

**REPUBLIC OF TURKEY
HACETTEPE UNIVERSITY
INSTITUTE OF HEALTH SCIENCES**

**3'END MATURATION OF
HUMAN TELOMERASE RNA COMPONENT (TERC)
AND ITS EFFECTS ON TELOMERE DYNAMICS**

Barış Boyraz, M.D.

**Tumor Biology and Immunology Program
Doctor of Philosophy Thesis**

ANKARA

2017

**REPUBLIC OF TURKEY
HACETTEPE UNIVERSITY
INSTITUTE OF HEALTH SCIENCES**

**3'END MATURATION OF
HUMAN TELOMERASE RNA COMPONENT (TERC)
AND ITS EFFECTS ON TELOMERE DYNAMICS**

Bariř Boyraz, M.D.

**Tumor Biology and Immunology Program
Ph.D. Thesis**

**Advisor of the Thesis:
Prof. Ayřegöl Üner, M.D., Ph.D.**

**Co-Advisor of the Thesis:
Assistant Prof. Suneet Agarwal, M.D., Ph.D.**

ANKARA

2017

CONFIRMATION PAGE

**3'end Maturation of Human Telomerase RNA Component (TERC)
and Its Effects on Telomere Dynamics.
Barış Boyraz, M.D.**

This study has been approved and accepted as a Ph.D. dissertation in the program of 'Tumor Biology and Immunology' by the examining committee, whose members are listed below, on 10/07/2017.

Chairman of the Committee: Prof. A. Lale Doğan, M.D., Ph.D.
Hacettepe University

Advisor of the Dissertation: Prof. Ayşegül Üner, M.D., Ph.D.
Hacettepe University

Member: Prof. Rengül Çetin Atalay, M.D., Ph.D.
Middle East Technical University

Member: Assoc. Prof. Aytekin Akyol, M.D.
Hacettepe University

Member: Assist. Prof. Özgür Şahin, Ph.D.
Bilkent University

APPROVAL

This dissertation has been approved by the committee above in conformity to the regulations and by laws of Hacettepe University Graduate Programs.

31 Temmuz 2017

(Signature)

Prof. Diclehan Orhan, M.D., Ph.D.

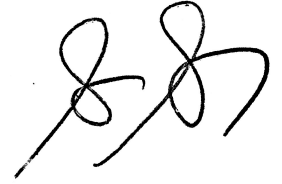
Institute Director

YAYIMLAMA VE FİKRİ MÜLKİYET HAKLARI BEYANI

Enstitü tarafından onaylanan lisansüstü tezimin tamamını veya herhangi bir kısmını, basılı (kağıt) ve elektronik formatta arşivleme ve aşağıda verilen koşullarla kullanıma açma iznini Hacettepe Üniversitesine verdiğimi bildiririm. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak, tezimin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanım hakları bana ait olacaktır.

Tezin kendi orijinal çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanılması zorunlu metinlerin yazılı izin alınarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

Tezimin tamamı dünya çapında erişime açılabilir ve bir kısmı veya tamamının fotokopisi alınabilir.



Dr. Barış Boyraz

10 Temmuz 2017

ETHICAL DECLARATION

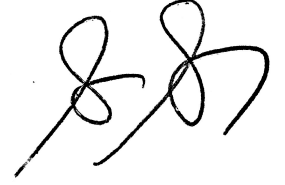
In this thesis study, I declare that all the information and documents have been obtained in the base of the academic rules and all audio-visual and written information and results have been presented according to the rules of scientific ethics. I did not do any distortion in data set. In case of using other works, related studies have been fully cited in accordance with the scientific standards. I also declare that my thesis study is original except cited references. It was produced by myself in consultation with Prof. Ayşegül Üner, M.D., Ph.D. and Assist. Prof. Suneet Agarwal, M.D., Ph.D. and written according to the rules of thesis writing of Hacettepe University Institute of Health Sciences .

A handwritten signature in black ink, consisting of stylized loops and a long horizontal stroke at the bottom.

Barış Boyraz, M.D.

ETİK BEYAN

Bu alıřmadaki bütn bilgi ve belgeleri akademik kurallar erevesinde elde ettiđimi, grsel, iřitsel ve yazılı tm bilgi ve sonuları bilimsel ahlak kurallarına uygun olarak sunduđumu, kullandıđım verilerde herhangi bir tahrifat yapmadıđımı, yararlandıđım kaynaklara bilimsel normlara uygun olarak atıfta bulunduđumu, tezimin kaynak gsterilen durumlar dıřında zgn olduđunu, Prof. Dr. Ayřegl ner ve Yrd. Do. Dr. Suneet Agarwal danıřmanlıđında tarafımdan retilildiđini ve Hacettepe niversitesi Sađlık Bilimleri Enstits Tez Yazım Ynergesine gre yazıldıđını beyan ederim.



Dr. Barıř Boyraz

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my mentors Dr. Ayşegül Üner and Dr. Suneet Agarwal, whose guidance and supervision has helped me conceive, carry on and complete my graduate work. I feel very lucky to have worked with these two great mentors. I believe I have learned a lot from them and they had a great impact on my personality and scientific approach. Both being early birds, I had the chance to go through interesting pathology cases with Dr. Üner before my classes during medical school and we had great discussions with Dr. Agarwal every morning almost an hour before the next person walked in. I realized I wanted to pursue a career as a pathologist as I spent time with Dr. Üner, who as a great teacher always shared her deep knowledge in hematopathology. Dr. Agarwal has taught me the basis of all the molecular biology I now know and his approach to scientific questions always fascinated me. And I believe, this combination fulfills the aim of this M.D.-Ph.D. program to raise physician-scientists. Yet, despite all the great work, both my mentors care more about people than projects, which is what made me feel like home all this time. I am so happy and proud to have worked with these great mentors.

Scientific projects need group effort and I am so lucky to have worked with the amazing people at Agarwal lab. I have worked most closely with Dr. Diane Moon- a smart, calm, perfectionist and caring post-doc- who has helped me develop lab skills with all her patience. Then comes Matthew F. Segal and Aidan Moriarty who have been great lab mates and helped me very much with the projects. I was lucky to work with two very smart summer interns, Maud Z. Muosieyiri and Asli Aykanat, both accomplishing very difficult tasks for my projects during their internship.

I also feel very lucky to share the lab space with Sankaran lab and Croker lab. Biggest thanks go to Aoi Wakabayashi, who has been a great friend both in and outside the lab. I already miss her sincerity, sense of humor and funny dances. During long lab hours Rajiv Khajuria, Leif Ludwig, Arshad Al-Obeidi and Jacob Ulirsch have been some of my neighbors who made the lab a better place.

I regard myself very lucky to be a part of the M.D.-Ph.D. program and Basic Oncology Department at Hacettepe. I would also like to thank my thesis dissertation committee Dr. A. Lale Doğan and Dr. Rengül Çetin Atalay for their continuous support and availability throughout my graduate education. The financial support from the Technological and Scientific Council of Turkey (TÜBİTAK) 2214-A scholarship has been very helpful during this process.

I work much better with friends around and have always been very lucky to be surrounded by intelligent people. Onur Can Akboyraz has been a mirror for me since high school. Fatih Özel helped me discover my sense of humor. İmre Özbay has always stimulated my critical thinking with his hard questions. Oğuzhan Duvar, a late-catch, made the last year of medical school the best with his wit. Deniz Aybaş and Emre Ateş have been my family away from my family in Boston. Our long dinners with jokes around Turkish political scene, afternoon coffees with thorough discussion of any situation, weekend breakfasts made me feel like home. John E. Yazbek has been a true comrade all this time. It is not easy for one to find such a great friend, it has been a privilege to have him as my best friend all along.

Every individual needs someone to guide them initially to find their own path. In my case, my grandmother Nilüfer Paydaş was the one to help me discover my capabilities. She used to teach me every day after school and every weekend all my childhood. It has been an amazing experience growing up with her.

Maslow's pyramid puts self-actualization at the top and I thank my mom, Ruken Paydaş, to help me achieve this potential. She has fully supported me with everything that was important to me. Some day, I can only hope to be as strong, sensitive, caring, hardworking and selfless as she is. I dedicate this thesis and all the accomplishments to her.

ABSTRACT

Boyras, B. 3'end Maturation of Human Telomerase RNA Component (TERC) and Its Effects on Telomere Dynamics, Hacettepe University Institute of Health Sciences, Ph.D. Thesis in Tumor Biology and Immunology, Ankara, 2017.

Telomerase RNA Component (TERC) is the non-coding RNA template and scaffold of the telomerase holoenzyme. TERC level has been shown to be a limiting factor for telomerase activity, lower levels resulting in telomere diseases such as dyskeratosis congenita (DC), idiopathic pulmonary fibrosis; and higher levels being observed in various cancers. In contrast to telomerase RNAs in other species, the mechanisms of human TERC biogenesis and accumulation has not been well understood. Recently, it has been shown that mutations in the gene encoding poly(A) specific ribonuclease (PARN) causes telomere disease. Here, using somatic cells and induced pluripotent stem cells (iPSC) from DC patients with *PARN* mutations, we show that PARN is required for the 3' end maturation of the TERC. Patient cells as well as immortalized cells in which PARN is disrupted show decreased levels of TERC. Deep sequencing of TERC RNA 3' termini reveals that PARN is required for removal of post-transcriptionally acquired oligo(A) tails that target nuclear RNAs for degradation. Previously, PARN has been shown to deadenylate certain non-coding RNAs such as small nucleolar RNAs and microRNAs that are adenylated by the non-canonical poly(A) polymerase. Here we find that PAPD5 is responsible for the post-transcriptional oligoadenylation of TERC and this oligo(A) tail destabilizes nascent TERC transcripts. Disrupting PAPD5 favors TERC 3' end maturation by the PARN and increases steady state TERC levels. In *PARN*-mutant iPSCs from DC patients, in which TERC biogenesis is impaired, PAPD5 knockdown restores TERC RNA levels, telomerase activity and telomere elongation. Our results show that TERC biogenesis and steady state levels depends on the balance of PARN and PAPD5. Here, we demonstrate a role for PARN and PAPD5 in regulating telomere length homeostasis and identify it as a potential target for the therapeutic manipulation of telomerase.

Keywords: telomere diseases, TERC biogenesis, iPSC modeling

ÖZET

Boyras, B. İnsan Telomeraz RNA Komponent (TERC) 3'ucunun Matürasyonu ve Telomer Dinamikleri Üzerindeki Etkisi, Hacettepe Üniversitesi Sağlık Bilimleri Enstitüsü, Tümör Biyolojisi ve İmmünolojisi Doktora Tezi, Ankara, 2017. Telomeraz RNA Komponenti (TERC), telomeraz holoenziminin şablon olarak kullandığı, enzimin tersiyer yapısının oluşması için kritik bir protein kodlamayan RNA'dır. TERC seviyeleri telomeraz aktivitesi için sınırlayıcı nitelikte olup düşük TERC seviyeleri dyskeratosis congenita (DC) ve idiyopatik pulmoner fibrozis gibi telomer hastalıklarında, yüksek TERC seviyeleri ise çeşitli malignitelere gözlemlenmektedir. Diğer organizmalardakinin aksine insan TERC biyogenezi tam olarak anlaşılamamıştır. Yakın zamanda, poli(A) spesifik ribonükleaz (PARN)'ı kodlayan gendeki mutasyonların telomer hastalıkları ile ilişkili olduğu gösterilmiştir. Bu tezde *PARN* mutasyonuna sahip DC hastalarından elde edilen somatik hücre ve indüklenmiş pluripotent kök hücreler (iPKH) kullanılarak PARN'ın TERC 3'ucunun matürasyonu için kritik olduğu gösterilmektedir. Hasta hücrelerinde ve PARN seviyesinin düşürüldüğü hücre hatlarında düşük TERC seviyeleri gözlemlenmiştir. TERC 3'ucunun derin dizilenmesi sonucu PARN eksikliğinde TERC 3'ucunda oligo(A) kuyruğunun varlığı gösterilmiştir. Önceki çalışmalarda PARN'ın protein kodlamayan RNA'lara bir Poli(A) polimeraz olan PAPD5 tarafından eklenen oligo(A) kuyruğunun deadenile ettiği gösterilmiştir. Bu tezde PAPD5'in TERC 3'ucunun post-transkripsiyonal oligoadenilasyonundan sorumlu olduğu ve bu oligo(A) kuyruğunun TERC'i destabilize ettiği gözlemlenmiştir. PAPD5 inhibisyonu TERC 3'ucunun matürasyonunu ve TERC seviyelerini artırmaktadır. *PARN*-mutant iPKH'lerde PAPD5 inhibisyonu TERC seviyelerini, telomeraz aktivitesini artırmış ve telomer uzamasına sebep olmuştur. Bu sonuçlar TERC biyogenezi ve seviyelerinin PARN ve PAPD5 arasındaki dengeye bağlı olduğunu göstermektedir. Bu iki molekül, hücrede telomeraz aktivitesi ve telomere uzunluğu dinamiklerini düzenlemek adına uygun tıropatik aday niteleğindedir.

Anahtar kelimeler: telomere hastalıkları, TERC biyogenezi, iPKH modelleme

CONTENTS

	Page
CONFIRMATION PAGE	iii
YAYIMLAMA VE FİKRİ MÜLKİYET HAKLARI BEYANI	iv
ETHICAL DECLARATION	v
ETİK BEYAN SAYFASI	vi
ACKNOWLEDGEMENTS	vii
ABSTRACT	ix
ÖZET	x
CONTENTS	xi
ABBREVIATIONS	xiv
FIGURES	xv
TABLES	xvii
1. INTRODUCTION	1
2. GENERAL INFORMATION	2
2.1 Introduction to telomeres	2
2.2 Human telomerase RNA component (TERC)	4
2.3 Telomere biology diseases	5
2.4 Using induced pluripotent stem cells to model diseases	7
2.5 Aim of this project	8
3. MATERIALS AND METHODS	9
3.1 Patient material	9
3.2 Primary cells and cell lines	9
3.3 <i>TERC</i> and <i>TERT</i> cDNA expression	10

3.4 TERT and Dyskerin Immunoprecipitation	10
3.5 RNA isolation, cDNA synthesis and quantitative RT-PCR	10
3.6 Western blot analysis	11
3.7 Northern blots	11
3.8 Telomerase activity assay	12
3.9 Telomere length measurements	12
3.10 RNA interference and <i>PAPD5</i> cDNA expression	12
3.11 <i>TERC</i> RNA decay	13
3.12 3' rapid amplification of cDNA ends (RACE)	13
3.13 MiSeq library preparation and analysis	14
3.14 RNA-Seq library preparation and analysis	15
4. RESULTS	18
4.1 A novel <i>TERC</i> CR4/CR5 domain mutation causes telomere disease via decreased TERT binding	18
4.2 Poly(A)-specific ribonuclease (PARN) mediates 3' end maturation of the telomerase RNA component	31
4.3 Post-transcriptional manipulation of <i>TERC</i> reverses molecular hallmarks of telomere disease	42
5. DISCUSSION	68
6. SUMMARY AND FUTURE DIRECTIONS	73
7. REFERENCES	74
8. APPENDIX	80
Appendix-1: Publications	
Appendix-2: Meetings attended	

Appendix-3: Collaborators' contributions

Appendix-4: Funding Sources

9. CURRICULUM VITAE

84

ABBREVIATIONS

3'RACE	3'Rapid amplification of cDNA ends
DC	Dyskeratosis congenita
DMEM	Dulbecco's modified eagle medium
HEK293	Human embryonic kidney 293
IPF	Idiopathic pulmonary fibrosis
iPSC	Induced pluripotent stem cells
ncRNA	Non coding RNA
PAPD5	Poly(A) RNA polymerase D5
PARN	Poly(A) specific ribonuclease
PCR	Polymerase chain reaction
PVDF	Polyvinylidene fluoride
qPCR	Quantitative PCR
RFLP	Restriction fragment length polymorphism
RT-PCR	Reverse transcription PCR
SDS-PAGE gel	sodium dodecyl sulfate polyacrylamide gel
shRNA	Short hairpin RNA
snoRNA	Small nucleolar RNA
TERC	Telomerase RNA component
TERT	Telomerase reverse transcriptase
TRAP	Telomere repeat amplification protocol
TRF	Terminal restriction fragment

FIGURES

Figure		Page
2.1.1	Telomeres and components of telomere maintenance	3
2.2.1	Human telomerase RNA component and known protein components of the telomerase holoenzyme	4
2.3.1	Telomere biology diseases: a spectrum	6
4.1.1	A family with telomere disease	19
4.1.2	Variant of unknown significance in <i>TERC</i> gene	20
4.1.3	Decreased telomerase activity with mutant <i>TERC</i>	22
4.1.4	Reprogramming of <i>TERC</i> G319A fibroblasts	24
4.1.5	Telomere length and <i>TERC</i> RNA levels in <i>TERC</i> G319A iPSCs	25
4.1.6	RT-PCR after Dyskerin and TERT Immunoprecipitation	27
4.1.7	<i>TERC</i> RT-PCR after Dyskerin and TERT Immunoprecipitation	28
4.1.8	<i>TERC</i> RT-PCR after Dyskerin and TERT Immunoprecipitation, other clones	29
4.1.9	RFLP of <i>TERC</i> amplicons	30
4.2.1	<i>PARN</i> mutations in two families with dyskeratosis congenita	32
4.2.2	Decreased <i>TERC</i> levels in <i>PARN</i> deficient iPS cells	34
4.2.3	Decreased <i>TERC</i> levels in <i>PARN</i> deficient 293 cells	35
4.2.4	<i>PARN</i> deficiency results in abnormal <i>TERC</i> 3' ends	37
4.2.5	Decreased mature <i>TERC</i> and increased 3' extended <i>TERC</i> species in <i>PARN</i> deficient patient cells	38
4.2.6	Decreased stability of <i>TERC</i> in <i>PARN</i> deficient cells	40
4.3.1	Telomere length attrition in <i>PARN</i> deficient cells	43
4.3.2	Telomere length elongation in <i>PARN</i> kd 293 cells after <i>PARN</i> overexpression	44
4.3.3	Telomere length elongation in immortalized <i>PARN</i> -mutant fibroblasts after <i>PARN</i> overexpression	45
4.3.4	<i>TERC</i> rescues telomere length in <i>PARN</i> deficient cells	47
4.3.5	<i>PAPD5</i> knockdown in HEK293 cells	49
4.3.6	<i>PAPD5</i> inhibition changes 3' of <i>TERC</i>	50

4.3.7	<i>PAPD5</i> inhibition decreases TERC oligoadenylation	53
4.3.8	Manipulation of <i>PAPD5</i> changes TERC levels	54
4.3.9	Overexpression of <i>PAPD5</i> in control and <i>PAPD5</i> knockdown cells	55
4.3.10	<i>PAPD5</i> inhibition restores TERC in <i>PARN</i> deficient cells	56
4.3.11	<i>PAPD5</i> inhibition restores telomerase activity in <i>PARN</i> deficient cells	57
4.3.12	<i>PAPD5</i> inhibition rescues TERC maturation and steady state levels in <i>PARN</i> -mutant patient cells	59
4.3.13	<i>PAPD5</i> inhibition decreases TERC decay rate in <i>PARN</i> -mutant patient cells	60
4.3.14	<i>PAPD5</i> inhibition rescues telomerase activity and telomere length in <i>PARN</i> -mutant patient cells	62
4.3.15	<i>PAPD5</i> inhibition rescues telomere length in <i>PARN</i> -mutant patient cells	63
4.3.16	<i>PAPD5</i> inhibition rescues telomere length in <i>PARN</i> -mutant patient cells, other clones	64
5.1	Model for TERC 3' end maturation by <i>PARN</i>	69
695.2	Model - reciprocal regulation of TERC maturation by <i>PARN</i> and <i>PAPD5</i>	71

TABLES

Table	Page
3.1 Oligo/Primer name and DNA sequences used in this thesis	16
4.2.1 RNA-seq transcriptome analysis in PARN deficient cells	41
4.3.1 RNA-Seq transcriptome analysis of the effects of PAPD5 knockdown: transcripts altered in excess of the fold increase in TERC RNA	65
4.3.2 Genes whose transcript levels are altered in excess of the fold-change in TERC levels, in pair-wise comparisons of PAPD5 knockdown versus control cells	66
4.3.3 Genes whose transcript levels are altered in excess of the fold-change in TERC levels, in any 2 of 3 pair-wise comparisons of PAPD5 knockdown versus control: HEK 293 cells and PARN-mutant Patient 1 clones 1 and 2	67

1. INTRODUCTION

The telomerase RNA component (TERC) is the essential non-coding RNA template and scaffold of the telomerase holoenzyme (1-3). TERC levels are limiting for telomerase activity and telomere maintenance, and thus are a critical determinant of self-renewal (4-5). The *TERC* locus is recurrently amplified in cancer (6-7), whereas reductions in *TERC* levels due to genetic mutations in *DKC1*, *NOP10*, *NHP2* or *TERC* itself result in telomere diseases such as dyskeratosis congenita and idiopathic pulmonary fibrosis (8-10).

PARN is a widely-expressed cap-dependent, poly(A) deadenylase with a canonical role in regulating global mRNA levels during development (11-14), and additional, more specialized functions including end-trimming of the Dicer independent microRNA (miR)-4517 (15) and deadenylation of small nucleolar (sno)RNAs (16). PARN loss-of-function mutations have recently been implicated in IPF (17) and dyskeratosis congenital (18), but why a deficiency of PARN, which has no known role in telomere biology, should mimic telomere diseases is unexplained (19). Here, using induced pluripotent stem cells derived from *PARN*-mutant patient somatic cells we define how *PARN* mutations are linked to telomere biology (20-21).

In another family with telomere disease, a novel variant of unknown significance in *TERC* gene (c.319G>A) has been detected. By using the same approach we define the pathogenicity of the novel variant (22).

This thesis includes three different projects all three being functional genetics studies in telomere biology diseases using induced pluripotent stem cells as models. TERC biology is central to all three projects, one of them functionally describing a novel variant (22) and the other two defining a new biogenesis pathway for TERC (20-21).

2. GENERAL INFORMATION

2.1 Introduction to telomeres

Unlike the circular chromosome of prokaryotes, eukaryotic chromosomes are arranged in a linear fashion in the cell. The linearity of the chromosomes brings several problems regarding genome integrity. Double strand breaks occur spontaneously in the cell and they induce DNA damage response leading to activation of cell cycle checkpoints to arrest cell cycle (23). The linear ends of chromosome termini should be differentiated from double strand breaks so as not to induce this response. Second problem is regarding DNA replication. DNA polymerase has a 5'-3' directionality and if there was not any protective mechanism at the ends of chromosome termini, one strand would get shorter in each DNA replication cycle and this would eventually lead to loss of genetic information. Thus, telomeres have evolved over the course of time to protect eukaryotic linear chromosome ends.

Telomeres consist of repetitive nucleotide sequences and various DNA binding proteins attached to these sequences (24-26). The repetitive sequences and length of telomeric DNA vary from organism to organism. In the case of humans, the repetitive sequence is 5'-TTAGGG-3' and this sequence extends several kilobases, averaging 10 kb in a newborn human's cord blood (27). Telomeric DNA sequences are bound by a group of six proteins (TRF1, TRF2, RAP1, TIN2, TPP1 and POT1) called the shelterin complex (28-34) (**Figure 2.1.1**). Proteins in this complex specifically recognize and bind to telomeric DNA sequence and are important for the physical protection of these termini. They also inhibit recombination and fusion events, and prevent DNA damage response (35-37).

The other complex involved in telomere maintenance is the telomerase holoenzyme, which is responsible for the elongation of chromosome termini (38-40). It is a ribonucleoprotein that consists of two conserved components, telomerase reverse transcriptase (TERT) and telomerase RNA component (TERC), and

additional structural proteins (DKC1, NHP2, GAR1, NOP10) that are important for the stability and trafficking of the complex (41-43) (**Figure 2.1.1**). TERC has a sequence complementary to telomeric repeats and is used as a template by TERT to further elongate telomeres and this solves the ‘end replication problem’(44).

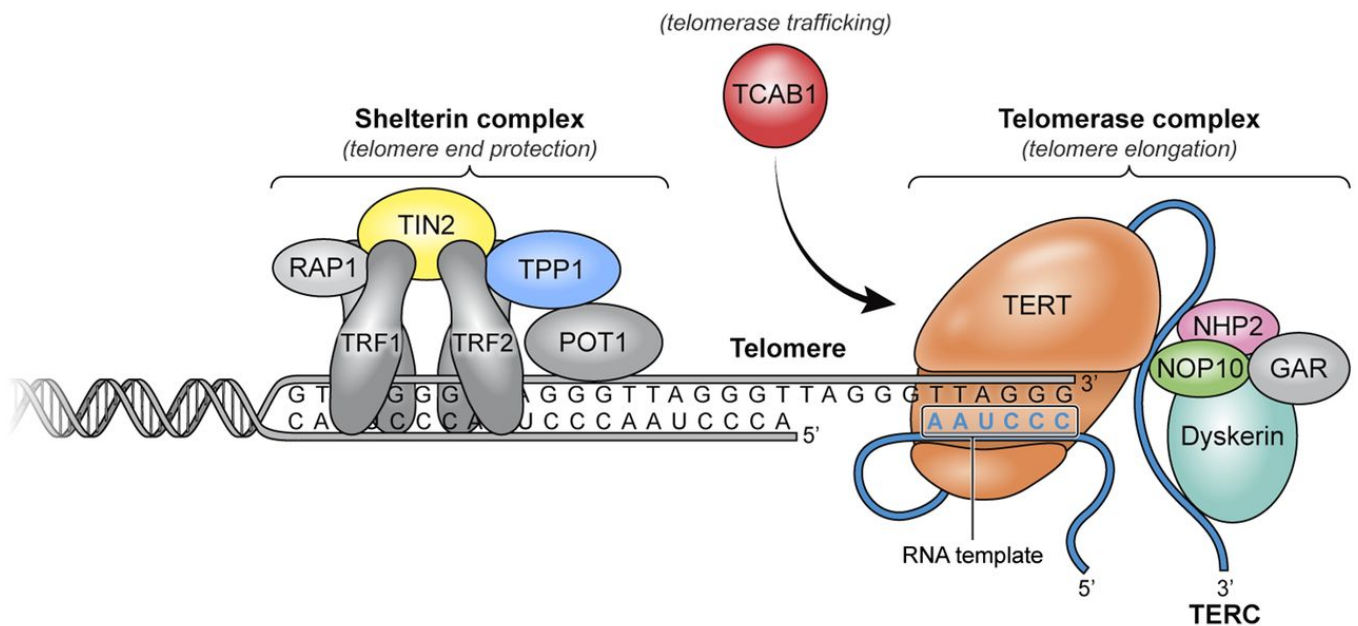


Figure 2.1.1: Telomeres and components of telomere maintenance

(From Townsley et al., *Blood*, 2014)

2.2 Human telomerase RNA component (TERC)

TERC is the RNA template used by TERT to elongate telomere ends (1-3). It is a 451 nucleotide-long RNA that has a complex secondary structure (45) (**Figure 2.2.1**). The main functional domains of TERC are conserved among vertebrates (45). The 3' end of TERC has an H/ACA motif (46) which consists of two hairpins and conserved nucleotide sequences in between. This motif is commonly seen at the 3' end of small nucleolar RNAs (16).

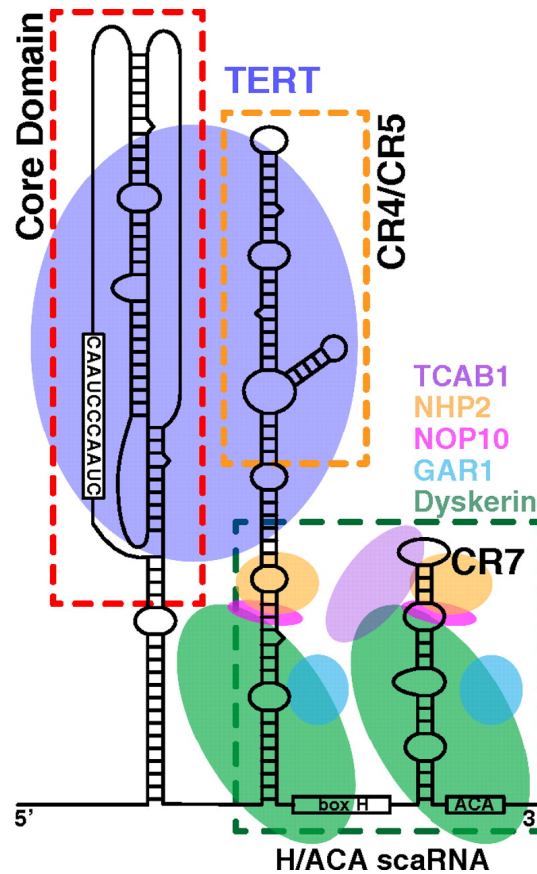


Figure 2.2.1: Human telomerase RNA component and known protein components of the telomerase holoenzyme

(From Zhang et al., *PNAS*, 2011)

It has been shown that TERC levels are limiting for the telomerase activity (4-5). Pathogenic mutations in *TERC* or genes encoding for the proteins attaching to the 3' end of TERC decrease TERC levels (8-10). This results in decreased telomerase activity and impaired telomere maintenance. Multiple loss-of-function mutations have been defined in *TERC*. It has been shown to be recurrently amplified in various cancers leading to increased telomerase activity (6-7).

Although telomerase RNAs are functionally similar in different organisms, their structures and thus their biogenesis pathways vary greatly. Splicesomal cleavage has been shown to be main mechanism by which yeast telomerase RNAs are generated (47-48). Human telomerase RNA component transcriptional regulation and biogenesis have been incompletely understood.

2.3 Telomere biology diseases

As telomeres are central to genome protection in humans, one can easily assume that disruption of these structures would lead to various pathologies, which happens to be the case. The prototypical Mendelian disease is called dyskeratosis congenita (DC). DC is a severe childhood syndrome that is mainly characterized by dermatological findings: hypopigmented skin, nail dystrophy and oral leukoplakia (49). The main mortality reason in DC is bone marrow failure in which the cellularity of the bone marrow is replaced with fat tissue and this loss of hematopoiesis leads to pancytopenia (49).

Although DC is the prototypical telomere disease, it is now appreciated that telomere diseases form a spectrum (**Figure 2.3.1**). Mutations in genes encoding for telomere related proteins and disruption of telomeres have been shown is aplastic anemia, idiopathic pulmonary fibrosis, liver cirrhosis, myelodysplastic syndrome and leukemia (8-9,50-57). The manifestations are thought to arise from loss of stem cell population and secondary damage in highly proliferative tissues (58-59). The specific mutation and telomere length of the individual determines the severity and age of onset of these conditions.

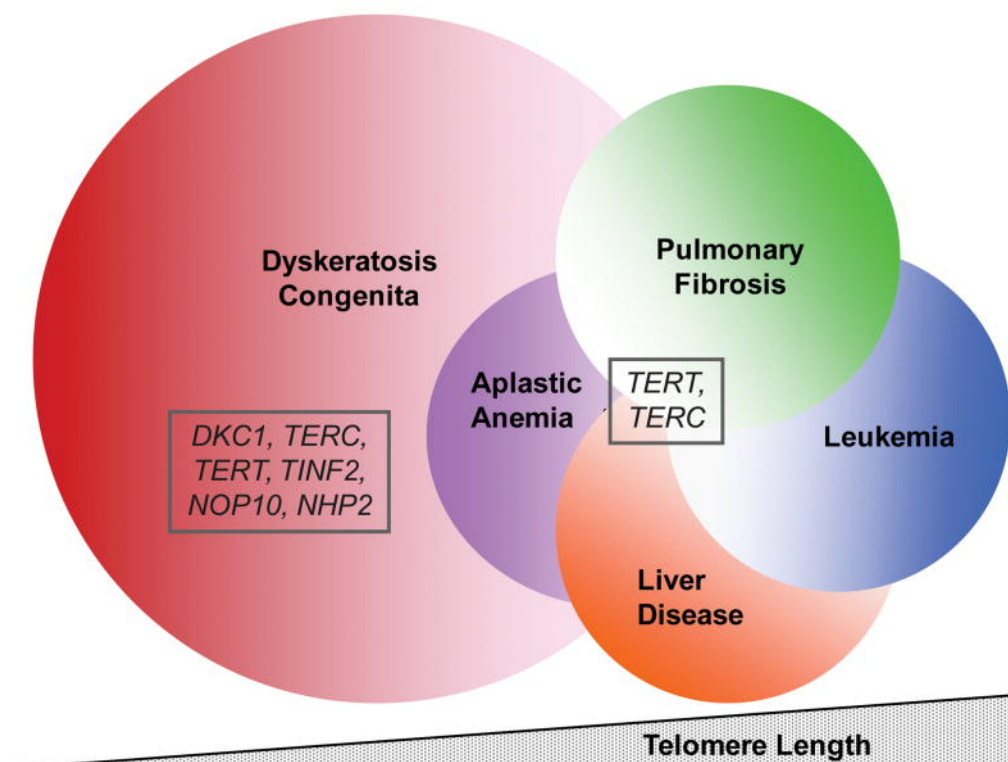


Figure 2.3.1: Telomere biology diseases: a spectrum

(From Savage et al, *Genet in Med*, 2010)

When families with telomere diseases were examined, genetic anticipation was shown in those pedigrees (60-61). The offspring inherits both the mutation and short telomeres from the parents. Because the mutation disrupts telomere maintenance, telomere length gets shorter in each subsequent generation. This leads to an evolving disease phenotype; where one can see adult-onset, mild forms of telomere diseases in earlier generations as opposed to more severe and early-onset forms like DC being seen in later generations.

DC has been a known entity for almost a century but its link to telomere biology has been recent. In 1980s linkage analysis studies defined a region in X-linked DC patients: Xq28 (62). When the patients in a DC registry in London were screened for that region, it was further refined to a 1.4 cM area in Xq28. A deletion in *DKC1* gene in that region has been found in one of the patients and the

pathogenicity of the gene was further supported by missense variants in 5 other DC patients in the registry (41). *DKC1* was known to be encoding for dyskerin, which at the time was known to bind H/ACA box small nucleolar RNAs and it was thought be involved in ribosomal RNA modification and ribosome biogenesis (41). Thus, it was thought that DC could be a ribosomal disorder. *TERC* also has an H/ACA box domain at the 3'end. It was later shown that *DKC1* mutations resulted in decreased *TERC* stability and it was concluded that *DKC1* mutations mainly disrupt dyskerin-*TERC* binding (63). After the detection of mutations in *TERC* and other molecules involved in telomere biology, it was concluded that DC is a telomere disease (10).

Since then, many mutations have been shown to cause telomere disease. There are now eleven genes known to be involved in telomere diseases: *DKC1*, *TERT*, *TERC*, *NHP2*, *NOP10*, *TPP1*, *TCAB1*, *TIN2*, *RTEL1*, *CTCI* and *PARN*. These mutations can cause telomere disease by various defined mechanisms: reducing telomerase activity, impairing telomerase trafficking and recruitment to telomeres, telomere instability and multifactorial.

2.4 Using induced pluripotent stem cells to model diseases

Traditionally, cell lines and model organisms have been used to understand human diseases. Although useful in many ways, they may not accurately model the human diseases due to lack of disease specificity and species-specific differences, respectively. Reprogramming of adult cells into induced pluripotent stem cells (iPSCs) holds great promise for biomedical research. Reprogramming of somatic cells into pluripotent stem cells was first reported in 2006 (64). A set of embryonic transcription factors; Klf4, Sox2, Oct4 and c-Myc was shown to be effective to revert the differentiation process.

Because iPS cells are derived directly from one's own somatic cells, they carry the exact genetic information and with their endless replicative capacity, they can be expanded to limitless quantities and thus, they are a great source for molecular and biochemical studies. In the era of clinical sequencing, with many

variants and new genes being defined, iPS cells allow one to overcome limitations of model organisms and cell lines to perform functional genetics studies (22).

2.5 Aim of this project

Our goal in this thesis is to investigate the pathogenicity of variants of unknown significance in known genes (22) and function of newly discovered genes that cause telomere diseases (20-21) using induced pluripotent stem cells derived from patient somatic cells as models.

3. MATERIALS AND METHODS

All of the reagents and devices used in this thesis are present at Suneet Agarwal Laboratory, Boston Children's Hospital, Harvard Medical School, Boston, MA, USA where experiments were performed.

3.1 Patient material

Biological samples were procured under protocols approved by Institutional Review Boards, and after written informed consent in accordance with the Declaration of Helsinki.

3.2 Primary cells and cell lines

Fibroblasts were cultured from skin biopsies obtained from patients and healthy volunteer subjects under approved protocols. Briefly 2 mm punch biopsies were diced and placed under a coverslip in DMEM media with 15% fetal calf serum (FCS) until keratinocyte and fibroblast outgrowths were apparent. Fibroblasts were sub-cultured and expanded using trypsin 0.05% and DMEM 15% FCS.

WT1 and WT2 iPSCs are derived from normal human fibroblast NHSF2 and hFib2, and WT3 iPSCs from BJ1 normal fibroblasts (20-22). Patient fibroblast cultures and iPSCs were generated from *PARN*-mutant skin fibroblasts (20) and *TERC*-mutant bone marrow fibroblasts (22). Fibroblasts were reprogrammed using the pRRL.PPT.SF.hOKSMco.idTomato.preFRT lentiviral reprogramming vector (a gift from A. Schambach, Hannover Medical School). For feeder-free culture, iPSCs were maintained in Essential 8 medium (Life Technologies) on hES-qualified Matrigel matrix (BD Biosciences). iPSC G-band karyotyping was performed by Cell Line Genetics (Madison, WI).

PARN-mutant patient fibroblasts were immortalized by transduction with pBABE-hTERT-puro retroviral vectors.

WI38- VA13 cells (ATCC) were sub-cultured and expanded using trypsin 0.05% and DMEM 10% FCS.

HEK 293 were sub-cultured and expanded using trypsin 0.05% and DMEM 10% FCS.

3.3 *TERC* and *TERT* cDNA expression

Retroviral and lentiviral vectors pBABE-hTERT-puro (Addgene #1171; gift of R. Weinberg) and pHIV7/SF-U3-TER-500 and controls were used. pHIV7/SF-U3-TER-500 G319A was created by site-directed mutagenesis with primers TERC_SDM_F/R (Table 3.1). VA13 cells were transfected with 1 µg of each *TERC* and *TERT*-encoding plasmids using Lipofectamine LTX with Plus Reagent.

3.4 TERT and Dyskerin Immunoprecipitation

TERC het.G319A iPSCs (50×10^6 cells) were lysed in 1000 µl lysis buffer (20 mM Tris-HCl, pH: 7.4, 150 mM NaCl, 0.5% NP-40, 0.1 mM dithiothreitol (DTT) and protease inhibitor cocktail). The lysate was homogenized by passage through a 23 g needle and kept on ice for 30 mins. The lysate was then pelleted by centrifugation (16,000 g) for 20 mins, 4°C. The supernatant was precleared with 50 µl Protein A Sepharose for 2 hours at 4°C. The precleared lysate was divided in to three aliquots and these were incubated with either rabbit-polyclonal anti-TERT antibody (Rockland, 2 µl), rabbit-polyclonal anti-dyskerin antibody (Santa Cruz, 10 µl) or rabbit-polyclonal IgG (Abcam, 2 µl), for 3 hours at 4°C, followed by incubation with 30 µl Protein A Sepharose overnight. The beads were washed 3 times with lysis buffer and RNA was extracted from the bound complexes using TRIzol (Ambion).

3.5 RNA isolation, cDNA synthesis and quantitative RT-PCR

RNA from cultured cells was recovered using TRIzol (Ambion). After DNase treatment (Turbo DNA-free, Ambion), cDNA was synthesized using 1 µg of total RNA, or the entire amount recovered from immunoprecipitated complexes, 50 ng random hexamers, and 1 µl SuperScript III Reverse Transcriptase (Invitrogen) in a

total volume of 20 μ l for 1 hr at 50°C. Primers used for RT-PCR are provided in Table 3.1.

3.6 Western blot analysis

Total cellular lysates were subjected to SDS-PAGE and transfer to PVDF membranes using standard procedures. Detection of TERT was performed using rabbit anti-TERT antibodies (Rockland, dilution 1:500) and horseradish peroxidase (HRP)- conjugated goat anti-rabbit secondary antibody (Bio-Rad, 170-5046; 1:15,000 dilution) followed by chemiluminescent detection using Clarity Western ECL substrate (Biorad). Detection of PARN and dyskerin was performed using primary antibodies to human PARN (Abcam ab-188333, dilution 1:5000) and dyskerin (Santa Cruz sc-48794, dilution 1:1000) and horseradish peroxidase (HRP)-conjugated goat anti-rabbit secondary antibody (Biorad 170- 5046, dilution 1:15,000), followed by chemiluminescent detection using Clarity Western ECL substrate (Biorad). Detection of PAPD5 was performed using rabbit anti-PAPD5 (Atlas, dilution 1:1000) and horseradish peroxidase (HRP)- conjugated goat anti-rabbit secondary antibody (Bio-Rad, 170-5046; 1:15,000 dilution) followed by chemiluminescent detection using Clarity Western ECL substrate (Biorad).

Protein loading was determined using HRP-conjugated actin antibodies (Santa Cruz sc-1615, dilution 1:1000) on the same membranes and used for normalization. Imaging of chemiluminescent signals was performed using the Biorad ChemiDoc Touch imaging system.

3.7 Northern blots

Formaldehyde/agarose gel electrophoresis. 5–10 μ g of total RNA was separated on 1.5% agarose/formaldehyde gels, transferred to Hybond N+ membranes (Amersham) in 10X SSC by capillary transfer, and hybridized with α -³²P-dCTP-labeled full-length *TERC* probe in ULTRAhyb buffer (Life Technologies). Signals were normalized to *18S* or *28S* rRNA based on ethidium bromide staining.

Quantification was performed using ImageJ. Primers used to amplify probe are *TERC*: TERC_F2/TERC_R1 (Sequences can be found in Table 3.1 below).

3.8 Telomerase activity assay

VA13 cells and iPSCs were lysed in CHAPS lysis buffer and protein was quantified by the Bradford assay (Bio-Rad). Three fold dilutions (starting from 75 ng) of cell extracts were subjected to the telomere repeat amplification protocol (TRAP) using the TRAPeze Telomerase Detection kit (Millipore). Products were resolved on 15% TBE polyacrylamide gels and visualized by staining with GelRed (Biotium).

3.9 Telomere length measurements

Terminal restriction fragment (TRF) length analysis. 2-3 µg of genomic DNA was digested with HinfI/RsaI restriction enzymes and separated on 0.8% agarose gels, followed by Southern blot using the TeloTAGGG telomere length assay kit (Roche).

3.10 RNA interference and *PAPD5* cDNA expression

short-hairpin RNA (shRNA) constructs. Duplex oligonucleotides encoding shRNA targeting human *PARN* and *PAPD5* or luciferase control were cloned into the pLKO.1-puro vector and pLKO.1-blast vectors (Addgene #10878 and #26655, gifts of D. Root). shRNA sequences can be found in Table 3.1 below.

cDNA expression constructs. Codon-optimized cDNA encoding the *PAPD5* open reading frame (ENA accession: FR872509) with an N-terminal FLAG tag was synthesized (Integrated DNA Technologies), and cloned into pLX304 (blasticidin^R; Addgene #25890, gift of D. Root). Lentiviral vectors encoding *EGFP* and *PARN* were cloned into pLX301 (puromycin^R; Addgene #25895; gift of Dr. D. Root) and pLX304 (17). Retroviral and lentiviral vectors pBABE-hTERT-puro (Addgene

#1171; gift of R. Weinberg), p-MIG-*DKC1* and pHIV7/SF-U3-TER-500 were also used.

Viral vector production and transduction. Retroviral particles were produced by co-transfection of HEK 293T cells with retroviral vectors, VSV-G and Gag-Pol packaging plasmids. Lentiviral particles were produced by co-transfection of HEK 293T cells with lentiviral vectors, pCMV_dR8.91 and pCMV_VSV-G. For knockdown and overexpression experiments, HEK 293 cells, fibroblasts and iPSCs were transduced with viral vectors in the presence of protamine sulfate (10 µg/ml) for 8-12 hours. For *TERC* and *DKC1* experiments, transduced HEK 293 cells and fibroblasts were sorted for mCherry or EGFP expression and cultured without further selection. For antibiotic selection in shRNA and other overexpression experiments, HEK 293 cells and fibroblasts were cultured in puromycin (1-2 µg/ml) and/or blasticidin (5-10 µg/ml), and iPSCs were cultured in puromycin (0.2 µg/ml) and/or blasticidin (3-5 µg/ml).

3.11 *TERC* RNA decay

10⁶ iPSCs were treated with 5 µg/ml of actinomycin D (Life Technologies) and harvested in TRIzol. Purified RNA was subjected to Northern blot. *TERC* signals were normalized to 18S rRNA. Decay slopes were determined by simple linear regression and transcript half-life was calculated as the x-intercept at y=0.5, using GraphPad Prism.

3.12 3' rapid amplification of cDNA ends (RACE)

600 ng of total RNA was ligated to 5 µM of 5'-adenylated, 3'-blocked adapter (Universal miRNA cloning linker, New England Biolabs) with 280 units of T4 RNA ligase truncated KQ (New England Biolabs), 25% PEG 8000, and 1 µl RNaseOUT (Life Technologies) in a 20 µl reaction at 25 degrees C for 16–24 hr. After cleanup with RNA Clean and Concentrator columns (Zymo Research), followed by DNase treatment, cDNA was synthesized with 5 pmol of universal RT primer and

SuperScript III. PCR amplification was carried out using 5 μ M of TERC_F1/universal RT primer sets (Table 3.1) with SsoAdvanced Universal SYBR Green Supermix (Bio-Rad). PCR products were directly analyzed run on 2.5% agarose gels to visualize mature and extended TERC transcripts, or subject to Qiaquick PCR purification columns (Qiagen) for library preparation for deep sequencing. For Sanger sequencing, 3' RACE PCR products were directly cloned into the pCR4_TOPO vector (Life Technologies), and individual clones were sequenced using the TERC_F1 primer.

3.13 MiSeq library preparation and analysis

RACE products were prepared for deep sequencing using the TruSeq Nano DNA LT Library Prep Kit (Illumina). Briefly, for each sample, linkers carrying unique barcodes were ligated to RACE products with DNA ligase, amplified using Illumina adapters, and size selected using magnetic beads (Ampure). The completed libraries were submitted to the Tufts University Genomics Core for sequencing and data analysis. The quality and quantity of each library was determined on an Advanced Analytical Fragment Analyzer. The libraries were then pooled to equimolar concentrations. The pooled library was sequenced on an Illumina MiSeq with paired-end 250 bases using the Illumina TruSeq V2 500 cycles kit. The reads were de-multiplexed with CASAVA 1.8.2 and the read 1 and read 2 fastq files for each sample were generated. For data analysis, the Illumina adapter sequence was first removed from the end of each read using Trimmomatic⁵⁰. The resulting unbroken pair was then joined using FLASH⁵¹. The joined-reads from each sample were then mapped to the TERC gene (NR_001566) with the Bowtie2 mapper⁵², and outputted to a SAM file. The SAM file from each sample was then used as input for custom developed Perl scripts to remove the 3' RACE universal adapter sequence, determine the position at which the genomic sequenced ended, and the position and length of the oligo(A) tail. Code availability: Perl scripts for TERC RNA end analysis will be made available upon request.

3.14 RNA-Seq library preparation and analysis

RNA integrity was verified using the Advanced Analytical Fragment Analyzer (FA). 0.5 – 1 µg of total RNA was used as input for library preparation using TruSeq Stranded Total RNA with RiboZero Gold kit (Illumina). The molar concentrations of the libraries were determined using the FA, and pooled at equimolar concentrations. The pooled libraries were sequenced on two lanes of HiSeq 2500 High Output single read 50 bases format. Sequence reads were aligned to the human transcriptome using HISAT version 0.1.653. The transcriptome file consisted of protein-coding and ncRNA sequences downloaded from ENSEMBL (release 80). To estimate transcript abundances, we applied Salmon version 0.3.254 to the aligned reads and summarized transcript abundances into gene-level expression levels by summing all transcript expression levels mapping to the same gene. Gene-to-transcript mappings, and transcript type annotations (e.g. assignment of transcript to categories such as snRNA, snoRNA, etc.) were downloaded also from ENSEMBL. Auto-annotation was manually adjusted by annotating TERC and other scaRNAs as snoRNAs, for the purpose of this analysis. We took two approaches to find genes that were commonly differentially expressed between normal and PARN-deficient cell lines. First, we performed the Wilcoxon signed-rank test on the 13,508 genes that had expression levels of at least 1 transcripts per million (TPM) in one sample, using the following pairs of samples: WT1 FIB versus Patient 1 FIB; WT1 IPS versus Patient 1 cl2 iPS; WT2 iPS versus Patient 2 cl1 iPS; 293 control KD versus 293 PARN KD; *PAPD5* versus luciferase control knockdown cell lines (HEK 293, WT iPSC, Patient 1 clone 1 and clone 2 iPSCs). No genes were nominally significantly differentially expressed across all pairs using this approach, or using the paired t-test. Second, because we have shown that TERC levels are decreased in PARN-deficient cell lines by quantitative RT-PCR and Northern blot, we reasoned that its fold change magnitude would be an appropriate threshold to define other genes as differentially expressed in a paired comparison manner. In this analysis, we asked whether the number of genes in each category of transcript was different than would be expected by chance given the total number of genes differentially expressed using the Chi-squared test.

Table 3.1 Oligo/Primer name and DNA sequences used in this thesis

Oligo/Primer name	DNA sequence
PAPD5_F (RT-PCR)	AGGGAGTCGTGGGTCTGCATGAA
PAPD5_R (RT-PCR)	ATATCTGGACGTCAGCGCTGGG
OCT4_F (RT-PCR)	AGCGAACCAGTATCGAGAAC
OCT4_R (RT-PCR)	TTACAGAACCACACTCGGAC
SOX2_F (RT-PCR)	AGCTACAGCATGATGCAGGA
SOX2_R (RT-PCR)	GGTCATGGAGTTGTACTGCA
NANOG_F (RT-PCR)	TGAACCTCAGCTACAAACAG
NANOG_R (RT-PCR)	TGGTGGTAGGAAGAGTAAAG
TERC_F (RT-PCR)	CGCCTTCCACCGTTCATTCTAGAG
TERC_R (RT-PCR)	GCATGTGTGAGCCGAGTCCTGG
TERT_F (RT-PCR)	TGTGCACCAACATCTACAAG
TERT_R (RT-PCR)	GCGTTCTTGGCTTTCAGGAT
ACTB_F (RT-PCR)	TGAAGTGTGACGTGGACATC
ACTB_R (RT-PCR)	GGAGGAGCAATGATCTTGAT
POL2A_F (RT-PCR)	GCTTGATGCGGGTGCTGAGTGA
POL2A_R (RT-PCR)	GTCCTGGCGGTTGACTCCGTGT
SNORA63_F	GCTGCTAAGTGCTGTGTTGTCGTTC
SNORA63_R	GACCAAGCGTCCCTGGCTGCTA
SNORD104_F	GGCCTGCTGTGATGACATTC
SNORD104_R	GGCAGGCTCAGACTCCAGTT
TERC_SDM_F	GACCCGCGGTTGACAGAGCCCAACTCT

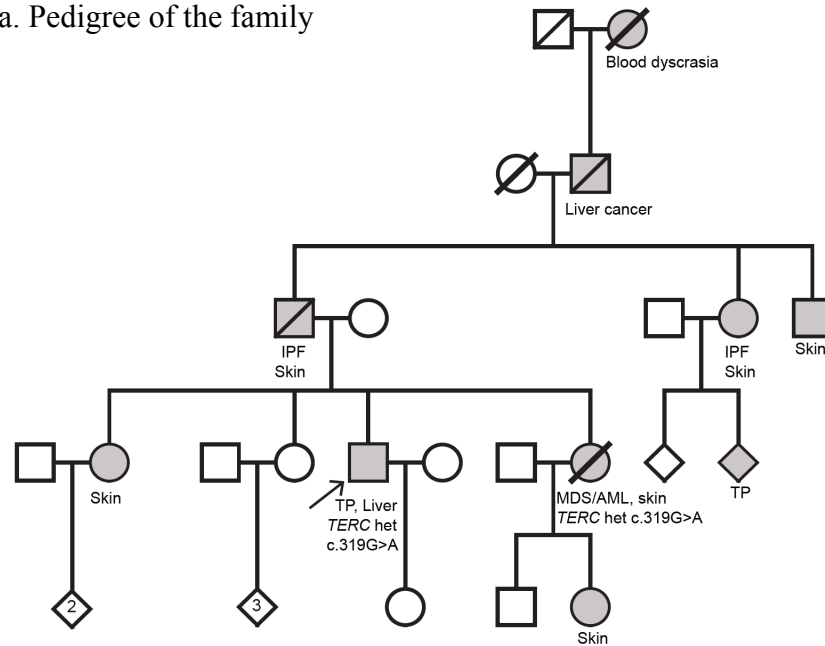
TERC_SDM_R	AGAGTTGGGCTCTGTCAACCGCGGGTC
TERC_F1 (3'RACE)	CTCTGTCAGCCGCGGGTCTCTC
Universal RT Primer	CTACGTAACGATTGATGGTGCCTACAG
TERC_F2 (Northern probe)	GGGTTGCGGAGGGTGGGCCT
TERC_R1 (Northern probe)	GCATGTGTGAGCCGAGTCCTGG
Luciferase shRNA	CCGGCGCTGAGTACTTCGAAATGTCCTCGA GGACATTTTCGAAGTACTCAGCGTTTTTG
PARN shRNA1	CCGGCCGCACTGTATTTAACTTAATCTCGA GATTAAGTTAAATACAGTGCGGTTTTTG
PARN shRNA2	CCGGGCCTTTGGTAACATTCAGATACTCGA GTATCTGAATGTTACCAAAGGCTTTTTG
PAPD5 shRNA1	CCGGCGATGTTGGAAGGAGTTCATACTCGA GTATGAACTCCTTCCAACATCGTTTTTG
PAPD5 shRNA2	CCGGGCCACATATAGAGATTGGATACTCGA GTATCCAATCTCTATATGTGGCTTTTTG

4. RESULTS

4.1 A novel TERC CR4/CR5 domain mutation causes telomere disease via decreased TERT binding: Functional analysis of a novel variant in a previously known gene to cause telomere disease (22)

The proband was a previously healthy 36 year old man with an incidental finding of moderate thrombocytopenia during routine evaluation. He had otherwise normal hematologic parameters and physical findings, and a bone marrow exam showed normal cellularity and adequate megakaryocytes. Further evaluation revealed hepatosplenomegaly and portal hypertension, and a liver biopsy showed fibrosis and nodular regenerative changes. He was deemed to have liver disease of unclear etiology and secondary thrombocytopenia. Several months later, the proband's previously healthy 34 year old sister presented with pancytopenia and was found to have refractory anemia with excess blasts, type 1 (**Figure 4.1.1a-b**). She was treated with DNA hypomethylating therapy, but progressed to AML with complex cytogenetics. She attained remission after anthracycline and cytarabine induction chemotherapy, and was referred for allogeneic bone marrow transplantation. A detailed history and examination revealed her brother's liver abnormalities, their father's death from pulmonary fibrosis at 63 years of age, and subtle stigmata of DC in her and several extended family members, raising suspicion for telomere disease (**Figure 4.1.1a**). Telomere length testing in her and the proband showed very low mean lymphocyte telomere length ($\leq 1^{\text{st}}$ percentile; **Figure 4.1.1c**). Genetic testing revealed a heterozygous *TERC* c.319G>A variant of unknown significance but no mutations in *TERT*, *DKC1*, *NHP2*, *NOP10*, *TINF2*, or *TCAB1* (**Figure 4.1.2**).

a. Pedigree of the family



b. Histology of bone marrow biopsy

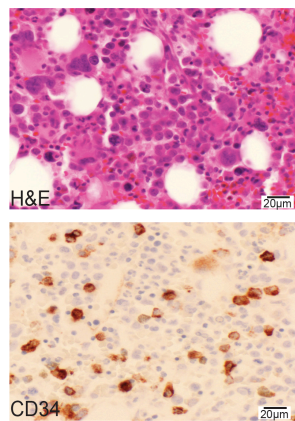
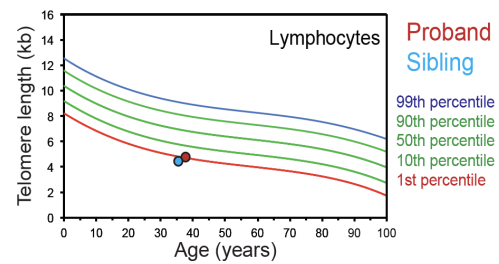
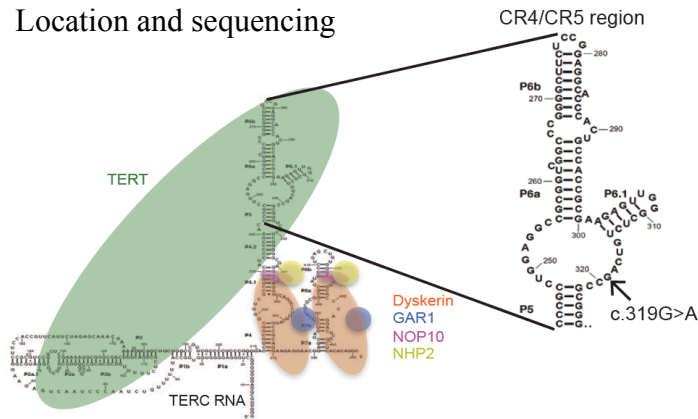
c. Telomere length measurement
Flow-FISH

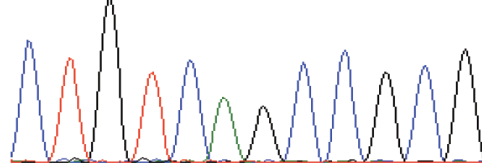
Figure 4.1.1: A family with telomere disease (a) Pedigree of the family; the proband is indicated with an arrow. Affected individuals are shown in gray and their symptoms are written below. Mutations in the individuals genotyped are shown. TP: thrombocytopenia, IPF: idiopathic pulmonary fibrosis **(b)** Histology image of the bone marrow biopsy (Scale bar showing 20 µm distance); (top) Hypercellular fibrotic bone marrow with dyspoiesis is observed with hematoxylin and eosin (H&E) staining, (bottom) CD34 staining shows increased percentage of CD34-positive cells **(c)** Telomere length measurement by flow-FISH in lymphocytes in the proband and his sister.

a. *TERC* c.319G>A mutation

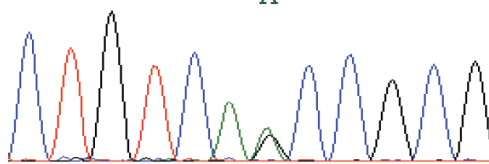
Location and sequencing



WT: C T G T C A G C C G C G



Proband: C T G T C A G C C G C G

b. *TERC* sequences in vertebrates

Human	CTCTGTCAGCCG--CGGGTCCT
Rhesus	CTCTGTCAGCCG--CGGGTCCT
Mouse	CTCTGTCAGCCG--CGGGGCC
Armadillo	CTCTGTCAGCCG--CGGATCC
Cow	CTCTGTCAGCCG--CGGGTCCT
Cat	CTCTGTCAGCCG--CGGGTCCT
Dog	CTCTGTCAGCCG--CGGGGCC
Elephant	CTCTGTCAGCCG--CGGGTCCT
Chicken	CTCTGTCAGCCT--CGGCG--
X. tropicalis	CTCTGTTAGCCCTTGGGG---
Stingray	CTCTGTTAGCCCT--GGCA---
Dogfish Shark	CTCTGTCAGCCC--GGCT---

Figure 4.1.2: Variant of unknown significance in *TERC* gene (a) (top) Schematic structure of telomerase ribonucleoprotein and the specific location of the variant in the TERC CR4/CR5 region is indicated with an arrow. (bottom) Sequencing of TERC in the proband shows the c.319G>A variant (b) Phylogenetic conservation of TERC nucleotide 319, which is mutated in this family

Given the implications of diagnosing telomere disease in this family, we undertook a detailed evaluation of the TERC G319A variant. The guanine at this position in the TERC CR4/CR5 domain is highly conserved amongst vertebrates (45) (**Figure 4.1.2**), and human genetic variation at this position is not evident in the Exome Aggregation Consortium database, suggesting an important albeit undefined role. We tested the function of TERC G319A in a commonly used cell-based assay. When we ectopically expressed normal or G319A TERC along with TERT in VA13 cells, which are TERT and TERC negative, we found that the G319A TERC was capable of reconstituting telomerase activity, but at levels ~40% lower than normal TERC (**Figure 4.1.3a**). Expression levels of TERC and TERT appeared equivalent across experimental samples (**Figure 4.1.3b**). We regarded these results as suggestive but inconclusive that the TERC G319A variant is pathological, given the modest decrease in telomerase activity and the unclear mechanism by which it would compromise TERC function.

a. Telomerase activity in VA13 cells- TRAP assay

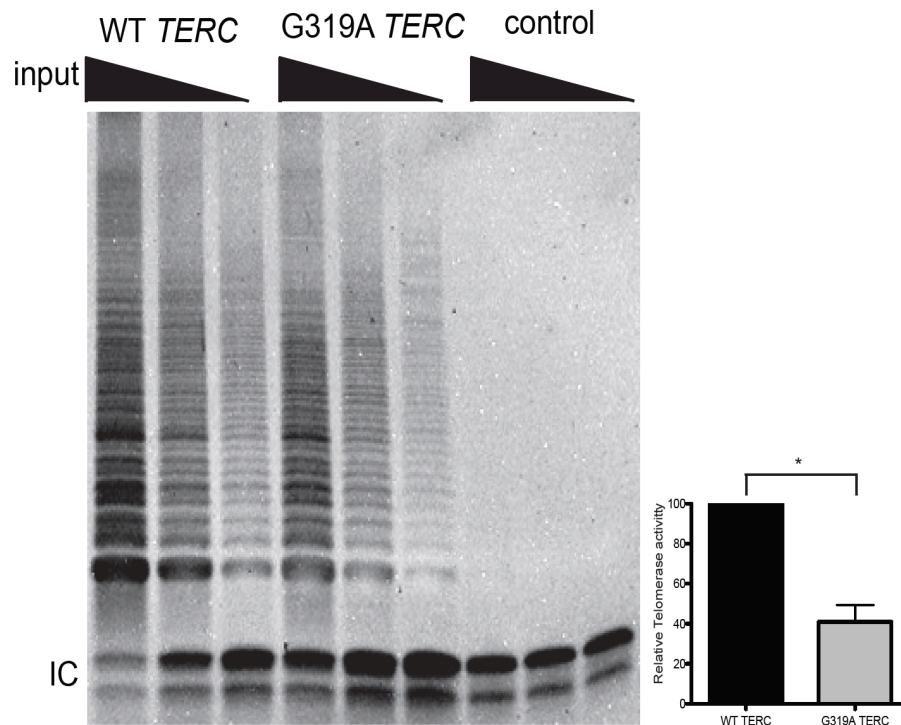
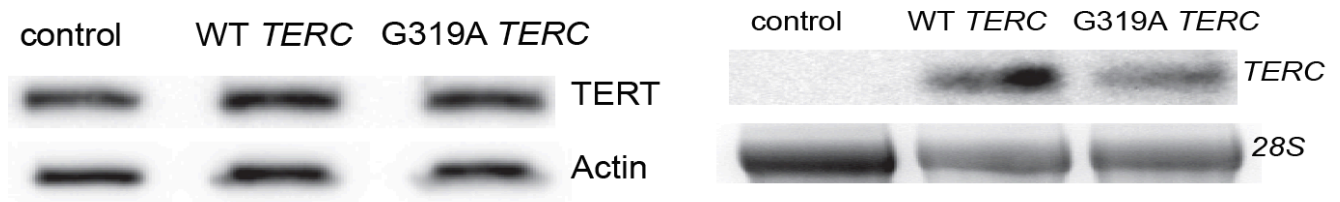
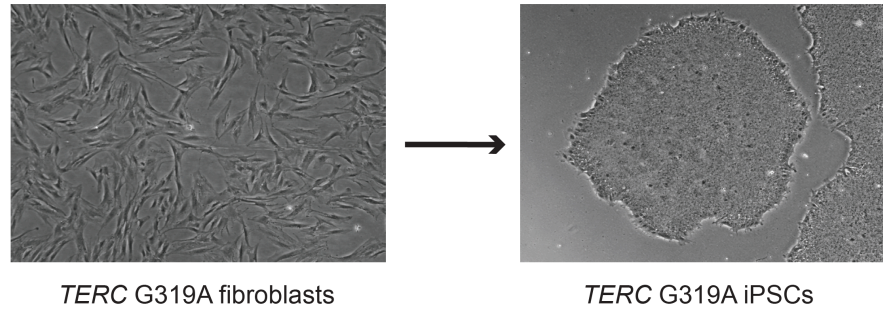
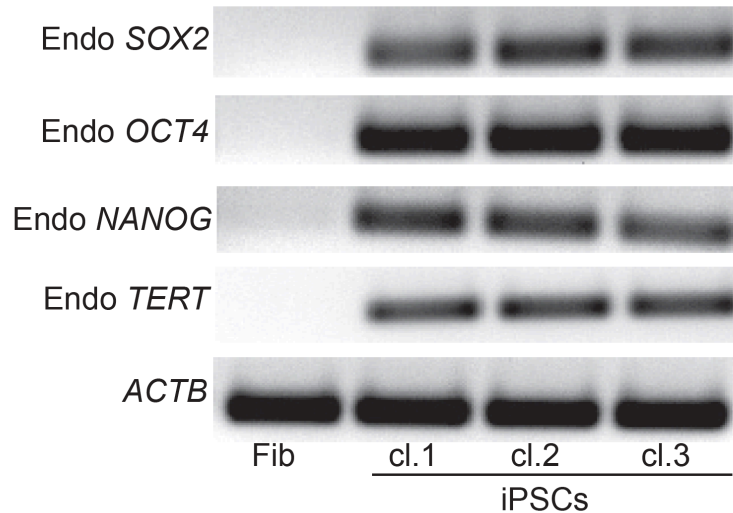
b. TERT and TERC levels in VA13 cells
Western and Northern blot

Figure 4.1.3: Decreased telomerase activity with mutant *TERC* (a) TRAP assay for telomerase activity in VA13 cells transfected with TERT and either of WT, G319A *TERC* or control plasmid, using 3-fold dilutions of input cell extract. Internal control (IC) amplification standard is indicated. Relative telomerase activities are shown in the graph. (b) (left) Immunoblot of TERT and actin protein levels in VA13 cells transfected with TERT. (right) Northern blot of *TERC* RNA from VA13 cells transfected with either WT, G319A *TERC* or control plasmid. Ethidium bromide staining of 28S rRNA is used as a loading control.

We turned our attention to patient cells by culturing bone marrow-derived fibroblasts from the proband. To overcome senescence and limitations of cell numbers, and to study the function of TERC G319A in TERT-expressing cells, we reprogrammed the proband's fibroblasts using a lentiviral vector encoding OCT4, SOX2, KLF4 and MYC, and isolated several iPSC clones (**Figure 4.1.4**). All *TERC* het.G319A iPSC clones showed low telomere length compared to WT iPSCs (**Figure 4.1.5a-b**), consistent with constitutionally short telomeres, but displayed continuous self-renewal in culture. We again found no differences in overall TERC levels in normal versus *TERC* het.G319A iPSCs (**Figure 4.1.5c**). By sequencing cDNA from patient iPSCs, we observed that transcripts from both the normal and G319A *TERC* alleles were equally represented (**Figure 4.1.5d**). These data indicate telomere maintenance defects in *TERC* het.G319A patient cells, but that neither transcription nor post-transcriptional accumulation of TERC is compromised by the G319A variant.

a. *TERC* G319A cells- morphology

b. RT-PCR for pluripotency associated genes



c. Karyotype of G319A iPSCs

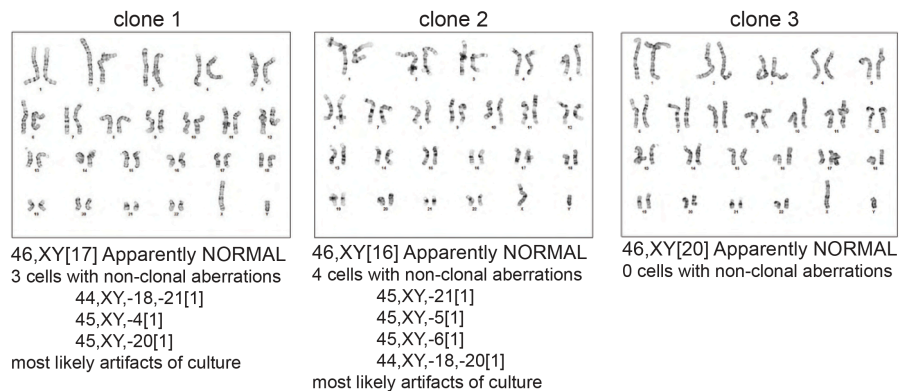
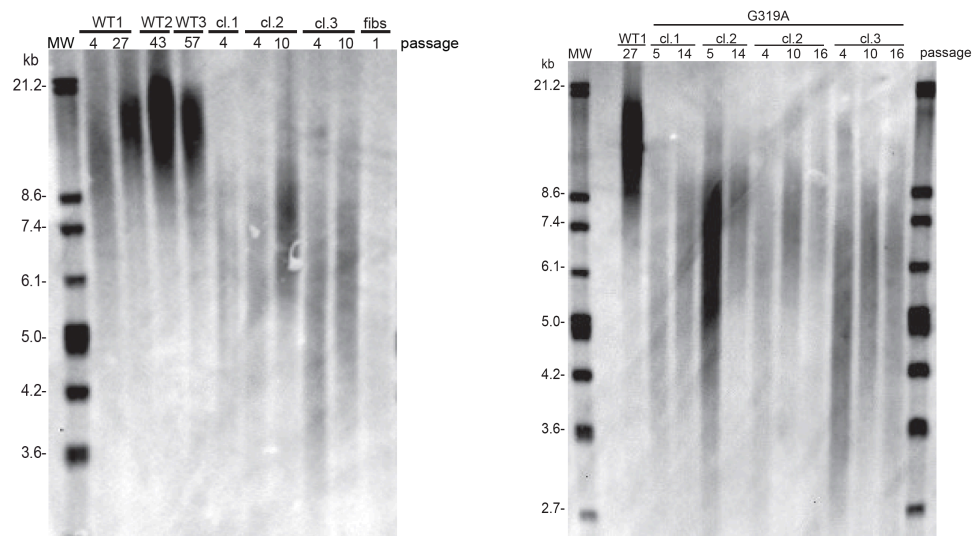


Figure 4.1.4: Reprogramming of *TERC* G319A fibroblasts (a) Fibroblasts and iPS colony morphology (40x magnification) (b) RT-PCR for endogenous (endo) *SOX2*, *OCT4*, *NANOG*, *TERT* and *ACTB* transcripts in three different *TERC*-mutant iPSC clones (cl.) and *TERC*-mutant fibroblasts (fib) (c) G-band karyotyping of G319A iPSCs with cytogenetic profile (below each image).

a. Telomere length analysis- Southern blot



b. TERC levels

Northern blot and cDNA sequencing

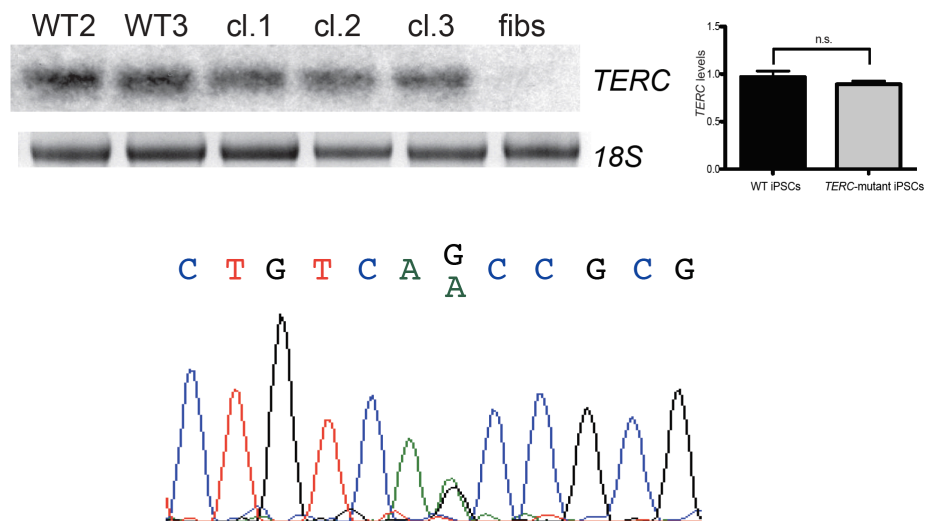


Figure 4.1.5: Telomere length and *TERC* RNA levels in *TERC* G319A iPSCs (a) Telomere restriction fragment length (TRF) analysis of WT iPSCs, 3 different clones (cl.) of *TERC*-mutant iPSCs and *TERC*-mutant fibroblasts **(b)** (top, left) Representative Northern blot of TERC RNA from WT, *TERC*-mutant iPSCs and *TERC*-mutant fibroblasts. (top, right) Relative *TERC* levels are shown in the graph. (bottom) Sequencing of *TERC* cDNA in *TERC*-mutant iPSCs.

Prior studies indicate that, in addition to the TERC template region, the CR4/CR5 domain serves as an independent site to which TERT binds in the telomerase holoenzyme (65-66). We hypothesized that, despite its location outside the P6b and P6.1 stem-loops in CR4/CR5 (**Figure 4.1.2**) that are known to be required for TERT association (65-67), the TERC G319 residue might also be important for binding to TERT. We exploited our ability to generate millions of patient iPSCs and the heterozygosity of the TERC variant to determine the relative association of normal versus G319A TERC with TERT, in the same cell. By immunoprecipitation (IP), we isolated TERC-containing ribonucleoprotein complexes (RNPs) either using anti-dyskerin or anti-TERT antibodies. After validating our ability to isolate dyskerin- and TERT-RNPs (**Figure 4.1.6**), we performed IPs from *TERC* het.G319A patient iPSCs. Sequencing of the RT-PCR products of dyskerin-bound RNA from patient iPSCs showed equal recovery of the G319A and normal TERC species (**Figure 4.1.7**). In contrast, when we pulled down TERT RNPs, we found a decreased proportion of G319A versus normal TERC, in three independent replicates (**Figure 4.1.7-8**). We exploited the gain of a HincII restriction site in the mutant TERC cDNA sequence to quantitate the relative abundance of G319A versus normal TERC associated with TERT, and found it to be diminished by ~75% (**Figure 4.1.9**). These results demonstrate the TERC G319A mutation compromises telomerase function via decreasing binding of TERC to TERT *in vivo*.

RT-PCR products after immunoprecipitation

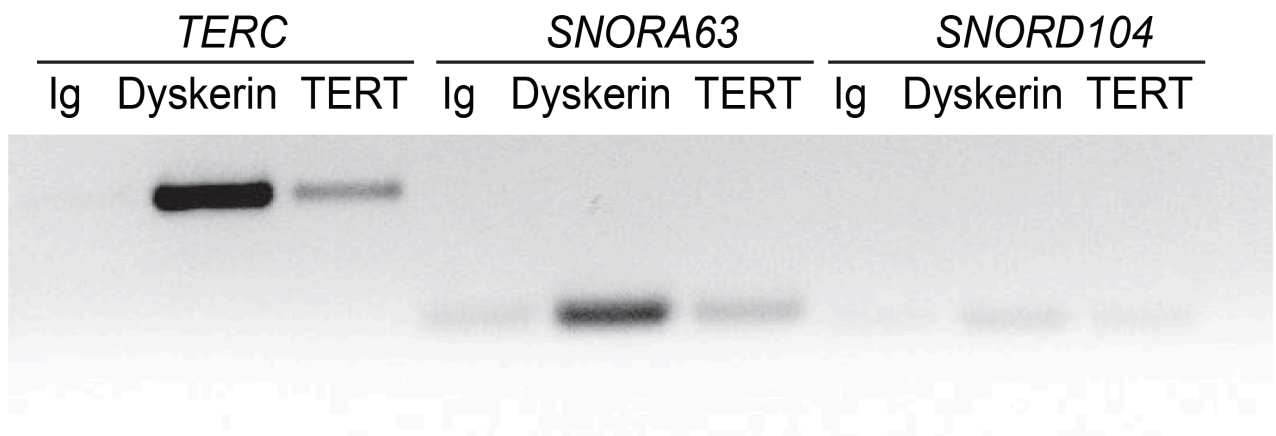


Figure 4.1.6: RT-PCR after Dyskerin and TERT Immunoprecipitation Agarose gel electrophoresis image of RT-PCR products for *TERC*, *SNORA63* and *SNORD104* following Dyskerin and TERT immunoprecipitation in *TERC*-mutant iPSCs; *SNORD63* RT-PCR used as a positive control for Dyskerin IP and *SNOR104* used as a negative control for all. Ig: Immunglobulin control

Dyskerin and TERT immunoprecipitation in iPSCs
TERC RT-PCR sequencing

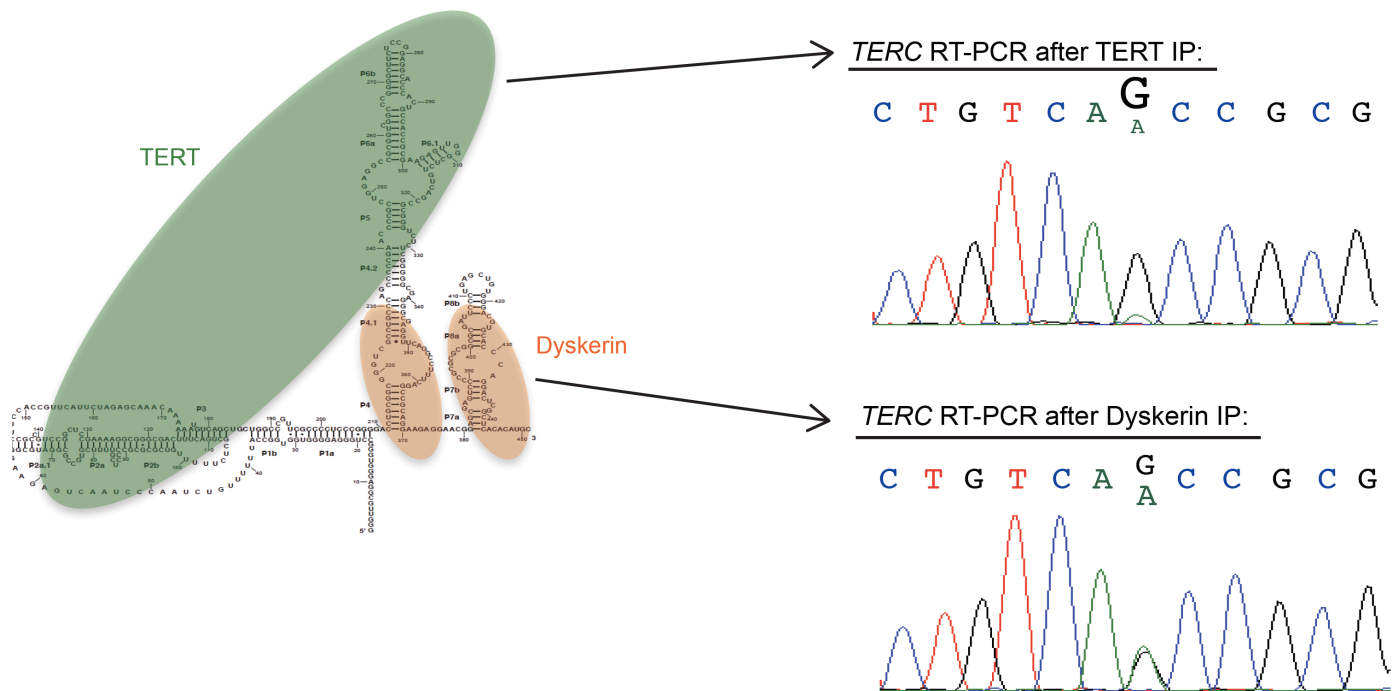


Figure 4.1.7: *TERC* RT-PCR after Dyskerin and TERT Immunoprecipitation

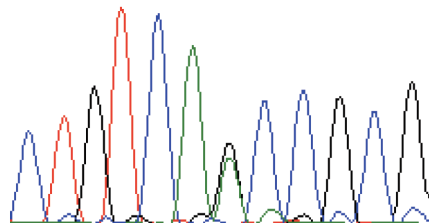
(left) Schematic structure of TERC RNA bound to TERT and dyskerin. (right) Sequencing of TERC RT-PCR amplicons after TERT immunoprecipitation (top) and dyskerin immunoprecipitation (bottom) in *TERC*-mutant iPSCs are shown.

Dyskerin and TERT immunoprecipitation in iPSCs *TERC* RT-PCR sequencing

TERC-mutant iPSCs clone 1:

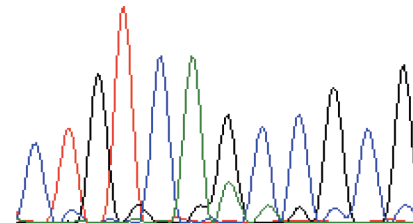
TERC RT-PCR after Dyskerin IP

C T G T C A ^G_A C C G C G



TERC RT-PCR after TERT IP

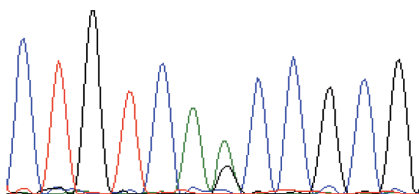
C T G T C A ^G_A C C G C G



TERC-mutant iPSCs clone 3:

TERC RT-PCR after Dyskerin IP

C T G T C A ^G_A C C G C G



TERC RT-PCR after TERT IP

C T G T C A ^G_A C C G C G

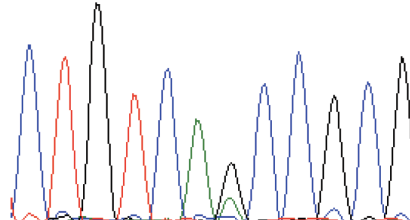


Figure 4.1.8: *TERC* RT-PCR after Dyskerin and TERT Immunoprecipitation, other clones Chromatograms of *TERC* RT-PCR products after Dyskerin and TERT immunoprecipitation of *TERC*-mutant iPSCs.

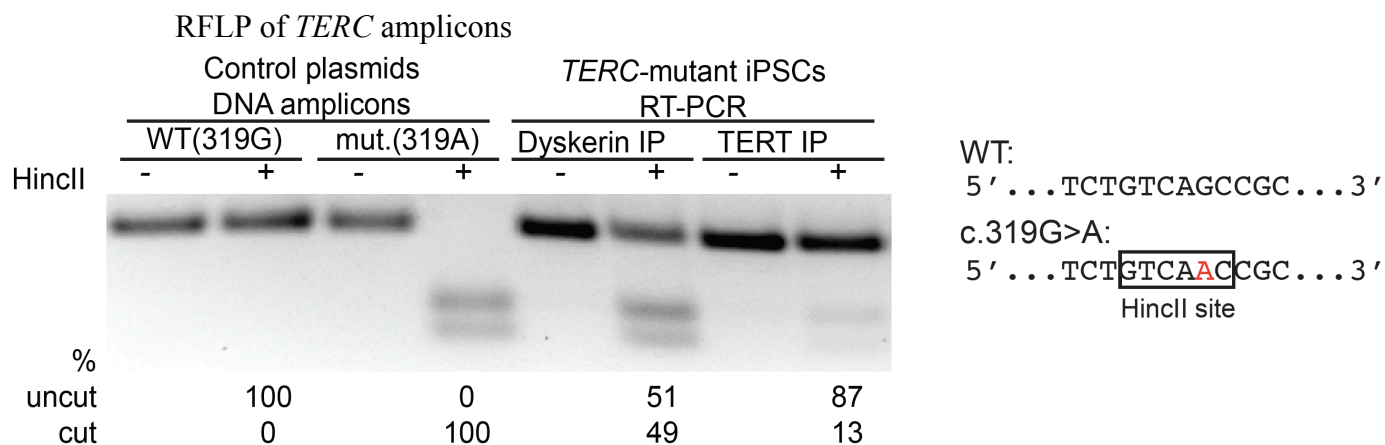
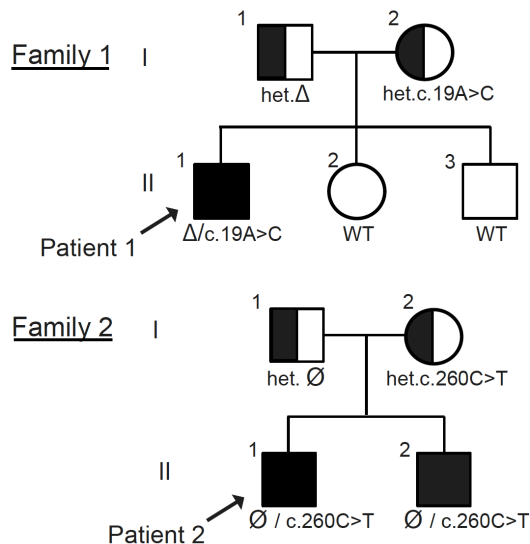


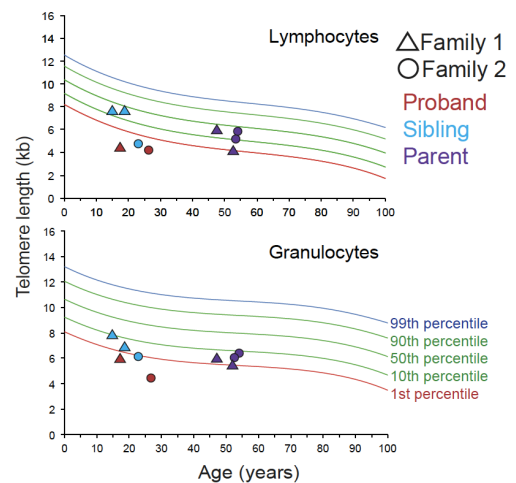
Figure 4.1.9: RFLP of *TERC* amplicons Restriction fragment length polymorphism (RFLP) of *TERC* amplicons after RT-PCR from Dyskerin and TERT IP of patient iPSCs (lanes 5-8). The HincII site created by the G319A mutation is depicted on the right side. Lanes 1-4 show digest patterns of amplicons from WT (319G) and mutant (319A) plasmid DNA. Products are separated by agarose gel electrophoresis and band intensity was quantified using ImageLab (Biorad)

4.2 Poly(A)-specific ribonuclease (PARN) mediates 3' end maturation of the telomerase RNA component: Functional study on how *PARN* mutations, a recent genetic discovery causing telomere diseases, are involved in telomere biology (20)

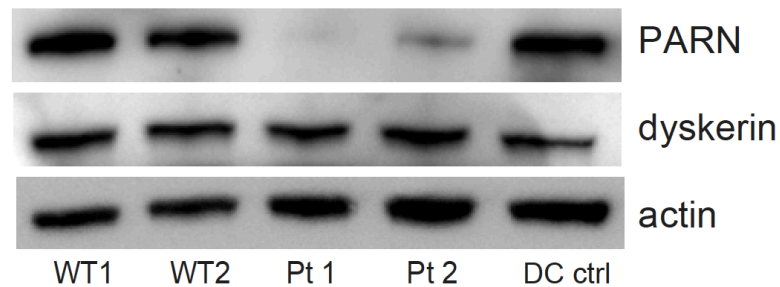
PARN loss-of-function mutations have recently been implicated in IPF and DC but why deficiency of *PARN*, which has no known role in telomere biology, should mimic telomere diseases is unexplained (17-18). By candidate gene sequencing in the subjects with genetically uncharacterized dyskeratosis congenita / telomere disease in the Pediatric Myelodysplastic Syndrome / Bone Marrow Failure registry at Boston Children's Hospital, we identified biallelic defects in *PARN* in two families (**Figure 4.2.1a**), representing ~15% of the unrelated families in this disease category. The probands manifested classic features of dyskeratosis congenita and associated findings, including bone marrow failure and very short peripheral blood cell telomere length (**Figure 4.2.1b**). Patient 1 was found to carry an undescribed missense variant c.19A>C on one allele, inherited from his mother, resulting in a substitution of a highly conserved amino acid p.Asn7His, and a large deletion encompassing the entire *PARN* gene on the other allele, inherited from his father. Patient 2 was found to carry a heterozygous missense variant c.260C>T, encoding the substitution of a highly conserved amino acid, p.Ser87Leu, inherited from his unaffected mother. He and his affected brother had no other potentially pathogenic exon-encoded variants. In keeping with these findings, *PARN* protein levels were compromised (**Figure 4.2.1c**). These data show compound heterozygous loss-of-function mutations in *PARN* in two families with dyskeratosis congenita and very short telomeres.

a. Pedigree of families with *PARN* mutations

b. Flow- FISH telomere length analysis



c. Protein levels in fibroblasts- Western blot

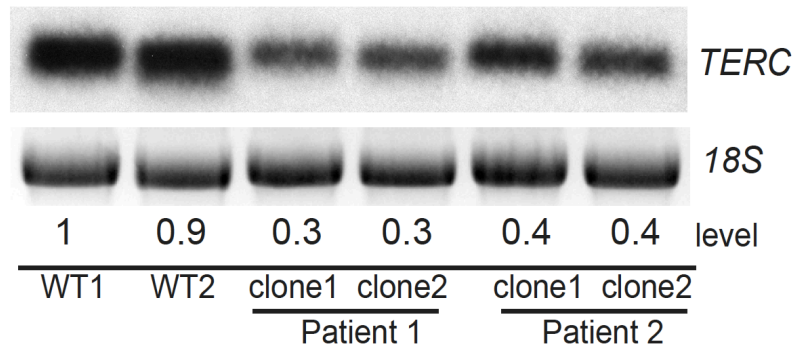
**Figure 4.2.1: *PARN* mutations in two families with dyskeratosis congenita**

(a) Segregation of compound heterozygous *PARN* mutations in families 1 and 2. Proband is indicated by arrows. Clinically affected individuals are shown as filled shapes and carriers as half-filled shapes. Nucleotide substitutions are indicated, as well as the inheritance of a large deletion (Δ) encompassing the *PARN* locus in family 1, and a non-coding defect affecting accumulation of *PARN* transcripts in family 2 ($\emptyset$). **(b)** Telomere length measurement by flow cytometry fluorescence in situ hybridization in lymphocytes (upper panel) and granulocytes (lower panel). Members of family 1 and family 2 are indicated by triangles and circles, respectively. Proband is in red, siblings in blue and parents in purple. **(f)** Western blot of *PARN*, dyskerin and actin proteins in fibroblasts from normal individuals, patients 1 and 2 with *PARN* mutations, and a dyskeratosis congenita patient without *PARN* mutations (DC ctrl).

To overcome potential artifacts and limitations in studying near-senescent patient cells, we generated iPS cells from the fibroblasts of patients 1 and 2 (**Figure 4.2.2**), which also showed PARN deficiency and reduced *TERC* levels compared to normal iPS cells, without diminished dyskerin protein (**Figure 4.2.2**). To further investigate how *PARN* affects *TERC* levels, we knocked down *PARN* in HEK293 (293) cells, and again found reductions in *TERC* levels but not *DKCI* or dyskerin (**Figure 4.2.3**). These results demonstrate that *PARN* deficiency results in diminished *TERC* levels, independent of *DKCI*/dyskerin.

Box H/ACA snoRNAs are intron-encoded and undergo maturation by exonucleolytic processing of spliced intermediates (68). Nascent snoRNAs are subject to oligo-adenylation by TRF4-2 (16,69) (also known as PAPD5), a component of the TRAMP (TRF4-2, AIR2, MTR4) complex (70), which stimulates degradation of nuclear RNAs by the exosome. PARN counteracts snoRNA oligo-adenylation and promotes 3' end maturation, which may prevent further TRAMP-mediated oligo(A) addition and subsequent degradation. *TERC* is transcribed by RNA polymerase II (RNA pol II) as an autonomous genetic unit, has 7-methyl-guanosine cap and a precise 3' end, but unlike other RNA pol II transcripts does not contain a long poly(A) tail, and the mechanisms of 3' end maturation are unknown (2, 46, 71, 72). Based on the box H/ACA architecture shared by snoRNAs and *TERC*, and the recent finding that *TERC* exists in oligo(A)-containing forms (73), we hypothesized that PARN participates in *TERC* 3' end maturation and stabilization by removing oligo(A) in nascent transcripts.

a. TERC levels in iPS cells- Northern blot



b. Protein levels in iPS cells- Western blot

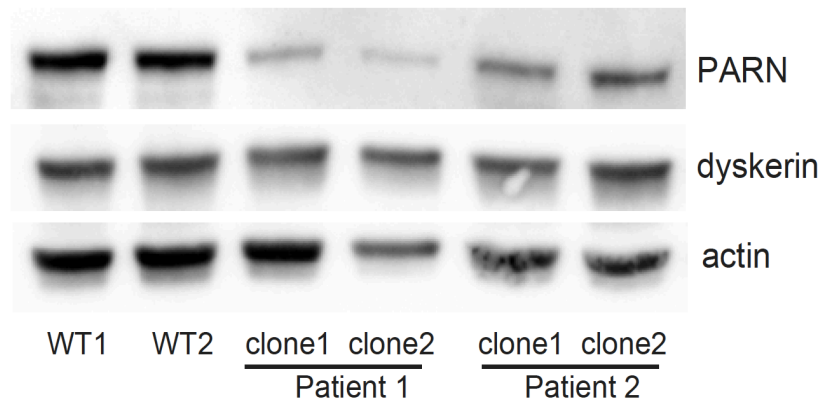
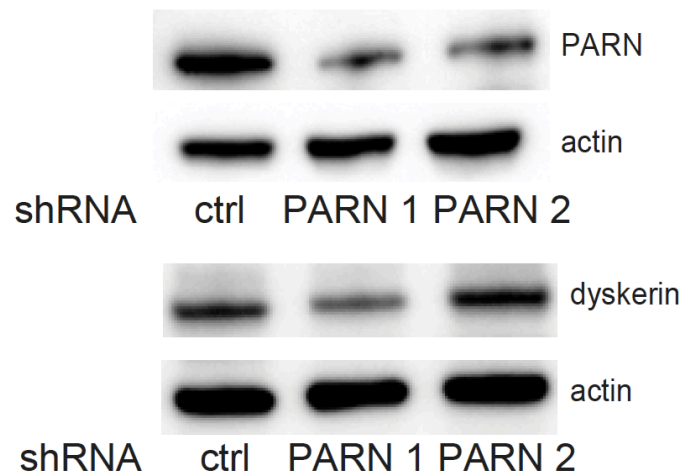


Figure 4.2.2: Decreased TERC levels in PARN deficient iPS cells (a) Northern blot of TERC following denaturing agarose gel electrophoresis of RNA from iPS cells. Ethidium bromide staining of 18S rRNA was used as a loading control. Normalized TERC levels relative to WT1 are indicated. **(b)** Representative western blot of PARN, dyskerin and actin proteins in iPS cells.

a. Protein levels in 293 cells- Western blot



b. Transcript levels in 293 cells- qPCR

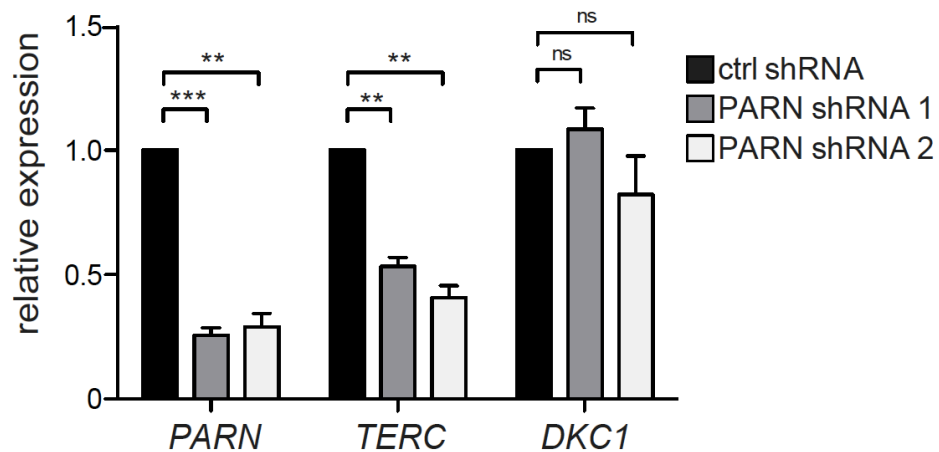
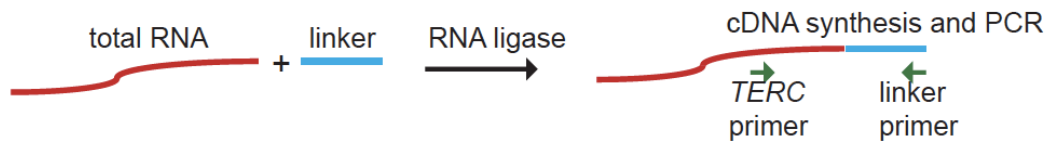


Figure 4.2.3: Decreased TERC levels in PARN deficient 293 cells (a) Representative western blot of PARN, dyskerin and actin proteins in 293 cells transduced with lentivirus encoding shRNA directed against PARN versus luciferase. **(b)** qPCR of PARN, TERC, and DKC1 transcripts from 293 cells transduced with lentivirus encoding shRNA directed against PARN versus luciferase as a control (ctrl) (n=3).

Because *TERC* oligo(A) tails average 6-7 nucleotides in length and may not be quantified accurately using oligo(dT)₁₀, we performed 3' rapid amplification of cDNA ends (RACE) in normal and *PARN*-mutant patient cells (**Figure 4.2.4a**). We observed an alteration in the size distribution of *TERC* 3' ends in both patient fibroblasts and iPS cells compared to normal cells, in that the intensity of the amplicon at the expected size was reduced, and larger, extended transcripts predominated (**Figure 4.2.4b**). These results were recapitulated in 293 cells upon transient *PARN* knockdown (**Figure 4.2.4b**). These data indicate that *PARN* deficiency results not only in diminished *TERC* steady-state levels but also in increased *TERC* RNA species with abnormal 3' ends.

To investigate this finding more closely, we profiled the 3' ends of *TERC* RNA in fibroblasts and iPS cells by deep sequencing (**Figure 4.2.5**). We found a significant increase in the proportion of *TERC* species other than mature *TERC* in patient cells versus normal cells, including transcripts extended beyond the canonical 3' end, and a diverse range of non-genomically encoded additions. We focused our attention on the distribution and frequency of two classes of *TERC* species: those containing up to 8 genomically-encoded bases beyond the canonical *TERC* end, and the corresponding oligo(A) forms, which together comprised the majority of *TERC* forms across all samples. In patient fibroblasts, we found that the percentage of mature *TERC* (i.e. ending at the canonical 5'-...AUGC-3' sequence) was markedly reduced compared to normal fibroblasts (59% versus 84%; Fig. 4a). Because the total amount of *TERC* in *PARN* deficient cells is ~45% of normal cells (**Figure 4.2.2-3**), the level of mature *TERC* in patient fibroblasts is approximately one-third of that found in normal fibroblasts (27% versus 84%). At the same time, we found a significant increase in the proportion of oligo(A) forms of *TERC*, including transcripts with genomically encoded sequences beyond the mature *TERC* 3' end (**Figure 4.2.5**). Collectively, the analyses performed in the setting of *PARN* deficiency reveal intermediates in *TERC* biogenesis, and indicate a critical and non-redundant function for *PARN* in counteracting the oligo-adenylation of nascent *TERC* transcripts.

a. 3'RACE strategy



b. 3'RACE TERC PCR amplicons

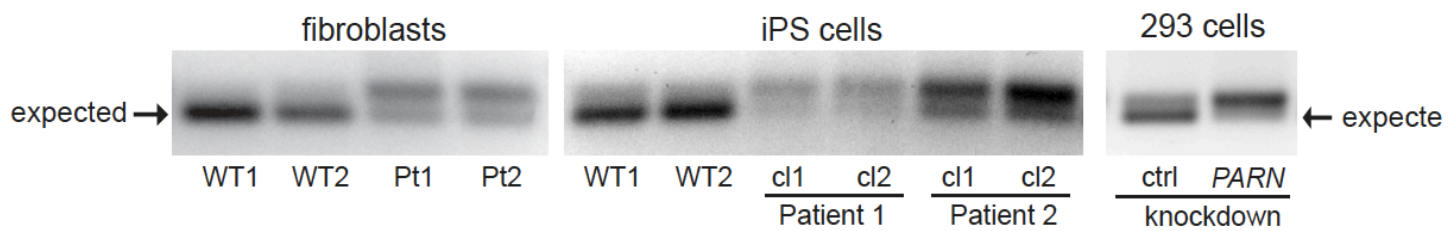


Figure 4.2.4: PARN deficiency results in abnormal TERC 3' ends (a) 3'RACE strategy: a universal linker is ligated to the 3' ends of total RNA. cDNA synthesis with a linker-specific primer followed by PCR using linker- and TERC-specific primers yields amplicons representing the diversity of TERC 3' ends. (b) 3'RACE products from normal and patient (Pt) fibroblasts (left) and iPS clones (cl; middle), and 293 cells subjected to shRNA-mediated knockdown of PARN versus luciferase (ctrl) (right). Amplicons were separated by agarose gel electrophoresis, and the expected size (184 bp) representing mature TERC is shown (arrow).

Deep sequencing of 3'RACE TERC PCR amplicons

TERC gene:

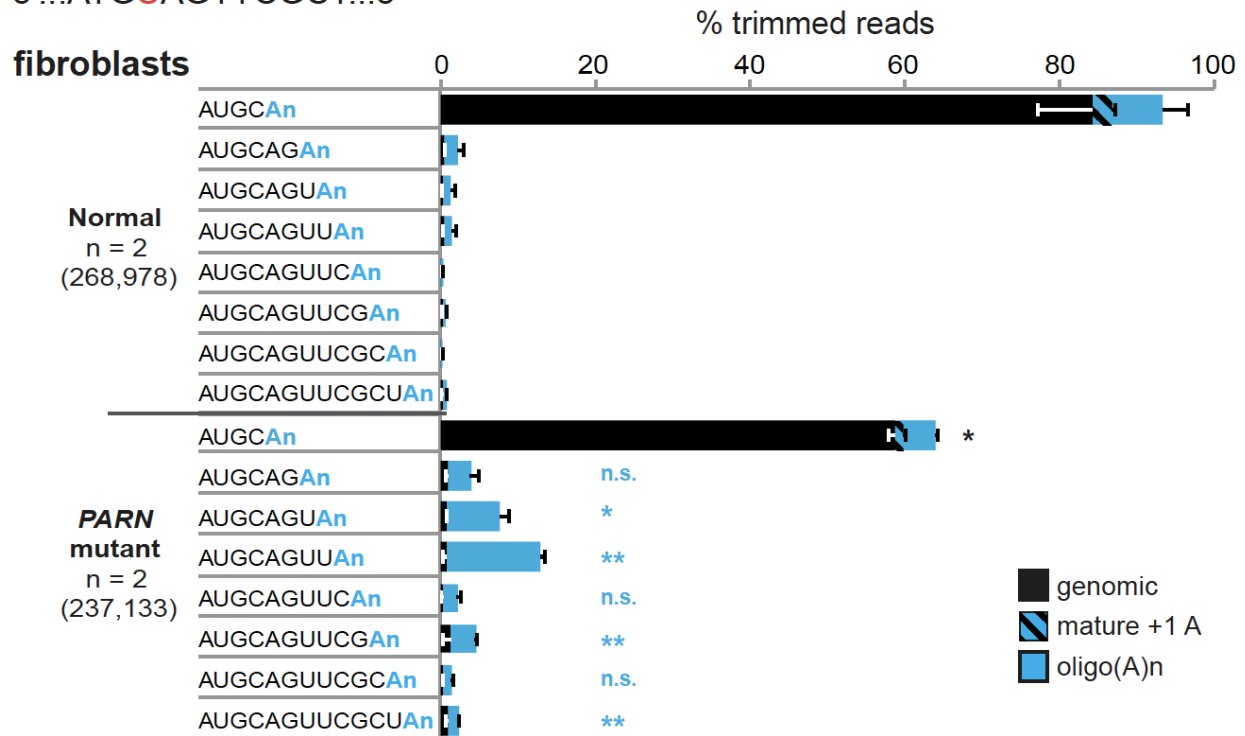
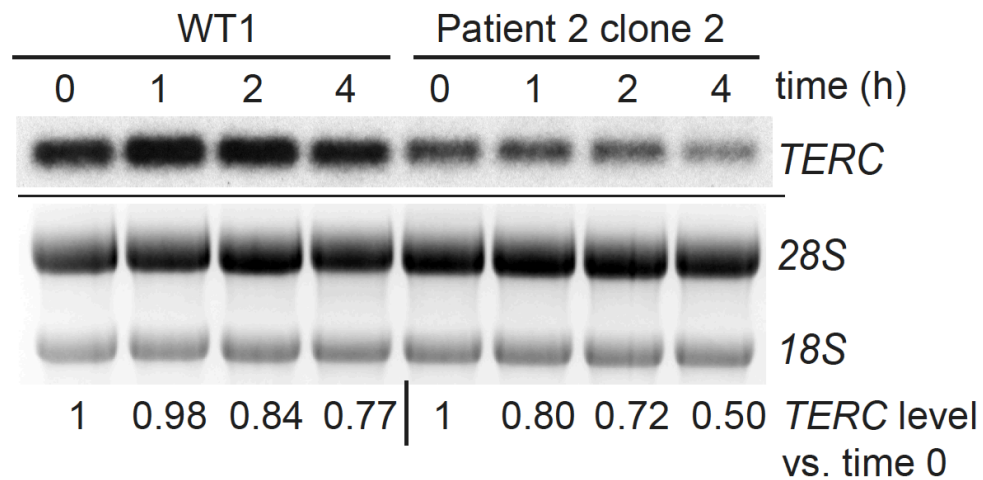
5'...ATG**C**AGTTCGCT...3'

Figure 4.2.5: Decreased mature TERC and increased 3' extended TERC species in PARN deficient patient cells Fibroblasts: 3' RACE PCR products from normal and PARN mutant patient fibroblasts were subjected to deep sequencing and aligned to the TERC gene. The canonical TERC 3' terminus is shown in red, in the context of the TERC gene sequence. The relative proportions of mature TERC and TERC RNA species extending up to 8 bases into the genomic sequence, with or without post-transcriptionally added oligo(A) tails, are depicted. Genomically-encoded termini are in black, mature TERC with a single A (which may or may not be genomically-encoded) is hatched, and oligo(A) additions of any length (n) are in solid blue. Proportions are averaged for normal fibroblasts and patient fibroblasts (n=2 for each group). The total number of trimmed reads for each group is shown in parentheses. Error bars represent standard deviations. For statistical evaluations, mature TERC forms and oligo(An) ends for all genomically-extended TERC species were compared between normal and patient cells in a two-tailed t-test. Significance is indicated by p-values (* ≤ 0.05 ; ** ≤ 0.01 ; *** ≤ 0.001 ; n.s. not significant; black for mature TERC, blue for extended, oligo(An) forms).

Oligo-adenylation is predicted to target nuclear RNA transcripts for degradation, and therefore we determined the decay rate of TERC RNA in PARN-deficient versus normal cells. We inhibited pol II transcription using actinomycin and quantified TERC RNA by Northern blot. We found a decreased half-life of TERC transcripts in patient iPS cells compared to normal iPS cells (**Figure 4.2.6**). These results indicate that PARN deficiency results in destabilization of TERC RNA, and provide a mechanism for the observed decrease in steady-state TERC levels.

Because PARN's canonical role is believed to be in mRNA turnover, we performed RNA-Seq to assess the effects of PARN deficiency on mRNAs globally. Remarkably, we found that no protein-coding mRNAs manifested a fold change (increase or decrease) in steady-state levels that consistently exceeded the change in TERC levels in all 7 sample pairs (**Table 4.2.1**). These results suggest that the degree of PARN deficiency that is sufficient to disrupt TERC RNA transcript levels, 3' end processing, and stability does not result in changes of a similar magnitude in the levels of mRNAs expressed across cell types.

- a. Actinomycin treatment of iPS cells
TERC levels- Northern blot



- b. *TERC* decay rate

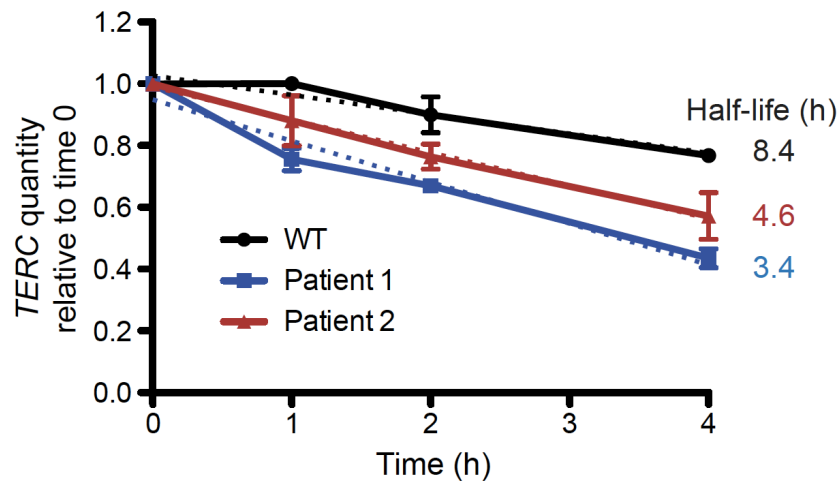


Figure 4.2.6: Decreased stability of TERC in PARN deficient cells

(a) Representative Northern blot of TERC RNA levels in normal versus patient iPS cells, at 0, 1, 2 and 4 hours (h) following actinomycin treatment. WT1 and patient 2 clone 2 iPS cells are shown here. Ethidium bromide staining of 28S and 18S rRNAs is shown, and 18S rRNA was used as a loading control. Normalized TERC levels relative to time 0 for each sample are indicated. **(b)** Graph of TERC decay rate in normal and patient iPS cells, calculated from Northern blot analysis as shown in (a) ($n=2$ for each group; error bar represent standard deviation). The dashed line reflects the slope determined by simple linear regression, and half life is the calculated x-intercept at $y=0.5$.

Table 4.2.1: RNA-seq transcriptome analysis in PARN deficient cells

		Altered across all 7 pair-wise comparisons			Altered in any 6 of 7 pair-wise comparisons			Altered in any 5 of 7 pair-wise comparisons		
	Total genes analyzed	Obs	Exp	p-value	Obs	Exp	p-value	Obs	Exp	p-value
Decreased										
mRNA	12937	0	0.96	0.02	1	3.83	7.0E-09	4	10.5	1.5E-19
rRNA	63	0	0	1	0	0.02	1	0	0.05	1
snoRNA	249	1	0.02	0.0003	3	0.07	1.9E-19	6	0.2	1.5E-32
snRNA	259	0	0.02	1	0	0.08	1	1	0.21	0.52
Increased										
mRNA	12937	0		NA	1	1.92	0.14	9	10.5	0.12
rRNA	63	0		NA	0	0.01	1	0	0.05	1
snoRNA	249	0		NA	1	0.04	0.01	2	0.2	0.004
snRNA	259	0		NA	0	0.04	1	0	0.21	1

mRNA: protein-coding mRNA; rRNA: ribosomal RNA; snoRNA: small nucleolar RNA (includes box H/ACA, C/D, scaRNA); snRNA: small nuclear RNA

Obs: observed; Exp: expected by chance; NA: not analyzed.

Bold: significantly more than expected; *Italic*: significantly fewer than expected

4.3 Post-transcriptional manipulation of TERC reverses molecular hallmarks of telomere disease: Identification of a novel molecule in TERC post-transcriptional processing pathway and how its manipulation affects telomere dynamics (21)

To directly assess the effects of PARN deficiency on telomere maintenance, we disrupted *PARN* in HEK293 cells using lentivirus-encoded short-hairpin RNAs (shRNAs), and continuously cultured the transduced cells under antibiotic selection. In *PARN* knockdown cells, using two independent shRNAs, we observed telomere attrition over time as compared to control cell lines (**Figure 4.3.1**). When we overexpressed a *PARN* cDNA that is not susceptible to knockdown by the shRNA, we observed increased telomere length as compared to control cells (**Figure 4.3.2**). Next, we evaluated the effects of restoring PARN expression on telomere maintenance in patient cells. In TERT-immortalized fibroblasts from patients with DC caused by *PARN* loss-of-function mutations, we found that PARN expression restored telomere elongation. These results confirm that PARN deficiency compromises telomere maintenance, and indicate that PARN is limiting for telomere maintenance in human cells, despite the presence of TERT (**Figure 4.3.3**).

Telomere length analysis- Southern blot

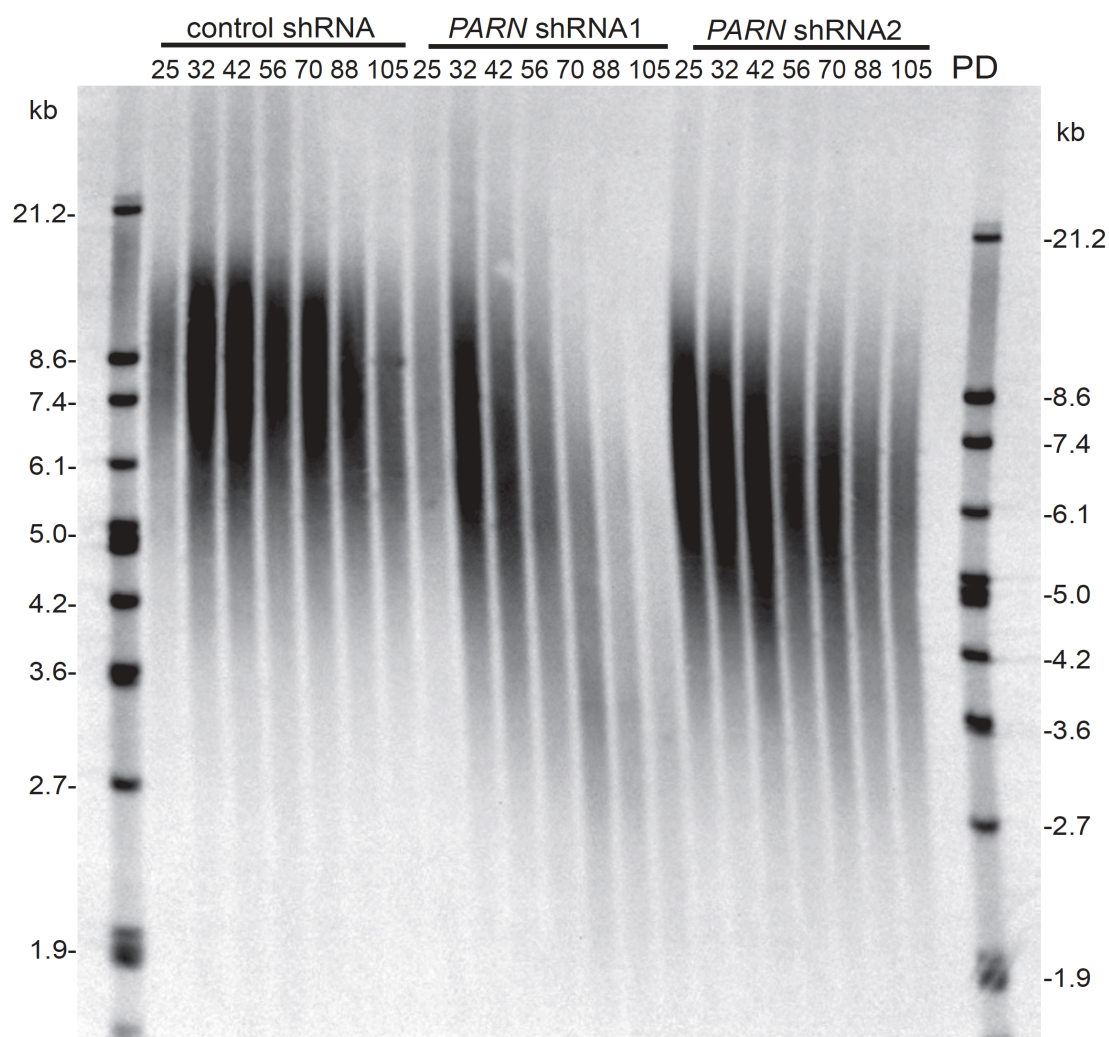


Figure 4.3.1: Telomere length attrition in PARN deficient cells. Telomere restriction fragment (TRF) length in HEK293 cells infected with shRNAs targeting luciferase (control) versus *PARN*.

Telomere length analysis- Southern blot

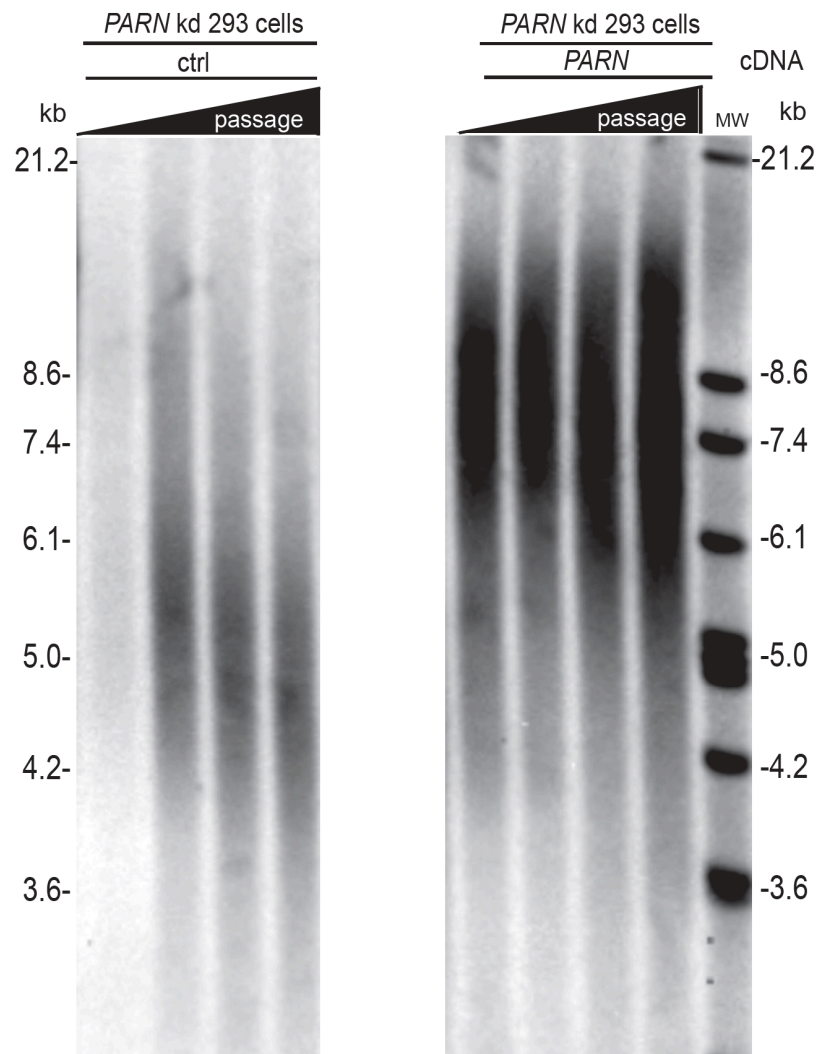


Figure 4.3.2: Telomere length elongation in *PARN* kd 293 cells after *PARN* overexpression Telomere restriction fragment (TRF) length in *PARN* kd HEK293 cells. Cells were transduced with lentivirus containing vector alone (ctrl) or encoding *PARN*. Telomere length was followed for 12 passages.

Telomere length analysis- Southern blot

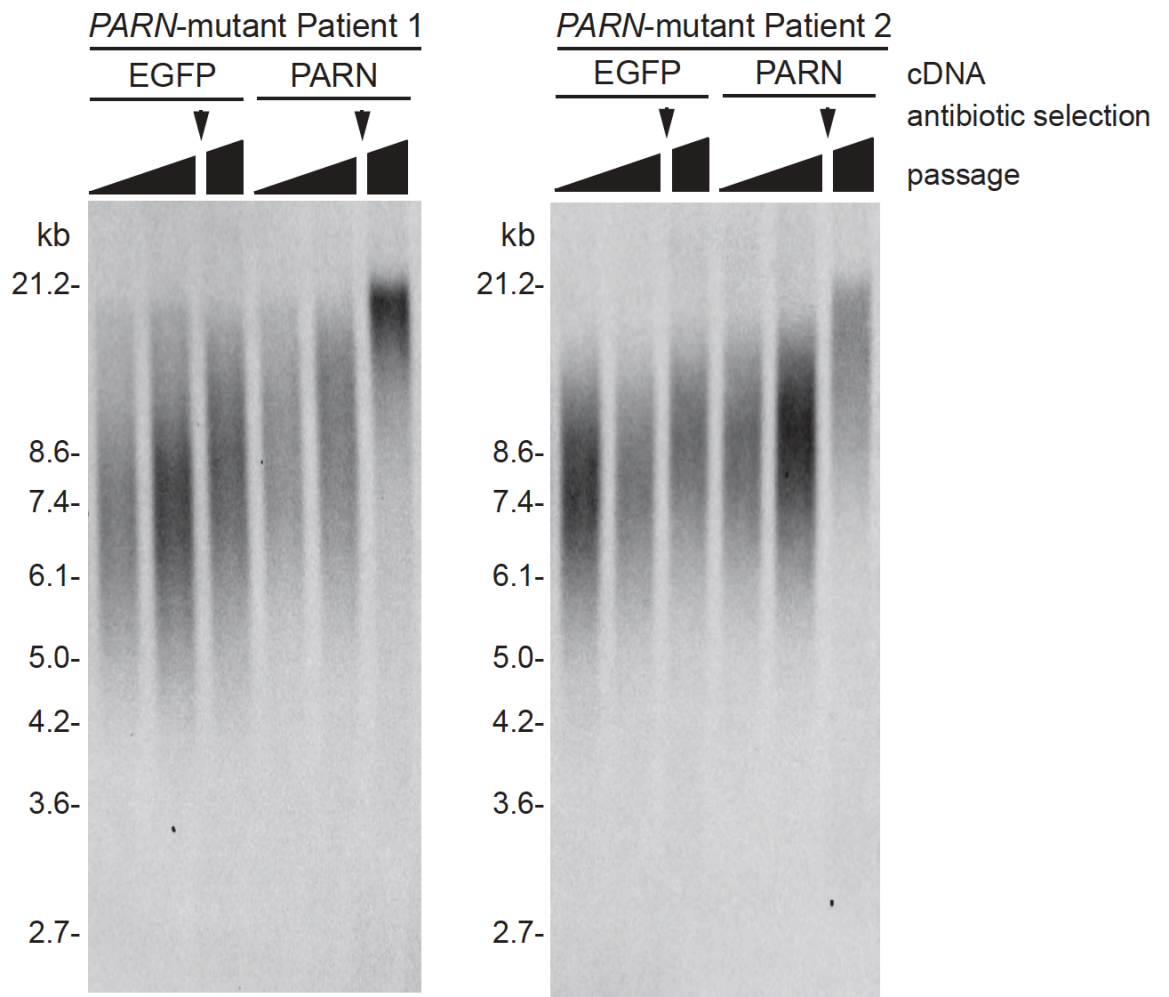
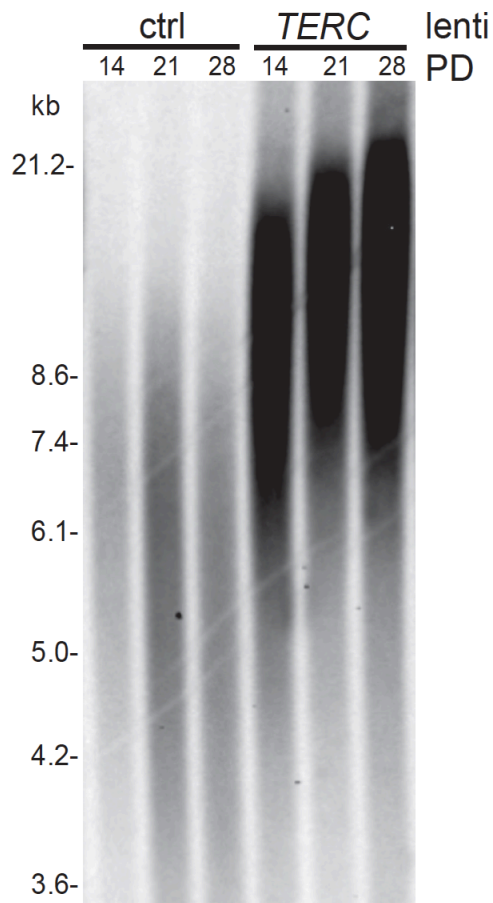


Figure 4.3.3: Telomere length elongation in immortalized *PARN*-mutant fibroblasts after *PARN* overexpression Telomere restriction fragment (TRF) length in *TERT*-immortalized, *PARN*-mutant fibroblasts. Cells were transduced with lentivirus containing *EGFP* or *PARN*, and also carrying blastocidin resistance cassette. Blastocidin selection was performed at time point indicated by the arrowhead.

We have previously suggested that PARN deficiency causes telomere disease primarily via direct effects on TERC (20). To determine whether PARN deficiency affects telomere elongation primarily via effects on TERC, we stably overexpressed *TERC* via lentiviral vectors in PARN-deficient cells and analyzed telomere length over time. We found that expression of *TERC* in *PARN*-knockdown HEK293 cells led to rapid elongation of telomeres (**Figure 4.3.4a**). Similarly, we found that overexpression of *TERC* in TERT-immortalized fibroblasts from patients with *PARN* mutations led to elongation of telomeres (**Figure 4.3.4b**). TERC has been shown to be limiting for telomere elongation in cells carrying DC-associated variants of dyskerin. These data demonstrate that PARN deficiency impairs telomere maintenance via a direct impact on TERC, and suggest that enhancing TERC will restore telomere elongation in *PARN*-mutant cells.

a. Telomere length- Southern blot
PARN-kd HEK293 cells



b. Telomere length-Southern blot
PARN-mutant fibroblasts

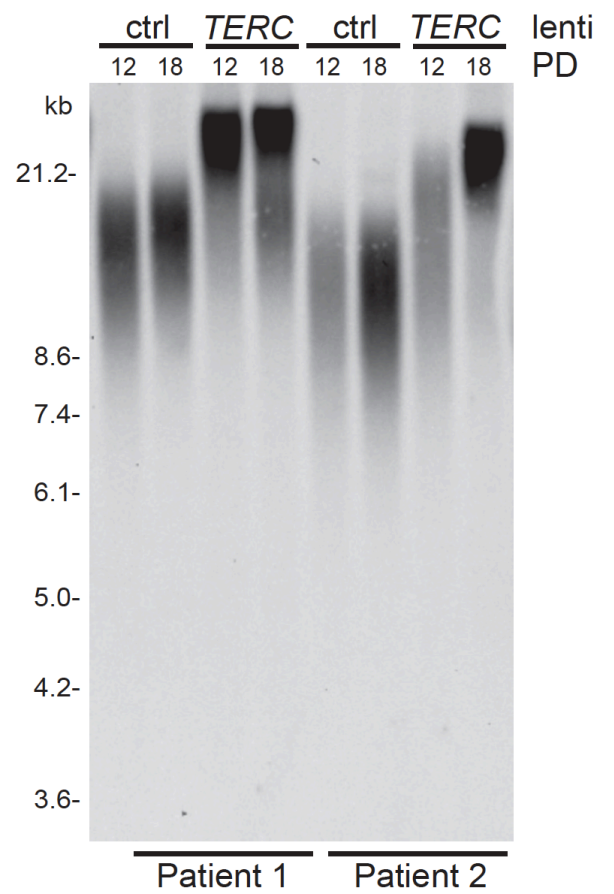


Figure 4.3.4: *TERC* rescues telomere length in *PARN* deficient cells (a) TRF in *PARN*-knockdown HEK293 cells with lentivirus encoding *TERC* versus vector (ctrl). (b) TRF of immortalized, *PARN*-mutant patient fibroblasts with lentivirus encoding *TERC* versus vector.

Prior studies indicate that PARN deadenylates non-coding RNAs that have been undergone post-transcriptional 3' adenylation by PAPD5, and thus promotes maturation rather than degradation of non-coding RNAs (15-16). We hypothesized that if the primary role of PARN in telomere biology is to counteract oligo-adenylation and degradation of TERC, then *PAPD5* inhibition should be sufficient restore TERC processing and increase telomerase activity in human cells. To investigate this, we disrupted *PAPD5* in HEK293 cells using lentivirus-encoded shRNAs and performed rapid amplification of complementary DNA ends (RACE) to characterize TERC 3' ends (**Figure 4.3.5 and Figure 4.3.6a**). With *PAPD5* knockdown, we observed an alteration in the size distribution of TERC 3' ends, in that the intensity of the amplicon corresponding to the mature form of TERC was increased relative to the immature, extended form of TERC that predominates in *PARN*-knockdown cells (**Figure 4.3.6b**).

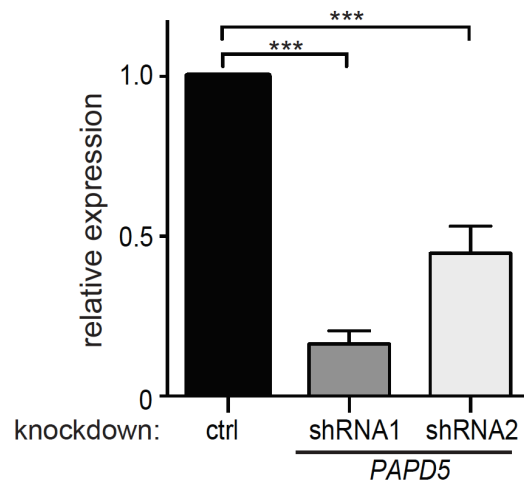
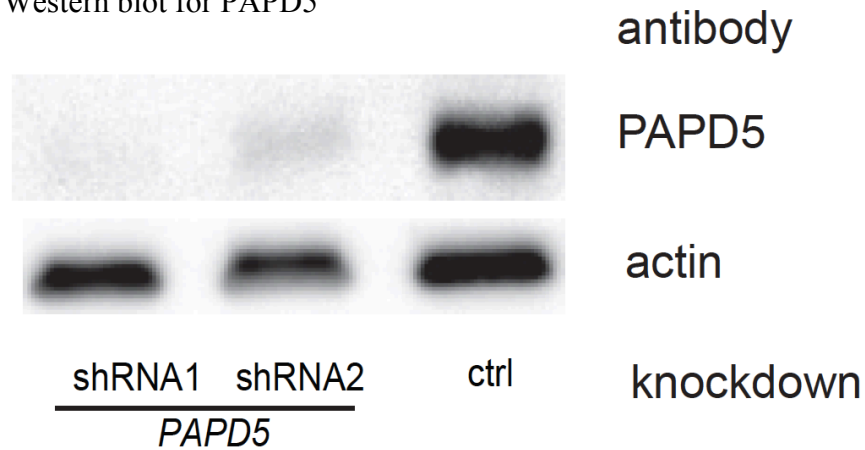
a. qPCR for *PAPD5*b. Western blot for *PAPD5*

Figure 4.3.5: *PAPD5* knockdown in HEK293 cells (a) Quantitative PCR (qPCR) of *PAPD5* transcripts from HEK 293 cells transduced with lentiviruses expressing shRNAs against luciferase (ctrl) versus *PAPD5*. $n = 3$ biological replicates. Error bar indicates standard deviation, and significance is indicated by P value ≤ 0.001 (***) (b) Immunoblot of *PAPD5* and actin protein levels in HEK293 cells transduced with lentivirus encoding shRNA against luciferase (ctrl) or *PAPD5*.

a. 3'RACE strategy

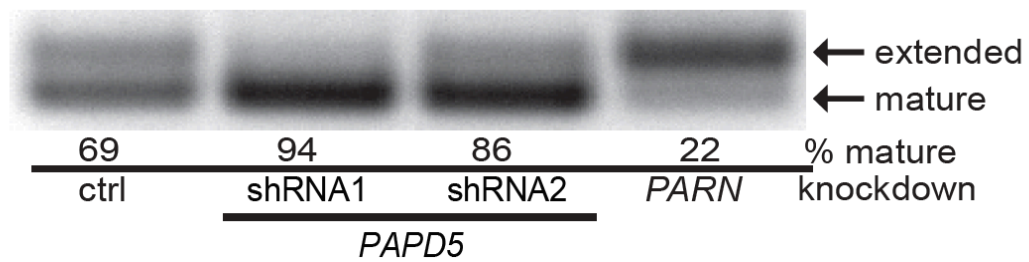
b. 3'RACE *TERC* PCR amplicons

Figure 4.3.6: *PAPD5* inhibition changes 3' of *TERC* (a) Strategy for 3' rapid amplification of cDNA ends (RACE). (b) Agarose gel electrophoresis of 3' RACE *TERC* products from HEK293 cells lentivirus encoding shRNA against *PAPD5*, *PARN* or luciferase (ctrl), n = 4.

By deep sequencing of 3' RACE products, we found that PAPD5 inhibition significantly reduced the proportion of genomically extended TERC species that were oligo-adenylated, as well as the average oligo(A) tail length of nascent TERC transcripts (**Figure 4.3.7**). Conversely, both the percentage of mature TERC (**Figure 4.3.7**) and steady-state levels of TERC transcripts were increased in *PAPD5*-knockdown cells (**Figure 4.3.8**; lanes 1 and 3). When we overexpressed *PAPD5* cDNA in cells (**Figure 4.3.9**), we found that total TERC transcript levels were decreased to approximately one-third of those found in control cells within 48 hours (**Figure 4.3.8**; lanes 1 and 2). In *PAPD5*-knockdown cells, expression of *PAPD5* cDNA (which is codon-optimized and insensitive to *PAPD5* shRNAs) also reduced TERC levels to less than those found in control cells (**Figure 4.3.8**; lanes 3 and 4). These data confirm that PAPD5 negatively regulates TERC via oligo-adenylation of nascent RNA transcripts.

To establish the relationship between PAPD5 and PARN in post-transcriptional regulation of TERC, we disrupted PAPD5 in the setting of PARN deficiency. By 3' RACE, *PARN* knockdown results in an accumulation of extended TERC forms (**Figure 4.3.10a**, lane 3 versus lane 1). We found that PAPD5 inhibition in PARN-deficient cells restored the distribution of TERC 3' ends to that observed in control cells (**Figure 4.3.10a**; compare lanes 4 and 1). By deep sequencing of 3' RACE products, we found that the proportion of mature TERC was significantly increased after *PAPD5* knockdown in PARN-deficient cells (65% versus 32%). Moreover, the increased proportion (~60%) of oligo(A) TERC species found in PARN-deficient cells was decreased by 3-fold after inhibiting PAPD5, and the average length of oligo(A) tails was also reduced. By Northern blot, we found that steady state levels of TERC transcripts were increased by PAPD5 inhibition in both normal and PARN-deficient cells (**Figure 4.3.10b**). Notably, TERC was restored in PARN-deficient cells by *PAPD5* knockdown to levels found in control cells (**Figure 4.3.10b**). To determine the effects of changes in TERC on telomerase function, we performed the telomerase repeat amplification protocol (TRAP) assay and found that *PAPD5* knockdown increased telomerase activity in PARN-deficient cells (**Figure 4.3.11**). These data indicate that PAPD5 and PARN mediate opposing and non-

redundant effects in the post-transcriptional maturation of nascent TERC transcripts. Despite the presence of other poly(A) polymerases and deadenylases in cells, the manipulation of these two factors significantly impacts TERC levels and cellular telomerase activity.

a. Deep sequencing of *TERC* RNA ends
TERC gene:

5'...ATG**C**AGTTCGCT...3'

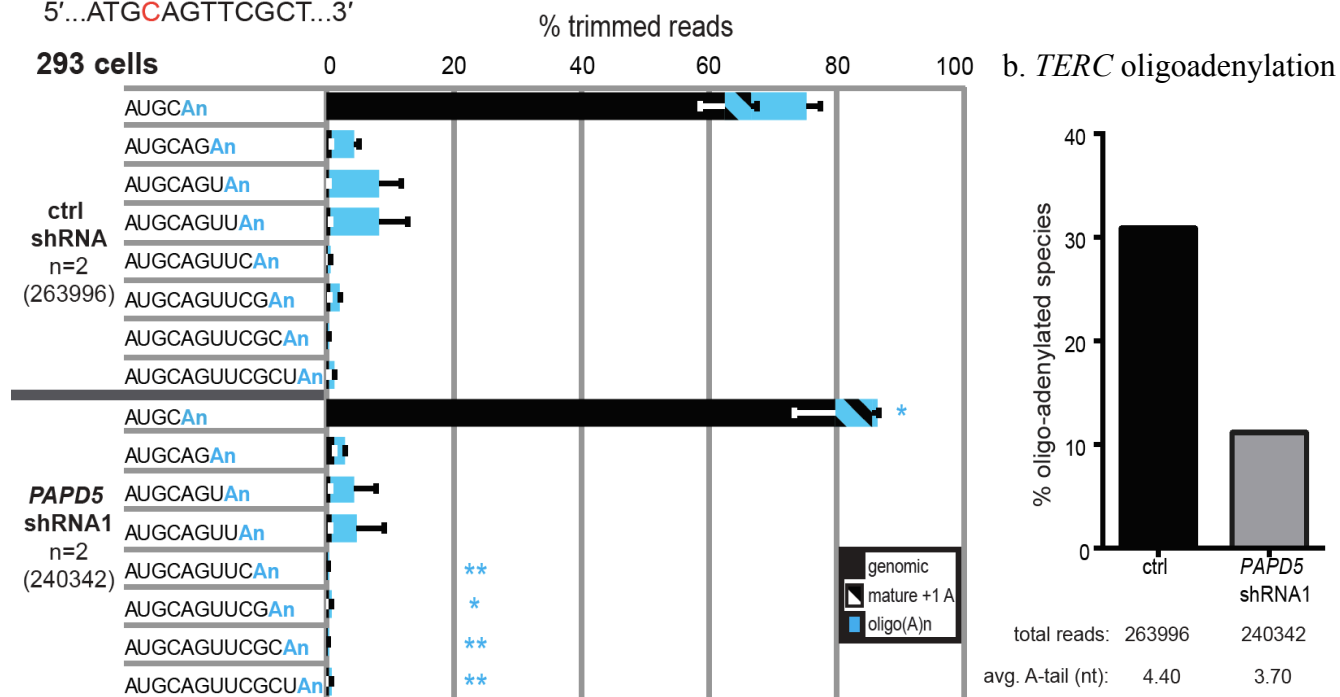


Figure 4.3.7: *PAPD5* inhibition decreases *TERC* oligoadenylation (a) RACE products were subjected to deep sequencing and aligned to the *TERC* gene. Canonical *TERC* terminus is shown in red. Genomically-encoded termini are in black, mature *TERC* with a single A (which may or may not be genomically-encoded) is hatched, and oligo(A) additions of any length (n) are in solid blue. Total trimmed reads in parentheses. n= 2. *P* values: * ≤ 0.05 ; ** ≤ 0.01 . (b) Oligo(A)_n species as a proportion of total reads in control versus *PAPD5* knockdown cells. The average A-tail length in nucleotides (nt) is indicated.

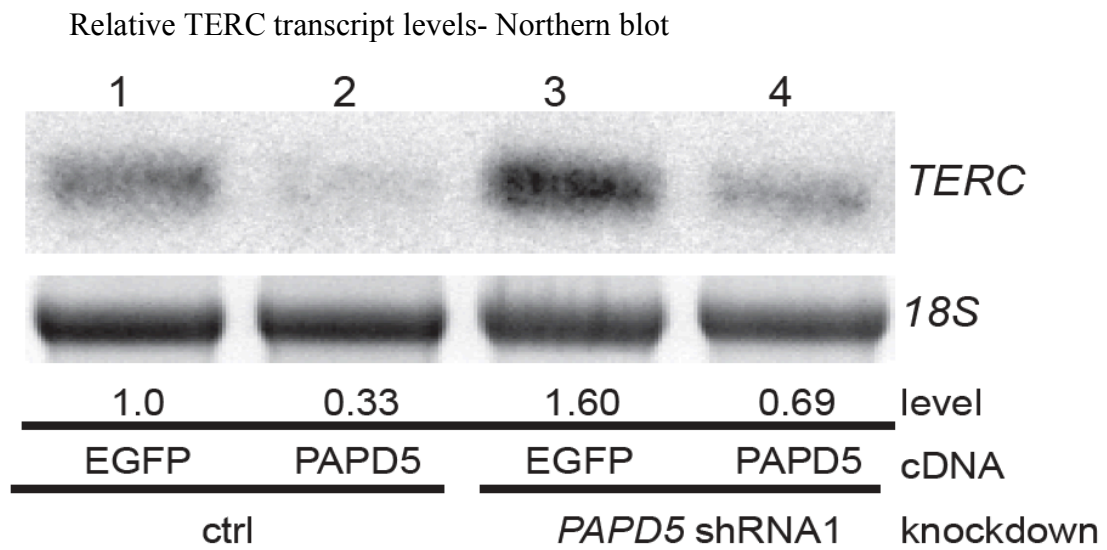


Figure 4.3.8: Manipulation of *PAPD5* changes TERC levels Northern blot of TERC from HEK293 cells with shRNA against *PAPD5* versus luciferase, followed by transfection with plasmids expressing *PAPD5* versus *EGFP* cDNAs. Loading control: ethidium bromide staining of 18S rRNA. n = 3.

Western blot for PAPD5 overexpression

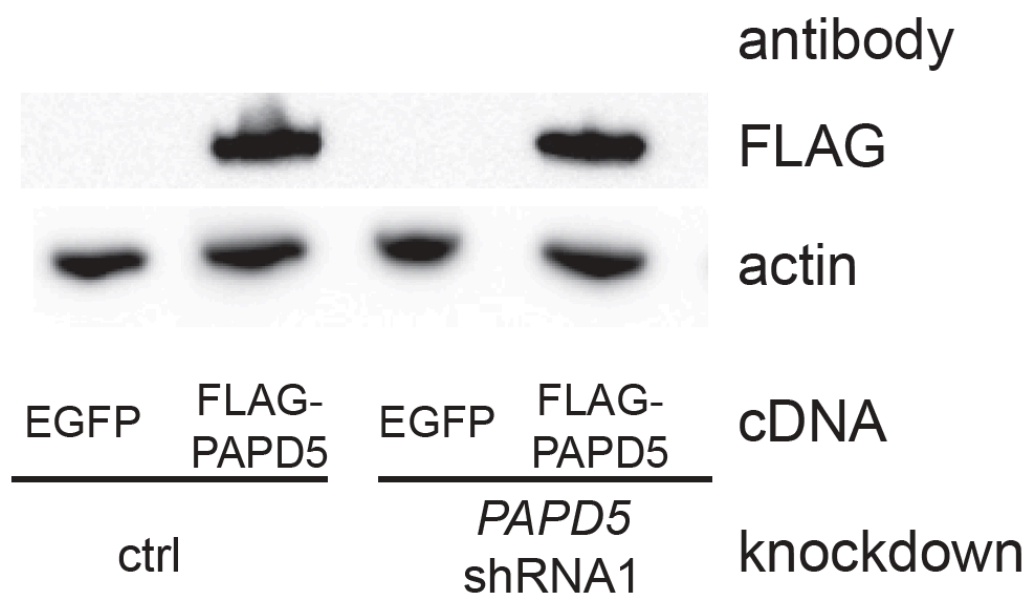


Figure 4.3.9: Overexpression of PAPD5 in control and PAPD5 knockdown cells

Western blots of cell lysates from control knockdown versus PAPD5 knockdown HEK 293 cells, with lentiviral overexpression of either a cDNA encoding FLAG-tagged, codon-optimized PAPD5 or EGFP. Actin is shown as a loading control. Cells are those used in Figure 4.3.5.

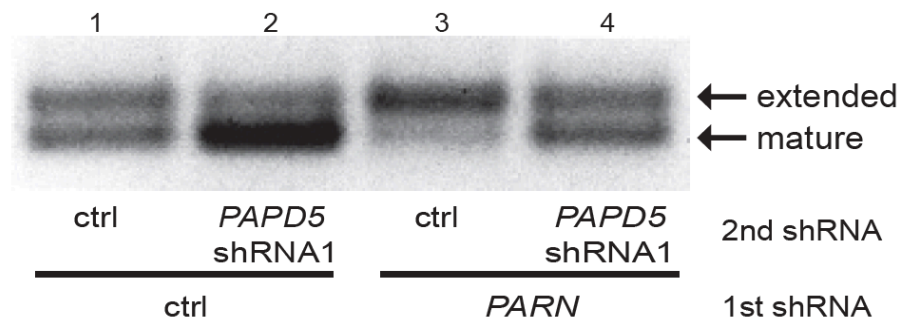
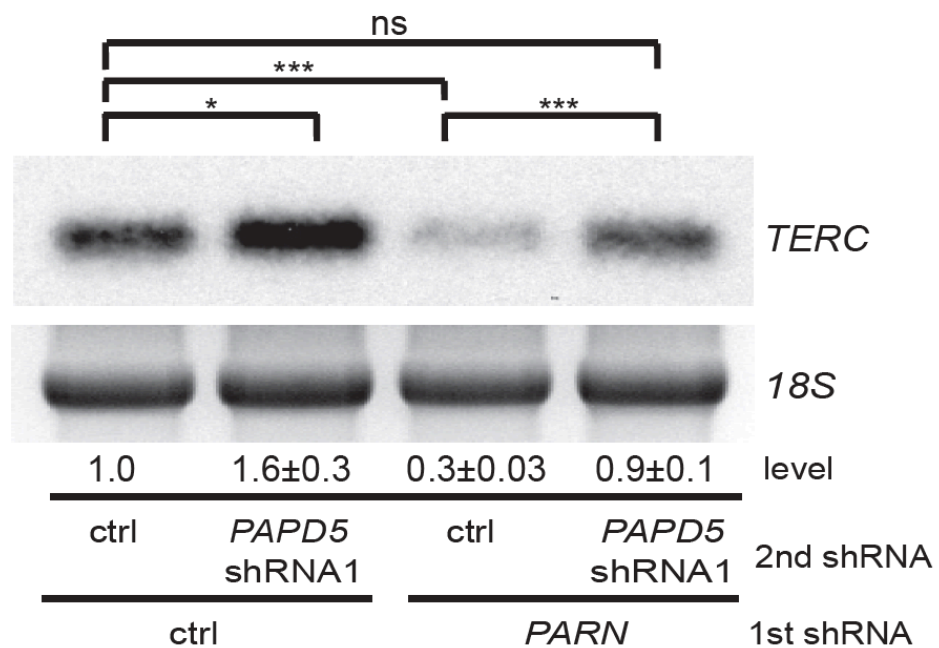
a. 3'RACE *TERC* PCR ampliconsb. Relative *TERC* transcript levels- Northern blot

Figure 4.3.10: *PAPD5* inhibition restores *TERC* in *PARN* deficient cells. (a) RACE amplicons from HEK293 cells with shRNA directed against *PARN* versus luciferase, followed by shRNA against *PAPD5* or luciferase, n = 3. (b) Northern blot of *TERC* from HEK293 cells with shRNA against *PARN* versus luciferase, followed by shRNA against *PAPD5* versus luciferase. n = 3.

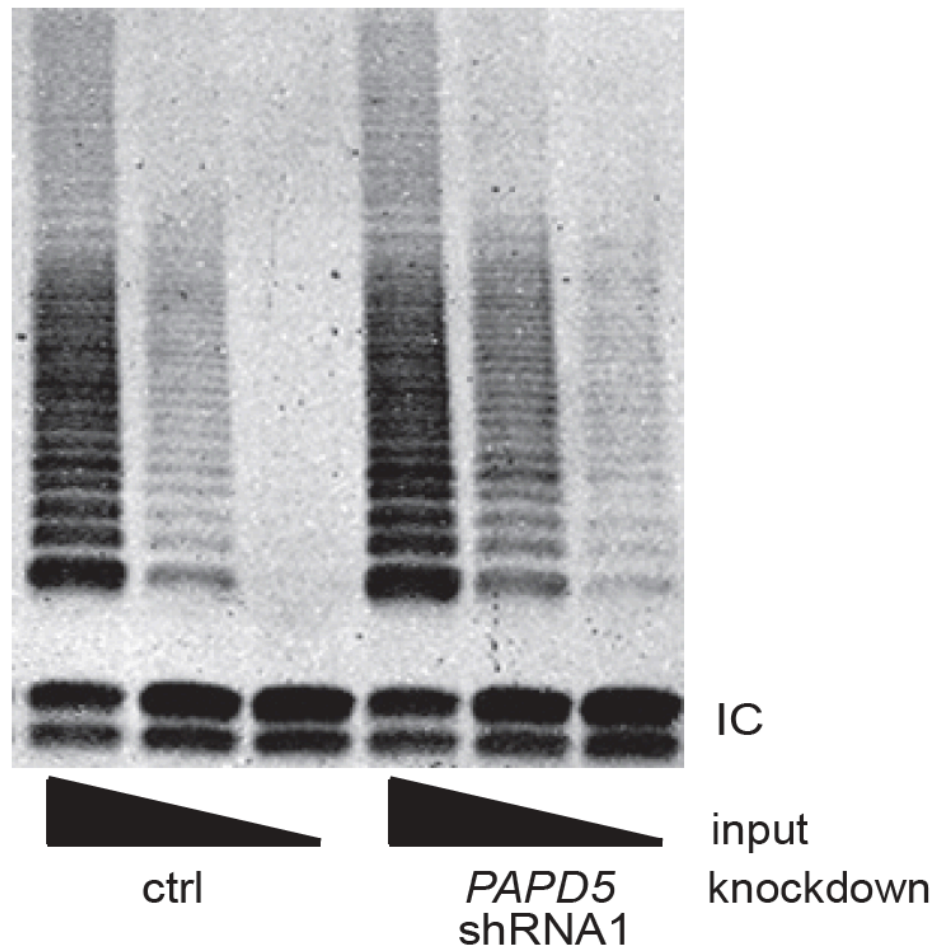
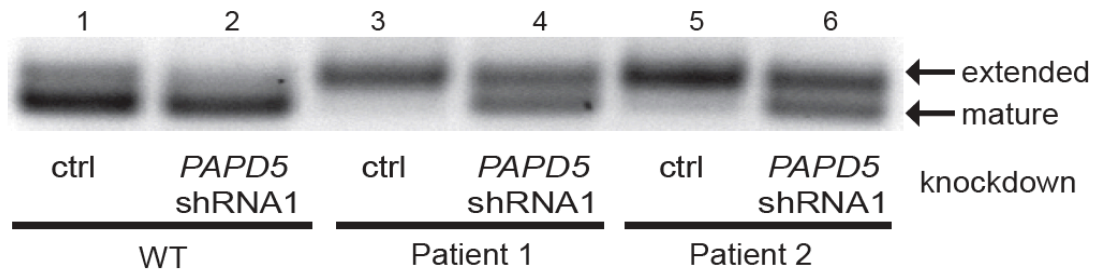
Telomerase activity- *PARN* kd cells

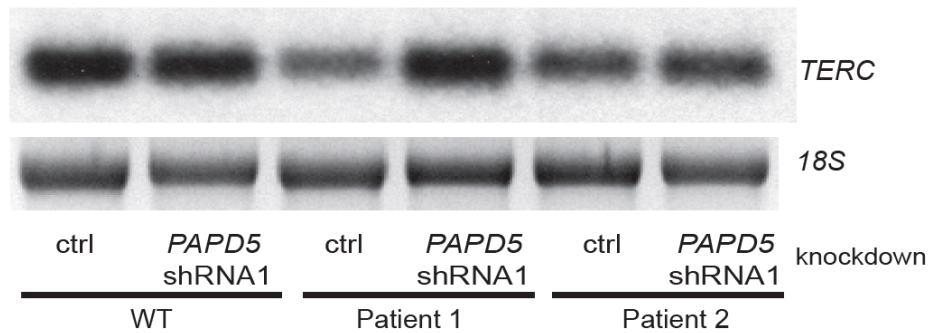
Figure 4.3.11: *PAPD5* inhibition restores telomerase activity in PARN deficient cells. Telomerase activity (TRAP) in PARN-deficient HEK293 cells after control versus *PAPD5* knockdown, using 3-fold dilutions of input cell extract. IC: internal control amplification standard. n = 3.

Our results indicate that TERC is limiting in *PARN*-deficient cells due to inadequate deadenylation of nascent TERC transcripts, and that this effect can be reversed by inhibiting the major adenylation activity, PAPD5. To determine whether inhibiting post-transcriptional adenylation by PAPD5 could reverse molecular features of telomere disease in *PARN*-mutant DC patient cells, we transduced induced pluripotent stem cells (iPSCs) from two DC patients harboring pathogenic *PARN* mutations with lentiviral vectors encoding *PAPD5* shRNA. By 3' RACE, *PARN*-mutant patient cells show a predominance of extended TERC forms (**Figure 4.3.12a**, lanes 1, 3, and 5). We found that this altered distribution of TERC species in *PARN*-mutant patient cells was normalized by *PAPD5* knockdown, as evidenced by an increase in the proportion of amplicons corresponding to the mature form of TERC (**Figure 4.3.12a**). Accordingly, by Northern blot, we found that PAPD5 inhibition rescued total TERC levels in patient iPSCs (**Figure 4.3.12b-c**). The increased TERC steady state levels were due to higher stability of RNA transcripts, which we demonstrated by inhibiting Pol II transcription using actinomycin D in patient iPSCs with or without PAPD5 inhibition. When we compared decay rates by Northern blot, we found that *PAPD5* knockdown increased the half-life of TERC transcripts in *PARN*-mutant patient iPSCs compared to cells expressing control shRNA (**Figure 4.3.13**). Collectively, these results indicate that PAPD5 inhibition is sufficient to restore TERC processing, stability and steady-state levels in the setting of *PARN* mutations.

a. 3'RACE TERC PCR amplicons- iPSCs



b. TERC levels in iPSCs- Northern blot



c. Relative TERC levels in iPSCs

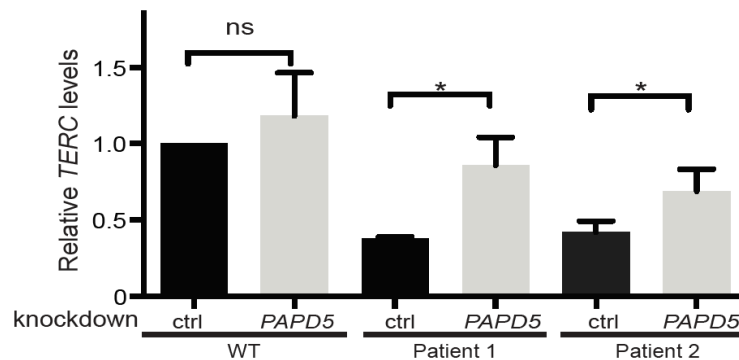
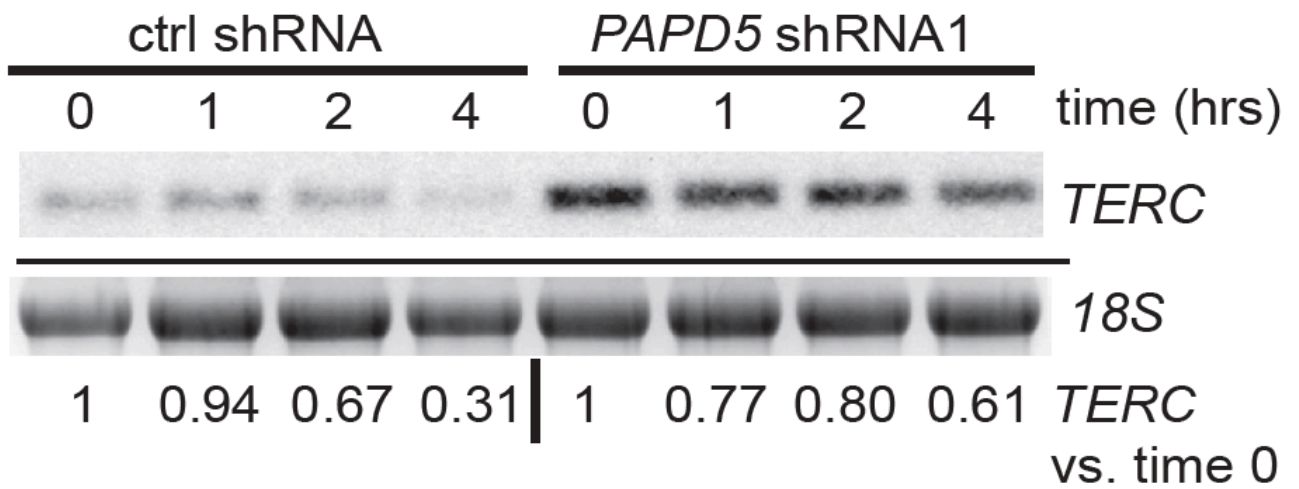


Figure 4.3.12: *PAPD5* inhibition rescues TERC maturation and steady state levels in *PARN*-mutant patient cells. (a) 3' RACE TERC amplicons from normal (WT) versus *PARN*-mutant iPSCs after lentiviral shRNA directed against *PAPD5* versus luciferase (ctrl). (b) Representative Northern blot of TERC RNA from WT versus *PARN*-mutant iPSCs after lentiviral shRNA directed against *PAPD5* versus luciferase. (c) Relative TERC levels from Northern blots (b) are quantified relative to WT cells with control knockdown, n = 3.

a. Actinomycin treatment of *PARN*-mutant iPSCs

b. TERC decay rate in iPSCs

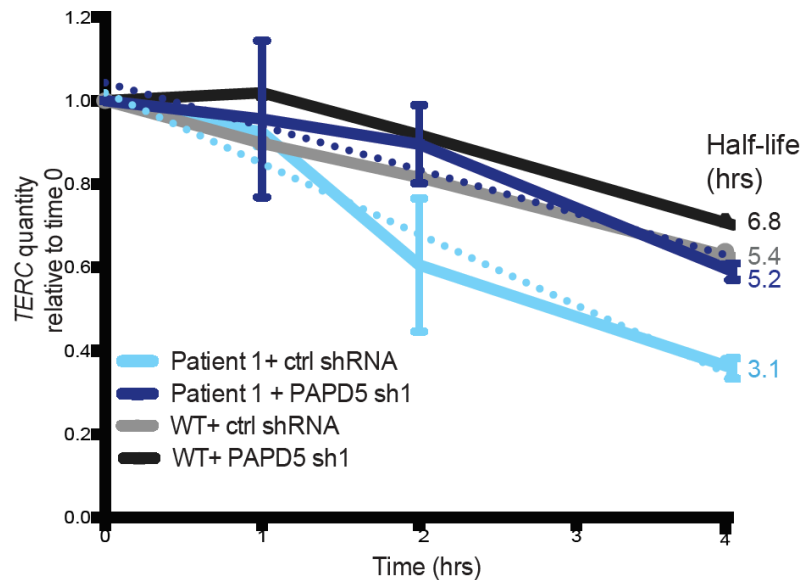


Figure 4.3.13: *PAPD5* inhibition decreases TERC decay rate in *PARN*-mutant patient cells. (a) Representative Northern blot of TERC levels in WT and *PARN*-mutant iPSCs after lentiviral shRNA directed against *PAPD5* versus luciferase, at 0-4 hours following actinomycin treatment. TERC levels are normalized relative to time 0. (b) Graph of TERC decay rate. Dashed line reflects slope determined by simple linear regression.

We next asked whether PAPD5 inhibition would be sufficient to rescue telomerase activity and telomere elongation. Using TRAP assays, we found that *PAPD5* inhibition in *PARN*-mutant patient iPSCs increased cellular telomerase activity (**Figure 4.3.14**). Remarkably, when we assayed telomere length by Southern blot, we found that *PAPD5* knockdown in *PARN*-mutant patient iPSCs led to a rapid increase in telomere lengths, approximating those found in control iPSCs (**Figure 4.3.15-16**). Patient iPSCs with stable *PAPD5* knockdown showed continuous self-renewal in culture. We next performed a global analysis of transcriptional changes after *PAPD5* knockdown by RNA sequencing. In patient iPSCs, we found only a small number of transcripts (11 genes) whose levels were altered to a degree that exceeded the increase in *TERC*; of those, only non-coding RNAs were commonly affected more than would be predicted by chance (**Tables 4.3.1-3**). Taken together our data indicate that PAPD5 inhibition restores *TERC* levels, telomerase activity and telomere length in cells from patients with *PARN* mutations.

Telomerase activity in iPSCs

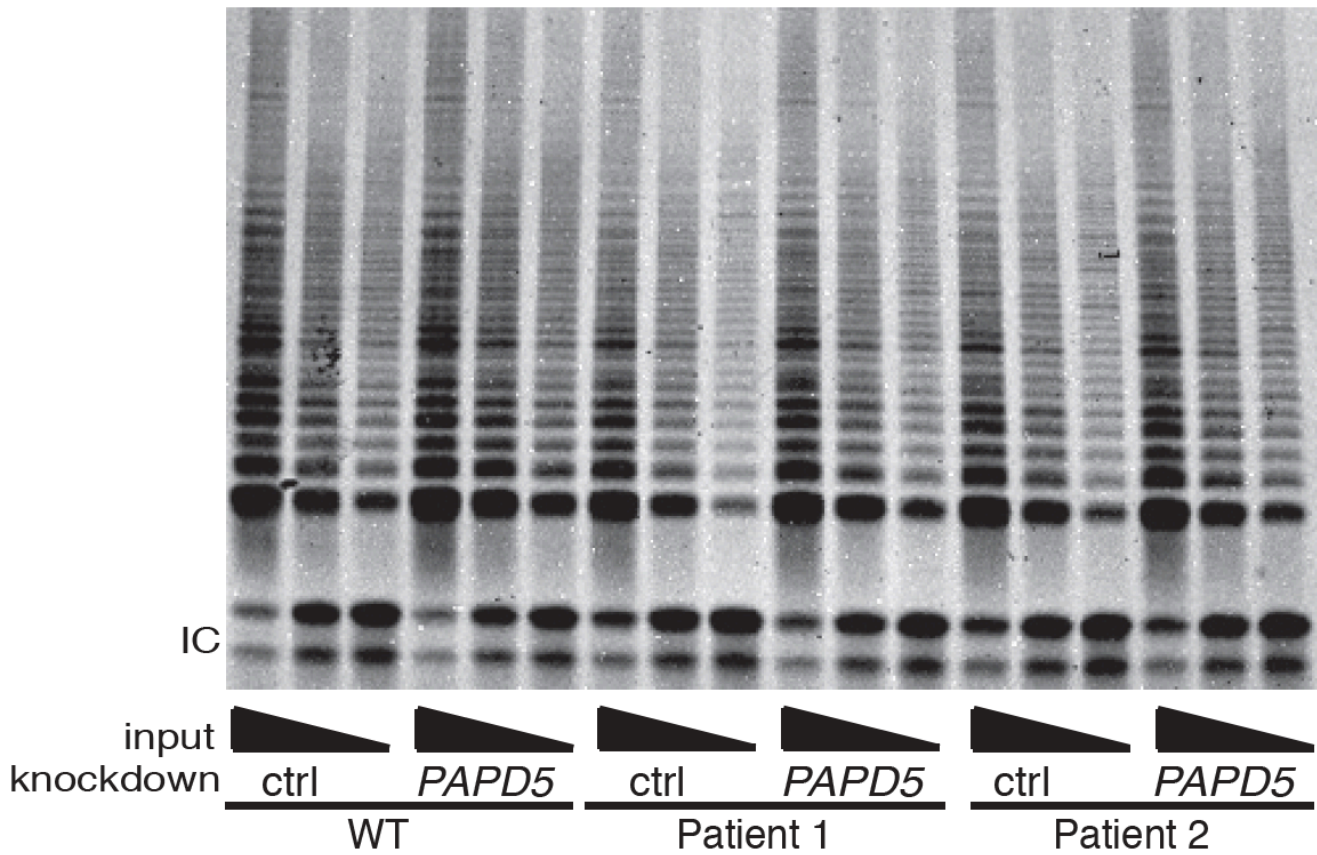


Figure 4.3.14 *PAPD5* inhibition rescues telomerase activity and telomere length in *PARN*-mutant patient cells TRAP assay in iPSCs after lentiviral shRNA against *PAPD5* versus luciferase, n = 3.

Telomere length in iPSCs-Southern blot

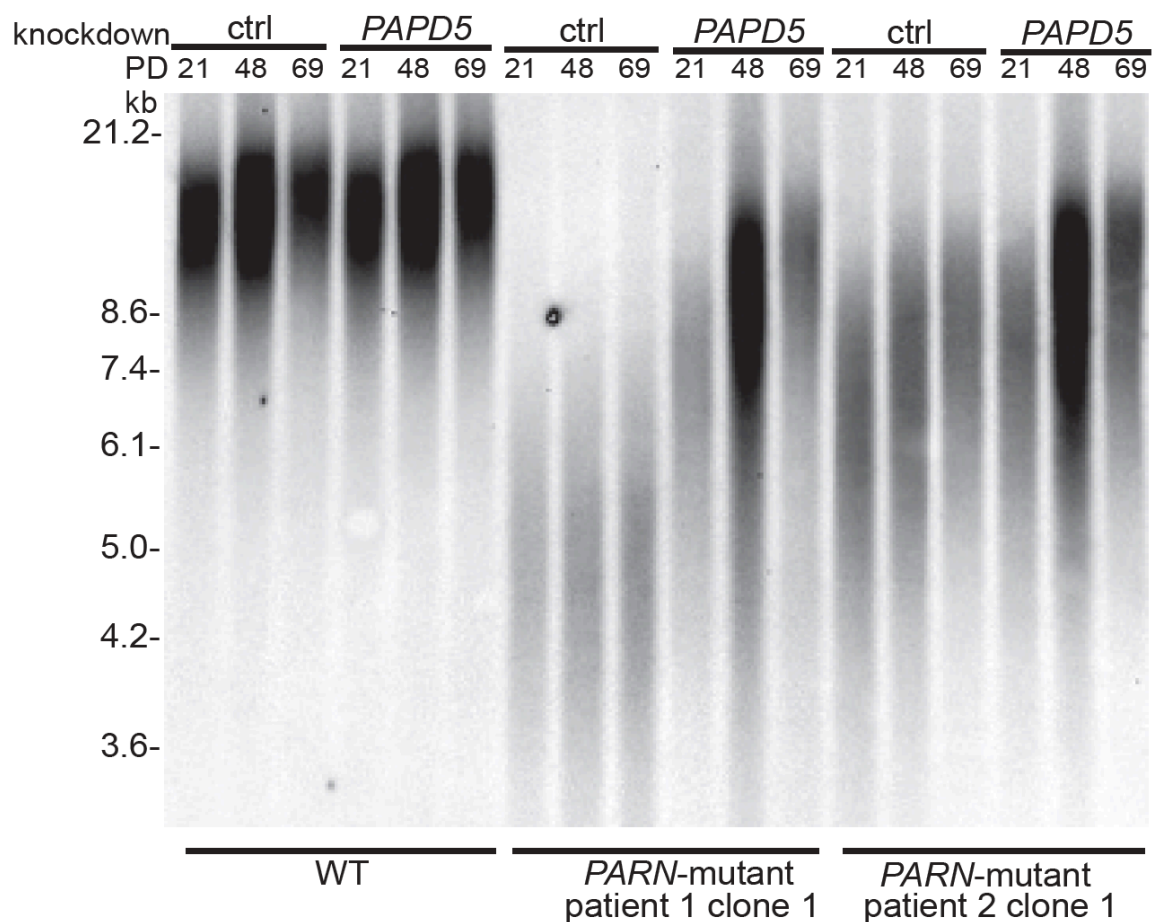


Figure 4.3.15 *PAPD5* inhibition rescues telomere length in *PARN*-mutant patient cells TRF of WT and *PARN*-mutant patient iPSCs after lentiviral shRNA against *PAPD5* versus luciferase, over time in culture.

Telomere length in iPSCs- Southern blot

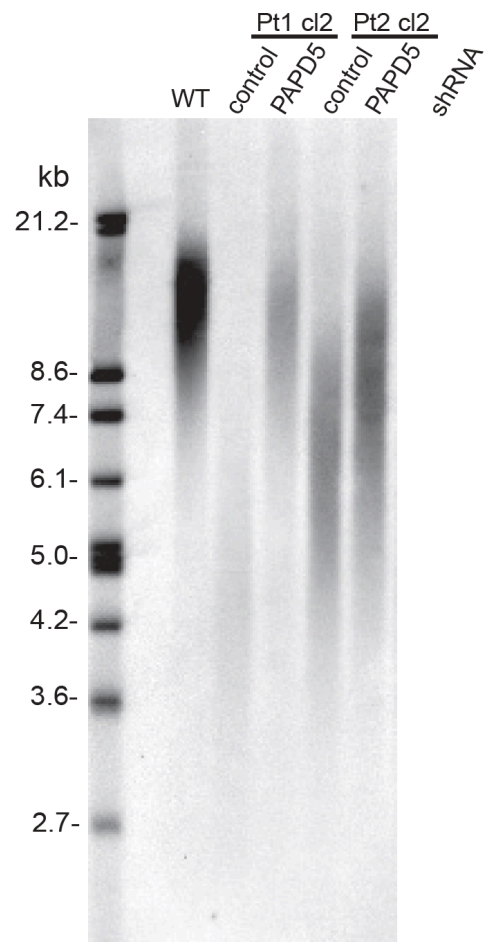


Figure 4.3.16 *PAPD5* inhibition rescues telomere length in *PARN*-mutant patient cells, other clones TRF of WT and *PARN*-mutant patient iPSCs after lentiviral shRNA against *PAPD5* versus luciferase, over time in culture.

Table 4.3.1: RNA-Seq transcriptome analysis of the effects of PAPD5 knockdown: transcripts altered in excess of the fold increase in TERC RNA

		Altered in any 2 of 3 pair-wise comparisons: HEK 293 and <i>PARN</i> -mutant Patient 1 iPSC clones 1 and 2			Altered in both pair-wise <i>PARN</i> -mutant comparisons: Patient 1 iPSC clones 1 and 2		
	Total genes analyzed	Obs	Exp	P-value	Obs	Exp	P-value
Decreased							
mRNA	19698	30	33.7	0.34	4	3.97	1
lincRNA	7670	7	13.1	0.07	1	1.55	0.97
snoRNA	454	7	0.78	5.9E-11	0	0.09	1
snRNA	1848	7	3.2	0.05	1	0.37	0.83
Increased							
mRNA	19698	47	53.6	0.15	3	3.31	1
lincRNA	7670	12	20.9	0.03	0	1.29	0.42
snoRNA	454	11	1.2	4.0E-17	2	0.08	2.0E-7
snRNA	1848	11	5.0	0.01	0	0.31	1

mRNA: protein-coding mRNA; lincRNA: long intergenic non-coding RNA; snoRNA: small nucleolar RNA (includes box H/ACA, C/D, scaRNA); snRNA: small nuclear RNA

Obs: observed; Exp: expected by chance

Bold: significantly more than expected by chance

Italic: significant fewer than expected by chance

Table 4.3.2: Genes whose transcript levels are altered in excess of the fold-change in TERC levels, in pair-wise comparisons of PAPD5 knockdown versus control cells

	Altered in all 3 of 3 pair-wise comparisons: HEK 293 and <i>PARN</i> -mutant Patient 1 iPSC clones 1 and 2	Altered in both pair-wise <i>PARN</i> -mutant comparisons: Patient 1 iPSC clones 1 and 2
Decreased	<i>none</i>	CHURC1-FNTB
		DEC1
		GARS
		RNU6-39P
		RP4-533D7.6
		RP4-794H19.1
Increased	SNORD100	AC018867.1
		GPAA1
		SNORA73
		SNORD100
		USP7

Table 4.3.3: Genes whose transcript levels are altered in excess of the fold-change in TERC levels, in any 2 of 3 pair-wise comparisons of PAPD5 knockdown versus control: HEK 293 cells and PARN-mutant Patient 1 clones 1 and 2

Decreased	Increased
AARSD1	AC004076.7
AC069499.1	AC018867.1
AL592183.1	AC112719.2
ANP32B	ANKMY1
AP001628.6	AP001059.7
BASP1	BPI
CD52	BRPF3
CHURC1-FNTB	C12orf79
CTC-273B12.7	C16orf59
DDX56	C16orf93
DEC1	C18orf8
DXO	C1orf127
FKSG63	C5orf22
G0S2	CGRRF1
GARS	CH507-513H4.5
HLA-DPA1	COMMD3-BMI1
MILR1	CTC-435M10.3
MTRR	CTD-2302E22.2
MVP	CTD-2313N18.7
NDUFS7	CYBA
NPIPA1	DNASE1L1
OFCC1	DPP7
POLR3GL	EIF4ENIF1
PRR4	FXVD1
PSMC3IP	GABARAP
RNU6-24P	GEMIN4
RNU6-307P	GPAA1
RNU6-314P	LINC01360
RNU6-341P	NDUFA7
RNU6-39P	NOMO3
RNU6-48P	PACRG
RNU6-606P	PIAS3
RP11-101E3.5	PMM2
RP11-2B6.3	PRKAB1
RP11-397H6.1	PTCD3
RP11-686D22.5	PVR
RP11-70F11.11	RFXANK
RP11-895M11.3	RMRP
RP4-533D7.6	RNU4-86P
RP4-794H19.1	RNU6-1099P
SCARNA1	RNU6-1275P
SNORA21	RNU6-373P
SNORA31	RNU6-45P
SNORA55	RNU6-714P
SNORA68	RNU7-140P
SNORD11	RNU7-154P
snoU13	RNU7-181P
SPRR2G	RNU7-77P
TARS	RNVU1-15
TMEM141	RP11-160E2.6
U2AF1L4	RP11-171I2.3
	RP11-245A18.1
	RP11-338N10.3
	RP11-38C17.1
	RP11-430H10.1
	RP11-577B7.1
	RP11-644F5.10
	RP3-526F5.2
	RP4-539M6.19
	RP6-24A23.6
	RPAP1
	SCARNA12
	SIGIRR
	SLC25A16
	SLC6A6
	SLX1B
	SMTNL2
	SNORA13
	SNORA5C
	SNORA73
	SNORD100
	SNORD116-15
	SNORD33
	SNORD38A
	SNORD46
	SNORD89
	SNORD92
	SULT1A3
	TRMT1
	UBE2F
	USP7

5. DISCUSSION

It is well established that the level of TERC in cells is important in health and disease – in tissue renewal, degenerative disorders, cancer, and possibly longevity – but mechanisms by which TERC levels are regulated have remained elusive. Our findings, emerging from genetic discovery in patients with telomere diseases (17-18), demonstrate that PARN mediates post-transcriptional maturation of TERC (20). We propose that PARN removes oligo(A) tails and/or genomically-encoded terminal nucleotides in nascent TERC transcripts, leading to maturation of the 3' end and protection from further PAPD5 mediated oligo-adenylation and degradation by the exosome (20-21) (**Figure 5.1**). We cannot exclude the possibility that PARN mutations impact the levels of other RNAs to contribute to disease manifestations. However, given the phenotypic identity between patients with PARN mutations and other forms of TERC deficiency, our results suggest that disrupting PARN's role in TERC biogenesis is sufficient to cause dyskeratosis congenita / telomere diseases. Based on our findings and those of others investigating the global effects of PARN disruption, we speculate that a major, non-redundant role of PARN in mammalian cells is not in mRNA metabolism, but in regulating the biogenesis of TERC and other non-coding RNAs.

We showed that PARN functions in the terminal stages of 3' deadenylation and end trimming for productive maturation of nascent TERC transcripts, suggesting that low TERC levels are sufficient to explain telomere disease in patients with *PARN* mutations. Consistent with this hypothesis, we show here that TERC expression can restore telomere length in PARN-deficient cell lines and *PARN*-mutant patient cells. Similar findings are seen in cells from patients with mutations in *TERC* itself or in *DKC1*. These results support the hypothesis that TERC deficiency underlies defective telomere maintenance in patients with *PARN* mutations. Given our findings of increased oligo-adenylation and destabilization of TERC RNA species in the setting of *PARN* mutations, we focused on reversing this post-transcriptional modification as a means of normalizing TERC levels.

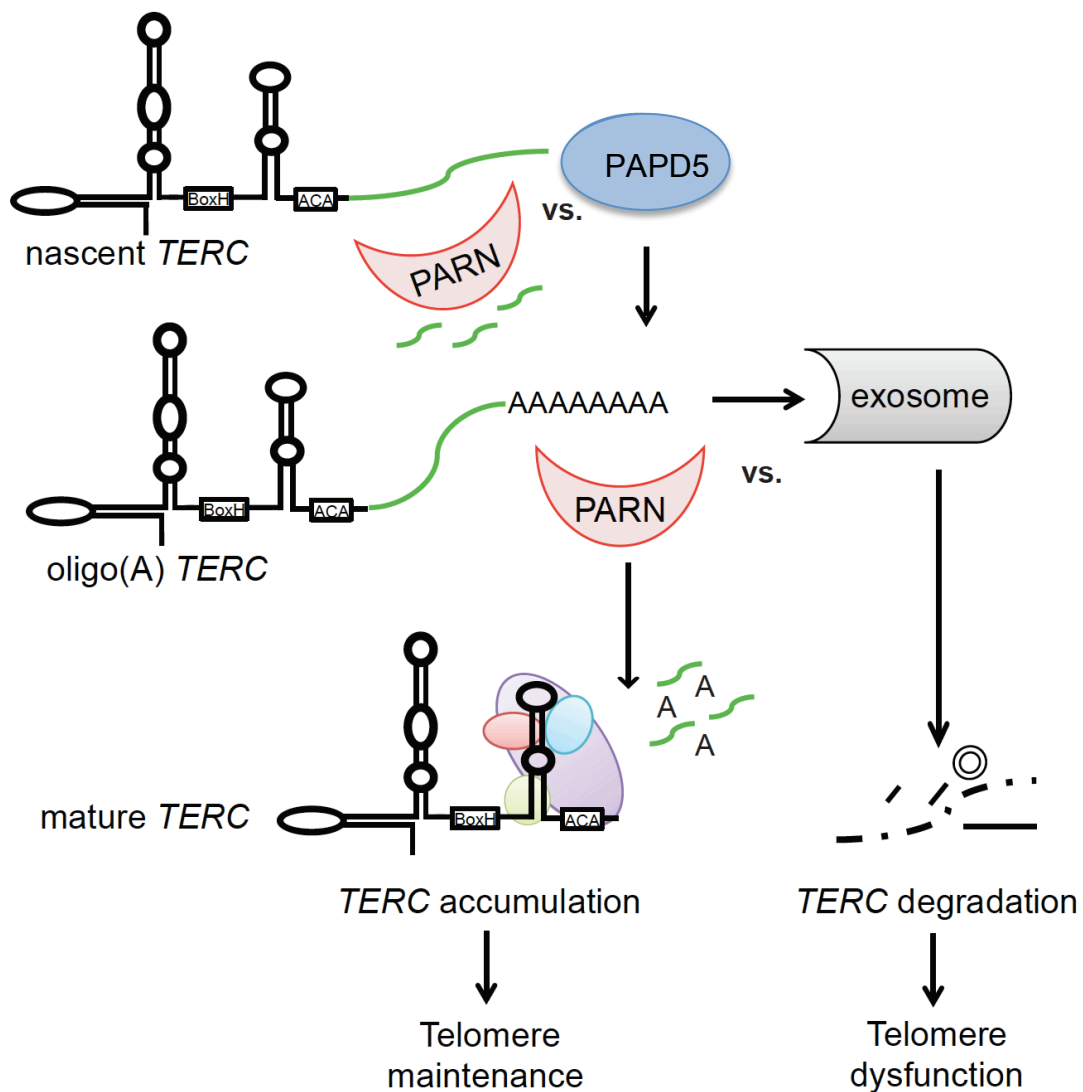


Figure 5.1: Model for TERC 3' end maturation by PARN We propose the following model: 3' ends of nascent TERC RNA are subject to TRF4-2 (PAPD5) / TRAMP-mediated oligo-adenylation, which targets transcripts for degradation by the exosome. PARN counteracts the degradation pathway by removing oligo(A) tails and/or trimming genomically-encoded bases (green) of nascent TERC to yield a mature 3' end. PARN deficiency tips the balance in favor of degradation, leading to reduced TERC levels and telomere dysfunction. We note that the order and timing of nascent TERC folding, 3' end processing, and assembly with the dyskerin complex /telomerase reverse transcriptase remains largely unknown.

(From Moon et al., *Nat Genet*, 2015)

Our results indicate a critical role for PAPD5 in regulating human TERC biogenesis, a discovery that stems from the findings of *PARN* mutations in patients with telomere diseases (21). We found that PAPD5 oligo-adenylates and destabilizes nascent TERC transcripts in patient cells, corroborating and extending recent findings in cell lines. From modulating PARN and PAPD5 individually or in combination, we find that their functions in TERC maturation are largely non-redundant, despite the presence of other poly(A) polymerases and deadenylases in the cell. Our data further show that PAPD5 inhibition increases TERC to levels sufficient to augment telomerase activity and telomere length in *PARN*-mutant patient cells, but without altering the majority of other RNAs across the transcriptome to a similar degree, and without compromising cellular viability. We propose a model where the balance of PARN and PAPD5 activities serves to establish TERC steady state levels, which in turn plays a major role in determining telomerase activity and telomere maintenance in TERT-expressing cells, such as stem cells and cancer cells (**Figure 5.1**).

Our results provide a novel example of how enzymes that regulate non-coding RNA processing might be targeted to reverse disease phenotypes. Although non-coding RNAs play a variety of roles in diseases, prospective therapies generally depend on delivery of the RNA or anti-sense strategies. We and others have shown that ectopically expressing TERC is sufficient to extend telomere length in cells from patients with various forms of DC, but that approach poses significant challenges for clinical translation. The results presented here identify PAPD5 as a component of the post-transcriptional machinery potentially amenable to pharmacological inhibition, to restore telomere maintenance in patients with DC and IPF caused by *PARN* mutations. More broadly, our results suggest an unanticipated therapeutic window for PARN and PAPD5 manipulation that might be exploited to alter telomerase activity and self-renewal in a range of degenerative and malignant diseases (**Figure 5.2**).

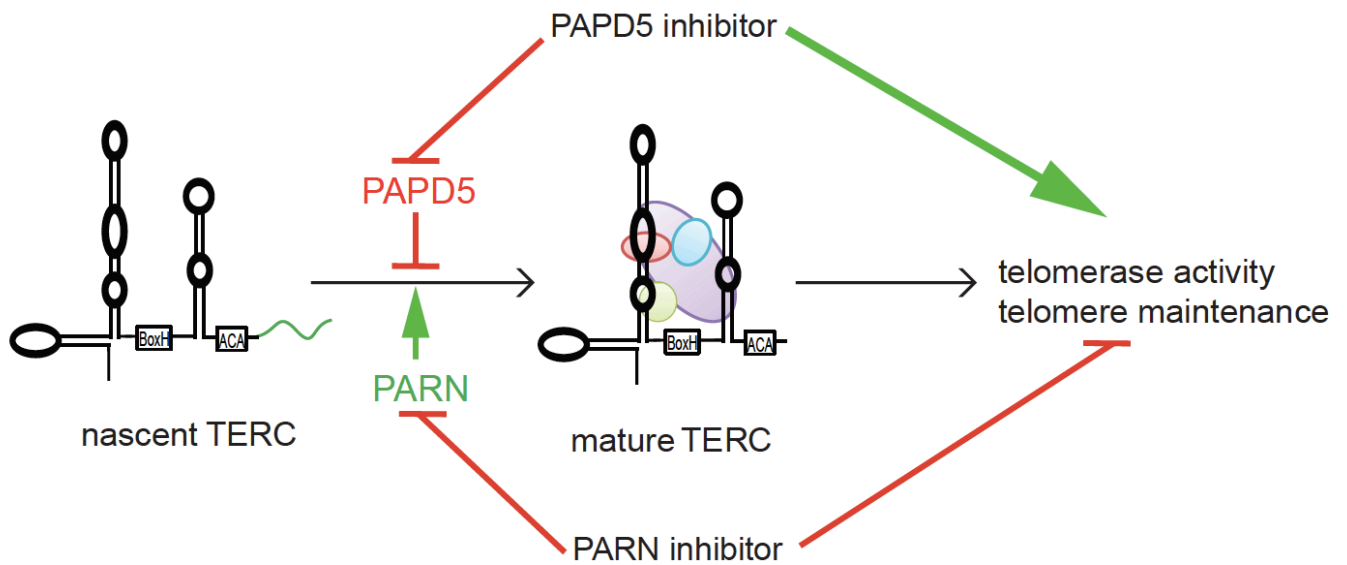


Figure 5.2: Model - reciprocal regulation of TERC maturation by PARN and PAPD5 Model demonstrating the reciprocal regulation of TERC levels by PAPD5 and PARN, and the potential for therapeutic manipulation of telomerase in degenerative or malignant disorders.
(From Boyraz et al., *JCI*, 2016)

In the era of clinical genome sequencing, determining the pathogenicity of variants of unknown significance or understanding the function of a new gene is an important challenge, often with immediate implications for patient diagnosis and counseling. In the first study, patient iPSCs allowed us to directly compare the binding of normal versus variant TERC to endogenous TERT using a biochemical approach (22). Our results reveal a sequence in TERC that is important for binding TERT, which when disrupted in a heterozygous state in humans is sufficient to cause telomere disease. Validating the TERC G319A mutation in patient iPSCs allows us to proceed with genetic testing and counseling of older and younger family members at risk.

In summary, our results illustrate the utility of patient-derived iPSCs for evaluating genetic variants and defining molecular mechanisms of disease in patients with bone marrow failure syndromes.

6. SUMMARY AND FUTURE DIRECTIONS

1. Disruption of TERC:TERT binding in a heterozygous state is sufficient to cause telomere disease.
2. iPSCs can be used as tools to perform extensive biochemical and molecular analyses to determine the pathogenicity of variants of unknown significance.
3. PARN mediates the post-transcriptional maturation of *TERC* via 3'-end processing.
4. *TERC* levels are limiting for telomere maintenance in PARN deficiency.
5. PAPD5 negatively regulates *TERC* via oligo-adenylation, a discovery that stems from the findings of patients with *PARN* mutations.
6. *PAPD5* knockdown reverses the molecular hallmarks of telomere disease in *PARN*-mutant patient cells.
7. PARN and PAPD5 manipulation can be used to alter telomerase activity and self-renewal in a range of degenerative and malignant diseases.

REFERENCES

1. Egan ED, and Collins K. An enhanced H/ACA RNP assembly mechanism for human telomerase RNA. *Mol Cell Biol.* 2012; 32(13):2428-39.
2. Feng J, et al. The RNA component of human telomerase. *Science.* 1995; 269 (5228):1236-41.
3. Venteicher AS, et al. A human telomerase holoenzyme protein required for Cajal body localization and telomere synthesis. *Science.* 2009; 323 (5914):644-8.
4. Cristofari G, and Lingner J. Telomere length homeostasis requires that telomerase levels are limiting. *EMBO J.* 2006; 25(3):565-74.
5. Greider CW. Telomerase RNA levels limit the telomere length equilibrium. *Cold Spring Harb Symp Quant Biol.* 2006; 71:225-9.
6. Cao Y, et al. Increased copy number of the TERT and TERC telomerase subunit genes in cancer cells. *Cancer Sci.* 2008; 99(6):1092-9.
7. Soder AI, et al. Amplification, increased dosage and in situ expression of the telomerase RNA gene in human cancer. *Oncogene.* 1997;14(9):1013-21.
8. Armanios MY, et al. Telomerase mutations in families with idiopathic pulmonary fibrosis. *N Engl J Med.* 2007;356 (13):1317-26.
9. Tsakiri KD, et al. Adult-onset pulmonary fibrosis caused by mutations in telomerase. *Proc Natl Acad Sci U S A.* 2007;104 (18):7552-7.
10. Vulliamy T, et al. The RNA component of telomerase is mutated in autosomal dominant dyskeratosis congenita. *Nature.* 2001;413(6854):432-5.
11. Dehlin, E., Wormington, M., Korner, C.G. & Wahle, E. Cap-dependent deadenylation of mRNA. *The EMBO journal.* 2000; 19(5):1079-86.
12. Korner, C.G. & Wahle, E. Poly(A) tail shortening by a mammalian poly(A)-specific 3'-exoribonuclease. *The Journal of biological chemistry.* 1997; 272(16):10448-56.
13. Korner, C.G. *et al.* The deadenylating nuclease (DAN) is involved in poly(A) tail removal during the meiotic maturation of *Xenopus* oocytes. *EMBO J.* 1998; 17(18):5427-37.

14. Virtanen, A., Henriksson, N., Nilsson, P. & Nissbeck, M. Poly(A)-specific ribonuclease (PARN): an allosterically regulated, processive and mRNA cap-interacting deadenylase. *Crit Rev Biochem Mol Biol* 2013; 48(2):192-209.
15. Yoda, M. *et al.* Poly(A)-specific ribonuclease mediates 3'-end trimming of Argonaute2-cleaved precursor microRNAs. *Cell Rep* 2013; 5(3):715-26.
16. Berndt, H. *et al.* Maturation of mammalian H/ACA box snoRNAs: PAPD5-dependent adenylation and PARN-dependent trimming. *RNA* 2012; 18(5):958-72.
17. Stuart, B.D. *et al.* Exome sequencing links mutations in PARN and RTEL1 with familial pulmonary fibrosis and telomere shortening. *Nat Genet* 2015; 47(5):512-7.
18. Tummala, H. *et al.* Poly(A)-specific ribonuclease deficiency impacts telomere biology and causes dyskeratosis congenita. *J Clin Invest* 2015; 125(5):2151-60.
19. Mason, P.J. & Bessler, M. mRNA deadenylation and telomere disease. *J Clin Invest.* 2015; 125(8):3304.
20. Moon D.H. *et al.* Poly(A)-specific ribonuclease (PARN) mediates 3'-end maturation of the telomerase RNA component. *Nat Genet.* 2015; 47(12):1482-8.
21. Boyraz B. *et al.* Posttranscriptional manipulation of TERC reverses molecular hallmarks of telomere disease. *JCI.* 2016; 126(9):3377-82.
22. Boyraz B. *et al.* A novel TERC CR4/CR5 domain mutation causes telomere disease via decreased TERT binding. *Blood.* 2016; 128(16):2089-2092.
23. Sancar, A., Lindsey-Boltz, L.A., Unsal-Kacmaz, K., and Linn, S. Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annu Rev Biochem* 2004; 73:39-85.
24. Henderson, E.R., and Blackburn, E.H. An overhanging 3' terminus is a conserved feature of telomeres. *Mol Cell Biol* 1989; 9(1):345-8.
25. Klobutcher, L.A., Swanton, M.T., Donini, P., and Prescott, D.M. All gene-sized DNA molecules in four species of hypotrichs have the same terminal sequence and an unusual 3' terminus. *Proc Natl Acad Sci U S A* 1981; 78(5):3015-9.

26. Palm, W., and de Lange, T. How shelterin protects mammalian telomeres. *Annu Rev Genet.* 2008; 42, 301-334.
27. Vaziri, H. *et al.* Loss of telomeric DNA during aging of normal and trisomy 21 human lymphocytes. *Am. J. Hum. Genet.* 1993; 52(4):661-7.
28. de Lange, T. Protection of mammalian telomeres. *Oncogene.* 2002; 21, 532-540.
29. de Lange, T. How telomeres solve the end-protection problem. *Science* 2009; 326, 948-952.
30. Baumann, P., and Cech, T.R. Pot1, the putative telomere end-binding protein in fission yeast and humans. *Science.* 2001; 292, 1171-1175.
31. Kim, S.H., Beausejour, C., Davalos, A.R., Kaminker, P., Heo, S.J., and Campisi, J. TIN2 mediates functions of TRF2 at human telomeres. *J Biol Chem.* 2004; 279, 43799-43804.
32. Li, B., Oestreich, S., and de Lange, T. Identification of human Rap1: implications for telomere evolution. *Cell.* 2000; 101, 471-483.
33. Liu, D., Safari, A., O'Connor, M.S., Chan, D.W., Laegeler, A., Qin, J., and Songyang, Z. PTOPI interacts with POT1 and regulates its localization to telomeres. *Nat Cell Biol.* 2004; 6, 673-680.
34. Ye, J.Z., Hockemeyer, D., Krutchinsky, A.N., Loayza, D., Hooper, S.M., Chait, B.T., and de Lange, T. POT1-interacting protein PIP1: a telomere length regulator that recruits POT1 to the TIN2/TRF1 complex. *Genes Dev.* 2004; 18, 1649-1654.35.
35. Bae, N.S., and Baumann, P. A RAP1/TRF2 complex inhibits nonhomologous end-joining at human telomeric DNA ends. *Mol Cell.* 2007; 26, 323-334.
36. Sarthy, J., Bae, N.S., Scrafford, J., and Baumann, P. (2009). Human RAP1 inhibits non-homologous end joining at telomeres. *EMBO J* 2009; 28, 3390-3399.
37. van Steensel, B., Smogorzewska, A., and de Lange, T. TRF2 protects human telomeres from end-to-end fusions. *Cell.* 1998; 92, 401-413.
38. Greider, C.W., and Blackburn, E.H. Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. *Cell.* 1985; 43, 405 413.

39. Greider, C.W., and Blackburn, E.H. The telomere terminal transferase of Tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity. *Cell*. 1987; 51, 887-898.
40. Greider, C.W., and Blackburn, E.H. A telomeric sequence in the RNA of Tetrahymena telomerase required for telomere repeat synthesis. *Nature*. 1989; 337, 331-337.
41. Heiss, N.S. *et al.* X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. *Nat Genet* 1998; 19, 32-8.
42. Vulliamy, T. *et al.* Mutations in the telomerase component NHP2 cause the premature ageing syndrome dyskeratosis congenita. *Proceedings of the National Academy of Sciences of the USA*. 2008; 105, 8073-8.
43. Walne, A.J. *et al.* Genetic heterogeneity in autosomal recessive dyskeratosis congenita with one subtype due to mutations in the telomerase-associated protein NOP10. *Human molecular genetics* 2007; 16, 1619-29.
44. Olovnikov, A.M. A theory of marginotomy. The incomplete copying of template margin in enzymic synthesis of polynucleotides and biological significance of the phenomenon. *Journal of theoretical biology*. 1973; 41, 181-190.
45. Chen JL, Blasco MA, Greider CW. Secondary structure of vertebrate telomerase RNA. *Cell*. 2000; 100(5):503-514.
46. Mitchell, J.R., Cheng, J. & Collins, K. A box H/ACA small nucleolar RNA-like domain at the human telomerase RNA 3' end. *Molecular and cellular biology* 1999; 19, 567-76.
47. Box, J.A., Bunch, J.T., Tang, W. and Baumann, P. Spliceosomal cleavage generates the 3' end of telomerase RNA. *Nature*. 2008; 456 (7224), 910-4.
48. Tang, W., Kannan, R., Blanchette, M., and Baumann, P. Telomerase RNA biogenesis involves sequential binding by Sm and Lsm complexes. *Nature*. 2012; 484 (7393), 260-4
49. de la Fuente, J. & Dokal, I. Dyskeratosis congenita: advances in the understanding of the telomerase defect and the role of stem cell transplantation. *Pediatr. Transplant* 2007; 11(6):584-94.

50. Alder, J. K. *et al.* Short telomeres are a risk factor for idiopathic pulmonary fibrosis. *Proc. Natl Acad. Sci. USA.* 2008; 105(35):13051-6.
51. Cronkhite, J.T. *et al.* Telomere shortening in familial and sporadic pulmonary fibrosis. *Am. J. Respir. Crit. Care Med.* 2008; 178(7):729-37.
52. Yamaguchi, H. *et al.* Mutations in *TERT*, the gene for telomerase reverse transcriptase, in aplastic anemia. *N. Engl. J. Med.* 2005; 352(14):1413-24.
53. Du, H. Y. *et al.* *TERC* and *TERT* gene mutations in patients with bone marrow failure and the significance of telomere length measurements. *Blood.* 2008; 113(2):309-16.
54. Yamaguchi, H. *et al.* Mutations of the human telomerase RNA gene (*TERC*) in aplastic anemia and myelodysplastic syndrome. *Blood.* 2003; 102(3):916-8.
55. Calado, R. T. *et al.* A spectrum of severe familial liver disorders associate with telomerase mutations. *PLoS ONE* 2009; 4(11):e7926.
56. Kirwan, M. *et al.* Defining the pathogenic role of telomerase mutations in myelodysplastic syndrome and acute myeloid leukemia. *Hum. Mutat.* 2009; 30(11):1567-73.
57. Calado, R. T. *et al.* Constitutional hypomorphic telomerase mutations in patients with acute myeloid leukemia. *Proc. Natl Acad. Sci. USA.* 2009; 106(4):1187-92.
58. Hao, L. Y. *et al.* Short telomeres, even in the presence of telomerase, limit tissue renewal capacity. *Cell.* 2005; 123(6):1121-31.
59. Goldman, F. D. *et al.* Characterization of primitive hematopoietic cells from patients with dyskeratosis congenita. *Blood.* 2008; 111(9):4523-31.
60. Vulliamy, T. *et al.* Disease anticipation is associated with progressive telomere shortening in families with dyskeratosis congenita due to mutations in *TERC*. *Nature Genet.* 2004; 36(5):447-9.
61. Armanios, M. *et al.* Haploinsufficiency of telomerase reverse transcriptase leads to anticipation in autosomal dominant dyskeratosis congenita. *Proc. Natl Acad. Sci. USA.* 2005; 102(44):15960-4.
62. Connor J.M. *et al.* Assignment of the gene for dyskeratosis congenita to Xq28; *Hum Genet.* 1986; 72(4):348-51.

63. Mitchell J.R. *et al.* A telomerase component is defective in the human disease dyskeratosis congenita. *Nature*. 1999; 402(6761):551-5.
64. Takahashi K., Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*. 2006; 126(4):663-76.
65. Mitchell JR, Collins K. Human telomerase activation requires two independent interactions between telomerase RNA and telomerase reverse transcriptase. *Mol Cell*. 2000; 6(2):361-371.
66. Tesmer VM, Ford LP, Holt SE, et al. Two inactive fragments of the integral RNA cooperate to assemble active telomerase with the human protein catalytic subunit (hTERT) in vitro. *Mol Cell Biol*. 1999; 19(9):6207-6216.
67. Chen JL, Opperman KK, Greider CW. A critical stem-loop structure in the CR4-CR5 domain of mammalian telomerase RNA. *Nucleic Acids Res*. 2002; 30(2):592-597.
68. Kiss, T., Fayet-Lebaron, E. & Jady, B.E. Box H/ACA small ribonucleoproteins. *Molecular cell*. 2010; 37(5):597-606.
69. Rammelt, C., Bilen, B., Zavolan, M. & Keller, W. PAPD5, a noncanonical poly(A) polymerase with an unusual RNA-binding motif. *RNA*. 2011; 17(9):1737-46.
70. LaCava, J. *et al.* RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell*. 2005; 121(5):713-24.
71. Fu, D. & Collins, K. Distinct biogenesis pathways for human telomerase RNA and H/ACA small nucleolar RNAs. *Molecular cell*. 2003; 11(5):1361-72.
72. Zaug, A.J., Linger, J. & Cech, T.R. Method for determining RNA 3' ends and application to human telomerase RNA. *Nucleic acids research*. 1996; 24(3):532-3.
73. Goldfarb, K.C. & Cech, T.R. 3' terminal diversity of MRP RNA and other human noncoding RNAs revealed by deep sequencing. *BMC molecular biology*. 2013; 14:23.

Appendix-1: Publications

Moon DH, Segal M, Boyraz B, Guinan E, Hofmann I, Cahan P, Tai AK, Agarwal S. Poly(A)-specific ribonuclease (PARN) mediates 3'-end maturation of the telomerase RNA component. *Nature Genetics*. Dec;47(12):1482-8, 2015.

Boyraz B, Moon DH, Segal M, Muosieyiri MZ, Aykanat A, Tai AK, Cahan P, Agarwal S. Post-transcriptional manipulation of TERC reverses molecular hallmarks of telomere disease. *Journal of Clinical Investigation*. Sep 1;126(9):3377- 82, 2016.

Boyraz B, Bellomo CM, Fleming MD, Cutler CS, Agarwal S. A novel TERC CR4/CR5 domain mutation causes telomere disease via decreased TERT binding. *Blood*. Oct 20;128(16):2089-2092, 2016.

Appendix-2: Meetings attended

April 2016, EMBO Telomeres, Telomerase and Disease Meeting. Liège, Belgium. Talk presented. Title: Post-transcriptional manipulation of TERC reverses molecular hallmarks of telomere disease.

June 2016, Boston Children's Hospital Dr. M. Judah Folkman Research Day. Boston, MA, USA. Talk presented. Title: Post-transcriptional processing of human telomerase RNA component regulates telomere length homeostasis. Won best presentation award.

Appendix-3: Collaborators' contributions

Matthew F. Segal (Boston Children's Hospital) contributed to Figures 4.2.1, 4.2.3, 4.3.1, 4.3.2, 4.3.4.

Diane H. Moon (Boston Children's Hospital) contributed to Figures 4.2.2, 4.2.4, 4.3.3, 4.3.4.

Suneet Agarwal (Boston Children's Hospital) contributed to Figure 4.2.6.

Maud Z. Muosieyiri (Wellesley College) contributed to Figure 4.3.9.

Aslı Aykanat (Koç University) contributed to Figure 4.3.13.

Albert K. Tai (Tufts Medical School) performed next generation sequencing, analyzed 3' RACE deep sequencing data.

Patrick Cahan (Johns Hopkins University) analyzed RNA-seq data.

Eva Guinan (Boston Children's Hospital) provided PARN-mutant patient information and materials.

Corey S. Cutler (Dana-Farber Cancer Institute) and Courtney Bellamo (Albany Medical Center) provided TERC-mutant patient information and materials.

Mark D. Fleming (Boston Children's Hospital) analyzed histology and provided images.

Appendix-4: Funding sources

The work was funded by Pedals for Pediatrics, Boston Children's Hospital; NIH; the Manton Center for Orphan Disease Research; and the Scientific and Technological Council of Turkey (TÜBİTAK 2214-A).

9. CURRICULUM VITAE

I- Personal Information

Surname, Name: Boyraz, Barış

Date of Birth: 27/09/1991

Place of Birth: Ankara, Turkey

Nationality: Turkish

Contact Address: Hoşdere Caddesi Cankaya Evleri B Blok No: 13 Cankaya/Ankara

Telephone Number: +90-538-570-5414

E-mail Address: barisboyraz91@yahoo.com

II- Education

Hacettepe University Medical School, Ankara, Turkey

M.D.-Ph.D. Program, 2009-2017, graduated with high honors

Tumor Biology and Immunology Program

Ph.D. Advisors: Dr. Ayşegül Üner, Dr. Suneet Agarwal

III- Working Experience

Boston Children's Hospital, Boston, MA, USA

Suneet Agarwal Lab

Visiting graduate student, 2014-2016

IV- Publications

Poly(A)-specific ribonuclease (PARN) mediates 3'-end maturation of the telomerase RNA component, *Nature Genetics* 2015, Diane H. Moon, Matthew Segal, **Barış Boyraz**, Eva Guinan, Inga Hofmann, Patrick Cahan, Albert K. Tai, Suneet Agarwal

Post-transcriptional manipulation of TERC reverses molecular hallmarks of telomere disease, *JCI* 2016, **Barış Boyraz**, Diane H. Moon, Matthew Segal, Maud Z. Muosieyiri, Aslı Aykanat, Albert K. Tai, Patrick Cahan, Suneet Agarwal

A novel TERC CR4/CR5 domain mutain causes telomere disease via decreased TERT binding, *Blood* 2016, **Bariş Boyraz**, Courtney M. Bellamo, Mark D. Fleming, Corey S. Cutler, Suneet Agarwal

V- Abstracts and Presentations

Post-transcriptional manipulation of TERC reverses molecular hallmarks of telomere disease, **Boyraz** et al. (*Oral presentation*), EMBO Telomeres, Telomerase and Disease Meeting 2016, Liège, Belgium

Post-transcriptional manipulation of TERC reverses molecular hallmarks of telomere disease, **Boyraz** et al. (*Oral presentation*), Boston Children's Hospital Dr. M. Judah Folkman Research Day 2016, Boston, MA, USA

Manipulating heteroplasmy in mitochondrial genetic disorders, **Bariş Boyraz**, Katelyn Gagne, Suneet Agarwal (*Poster*) Harvard Stem Cell Institute Malkin Retreat 2015, Cambridge, MA, USA

VI- Selected Honors and Awards

Medical Student Resarch Award, Eczacıbaşı Medical Awards 2017, Turkey

Best Presentation Award, Dr. M. Judah Folkman Research Day 2016, Boston, MA, USA

International PhD Fellowship/ 2214-A, 2015-2016, The Scientific and Technological Research Council of Turkey (TÜBİTAK)

Bayer Fellowship, 2015, 65th Lindau Nobel Laureate Meeting, Lindau, Germany