ncTER1 lacks a 5' TMG cap, we attempted to immunoprecipate ncTER1 from total RNA with an anti-TMG antibody. The *N.crassa* U2 snRNA, a small nuclear RNA containing an Sm binding site and a 5' TMG cap (Tollervey and Mattaj, 1987), was recovered from the anti-TMG IP, whereas no ncTER1 signal was detected in the immunoprecipitated fraction. The supernatant fraction retained similar ncTER1 intensity, thus the ncTER1 remained intact. This suggests the mature ncTER1 lacks a 5' TMG cap. The precursor ncTER1 harboring a polyA tail implies ncTER1 is transcribed by RNA polymerase II (RNA pol II), similar to yeast and vertebrate telomerase RNAs. We thus predicted ncTER1 has a 7-methylguanosine cap; however, this has yet to be confirmed or disavowed. The RNA processing

mechanism appears novel in the filamentous fungi lineage, whereby the hypermethylated cap was lost during the evolution.

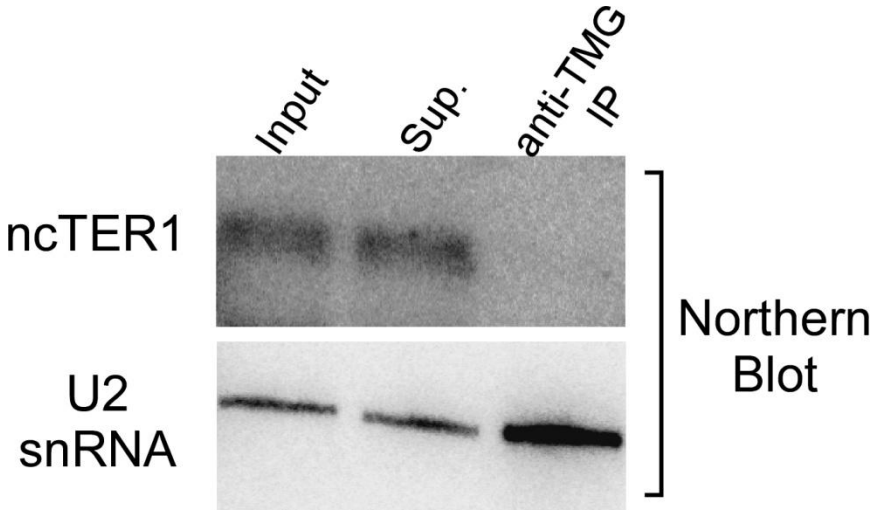


Figure 3.10 Anti-TMG immunoprecipitation from *N. crassa* total RNA

Anti-TMG antibody (Santa Cruz Biotechnology) immuno-precipitation was carried out with *N. crassa* total RNA. Northern blot analysis for ncTER1 and *N. crassa* U2 snRNA was shown with 1μg of total RNA as input and supernatant (Sup.) after IP, and 150ng of immuno-precipitated RNA.

3.5 Discussion

The sequences of 69 filamentous fungal putative TER1s have been identified within this study, providing strong supports from phylogenetic comparative analysis. The derived ncTER1 secondary structure affords novel insights into the structural evolution of filamentous fungal TER1s, a contrast to yeast TERs, and a more comprehensive view of telomerase RNP biogenesis.

3.5.1 Filamentous fungal TER1 3’ end processing

To date, telomerase RNAs identified from distinct groups of species exhibit striking evolutionarily divergent 3’ end processing. In ciliates, telomerase RNA is transcribed by RNA polymerease III (RNA pol III) (Romero and Blackburn,

1991), whereas in yeasts and vertebrates, telomerase RNAs are RNA pol II transcripts (Blasco et al., 1995; Chapon et al., 1997; Feng et al., 1995; Zaug et al., 1996). Yeast, including *S. pombe* and several Candida species, telomerase RNA precursors harbor a polyA tail, which is cleaved by a spliceosome mediated mechanism for TER maturation (Box et al., 2008; Gunisova et al., 2009). These yeast species universally contain a conserved 5' splicing signal sequence, 5'-GUAUGU-3', while filamentous fungi employ a non-canonical splicing signal sequence, 5'-AUAAGU-3', suggesting a unique evolution. The saccharomycetes species, including *S. cerevisiae*, however, lack a conserved splicing signal, suggesting this 3' processing pathway was lost during evolution. Although vertebrate telomerase RNAs are RNA pol II transcripts, they appear to be fused with a preexisting RNA domain: the H/ACA snoRNA. The 3' end processing of vertebrate telomerase RNAs undergoes a unique snoRNA biogenesis pathway with the 3' end cleaved 3 nucleotides downstream of the ACA motif and the snoRNA domain associating with the H/ACA binding protein complex, dyskerin protein complex (Chen et al., 2000; Mitchell et al., 1999). The cause for telomerase RNAs evolutionarily divergent 3' end processing pathway remains elusive, requiring investigation.

3.5.2 *N. crassa* TER1 secondary structure

The ncTER1 sequence contains a canonical 9 nt template sequence, 5'-UAACCCUAA-3', similar to vertebrate TERs. The predicted template boundary helix located immediately upstream of the template, indicates a structural and functional similarity to yeast TERs. Several non-canonical template sequences

corresponding to irregular telomere DNA repeats were identified in filamentous fungal TER1s. The majority of non-canonical template sequences were found in the *Aspergillus* genus, suggesting a rapid template and telomere DNA co-evolution. The TTAGGG telomere repeat sequence was discovered in several distantly related groups, including all vertebrates and a portion of invertebrates (Lejnine et al., 1995; Sinclair et al., 2007), a few plants (Adams et al., 2001), several protists such as *Trypanosoma brucei* (Blackburn and Challoner, 1984), and most filamentous fungi. Most other groups contain similar a TG rich simple repeat, such as TTTAGGG in most plants (Fuchs et al., 1995), and TTAGG in a majority of insects (Frydrychova et al., 2004; Okazaki et al., 1993; Sahara et al., 1999). This evidence suggests the TTAGGG-type telomere DNA repeat is the eukaryotic ancestral repeat. The early branching yeast species in ascomycetes contain irregular telomere repeats (Supplemental Table S2.1), likely due to their rapid evolution rate.

The pseudoknot structure is highly conserved within filamentous fungal TER1, containing A:U:U triple-helix similar across all telomerase RNAs. This indicates the conserved pseudoknot structure is ancient, a component of a common ancestral telomerase RNA. Although separated by more than 900 nt, secondary folding places the pseudoknot domain from filamentous fungal TER1 in close proximity to the template sequence, indicating the pseudoknot structure positions the template into telomerase active site (Zhang et al., 2010). The P6/6.1 (Three-way junction) is highly conserved among all filamentous fungal TER1s identified. Unlike other species, the length of this domain is quite conserved in

filamentous fungi, with most ranging from only 38 to 39nt, a small size similar to teloest fish TER. The P6.1 shares sequence and structural conservation with vertebrate TER P6.1, suggesting, unlike yeast TERs, the compacted P6/6.1 is essential for telomerase enzymatic activity, similar to the vertebrate TERs. The *in vitro N. crassa* telomerase reconstitution telomerase activity assay is required to verify this theory.

3.5.3 *N. crassa* telomerase RNP biogenesis

N.crassa telomerase RNP biogenesis

Est1 is the only common protein identified to associate with active telomerase from budding yeast, fission yeast, and *N. crassa*. Putative Est1 homologs in other filamentous fungal species were identified through a bioinformatic search (data not shown), suggesting the Est1 protein is part of the telomerase holoenzyme for all three major fungal groups. Studies in budding yeast telomerase RNP biogenesis indicate that Est1 is an essential telomerase holoenzyme component, directly associating with the TLC1 RNA (Zhou et al., 2000) and recruits telomerase to the telomeres via interactions with Cdc13 protein. Also in *S. pombe*, Est1 association with spTER1 is spTrt1 (*S. pombe* telomerase reverse transcriptase) independent (Leonardi et al., 2008; Webb and Zakian, 2008). Although an RNA binding domain has yet to be identified in either spEst1 or ncEst1, it is probable for Est1 in fission yeast and filamentous fungi to have a similar function as in budding yeast. The filamentous fungal TER1 binding domain for Est1 and ncEst1 RNA binding domain require further investigation to better understand the Est1-TER interactions.

The Ku70/Ku80 heterodimer associates with a conserved stem-loop structure within budding yeast species (Peterson et al., 2001), a separate pathway for telomerase localization to telomeres. However, Ku70/Ku80 is not observed in either *S. pombe* or *N. crassa* suggesting budding yeast TLC1 acquired this unique mechanism during evolution.

The TMG cap and Sm binding site are important for telomerase RNA stability and processing in both budding and fission yeast, however neither was found in ncTER1. Previous telomerase RNAs report proteins, or protein complexes, associate with the 3' end and are important for RNP processing and stability: P65 in ciliate, dyskerin protein complex in vertebrate and plant, and Sm proteins in yeast. As filamentous fungal TER1s are generally larger than all of other telomerase RNAs, and little of the total RNA sequences (~300 nt in length) are required for telomerase activity reconstitution *in vitro*, suggests TER1 has several potential regions for additional proteins binding and may bind a greater number of proteins compared to the telomerase RNAs from other groups. However, other than Est1 and the spliceosome signal, mediating 3' end processing, no other protein, or protein complex, binding sites to facilitate filamentous fungal telomerase RNP processing have been identified. The novelty in the sequence suggests a unique telomerase RNP biogenesis pathway is present in filamentous fungi. The identification of pathway critical for filamentous fungal RNP will provide insight into telomerase RNPs and other RNPs evolution.

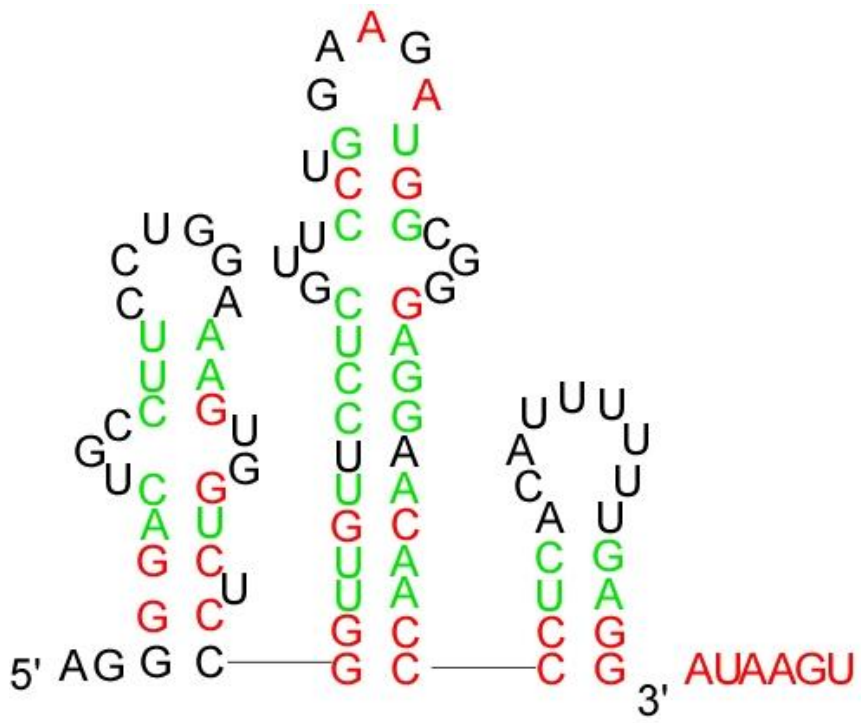


Figure 3.11 Predicted secondary structure of ncTER1 3' end region.

Nucleotides with 90% or more identities within Sordariomycetes TER1 are shown in red. Green nucleotides indicate the base pairings contain covariation. The AUAAGU sequence downstream of the 3' end is the 5' splice signal.

In budding yeast, Est3 is also a telomerase holoenzyme component essential for telomerase function *in vivo*, but is dispensable for *in vitro* catalytic activity (Lendvay et al., 1996; Lingner et al., 1997a). Due to the strong sequence variation of Est3 within different species, we were not able to identify its homolog in any filamentous fungal species. It would, however, be interesting to investigate if Est3 also involves in filamentous fungal telomerase RNP biogenesis.

3.5 References

Adams, S.P., Hartman, T.P., Lim, K.Y., Chase, M.W., Bennett, M.D., Leitch, I.J., and Leitch, A.R. (2001). Loss and recovery of Arabidopsis-type telomere repeat sequences 5'-(TTTAGGG)(n)-3' in the evolution of a major radiation of flowering plants. Proc Biol Sci *268*, 1541-1546.

Bhattacharyya, A., and Blackburn, E.H. (1997). Aspergillus nidulans maintains short telomeres throughout development. Nucleic Acids Res *25*, 1426-1431.

Blackburn, E.H., and Challoner, P.B. (1984). Identification of a telomeric DNA sequence in Trypanosoma brucei. Cell *36*, 447-457.

Blackburn, E.H., and Collins, K. (2010). Telomerase: An RNP Enzyme Synthesizes DNA. Cold Spring Harb Perspect Biol.

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

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

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

Chakrabarti, K., Pearson, M., Grate, L., Sterne-Weiler, T., Deans, J., Donohue, J.P., and Ares, M., Jr. (2007). Structural RNAs of known and unknown function identified in malaria parasites by comparative genomics and RNA analysis. RNA *13*, 1923-1939.

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

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

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

Evans, S.K., Sistrunk, M.L., Nugent, C.I., and Lundblad, V. (1998). Telomerase, Ku, and telomeric silencing in Saccharomyces cerevisiae. Chromosoma *107*, 352-358.

Feng, J., Funk, W.D., Wang, S.S., Weinrich, S.L., Avilion, A.A., Chiu, C.P., Adams, R.R., Chang, E., Allsopp, R.C., Yu, J., *et al.* (1995). The RNA component of human telomerase. Science *269*, 1236-1241.

Frydrychova, R., Grossmann, P., Trubac, P., Vitkova, M., and Marec, F. (2004). Phylogenetic distribution of TTAGG telomeric repeats in insects. Genome *47*, 163-178.

Fuchs, J., Brandes, A., and Schubert, I. (1995). Telomere sequence localization and karyotype evolution in higher plants Plant Systematics and Evolution *196*, 227-241.

Gunisova, S., Elboher, E., Nosek, J., Gorkovoy, V., Brown, Y., Lucier, J.F., Laterreur, N., Wellinger, R.J., Tzfati, Y., and Tomaska, L. (2009). Identification and comparative analysis of telomerase RNAs from Candida species reveal conservation of functional elements. RNA *15*, 546-559.

Honda, S., and Selker, E.U. (2009). Tools for fungal proteomics: multifunctional neurospora vectors for gene replacement, protein expression and protein purification. Genetics *182*, 11-23.

Hsu, M., McEachern, M.J., Dandjinou, A.T., Tzfati, Y., Orr, E., Blackburn, E.H., and Lue, N.F. (2007). Telomerase core components protect Candida telomeres from aberrant overhang accumulation. Proc Natl Acad Sci U S A *104*, 11682-11687.

Jady, B.E., Bertrand, E., and Kiss, T. (2004). Human telomerase RNA and box H/ACA scaRNAs share a common Cajal body-specific localization signal. J Cell Biol *164*, 647-652.

Kachouri-Lafond, R., Dujon, B., Gilson, E., Westhof, E., Fairhead, C., and Teixeira, M.T. (2009). Large telomerase RNA, telomere length heterogeneity and escape from senescence in Candida glabrata. FEBS Lett *583*, 3605-3610.

Kusumoto, K.I., Suzuki, S., and Kashiwagi, Y. (2003). Telomeric repeat sequence of Aspergillus oryzae consists of dodeca-nucleotides. Appl Microbiol Biotechnol *61*, 247-251.

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

Lejnine, S., Makarov, V.L., and Langmore, J.P. (1995). Conserved nucleoprotein structure at the ends of vertebrate and invertebrate chromosomes. Proc Natl Acad Sci U S A *92*, 2393-2397.

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

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

Lin, J., Ly, H., Hussain, A., Abraham, M., Pearl, S., Tzfati, Y., Parslow, T.G., and Blackburn, E.H. (2004). A universal telomerase RNA core structure includes

structured motifs required for binding the telomerase reverse transcriptase protein. Proc Natl Acad Sci U S A *101*, 14713-14718.

Lingner, J., Cech, T.R., Hughes, T.R., and Lundblad, V. (1997). Three Ever Shorter Telomere (EST) genes are dispensable for in vitro yeast telomerase activity. Proc Natl Acad Sci U S A *94*, 11190-11195.

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

Mitchell, J.R., and Collins, K. (2000). Human telomerase activation requires two independent interactions between telomerase RNA and telomerase reverse transcriptase. Mol Cell *6*, 361-371.

Ninomiya, Y., Suzuki, K., Ishii, C., and Inoue, H. (2004). Highly efficient gene replacements in Neurospora strains deficient for nonhomologous end-joining. Proc Natl Acad Sci U S A *101*, 12248-12253.

O'Connor, C.M., and Collins, K. (2006). A novel RNA binding domain in tetrahymena telomerase p65 initiates hierarchical assembly of telomerase holoenzyme. Mol Cell Biol *26*, 2029-2036.

O'Connor, C.M., Lai, C.K., and Collins, K. (2005). Two purified domains of telomerase reverse transcriptase reconstitute sequence-specific interactions with RNA. J Biol Chem *280*, 17533-17539.

Okazaki, S., Tsuchida, K., Maekawa, H., Ishikawa, H., and Fujiwara, H. (1993). Identification of a pentanucleotide telomeric sequence, (TTAGG)n, in the silkworm Bombyx mori and in other insects. Mol Cell Biol *13*, 1424-1432.

Peterson, S.E., Stellwagen, A.E., Diede, S.J., Singer, M.S., Haimberger, Z.W., Johnson, C.O., Tzoneva, M., and Gottschling, D.E. (2001). The function of a stem-loop in telomerase RNA is linked to the DNA repair protein Ku. Nat Genet *27*, 64-67.

Richards, R.J., Wu, H., Trantirek, L., O'Connor, C.M., Collins, K., and Feigon, J. (2006). Structural study of elements of Tetrahymena telomerase RNA stem-loop IV domain important for function. RNA *12*, 1475-1485.

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

Sahara, K., Marec, F., and Traut, W. (1999). TTAGG telomeric repeats in chromosomes of some insects and other arthropods. Chromosome Res *7*, 449-460.

Seto, A.G., Livengood, A.J., Tzfati, Y., Blackburn, E.H., and Cech, T.R. (2002). A bulged stem tethers Est1p to telomerase RNA in budding yeast. Genes Dev *16*, 2800-2812.

Seto, A.G., Zaug, A.J., Sobel, S.G., Wolin, S.L., and Cech, T.R. (1999). Saccharomyces cerevisiae telomerase is an Sm small nuclear ribonucleoprotein particle. Nature *401*, 177-180.

Shefer, K., Brown, Y., Gorkovoy, V., Nussbaum, T., Ulyanov, N.B., and Tzfati, Y. (2007). A triple helix within a pseudoknot is a conserved and essential element of telomerase RNA. Mol Cell Biol *27*, 2130-2143.

Sinclair, C.S., Richmond, R.H., and Ostrander, G.K. (2007). Characterization of the telomere regions of scleractinian coral, Acropora surculosa. Genetica *129*, 227-233.

Singer, M.S., and Gottschling, D.E. (1994). TLC1: template RNA component of Saccharomyces cerevisiae telomerase. Science *266*, 404-409.

Theimer, C.A., Blois, C.A., and Feigon, J. (2005). Structure of the human telomerase RNA pseudoknot reveals conserved tertiary interactions essential for function. Mol Cell *17*, 671-682.

Tollervey, D., and Mattaj, I.W. (1987). Fungal small nuclear ribonucleoproteins share properties with plant and vertebrate U-snRNPs. EMBO J *6*, 469-476.

Webb, C.J., and Zakian, V.A. (2008). Identification and characterization of the Schizosaccharomyces pombe TER1 telomerase RNA. Nat Struct Mol Biol *15*, 34-42.

Xie, M., Mosig, A., Qi, X., Li, Y., Stadler, P.F., and Chen, J.J. (2008). Structure and function of the smallest vertebrate telomerase RNA from teleost fish. J Biol Chem *283*, 2049-2059.

Zappulla, D.C., Goodrich, K., and Cech, T.R. (2005). A miniature yeast telomerase RNA functions in vivo and reconstitutes activity in vitro. Nat Struct Mol Biol *12*, 1072-1077.

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

Zhang, Q., Kim, N.K., Peterson, R.D., Wang, Z., and Feigon, J. (2010). Structurally conserved five nucleotide bulge determines the overall topology of the core domain of human telomerase RNA. Proc Natl Acad Sci U S A *107*, 18761-18768.

Zhou, J., Hidaka, K., and Futcher, B. (2000). The Est1 subunit of yeast telomerase binds the Tlc1 telomerase RNA. Mol Cell Biol *20*, 1947-1955.

Chapter 4

FUNCTIONAL STUDY OF N. CRASSA TELOMERASE RNA

4.1 Abstract

In the vast majority of eukaryotic cells, telomerase is the principle means of extending telomeric DNA found in linear chromosome ends. Telomerase *in vitro* activity assay was developed to study telomerase activity regulation. Herein, *Neurospora crassa* telomerase was reconstituted in rabbit reticulocyte lysate (RRL) from *in vitro* expressed ncTERT and *in vitro* transcribed ncTER1. *N. crassa* telomerase reconstituted *in vitro* generates significant telomerase activity by adding 6-nt repeat GGGTTA to the telomeric primer, similar to the human enzyme. Unlike yeast telomerase, *N. crassa* telomerase reconstituted *in vitro* exhibited highly processive repeat-addition activity. Although the repeat-addition processivity of wildtype *N. crassa* telomerase was initially less than human telomerase, the processivity was dramatically increased with the augmentation of ncTER1 template sequence, increasing the length of the realignment region. The template-pseudoknot and P6/6.1 RNA fragments of ncTER1 were sufficient and essential to reconstitute telomerase activity *in trans*, identical to *in vitro* behavior of vertebrate telomerase RNAs.

4.2 Introduction

Telomerase is a unique reverse transcriptase (RT) responsible for telomeric DNA 3' extension. Telomerase activity is essential for cell immortality, counterbalancing telomere erosion arising from cellular repilcation. Enzymatic activity was identified in the germ line, stem cells and approximately 90% of all human cancer cells (Kim et al., 1994), while little or no activity was observed in somatic cells. The lack of telomerase activity results in telomere attrition, leading to cell senescence. In order to study the telomerase enzymatic mechanism and telomerase regulation, *in vitro* reconstitution of telomerase activity is required. Two core components are essential for reconstitution of enzymatic activity *in vitro*: the catalytic protein component, telomerase reverse transcriptase (TERT), and the telomerase RNA (TER) (Mitchell and Collins, 2000).

Telomerase RNA contains a short template sequence reverse transcribed into telomeric DNA repeats. The template in all TERs includes a 5' region with one repeat complementary to telomeric DNA repeat for DNA synthesis, and a 3' region responsible for telomeric primer annealing, typically less than one complete telomeric DNA repeat. In ciliates and vertebrates, a striking feature of the enzyme is the addition of multiple telomeric DNA repeats to a given telomeric primer before complete enzyme-dissociation from the primer, denoted as repeat addition processivity (Figure 4.1). Conventional reverse transcriptases, such as HIV RT, have long RNAs as template and synthesize a long DNA before dissociation. After reaching the 5' end of template region, telomerase, employs a unique mechanism, whereby the DNA-RNA duplex dissociates and undergoes a

template translocation step before re-aligning the telomeric DNA primer to the template realignment region without complete releasing the DNA primer (Greider and Blackburn, 1989). During a telomerase activity assay, the telomeric primer can complete dissociate from telomerase during the translocation step halting DNA synthesis to the given primer, this results in a 6-nt ladder pattern of telomerase products. Remarkably, telomerase from most budding yeast species lack this unique repeat addition processivity. Only a single telomeric DNA repeat could be synthesized by yeast telomerase before complete dissociation from the primer (Cohn and Blackburn, 1995; Fulton and Blackburn, 1998).

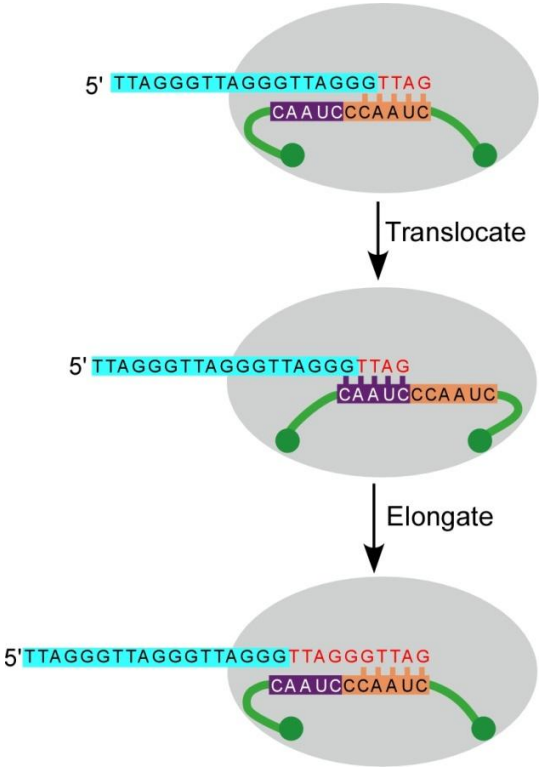


Figure 4.1 Schematic of telomerase translocation

Telomeric primer $(TTAGGG)_3$ (shaded with light blue) is elongated by telomerase to the hTER template 5' end (green line, template sequence is shown). The translocation step happens and telomeric primer anneals to the template re-alignment region (shaded with purple). Telomerase elongates the telomeric primer again. Nucleotides synthesized by telomerase are shown in red.

In human, two independent RNA-protein interactions (Pseudoknot and CR4-CR5 fragments) are essential for telomerase activity *in vitro* (Lai et al., 2001; Mitchell and Collins, 2000). Telomerase activity could be reconstituted *in vitro* with recombinant TERT protein and those two RNA fragments *in trans*. Similar mechanism was observed in ciliate telomerase that stem loop IV of TER significantly stimulate the telomerase activity and processivity *in trans* (Mason et al., 2003). In contrast, yeast TER three-way junction is not required for telomerase activity *in vitro* or *in vivo*, according to the truncation analysis (Zappulla et al., 2005).

In the previous chapters, filamentous fungal TER1 conserved secondary structure has been determined with both similarities to vertebrate and yeast TERs. In order to functionally analyze the *N. crassa* telomerase, the *in vitro* telomerase reconstitution system is developed for *N. crassa*. Similar to human telomerase, *N. crassa* telomerase reconstituted *in vitro* exhibits significant and processive enzymatic activity. Two truncated RNA fragments, template-pseudoknot and P6/6.1, could also reconstitute active *N. crassa* telomerase *in trans*.

4.3 Materials and Methods

4.3.1 Plasmid construction

To make a plasmid for expressing ncTERT in rabbit reticulocyte lysate (RRL), *N. crassa tert* was RT-PCR amplified from total RNA and cloned into pNFLAG-hTERT vector (generous gift from Dr. Vinayaka Prasad) through *Not* I and *Bgl* II sites to replace hTERT gene.

4.3.2 RNA manipulation

ncTER1 truncations and mutations were constructed by overlapping PCR. T7 promoter based RNA *in vitro* transcription was carried out using PCR DNA as template as described previously (Xie et al., 2008).

4.3.3 *In vitro* reconstitution of human and *N. crassa* telomerase

Human and *N. crassa* telomerase was reconstituted using TnT T7 quick coupled rabbit reticulocyte lysate system (RRL, Promega). *N. crassa* TERT was synthesized in 10μl RRL using pNFLAG-ncTERT vector following manufacturer's instruction. *In vitro* transcribed ncTER1 (0.1μM) or ncTER1 fragments (1μM each) were added into the recombinant ncTERT in RRL and incubated at 30^{o}C for 30min. Human telomerase was reconstituted in 10μl RRL with hTERT expressed from pNFLAG-hTERT vector and *in vitro* transcribed full length hTER (1μM final concentration) following the same procedure as *N. crassa* telomerase reconstitution.

4.3.4 Conventional telomerase activity direct assay

Direct telomerase activity assay was carried out for 60 min at 30 ℃ with 2 μL of *in vitro* reconstituted telomerase in a 10 μL reaction with 1x PE buffer (50mM Tris-HCl pH 8.3, 2mM DTT, 0.5mM $MgCl_2$, and 1mM spermidine), 1 mM dTTP, 1 mM dATP, 5 μM dGTP, 0.165 μM α-^{32}P-dGTP (3000 Ci/mmole, 10 mCi/mL, PerkinElmer), and 1 μM $(TTAGGG)_3$ telomeric primer. The reaction was then terminated by phenol/chloroform extraction, followed by ethanol precipitation. The reaction products were resolved onto a 10% polyacrylamide 8M urea denaturing gel, which was dried and analyzed by a Phosphor Imager

(Bio-Rad). Quantitation of telomerase activity and repeat addition processivity was described previously (Xie et al., 2010).

4.4 Results

4.4.1 *N. crassa* telomerase reconstituted *in vitro* is active and processive

In ciliate, yeast, and vertebrate, the telomerase core enzyme only consists of catalytic TERT protein plus the telomerase RNA subunit, as shown with enzyme activity reconstitution *in vitro*. Previous study showed that the 1.2-kilobase *S. cerevisiae* TLC1 RNA failed to reconstitute activity with yeast TERT protein *in vitro*, probably due to the large RNA misfolding, as the truncated TLC1 RNA (mini-T) exhibit robust telomerase activity upon reconstitution *in vitro* (Zappulla et al., 2005). To our surprise, the *in vitro* transcribed 2,049nt ncTER1 mature RNA, when reconstituted with ncTERT in Rabbit Reticulocyte Lysate (RRL), yielded significant and processive telomerase activity similar to human telomerase activity (Fig 4.2). Human telomerase synthesizes GGTTAG DNA repeats to the telomeric primer. The major bands from the *N. crassa* telomerase activity were one nucleotide below that from human telomerase activity, further confirming 5'-UAACCCUAA-3' is utilized as template to synthesize telomeric DNA repeats, GGGTTA. The predicted +3nt band was not visible from *N. crassa* telomerase activity, resulting from that no α-^{32}P-dGTP was incorporated for the first repeat as predicted. The *in vitro* transcribed precursor and spliced ncTER1 with 10 Adenosine at the 3' end to mimic the polyA tail also generated visible but much lower telomerase enzymatic activity, with the same pattern as that from mature RNA. This suggested that the exon2 or exon2+intron sequences might

inhibit telomerase activity by disrupting the secondary structure. The processivity of *N. crassa* telomerase reconstituted *in vitro*, however, was apparently lower than that of human telomerase, which was probably due to the short length of ncTER1 template (9nt vs 11nt in hTER). To test our hypothesis, we constructed three ncTER1 template mutants by extending the template by 1 or 2 nucleotides (Fig 4.3).

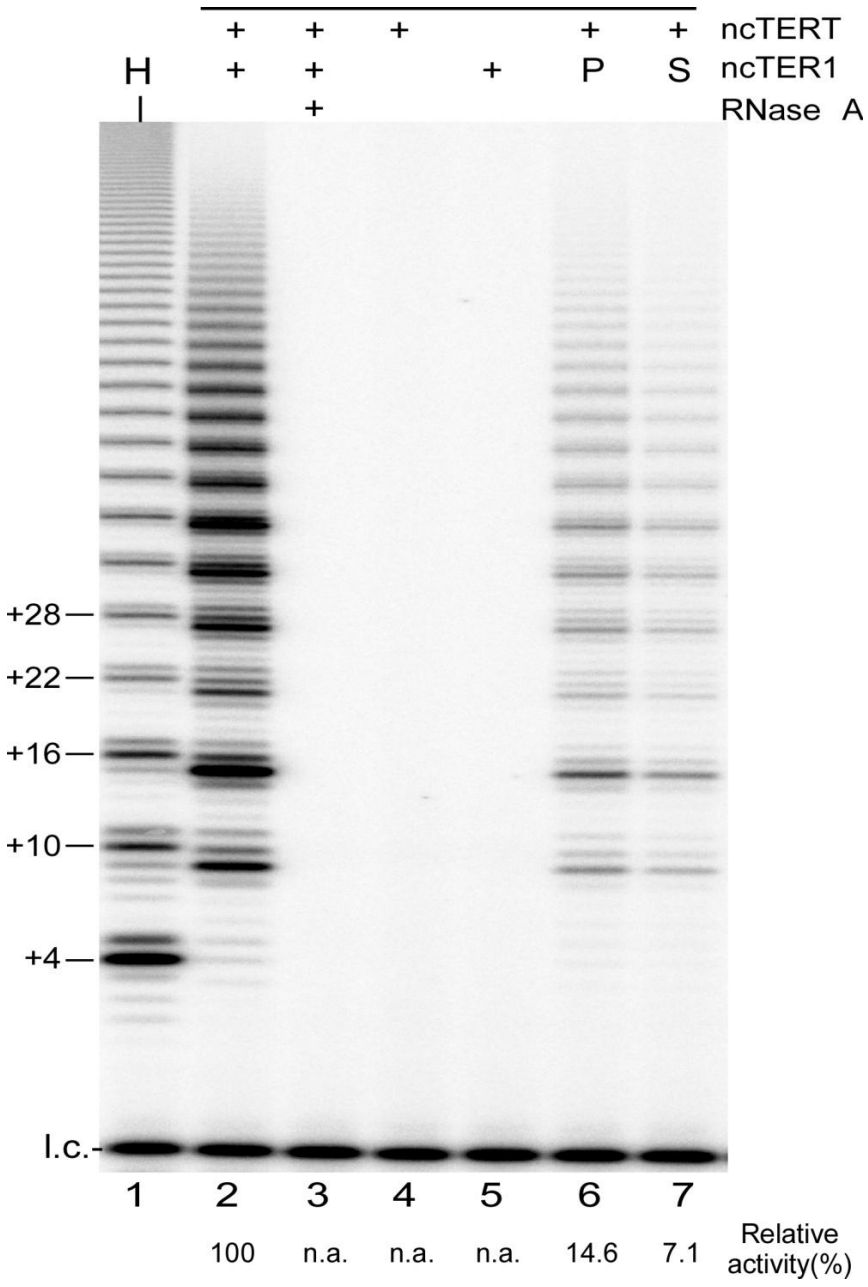


Figure 4.2 Human and *N. crassa* telomerase activity assay reconstituted in RRL

Telomerase activity was reconstituted in Rabbit Reticulocyte Lysate with telomeric primer $(TTAGGG)_3$. Lane1: hTERT+hTER full length, lane 2: ncTERT+ncTER1, lane 3: ncTERT+ncTER1 treated with RNase A, lane4: ncTERT only, lane5: ncTER1 only, lane 6: ncTERT+ncTER1 precursor form, lane 7: ncTERT+ncTER1 spliced form. *In vitro* transcribed *N. crassa* telomerase RNAs were added with 0.1uM concentration for reconstitution whereas hTER was 1uM. L.c.: ^{32}p end-labeled 15mer loading control. Relative activities as normalized to mature ncTER1 (lane 2) are shown under the gel image.

4.4.2 ncTER1 with longer template alignment exhibit higher repeat addition processivity

Three ncTER1 template mutants were constructed and *in vitro* transcribed for *N. crassa* telomerase reconstitution (Fig 4.3). The t1 mutant was designed to insert a single 'C' nucleotide upstream of the template sequence. As expected, the major bands of the activity were shifted one nucleotide up comparing to that of wildtype enzyme. The t2 and t3 mutants contained one or two 'C' nucleotide insertion downstream of the template sequence respectively. All three ncTER1 template mutants exhibited significant repeat addition processivity increase comparing to the wildtype enzyme, with the most increase shown in t3 mutant which harbored 2nt insertion. This could be explained that the longer template provides larger template realignment region which stabilizes the DNA-RNA duplex during the telomerase translocation event, consistent with the result seen in human and mouse telomerase RNA (Chen and Greider, 2003). Comparing all filamentous fungal TER1 template region, most of them use only 9nt template 5'-UAACCCUAA-3' with several to be only 8nt template 5'-UAACCCUA-3'. The phenomena may correlate with the short telomere length discovered in filamentous fungal species. Their telomerase may evolve to be less processive resulting from no requirement to synthesize long telomere repeats during cell reproduction.

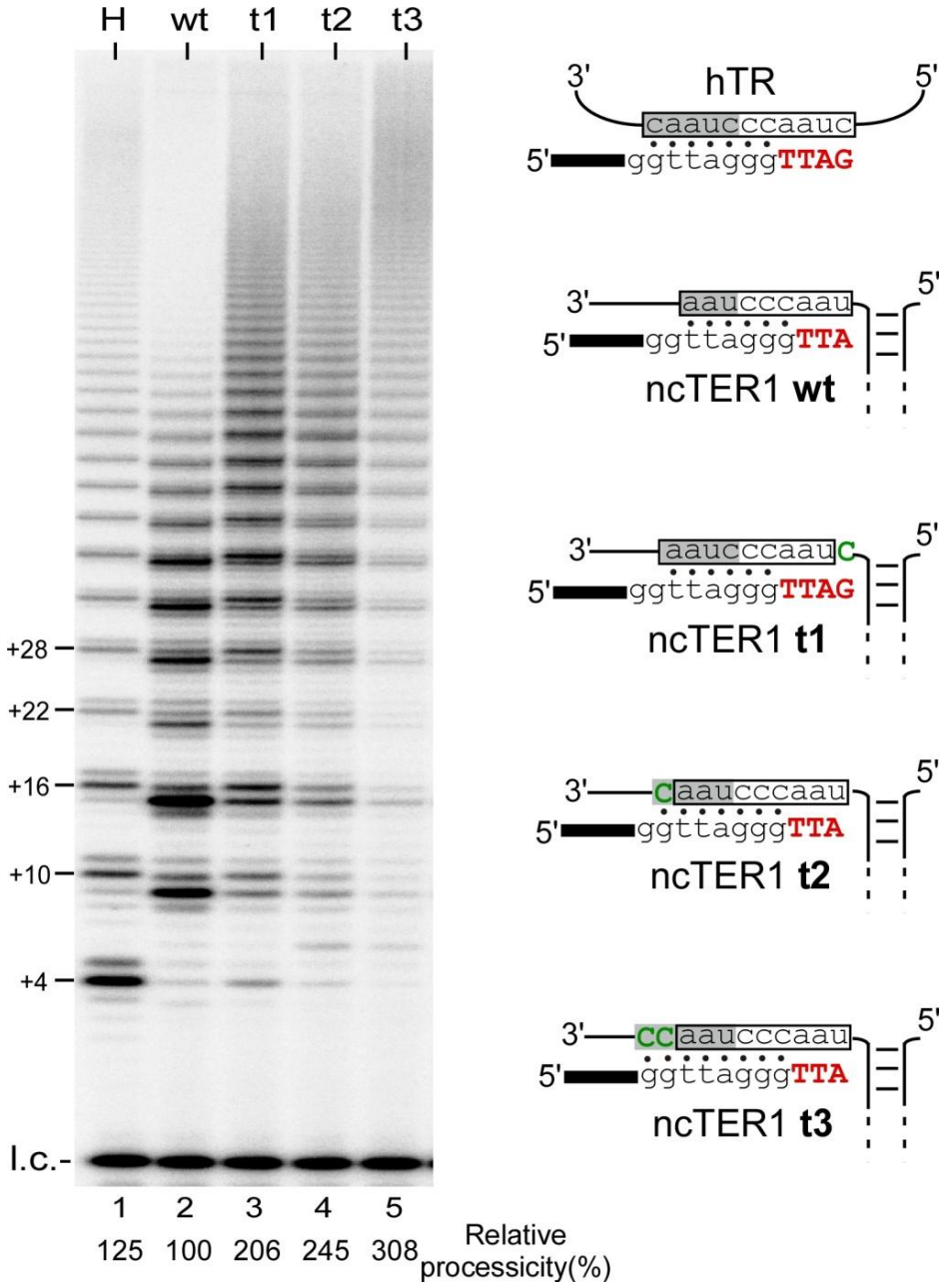


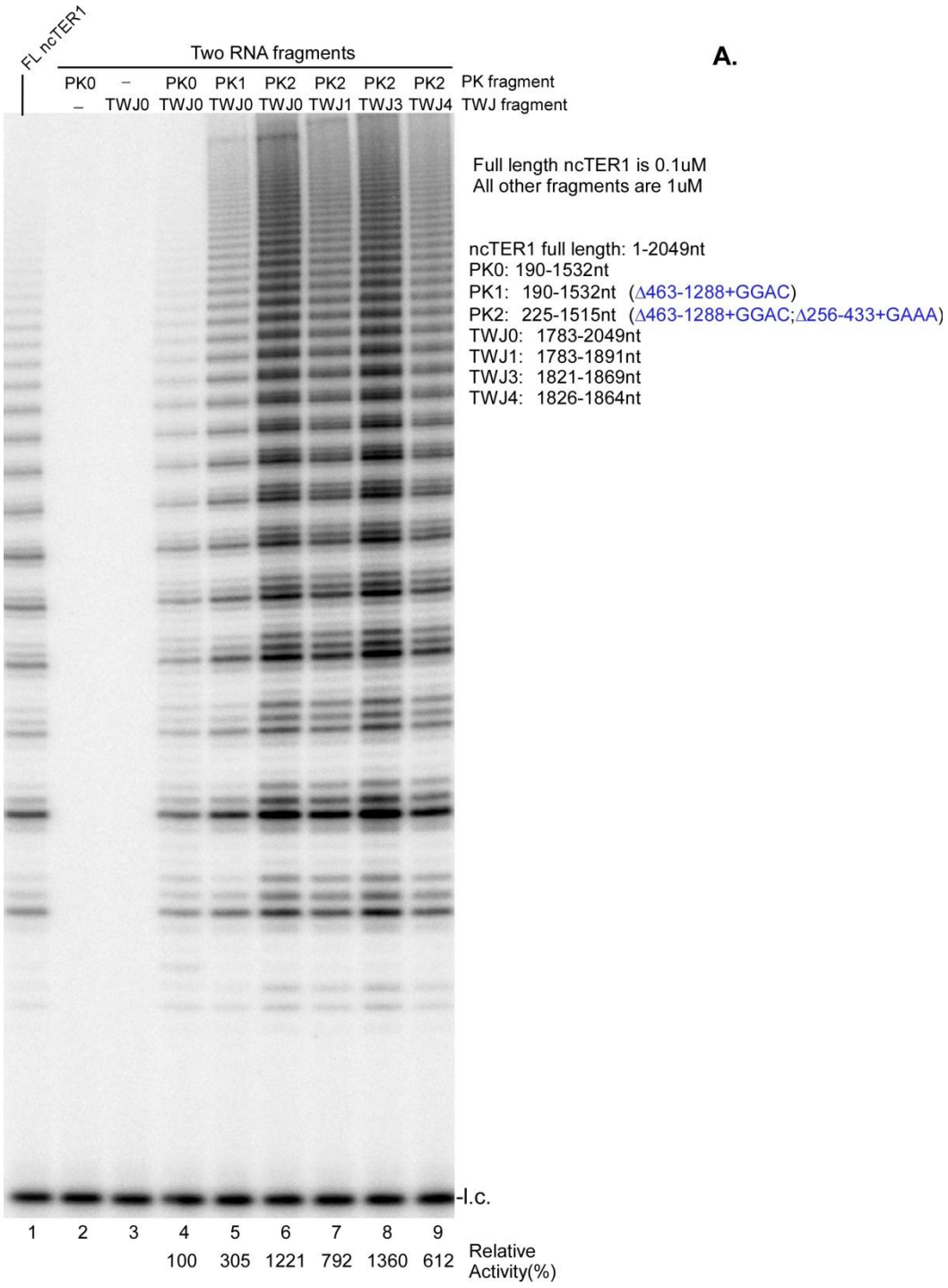
Figure 4.3 *N. crassa* telomerase activities with ncTER1 template mutants.

Each template region for hTER, ncTER1 wt, and ncTER1 t1, t2 and t3 mutants are shown on the right. The solid black box indicates the DNA primer sequence: 5'-TTAGGGTTAG-3'. Normal telomerase RNA template regions are boxed. The inserted nucleotide(s) to either 5' or 3' end of the ncTER1 template are shown as green, while template realignment regions are shaded in grey. The newly synthesized DNA nucleotides for first round of telomerase reaction are shown as red color. Relative processivities as normalized to wildtype ncTER1 are shown under the gel.

4.4.3 ncTER1 template-pseudoknot and TWJ (P6/6.1) can reconstitute telomerase activity *in trans* with ncTERT in RRL

The catalytic TERT protein harbors two separate RNA binding domains. In ciliate and vertebrate telomerase, the telomerase activity can be reconstituted *in vitro* with TERT and two telomerase RNA fragments *in trans* (Lai et al., 2003; Mason et al., 2003; Mitchell and Collins, 2000). While in budding yeast *S. cerevisiae*, the mini-T fragment containing template-pseudoknot and other proteins binding sites is sufficient for both *in vitro* and *in vivo* telomerase reconstitution (Zappulla et al., 2005). With the functional structure and telomerase activity similarities to vertebrate telomerase RNA, we propose that ncTER1 could also be separated into two segments which have independent binding to TERT protein. To verify this, a serial of ncTER1 truncations were generated for telomerase reconstitution in RRL (Fig 4.4). The template-pseudoknot domain PK0 and TWJ domain TWJ0 could reconstitute active telomerase with ncTERT in RRL only when both domains existed, suggesting functional similarities to vertebrate telomerase RNA. When truncating the two big arms (463-1288 and 256-433) of the ncTER1 PK0 domain, strongest activity was seen with PK2 RNA fragment (295nt). This suggested that truncated RNAs might fold better than the larger form, thus generate stronger activity. Similar phenomena were observed in TWJ domain truncations. The strongest activity was observed when reconstitution with PK2 and TWJ3. TWJ4, however, exhibited around half of telomerase activity as compared to TWJ3, indicating the bottom stem (P5 in vertebrate TERs)

in TWJ3, although not essential for activity *in vitro*, may provide support for maintaining TWJ RNA secondary structure.



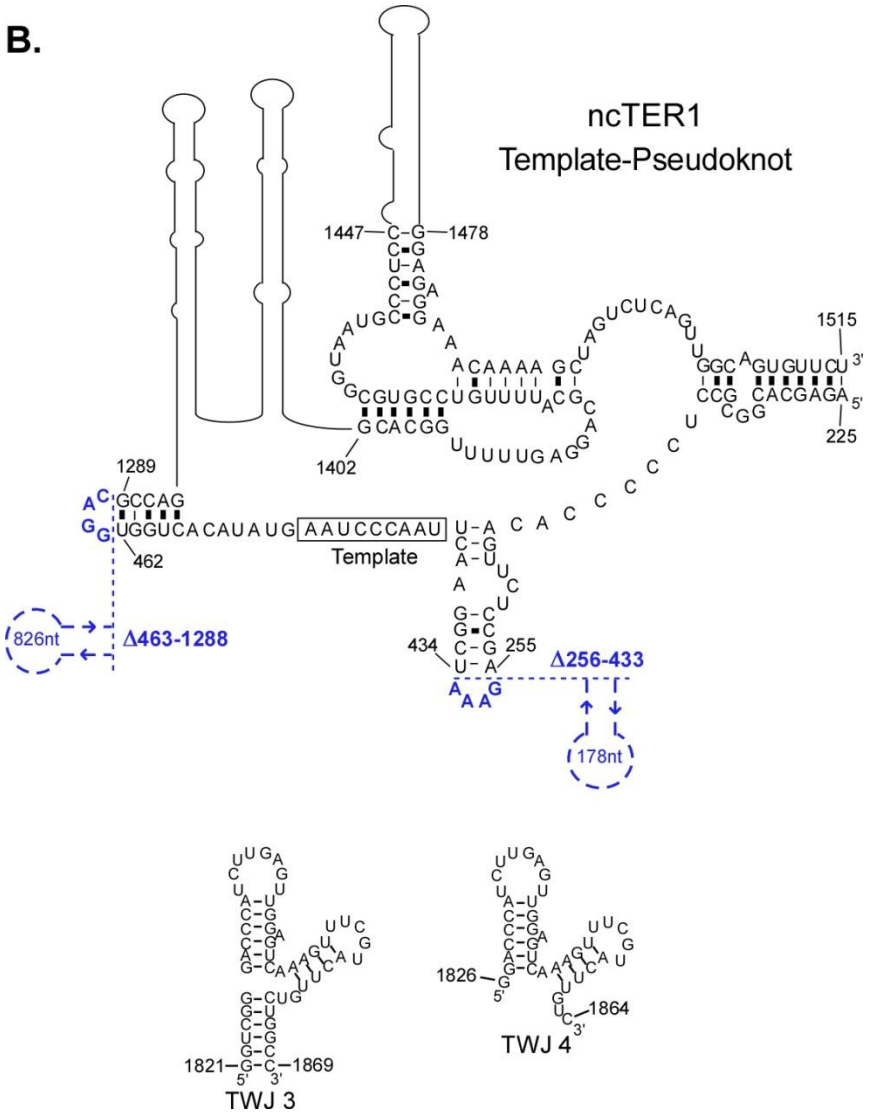


Figure 4.4 Telomerase activity assay with truncated ncTER1 fragments.

(A) *N. crassa* telomerase activity assay with full length and truncated ncTER1 fragment(s) reconstituted with ncTERT in RRL. The full length ncTER1 activity was shown in lane 1. Single fragment pseudoknot PK0 and TWJ0 reconstituted with ncTERT in RRL respectively were shown in lane 2 and 3. Various two fragments reconstitution with ncTERT *in trans* were shown from lane 4 to lane 9. A 15mer 32P end-labeled oligonucleotide was served as the loading control (l.c.). Relative activities (%) normalized to lane 4 was shown under the gel. Sequence numbers for each ncTER1 truncation were shown with numbers corresponding to the full length mature ncTER1. (B) Secondary structure of ncTER1 template-pseudoknot and TWJ truncations. Predicted ncTER1 structure from 225-1515nt was presented, with truncations indicated in blue. ncTER1 TWJ3 and TWJ4 secondary structure prediction were shown.

4.4.4 ncTER1 P6/6.1 functions the same as vertebrate CR4/CR5 domain

The p6.1 loop consists of two highly conserved nucleotides U1854 and G1856 which also exist in vertebrate p6.1 loop, and play an important function in telomerase activity regardless of protein-RNA binding as previously examined in mouse telomerase RNA (Chen et al., 2002). TWJ4, the smallest TWJ form, was then mutated accordingly and as predicted, dramatic telomerase activity reduction was seen with both single and double mutations. This suggests the conservation of P6/6.1 between filamentous fungal TER1 and vertebrate telomerase RNA.

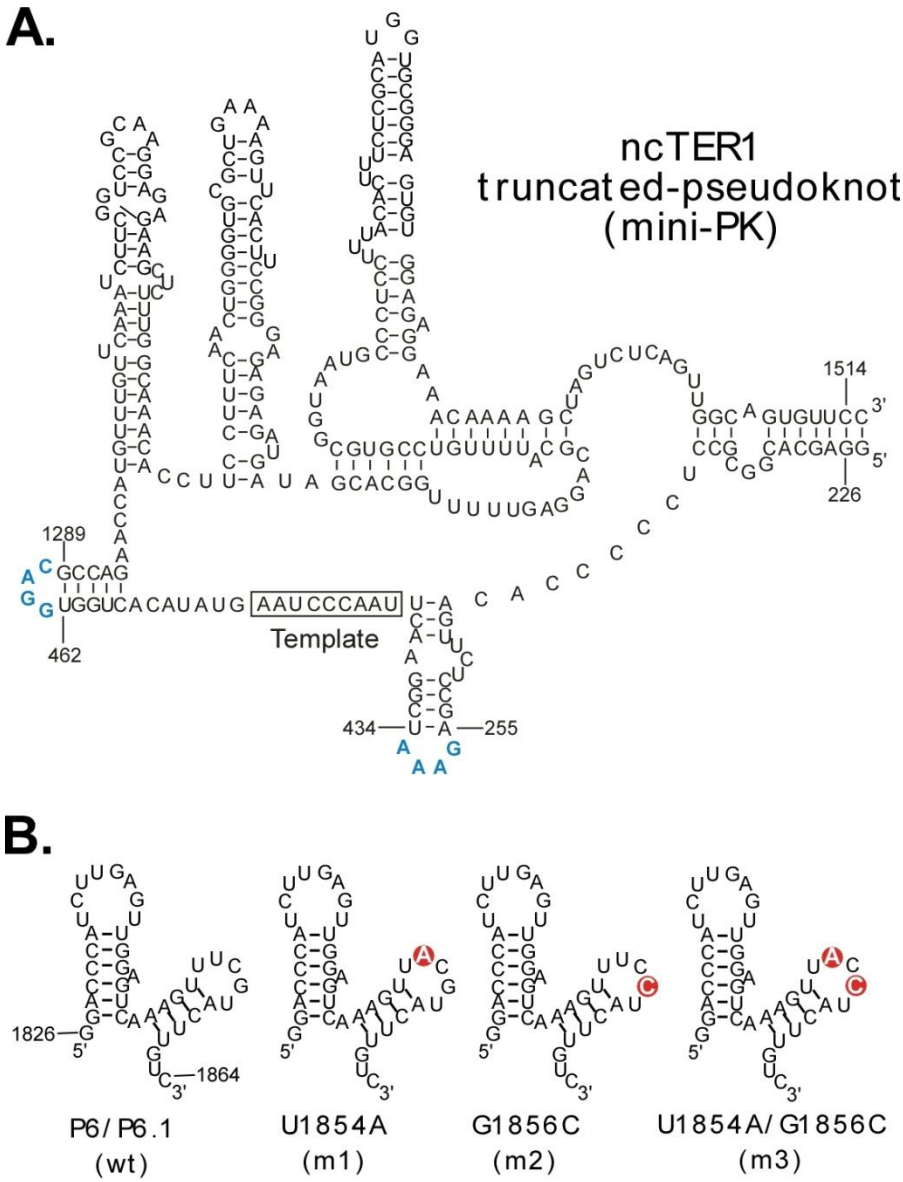


Figure 4.5 Secondary of ncTER1 template-pseudoknot and P6/6.1

(A) Truncated ncTER1 template-pseudoknot structure (mini-PK). Each truncated region (256-433 and 463-1288) is replaced with a loop as indicated with blue. The ncTER1 template region is boxed. Nucleotide numbers from full length ncTER1 are labeled. (B) The structures of mini P6/6.1 together with the mutants m1 (U1854A), m2 (G1856C) and m3 (U1854A+G1856C) are shown. The mutated nucleotide(s) are shaded in red color.

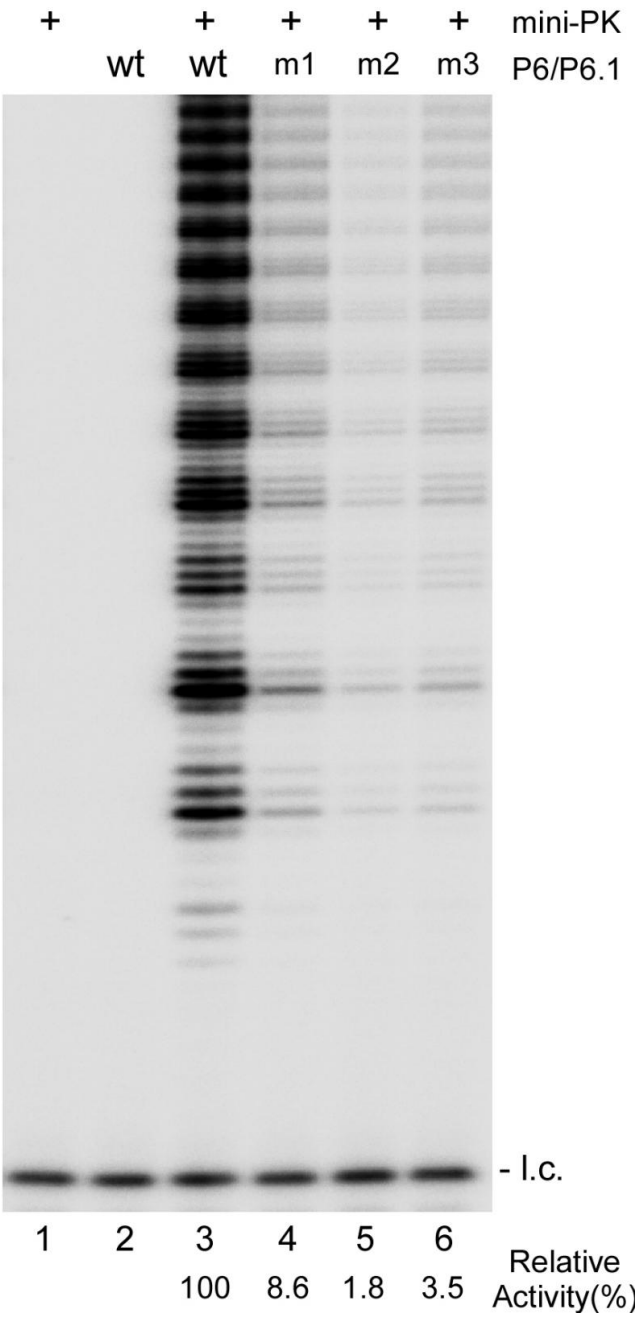


Figure 4.6 Telomerase activity assay with ncTER1 truncated fragments

Telomerase activity assay was carried out with RRL expressed ncTERT assembled with individual fragment mini-PK (lane 1) or P6/6.1 RNA (lane 2) or both fragments as mini-PK plus P6/6.1 wt (lane 3), m1 (lane 4), m2 (lane 5) and m3 (lane6). The relative telomerase activities normalized to lane 3 are shown under the gel. L.c.: a 15mer ^{32}P end-labeled primer used as loading control.

4.4.5 Telomerase activity analysis of ncTER1 pseudoknot stem loop truncations

Several hairpin structures were observed within the conserved ncTER1 template-pseudoknot, according to the predicted ncTER1 secondary structure in chapter 3. Two of them locate at the upstream region of the pseudoknot, with little sequence conservation (data not shown). The other one locates within the pseudoknot region, and was named as stem PK2.1, with highly conserved bottom stem sequences. The only exception of this conserved stem sequences is *Magnaporthe oryzae* TER1 which still contains a hairpin structure at the corresponding location (Supplemental fig S4.1). Such hairpin structure was only discovered in *Kluyveromyces lactis* telomerase RNA which harbors two hairpins within its pseudoknot region (Chappell and Lundblad, 2004; Podlevsky et al., 2008). We asked the question that whether these hairpins are important for the telomerase enzymatic activity.

To investigate the function of these hairpin structures, we performed the telomerase activity assay using template/pseudoknot truncations. The template/pseudoknot truncation PK1 was utilized for making further truncations. The PK1.4 deleted the two variable hairpins (1300-1346 and 1352-1398), connecting the core enclosing helix 2 with pseudoknot with a short single stranded region: 5'-AACCAUACCUUAUA-3'. The PK1.5 and PK1.6 were truncations for the conserved hairpin structure, with PK1.5 deleting the whole hairpin, while PK1.6 remaining the conserved stem sequences with an artificial tetraloop (GAAA). *N. crassa* telomerase with truncated ncTER1 PK1 was reconstituted in RRL and activity assay was performed as described previously.

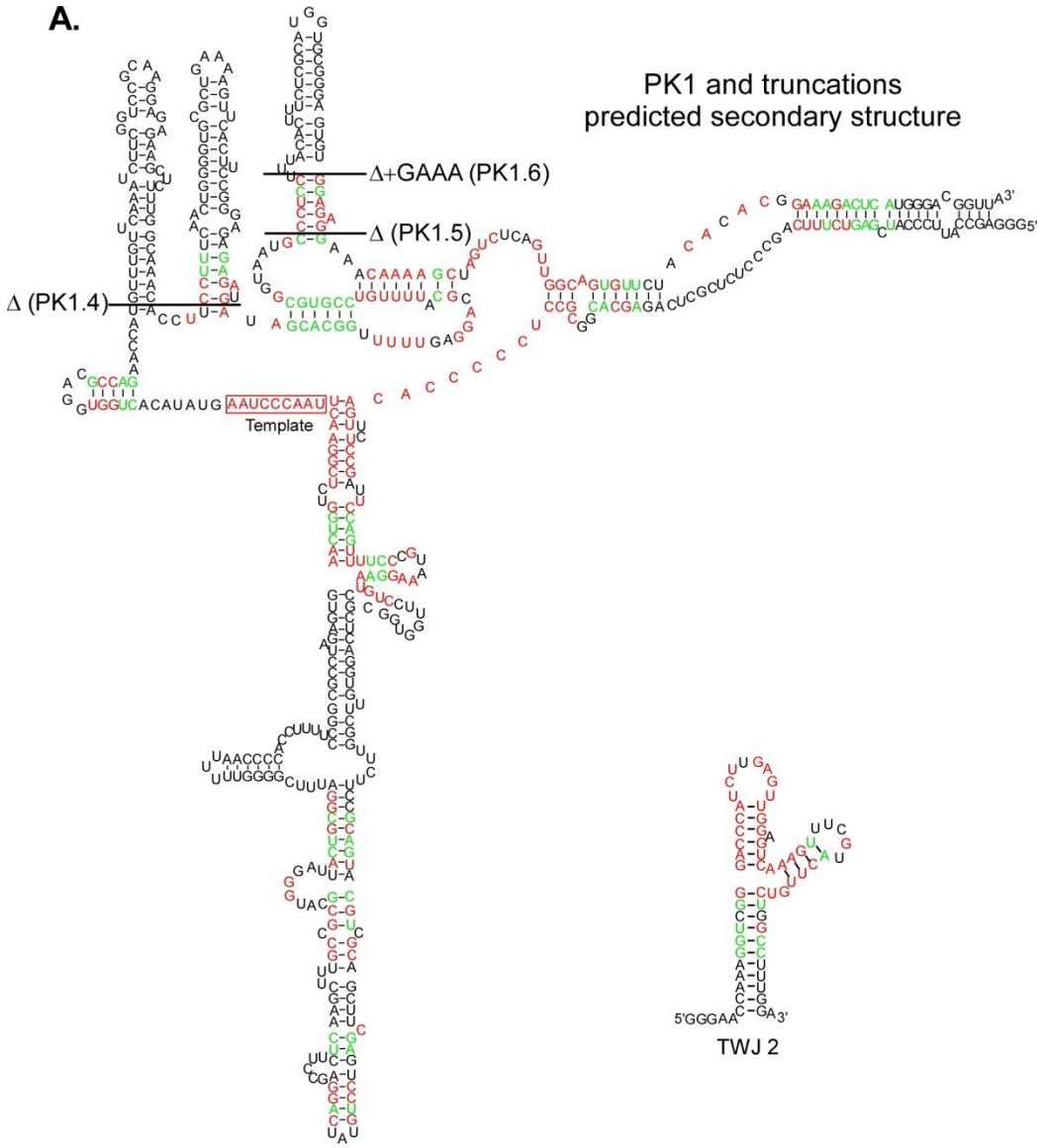


Figure 4.7 Telomerase activity assay with ncTER1 pseudoknot hairpin truncations

(A) Secondary structures were shown for mini-PK1 and its hairpin truncations together with TWJ2. Residues with more than 90% conservation are highlighted in red while base pairings with covariations are labeled as green. Deletions are made with the sequences above the lines either deleted (PK1.4 and PK1.5) or replaced by a tetraloop sequence GAAA (PK1.6).

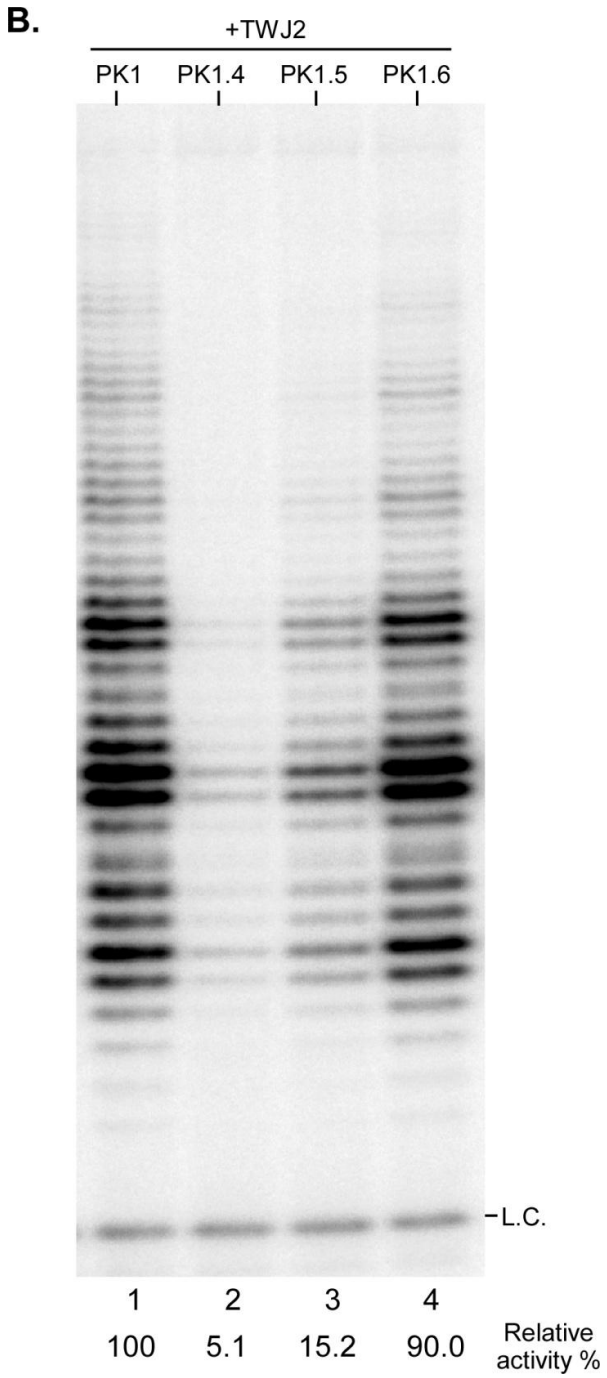


Figure 4.7 continue

(B) *N. crassa* telomerase activity assay performed with TWJ2 and mini-PK1 wildtype (lane 1) or truncations (lane 2 to 4). Activities were quantitated and normalized to lane1 and indicated at the bottom of the gel. L.c.: a 15mer ^{32}P end-labeled DNA oligo served as loading control.

4.5 Discussion

It was proved in other species that telomerase enzyme core only consists of a catalytic TERT protein and a telomerase RNA component. Our *in vitro* reconstituted *N. crassa* telomerase also confirmed this conclusion, and proved that Est1 protein is not essential for *in vitro* activity, similar to yeast telomerase (Cohn and Blackburn, 1995). Whether Est1 in *N. crassa* is essential for *in vivo* telomerase biogenesis and function, however, still remains unknown. ncEst1 knock out strain needs to be investigated for telomere maintenance, as Est1 knock out *S. cerevisiae* cells showed senescence phenotype (Lundblad and Szostak, 1989).

S. cerevisiae full length telomerase RNA TLC1 failed to reconstitute telomerase activity *in vitro,* due to large RNA misfolding (Zappulla et al., 2005). Instead, telomerase RNA truncations were required for successful yeast telomerase *in vitro* reconstitution. The 2,049nt ncTER1 mature RNA, to our surprise, reconstitutes active telomerase *in vitro*. This suggests that ncTER1 tends to fold a correct structure independently. This reconstitution system provides a useful tool for the study of telomerase function and regulation.

Unlike yeast, two RNA fragments, a template-pseudoknot and a P6/6.1 (three way junction) are essential for *N. crassa* telomerase enzymatic activity *in vitro*, similar to distantly related vertebrate species. This suggests that these two domains might be present in the telomerase RNA ancestor. The three way junction domain, although still remains in all yeast telomerase RNAs, might lost its function of stimulating telomerase activity during their fast evolution.

4.6 References

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

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

Chen, J.L., Opperman, K.K., and Greider, C.W. (2002). A critical stem-loop structure in the CR4-CR5 domain of mammalian telomerase RNA. Nucleic Acids Res *30*, 592-597.

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

Fulton, T.B., and Blackburn, E.H. (1998). Identification of Kluyveromyces lactis telomerase: discontinuous synthesis along the 30-nucleotide-long templating domain. Mol Cell Biol *18*, 4961-4970.

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

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

Lai, C.K., Miller, M.C., and Collins, K. (2003). Roles for RNA in telomerase nucleotide and repeat addition processivity. Mol Cell *11*, 1673-1683.

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

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

Mason, D.X., Goneska, E., and Greider, C.W. (2003). Stem-loop IV of tetrahymena telomerase RNA stimulates processivity in trans. Mol Cell Biol *23*, 5606-5613.

Mitchell, J.R., and Collins, K. (2000). Human telomerase activation requires two independent interactions between telomerase RNA and telomerase reverse transcriptase. Mol Cell *6*, 361-371.

Podlevsky, J.D., Bley, C.J., Omana, R.V., Qi, X., and Chen, J.J. (2008). The telomerase database. Nucleic Acids Res *36*, D339-343.

Xie, M., Mosig, A., Qi, X., Li, Y., Stadler, P.F., and Chen, J.J. (2008). Structure and function of the smallest vertebrate telomerase RNA from teleost fish. J Biol Chem *283*, 2049-2059.

Xie, M., Podlevsky, J.D., Qi, X., Bley, C.J., and Chen, J.J. (2010). A novel motif in telomerase reverse transcriptase regulates telomere repeat addition rate and processivity. Nucleic Acids Res *38*, 1982-1996.

Zappulla, D.C., Goodrich, K., and Cech, T.R. (2005). A miniature yeast telomerase RNA functions in vivo and reconstitutes activity in vitro. Nat Struct Mol Biol *12*, 1072-1077.

Chapter 5

CONCLUSION

Various approaches have been employed for the identification of telomerase RNAs to overcome the strikingly divergent nature of telomerase RNA during evolution. Our approach of combining purification of telomerase holoenzyme, deep sequencing and bioinformatic search has universal applications for the identification of novel telomerase RNAs from a variety of species, and can be amended for novel non-telomerase RNAs as well. With significant improvements in sequencing technology, additional telomeres and telomerases will be characterized from evolutionarily distinct species, expanding the body of knowledge for conserved domains and stimulating the telomerase biochemical research.

Neurospora crassa telomerase RNA was identified and characterized in this study and presents structural and functional similarities with both vertebrate and yeast telomerase RNAs (Figure 5.1). The ncTER1 contains a vertebrate-like telomerase RNA template and telomeric DNA repeat, thus *N. crassa* would serve as a better model organism for the study of telomere biology and telomerase regulation. Additional filamentous fungal species have been well characterized and have many genetic tools available, including several *Aspergillus* species, e.g. *Aspergillus nidulans* and *Aspergillus niger*. The strikingly quick evolution rate of telomeric DNA repeat as observed in *Aspergillus* species serves as a model system for the co-evolution of telomere and telomere binding proteins. All of

these evidences suggest that filamentous fungus is a great system for the study of telomerase biochemistry and telomerase RNA evolution.

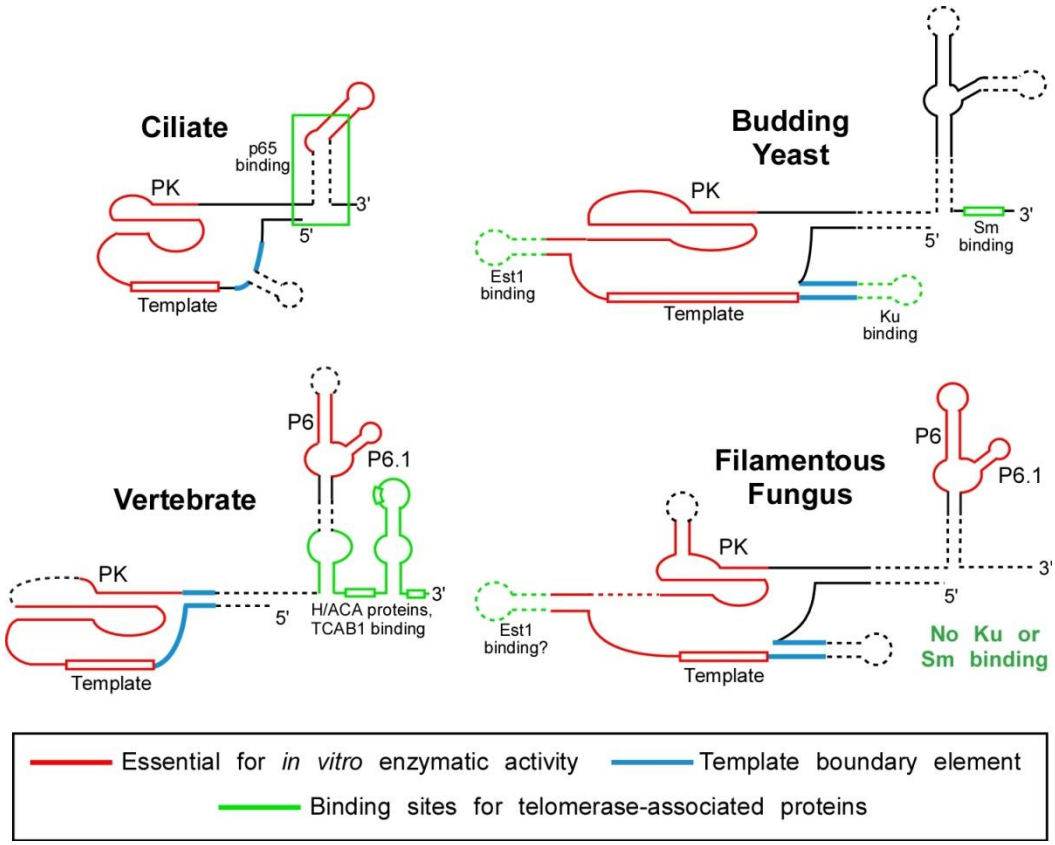


Figure 5.1 Comparison of four major telomerase RNAs

The red boxes represent the template region, whereas dashed lines represent variable sequences within the same phylogenetic group. PK indicates the pseudoknot domain. The structures essential for *in vitro* enzymatic activity were labeled in red. TER template boundary elements were indicated with blue color. Telomerase associated proteins binding sites were highlighted with green color, with proteins indicated adjacent. Est1 binding site of filamentous fungal TER1 is predicted to be the same position as that in budding yeast. No ku or Sm binding sites were identified from *N. crassa* telomerase RNA.

REFERENCES

Adams, S.P., Hartman, T.P., Lim, K.Y., Chase, M.W., Bennett, M.D., Leitch, I.J., and Leitch, A.R. (2001). Loss and recovery of Arabidopsis-type telomere repeat sequences 5'-(TTTAGGG)(n)-3' in the evolution of a major radiation of flowering plants. Proc Biol Sci *268*, 1541-1546.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res *25*, 3389-3402.

Baltimore, D. (1970). RNA-dependent DNA polymerase in virions of RNA tumour viruses. Nature *226*, 1209-1211.

Baumann, P., Podell, E., and Cech, T.R. (2002). Human Pot1 (protection of telomeres) protein: cytolocalization, gene structure, and alternative splicing. Mol Cell Biol *22*, 8079-8087.

Bhattacharyya, A., and Blackburn, E.H. (1997). Aspergillus nidulans maintains short telomeres throughout development. Nucleic Acids Res *25*, 1426-1431.

Blackburn, E.H., and Challoner, P.B. (1984). Identification of a telomeric DNA sequence in Trypanosoma brucei. Cell *36*, 447-457.

Blackburn, E.H., and Collins, K. (2010). Telomerase: An RNP Enzyme Synthesizes DNA. Cold Spring Harb Perspect Biol.

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

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

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

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

Bryan, T.M., Sperger, J.M., Chapman, K.B., and Cech, T.R. (1998). Telomerase reverse transcriptase genes identified in Tetrahymena thermophila and Oxytricha trifallax. Proc Natl Acad Sci U S A *95*, 8479-8484.

Chakrabarti, K., Pearson, M., Grate, L., Sterne-Weiler, T., Deans, J., Donohue, J.P., and Ares, M., Jr. (2007). Structural RNAs of known and unknown function identified in malaria parasites by comparative genomics and RNA analysis. RNA *13*, 1923-1939.

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

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

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

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

Chen, J.L., Opperman, K.K., and Greider, C.W. (2002). A critical stem-loop structure in the CR4-CR5 domain of mammalian telomerase RNA. Nucleic Acids Res *30*, 592-597.

Cifuentes-Rojas, C., Kannan, K., Tseng, L., and Shippen, D.E. (2011). Two RNA subunits and POT1a are components of Arabidopsis telomerase. Proc Natl Acad Sci U S A *108*, 73-78.

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

Connelly, J.C., and Arst, H.N., Jr. (1991). Identification of a telomeric fragment from the right arm of chromosome III of Aspergillus nidulans. FEMS Microbiol Lett *64*, 295-297.

Cristofari, G., Adolf, E., Reichenbach, P., Sikora, K., Terns, R.M., Terns, M.P., and Lingner, J. (2007). Human telomerase RNA accumulation in Cajal bodies facilitates telomerase recruitment to telomeres and telomere elongation. Mol Cell *27*, 882-889.

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

Danner, S., and Belasco, J.G. (2001). T7 phage display: a novel genetic selection system for cloning RNA-binding proteins from cDNA libraries. Proc Natl Acad Sci U S A *98*, 12954-12959.

Davis, R.H. (2000). Neurospora : contributions of a model organism (Oxford ; New York, Oxford University Press).

Evans, S.K., Sistrunk, M.L., Nugent, C.I., and Lundblad, V. (1998). Telomerase, Ku, and telomeric silencing in Saccharomyces cerevisiae. Chromosoma *107*, 352-358.

Feng, J., Funk, W.D., Wang, S.S., Weinrich, S.L., Avilion, A.A., Chiu, C.P., Adams, R.R., Chang, E., Allsopp, R.C., Yu, J., *et al.* (1995). The RNA component of human telomerase. Science *269*, 1236-1241.

Forstemann, K., Hoss, M., and Lingner, J. (2000). Telomerase-dependent repeat divergence at the 3' ends of yeast telomeres. Nucleic Acids Res *28*, 2690-2694.

Frydrychova, R., Grossmann, P., Trubac, P., Vitkova, M., and Marec, F. (2004). Phylogenetic distribution of TTAGG telomeric repeats in insects. Genome *47*, 163-178.

Fuchs, J., Brandes, A., and Schubert, I. (1995). Telomere sequence localization and karyotype evolution in higher plants Plant Systematics and Evolution *196*, 227-241.

Fulton, T.B., and Blackburn, E.H. (1998). Identification of Kluyveromyces lactis telomerase: discontinuous synthesis along the 30-nucleotide-long templating domain. Mol Cell Biol *18*, 4961-4970.

Galagan, J.E., Calvo, S.E., Borkovich, K.A., Selker, E.U., Read, N.D., Jaffe, D., FitzHugh, W., Ma, L.J., Smirnov, S., Purcell, S., *et al.* (2003). The genome sequence of the filamentous fungus Neurospora crassa. Nature *422*, 859-868.

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

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

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

Gunisova, S., Elboher, E., Nosek, J., Gorkovoy, V., Brown, Y., Lucier, J.F., Laterreur, N., Wellinger, R.J., Tzfati, Y., and Tomaska, L. (2009). Identification and comparative analysis of telomerase RNAs from Candida species reveal conservation of functional elements. RNA *15*, 546-559.

Hinkley, C.S., Blasco, M.A., Funk, W.D., Feng, J., Villeponteau, B., Greider, C.W., and Herr, W. (1998). The mouse telomerase RNA 5"-end lies just upstream of the telomerase template sequence. Nucleic Acids Res *26*, 532-536.

Hoffmann, S., Otto, C., Kurtz, S., Sharma, C.M., Khaitovich, P., Vogel, J., Stadler, P.F., and Hackermuller, J. (2009). Fast Mapping of Short Sequences with Mismatches, Insertions and Deletions Using Index Structures. Plos Comput Biol *5*.

Honda, S., and Selker, E.U. (2008). Direct interaction between DNA methyltransferase DIM-2 and HP1 is required for DNA methylation in Neurospora crassa. Mol Cell Biol *28*, 6044-6055.

Honda, S., and Selker, E.U. (2009). Tools for fungal proteomics: multifunctional neurospora vectors for gene replacement, protein expression and protein purification. Genetics *182*, 11-23.

Hsu, M., McEachern, M.J., Dandjinou, A.T., Tzfati, Y., Orr, E., Blackburn, E.H., and Lue, N.F. (2007). Telomerase core components protect Candida telomeres from aberrant overhang accumulation. Proc Natl Acad Sci U S A *104*, 11682-11687.

Jady, B.E., Bertrand, E., and Kiss, T. (2004). Human telomerase RNA and box H/ACA scaRNAs share a common Cajal body-specific localization signal. J Cell Biol *164*, 647-652.

Kachouri-Lafond, R., Dujon, B., Gilson, E., Westhof, E., Fairhead, C., and Teixeira, M.T. (2009). Large telomerase RNA, telomere length heterogeneity and escape from senescence in Candida glabrata. FEBS Lett *583*, 3605-3610.

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

Kusumoto, K.I., Suzuki, S., and Kashiwagi, Y. (2003). Telomeric repeat sequence of Aspergillus oryzae consists of dodeca-nucleotides. Appl Microbiol Biotechnol *61*, 247-251.

Lai, C.K., Miller, M.C., and Collins, K. (2003). Roles for RNA in telomerase nucleotide and repeat addition processivity. Mol Cell *11*, 1673-1683.

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

Lejnine, S., Makarov, V.L., and Langmore, J.P. (1995). Conserved nucleoprotein structure at the ends of vertebrate and invertebrate chromosomes. Proc Natl Acad Sci U S A *92*, 2393-2397.

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

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

Lin, J., Ly, H., Hussain, A., Abraham, M., Pearl, S., Tzfati, Y., Parslow, T.G., and Blackburn, E.H. (2004). A universal telomerase RNA core structure includes structured motifs required for binding the telomerase reverse transcriptase protein. Proc Natl Acad Sci U S A *101*, 14713-14718.

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

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

Mason, D.X., Goneska, E., and Greider, C.W. (2003). Stem-loop IV of tetrahymena telomerase RNA stimulates processivity in trans. Mol Cell Biol *23*, 5606-5613.

McClintock, B. (1939). The Behavior in Successive Nuclear Divisions of a Chromosome Broken at Meiosis. Proc Natl Acad Sci U S A *25*, 405-416.

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

Meier, U.T., and Blobel, G. (1990). A nuclear localization signal binding protein in the nucleolus. J Cell Biol *111*, 2235-2245.

Meier, U.T., and Blobel, G. (1992). Nopp140 shuttles on tracks between nucleolus and cytoplasm. Cell *70*, 127-138.

Meier, U.T., and Blobel, G. (1994). NAP57, a mammalian nucleolar protein with a putative homolog in yeast and bacteria. J Cell Biol *127*, 1505-1514.

Metzker, M.L. (2010). Sequencing technologies - the next generation. Nat Rev Genet *11*, 31-46.

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

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

Mitchell, J.R., and Collins, K. (2000). Human telomerase activation requires two independent interactions between telomerase RNA and telomerase reverse transcriptase. Mol Cell *6*, 361-371.

Morin, G.B. (1989). The human telomere terminal transferase enzyme is a ribonucleoprotein that synthesizes TTAGGG repeats. Cell *59*, 521-529.

Muller, H.J. (1938). The remaking of chromosomes. Collecting Net *13*, 181-198.

Ninomiya, Y., Suzuki, K., Ishii, C., and Inoue, H. (2004). Highly efficient gene replacements in Neurospora strains deficient for nonhomologous end-joining. Proc Natl Acad Sci U S A *101*, 12248-12253.

O'Connor, C.M., and Collins, K. (2006). A novel RNA binding domain in tetrahymena telomerase p65 initiates hierarchical assembly of telomerase holoenzyme. Mol Cell Biol *26*, 2029-2036.

O'Connor, C.M., Lai, C.K., and Collins, K. (2005). Two purified domains of telomerase reverse transcriptase reconstitute sequence-specific interactions with RNA. J Biol Chem *280*, 17533-17539.

Okazaki, S., Tsuchida, K., Maekawa, H., Ishikawa, H., and Fujiwara, H. (1993). Identification of a pentanucleotide telomeric sequence, (TTAGG)n, in the silkworm Bombyx mori and in other insects. Mol Cell Biol *13*, 1424-1432.

Peterson, S.E., Stellwagen, A.E., Diede, S.J., Singer, M.S., Haimberger, Z.W., Johnson, C.O., Tzoneva, M., and Gottschling, D.E. (2001). The function of a stem-loop in telomerase RNA is linked to the DNA repair protein Ku. Nat Genet *27*, 64-67.

Podlevsky, J.D., Bley, C.J., Omana, R.V., Qi, X., and Chen, J.J. (2008). The telomerase database. Nucleic Acids Res *36*, D339-343.

Richard, P., Darzacq, X., Bertrand, E., Jady, B.E., Verheggen, C., and Kiss, T. (2003). A common sequence motif determines the Cajal body-specific localization of box H/ACA scaRNAs. EMBO J *22*, 4283-4293.

Richards, R.J., Wu, H., Trantirek, L., O'Connor, C.M., Collins, K., and Feigon, J. (2006). Structural study of elements of Tetrahymena telomerase RNA stem-loop IV domain important for function. RNA *12*, 1475-1485.

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

Sahara, K., Marec, F., and Traut, W. (1999). TTAGG telomeric repeats in chromosomes of some insects and other arthropods. Chromosome Res *7*, 449-460.

Schadt, E.E., Turner, S., and Kasarskis, A. (2010). A window into third-generation sequencing. Hum Mol Genet *19*, R227-240.

Schechtman, M.G. (1987). Isolation of telomere DNA from Neurospora crassa. Mol Cell Biol *7*, 3168-3177.

Selker, E.U. (1990). Premeiotic instability of repeated sequences in Neurospora crassa. Annu Rev Genet *24*, 579-613.

Seto, A.G., Livengood, A.J., Tzfati, Y., Blackburn, E.H., and Cech, T.R. (2002). A bulged stem tethers Est1p to telomerase RNA in budding yeast. Genes Dev *16*, 2800-2812.

Seto, A.G., Zaug, A.J., Sobel, S.G., Wolin, S.L., and Cech, T.R. (1999). Saccharomyces cerevisiae telomerase is an Sm small nuclear ribonucleoprotein particle. Nature *401*, 177-180.

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

Shefer, K., Brown, Y., Gorkovoy, V., Nussbaum, T., Ulyanov, N.B., and Tzfati, Y. (2007). A triple helix within a pseudoknot is a conserved and essential element of telomerase RNA. Mol Cell Biol *27*, 2130-2143.

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

Sinclair, C.S., Richmond, R.H., and Ostrander, G.K. (2007). Characterization of the telomere regions of scleractinian coral, Acropora surculosa. Genetica *129*, 227-233.

Singer, M.S., and Gottschling, D.E. (1994). TLC1: template RNA component of Saccharomyces cerevisiae telomerase. Science *266*, 404-409.

Smith, C.M., and Steitz, J.A. (1997). Sno storm in the nucleolus: new roles for myriad small RNPs. Cell *89*, 669-672.

Spatafora, J.W., Sung, G.H., Johnson, D., Hesse, C., O'Rourke, B., Serdani, M., Spotts, R., Lutzoni, F., Hofstetter, V., Miadlikowska, J., *et al.* (2006). A five-gene phylogeny of Pezizomycotina. Mycologia *98*, 1018-1028.

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

Temin, H.M., and Mizutani, S. (1970). RNA-dependent DNA polymerase in virions of Rous sarcoma virus. Nature *226*, 1211-1213.

Theimer, C.A., Blois, C.A., and Feigon, J. (2005). Structure of the human telomerase RNA pseudoknot reveals conserved tertiary interactions essential for function. Mol Cell *17*, 671-682.

Tollervey, D., and Mattaj, I.W. (1987). Fungal small nuclear ribonucleoproteins share properties with plant and vertebrate U-snRNPs. EMBO J *6*, 469-476.

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

Watson, J.D. (1972). Origin of concatemeric T7 DNA. Nat New Biol *239*, 197-201.

Webb, C.J., and Zakian, V.A. (2008). Identification and characterization of the Schizosaccharomyces pombe TER1 telomerase RNA. Nat Struct Mol Biol *15*, 34-42.

Xie, M., Mosig, A., Qi, X., Li, Y., Stadler, P.F., and Chen, J.J. (2008). Structure and function of the smallest vertebrate telomerase RNA from teleost fish. J Biol Chem *283*, 2049-2059.

Xie, M., Podlevsky, J.D., Qi, X., Bley, C.J., and Chen, J.J. (2010). A novel motif in telomerase reverse transcriptase regulates telomere repeat addition rate and processivity. Nucleic Acids Res *38*, 1982-1996.

Yang, Y., Isaac, C., Wang, C., Dragon, F., Pogacic, V., and Meier, U.T. (2000). Conserved composition of mammalian box H/ACA and box C/D small nucleolar ribonucleoprotein particles and their interaction with the common factor Nopp140. Mol Biol Cell *11*, 567-577.

Yao, M.C., Blackburn, E., and Gall, J. (1981). Tandemly repeated C-C-C-C-A-A hexanucleotide of Tetrahymena rDNA is present elsewhere in the genome and may be related to the alteration of the somatic genome. J Cell Biol *90*, 515-520.

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

Zappulla, D.C., Goodrich, K., and Cech, T.R. (2005). A miniature yeast telomerase RNA functions in vivo and reconstitutes activity in vitro. Nat Struct Mol Biol *12*, 1072-1077.

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

Zhang, Q., Kim, N.K., Peterson, R.D., Wang, Z., and Feigon, J. (2010). Structurally conserved five nucleotide bulge determines the overall topology of the core domain of human telomerase RNA. Proc Natl Acad Sci U S A *107*, 18761-18768.

Zhou, J., Hidaka, K., and Futcher, B. (2000). The Est1 subunit of yeast telomerase binds the Tlc1 telomerase RNA. Mol Cell Biol *20*, 1947-1955.

APPENDIX A

NEUROSPORA CRASSA STRAINS UTILIZED IN THIS STUDY

Supplemental Table A.1

List of *Neurospora crassa* strains utilized in this thesis.

Strain	Genotype	Reference
NC1	*mat A his-3 Δmus-52::bar*$^{+}$	(Honda and Selker, 2008)
NC2	*mat A his-3 Δmus-52::bar*$^{+}$*, tert-3xFLAG::hph::loxP*	this study
NC3	*mat a, tert-3xFLAG::hph::loxP*	this study
NC5	*mat A his-3*$^{+}$*::Pccg-1::SmD3-3xFLAG, Δmus-52::bar*$^{+}$	this study
NC6	*mat A his-3*$^{+}$*::Pccg-1::Ku80-3xFLAG, Δmus-52::bar*$^{+}$	this study
NC7	*mat A his-3*$^{+}$*::Pccg-1::Est1-3xFLAG, Δmus-52::bar*$^{+}$	this study
NC8	*mat A his-3*$^{+}$*::Pccg-1::ncTER1 (template 5'-UAA<u>C</u>CCCUAA-3'), Δmus-52::bar*$^{+}$	this study
NC9	*mat A his-3*$^{+}$*::Pccg-1::ncTER1 (template 5'-UAA<u>A</u>CCCUAA<u>A</u>-3'), Δmus-52::bar*$^{+}$	this study
NC11	*mat A his-3*$^{+}$*::Pccg-1::ncTER1 (template 5'-UAA<u>G</u>CCCUAA-3'), Δmus-52::bar*$^{+}$	this study

APPENDIX B

SUPPLEMENTAL INFORMATION FOR CHAPTER 2

Supplemental Table S2.1

Comparison of telomeric repeats from three groups

	Species	Telomeric repeat sequence
Vertebrate		TTAGGG
Yeasts	*S. cerevisiae*	G_{2-8}TTAC(A)
	S. pombe	$T(G)_{2-3}(TG)_{1-6}$
	K. lactis	ACGGATTTGATTAGGTATGTGGTGT
Filamentous fungi	*N. crassa*	TTAGGG
	A. nidulans	TTAGGG

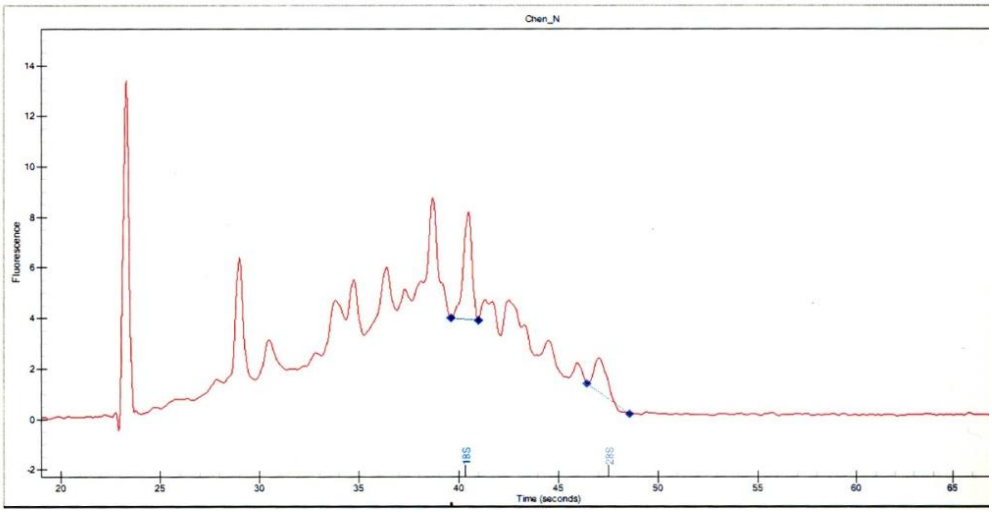
Supplemental Table S2.2

Telomere length of *N. crassa* chromosome VR

	Telomere repeat #	Telomere length (bp)
Wildtype telomeres	23.5	141
	26	156
	21	126
	26.5	159
	29.5	177
	26.5	159
	26	156
	22.5	135
	25.5	153
	24	144
	27.5	165
	24.5	147
	22	132
	24.5	147
	27.5	165
	27.5	165
	24.5	147
	25.5	153
	19.5	117
	26.5	159
	28.5	171
	26.5	159
wt average	25.25	151.5
wt standard deviation	2.457980201	14.7478812
Mutant telomeres	26.5	159
	23.5	141
	25.5	153
	21	126

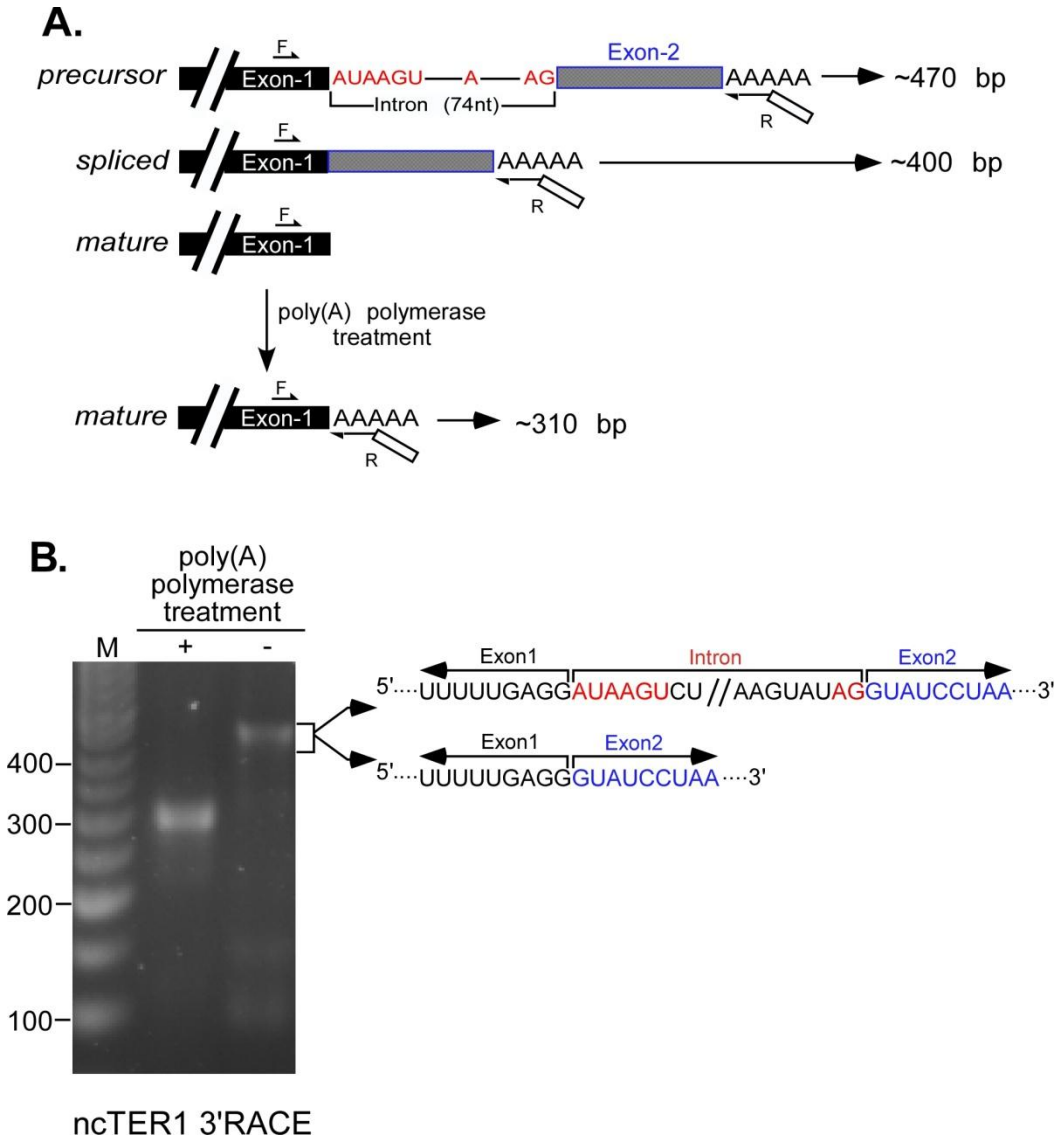
roject:	Jin_Misc	**Acq. Analyst:**	DefaultUser
ssay:	Eukaryote Total RNA StdSens	**Acq. Time:**	8/13/2009 2:5
un:	Run_8-13-2009_2-52-01 PM	**Signature:**	N/A
un Version:	N/A		

Vell# 10 Chen_N



Supplemental figure S2.1 Gel filtration analysis of RNAs for Solexa sequencing.

RNAs were analyzed by size before Solexa deep sequencing. The 18s and 28s rRNA positions are shown.

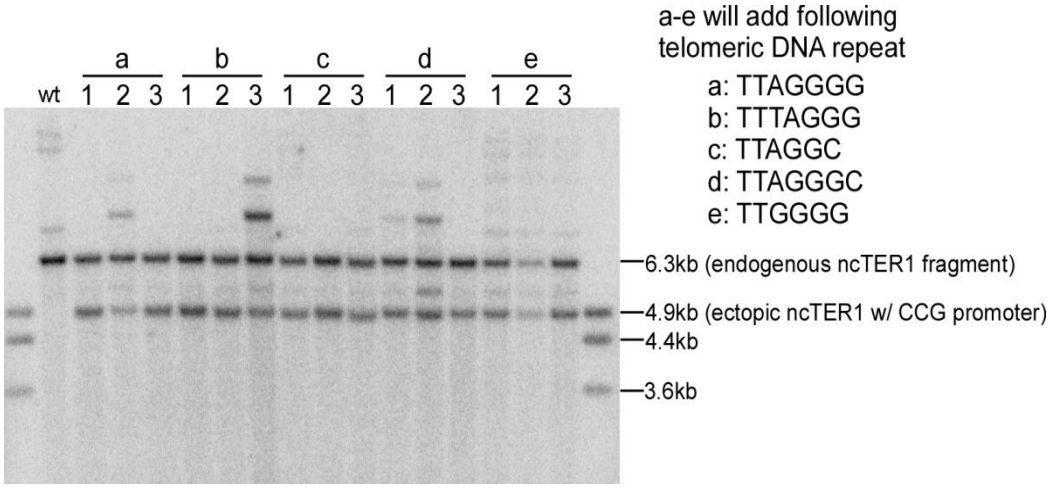


Supplemental Figure S2.2 ncTER1 3' RACE

(A) Schematic of ncTER1 3' RACE. Without the yeast polyA polymerase treatment, the precursor and spliced RNA are amplified. The mature RNA could only be amplified upon polyA polymerase treatment. The size of each product is shown. (B) The *N. crassa* total RNA was initially treated with yeast polyA polymerase before the RT reaction in lane1, whereas no RNA treatment was performed before the RT reaction in lane2. The bands shown in this gel were cloned into pZero vector and sequenced. The band from lane2 has two populations of RNA molecules with partial sequences shown on the right. The larger one (precursor form) harbors an intron between two exons which was completely spliced in the shorter form (spliced form). The potential 5' and 3' splice signal was shown in red. Partial sequences from exon 2 were shown in blue. M indicates the 50bp DNA marker (New England Biolabs).

Supplemental Figure S2.3 ncTER1 full length sequence

ncTER1 precursor full length sequence is shown. The 5' and 3' end are indicated by arrows from RACE results. The mature RNA termination sites are indicated with red arrows, while black arrows are for precursor and spliced form. Numbers next to the arrows indicate the frequency of each termination according to the sequencing result. The template sequence is boxed, and 5' splicing signal (5'ss), branch point (bp) and 3' splicing signal (3' ss) are labeled with red color. Nucleotide positions are shown on the right.

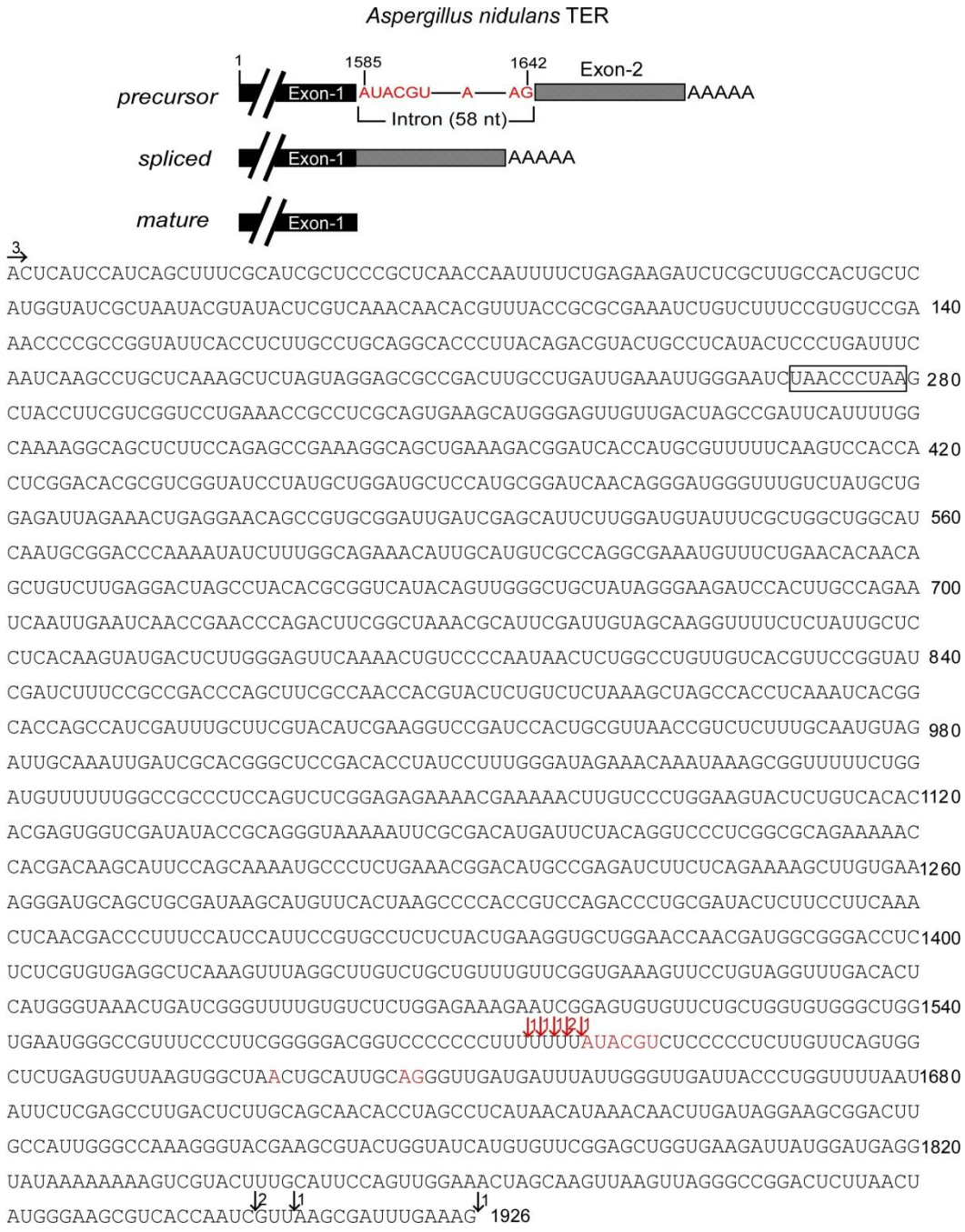


Supplemental Figure S2.4 Southern blot of ncTER1 template mutant strains.

Three individual clones were screened for each ncTER1 template mutant strain. Genomic DNAs were digested with Bsp HI and resolved in a 1% agarose gel. Southern blot was carried out as described in materials and methods. Linearized plasmid DNAs containing *ter1* sequences were utilized as the size marker. Radiolabeled oligonucleotides specific for *N. crassa ter1* gene were utilized as the hybridization probes. The probe sequences are: 5'-CCAACCACAAGTCGGTCTCTGTTTTG-3', 5'-GAGAGGGCTGAAAGACT-CGATG-3' and 5'-GACCAAACTCGTTGAAAACCAATGC-3'. The sizes of the marker were listed on the right side of the gel. Transformants a-1, b-1 and d-3 were utilized for telomeric DNA cloning.

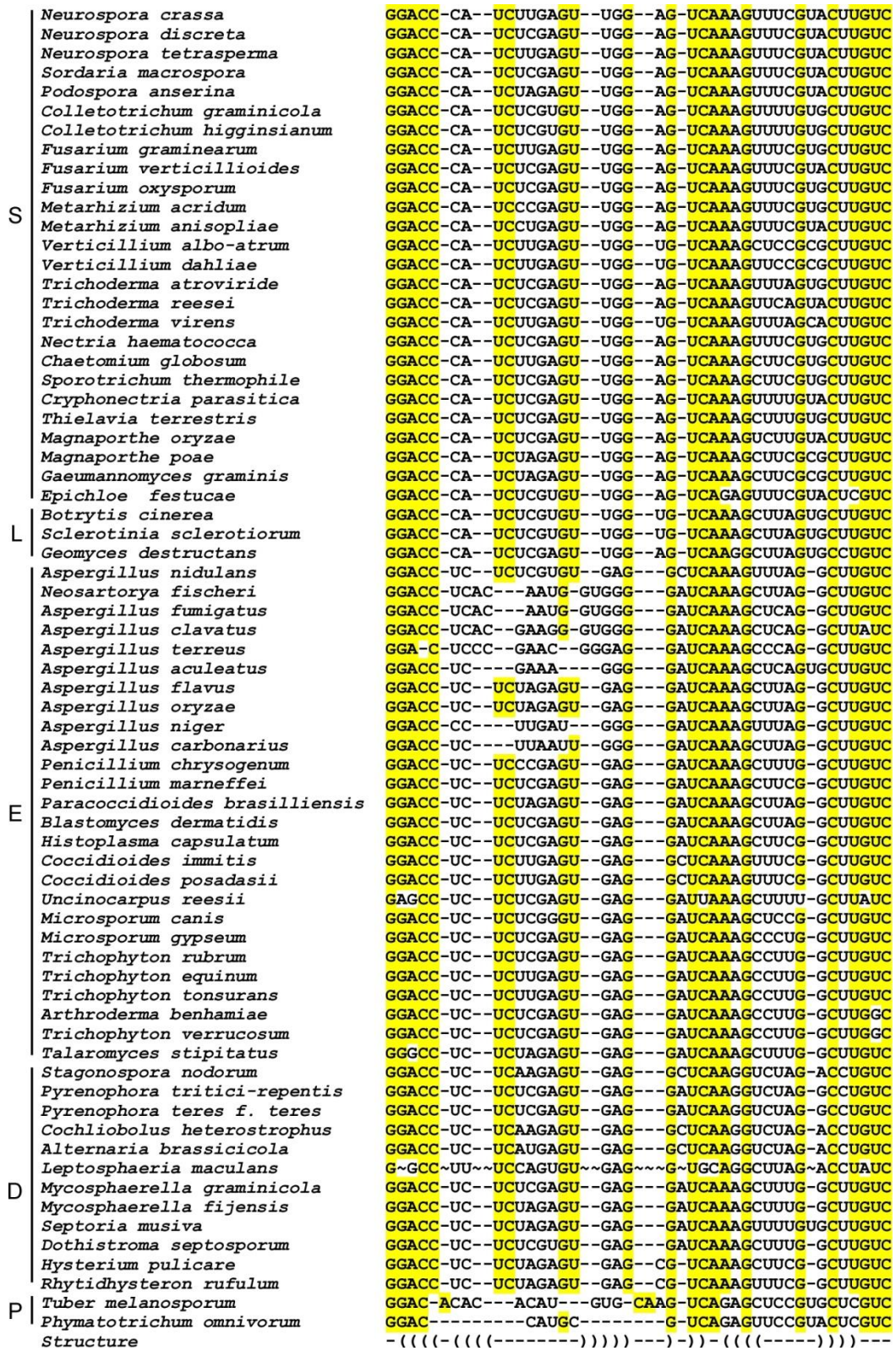
APPENDIX C

SUPPLEMENTAL INFORMATION FOR CHAPTER 3



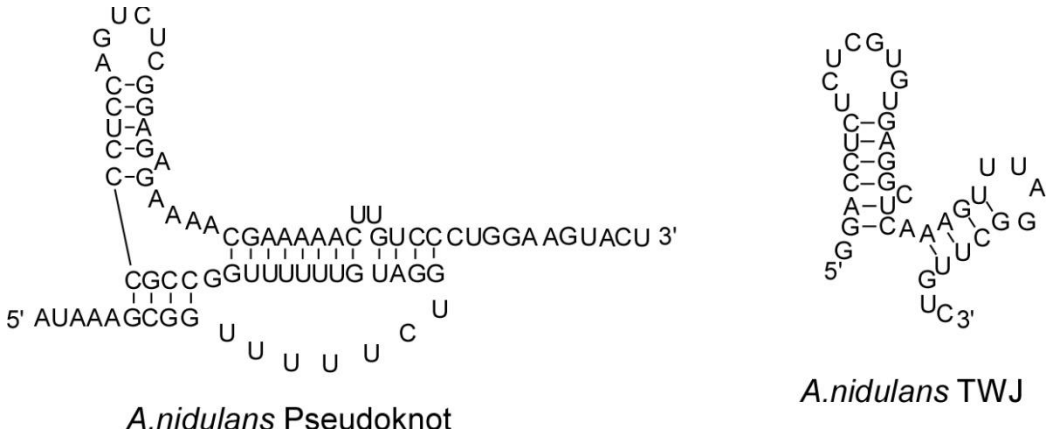
Supplemental Figure S3.1 *Aspergillus nidulans* TER1 full length sequence

The 5' and 3' end are indicated by arrows from RACE results. The mature RNA termination sites are indicated with red arrows, while black arrows are for precursor and spliced form. Numbers next to the arrows indicate the frequency of each termination according to the sequencing result. The template sequence is boxed, and 5' splicing signal, branch point and 3' splicing signal are labeled with red color. Nucleotide positions are shown on the right.


S
L
E
D
P
Neurospora crassa GGACC-CA--UCUUGAGU--UGG--AG-UCAAAGUUUCGUACUUGUC
Neurospora discreta GGACC-CA--UCUCGAGU--UGG--AG-UCAAAGUUUCGUACUUGUC
Neurospora tetrasperma GGACC-CA--UCUUGAGU--UGG--AG-UCAAAGUUUCGUACUUGUC
Sordaria macrospora GGACC-CA--UCUCGAGU--UGG--AG-UCAAAGUUUCGUACUUGUC
Podospora anserina GGACC-CA--UCUAGAGU--UGG--AG-UCAAAGUUUCGUACUUGUC
Colletotrichum graminicola GGACC-CA--UCUCGUGU--UGG--AG-UCAAAGUUUUGUGCUUGUC
Colletotrichum higginsianum GGACC-CA--UCUCGUGU--UGG--AG-UCAAAGUUUUGUGCUUGUC
Fusarium graminearum GGACC-CA--UCUUGAGU--UGG--AG-UCAAAGUUUCGUGCUUGUC
Fusarium verticillioides GGACC-CA--UCUCGAGU--UGG--AG-UCAAAGUUUCGUACUUGUC
Fusarium oxysporum GGACC-CA--UCUCGAGU--UGG--AG-UCAAAGUUUCGUGCUUGUC
Metarhizium acridum GGACC-CA--UCCCGAGU--UGG--AG-UCAAAGUUUCGUGCUUGUC
Metarhizium anisopliae GGACC-CA--UCCUGAGU--UGG--AG-UCAAAGUUUCGUACUUGUC
Verticillium albo-atrum GGACC-CA--UCUUGAGU--UGG--UG-UCAAAGCUCCGCGCUUGUC
Verticillium dahliae GGACC-CA--UCUUGAGU--UGG--UG-UCAAAGUUCCGCGCUUGUC
Trichoderma atroviride GGACC-CA--UCUCGAGU--UGG--AG-UCAAAGUUUAGUGCUUGUC
Trichoderma reesei GGACC-CA--UCUCGAGU--UGG--AG-UCAAAGUUCAGUACUUGUC
Trichoderma virens GGACC-CA--UCUUGAGU--UGG--UG-UCAAAGUUUAGCACUUGUC
Nectria haematococca GGACC-CA--UCUCGAGU--UGG--AG-UCAAAGUUUCGUGCUUGUC
Chaetomium globosum GGACC-CA--UCUUGAGU--UGG--AG-UCAAAGCUUCGUGCUUGUC
Sporotrichum thermophile GGACC-CA--UCUCGAGU--UGG--AG-UCAAAGCUUCGUGCUUGUC
Cryphonectria parasitica GGACC-CA--UCUCGAGU--UGG--AG-UCAAAGUUUUGUACUUGUC
Thielavia terrestris GGACC-CA--UCUCGAGU--UGG--AG-UCAAAGCUUUGUGCUUGUC
Magnaporthe oryzae GGACC-CA--UCUCGAGU--UGG--AG-UCAAAGUCUUGUACUUGUC
Magnaporthe poae GGACC-CA--UCUAGAGU--UGG--AG-UCAAAGCUUCGCGCUUGUC
Gaeumannomyces graminis GGACC-CA--UCUAGAGU--UGG--AG-UCAAAGCUUCGCGCUUGUC
Epichloe festucae GGACC-CA--UCUCGUGU--UGG--AG-UCAGAGUUUCGUACUCGUC
Botrytis cinerea GGACC-CA--UCUCGUGU--UGG--UG-UCAAAGCUUAGUGCUUGUC
Sclerotinia sclerotiorum GGACC-CA--UCUCGUGU--UGG--UG-UCAAAGCUUAGUGCUUGUC
Geomyces destructans GGACC-CA--UCUCGAGU--UGG--AG-UCAAGGCUUAGUGCCUGUC
Aspergillus nidulans GGACC-UC--UCUCGUGU--GAG---GCUCAAAGUUUAG-GCUUGUC
Neosartorya fischeri GGACC-UCAC---AAUG-GUGGG---GAUCAAAGCUUAG-GCUUGUC
Aspergillus fumigatus GGACC-UCAC---AAUG-GUGGG---GAUCAAAGCUCAG-GCUUGUC
Aspergillus clavatus GGACC-UCAC--GAAGG-GUGGG---GAUCAAAGCUCAG-GCUUAUC
Aspergillus terreus GGA-C-UCCC--GAAC--GGGAG---GAUCAAAGCCCAG-GCUUGUC
Aspergillus aculeatus GGACC-UC----GAAA----GGG---GAUCAAAGCUCAGUGCUUGUC
Aspergillus flavus GGACC-UC--UCUAGAGU--GAG---GAUCAAAGCUUAG-GCUUGUC
Aspergillus oryzae GGACC-UC--UCUAGAGU--GAG---GAUCAAAGCUUAG-GCUUGUC
Aspergillus niger GGACC-CC----UUGAU---GGG---GAUCAAAGUUUAG-GCUUGUC
Aspergillus carbonarius GGACC-UC----UUAAUU--GGG---GAUCAAAGCUUAG-GCUUGUC
Penicillium chrysogenum GGACC-UC--UCCCGAGU--GAG---GAUCAAAGCUUUG-GCUUGUC
Penicillium marneffei GGACC-UC--UCUCGAGU--GAG---GAUCAAAGCUUCG-GCUUGUC
Paracoccidioides brasilliensis GGACC-UC--UCUAGAGU--GAG---GAUCAAAGCUUAG-GCUUGUC
Blastomyces dermatidis GGACC-UC--UCUCGAGU--GAG---GAUCAAAGCUUAG-GCUUGUC
Histoplasma capsulatum GGACC-UC--UCUCGAGU--GAG---GAUCAAAGCUUCG-GCUUGUC
Coccidioides immitis GGACC-UC--UCUUGAGU--GAG---GCUCAAAGUUUCG-GCUUGUC
Coccidioides posadasii GGACC-UC--UCUUGAGU--GAG---GCUCAAAGUUUCG-GCUUGUC
Uncinocarpus reesii GAGCC-UC--UCUCGAGU--GAG---GAUUAAAGCUUUU-GCUUAUC
Microsporum canis GGACC-UC--UCUCGGGU--GAG---GAUCAAAGCUCCG-GCUUGUC
Microsporum gypseum GGACC-UC--UCUCGAGU--GAG---GAUCAAAGCCCUG-GCUUGUC
Trichophyton rubrum GGACC-UC--UCUCGAGU--GAG---GAUCAAAGCCUUG-GCUUGUC
Trichophyton equinum GGACC-UC--UCUUGAGU--GAG---GAUCAAAGCCUUG-GCUUGUC
Trichophyton tonsurans GGACC-UC--UCUUGAGU--GAG---GAUCAAAGCCUUG-GCUUGUC
Arthroderma benhamiae GGACC-UC--UCUCGAGU--GAG---GAUCAAAGCCUUG-GCUUGGC
Trichophyton verrucosum GGACC-UC--UCUCGAGU--GAG---GAUCAAAGCCUUG-GCUUGGC
Talaromyces stipitatus GGGCC-UC--UCUAGAGU--GAG---GAUCAAAGCUUUG-GCUUGUC
Stagonospora nodorum GGACC-UC--UCAAGAGU--GAG---GCUCAAGGUCUAG-ACCUGUC
Pyrenophora tritici-repentis GGACC-UC--UCUCGAGU--GAG---GAUCAAGGUCUAG-GCCUGUC
Pyrenophora teres f. teres GGACC-UC--UCUCGAGU--GAG---GAUCAAGGUCUAG-GCCUGUC
Cochliobolus heterostrophus GGACC-UC--UCAAGAGU--GAG---GCUCAAGGUCUAG-ACCUGUC
Alternaria brassicicola GGACC-UC--UCAUGAGU--GAG---GCUCAAGGUCUAG-ACCUGUC
Leptosphaeria maculans G~GCC~UU~~UCCAGUGU~~GAG~~~G~UGCAGGCUUAG~ACCUAUC
Mycosphaerella graminicola GGACC-UC--UCUCGAGU--GAG---GAUCAAAGCUUUG-GCUUGUC
Mycosphaerella fijensis GGACC-UC--UCUAGAGU--GAG---GAUCAAAGCUUUG-GCUUGUC
Septoria musiva GGACC-UC--UCUAGAGU--GAG---GAUCAAAGUUUUGUGCUUGUC
Dothistroma septosporum GGACC-UC--UCUCGUGU--GAG---GAUCAAAGCUUUG-GCUUGUC
Hysterium pulicare GGACC-UC--UCUAGAGU--GAG--CG-UCAAAGCUUCG-GCUUGUC
Rhytidhysteron rufulum GGACC-UC--UCUAGAGU--GAG--CG-UCAAAGUUUCG-GCUUGUC
Tuber melanosporum GGAC-ACAC---ACAU---GUG-CAAG-UCAGAGCUCCGUGCUCGUC
Phymatotrichum omnivorum GGAC---------CAUGC--------G-UCAGAGUUCCGUACUCGUC
Structure -((((-((((--------)))))---)-))-((((-----))))---

Supplemental Figure S3.2 Sequence alignment of P6/6.1 from 69 filamentous fungal TER1s

The nucleotides with 90% or more identities are shaded in yellow. Helix structures are indicated under the sequence with parenthesis. S: Sordariomycetes; L: Leotiomycetes; E: Eurotiomycetes; D: Dothidiomycetes; P: Pezizomycetes.

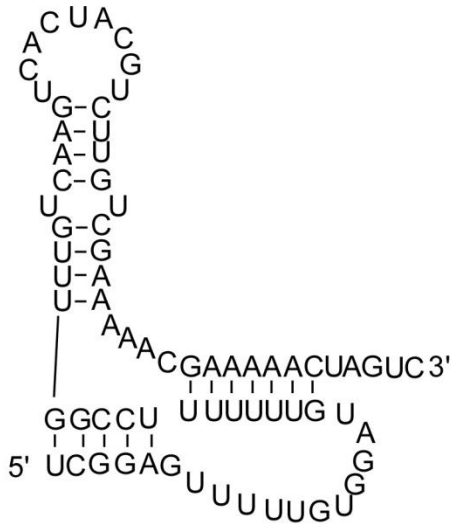


Supplemental Figure S3.3 Secondary structure of *Aspergillus nidulans* TER1 conserved regions.

A. nidulans TER1 pseudoknot and three way junction (TWJ) structures were determined from sequence alignment.

APPENDIX D

SUPPLEMENTAL INFORMATION FOR CHAPTER 4



Magnaporthe oryzae Pseudoknot

Supplemental Figure S4.1 Predicted secondary structure of *M.oryzae* pseudoknot.

The pseudoknot structure of *M.oryzae* TER1 is predicted from the phylogenetic comparative analysis (Figure 3.4).

APPENDIX E

IDENTIFICATION OF A PUTATIVE CAB BOX BINDING PROTEIN NOPP140

E.1 Abstract:

Telomerase is a unique reverse transcriptase responsible for telomere maintenance in most eukaryotic organisms including human. Telomerase core enzyme contains a catalytic protein component: telomerase reverse transcriptase (TERT) and a telomerase RNA component (TER). Telomerase holoenzyme contains proteins other than TERT, which are essential for telomerase in vivo stability and function. In human, dyskerin protein complex was identified to be associated with hTER *in vivo*, and plays an important role in telomerase RNP biogenesis. A CAB box within the Conserved Region 7 (CR7) of hTER was discovered to recruit telomerase holoenzyme into the Cajal bodies in nuclei. A hypothesis was proposed that a protein(s) interacted with the CAB box and shuttled human telomerase within nucleolus and Cajal bodies. To isolate the CAB box association protein(s), phage display system was performed for RNA binding proteins selection from human tumor cDNA libraries. C-terminal region of human Nopp140 protein, was then selected from phage display with binding affinity to the wild-type CR7 RNA, but not to CR7 mutant RNA (CR7m). Nopp140 was then expressed and purified in *E.coli* for testing binding affinity with CR7 and CR7m RNAs. An *in vivo* immunoprecipitation assay was also performed to investigate the association between hTER, dyskerin and Nopp140.

E.2 Introduction:

Telomerase is an RNA-dependent DNA polymerase that adds telomeric DNA repeats to chromosome ends (Greider and Blackburn, 1985, 1987). Human telomerase consists of two core components, a catalytic human Telomerase Reverse Transcriptase (hTERT) and the human Telomerase RNA (hTER) which serves as the template (Greider and Blackburn, 1989). hTER harbors three conserved domains (Chen et al., 2000), (Fig E.1) in which Pseudoknot and CR4-CR5 domains interact with hTERT independently and are essential for enzymatic activity *in vitro* (Mitchell and Collins, 2000). The interaction between hTERT and hTER was reported to be the only requirement for telomerase activity *in vitro* (Mitchell and Collins, 2000). However, other telomerase holoenzyme components, e.g. dyskerin protein complex, are essential for *in vivo* telomerase biogenesis and stability. hTER ScaRNA domain, originally named as SnoRNA domain, is a conserved among other ScaRNAs. Box H/ACA SnoRNA was recognized as a small RNA family which localize in the nucleolus and function for pseudouridylation formation (Smith and Steitz, 1997). A conserved CAB box (with consensus sequence 5'-UGAG-3') was identified within several SnoRNAs and was responsible for Cajal body-specific localization (Richard et al., 2003). This new group of RNA was named as ScaRNA (Small Cajal body RNA). A conserved CAB box was also identified in hTER (Jady et al., 2004), and Cajal body-specific localization was observed for human telomerase. Cajal bodies are small nuclear organelles in the nucleus, which contain proteins involved in transcription and processing of nuclear RNAs. Cajal body localization of hTER

was then proposed to be associated with hTER and human telomerase biogenesis. Further study revealed that the Cajal body-specific localization of human telomerase was essential for human telomerase *in vivo* function as elongating the telomeres (Cristofari et al., 2007).

In order to study the function of human telomerase Cajal body-specific localization, a hypothesis was proposed that a protein(s) is responsible for the recognition of CAB box and transferring ScaRNAs including hTER to Cajal bodies. Several techniques were developed for selection of novel RNA binding proteins, including phage display, mRNA display and yeast three-hybrid. The T7 phage display system (Danner and Belasco, 2001) was utilized in this study for the selection of CAB box binding protein. Phage display is a powerful technique to study Protein-Ligand interactions *in vitro*. A gene of interest is fused to that of a phage capsid protein (capsid 10 protein in T7 phage), resulting in phage particles displaying the encoded protein as well as harboring its coding sequence. The phage bound to the ligand will then be amplified in *E.coli* cells and used for the next round of selection.

In our system, T7 phage human tumor cDNA libraries (Novagen) were used in which the protein library was fused to its capsid protein in order to be displayed on the surface of the phage. A putative CAB box binding protein, Nopp140, was selected independently from two T7 phage libraries responsible for hTER CR7 specific interaction. However, when other systems were performed for the verification of RNA-protein specific interaction, only non-specific interaction was observed.

E.3 Materials and methods:

E.3.1 RNA manipulation

T7 promoter based RNA *in vitro* transcription using PCR DNA as template was described previously. RNA was purified from 8M urea denaturing gel and dissolved in dH_2O. Concentration of RNAs was measured with UV260nm absorbance by using Nanodrop-1000 (NanoDrop). DNA-RNA duplex was obtained by heat denaturing 1.25μM of each DNA oligo and purified RNA in 1x annealing buffer containing 10mM Tris-HCl pH 7.5, 100mM NaCl and 1mM EDTA at 80^oC for 2min and slowly cools to room temperature. To measure the duplex forming efficiency, DNA-RNA duplex and DNA, RNA alone were resolved in 8% polyacrylamide 8M urea gel, and stained with Ethidium Bromide (EtBr). To investigate the RNA stability in phage library (*E.coli* cell lysate), 1μM of RNAs were incubated with 18μl of cell lysate (phage) at room temperature for 1 hour, and resolved onto a urea denaturing gel.

E.3.2Phage libraries amplification

Two T7 phage libraries: T7select human breast tumor cDNA library and T7select human colon tumor cDNA library were obtained from Novagen (EMD Millipore now), with primary clones as $1.6x10^7$ and $1.3x10^7$ respectively. *E.coli* strain *BLT5615rna* (Genotype: F^- *rna::kan ompT* gal $hsdS_B$ *dcm lac* pAR5615) was inoculated in 200ml of LB liquid media until OD600 reaching 0.6, and IPTG was added to 1.0mM. T7 phage library was inoculated into the *E.coli* cells 0.5 hour after IPTG was added. The LB media was centrifuged at 4,000g for 20min

after *E.coli* cells were lysed, and supernatant containing T7 phage was stored at 4^{o}C. Phage plaque assay was then performed to measure the phage titer.

E.3.3 T7 phage Plaque assay

Ten fold serial dilutions of T7 phage libraries amplified from *E.coli* strain *BLT5615rna* was performed with sterile LB medium as diluent until reaching 10^{-10} dilution (Fig E.1). 100µl of each T7 phage dilution (10^{-10} to 10^{-8}), 100µl of 100mM IPTG and 3ml of 50^{o}C pre-melted top agarose (0.5% agarose autoclaved) was mixed in a sterilized 12ml tube containing 250 µl *E.coli BLT5615rna* cells (OD600nm=1.0) and poured onto the LB plates with corresponding antibiotics (50µg/ml ampicillin and 15µg/ml kanamycin). Plates were incubated at 37^{o}C for 3-4hours, and T7 phage plaques were checked and counted for calculation of phage titer of each dilution.

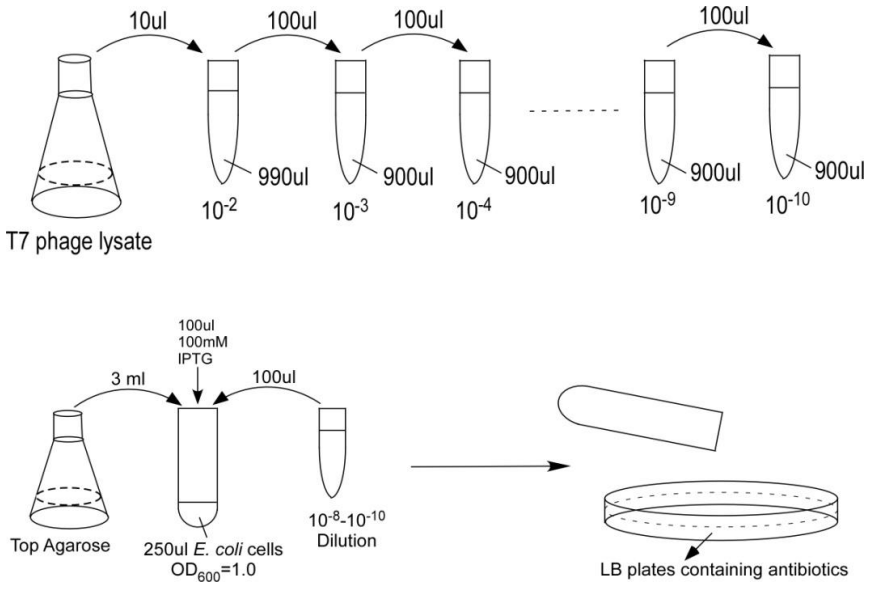


Figure E.1 Schematic of T7 phage plaque assay

T7 phage amplified from *E.coli* was serially diluted as indicated with LB medium. 100µl of diluted phage, 100µl of 100mM IPTG, and 3ml of top agarose were mixed with 250µl *E.coli* cells and plated onto the LB plate.

E.3.4 Phage display selection

E.coli strain *BLT5615rna* was inoculated in LB liquid medium from an overnight culture (1:100 dilution) and grown at 37^{o}C until the OD600nm reaches 0.5. IPTG was added to the culture to final concentration 1mM and cells were grown for another 30 min before phage infection. 10-120nM RNA-DNA duplex (concentration of RNA-DNA duplex decreases about 20nM with each round of selection) prepared as described in RNA manipulation was mixed in a low retention siliconized tube with 1-5x10^{9} pfu phage (number of phage decreases with each round of selection cycle), 50 μg of *E.coli* tRNA and 20 units of SUPERase IN (Ambion) in 1x TENT buffer: 10mM Tris-HCl, pH 8.0, 1mM EDTA, 250mM NaCl and 0.5% Triton X-100. The selection mixture was incubated at room temperature for 50min with gentle rotation, with 10μl of prewashed Dynabeads M-280 Streptavidin (10mg/ml) added after 20min. Dynabeads were separated with a magnet and washed twice and resuspended in 1xTENT buffer. The bound phage was used without release from the beads to infect a log-phage *E.coli* culture (5ml, described previously). The amplified T7 phage was harvested by centrifugation at 8,000g for 10min after *E.coli* cells were completely lysed, and was utilized for the next round of selection cycle. 1μl of phage harvested from each round of selection was directly utilized without purification of DNA for PCR amplification with 0.5μM of each primer flanking the cDNA insertion: 5'-AGGAGCTGTCGTATTCCAGTC-3' and 5'-AACCCCTCAAGACCCGTTTAG-3' in 1x NEB Taq buffer and 1U of NEB Taq DNA polymerase (New England Biolabs). PCR was carried out with the initial

94^{o}C denaturation for 4min following with 30 cycles of 94 ^{o}C 30s, 51 ^{o}C 30s and 72 ^{o}C 2min, and a final extension at 72 ^{o}C for 6min. PCR products were resolved in a 1% agarose gel. Specific PCR DNA bands were gel purified with Wizard SV Gel and PCR clean-up system (Promega), and cloned into pCR4-TOPO vector (Invitrogen) and sequenced.

E.3.5 Protein expression and purification

Nopp140 and Nopp140-C34 (C-terminal domain of Nopp140 identified in phage display with 34kDa) gene was RT-PCR amplified from human total RNA and cloned into pRSF vector (EMD). Nopp140 protein was expressed in BL21(DE3) cells with T7 promoter and purified as described previously. Briefly, E.coli cells were harvested after 18^{o}C 3 hours IPTG induction when OD600nm reaching 0.6. Cell lysate, obtained as supernatant after sonification, was heat denatured at 90^{o}C for 30min. Supernatant containing Nopp140 protein was collected by centrifuging at 14,000g for 20min, and loaded onto the SP-Sepharose column (16x26) equilibrated with 20mM Sodium phosphate pH 7.5 and 200mM NaCl. The SP-Sepharose colum was then washed with 400ml of 0.2M-0.7M NaCl gradient buffer containing 20mM sodium phosphate pH 7.5, with 5ml of each fraction collected. 10μL of each fraction was resolved onto an 8% SDS-PAGE gel and stained with PAGE blue (Fermentas). Fractions with Nopp140 protein were combined and dialyzed against 20mM sodium phosphate pH 7.5, and 50mM NaCl overnight at 4^{o}C. The dialyzed protein sample was loaded onto Hydroxyapatite column (16x26) equilibrated with 20mM sodium phosphate buffer pH 7.5. The column was then washed with 400ml of 20mM-500mM gradient sodium

phosphate buffer pH 7.5, and 5ml of each fraction was collected. Each fraction was checked again with 8% SDS-PAGE gels and fractions harboring fairly pure Nopp140 protein was combined and concentrated with Centriplus YM-10 (Amicon). Nopp140-C34 protein was expressed and purified the same method as Nopp140 protein except only SP-Sepharose column was utilized as Hydroxyapatite column was not necessary.

E.3.6 Gel mobility shift assay (EMSA)

T7 promoter *in vitro* transcribed RNAs were dephosphorylated by 1 unit of Calf Intestine Phosphatase (New England Biolabs) in 1x NEB buffer 3 at 37^{o}C for 30 min. Dephosphorylated RNAs recovered from phenol/chloroform extraction and ethanol precipitation were then kinased by T4 polynucleotide kinase (T4 PNK, New England Biolabs) with 1x T4 PNK buffer and 50μCi γ-^{32}P-ATP (Specific acitivity: 7000Ci/mmol, MP biomedicals). Radiolabeled RNA oligonucleotides were purified with quick spin G-25 columns (GE Healthcare) following manufacturer's instruction. The 5' ^{32}P end-labeled RNAs (~10nM) were incubated with different concentration of purified Nopp140 protein (1-10μM) and 5mg of *E.coli* tRNA in 1x binding buffer (20mM Tris-HCl pH 8.0, 50mM KCl, 5mM $MgCl_2$, 1mM β-mercaptoethanol, 10% glycerol and 0.1mg/ml BSA) at room temperature for 20min. The RNA-protein mixtures were resolved in a 4% native PAGE gel, which was then dried in a gel drier and exposed with Phosphorimager (Bio-Rad).

E.3.7 UV crosslinking assay

5’ ^{32}P end-labeled RNAs (~10nM CR7 and CR7m) were incubated with 2μM of purified Nopp140 protein in 1x binding buffer. 300nM of cold CR7 and CR7m RNAs were added to some of the mixtures as RNA-protein binding competitors. Mixtures were incubated at room temperature for 5min, and UV crosslinking was carried out by UV-crosslinker for 2min. Samples were heated at 90^{o}C for 5min with 1x SDS loading buffer added, and were resolved in an 8% SDS-PAGE gel. Gel was then dried in a gel drier and exposed and analyzed by Phosphorimager.

E.3.8 Phage pull down assay

Phage pull down assay was performed as a repeat experiment for investigation of interactions between Nopp140 displaying T7 phage and CR7 bait RNA. T7 phage with Nopp140 protein coated on the surface was obtained by screening the single phage plaque with Nopp140 gene specific PCR analysis. Nopp140 displaying T7 phage was then amplified and diluted 100fold with the original T7 phage library (human breast tumor cDNA library). This Nopp140 T7 phage diluted phage library was input for the T7 phage display together with CR7 and CR7m bait RNAs as well as various competitors. The bound phage after washing was then utilized for PCR analysis specifically for Nopp140 cDNA insert.

E.3.9 293FT cell transfection

293FT cell (Invitrogen) transfection was carried out by using Fugene-HD transfection reagent (Roche) as described previously. Briefly, 293FT cells were seeded onto a 6-well plate in DMEM medium supplemented with 10% FBS and 1x antibiotics mix (Penicillin, Streptomycin and Amphotericin B, from Lonza)

and grown at 37^{o}C with 5% CO_2. After reaching 80-90% confluency, 293FT cells from each well were transfected with 2μg of plasmid DNA (pIRES2-EGFP-2xHA-Nopp140) using Fugene-HD transfection reagent following manufacturer's instruction. DMEM medium was changed once 24 hours post transfection, and cells were collected 48 hours post transfection by 1X trypsin (0.05% trypsin, Invitrogen).

E.3.10 Immunoprecipitation and western blot

Transient transfected 293FT cells were lysed with 1xCHAPS lysis buffer (10mM Tris-HCl, pH 7.5, 400mM NaCl, 1mM EGTA, 0.5% CHAPS, 10% glycerol, supplemented with 1x protease inhibitor cocktail (Roche) and 5mM β-mercaptoethanol before use) at 4^{o}C for 10min with gentle rotation. Cell lysate was obtained by centrifugation of cell resuspension at 16,000g for 10min. 2xHA-Nopp140 expressed in 293FT cells was immunoprecipitated by using anti-HA F7 antibody conjugated agarose beads (Santa Cruz biotechnology). SBP2-hTERT was immunoprecipitated with streptavidin coated magnetic beads. 300μl of cell lysate (in 1xCHAPS lysis buffer) was incubated with 20μl of beads pre-washed with 1xCHAPS lysis buffer for 2 hours at 4^{o}C. The beads were then washed three times with CHAPS lysis buffer and heat at 90^{o}C with 1x SDS loading buffer. Bound protein samples were resolved in an 8% SDS-PAGE gel and electro-transferred onto PVDF membrane (Bio-rad). Western blot analysis was carried out as described previously by using anti-nopp140, anti-dyskerin antibodies (Santa Cruz Biotechnology) respectively.

E.4 Results:

E.4.1 Bait RNAs design and RNA stability and secondary structure check

Two RNA baits were designed for phage display: wild-type CR7 (hTER 395-432) and CR7 mutant (CR7m). C401was specifically mutated to U, which will support better base pairing with A422. Compensatory mutations were made for the conserved sequences of top stem loop (from hTER 407-422) in CR7m RNA. A single stranded RNA tail was designed to the 3' end of both RNAs in order to form a duplex with a biotinylated DNA primer. The RNAs were *in vitro* transcribed and annealed to the biotinylated DNA primer as described in materials and methods.

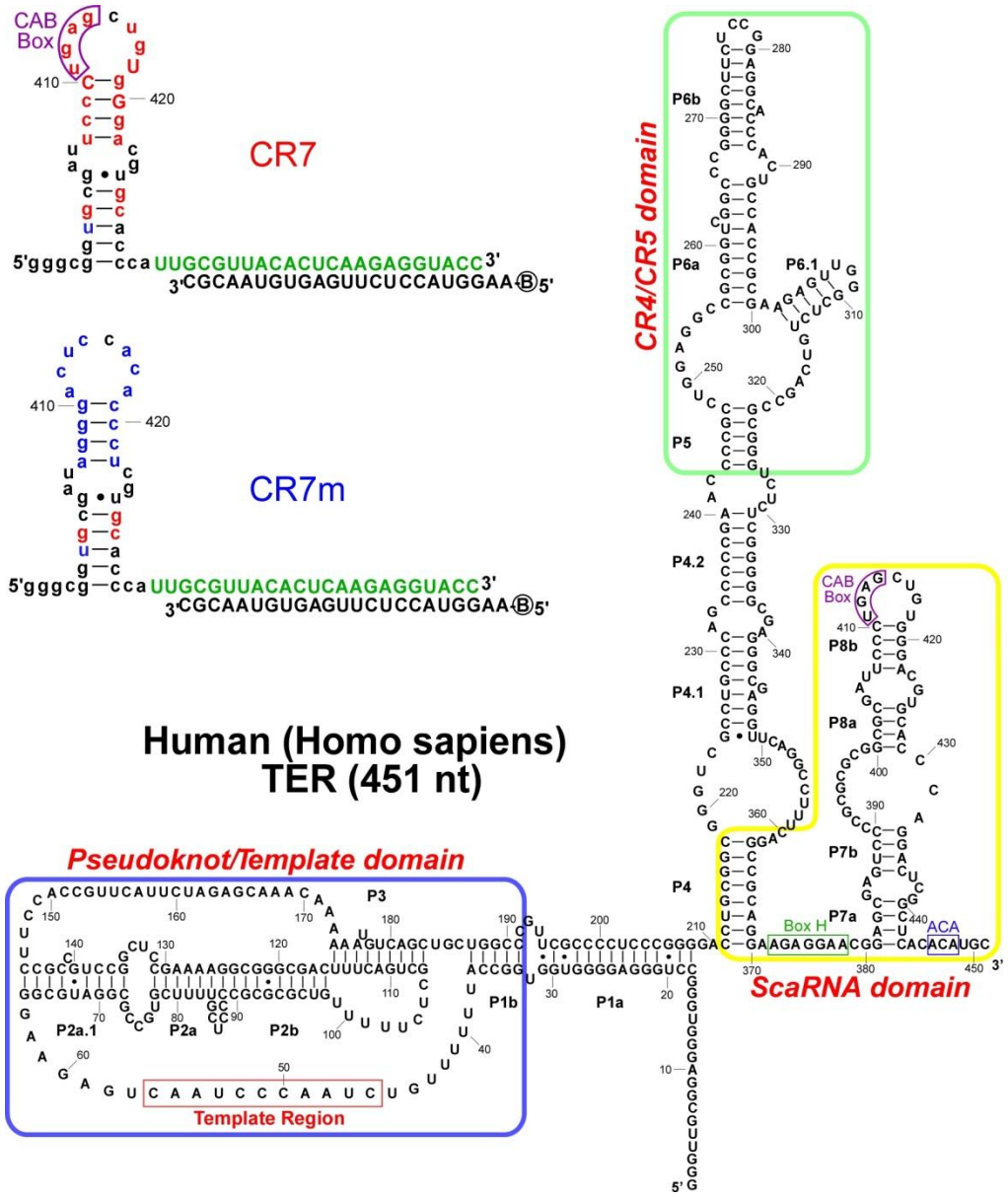


Figure E.2 Secondary structure of hTER and CR7, CR7m.

hTER full length secondary structure was shown (Chen et al., 2000). Three conserved domains were shown as red color. The CR7 and CR7m bait RNA structures were shown on the top right with biotinylated DNA oligonucleotide annealed. Nucleotides with 90% or more identity were labeled as red. The RNA nucleotides mutated in CR7m were shown in blue. 3' tail sequences were indicated as green. Circled B indicates the biotin attached to the 5' end of DNA oligonucleotide.

Amplification of phage libraries was carried out with *E.coli BLT5615rna*, a strain derived from *BLT5615* in which the *E.coli* RNase I gene was disrupted. RNAs in the lysate of this *E.coli* strain was reported to be more stable, as phage libraries were collected after *E.coli* cells lysis. To check if CR7 and CR7m RNAs remain intact within the cell lysate, we performed the RNA stability test. RNAs were incubated with T7 phage amplified and harvest in LB medium as described in materials and methods. RNAs were resolved in a 6% polyacrylamide 8M urea gel, and stained with EtBr. CR7 and CR7m RNAs were stable in the phage lysate, as the major RNA bands remain intact. Other DNA or RNA bands shown in the phage lysate are *E.coli* total RNA components. This result verifies that *E.coli* strain *BLT5615rna* is indeed suitable for phage display to select RNA binding proteins.

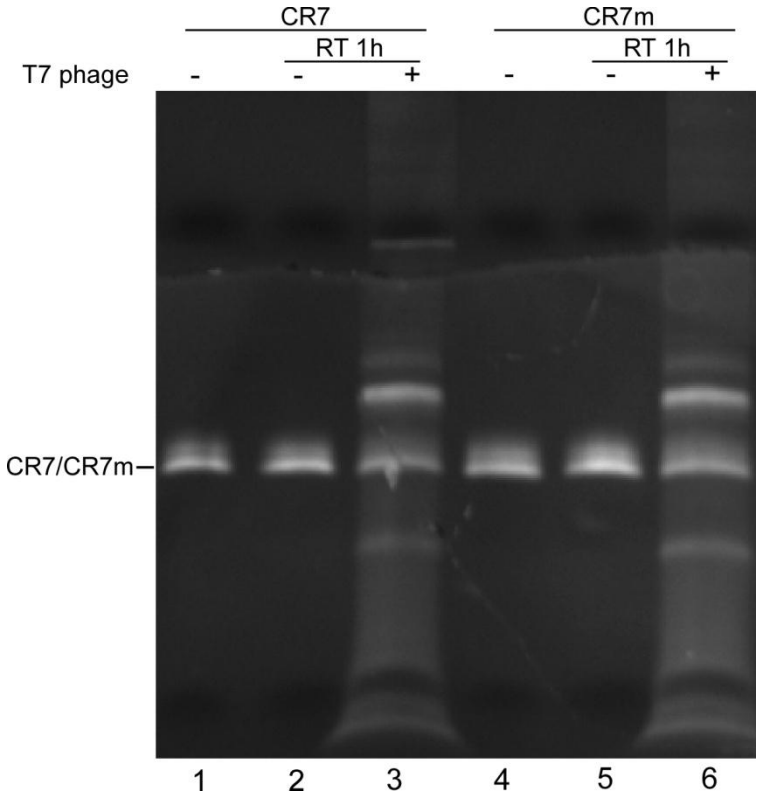


Figure E.3 RNA stability test in T7 phage lysate

CR7 and CR7m were mixed with dH_2O (lane 2 and 5) or T7 phage (lane 3 and 6) and incubated at 37^oC for 1h. RNAs were resolved onto a 6% polyacrylamide 8M urea gel and stained with Ethidium Bromide. All RNAs were intact comparing to the untreated RNAs (lane 1 and 4). Some RNA or DNA bands in lane 3 and 6 were *E.coli* RNA or DNA fragments.

RNA secondary structure is usually crucial for its function, thus it is important to verify the bait RNAs secondary structure as a hairpin. To test this, bait RNAs as well as annealed RNA-DNA duplexes were resolved on a 6% native PAGE gel (Fig E.4). The RNA-DNA duplexes from both CR7 and CR7m RNA form a single band on the gel with a weak and smaller biotinylated DNA oligonucleotide band. The CR7 and CR7m alone, however, generate two smear bands, with the smaller one migrating similarly to the RNA-DNA duplex. The larger band predicted as the RNA dimer disappeared upon DNA annealing. This indicated that the RNA-DNA duplexes tend to form monomer with the correct hairpin structure as predicted.

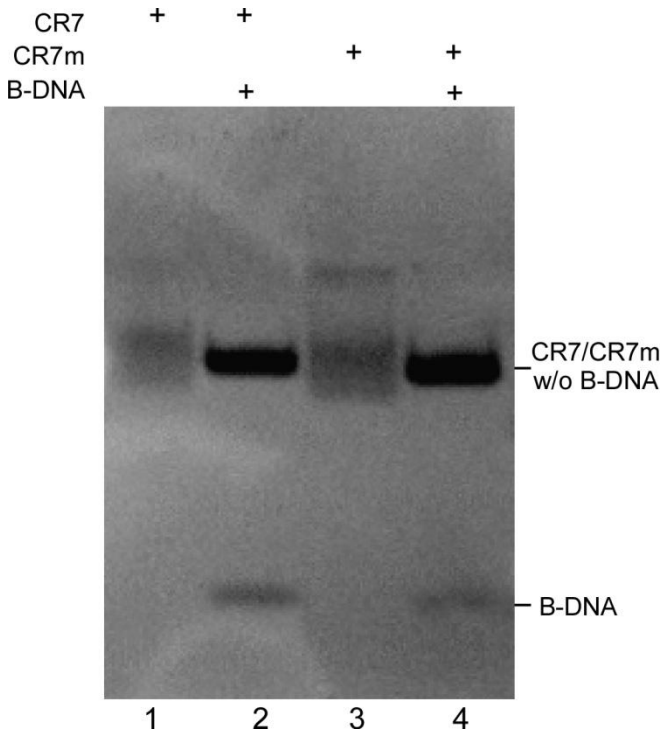


Figure E.4 CR7 and CR7m RNA secondary structure test

CR7 and CR7m RNA alone (lane 1 and 3) or annealed with biotinylated DNA primer (lane 2 and 4) were resolved in a 6% native PAGE gel and stained with Ethidium Bromide. Free biotinylated DNA primer was shown at the bottom as well as DNA-RNA hybrid.

E.4.2 Phage display with CR7 and CR7m RNAs as bait RNAs

Bait RNAs were obtained by annealing CR7 and CR7m with biotinylated DNA oligonucleotide as described in the materials and methods. Phage display was initially carried out using bait RNAs for T7 select breast tumor cDNA library phage selection. Phage display procedures are described in detail in the materials and methods, and also shown in figure E.5.

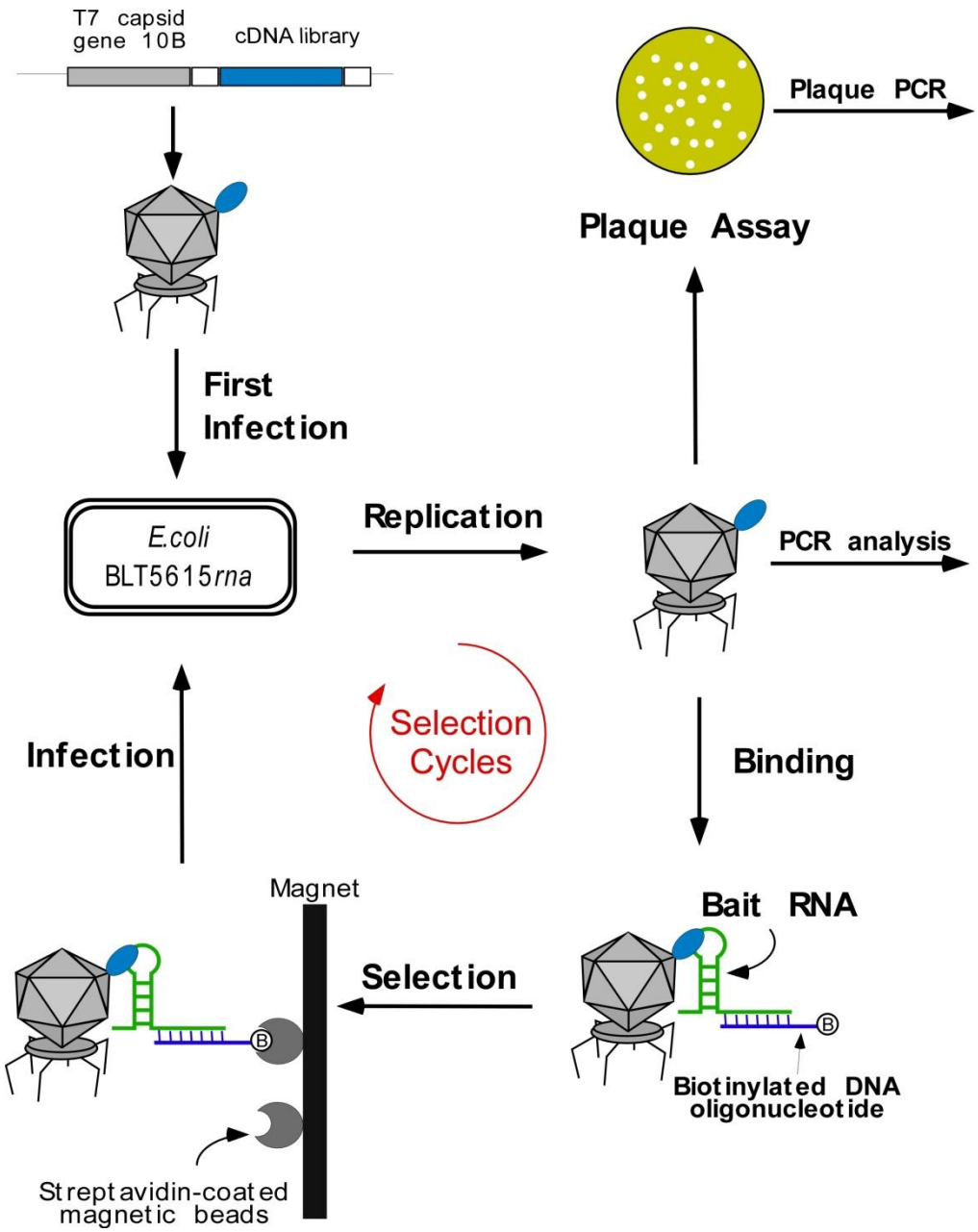


Figure E.5 Schematic of T7 phage display selection cycle

T7 phage libraries were originally amplified in *E.coli* strain *BLT5615rna*. T7 phage was then utilized for bait RNA selection. Biotinylated DNA oligonucleotide was separated with streptavidin-coated magnetic beads and the bound phage was utilized for E.coli infection and phage amplification. Plaque assay was performed after the T7 phage amplification for verification.

T7 phage titer was obtained for the first three rounds of selection with CR7 bait, and was shown in table 1. Consistent phage titer suggested promising phage selection for CR7 RNA bait binding. After six rounds of selection cycles, T7 phage was amplified for the cDNA insertion and result was shown in fig E.6. Two DNA bands were identified with CR7 bait RNA selection after 4-5 rounds of selections, whereas no significant bands were detected with CR7m bait RNA. The only strong band showing with CR7m bait RNA was seen in 4^{th} round, which disappeared ever since, suggesting a non-specific binding. The two DNA bands labeled as L and S, were cloned and sequenced respectively. The cDNA sequences from those two bands were blasted in NCBI. DNA band S only encodes for a 17 amino acid residue peptide which was then believed to be a non-specific product. The PCR DNA band L, according to its open reading frame, expresses the C-terminal region (from 386 to 699 amino acid residues out of 699) of human nucleolar phosphoprotein (hNopp140, refers to Nopp140 hereafter), a protein associated with dyskerin in human.

Table E.1

T7 phage titer from plaque assay

# of selection cycles	T7 phage plaque numbers			Phage Titer (pfu/ml)
	10^{-8} dilution	10^{-9} dilution	10^{-10} dilution	
1	59	5	1	$5.5x10^{10}$
2	88	7	0	$7.9x10^{10}$
3	157	13	1	$1.4x10^{11}$

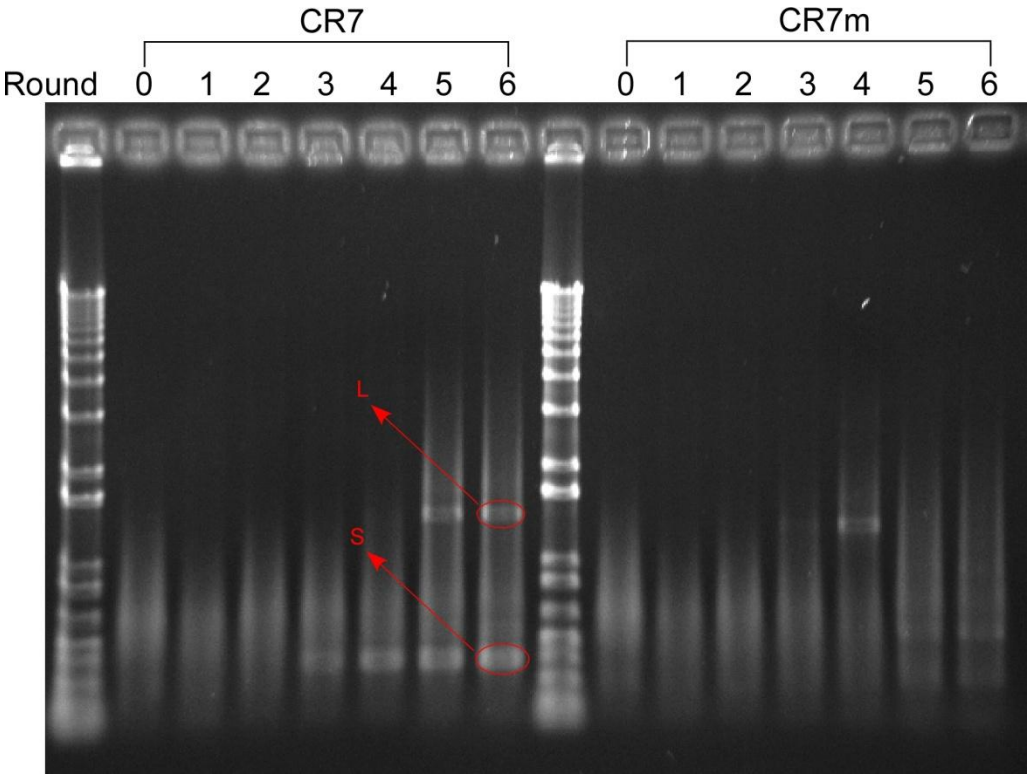


Figure E.6 PCR analysis for phage display with human breast tumor cDNA library

PCR was carried out after 6 rounds of selections with primers flanking the cDNA inserts. Specific bands from CR7 bait RNA selection (L and S) were gel extracted and cloned into pZero vector. Sequences of insert were blasted in NCBI.

In order to confirm this phage display result, another phage library: T7 select human colon tumor cDNA library was utilized for another set of phage display selection cycle. Conditions were maintained the same, and PCR analysis was carried out after six rounds of selection cycles. Similar results were observed with 3 bands (CR7-L, CR7-M and CR7-S) shown for CR7 bait RNA selection, while 3 different-size bands (CR7m-L, CR7m-M and CR7m-S) were also present for CR7m bait RNA selection. Each PCR product was gel extracted, cloned and sequenced the same method as described previously. The blast result was shown in Table E.2.

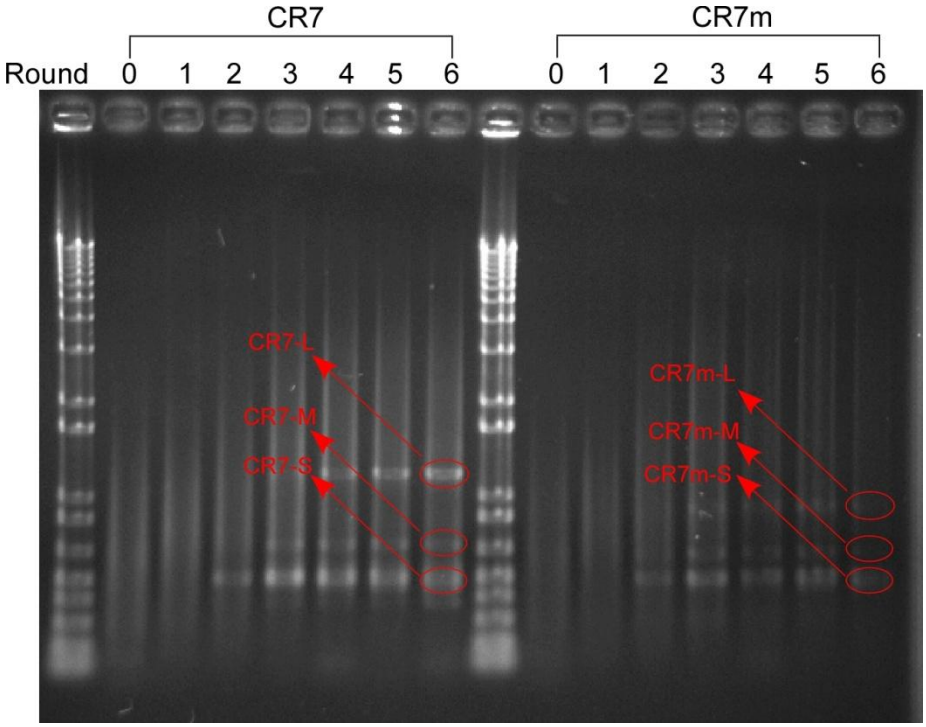


Figure E.7 PCR analysis for phage display with human colon tumor cDNA library

PCR was carried out after 6 rounds of selections with primers flanking the cDNA inserts. Three bands from CR7 and CR7m bait RNA selection (L, M and S) were gel extracted and cloned into pZero vector. Sequences of insert were blasted in NCBI.

Table E.2

Blast results for each DNA band

DNA Band	Blast results
CR7-L	hNopp140
CR7-M	N/A
CR7-S	Pinin, propionyl CoA Carboxylase
CR7m-L	PI-3-kinase-related kinase SMG-1 isoform 1homolog
CR7m-M	CP110
CR7m-S	Pinin, propionyl CoA Carboxylase

N/A: no significant blast result.

The CR7-L, again, encodes for the same C-terminal region of Nopp140 protein, with shorter 3' UTR region in the cDNA insert. This strongly suggests that Nopp140 specifically associates with the hairpin region of CR7 RNA but not CR7m RNA. Nopp140 protein was first discovered in rat as a nuclei localization signal (NLS) binding protein (Meier and Blobel, 1990), which shuttles between cytoplasm, nucleus and Cajal bodies (Meier and Blobel, 1992). It interacts with dyskerin protein complex (Meier and Blobel, 1994), which associates with hTER and plays an important role in human telomerase RNA biogenesis (Mitchell et al., 1999). This strongly indicates that Nopp140 protein is the key component between hTER and dyskerin complex.

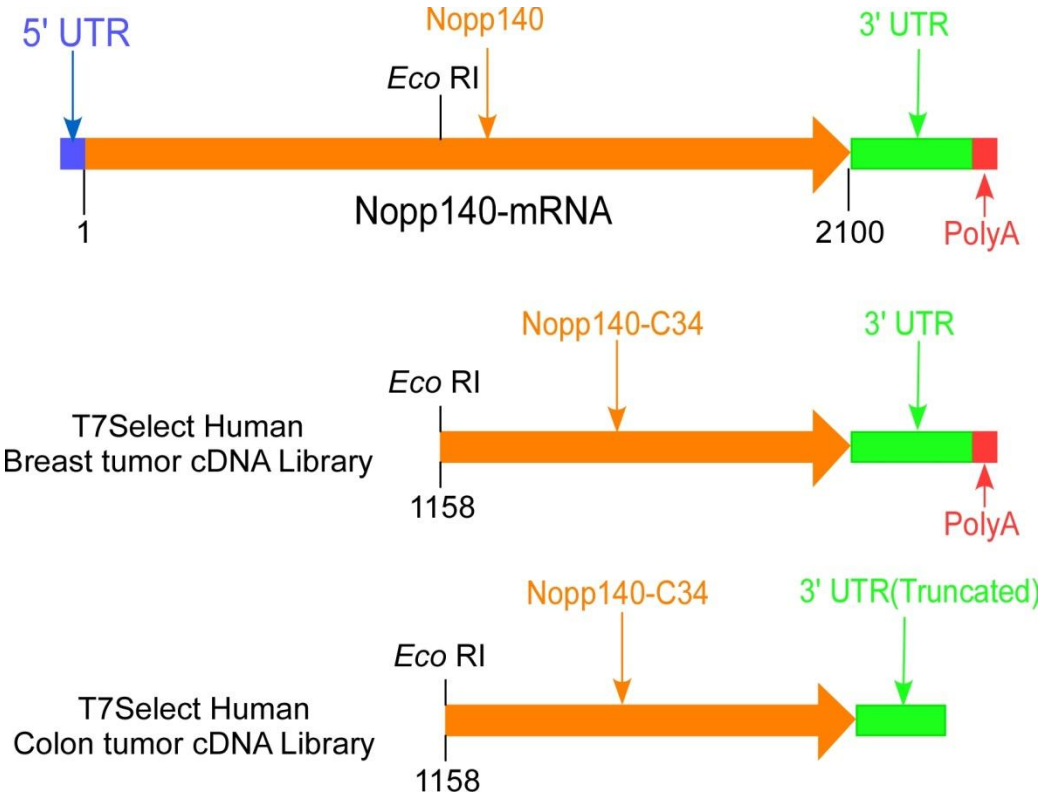


Figure E.8 Schematic of Nopp140 cDNA inserts in phage libraries.

Nopp140 mRNA structure was shown with 5' UTR labeled as blue, coding region as orange, 3' UTR as green and red for poly A tail. EcoR I site in the Nopp140 gene was indicated and utilized for T7 phage libraries construction. The breast cDNA insert in the T7 phage started from Eco RI site and ended with the polyA tail, whereas the colon tumor cDNA insert in the phage ended with truncated 3' UTR sequence.

E.4.3 Expression and purification of Nopp140 and Nopp140-C34 proteins in *E.coli*

Full length Nopp140 and Nopp140-C34 protein genes were amplified from human total RNA and cloned into pRSF vector. Expression was carried out in *BL21(DE3) E.coli* strain with IPTG induction. The entire purification procedure was performed following the previous report and also described in materials and methods part. Nopp140 protein expressed in human cells is normally hyper-phosphorylated by Casein Kinase II (CKII), resulting in ~80 phosphorylation cites. Purified Nopp140 was then kinased with CKII in order to confirm the correct protein folding and property. The mobility of Nopp140 decreased upon CKII kinasing, suggesting the purified Nopp140 protein could be phosphorylated in vitro. The un-phosphorylated Nopp140 was then used to investigate the protein-RNA interactions, since this is the same form displayed on the surface of the T7 phage library.

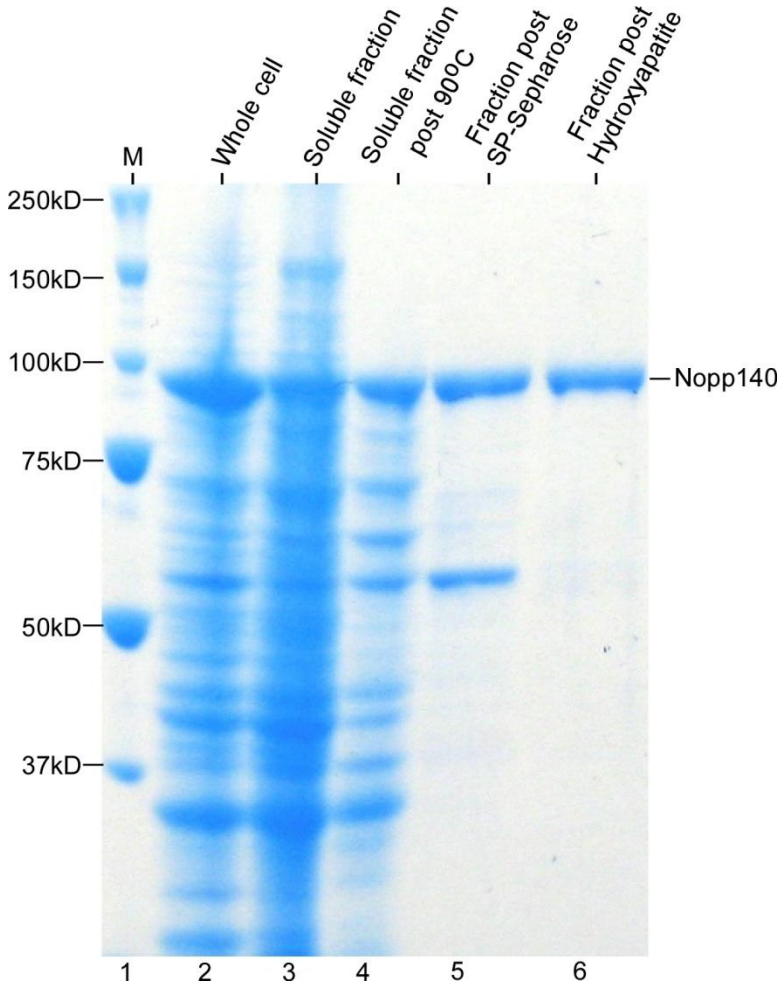


Figure E.9 Purification of Nopp140 from *E.coli*

Nopp140 expressed and purified from *E.coli* was shown for each purification step (lane 2 to 6). Size marker was shown in lane 1 and with sizes labeled on the left.

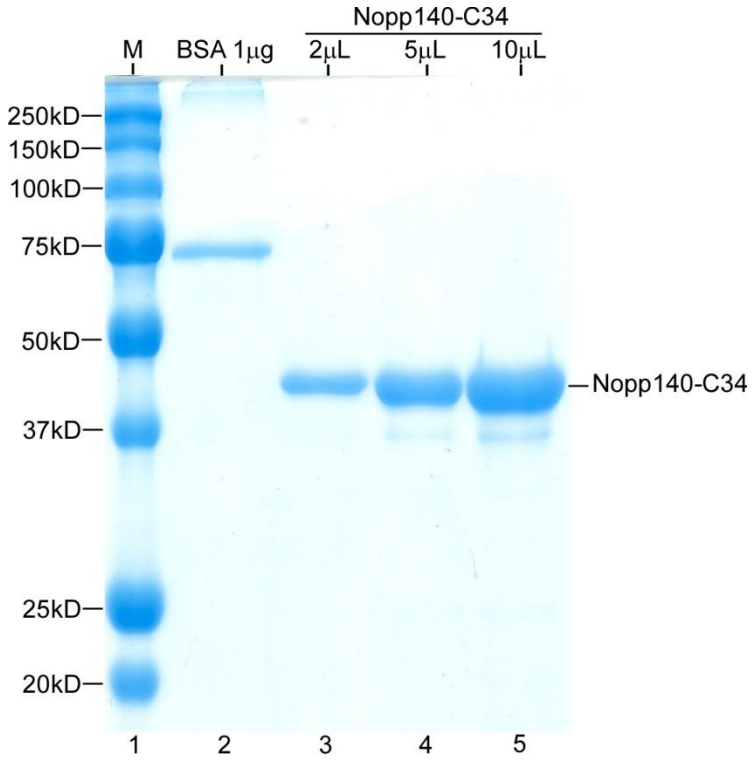


Figure E.10 Purified Nopp140-C34 protein from *E.coli*

Different amount of purified Nopp140-C34 protein was loaded onto the 12% SDS-PAGE gel (lane 3 to 5) as described in materials and methods. 1mg of BSA protein was loaded in lane 2 as the comparison of protein concentration. Size marker was shown in lane 1 with sizes labeled on the left.

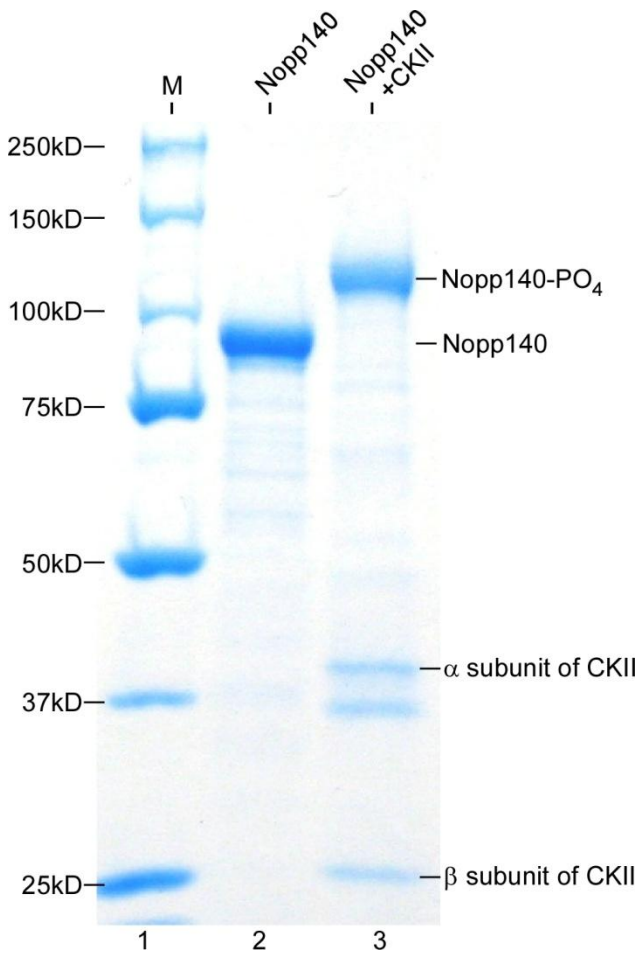


Figure E.11 Phosphorylation of Nopp140 by Casein Kinase II

Purified Nopp140 protein treated with CKII (lane 3) was loaded onto the 8% SDS-PAGE gel with the comparison of un-treated protein (lane 2). A clear gel mobility shift was seen upon the phosphorylation. The α and β subunits of CKII were shown as well as sizes of protein marker.

E.4.4 Gel mobility shift assay

Gel mobility shift assay is widely utilized for protein-DNA and protein-RNA analysis. To investigate the specific interactions between Nopp140 and CR7 RNA, we carried out this gel shift assay by using radiolabeled CR7 RNA and *E.coli* purified Nopp140 protein as described in materials and methods. With the increase of the Nopp140 protein concentration, more free CR7 RNA shifted up,

indicating more Nopp140-CR7 complex was formed. The CR7 and biotinylated DNA duplex generate the same mobility shift suggesting the binding between CR7 and Nopp140 protein is independent of 3' single-stranded tail. The CR7m RNA which acts a negative control, however, generated the same mobility shift pattern. Both RNA-protein complexes, as seen in fig E.12, located inside of the well, which suggests that the RNA-protein forms a large molecular weight complex before or after loading into the native gel. Precipitation of purified Nopp140 protein was observed upon dialysis with low salt buffer, which could be the cause of this precipitation of RNA-protein complex inside of the well, as the native gel running buffer (0.5xTBE) harbors low ionic strength condition. Gel shift assay performed with Nopp140-C34 protein generated similar result. This could be explained that Nopp140 contains 10 repeats of acidic and basic regions which might require high ionic strength to be stabilized in solution. To overcome this problem, UV crosslinking assay was performed with Nopp140 and CR7 in a high ionic strength buffer.

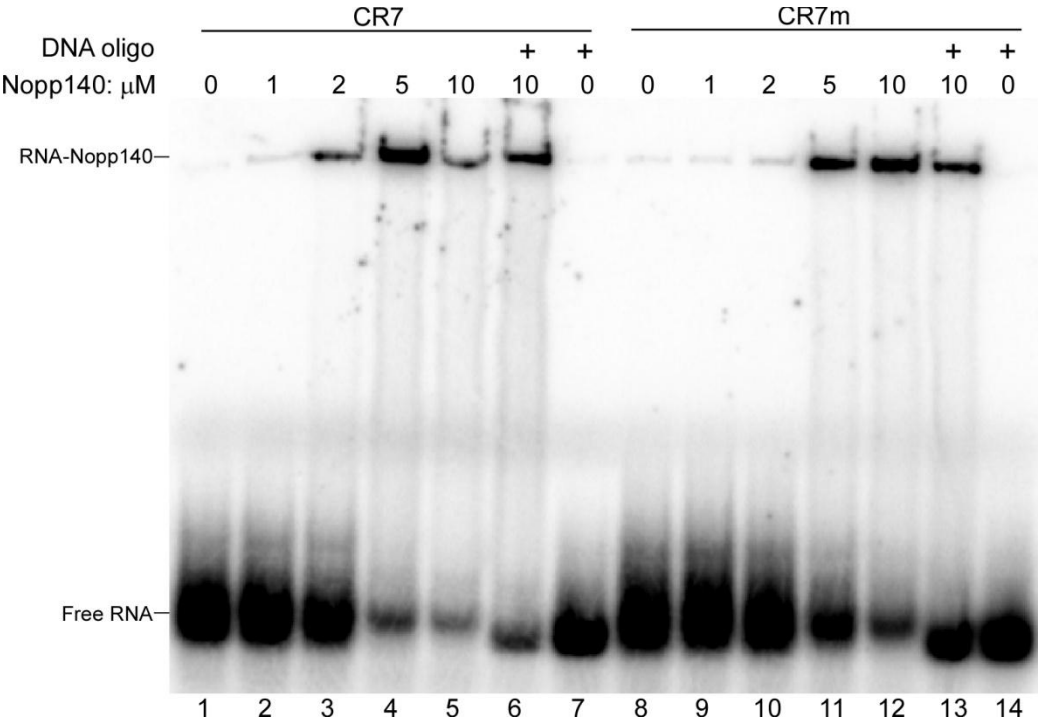


Figure E.12 Gel mobility shift assay of CR7 and CR7m with purified Nopp140 protein

5' end radiolabeled CR7 and CR7m RNAs were incubated with different amount of purified Nopp140 protein (lane 1-5 and 8-12). Biotinylated primer was annealed with CR7 and CR7m RNA for Nopp140 binding in lane 6 and 13. Samples were resolved in a 6% native PAGE gel. Free RNAs were labeled as well as RNA-Nopp140 complex.

E.4.5 UV crosslinking

UV crosslinking was performed using 5' ^{32}P end-labeled CR7 and CR7m and purified Nopp140 protein as described in the materials and methods. Both CR7 and CR7m crosslinked to the purified Nopp140 with similar intensity, suggesting the interaction between protein and RNA was likely to be non-specific (Fig E.13 lane 1 and 4). Also, with 30-fold non-radiolabeled RNA presence, the intensity of both crosslinked bands reduced dramatically, suggesting CR7 and CR7m RNA bound to the same region(s) of Nopp140 non-specifically. Purified

Nopp140-C34 protein was also utilized for crosslinking experiment and similar result was observed (data not shown).

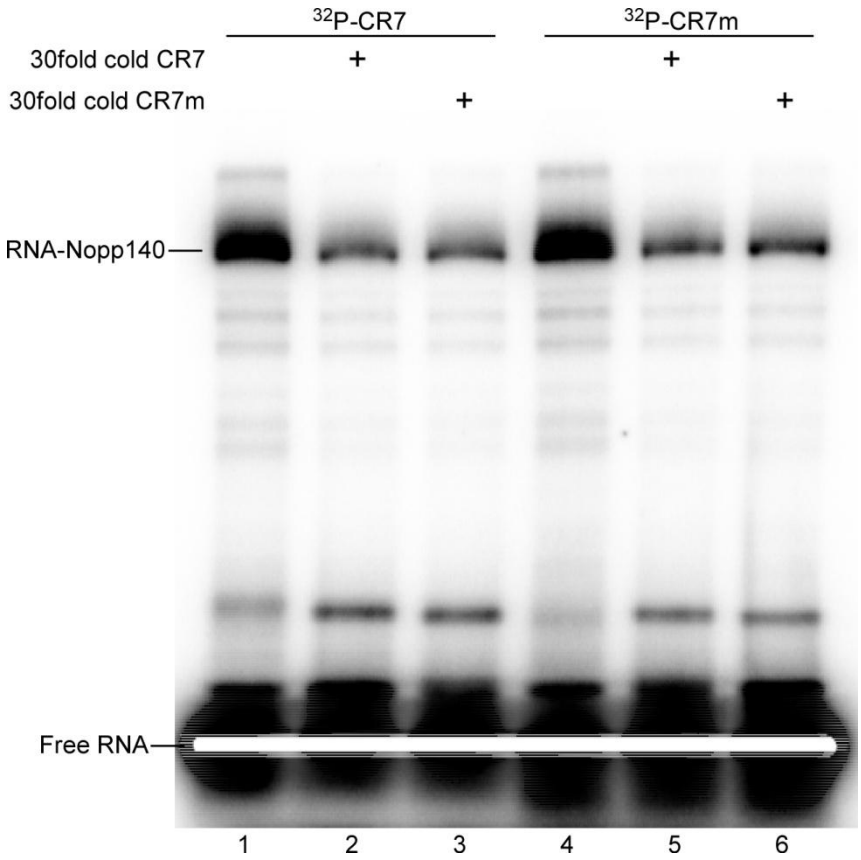


Figure E.13 UV crosslinking of Nopp140 and RNAs

5' end ^{32}P labeled CR7 and CR7m were crosslinked to the purified Nopp140 (lane 1 and 4) with 30 fold non-radiolabeled CR7 or CR7m as competitors (lane2, 3, 5 and 6). Crosslinked samples were resolved in an 8% SDS PAGE gel. Free RNA and RNA-Nopp140 complex were indicated.

E.4.6 Phage pull down assay

As both gel mobility shift assay and UV crosslinking assay failed to prove the specific interaction between Nopp140 and CR7, we designed a phage pull down assay, a similar assay to phage display, to investigate the associations. Nopp140 displaying T7 phage was screened from phage plaques and diluted in the original T7 phage library and one round of selection cycle was performed.

PCR amplification of Nopp140 gene from the bound phage was carried out as described in the materials and methods. Nopp140 gene was amplified with CR7 as the bait but not with CR7m. 20fold CR7 competitor, but not CR7m, inhibited the Nopp140 phage and CR7 binding suggesting CR7 but CR7m specifically bound to Nopp140 phage. The single stranded biotinylated DNA primer, however, exhibited non-specific binding with Nopp140 phage, as its binding could be inhibited by other single stranded DNA primers.

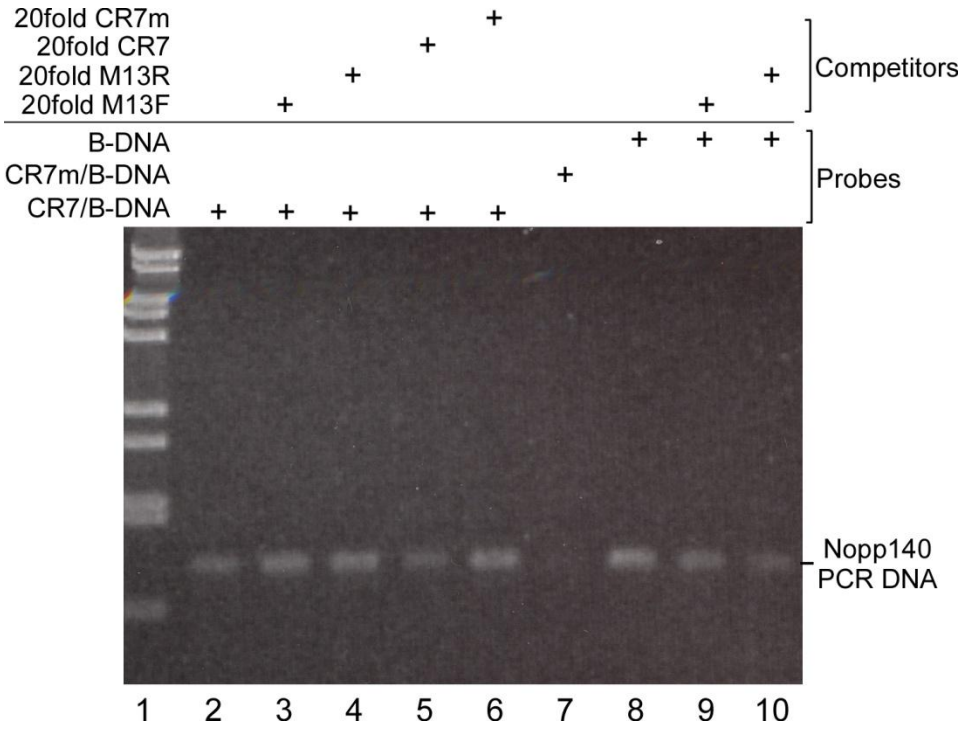


Figure E.14 T7 phage pull down assay

CR7 and CR7m bait RNAs were utilized for phage pull down assay. Nopp140 phage could be detected with CR7 bait RNA (lane 2) but not with CR7m (lane 7), and this association could only be inhibited by CR7 RNA (lane 3-6). Biotinylated DNA primer showed non-specific association with Nopp140 phage as their association was inhibited by M13F and M13R DNA primers.

E.4.7 Nopp140, hTERT and Dyskerin association analysis *in vivo*

As majority in vitro experiment failed to prove the specific interactions between Nopp140 protein and CR7 RNA, an *in vivo* immunoprecipitation (IP) assay was utilized for the investigation of the association between Nopp140 and hTER. Nopp140 protein was previously identified to be associated with dyskerin protein using the co-IP system (Meier and Blobel, 1994; Yang et al., 2000). However, whether their interaction is direct or indirect remains exclusive. Considering that Nopp140 is a putative CAB box binding protein, we hypothesized that the interaction between Nopp140 and dyskerin was scaRNA dependent, including hTER which harbors a CAB box. To verify this, Nopp140 gene was then cloned into pIRES2-2xHA-EGFP vector downstream of the 2xHA tag which could be used for immunoprecipitation. pIRES2-2xHA-Nopp140-EGFP vector was transiently transfected into 293FT cells as described in the materials and methods. Immunoprecipitation was carried out using anti-HA F7 antibody conjugated agarose beads (Santa Cruz Biotechnology), and proteins bound to the beads were detected with western blot. SBP2-hTERT was also expressed in 293FT cells, and immunoprecipitated with Streptavidin coated magnetic beads. Coilin and dyskerin proteins were both co-IPed with Nopp140 protein as expected. However they both associated with Nopp140 upon RNase A treatment, suggesting their interaction with Nopp140 is RNA independent. hTERT was also associated with dyskerin protein, as these two proteins both directly interact with hTER. The dyskerin protein, however, could not be detected upon RNase A treatment after SBP2-hTERT IP, further confirming that their interaction is hTER mediated.

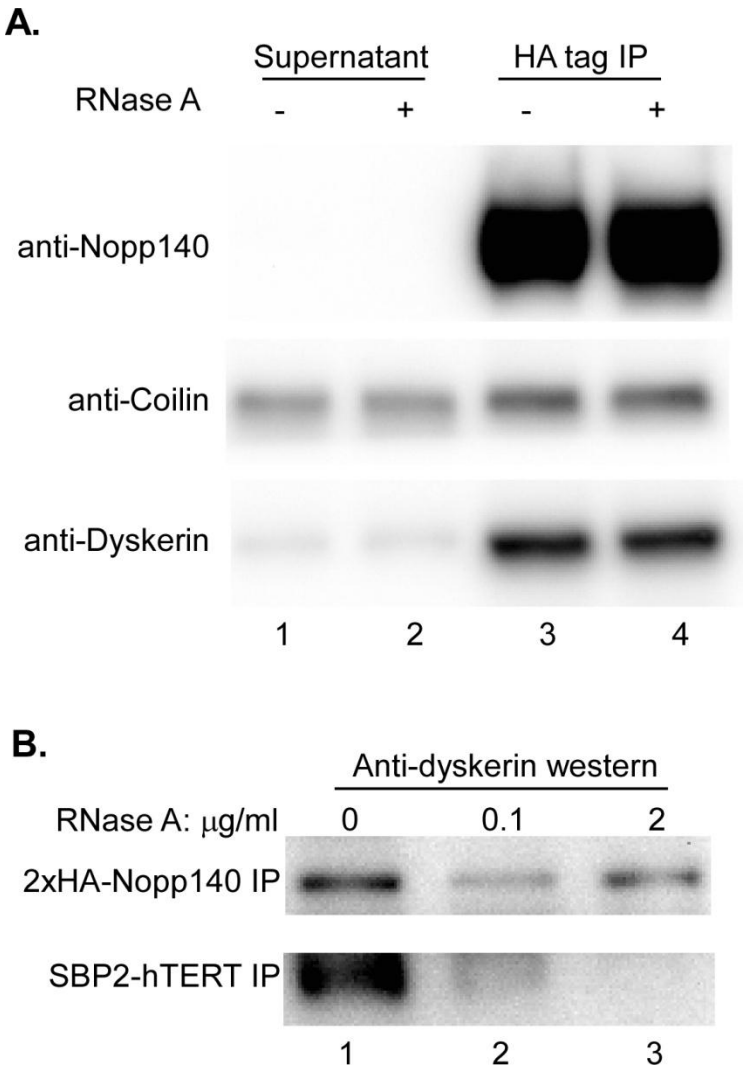


Figure E.15 Nopp140 and dyskerin protein association *in vivo*

(A) Immunoprecipitation was performed for 2xHA-Nopp140 expressed in 293FT cells with anti-HA F7 agarose beads (Santa Cruz Biotechnology) with or without RNase A treatment. Western blot was performed with anti-Nopp140, anti-coilin and anti-dyskerin antibodies respectively. (B) Immunoprecipitation was performed for 2xHA-Nopp140 and SBP2-hTERT proteins expressed in 293FT cells respectively. Western blot analysis was carried out with anti-dyskerin antibody. Both Nopp140 and hTERT showed association with dyskerin protein without the RNase A treatment (lane 1). The association between hTERT and dyskerin disappeared upon RNase A treatment while Nopp140 and dyskerin still associated with the same condition (lane 2 and 3).

E.5 Discussion

C-terminal region of Nopp140 protein with two different cDNA insertions was selected from two individual T7 select human tumor cDNA libraries through phage display to be a putative CAB box binding protein. Previous study also reveals that Nopp140 shuttles between nucleolus and Cajal bodies, and associates with dyskerin protein, part of the human telomerase and snoRNP component. All of these strongly support our hypothesis that Nopp140 is a CAB box binding protein and localizes scaRNAs to the Cajal bodies. However, with numerous experiments performed for the CR7 and Nopp140 protein interaction, we still couldn't make a clear conclusion about this. Instead, all of our results indicate that Nopp140 and CR7 interactions are non-specific and their *in vivo* associations are mediated by dyskerin protein. The reasons that we selected out the Nopp140 displaying T7 phage from two independent phage display experiments, however, still remain elusive.

A novel protein named TCAB1 was discovered and proved to be the CAB box interaction protein recently (Venteicher et al., 2009). Telomerase holoenzyme was purified and TCAB1 was discovered as one of the components. The specific interaction between TCAB1 and CAB box was also presented, suggesting that TCAB1 was indeed the CAB box binding protein.

E.6 References

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

Cristofari, G., Adolf, E., Reichenbach, P., Sikora, K., Terns, R.M., Terns, M.P., and Lingner, J. (2007). Human telomerase RNA accumulation in Cajal bodies facilitates telomerase recruitment to telomeres and telomere elongation. Mol Cell *27*, 882-889.

Danner, S., and Belasco, J.G. (2001). T7 phage display: a novel genetic selection system for cloning RNA-binding proteins from cDNA libraries. Proc Natl Acad Sci U S A *98*, 12954-12959.

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

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

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

Jady, B.E., Bertrand, E., and Kiss, T. (2004). Human telomerase RNA and box H/ACA scaRNAs share a common Cajal body-specific localization signal. J Cell Biol *164*, 647-652.

Meier, U.T., and Blobel, G. (1990). A nuclear localization signal binding protein in the nucleolus. J Cell Biol *111*, 2235-2245.

Meier, U.T., and Blobel, G. (1992). Nopp140 shuttles on tracks between nucleolus and cytoplasm. Cell *70*, 127-138.

Meier, U.T., and Blobel, G. (1994). NAP57, a mammalian nucleolar protein with a putative homolog in yeast and bacteria. J Cell Biol *127*, 1505-1514.

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

Mitchell, J.R., and Collins, K. (2000). Human telomerase activation requires two independent interactions between telomerase RNA and telomerase reverse transcriptase. Mol Cell *6*, 361-371.

Richard, P., Darzacq, X., Bertrand, E., Jady, B.E., Verheggen, C., and Kiss, T. (2003). A common sequence motif determines the Cajal body-specific localization of box H/ACA scaRNAs. EMBO J *22*, 4283-4293.

Smith, C.M., and Steitz, J.A. (1997). Sno storm in the nucleolus: new roles for myriad small RNPs. Cell *89*, 669-672.

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

Yang, Y., Isaac, C., Wang, C., Dragon, F., Pogacic, V., and Meier, U.T. (2000). Conserved composition of mammalian box H/ACA and box C/D small nucleolar ribonucleoprotein particles and their interaction with the common factor Nopp140. Mol Biol Cell *11*, 567-577.

APPENDIX F

CO-AUTHOR APPROVAL

I verify that the following co-authors have approved of my use of our collaborative work in my dissertation.

Julian J.L. Chen (Arizona State University)

Yang Li (Arizona State University)

Eric Selker (University of Oregon)

Shinji Honda (University of Oregon)

Peter F. Stadler (University of Leipzig)

Steve Hoffmann (University of Leipzig)

Manja Marz (Philipps University of Marburg)