2020

Testing the T-loop Model of Telomeric End Protection

Leonid Alexeevich Timashev

Testing the T-loop Model of Telomeric End Protection

A Thesis Presented to the Faculty of

The Rockefeller University

in Partial Fulfillment of the Requirements for

the degree of Doctor of Philosophy

by

Leonid Alexeevich Timashev

June 2020

Testing the T-loop Model of Telomeric End Protection

Leonid Alexeevich Timashev, Ph.D.
The Rockefeller University 2020

Telomeres are the key structures that protect the ends of linear chromosomes. Although they are often thought about in the context of cellular aging, their most important role is actually to protect the end of the DNA from being mis-interpreted as a site of DNA damage. Telomeres are thought to accomplish this through the action of shelterin. Shelterin is a multi-protein complex where each subunit is dedicated to a specific role in repressing a form of DNA damage signaling, repair, or recruiting telomerase to extend the telomere end.

One of the critical anchor points of shelterin is a protein known as TRF2. TRF2 is necessary to protect telomeres from becoming fused by non-homologous end-joining, and from one of the two main DNA damage signaling pathways, in this case the one driven by ATM and CHK2. It is thought to do this by rearranging the very 3' end of the DNA into a duplex loop – known as the t-loop. Together, TRF2 and t-loops are the main pillars of the t-loop model of end-protection, which is the focus of this thesis.

The first part of this thesis presents an overview of telomere protection and focuses specifically on what is known about end-protection in mammalian cells. From there, we test an alternative model of telomere end-protection, and find it to be unsubstantiated. We next analyze how TRF2 contributes to t-loop formation, including whether TRF2 cooperates with other shelterin components, uses non-shelterin factors, and which domains of TRF2 contribute. Finally, we try to understand how t-loops are made, and whether there are any external factors that

assist TRF2, or whether TRF2 is self-sufficient in repressing signaling, fusions, and forming t-loops. We then discuss the evolution of telomeres which serves as an important reference point towards understanding the greater context of the t-loop model, and its plausibility. The appendix discusses attempts to push the resolution of t-loop imaging in the context of whole cells.

The work presented here is of relevance to understanding the central mechanism of telomere end protection. What t-loops do – if anything – and how they are made is a question that is at the heart of telomere biology.

Acknowledgements

First and foremost I would like to thank my advisor, Titia de Lange, for helping to guide me through this challenging and trying process. Although one writes a thesis only once in their life, you have seen many people go through this experience, and therefore your guidance has been invaluable. You have been able to ground me, focus me, and most importantly, reassure me that I am on the right path throughout my time in your lab. Of course, your endless knowledge of telomeres, DNA damage, and help troubleshooting everything from PCRs to DNA preps has been crucial.

I would also like to thank the members of my committee, Frederick Cross and Sanford Simon in critiquing my work and providing an outside eye. I appreciate the lively discussions, as well as the questions you both posed to me, which though at times frustrating, ultimately increased the quality of my work.

Naturally I would also like to thank all the members of the de Lange lab for their help, both past and present. In particular I would like to thank Adriana, Kaori, Rosie and Devon for making sure that the lab runs: that we have reagents, glassware, and that nothing is exploding or breaking. A special shout out goes to Logan Myler who appeared at the end of my PhD as my bay mate (the first I had in over two years) to give me some life and bring me out of science depression.

I also want to thank my funding, from the Department of Defense, for making this work possible. In addition, the Rockefeller Bio-Imaging Resource Center, and in particular Alison North, has been invaluable in teaching me to use the OMX, keeping me from breaking it, and keeping it running. Xiaowei Zhuang and Hazen

Babcock were also instrumental in making this study happen, and were wonderfully responsive and cooperative collaborators.

Finally, I would like to thank my friends, family, and fencing community for doing your in encouraging me, supporting me, and keeping me sane over the course of my PhD. Last, I want to thank my fiancée, Isabelle Burden, for finding something loveable and tender behind my grim outer shell.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

53BP1	p53 binding protein 1
alt-NHEJ	alternative non-homologous end-joining
ATAC	Assay for Transposase-Accessible Chromatin
ATM	Ataxia telangiectasia mutated
ATR	ATM- and Rad3- related
BRCA	Breast Cancer early onset
ChIP	Chromatin Immunoprecipitation
CoIP	Co-immunoprecipitation
CST	Ctc1/Stn1/Ten1
DDR	DNA Damage Response
DNA	Deoxyribonucleic acid
DSB	Double stand break
EM	Electron Microscopy
EMSA	electrophoresis mobility shift assay
ExM	Expansion Microscopy
FISH	Fluorescence in-situ hybridization
HDR	Homology Directed Repair
HR	Homologous Recombination
iDDR	Inhibitor of the DNA Damage Response
IR	Ionizing Radiation
Lig	Ligase
MEF	Mouse Embryonic Fibroblast
MRN	Mre11/Rad50/Nbs1
NHEJ	non-homologous end-joining
PALM	photo activated localization microscopy
PARP	Poly-ADP RIbose
POT1	Protector of Telomeres 1
Rg	Radius of Gyration
RNA	Ribonucleic acid
SIM	Structured Illumination Microscopy
STORM	Stochastic optical reconstruction microscopy
ST/Q	Serine or Threonine and Glutamine
TIF	Telomere Dysfunction Induced Foci
TIN2	TERF1-interacting nuclear 2
TPP1	TINT2, PTOP, PIP1
TRF1	Telomeric repeat-binding factor 1
TRFH	Telomere repeat factor homology

Chapter 1

Introduction

The presence of linear chromosomes in eukaryotic cells raises several biological questions: how do cells faithfully duplicate their DNA without the loss of genetic information each replication cycle? How do cells distinguish between a deleterious DSB and the natural end of a linear chromosome? To address these challenges, cells have evolved a strategy to cap their chromosome ends with telomeres – composed of many kbs (as little as 1 kb in yeast up to 50 kb in rodents) of highly conserved, tandemly repeated DNA sequence (5’ TTAGGG 3’ in vertebrates). The first problem cells face regarding telomeres is referred to as the end-replication problem which is solved by an RNA templated DNA polymerase – telomerase – which extends the 3’ ends of telomeres (*1*–*4*). The second problem eukaryotes face is known as the end-protection problem and is solved by telomeric proteins collectively known as shelterin.

End Replication

The end-replication problem at chromosome ends can lead to gradual shortening of telomeric DNA. It occurs as a result of three different processes. First, RNA primase may fail to deposit an RNA primer perfectly at the end of the telomere, which will leave an unreplicated region at the DNA end (*5*, *6*). Secondly, even if primase were to accurately deposit the primer at the end, there is no mechanism to replace the actual RNA primer with DNA, which would ultimately cause DNA shortening. The third process that results in telomere shortening is exonucleolytic degradation of the telomere end which is required for proper maintenance of the 3’ overhang (*7*–*11*). After every replication the blunt ends at the telomeres replicated by leading-strand DNA synthesis need to be converted

into 3’ overhangs, which requires nuclease activity (*12*). The length of the overhang corresponds well with the amount of telomeric DNA lost per replicative cycle and repeated cycles of replication would cause progressive telomere shortening, leading to the threat of gene loss close to the telomere, but could also remove sufficient amounts of telomeric DNA so that the shelterin complex can no longer bind, causing the telomere to be recognized as a broken end by ATM.

Telomerase

The solution to the end-replication was found to be a reverse transcriptase enzyme that adds TTAGGG repeats in an RNA template dependent manner. This enzyme is called telomerase and its activity was first identified in extracts of *Tetrahymena* (*1*, *2*). Telomerase is actually a ribonucleoprotein and contains two components: the reverse transcriptase enzyme and the RNA template. The RNA part of the enzyme not only carries the telomeric template, but also provides essential structural elements that allow telomerase to function. Recently, two cryo-EM structures of telomerase have been solved, which will greatly aid our understanding of how telomerase is regulated and assembled (*14*, *15*).

Telomere shortening in human cells that lack telomerase plays an important role in determining their replicative lifespan. This limit to the proliferation of primary human cells is thought to act as a tumor suppressor pathway. Due to its essential function in maintaining telomere length, telomerase is an important factor in cancer progression. Most human cancers break through the telomere shortening tumor suppressor effect by activating telomerase or other telomere maintenance systems

as cancers are unable to grow without finding a method to extend their telomeres. It was first shown in 1994 that tumor cells had a high incidence of active telomerase (*16*). Further proof of the critical role of telomerase came when the Shay and Wright group expressed telomerase in normal cells and were able to significantly increase the number of divisions of the treated populations (*17*). In parallel, it was also shown that an inhibitor of telomerase slowed the growth of tumor cells in culture, suggesting that telomerase is important for tumor viability (*18*). Furthermore, the Weinberg lab was able to generate immortalized cancer cells from previously untransformed cells using telomerase in addition to other factors, such as the SV40 large and small T antigen, and Ras V12 (*19*, *20*). Current work in the field focuses on the use of telomerase inhibitors to treat cancers as an adjuvant to standard of care therapy (*21*, *22*).

A telomerase-independent pathway of telomere maintenance exists as well, called alternative lengthening of telomeres (ALT), which may account for telomere maintenance in the 15% of tumors that lack detectable telomerase activity (*23*, *24*). ALT is a recombination based mechanism whereby shorter telomeres are copied off of longer telomeres, and depends on recombination proteins such as RAD50, RAD51, RAD51AP, and RAD52 (*25–28*). Other hallmarks of ALT are C-circles, PML bodies, and lack of ATRX or DAXX expression (*29–33*). Cancers that utilize ALT as the main pathway of telomere maintenance tend to be more malignant and have poorer prognoses (*34*).

Maintenance of the Telomeric Overhang

Mammalian telomeres end in a long 3' overhang, and maintenance and creation of this overhang is regulated by shelterin. The overhang is typically hidden within the D-loop at the base of the t-loop, and is thought to be required for t-loop formation. Apollo, recruited by TRF2, functions at the initial step of overhang maintenance: it is responsible for short-range resection at the leading-end telomere to generate a substrate for Exo1 loading (9, 10). Loss of Apollo leads to blunt leading-end telomeres which can form telomere-telomere fusions in S or G2. After extensive resection by Exo1, POT1 recruits the CST complex for fill-in of the overhang which is a crucial step in proper regulation of overhang length and to avoid excessive telomere shortening (11).

CST is an essential complex, loss of which leads to catastrophic telomere dysregulation and cell death (*35*, *36*). Mammalian CST was originally identified based on sequence similarity of the mammalian and yeast Stn1 protein, and other complex members were identified as protein interactors (*37*, *38*). The CST complex is related to the RPA complex and contains multiple OB folds which are used to bind the telomeric DNA. The CST complex has three critical roles at telomeres, the first of which is to maintain a proper G rich overhang. CST accomplishes this task by recruiting Polα/Primase to the telomere to begin fill-in of the resected overhang (*11*, *39–41*). The second role of the CST complex is to regulate telomerase. CST acts as an inhibitor of telomerase activity both by sequestering the 3' end of the DNA thereby preventing telomerase from accessing it. In addition, the fill-in reaction itself may also be inhibitory (*42*, *43*). The final role

of the CST complex at telomeres is to facilitate replication of the telomere, which it does either by facilitating restart of the replication fork, or by recruiting Polα to fill in gaps caused by a stalled replication fork (*44*, *45*). CST also plays an important role in the regulation of resection at sites of chromosome internal DNA damage (*46–49*).

Shelterin

The shelterin complex is comprised of TRF1, TRF2, Rap1, TIN2, TPP1, and POT1, and interactions between these subunits have been identified by co-immuno precipitations (Fig. 1). TIN2 uses its N-terminal TRFH domain to bind TRF2 and TPP1, and it has a centrally located binding patch for TRF1. TIN2 can use its TRFH domain to bind TRF2 two different ways – in the linker region of TRF2 or in the TRFH domain of TRF2, however whether this latter interaction is physiologically relevant is unknown (128). In addition, TIN2 binds directly to TPP1 through an interaction of its TRFH domain with the C-terminus of TPP1(*50–52*). TPP1 itself binds to POT1 at a centrally located binding site. Finally, Rap1 was discovered via its interaction with TRF2 and later shown to be homologous to yeast Rap1 (115). This interaction is maintained between Rap1's RCT domain and a short sequence within the TRF2 hinge domain (111–113).

Within the shelterin complex, TRF2, TRF1, and POT1 bind DNA. TRF1 and TRF2 bind telomeric dsDNA with nanomolar affinity using their Myb/SANT domains (*53*). Both TRF1 and TRF2 bind dsDNA as dimers, and the TRFH domain is required for their homodimerization (*54–58*). Each TRF1 or TRF2 dimer can bind two TAGGGTT sequences. In addition, TRF2 also has two additional sites of DNA

binding. First, the N-terminal basic domain of TRF2 can bind structured DNA, such as is found at the base of the t-loop (*59*, *60*). The second is a set of lysines which make up a DNA binding feature in the TRFH domain. These lysines allow TRF2 to wind DNA around its TRFH domain in a nucleosome like manner (*61*, *62*). In contrast to TRF2, POT1 preferentially binds ssDNA (*63–65*). POT1 binds with nanomolar affinity to TTAGGGTTAG sites at the end of the telomere (*66*).

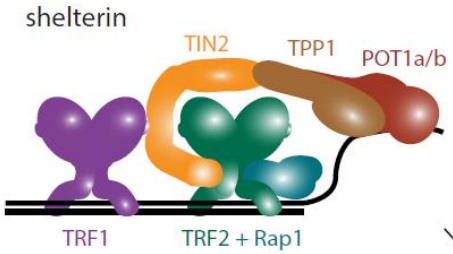
The structures of some of the shelterin components has been solved as well. The Myb/SANT domains of both TRF2 and TRF1 have been solved in complex with DNA (*67*, *68*). In addition, the binding interaction between TPP1 and POT1 has also been solved (*69*, *70*). The RCT domain of Rap1, which binds TRF2, has also been solved (*71*). Finally, the N-terminus of TIN2 has been solved, as has been its docking with TRF1 (*72*, *73*).

The individual components of shelterin do not show strong cooperativity when binding to DNA (*53*). In addition, the shelterin complex itself does not show cooperativity when binding to DNA, suggesting that it likely binds as independent units to telomeric DNA. Shelterin binds to DNA using a either a 3D search or a 2D search by sliding along the DNA (*53*).

The core of shelterin (TRF1, TRF2, Rap1, TIN2) is thought to be the most abundant part of shelterin, and there is sufficient amounts of it to coat the entire telomere (*74*). This also suggests that there may be subcomplexes of shelterin at the telomere, though where and in what proportion is still unknown (*75*, *76*).

Figure 1. The Shelterin Complex. A) The shelterin complex consists of six subunits that each have a distinct role in repressing either DNA damage signaling or DNA repair at the telomere end.

A



Pathway	Repressor (in shelterin)
ATM signaling	TRF2
ATR signaling	POT1a (b)
C-NHEJ	TRF2 (POT1a)
HDR	Rap1 + POT1a (or b)
alt-NHEJ	redundant
5′ resection	redundant?
Replication problems	TRF1

Although TIN2 is required to stabilize TRF2 on telomeres, shelterin is known to be stable without certain subunits, such as POT1, which would suggest that subcomplexes should be functional (*77*, *78*). Despite the fact that the entire telomere is coated in shelterin, nucleosomes do not appear to be perturbed by shelterin, though in vitro results show that shelterin may affect chromatin compaction (*79–81*).

The End Protection Problem

The creation of shelterin-free telomeres by simultaneous deletion of TRF1 and TRF2, as well as other experiments have revealed the end-protection problem at telomeres (*82*). These experiments have revealed seven distinct pathways that components within shelterin are required to repress to maintain telomeric integrity.

Ataxia telangiectasia Mutated (ATM) Signaling

ATM is the cellular DNA damage factor that responds to double strand-breaks (*79*). Telomeres resemble double-strand breaks and therefore are constantly at risk of activating the ATM response. Mre11-RAD50-Nbs1 (MRN) is the primary sensor of double strand breaks, which recruits and activates the ATM kinase (*83*). Persistent ATM signaling would result in CHK2-mediated cell cycle arrest through activation of p53 and inhibition of CDC25 (*84*, *85*). Activation of ATM at sites of DNA damage leads to a signaling cascade that begins with phosphorylation of Histone 2AX. (*86*). This leads recruitment of MDC1, RNF8, and RNF168, which ultimately leads to the ubiquitination of H2A that recruits 53BP1

(*87–89*). 53BP1 plays an important role in resection and repair of DSB breaks, and will be discussed in more detail below. Activation of ATM signaling at telomeres is primarily repressed by TRF2, although TIN2 contributes (*78*, *90*). The exact mechanism will be discussed below. Cells that are unable to arrest in response to ATM signaling, such as those that lack p53, will sometimes bypass mitosis and tetraploidize, which is a hallmark of cancer genomes (*91*, *92*).

Ataxia telangiectasia and Rad3 related (ATR) Signaling

ATR is activated by the presence of single stranded DNA that is bound by its sensor, RPA (*93*). RPA recruits ATRIP as well as the 9-1-1 complex, which activates ATR kinase activity through TopBP1 (*94*). ATR is activated upon loss of POT1, TIN2, or TPP1, the latter two being required to localize POT1 to the telomere. Like ATM, ATR activation would lead to cell cycle arrest through p53 and inhibition of CDC25, though in a CHK1 mediated fashion (*95*). Like ATM, ATR activation leads to the accumulation of DNA damage foci containing γH2AX, MDC1, and 53BP1.

ATR at telomeres is primarily repressed by POT1 (65). It is thought to prevent ATR activation by outcompeting RPA at the single stranded telomeric DNA. Because POT1 is locally tethered it has a higher local concentration and therefore outcompetes the binding of RPA and prevents recruitment of ATR and ATRIP. Protection by POT1 would then also prevent recruitment of RAD17 and the 9-1-1 complex, which would prevent ATR activation (147, 154, 155).

ATR activation at telomeres can also arise from replication problems. Due to the G-rich repetitive nature of the telomeric DNA, telomeres are prone to replication fork arrest (*96–99*). Within shelterin, TRF1 serves as the primary component responsible for promoting telomere replication in part by recruiting helicases such as BLM to assist with replication and unwinding of G4 structures (*98*, *100*). When TRF1 is deleted, replication fork stalling occurs at telomeres, which activates ATR (*100*).

Poly [ADP-ribose] polymerase 1

PARP1 is a sensor of ssDNA breaks and plays an important role in the recognition and repair of single strand breaks, including ssDNA breaks formed by base excision repair (*101*, *102*). PARP1 specifically recognizes a free 5’ end with a 3’ overhang (*103*, *104*). Accumulation of PARP1 results in deposition of PAR on chromatin which can cause it to relax, and on DNA bound proteins, which can cause them to unbind (*105*, *106*). PAR also recruits a number of factors to repair the damage and extended activity of PARP1 can lead to recruitment of MRN and activation of the DNA damage response as well as H2AX phosphorylation (*107*, *108*). PARP1 and Ku70/80 both compete for DNA ends and in the absence of Ku70/80 PARP1 signaling can lead to alt-NHEJ being used to mis repair telomere ends (*109*).

Repressing PARP1 is a challenge at telomeres since every telomere carries a 5’ end with a ssDNA overhang (*110*, *111*). Failure to repress PARP1 activation

would result in PARsylation of shelterin proteins and their removal from telomeric DNA.

c-NHEJ

c-NHEJ is the fastest and most common way by which cells repair their double strand breaks. c-NHEJ is initiated at breaks by the Ku70/80 heterodimer which synapses the ends and recruits DNA-PKcs (*112*). DNA-PKcs activates end processing by Artemis (*113*). Finally, XRCC4, XLF and Ligase 4 cooperate to bridge the ends and fuse the two free ends together (*114*).

As telomeres resemble double strand breaks, they are always at risk of being aberrantly repaired by c-NHEJ if they are not protected by the shelterin complex. Failure to protect from c-NHEJ results in a striking phenotype of long trains of chromosomes all fused at their telomeres (*115*). Such fusions prevent cells from undergoing proper cell division, can lead to chromosome bridges, cell death, or genome instability (*116*).

Telomere fusions can present differently depending on whether they occurred in G1 or S/G2. G1 fusions result in chromosome type fusions while S/G2 fusions result in chromatid type fusions which can be converted to chromosome-type fusions in the next cell cycle (*9*, *11*).The timing of fusions by c-NHEJ is regulated by CYREN, which binds and represses Ku70/80 in S/G2 to cause preferential repair of DSBs by HR (*117*). Interestingly, telomere fusions are not inhibited by the 3’ overhang, and unlike genomic DSB, telomere fusions require ATM activation, likely due to the role that 53BP1 plays (*78*, *118*, *119*). 53BP1

promotes fusion of deprotected telomere ends in G1 by stimulating the mobility of double-strand breaks, which is critical to allowing broken ends to find each other (*120*).

Alt-NHEJ

alt-NHEJ, also known as microhomology mediated end-joining, is often regarded as a backup pathway for DNA repair (*121*). This pathway repairs double strand breaks by finding small regions of homolog of 1-5 nt, and extending and ligating the breaks sites together using Pol θ and Lig3 (*122*). Alt-NHEJ requires PARP1 signaling for recruitment of Pol θ and Lig3 (*123*, *124*). Alt-NHEJ is typically an error prone pathway that can result in significant genetic instability.

Alt-NHEJ poses a threat to telomeres because of the intrinsic homology found within the TTAGGG sequence of the 3' telomeric overhang. Alt-NHEJ is repressed genome-wide by Ku70/80, presumably by competing with PARP1 for DNA ends (*109*). At telomeres, TRF2 and TIN2 repress alt-NHEJ, and failure to do so leads to fusions (*82*, *124*, *125*).

Hyper-Resection

Resection is a constant threat at telomeres because they resemble a double-strand break that needs to be repaired. Resection poses an additional challenge because it is a process that must occur at telomeres to properly maintain them – therefore blocking resection altogether is not a viable option (*11*). Instead, it must be tightly regulated to produce an overhang of the correct length.

Hyper-resection at the telomere can produce overhangs that are too long and the shortened C-rich strand can lead to further telomere shortening in the next cell cycle. The 53BP1 pathway is one of the major regulators of resection: 53BP1 acts to block resection, and loss of 53BP1 along with TRF2, TIN2, TPP1 or POT1 produces telomeres with over-resected ends (*82*).

Homologous Recombination

Homologous recombination is a threat to telomeres because of their resemblance to double strand breaks, which can be processed and repaired via homology directed repair. Additionally, telomeres already contain a 3' overhang, making them an ideal substrate for RAD51 binding and the initiation of homologous recombination (*126*).

Homologous recombination at telomeres can be identified by the presence of telomeric sister chromatid exchanges, that can be equal or unequal. Telomere exchanges appear on metaphase spreads as bi-color staining at telomere ends after CO-FISH (*127*).

Unequal exchanges can lead to a cell acquiring a very short telomere, which would induce cell cycle arrest due to depletion of shelterin at the critically short telomere. In addition, homologous recombination can cause loss of telomeric DNA through processing of the t-loop base and cleavage of the loop part (resulting in t-circles). This t-loop cleavage is counteracted by TRF2 (*59*). As mentioned earlier, cells using the ALT pathway of telomere maintenance rely heavily on homologous recombination to maintain their telomere length.

Telomeric Repeat Binding Factor 2 (TRF2)

TRF2 is one of the most important members of the shelterin complex and the subject of the work presented here. It is a MYB/SANT domain containing protein, which uses its Myb domain to recognize and bind specifically to the TTAGGG repeat (*56*). TRF2 acts as a homodimer, interacting with another TRF2 protein through the TRFH domain (*68*). Dimerization is essential to the function of TRF2, as is localization to telomeres via the Myb domain – loss of the Myb domain creates a dominant negative allele of TRF2 that leads to chromosome fusions and cell death (*115*, *128*, *129*). TRF2 also acts as a scaffold for other shelterin components, notably TIN2 and Rap1.

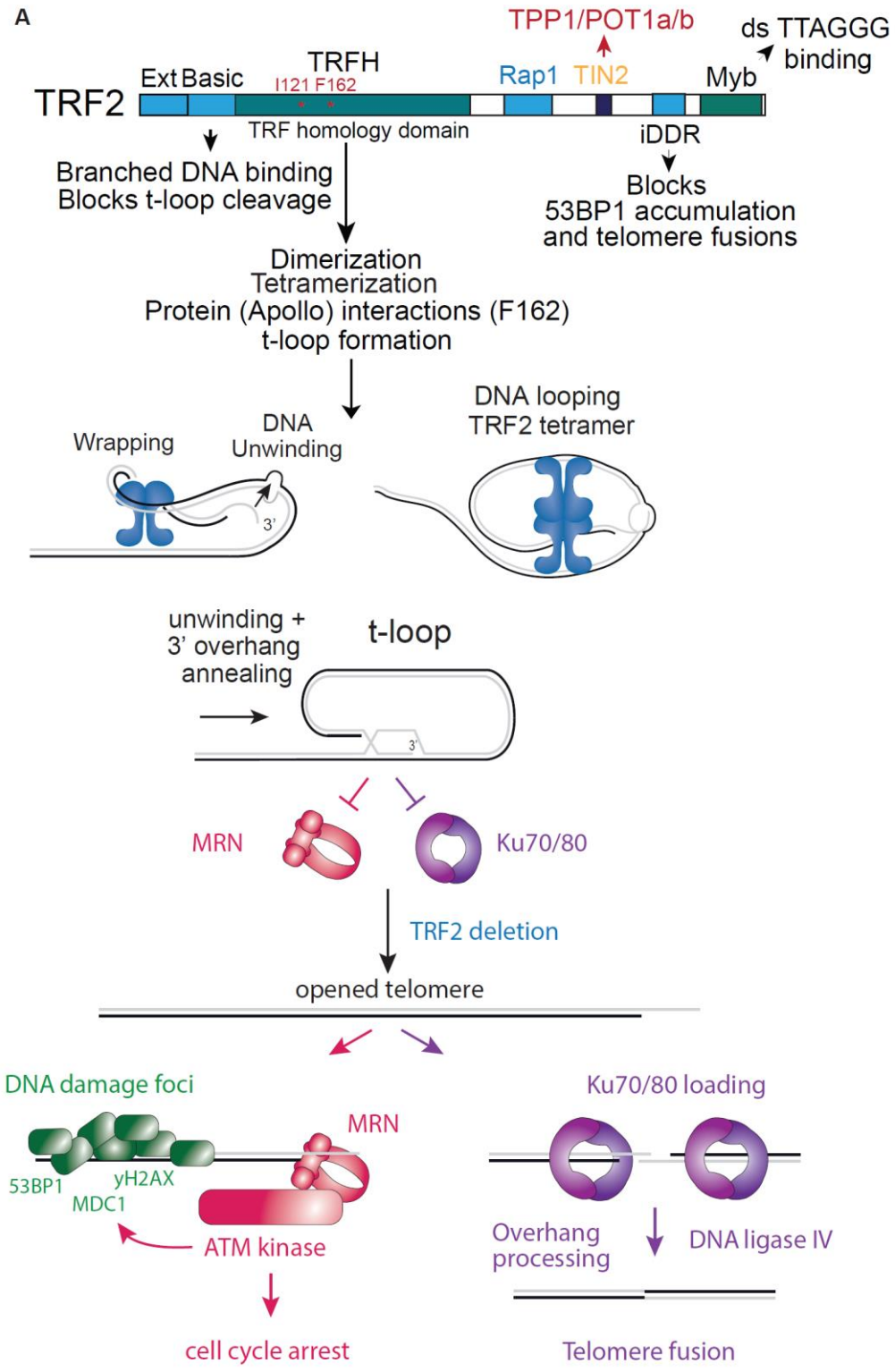
The TRF2 TRFH domain is primarily a protein interacting site. In addition to mediating dimerization, the TRFH domain contains the binding site for Apollo (at amino acid F162) and a binding site for RTEL1 (at amino acid I124). The TRFH domain also contains residues that might promote tetramerization of TRF2, as will be discussed later. TRF2 tetramerization has been proposed as a mechanism of t-loop formation (*130*, *131*). In addition the TRFH domain also includes the “Top domain”, a DNA binding feature that allows TRF2 to wrap DNA in a histone like manner around the TRFH domain (*61*, *132*). This action is mediated by a set of conserved lysines and arginine. This ability of TRF2 to wrap DNA may contribute to TRF2’s strand invasion ability by melting the DNA prior to D-loop formation.

TRF2’s basic domain is so called because it is rich in basic residues. This domain is located at the very N-terminus of the protein and is the first domain, except in the case of the long isoform of TRF2, which will be discussed in detail in

subsequent chapters. The basic domain is responsible for binding branched DNA, such as is found at the base of the t-loop, and repressing t-loop cleavage (*60*, *65*).

The hinge domain of TRF2 is a long flexible region that connects the Myb domain to the rest of the protein. In addition, the hinge domain contains binding sites for two shelterin proteins: Rap1 and TIN2. The Rap1 binding site is located at amino acids 284 - 297, and the TIN2 binding site at amino acids 352 – 367. In addition to these binding domains, the hinge domain also contains the iDDR region. This region prevents 53BP1 signaling but not γH2AX accumulation by interrupting the RNF168 pathway by directly binding BRCC3 and UBR5 (*131*). BRCC3 is a DUB which opposes RNF8 and RNF168 activity, and UBR5 is a ubiquitin ligase which targets RNF168 for degradation (*133*).

Figure 2. TRF2 and the T-loop Model. A) Schematic of the Domains of TRF and the t-loop model of telomere end-protection. The model suggests that the t-loop sequesters the telomere end and prevents loading of Ku70/80 and MRN. This prevents downstream telomere fusions and ATM activation. TRF2 is proposed to form t-loops through either DNA wrapping or through tetramerization mediated looping.


A
TPP1/POT1a/b
ds TTAGGG binding
TRF2
Ext Basic
TRFH
I121 F162
Rap1
TIN2
Myb
TRF homology domain
iDDR
Branched DNA binding
Blocks t-loop cleavage
Blocks
53BP1 accumulation
and telomere fusions
Dimerization
Tetramerization
Protein (Apollo) interactions (F162)
t-loop formation
Wrapping
DNA
Unwinding
DNA looping
TRF2 tetramer
3'
unwinding +
3' overhang
annealing
t-loop
3'
MRN
Ku70/80
TRF2 deletion
opened telomere
DNA damage foci
MRN
53BP1
yH2AX
MDC1
ATM kinase
Ku70/80 loading
Overhang
processing
DNA ligase IV
cell cycle arrest
Telomere fusion

The Myb domain is the most C-terminal domain of TRF2. The Myb domain specifically binds and recognizes the telomeric TTAGGG sequence, and is responsible for localizing TRF2 to the telomere (*56*). The TRF2 Myb domain structure has been solved, and it binds DNA in a non-distorting manner, suggesting that it alone should not be responsible for t-loop formation (*67*, *134*). Loss of the Myb domain creates a dominant negative allele of TRF2, which is unable to localize to telomeres but can still dimerize and poison wild-type TRF2 (*128*).

Out of the many DNA damage and repair pathways found at telomeres, TRF2's role is to blocking four: ATM signaling, c-NHEJ, PARP1 activation, and alt-NHEJ. Loss of TRF2 leads to ATM signaling and extensive chromosome fusions within 72 hours (*86*, *115*, *127*, *135*). Finally, TRF2 is responsible for repressing alt-NHEJ at telomeres in a redundant manner alongside the general repressor Ku70/80 (*82*, *124*, *125*, *136*).

T-loop Model

Genetic deletion and complementation has shown that TRF2 is a repressor of the c-NHEJ DSB repair pathway and ATM signaling at telomeres (Figure 3A). The exact mechanism by which TRF2 acts to repress ATM and block c-NHEJ is not known, but the leading hypothesis in the field is the t-loop model (*137*, *138*) (Figure 3B). The model proposes that TRF2 is able to sequester the end of the telomere by forming a lariat structure in which the 3' overhang strand invades into the telomeric DNA, forming a D-loop. The loop of the t-loop is dsDNA, ssDNA is only found at the D-loop. This entire structure, with a D-loop at the base, is called

a t-loop. The loop size is not fixed and can vary dramatically, and what factors regulate the invasion site are not known. The actual structure at the base of the loop itself is not known either: although hypothesized to be a simple D-loop, it may also be a Holliday Junction, a double Holliday Junction, or may involve any number of G-quadruplexes to hold the structure in place (*139*, *140*). Telomerase would hypothetically have access to the telomere end during telomere replication, when the t-loop is naturally unwound for replication and overhang maintenance. Finally, the 3' end at the base of the D-loop is a structure that can be extended by replication, in a hypothetical process called t-loop extension. Whether t-loop extension occurs in cells is not known.

The t-loop model is rooted in two critical experiments, the first is electron microscopy of looped structures in psoralen/UV crosslinked telomeres extracted from cells (*137*). Looped DNA structures were detected when baculovirus purified TRF2 was incubated with a prepared telomeric template that had a ~100 nt overhang, lending strong support to TRF2 being the primary factor responsible for t-loop formation. In the same assay, TRF1 did not produce t-loops. Next, the authors purified DNA from psoralen/UV crosslinked HeLa cells, liver hepatocytes, and peripheral blood, isolated the telomeres by gel purification, and once again detected large looped structures by EM. In a final experiment, SSB was added to the prepared DNA and visualized on EM, which suggested the presence of a D-loop at the junction.

The second experiment was detection of t-loops by the combination of fluorescence in-situ hybridization (FISH) and STORM imaging of spread and

UV/psoralen crosslinked DNA from nuclei (*138*). Importantly, this experiment showed that loss of TRF2 converted looped telomeres into linear telomeres, supporting the model that t-loops require TRF2 for their establishment and/or maintenance. In contrast, genetic deletion of TRF1, Rap1 and POT1 showed that neither of these proteins is required for t-loop formation. In these experiments, as in the earlier EM experiments, the percentage of molecules that ended in a t-loop never reached 100%. The STORM imaging showed that between 15 and 30% of telomeric molecules contained a t-loop, which is consistent with EM data that showed between 15 and 40% of molecules in a looped conformation. Why not all molecules contain a t-loop is uncertain. One possibility is that a large percentage of t-loops are lost during sample preparation, and that telomeric DNA may snap during the spreading procedure, which will create additional linear molecules from molecules that would have contained a t-loop. A second possibility is that not all telomeres in cells end in a t-loop, or that they spend a significant proportion of the time transiently unwound. If this is the case, it would challenge the mechanism underlying the t-loop model because ATM and c-NHEJ should be activated at telomeres not capped in a t-loop. A third and final possibility is that the crosslinking frequency is not sufficient to crosslink every t-loop; this may be the case in situations where the overhang is small and therefore less likely to incorporate a crosslink.

Another clue towards the importance of TRF2 is that it has been observed preferentially binding the base of the loop, which suggests it is physically involved in the strand invasion step (*141*). It is important to note however, that current data

shows that TRF2 is required to generate t-loops, but it is still not known whether it is sufficient to generate the t-loop in vivo.

Interestingly, although it is known that TRF2 is required for t-loop formation, it was TRF1 that was first identified as being able to pair DNA strands together (*142*). TRF1 was also found to have the ability to bind distant telomeric sequences, bring them together, and actually loop them as well, though this activity appears to be unrelated to t-loop formation, or at least not required (*55*). Interestingly, TRF1 like TRF2, can condense and wind chromatinized DNA around itself as well as increase the rate of strand invasion (*132*). However, the acidic domain of TRF1 appears to counteract both of these properties, perhaps explaining why in vivo only TRF2 is required for t-loop formation.

Further evidence that t-loops are biologically relevant comes from the observation that they can be found on chromatinized DNA, rather than preparations of naked DNA. Purified chromatin from mouse lymphocytes was observed to have distinct loops that contained telomeric DNA (*143*). In addition, the polytene chromosomes of *Oxytricha* were observed to end in loops under electron microscopy, clearly demonstrating the in vivo presence of t-loops (*144*).

How TRF2 creates a t-loop and how it maintains it and prevents it from being unwound except during DNA replication remains an outstanding question. Several studies have implicated different attributes of TRF2 in t-loop formation. One study has implicated a conserved set of basic residues in the TRFH domain – termed the Top DNA binding feature – in being essential for the wrapping role of TRF2, as well as important for strand invasion (*61*). This wrapping is proposed to

induce torsional stress, which melts the DNA and allows for strand invasion to create a D-loop and a t-loop (Fig. 2) (*62*). The rest of the TRFH domain may also be responsible for t-loop formation. The TRFH domain is necessary for dimerization, and swapping of TRF2's TRFH domain for that of TRF1 abolishes t-loop formation and allows for DNA damage signaling (*131*, *145*). There is also a third mechanism by which TRF2 may generate t-loops: via its myb domain, which has been shown to be able to alter chromatin structure (*80*). The myb domain, combined with TRF2's proposed ability to tetramerize (prior to which dimerization via the TRFH domain is a necessary prerequisite), may be enough to tie together telomeric DNA and generate higher order structures at telomeres, such as t-loops (Fig. 2) (*62*, *130*). Tetramerization would not be a novel mechanism to stimulate DNA looping, as binding of the Lac repressor tetramer has been proposed to generate a dsDNA loop (*146*, *147*).

The basic domain of TRF2 is thought to play a role in stabilizing the t-loop by binding the D-loop and consequently preventing t-loop cleavage. In addition, the basic domain prevents the association of PARP1 with telomeres, perhaps by occluding the 5' end. Loss of TRF2's basic domain results in branch migration and t-loop cleavage (*60*). This t-loop cleavage results in stochastic telomere shortening, repeated cycles of which can trigger ATM activation after telomeres reach a critically short length (*59*).

An alternative role that has been proposed for the shelterin complex is in the compaction of telomeric DNA (*148*). A study by the Yildiz group has presented in vivo results that show that all members of the shelterin complex have a role in

compacting telomeric DNA by crosslinking the DNA to shelterin and shelterin components to each other. Some of these results challenge the notion that the t-loop is the primary method of circumventing ATM signaling, but rather that compaction of the DNA prevents access to the telomere end (26). Our work, and work from other groups, conflicts with these published results.

Ultimately, the role of the t-loop and TRF2 is to prevent the access of Ku70/80 or MRN to the telomere end. Both Ku70/80 and MRN are present at the telomere constitutively, but their function there is not known (*112*, *149–152*). MRN is required to activate the ATM signaling pathway, and Ku70/80 is essential for beginning the process of c-NHEJ; if either of these were to take place at telomeres, the results would be disastrous to genome integrity and the viability of the cell (*112*, *153–155*).

Chapter 2

The DNA Damage Response at Telomeres Lacking Shelterin Does Not Require Substantial Chromatin Decompaction

Telomeres are protected by shelterin, a six-subunit protein complex that represses the DNA damage response (DDR) at chromosome ends. Extensive data suggest that TRF2 in shelterin remodels telomeres into the t-loop structure, thereby hiding telomere ends from double-stranded break repair and ATM signaling, whereas POT1 represses ATR signaling by excluding RPA. An alternative protection mechanism was suggested recently by which shelterin subunits TRF1, TRF2, and TIN2 mediate telomeric chromatin compaction, which was proposed to minimize access of DDR factors. We performed super resolution imaging of telomeres in mouse cells after conditional deletion of TRF1, TRF2, or both, the latter of which results in the complete loss of shelterin. Upon removal of TRF1 or TRF2, we observed only minor changes in the telomere volume in most of our experiments. Upon codeletion of TRF1 and TRF2, the telomere volume increased by varying amounts, but even those samples exhibiting small changes in telomere volume showed DDR at nearly all telomeres. Upon shelterin removal, telomeres underwent 53BP1-dependent clustering, potentially explaining at least in part the apparent increase in telomere volume. Furthermore, chromatin accessibility, as determined by ATAC-seq (assay for transposase-accessible chromatin [ATAC] with high throughput sequencing), was not substantially altered by shelterin removal. These results suggest that the DDR induced by shelterin removal does not require substantial telomere decompaction.

Introduction

The essential function of telomeres is to protect chromosome ends from being recognized as damaged DNA by the cellular DNA damage response machinery (*212*). The mechanism by which telomeres solve this end protection problem is of interest because loss of telomere protection plays an important role in cancer development and can lead to a myriad of human diseases (*213*–*215*). Human and mouse telomeres are comprised of a double-stranded telomeric TTAGGG repeat array that extends over many kilobases and terminates in a 3′ protrusion of single-stranded TTAGGG repeats. These sequences can be maintained by telomerase, the reverse transcriptase that counteracts terminal sequence loss during DNA replication (*1*). The telomeric DNA protects chromosome ends through its interaction with the six-subunit shelterin complex (*216*).

Shelterin is bound to telomeres through the interaction of TRF1 and TRF2 with the double-stranded telomeric DNA. TRF1 and TRF2 are linked by the central shelterin subunit TIN2, which stabilizes TRF1 and TRF2 on the DNA and also recruits the POT1/TPP1 heterodimer. POT1 is the ssDNA-binding protein in shelterin that recognizes TTAGGG repeats. Shelterin is compartmentalized with different subunits dedicated to distinct DDR pathways (*217*). TRF2 represses ATM kinase signaling and nonhomologous end-joining at telomeres, whereas POT1 prevents the activation of the ATR kinase. TRF1 does not contribute to the protection of telomere ends per se but is important for the efficient replication of the double-stranded telomeric DNA (*100*, *218*).

Shelterin has been proposed to solve the telomere end protection problem through distinct mechanisms involving primarily TRF2 and POT1. TRF2 has been proposed to protect telomeres by changing the structure of the telomeric DNA into the t-loop configuration, a lariat structure that results from strand invasion of the telomeric 3′ overhang into the double-stranded telomeric DNA (*137*, *138*). TRF2 can promote t-loop formation in vitro, probably due to its ability to wrap DNA (*61*, *141*). T loops have been proposed to represent an architectural solution to the end protection problem by sequestering the telomere end from proteins that load onto double-stranded breaks in DNA. The two main pathways that are repressed by TRF2, ATM kinase signaling and classical NHEJ, are initiated by end-loading factors, the Mre11– Rad50–Nbs1 complex and Ku70/80, respectively. Thus, by hiding the chromosome end from the MRN complex and Ku70/80, t-loop formation by TRF2 could prevent ATM kinase signaling and c-NHEJ. In contrast to the architectural mechanism by which TRF2 is proposed to act, POT1 (POT1a and POT1b in mouse shelterin) has been proposed to repress ATR signaling by rendering the single-stranded telomeric DNA inaccessible to RPA, the ssDNA sensor in the ATR pathway (*78*, *219*, *220*).

Recently, an alternative model was proposed for the protection of telomeres from all aspects of the DDR, including ATM and ATR kinase signaling (*148*). This model, referred to as the compaction model, is based on the observation that the telomere volume of human telomeric chromatin is increased upon knockdown or inhibition of shelterin subunits. Specifically, siRNA-induced knockdown of TRF1 and TIN1 resulted in an eightfold and a six-fold increase in telomere volume,

respectively. Similarly, fivefold greater telomeric volume was reported upon expression of a dominant-negative allele of TRF2, while siRNAs to POT1 or TPP1 had a modest (twofold) effect. The expansion of the telomeric chromatin was proposed to facilitate entry of DDR factors into the telomeric domain, thus allowing the DDR machinery to detect and respond to the telomere end, whereas, at functional telomeres, the shelterin-dependent compaction of the chromatin is proposed to block DDR factors from accessing the telomere terminus (*148*).

The chromatin state of genomic DNA and its compaction affect many nuclear processes, including transcription and replication (*221*). However, the role of chromatin compaction in the DDR is complex. Whereas the rate of DNA repair is slower in heterochromatin than in euchromatin, DDR signaling is not inhibited by the greater compaction of heterochromatin (*222–224*). In fact, decompaction of chromatin can diminish ATM kinase signaling, and, conversely, the induction of chromatin condensation can lead to activation of the ATM kinase in the absence of DNA damage (*225*). These data make it difficult to predict a priori whether chromatin compaction could have a protective role at telomeres.

TRF1 and TRF2 exhibit a number of properties in vitro that could potentially allow shelterin to compact the telomeric chromatin. TRF1 can bend telomeric DNA, pair two stretches of telomeric DNA, and form loops by binding to two distant half-sites using the two Myb domains in the TRF1 dimer (*54*, *142*). Furthermore, the N-terminal basic domain of TRF2 interacts with core histones and facilitates the condensation of naked telomeric DNA in vitro (*158*, *226*). In addition, TRF2 can wrap DNA and thus change the topology of telomeric DNA in vitro (*61*, *161*). It is

not known whether these attributes of TRF1 and TRF2 affect the compaction of the telomeric chromatin in vivo.

Here, we used super resolution stochastic optical reconstruction microscopy (STORM) to determine whether shelterin removal caused decompaction of mouse telomeres by measuring their radius of gyration (Rg) or volume. We used conditional knockouts to remove shelterin subunits TRF1 or TRF2 or the whole shelterin complex by codeletion of both TRF1 and TRF2 from mouse embryonic fibroblast (MEF) cells. Upon deletion of TRF1 or TRF2, we observed relatively small changes in the telomere Rg or volume in the majority of our experiments, although these conditions induced DDR at most telomeres in all of our experiments. Upon deletion of both TRF1 and TRF2, the telomere Rg/volume increased by varying amounts from experiment to experiment, but even samples with relatively small changes in telomere Rg/volume showed activation of the DDR at nearly all telomeres. The apparent increase in telomere Rg/volume could be explained at least in part by the clustering of dysfunctional telomeres, which we documented and showed to occur in a 53BP1-dependent manner. Moreover, the chromatin accessibility, as measured by ATAC-seq (assay for transposase-accessible chromatin [ATAC] with high throughput sequencing), was not altered substantially by shelterin removal. Taken together, these data suggest that removal of TRF1 and/or TRF2 does not necessarily cause substantial chromatin decompaction. Since all of these conditions reliably cause a DDR at most telomeres, our results suggest that chromatin compaction is not a primary mechanism by which shelterin protects telomeres from the DDR.

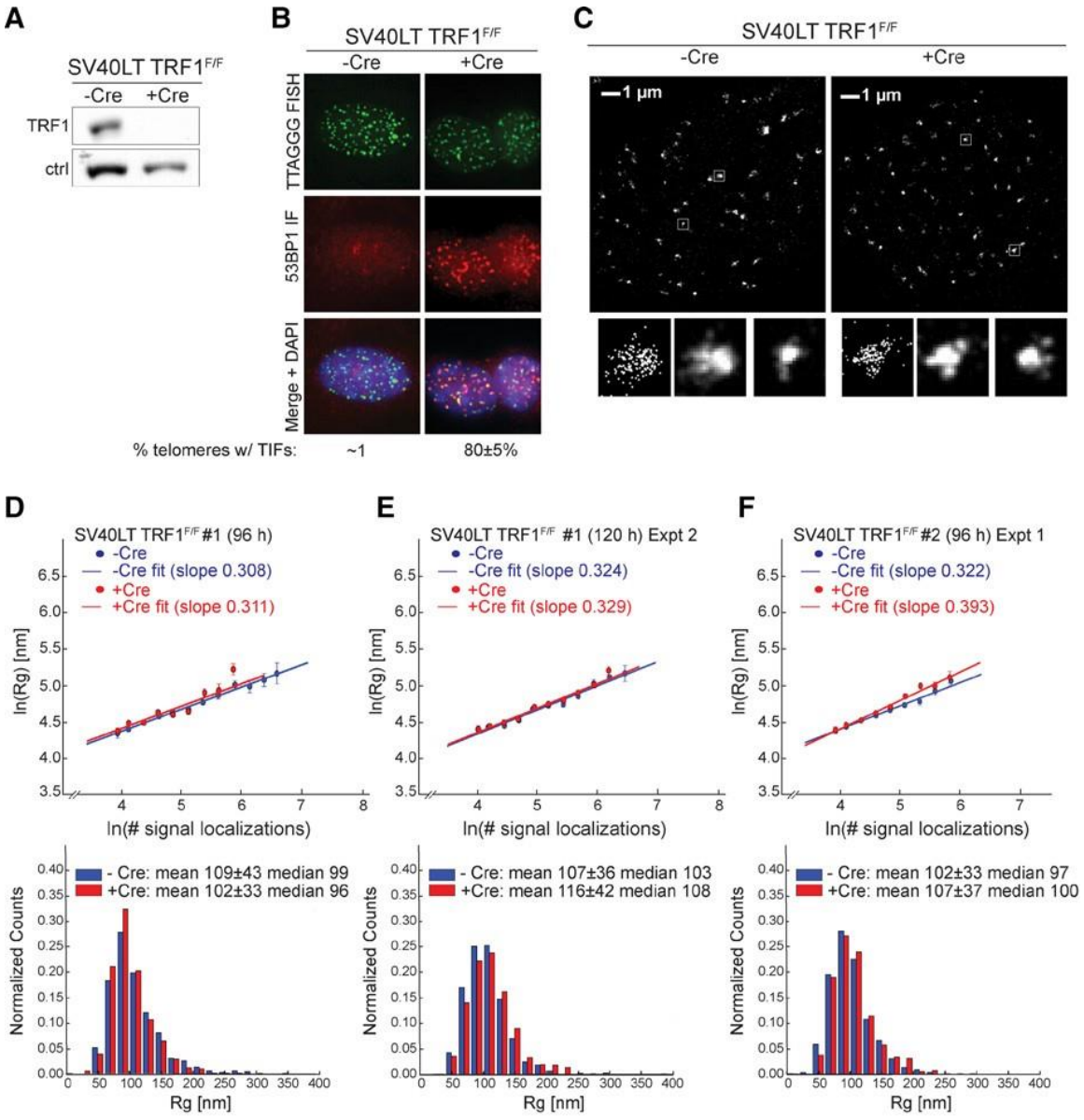
Results

The effect of TRF1 deletion on telomere volume

To probe the role of TRF1 in chromatin compaction at telomeres, we used SV40 large T (SV40LT) immortalized MEFs with floxed TRF1 alleles (SV40LT TRF1$^{F/F}$ CreERT2) in which Cre-mediated deletion of TRF1 can be induced with tamoxifen. In this and all experiments described below, the Cre-mediated deletion of shelterin components was verified by immunoblotting (Fig. 3A). Cells lacking TRF1 exhibit telomere dysfunction due to difficulties in replicating the telomeric sequence (*100*, *218*, *227*). This defect leads to the activation of the ATR kinase and telomere dysfunction-induced foci (TIFs), which represent the accumulation of DDR factors at telomeres (*205*). As expected, after induction of Cre, the vast majority of telomeres showed TIFs containing 53BP1 (Fig. 3B). To detect changes in telomere compaction, we imaged cells with and without TRF1 using telomeric fluorescence in situ hybridization (FISH) and two-dimensional (2D) STORM imaging with ~25-nm resolution (Fig. 3C) (*228*–*230*).

Figure 3. The effect of TRF1 deletion on telomere volume.

(A) Representative immunoblot showing deletion of TRF1 at 120 h after in- duction of Cre with tamoxifen in SV40LT immortalized TRF1$^{F/F}$ Rosa-CreERT2 MEFs. (Ctrl) Nonspecific band used as loading control. (B) Projected z-stack immunofluorescence (IF) images showing the presence of TIFs in the cells shown in A. (Green) Telomeric FISH with TelG-A647; (red) IF for 53BP1; (merge) green and red channels merged with DAPI DNA stain (blue). The percentage of telomeres with a 53BP1-positive TIF is shown below (average and SD from six experiments analyzed at 96 or 120 h). (C) Representative STORM images showing telomeric foci in cells with and without TRF1 (at 96 h after tamoxifen). Enlarged images of selected foci are shown below, and two of the enlarged images are accompanied by a localization presentation at the left that displays individual signal localizations as dots. (D–F, top) Graphs showing the natural log of Rg plotted versus the natural log of the number of telomere signal localizations obtained as in C from the indicated cells with and without Cre and processed in parallel. n ≥ 10 cells for each condition in each independent experiment. (Bottom) Accompanying histograms of the distribution of Rg values with the means ± SDs and median values given. Cells were treated with Cre for 96 or 120 h. D–F represent three independent experiments. (G) Summary of data obtained as in D–F and the measured changes in average Rg values and average convex hull volumes from seven independent TRF1 deletion experiments. The mean Rg values are presented in nanometers.



G

TRF1F/F cell line (time)	Mean Rg (nm) -Cre	Mean Rg (nm) +Cre	Rg change	Convex Hull volume change
#1 (96 h)	109±43	102±33	-6%	-36%
#1 (120 h) Expt 1*	109±44	118±41	+8%	+7%
#1 (120 h) Expt 2*	107±36	116±42	+8%	+30%
#2 (120 h) Expt 1	118±55	128±57	+8%	+34%
#2 (120 h) Expt 2	89±33	117±47	+31%	+144%
#2 (96 h) Expt 1*	102±33	107±37	+5%	+16%
#2 (96 h) Expt 2*	108±39	110±35	+2%	-19%
		Average±SD:	+8±11%	+25±58%

* biological replicates

To quantitatively evaluate changes in compaction, the telomeric signal localizations detected by STORM imaging, each corresponding to a detected activation event of the photo-switchable dye molecules labeling the telomere, were clustered using DBSCAN to segment individual telomeric foci (*231*). The Rg values of individual telomeres were then calculated based on the localizations in each telomeric focus. The Rg was plotted against the number of signal localizations per telomeric focus (Fig. 3D-F), and, in such plots, a decrease in compaction was expected to result in a positive vertical translocation of the plotted line. In addition to the plots of Rg values versus number of localizations, we also displayed the overall Rg distributions under each condition for each independent experiment (Fig. 3D-F).

The average Rg can be influenced by the length of the telomeres (which, in the MEFs used here, ranged from 20 to 50 kb) and is expected to be altered when telomeres fuse and thus double in size. However, in the case of TRF1 deletion, telomere fusions are not frequent, and we do not expect substantial changes in the length of mouse telomeres within the time frame studied here (*100*).

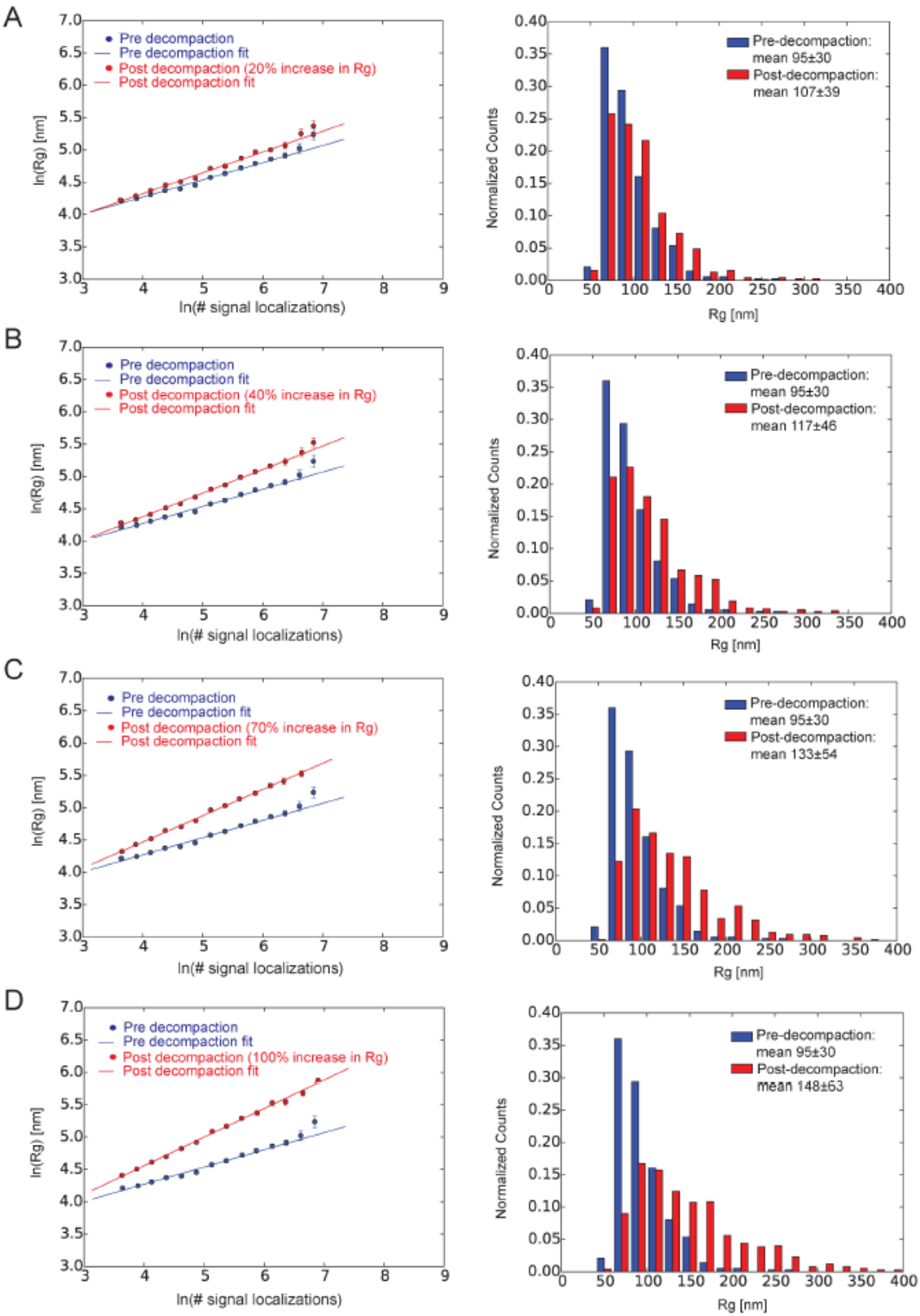
We performed seven independent Cre-mediated deletion experiments to analyze the effect of TRF1 loss. In each experiment, we measured the Rg distributions of telomeres in ≥10 cells for the plus and minus Cre conditions, with the plus Cre samples and the corresponding control minus Cre samples always prepared in parallel on the same day. The measured changes in average Rg values upon Cre-induced deletion of TRF1 were −6%, 8%, 8%, 8%, 31%, 5%, and 2% (average +8%) for the seven independent experiments (Fig. 3D–G). Similar

results were observed for the median Rg values (Fig. 3D–F). Although the measured Rg values can vary by a small amount among different batches of samples, results from the biological replicates (samples treated under the same condition on the same day) (Fig. 3G) and technical replicates (the same batch of cells plated on multiple coverslips) (data not shown) were nearly identical.

We considered the possibility that experimental noise may affect the measured extent of decompaction. To assess such effects, we simulated STORM images of telomeres in the predecompaction state using experimentally measured Rg values and number of FISH signal localizations in the telomeric foci and the measured background localization density under the minus Cre condition (Fig. 3A-D). In the post-decompaction simulations, the same values were applied except that the Rg values were changed based on the specified extent of decompaction. We considered various levels of decompaction; i.e., 20%, 40%, 70%, and 100% increases in Rg values, which correspond to 1.7-fold, 2.7-fold, fivefold, and eightfold increase in volume. Our simulations showed that the measured extent of decompaction in Rg is ~60% of the originally specified value (Fig. 4A-D). Based on these results, the measured changes in Rg values in most of our experiments (six out of seven) are consistent with a real Rg change of <14%, whereas one of the seven experiment is consistent with a real Rg change of ~50%.

Figure 4. Simulation of Telomere Decompaction

(A-D) Simulations using the experimentally measured Rg distribution, average localization number per telomeric focus, and background localization density to simulation STORM images of telomeres in the pre-decompaction condition and similar parameters to simulate STORM images of telomeres in the post-decompaction condition except that the average Rg values are increased by 20% (A), 40% (B), 70% (C), and 100% (D), respectively. Left panels: Plots of Rg values as a function of the localization number for the pre-decompaction condition (blue) and post decompaction conditions (red). Right panels: Bar graphs of distributions of Rg values for the pre-decompaction condition (blue) and post decompaction conditions (red). The means ± SDs are given in each case.


A
Pre decompaction
Pre decompaction fit
Post decompaction (20% increase in Rg)
Post decompaction fit
ln(Rg) [nm]
ln(# signal localizations)
Pre-decompaction: mean 95±30
Post-decompaction: mean 107±39
Normalized Counts
Rg [nm]
B
Pre decompaction
Pre decompaction fit
Post decompaction (40% increase in Rg)
Post decompaction fit
ln(Rg) [nm]
ln(# signal localizations)
Pre-decompaction: mean 95±30
Post-decompaction: mean 117±46
Normalized Counts
Rg [nm]
C
Pre decompaction
Pre decompaction fit
Post decompaction (70% increase in Rg)
Post decompaction fit
ln(Rg) [nm]
ln(# signal localizations)
Pre-decompaction: mean 95±30
Post-decompaction: mean 133±54
Normalized Counts
Rg [nm]
D
Pre decompaction
Pre decompaction fit
Post decompaction (100% increase in Rg)
Post decompaction fit
ln(Rg) [nm]
ln(# signal localizations)
Pre-decompaction: mean 95±30
Post-decompaction: mean 148±63
Normalized Counts
Rg [nm]

We also quantified the experiments using an alternative measure: the convex hull volume. We note that the convex hull analysis is more sensitive to the edge points of a cluster and hence to the chromatin segments located at the three-dimensional (3D) boundary of the telomere foci, whereas Rg measures the overall compaction state of the entire telomere. For the seven independent experiments described above, the changes in the convex hull volume upon deletion of TRF1 were −36%, 7%, 30%, 34%, 144%, 16%, and −19% (average +25%) (Fig. 3G). Considering the effect of the experimental background noise as described above, this measured average change is consistent with a real convex hull volume change of ~50%.

Despite the relatively small changes in the 3D size of telomeres, we observed a strong DDR in all of these experiments: Approximately 80% of the telomeres exhibited TIF signals in all experiments (e.g., Fig. 3B), including the six experiments that exhibited minimal change in telomere Rg or convex hull volume. Hence, DDR at telomeres upon removal of TRF1 does not appear to require substantial chromatin decompaction.

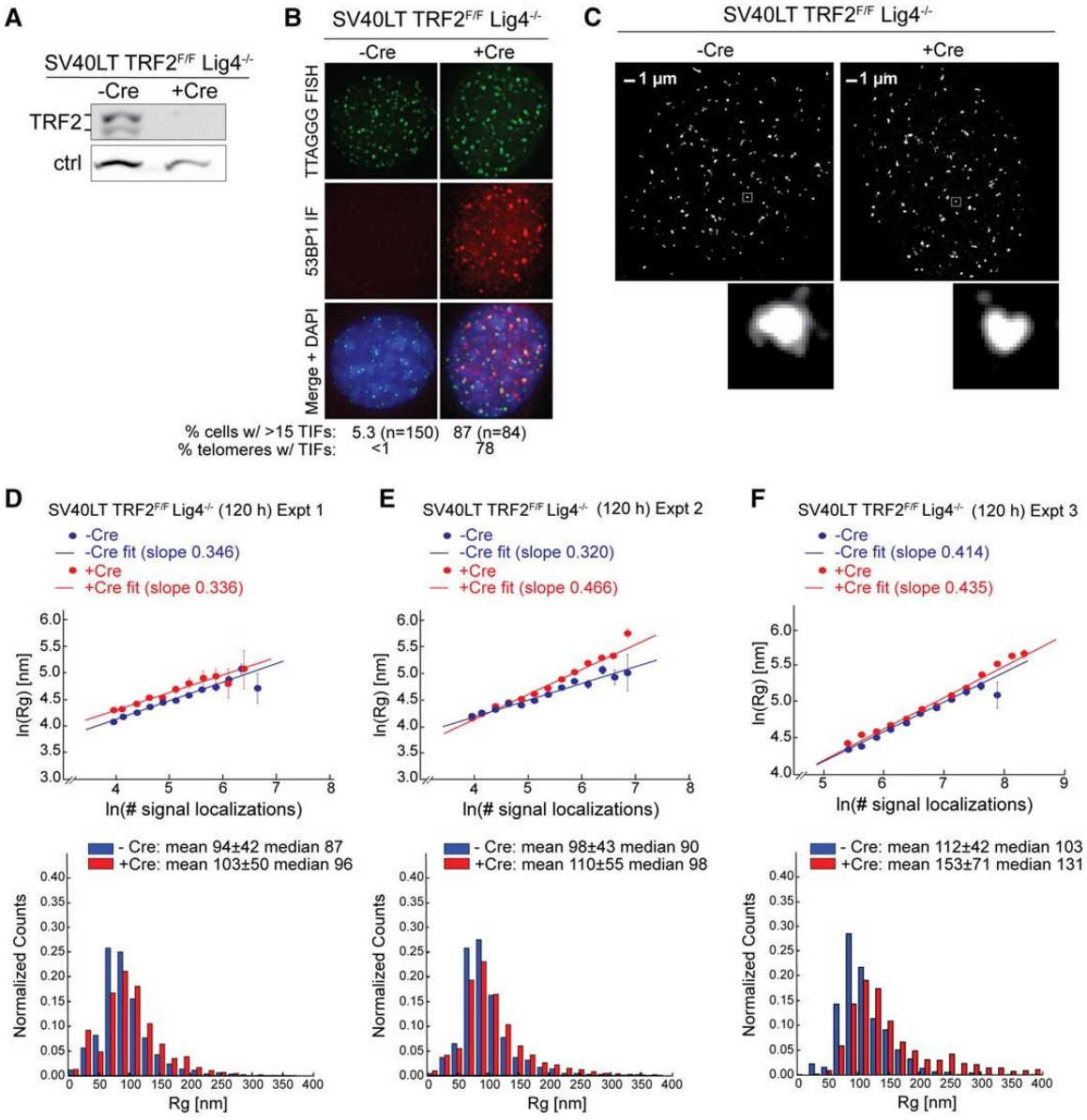
The effect of TRF2 deletion on telomere volume

Next, we investigated the role of TRF2 in chromatin compaction at telomeres. As deletion of TRF2 leads to telomere fusions, which will confound compaction measurements, we performed TRF2 deletion experiments in DNA ligase IV-negative MEFs (SV40LT TRF2$^{F/F}$ Lig4$^{-/-}$), in which telomere fusions are repressed (*86*). The deletion of TRF2 was confirmed by immunoblotting, and TIF

analysis revealed the expected telomere dysfunction following induction of Cre (Fig. 5A, B). STORM imaging and Rg analysis of telomeres with and without TRF2 was performed as described above (Fig. 5C–F). In the three sets of independent experiments conducted with −Cre and +Cre done in parallel, the average Rg values changed by +10%, +12%, and +37% upon Cre-induced deletion of TRF2, respectively (Fig. 5D–F). The measured convex hull volume changed by +9%, +31% and +142%, respectively. Based on our simulation results (Fig. 5A-D), the measured changes in the majority of our experiments (two out of three) are consistent with a real Rg change of ≤20% and a real convex hull volume change of ≤60%, whereas the third experiment showed a much larger increase in Rg and convex hull volume. We note that, in the last experiment, the large increase in Rg and volume was influenced by a small fraction (~20%) of telomeric foci that exhibited both large localization numbers and large Rg/volume values. This fraction could potentially be caused by close

Figure 5. The effect of TRF2 deletion on telomere volume.

(A) Representative immunoblot showing deletion of TRF2 120 h after tamoxifen treatment of SV40LT immortalized TRF2$^{F/F}$ Lig4$^{-/-}$ Rosa-CreERT2 MEFs to induce Cre. (Ctrl) Nonspecific band used as loading control. (B) Projected z-stack IF images showing the induction of TIFs in the Cre-treated cells shown in A. (Green) Telomeric FISH with TelG- A647; (red) IF for 53BP1; (merge) green and red channels merged with DAPI DNA stain (blue). The percentage of cells with >15 TIFs is shown as well as the percentage of telomeres containing TIFs. n = 20 nuclei. (C) Representative STORM images showing telomeric foci in cells with and without TRF2 as in A. Enlarged images of selected foci are shown below. (D–F, top) Graphs showing the natural log of Rg plotted versus the natural log of the number of telomere signal localizations in the indicated cells with and without Cre obtained as in C and processed in parallel. n ≥ 10 cells for each condition in each independent experiment. (Bottom) Accompanying histograms of the distribution of Rg values with the means ± SDs and median values given. Cells were treated with Cre for 120 h. D–F represent three independent experiments.

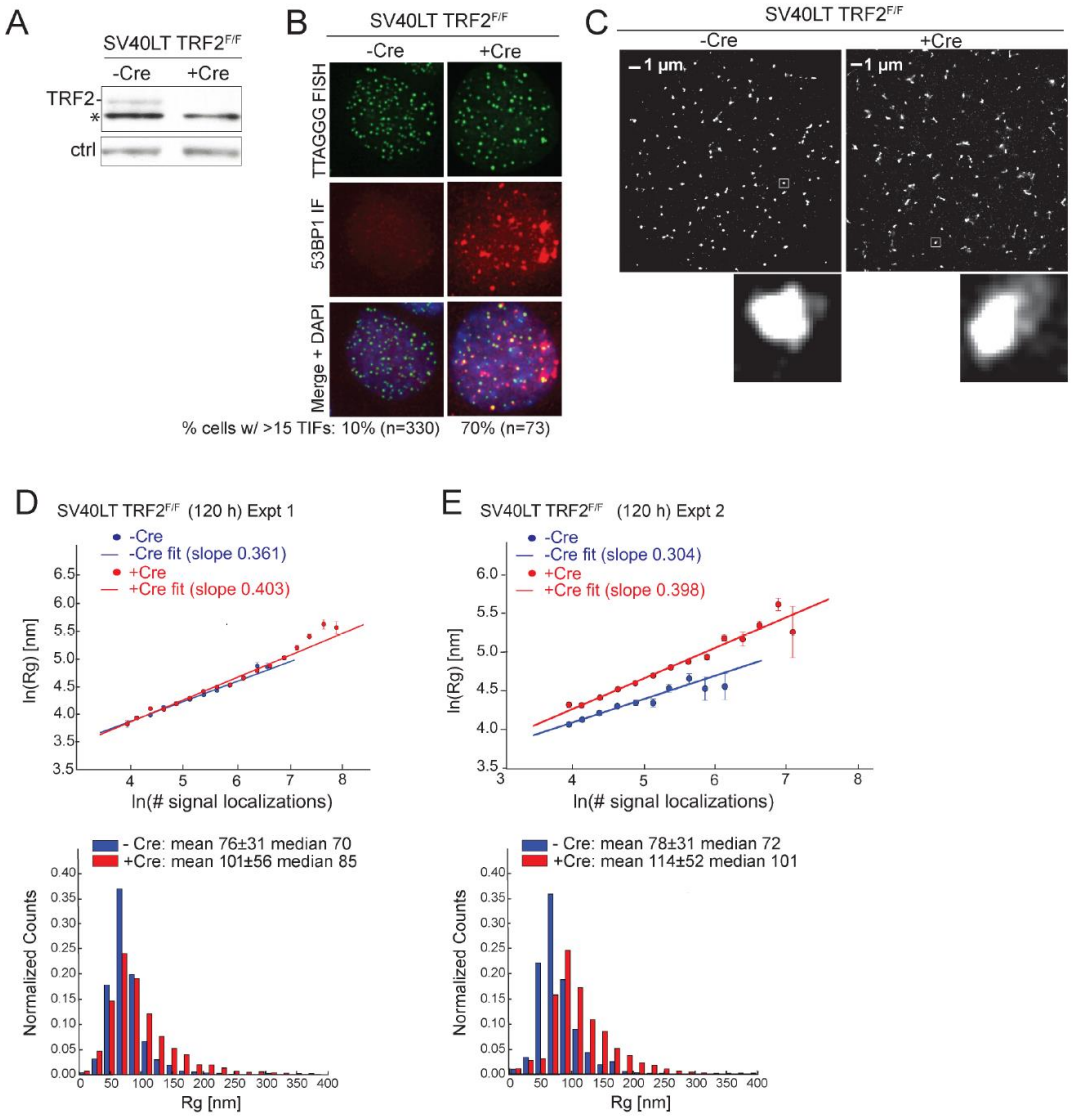

A
SV40LT TRF2F/F Lig4-/-
-Cre
+Cre
TRF2
ctrl
B
SV40LT TRF2F/F Lig4-/-
-Cre
+Cre
TTAGGG FISH
53BP1 IF
Merge + DAPI
% cells w/ >15 TIFs: 5.3 (n=150) 87 (n=84)
% telomeres w/ TIFs: <1 78
C
SV40LT TRF2F/F Lig4-/-
-Cre
+Cre
1 μm
1 μm
D
SV40LT TRF2F/F Lig4-/- (120 h) Expt 1
-Cre
-Cre fit (slope 0.346)
+Cre
+Cre fit (slope 0.336)
ln(Rg) [nm]
ln(# signal localizations)
- Cre: mean 94±42 median 87
+Cre: mean 103±50 median 96
Normalized Counts
Rg [nm]
E
SV40LT TRF2F/F Lig4-/- (120 h) Expt 2
-Cre
-Cre fit (slope 0.320)
+Cre
+Cre fit (slope 0.466)
ln(Rg) [nm]
ln(# signal localizations)
- Cre: mean 98±43 median 90
+Cre: mean 110±55 median 98
Normalized Counts
Rg [nm]
F
SV40LT TRF2F/F Lig4-/- (120 h) Expt 3
-Cre
-Cre fit (slope 0.414)
+Cre
+Cre fit (slope 0.435)
ln(Rg) [nm]
ln(# signal localizations)
- Cre: mean 112±42 median 103
+Cre: mean 153±71 median 131
Normalized Counts
Rg [nm]

apposition of dysfunctional telomeres, which occurs in cells that have undergone endoreduplication in response to DNA damage signaling, a low frequency of telomere fusions mediated by Lig4-independent (alternative) NHEJ, or the 53BP1-dependent clustering of dysfunctional telomeres as described below (*91*). However, it is also possible that this small fraction of telomeres underwent decompaction and became more accessible to FISH probes, hence exhibiting larger numbers of FISH signal localizations. In any case, the majority of telomeres lacking TRF2 did not exhibit substantial decompaction in these experiments, and yet we reliably observed DDR signaling at ~80% of telomeres in TRF2-deleted Lig4$^{-/-}$ MEF cells.

We also examined the effect of TRF2 deletion in NHEJ-proficient (DNA ligase IV-positive) MEFs (Fig. 6A-E), expecting a change in telomere size due to telomere fusions. At the time point analyzed, ~50% of the telomeres were fused. Indeed, in both sets of independent experiments (Fig. 6A-E), we observed telomeric foci with greater numbers of FISH signal localizations and larger Rg values (Rg increased by 33% and 46%), consistent with the expectation that fused telomeres give rise to larger telomeric foci.

Figure 6. Telomere measurements after deletion of TRF2 from Lig4-proficient MEFs

(*A*) Representative immunoblot showing deletion of TRF2 96 h after treatment with tamoxifen from $TRF2^{F/F}$ Rosa-CreERT2 MEFs. Ctrl: non-specific band. Asterisk: non-specific band. (*B*) Projected z-stack immunofluorescence images showing the presence of TIFs 96 h after treatment with tamoxifen. Green, telomeric FISH with TelG-A647; red, IF for 53BP1. Merge, green and red channels merged with DAPI (Blue). Percentage of cells with >15 TIFs is shown below. (*C*) Representative STORM images showing telomeric foci in cells with and without TRF2 as in (A) and enlarged images of selected foci shown below. (*D,E*) Graphs showing the natural log of Rg plotted vs the natural log of the number of telomere signal localizations in the indicated cells with and without Cre as in (A) obtained as in (C) and processed in parallel (n = 8 cells for -Cre; n = 6 cells for +Cre in each independent experiment). Each graph is paired with the accompanying histograms of distribution of Rg values with the means ± SDs and median values given. Cells were treated with Cre for 96 h. D and E represent two independent experiments.

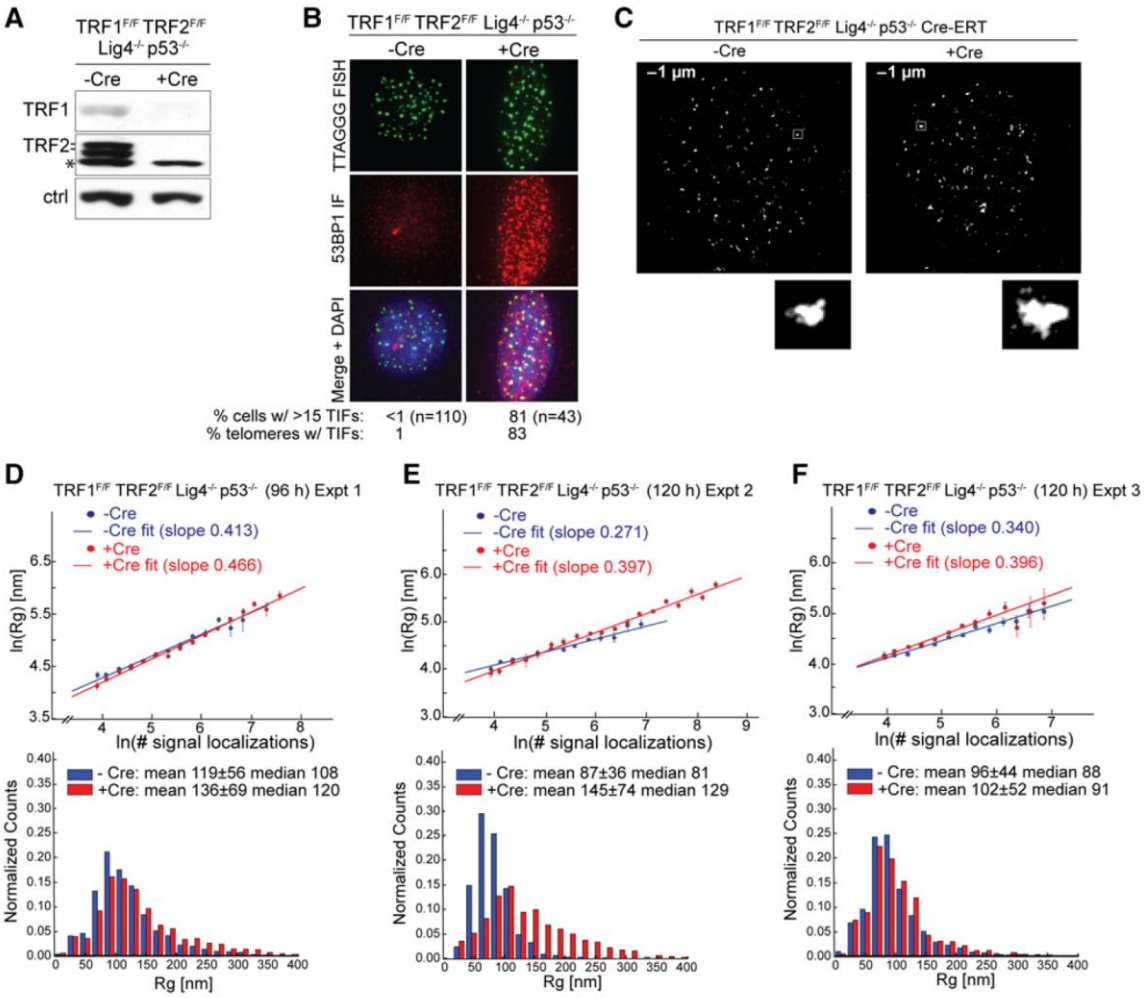

A
SV40LT TRF2F/F
-Cre
+Cre
TRF2
*
ctrl
B
SV40LT TRF2F/F
-Cre
+Cre
TTAGGG FISH
53BP1 IF
Merge + DAPI
% cells w/ >15 TIFs: 10% (n=330) 70% (n=73)
C
SV40LT TRF2F/F
-Cre
+Cre
1 μm
1 μm
D
SV40LT TRF2F/F (120 h) Expt 1
-Cre
-Cre fit (slope 0.361)
+Cre
+Cre fit (slope 0.403)
ln(Rg) [nm]
ln(# signal localizations)
- Cre: mean 76±31 median 70
+Cre: mean 101±56 median 85
Normalized Counts
Rg [nm]
E
SV40LT TRF2F/F (120 h) Expt 2
-Cre
-Cre fit (slope 0.304)
+Cre
+Cre fit (slope 0.398)
ln(Rg) [nm]
ln(# signal localizations)
- Cre: mean 78±31 median 72
+Cre: mean 114±52 median 101
Normalized Counts
Rg [nm]

The effect of shelterin deletion on telomere volume

Next, we examined the effect of dual deletion of TRF1 and TRF2, which results in removal of all shelterin subunits from telomeres ("shelterin-free" telomeres) (*82*). TRF1 and TRF2 were deleted from Lig4 deficient $TRF1^{F/F}$ $TRF2^{F/F}$ $Lig4^{-/-}$ $p53^{-/-}$ MEFs, in which telomere fusions are rare (*82*). The Cre-mediated deletion of TRF1 and TRF2 was confirmed by immunoblotting, and the resulting telomere dysfunction was evident from TIF analysis (Fig. 7A, B). We performed five independent experiments, each with −Cre and +Cre samples done in parallel. STORM imaging (Fig. 7C) showed varying amounts of Rg change upon Cre-induced deletion of TRF1 and TRF2: The average Rg values increased by 14%, 67%, 6%, 22%, and 36% in these five experiments (average 29%) (Fig. 7D–G; Fig. 8), which is consistent with an average of ~50% real Rg change when background is considered (Fig. 8).

Figure 7. The effect of shelterin removal on telomere volume.

(A) Representative immunoblot showing deletion of both TRF1 and TRF2 120 h after treatment of TRF1$^{F/F}$ TRF2$^{F/F}$ Lig4$^{-/-}$ p53$^{-/-}$ Rosa-Cre-ERT2 MEFs with tamoxifen to induce removal of shelterin. (Ctrl) Nonspecific band used as loading control; (∗) nonspecific band. (B) Projected z-stack IF images showing the induction of TIFs in the Cre-treated cells shown in A. (Green) Telomeric FISH with TelG-A647; (red) IF for 53BP1; (merge) green and red channels merged with DAPI DNA stain (blue). The percentage of cells with >15 TIFs is shown as well as the percentage of telomeres containing TIFs. n = 20 nuclei. (C) Representative STORM images showing telomeric foci in cells with and without shelterin at their telomeres. Enlarged images of selected foci are shown below. (D–F) Graphs showing the natural log of Rg plotted versus the natural log of the number of signal localizations per telomeric focus in the indicated cells with and without Cre as in A, imaged as in C, and processed in parallel. n ∼ 10 or more cells for each condition in each independent experiment. Each graph is paired with the accompanying histogram of distribution of Rg values, with the means ± SDs and median values given. Cells were treated with Cre for 96 or 120 h. D–F represent three independent experiments. (G) Summary of data obtained as in D–F and the measured changes in average Rg values and average convex hull volumes from five independent TRF1/TRF2 codeletion experiments. The mean Rg values are presented in nanometers.

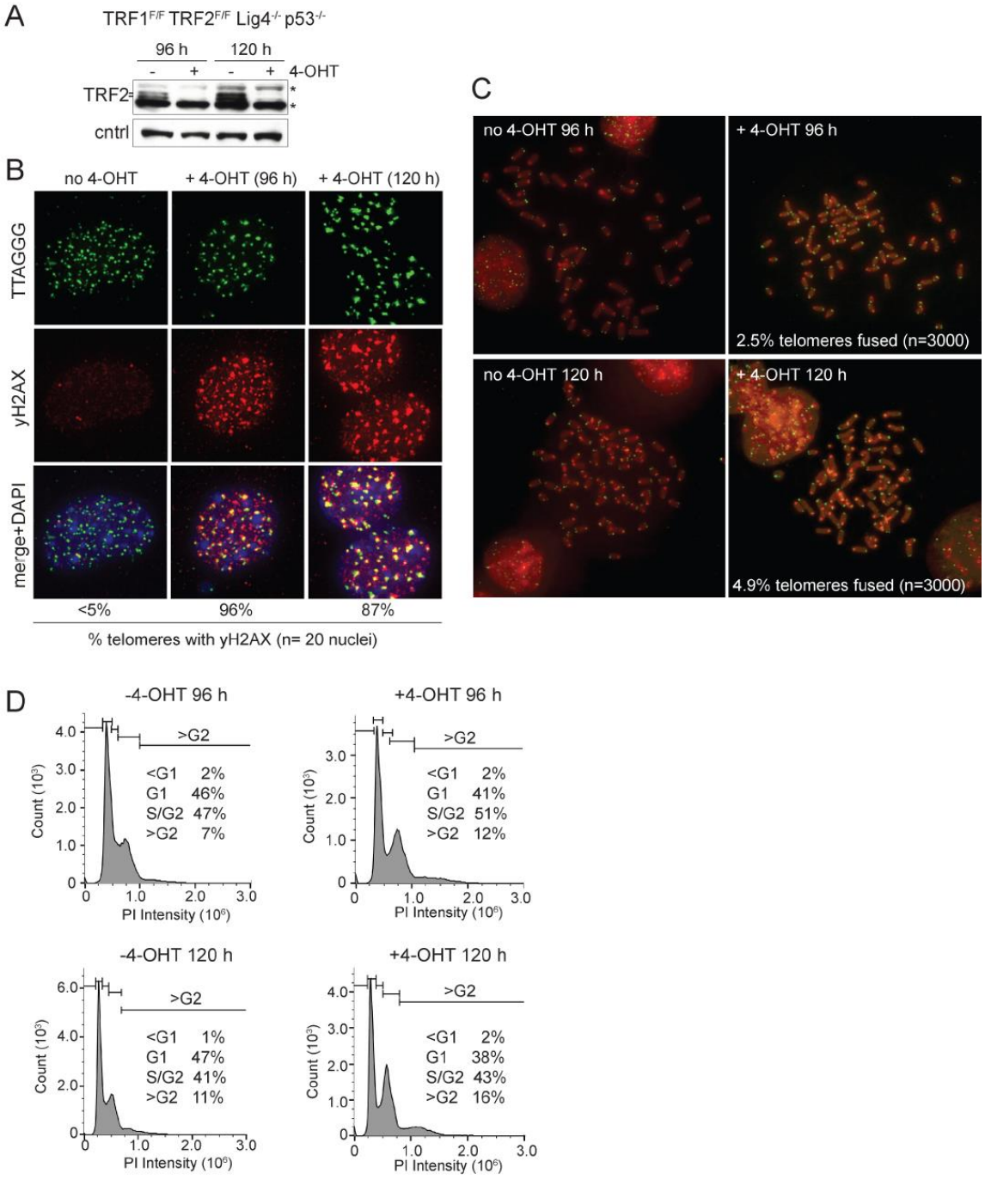

A
TRF1F/F TRF2F/F Lig4-/- p53-/-
-Cre
+Cre
TRF1
TRF2
*
ctrl
B
TRF1F/F TRF2F/F Lig4-/- p53-/-
-Cre
+Cre
TTAGGG FISH
53BP1 IF
Merge + DAPI
% cells w/ >15 TIFs: <1 (n=110) 81 (n=43)
% telomeres w/ TIFs: 1 83
C
TRF1F/F TRF2F/F Lig4-/- p53-/- Cre-ERT
-Cre
+Cre
–1 µm
–1 µm
D
TRF1F/F TRF2F/F Lig4-/- p53-/- (96 h) Expt 1
-Cre
-Cre fit (slope 0.413)
+Cre
+Cre fit (slope 0.466)
ln(Rg) [nm]
ln(# signal localizations)
- Cre: mean 119±56 median 108
+Cre: mean 136±69 median 120
Normalized Counts
Rg [nm]
E
TRF1F/F TRF2F/F Lig4-/- p53-/- (120 h) Expt 2
-Cre
-Cre fit (slope 0.271)
+Cre
+Cre fit (slope 0.397)
ln(Rg) [nm]
ln(# signal localizations)
- Cre: mean 87±36 median 81
+Cre: mean 145±74 median 129
Normalized Counts
Rg [nm]
F
TRF1F/F TRF2F/F Lig4-/- p53-/- (120 h) Expt 3
-Cre
-Cre fit (slope 0.340)
+Cre
+Cre fit (slope 0.396)
ln(Rg) [nm]
ln(# signal localizations)
- Cre: mean 96±44 median 88
+Cre: mean 102±52 median 91
Normalized Counts
Rg [nm]

The measured convex hull volumes changed by 63%, 436%, 11%, 90%, and 187% (average 157%). The substantial changes in the complex hull volume observed in some of these experiments are consistent with the previous observation that TRF1/TRF2 double knockdown induces a substantial increase in telomere volume in human cells (*148*). However, it is worth noting that, regardless of whether a small or large change was measured in telomere Rg/volume, maximal DDR is reliably detected upon codeletion of TRF1 and TRF2, with TIF signal being present at nearly all telomeres (Fig. 7B; Fig. 8B); for example, the experiment with only 14% increase in average Rg showed TIFs at 96% of the telomeres, and the experiment with a greater change in average Rg (36%) showed TIFs at 87% of the telomeres (Fig. 8B).

In most of these experiments, the increase in the average Rg value was substantially influenced by a subset of telomeric foci that exhibited both a greater number of FISH signal localizations and larger Rg/volume values, which could potentially be caused by telomere fusion, apposition of telomeres in cells undergoing endoreduplication, and/or clustering of dysfunctional telomeres (see below), although we cannot exclude the possibility that chromatin decompaction may have also contributed to the observed increase in Rg/volume.

Figure 8. Telomere fusions and endoreduplication after shelterin removal

(*A*) TRF2 immunoblot to verify the efficacy of Cre treatment in TRF1$^{F/F}$ TRF2$^{F/F}$ Lig4$^{-/-}$ p53$^{-/-}$ Rosa-Cre-ERT2 MEFs at 96 and 120 h after addition of tamoxifen. Cntrl: non-specific band used as loading control. Asterisks, non-specific bands. (*B*) Determination of the TIF response in the cells at the indicated time points. Green, telomeric FISH with TelG-A647; red, IF for γ-H2AX. Merged: green and red channels merged with DAPI DNA stain (blue). Quantification is given below the images. (*C*) Metaphases showing a low level of telomere fusions at 96 h and 120 h after induction of Cre. Quantification of the % of telomeres that are fused is indicated in the micrographs. (*D*) FACS analysis of the indicated cells to determine endoreduplication. Frequency of endoreduplication is approximated based on the % of cells with a DNA content >G2.


A
TRF1F/F TRF2F/F Lig4-/- p53-/-
96 h
120 h
- + - +
4-OHT
TRF2
cntrl
B
no 4-OHT
+ 4-OHT (96 h)
+ 4-OHT (120 h)
TTAGGG
γH2AX
merge+DAPI
<5%
96%
87%
% telomeres with γH2AX (n= 20 nuclei)
C
no 4-OHT 96 h
+ 4-OHT 96 h
2.5% telomeres fused (n=3000)
no 4-OHT 120 h
+ 4-OHT 120 h
4.9% telomeres fused (n=3000)
D
-4-OHT 96 h
<G1 2%
G1 46%
S/G2 47%
>G2 7%
+4-OHT 96 h
<G1 2%
G1 41%
S/G2 51%
>G2 12%
-4-OHT 120 h
<G1 1%
G1 47%
S/G2 41%
>G2 11%
+4-OHT 120 h
<G1 2%
G1 38%
S/G2 43%
>G2 16%
>G2
Count (10³)
PI Intensity (10⁶)

53BP1-dependent clustering of dysfunctional telomeres

To investigate the potential cause of the larger telomeric foci in the Cre-treated TRF1$^{F/F}$ TRF2$^{F/F}$ Lig4$^{-/-}$ MEFs described above, we first determined the extent of telomere fusion in these experiments (Fig. 8C). Analysis of metaphase spreads showed that ~5% of the telomeres had undergone fusion at the later (120-h) time point, which could contribute, but only in a small part, to the observed increase in telomere size. We also considered endoreduplication as a potential source of apparently larger telomeres. Endoreduplication takes place in cells with persistent dysfunctional telomeres and leads to close apposition of telomeric signals (*91*, *232*). FACS analysis showed that endoreduplication occurred in the experiment, potentially leading to telomere apposition in up to ~15% of the cells, but the difference between +Cre and −Cre samples was relatively small (Fig. 8D); hence, endoreduplication is unlikely to be a major contributing factor to the observed increase in telomere size.

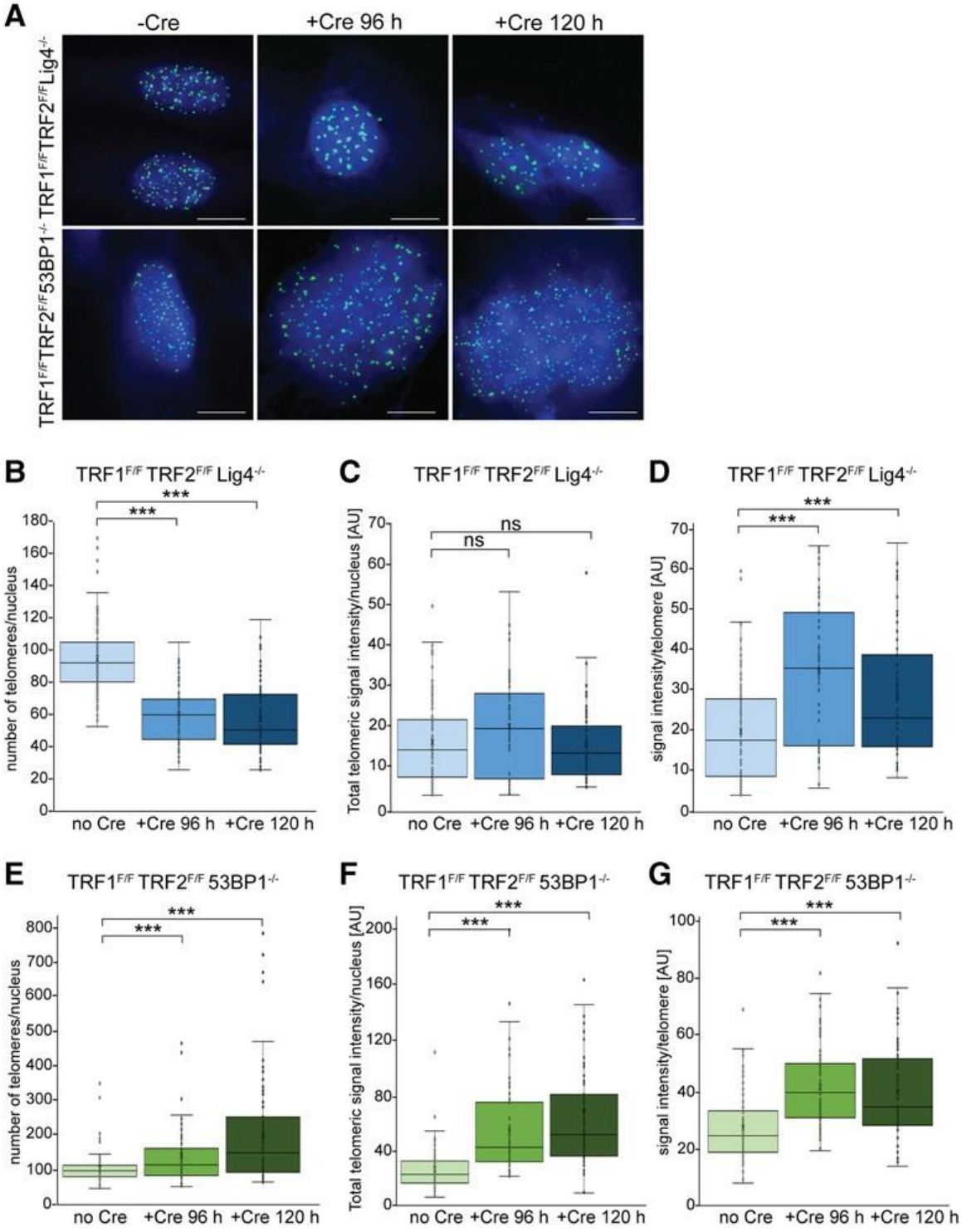
We next considered the clustering of dysfunctional telomeres as a source of the phenomenon observed. DSBs induced by a variety of treatments have been shown to undergo clustering, but the mechanism underlying these associations has not been determined (*233–237*). We argued that if dysfunctional telomeres become clustered, the number of telomeres detected per nucleus should decrease after treatment of the TRF1$^{F/F}$ TRF2$^{F/F}$ Lig4$^{-/-}$ MEFs with Cre. Indeed, the number of telomeric loci detected by FISH showed a substantial reduction at 96 and 120 h after Cre-mediated removal of shelterin (Fig. 9A). Compared with nuclei in untreated cells, nuclei in cells at 96 and 120 h after Cre treatment contained 37%

and 40% fewer telomeric foci, respectively (Fig. 9B), and the average FISH intensity per telomere increased after shelterin removal (Fig. 9C, D), consistent with clustering of dysfunctional telomeres.

As 53BP1 has been implicated in holding distant DSBs together within the context of immunoglobulin gene class switch recombination, we asked whether 53BP1 is involved in the clustering of dysfunctional telomeres (*238*). To this end, we used TRF1$^{F/F}$ TRF2$^{F/F}$ MEFs that lacked 53BP1 (Fig. 9A). As 53BP1 is required for the NHEJ of dysfunctional telomeres, the TRF1F/F TRF2F/F 53BP1$^{-/-}$ MEFs do not develop confounding telomere fusions (*82*, *239*). The data on these 53BP1-deficient cells showed that after Cre-mediated removal of shelterin, the reduction in the number of telomeric loci did not occur (Fig. 9A, 7E–G). In fact, the number of telomeric loci detected increased, potentially due to endoreduplication, which appears to be frequent (~30% of cells with 8N DNA content at 96 h) in the Cre-treaded TRF1$^{F/F}$ TRF2$^{F/F}$ 53BP1$^{-/-}$ MEFs (*82*). The appearance of some telomeric foci with larger sizes in the Cre-treated cells could be due to close apposition of telomeres, which tends to occur in cells undergoing endoreduplication. Overall, these results showed that dysfunctional telomeres lacking shelterin undergo considerable 53BP1-mediated clustering. This clustering can at least in part explain the occurrence of telomeric loci with a larger Rg/volume and a higher number of telomeric signal localizations after shelterin removal.

Figure 9. Clustering of dysfunctional telomeres.

(A) Examples of telomeric FISH used to count the number of detectable telomeric foci in the indicated TRF1$^{F/F}$ TRF2$^{F/F}$ Lig4$^{-/-}$ p53$^{-/-}$ Rosa-Cre-ERT2 and TRF1$^{F/F}$ TRF2$^{F/F}$ 53BP1$^{-/-}$ p53$^{-/-}$ Rosa-Cre-ERT2 MEFs with and without Cre treatment for the indicated times. Note that the Cre-treated cells lacking 53BP1 show evidence of endoreduplication (larger nuclei and greater number of telomeres). (B) Box and whisker plots of the number of telomeric foci detected per nucleus for the TRF1$^{F/F}$ TRF2$^{F/F}$ Lig4$^{-/-}$ p53$^{-/-}$ Rosa- Cre-ERT2 MEFs before and after treatment with tamoxifen for the indicated times. Data were obtained from three independent experiments, with >50 nuclei scored in each experiment. Imaging was done with 20 0.2-µm z-stacks with settings that were identical for minus and plus Cre samples of the same genotype. Detectable telomeric foci were counted in deconvolved collapsed z-stacks. All minus and plus Cre samples were processed in parallel. (∗∗∗) $P < 0.0001$, derived from a two-tailed Mann-Whitney U-test. (C) Box and whisker plots of the total telomere intensity per nucleus in TRF1$^{F/F}$ TRF2$^{F/F}$ Lig4$^{-/-}$ p53$^{-/-}$ Rosa-Cre-ERT2 cells before and after treatment with tamoxifen for the indicated times. Data were obtained on the data set as in B. The total FISH intensity was integrated over the whole nucleus as identified by DAPI. P-values are as in B. (D) Box and whisker plots of the FISH signal intensity per telomeric focus in TRF1$^{F/F}$ TRF2$^{F/F}$ Lig4$^{-/-}$ p53$^{-/-}$ Rosa-Cre-ERT2 cells. Data were obtained on the data set as in B. For each detectable telomeric focus, the total FISH intensity was determined. P-values are as in B. (E–G) Data are as in B–D but for TRF1$^{F/F}$ TRF2$^{F/F}$ 53BP1$^{-/-}$ p53$^{-/-}$ Rosa-Cre-ERT2 MEFs.

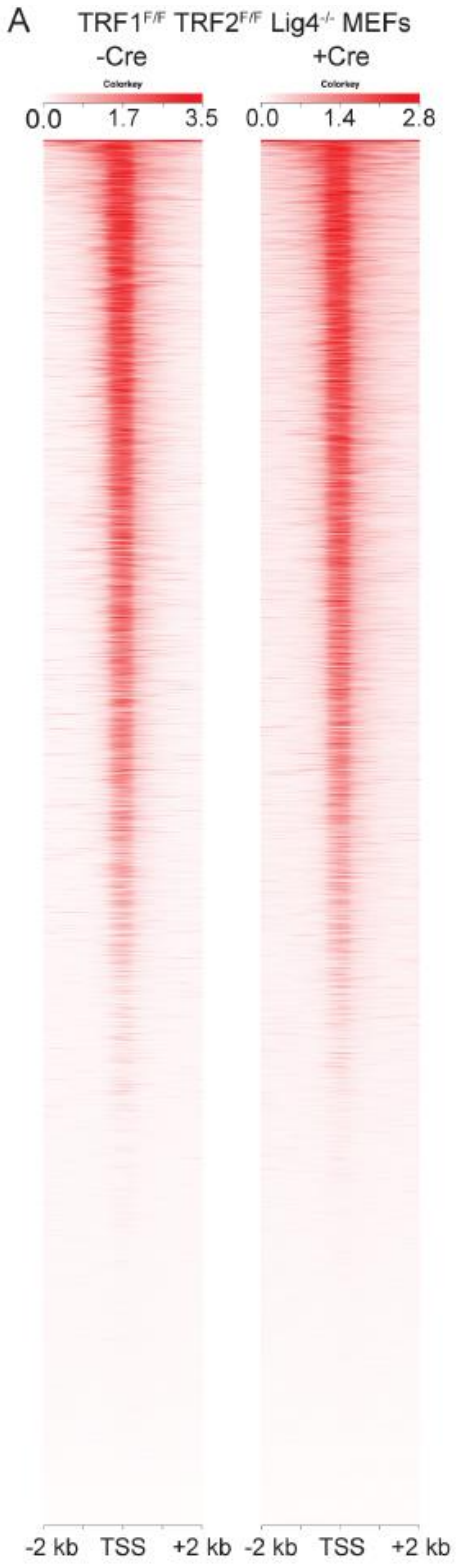

A
-Cre
+Cre 96 h
+Cre 120 h
TRF1F/F TRF2F/F Lig4-/-
TRF1F/F TRF2F/F 53BP1-/-
B
TRF1F/F TRF2F/F Lig4-/-
number of telomeres/nucleus
***
***
no Cre
+Cre 96 h
+Cre 120 h
C
TRF1F/F TRF2F/F Lig4-/-
Total telomeric signal intensity/nucleus [AU]
ns
ns
no Cre
+Cre 96 h
+Cre 120 h
D
TRF1F/F TRF2F/F Lig4-/-
signal intensity/telomere [AU]
***
***
no Cre
+Cre 96 h
+Cre 120 h
E
TRF1F/F TRF2F/F 53BP1-/-
number of telomeres/nucleus
***
***
no Cre
+Cre 96 h
+Cre 120 h
F
TRF1F/F TRF2F/F 53BP1-/-
Total telomeric signal intensity/nucleus [AU]
***
***
no Cre
+Cre 96 h
+Cre 120 h
G
TRF1F/F TRF2F/F 53BP1-/-
signal intensity/telomere [AU]
***
***
no Cre
+Cre 96 h
+Cre 120 h

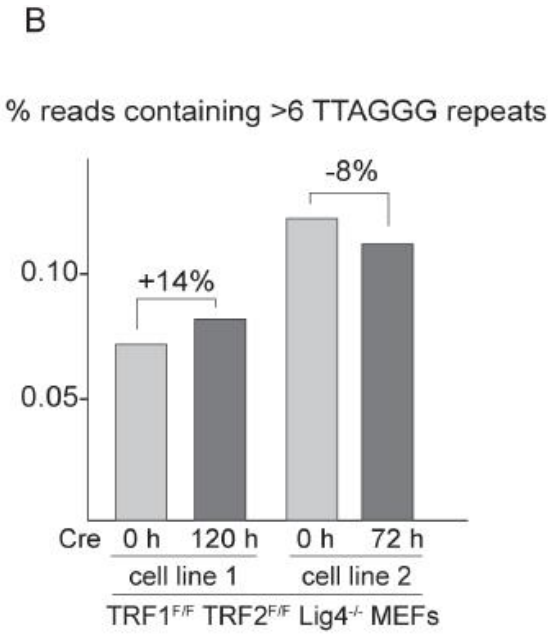
Removal of shelterin did not change telomeric chromatin accessibility significantly

The compaction state of chromatin has been studied with DNaseI and MNase, revealing a greater rate of digestion by these enzymes of DNA in “open” chromatin. Telomeric DNA in mammals is packaged in nucleosomes that appear to be present throughout the telomeric repeat array (*240*, *241*). The MNase sensitivity of telomeric DNA was found to be unaltered upon removal of shelterin or shelterin subunits, potentially indicating that the chromatin compaction of telomeres is not altered (*79*, *82*, *242*). However, it could be argued that the small size of MNase (17 kDa) might allow the enzyme access to the telomeric DNA regardless of its compacted state. We therefore turned to ATAC-seq, which reveals the accessibility of chromatin to the much larger 100kDa dimeric Tn5 transposase (*243*).

Figure 10. ATAC-seq to determine accessibility of telomeric chromatin

(*A*) ATAC-seq on TRF1F/F TRF2F/F Lig4-/- p53-/- Rosa-Cre-ERT2 MEFs before and after (120 h) treatment with tamoxifen to induce removal of shelterin. The expected preferential integration of Tn5 into regions surrounding the transcription start site (TSS) serves as a positive control for the efficacy of the ATAC-seq. (*B*) Bar graph representing the percentage of total paired-end reads (50-100 million) that contained at least 7 tandem TTAGGG (or CCCTAA) repeats in the indicated cells with and without Cre treatment for the indicated time. Cell lines 1 and 2 are independent MEFs lines with the same genotype. Between 50 and 100 million reads were obtained in each ATAC-seq experiment.


A
TRF1F/F TRF2F/F Lig4-/- MEFs
-Cre
+Cre
Colorkey
0.0 1.7 3.5
Colorkey
0.0 1.4 2.8
-2 kb TSS +2 kb
-2 kb TSS +2 kb


B
% reads containing >6 TTAGGG repeats
-8%
+14%
0.10
0.05
Cre 0 h 120 h 0 h 72 h
cell line 1
cell line 2
TRF1F/F TRF2F/F Lig4-/- MEFs

We performed ATAC-seq on nuclei from TRF1$^{F/F}$ TRF2$^{F/F}$ Lig4$^{-/-}$ MEFs before and after Cre-mediated deletion of TRF1 and TRF2. The data obtained showed the expected preferential insertion of Tn5 into transcription start sites (TSSs) (Fig. 10A), providing a control for the efficacy of the experiments. To evaluate the accessibility of telomeric chromatin, we determined the ratio of sequence reads with at least seven tandem TTAGGG (or CCCTAA) repeats to the total number of reads in each sample (Fig. 10B). The percentage of telomeric reads in the (−Cre) control samples (0.07% and 0.11%) is consistent with the approximate abundance of telomeric DNA in these MEFs (~0.05%–0.15% of the genome, depending on the length of the telomeres in the MEF line). In two independent ATAC-seq experiments, the Cre-induced deletion of TRF1 and TRF2 resulted in only a minor change (+14% and −8%) in the percentage of telomeric DNA reads (Fig. 10B), whereas the Cre induction conditions used in both experiments reliably caused DDR signaling at most telomeres. Such minor changes in the ATAC-seq results are unlikely to be meaningful, since ATAC-seq of repressed versus active chromatin usually shows large (more than fourfold) differences (*243*). Thus, based on Tn5 integration, the accessibility of the telomeric chromatin does not appear to be substantially altered upon shelterin removal.

Discussion

Here, we probed telomere decompaction in MEFs under three knockout conditions: TRF1 deletion, TRF2 deletion, and deletion of both TRF1 and TRF2, which removes the whole shelterin complex. These experiments were performed

under conditions where confounding telomere–telomere fusions are minimized. In all cases, the complete loss of the specified proteins induced by Cre was reliably observed. In the experiments where TRF1 or TRF2 was deleted, the majority of our experiments showed only small changes in telomere volume, and, in the experiments where both TRF1 and TRF2 were deleted, we observed varying degrees of telomere volume increase from experiment to experiment, ranging from minimal to substantial, yet all of these knockout conditions reliably induced a DDR at most telomeres in all experiments.

In the experiments where a substantial increase in average telomere volume was observed, the changes were influenced by a subset of the telomeric foci detected by STORM imaging that showed both greater Rg/volume values and a greater number of FISH signal localizations. We provide evidence that such large telomeric foci are likely formed at least in part through 53BP1-dependent clustering of dysfunctional telomeres, although we cannot exclude the possibility that chromatin decompaction also contributed to the observed telomere volume increase.

Consistent with our results on mouse cells, an accompanying study by Vancevska et al. (2017) shows that the majority of deprotected telomeres (i.e., telomeres exhibiting TIF signals), upon shelterin inhibition in human cells, also did not show substantial decompaction (*244*). Taken together, these results suggest that DDR at telomeres caused by removal of shelterin subunits does not require substantial chromatin decompaction; hence, chromatin compaction is unlikely to be a primary mechanism used by shelterin to repress DNA repair or DNA damage

signaling at telomeres.

Several additional observations support our notion above. First, in vitro studies with the six-subunit mouse shelterin complex have not revealed shelterin–shelterin or shelterin–DNA interactions that could mediate the proposed compaction (*53*). Second, the DDR is activated when DSBs are made inside the telomeric repeat array (*125*, *245*). According to the compaction model, the telomere compaction should also protect the DNA ends of the induced DSBs. To explain the activation of the DDR by DSBs inside the telomeric repeat array, the compaction model for protecting telomeres would require modification. For instance, a mechanism could be proposed that places the telomere terminus at an internal position in the compacted domain, thereby providing greater protection (*148*). Third, the compaction model does not readily explain how large protein complexes, such as the ~600kDa MRN complex, which acts as a shelterin-bound accessory factor at functional telomeres, can enter the telomeric chromatin when shelterin is intact (*151*). Finally, when shelterin is removed, the telomeric chromatin does not display a greater sensitivity to MNase, and chromatin accessibility as measured by ATAC-seq is also not significantly changed (this study) (*82*). However, our data do not exclude a role for chromatin compaction in slowing the DDR at telomeres. If such protective compaction occurs, it would likely be enforced primarily by factors other than shelterin and its associated proteins.

Chapter 3

TRF2 Is Sufficient To Form T-Loops In Vivo

Telomeric repeat binding factor 2 (TRF2) is a critical component of the telomere bound shelterin complex, responsible for suppressing ATM, c-NHEJ, and forming t-loops. T-loops are thought to protect the telomere end by sequestering the 3' end into a D-loop to protect it from ATM, Ku70/80 and MRN. We wished to address two issues regarding the function of TRF2: first, to determine whether TRF2 is sufficient to generate t-loops, and second, to investigate whether t-loops and TRF2 are sufficient to repress c-NHEJ, ATM, and alt-NHEJ. Here we show that cells containing only TRF2 at their telomeres showed the presence of t-loops, indicating that TRF2 is not only required, but also sufficient for t-loop formation. By creating shelterin free telomeres and replacing endogenous TRF2 with a mutant unable to recruit the rest of the shelterin complex, we demonstrated that TRF2 can partially repress the ATM response and protect telomeres from c-NHEJ. In addition, using an ATRi inhibitor in combination with cells lacking shelterin at their telomeres, we were able to show that TRF2 is the dominant factor in repressing ATM at telomeres. We also demonstrate that t-loops are not sufficient to protect telomeres from alternative non-homologous end-joining, though whether they offer any protection from alt-NHEJ has not been explored.

Introduction

To create telomeres containing TRF2 as the only shelterin component, we used complementation of $TRF1^{F/F}$ $TRF2^{F/F}$ MEFs with an allele of TRF2 that is incapable of binding to TIN2 and Rap1. Cells with a deletion of both TRF1 and TRF2 are referred to as having shelterin free telomeres, because their telomeres

lack all shelterin components and shelterin interacting factors (*82*). TRF1 and TRF2 can both independently anchor the shelterin complex, but loss of both of these factors prevents TIN2, TPP1, and POT1 from localizing to telomeres. These cells exhibit striking phenotypes associated with total deprotection of the telomere. When TRF1 and TRF2 are missing from telomeres these cells show telomere end fusions by c-NHEJ, as well as both ATM- and ATR-dependent DNA damage signaling (*82*). If Ku70/80 is missing in addition to TRF1 and TRF2, then alt-NHEJ becomes the predominant mechanism of telomere fusions because Ku70/80 is a general repressor of alt-NHEJ as well as being required for c-NHEJ (*82*). In the absence of Lig4, telomeres exhibit a low level of alt-NHEJ fusions and no c-NHEJ because Lig4 is required for c-NHEJ. (*82*).

Despite the fact that TRF2 is known to be involved in many of the end-protection processes at telomeres, we still do not have a clear understanding of whether or not TRF2 is capable of these functions on its own. In this series of experiments, we set out to test whether or not TRF2 is capable of performing many of the functions ascribed to it without the help of the rest of the shelterin complex, using shelterin-free telomeres and a mutant version of TRF2 lacking interaction sites with the rest of shelterin as a model. We test whether TRF2 is sufficient to make t-loops, whether it is sufficient to repress ATM, and whether it is sufficient to repress alt-NHEJ.

Results

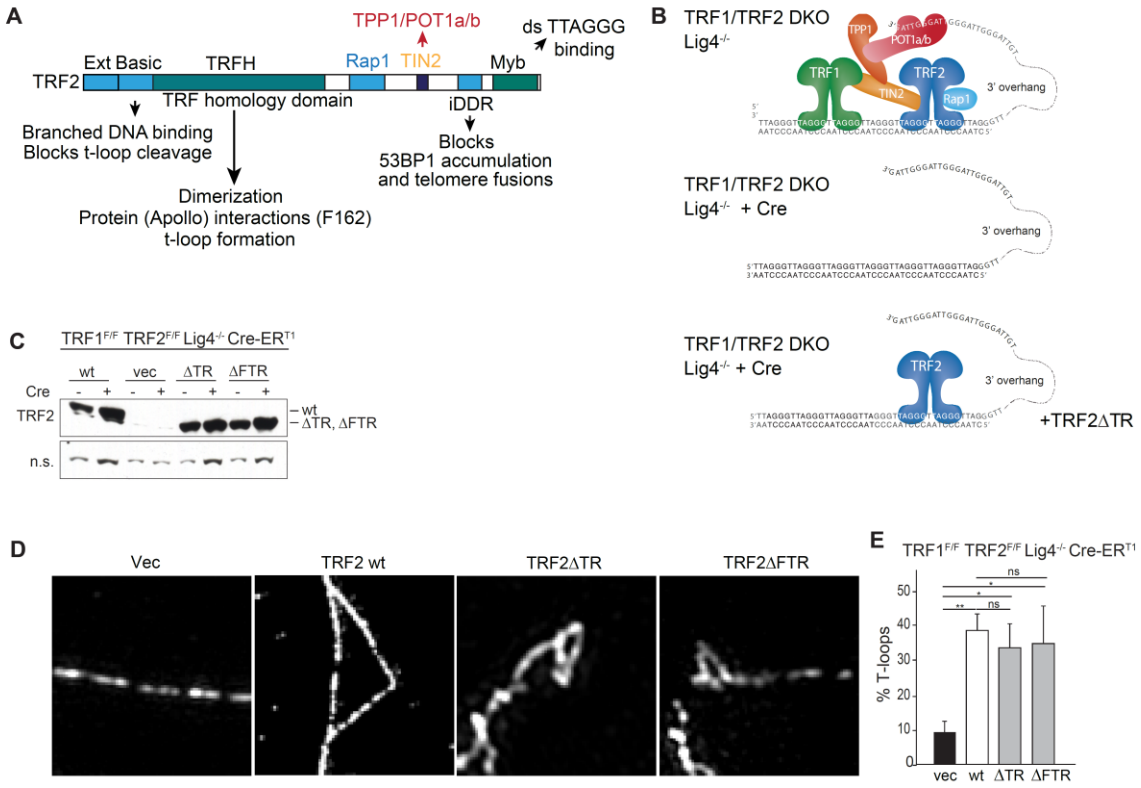
TRF2 Is Sufficient to Form T-loops

We used complementation of TRF1$^{F/F}$ TRF2$^{F/F}$ p53$^{-/-}$ Cre-ER^{T1} MEFs, which generate shelterin-free telomeres when gene deletion is induced using Cre-recombinase (Fig. 11C). These telomeres are chromatinized but are not bound by any members of the shelterin complex. For complementation we used a mutant form of TRF2 that lacks the TIN2 interaction site (amino acids 352 – 367) followed by deletion of endogenous TRF1 and TRF2, resulting in telomeres containing TRF2 and RAP1 but no other shelterin components (Fig. 11A,B) (*73*). Without the bridging properties of TIN2, TRF2 is unable to anchor the rest of the shelterin complex. In addition to removing the TIN2 interface, we removed the Rap1 binding site (amino acids 284 – 297) so that Rap1 would not be recruited to telomeres. We called this mutant TRF2△TR. Finally, we also created a mutant (TRF2△FTR) that also lacks the critical phenylalanine at position 162 where Apollo binds (Fig. 11A).

We wanted to test whether the rest of shelterin is required for t-loop formation, and whether t-loops were required for the suppression of ATM signaling and c-NHEJ. To determine whether TRF2 is sufficient for t-loop formation, the two mutant alleles as well as the wild type TRF2 were introduced into TRF1$^{F/F}$ TRF2$^{F/F}$ Lig4$^{-/-}$ Cre-ER^{T1} MEFs and deletion of TRF1 and TRF2 was induced by recombination between loxP sites driven by a tamoxifen inducible Cre-recombinase. Cre-recombinase expression was transiently driven by addition of tamoxifen to the medium for 12 hours.

Figure 11. TRF2 Alone is Sufficient to Form T-loops.

(A) Schematic of wild type TRF2 and the TIN2, Rap1, and Apollo interacting sites. (B) Schematic of telomeres after loss of TRF1 and TRF2, causing the telomeres to become shelterin free. Replacing TRF2 with interaction-deficient mutants of TRF2 allows one to test which functions TRF2 is sufficient for. (C) Western blot showing expression of TRF2, TRF2ΔTR, and TRF2ΔFTR in MEFs immortalized with p53-loss and floxed for TRF1 and TRF2 and deficient in Lig4. TRF2 in the top panel, n.s. non-specific loading control in the lower panel. (D) Representative images of DNA spreads showing t-loops in $TRF1^{F/F}$ $TRF2^{F/F}$ $Lig4^{-/-}$ MEFs, imaged using super-resolution OMX microscopy, 72 hours after tamoxifen induced Cre deletion. (E) Quantification of H, n = 3 independent experiments, 100 countable molecules per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$).

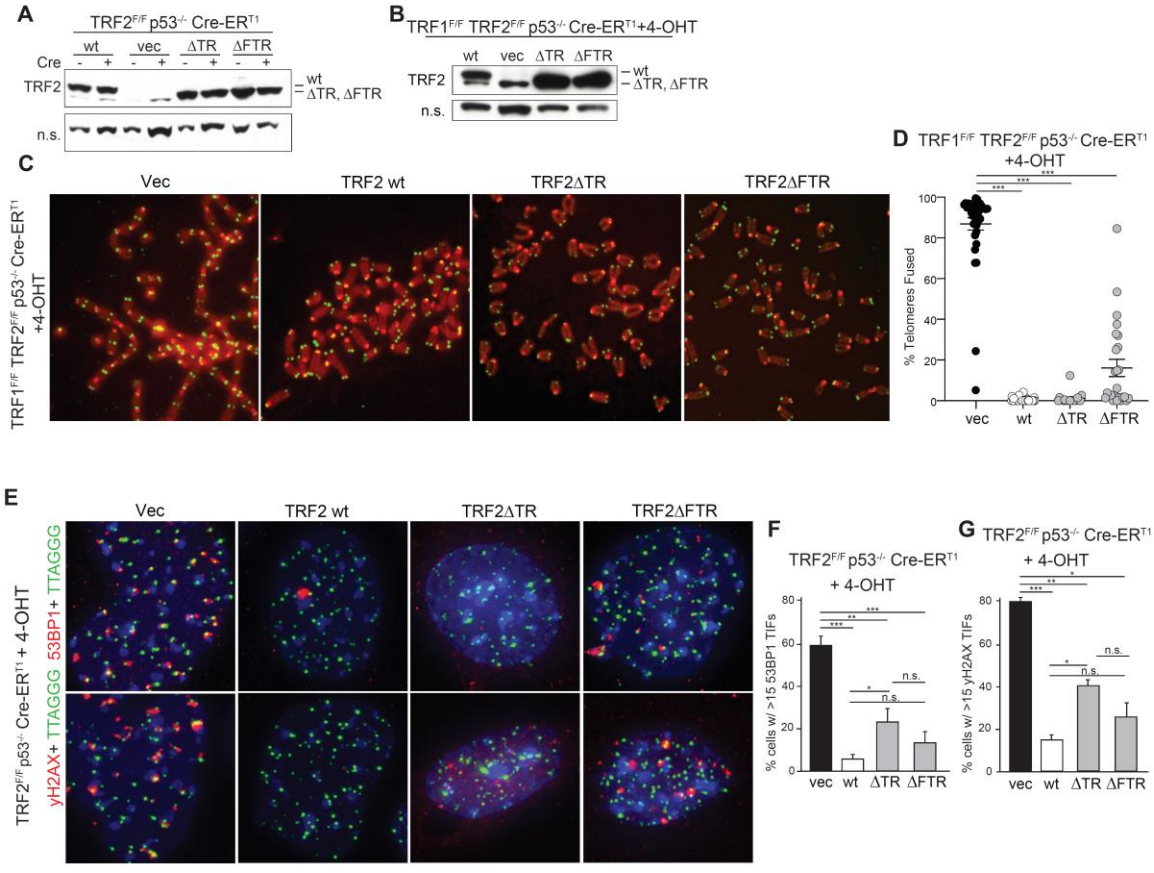

A
TPP1/POT1a/b
ds TTAGGG binding
Ext Basic
TRFH
Rap1
TIN2
Myb
TRF2
TRF homology domain
iDDR
Branched DNA binding
Blocks t-loop cleavage
Blocks
53BP1 accumulation
and telomere fusions
Dimerization
Protein (Apollo) interactions (F162)
t-loop formation
B
TRF1/TRF2 DKO
Lig4-/-
TPP1
POT1a/b
TRF1
TRF2
TIN2
Rap1
3' overhang
TRF1/TRF2 DKO
Lig4-/- + Cre
3' overhang
TRF1/TRF2 DKO
Lig4-/- + Cre
TRF2
3' overhang
+TRF2ΔTR
C
TRF1F/F TRF2F/F Lig4-/- Cre-ERT1
wt
vec
ΔTR
ΔFTR
Cre
TRF2
wt
ΔTR, ΔFTR
n.s.
D
Vec
TRF2 wt
TRF2ΔTR
TRF2ΔFTR
E
TRF1F/F TRF2F/F Lig4-/- Cre-ERT1
ns
*
*
**
ns
% T-loops
0
10
20
30
40
50
vec
wt
ΔTR
ΔFTR

For this experiment Lig4$^{-/-}$ cells were used to avoid the complication of telomere fusions. Telomere fusions can confound the scoring of t-loops because end-fusions would remove the t-loop (*138*). Since Ligase 4 is required for c-NHEJ, this would greatly reduce the number of fusions seen in our experiment and would allow us to more accurately assay for t-loop formation.

The TRF2△TR and TRF2△FTR mutant constructs were expressed at equal levels to the TRF2 wild type protein (Fig. 11C). However, all three constructs are highly over-expressed compared to endogenous TRF2. Endogenous TRF2 is difficult to detect in these by western blots and cannot be seen in the vector lanes. It is unlikely that high expression of TRF2 affects the results of a t-loop assay, as the endogenous levels of TRF2 are already sufficient to coat the telomere, and overexpression would not be expected to bring more TRF2 to the telomere (*74*). However, it is possible that a hypomorph in DNA binding or end-protection might be able to function if overexpressed because it would drive the protein equilibrium towards a telomere bound state.

Figure 12. TRF2 Can Suppress c-NHEJ and ATM

(A) Western blot showing expression of TRF2, TRF2△TR, and TRF2△FTR in MEFs immortalized with p53-loss and floxed for TRF1 and TRF2. TRF2 in the top panel, n.s. non-specific loading control in the lower panel. (B) Western blot showing expression of TRF2, TRF2△TR, and TRF2△FTR in MEFs immortalized with p53-loss and floxed for TRF2. TRF2 in the top panel, n.s. non-specific loading control in the lower panel. (C) Representative images of metaphases in $TRF1^{F/F}$ $TRF2^{F/F}$ MEFs, 72 hours after Cre-induction with tamoxifen. Red – DAPI. Green – Alexa647-$(TTAGGG)_3$. (D) Quantification of F, n = 3 independent experiments, 10 metaphases per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * p > 0.05, ** p >0 .01, *** p > 0.001). (E) Representative images of TIFs in $TRF2^{F/F}$ MEFs 72 hours after Cre-induction with tamoxifen. Top panel: Blue – DAPI. Green – Telomeres. Red – 53BP1. Bottom panel: Blue – DAPI. Green – Telomeres. Red – γH2AX. (F) Quantification of E, showing cells with 15 or more 53BP1 TIFs. n = 3 independent experiments, with 100 cells each. Significance is shown using a t-test where relevant. (G) Quantification of E, showing cells with 15 or more γH2AX TIFs. n = 3 independent experiments, 100 cells per replicate. Significance is determined using a two-tailed unpaired t-test where relevant (n.s. not significant, * p > 0.05, ** p > 0.01, *** p > 0.001).


A
TRF2F/F p53-/- Cre-ERT1
wt
vec
ΔTR
ΔFTR
Cre
TRF2
n.s.
wt
ΔTR, ΔFTR
B
TRF1F/F TRF2F/F p53-/- Cre-ERT1+4-OHT
wt
vec
ΔTR
ΔFTR
TRF2
n.s.
wt
ΔTR, ΔFTR
C
Vec
TRF2 wt
TRF2ΔTR
TRF2ΔFTR
TRF1F/F TRF2F/F p53-/- Cre-ERT1 +4-OHT
D
TRF1F/F TRF2F/F p53-/- Cre-ERT1 +4-OHT
% Telomeres Fused
vec
wt
ΔTR
ΔFTR
E
Vec
TRF2 wt
TRF2ΔTR
TRF2ΔFTR
TRF2F/F p53-/- Cre-ERT1 + 4-OHT
53BP1+ TTAGGG
γH2AX+ TTAGGG
F
TRF2F/F p53-/- Cre-ERT1 + 4-OHT
% cells w/ >15 53BP1 TIFs
vec
wt
ΔTR
ΔFTR
G
TRF2F/F p53-/- Cre-ERT1 + 4-OHT
% cells w/ >15 γH2AX TIFs
vec
wt
ΔTR
ΔFTR

T-loop formation was scored using the protocol developed by Doksani et al. in which nuclei are isolated and treated with psoralen and UV to crosslink the DNA (see methods) (*138*). Subsequently the chromatin is spread and t-loops are visualized with a $CCCTAA_3$-Alexa488 PNA probe and super resolution imaging using OMX. Cells infected with an empty vector had significantly reduced levels of t-loops compared to cells expressing wild type TRF2 (Fig. 11D,E). Both the TRF2△TR and TRF2△FTR mutant were able to sustain t-loop formation at the same level as wild type TRF2. These data indicate that TRF2 is sufficient to form t-loops.

TRF2 Is Sufficient to Protect Against c-NHEJ

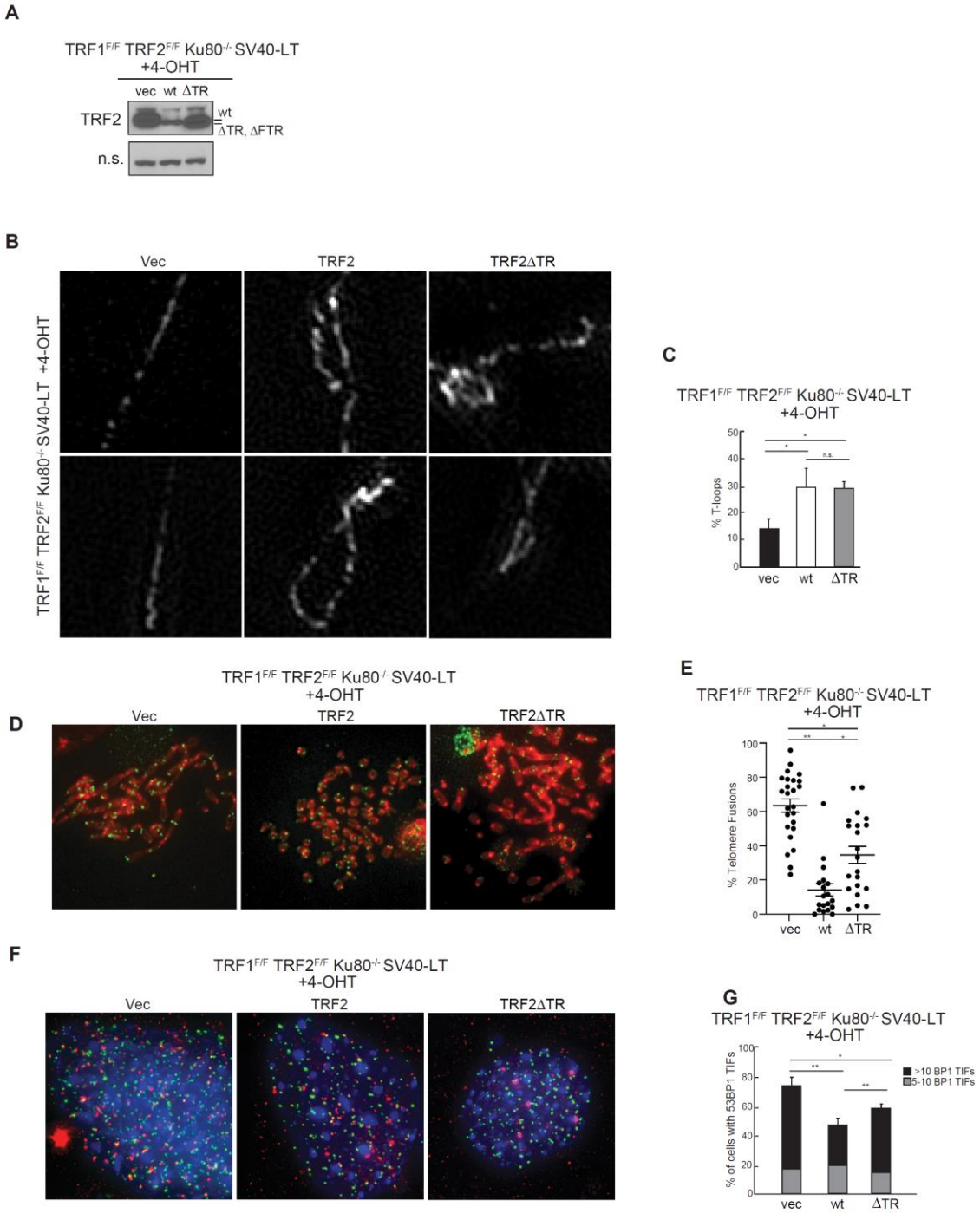
Having established that TRF2 is sufficient to form t-loops we wished to determine whether TRF2 is also sufficient to protect against telomere fusions and ATM signaling. To accomplish this, we analyzed the level of telomere fusions visible on metaphase spreads in $TRF1^{F/F}$ $TRF2^{F/F}$ cells expressing TRF2△TR or TRF2△FTR. The expression level of the TRF2 alleles was high and approximately equal, which is critical, as low levels of TRF2 can lead to a partial telomere fusion phenotype. 72 hours after Cre induction, we observed a significant number of fusions in the empty vector control, and no (<1%) fusions in wild type TRF2 and TRF2△TR (Fig. 12A-C). Performing this experiment at the earlier 72 hour timepoint minimizes the presence of chromatid-type fusions due to the failure of the TRF2△FTR to properly recruit Apollo for overhang processing (*11*). We observed a significant suppression of telomere fusions in the TRF2△FTR mutant, consistent

with the hypothesis that TRF2 alone is capable of suppressing c-NHEJ. The slight elevation of fusions in the TRF2△FTR mutant is consistent with the role of Apollo in preventing chromatid-type fusions on the leading strand.

Finally, we tested the ability of TRF2△TR and TRF2△FTR to block ATM. To do this, we expressed TRF2△TR and TRF2△FTR in $TRF2^{F/F}$ cells. We chose to use $TRF2^{F/F}$ cells because loss of other shelterin components (such as TRF1, TIN2, TPP1, or POT1) would lead to ATR-mediated DNA damage signaling, which would confound the interpretation of any TIFs we observe and make it impossible to conclude whether or not TRF2 can suppress ATM. At 72 hours after TRF2 deletion, we performed IF-FISH on mouse embryonic fibroblasts and assayed for TIFs. We tested both for 53BP1, and the upstream chromatin mark γH2AX. At 72 hours, cells not expressing any TRF2 had highly elevated levels of TIFs compared to cells expressing wild type TRF2, TRF2△TR or TRF2△FTR (Fig. 12E-G). Although we observed a slightly elevated level of TIFs in the two TRF2 mutants compared to wild type TRF2, the difference was only significant between wild type TRF2 and TRF2△TR; there was no significant difference in the level of γH2AX or 53BP1 TIFs between TRF2△TR and TRF2△FTR. These data suggest that TRF2△TR and TRF2△FTR are both capable of suppressing the ATM mediated DNA damage response.

Figure 13. T-loops are not sufficient to prevent alt-NHEJ.

(A) Western blot showing expression of wild type TRF2, or TRF2ΔTR (top panel) or a non-specific loading control (bottom panel) in TRF1$^{F/F}$ TRF2$^{F/F}$ Ku80$^{-/-}$ SV40-LT MEFs 96 hours after addition of tamoxifen. (B) Representative images of DNA spreads showing t-loops in TRF1$^{F/F}$ TRF2$^{F/F}$ Ku80$^{-/-}$ SV40-LT MEFs, imaged using super-resolution OMX microscopy, 96 hours after tamoxifen induced Cre deletion. (C) Quantification of B, n = 3 independent experiments, 100 countable molecules per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$). (D) Representative images of metaphases in TRF1$^{F/F}$ TRF2$^{F/F}$ Ku80$^{-/-}$ MEFs, 96 hours after Cre-induction with tamoxifen. Red – DAPI. Green – Alexa647-(TTAGGG)3 (E) Quantification of D, n = 3 independent experiments, 2000 telomere ends per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$). (F) Representative images of TIFs in TRF1$^{F/F}$ TRF2$^{F/F}$ Ku80$^{-/-}$ SV40-LT MEFs 96 hours after Cre-induction with tamoxifen. Top panel: Blue – DAPI. Green – Telomeres. Red – 53BP1. Bottom panel: Blue – DAPI. Green – Telomeres. Red – γH2AX. (G) Quantification of F, showing cells with 5-10 or 10 or more 53BP1 TIFs. n = 3 independent experiments, 100 cells per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$).


A
TRF1F/F TRF2F/F Ku80-/- SV40-LT
+4-OHT
vec wt ΔTR
TRF2
wt
ΔTR, ΔFTR
n.s.
B
Vec
TRF2
TRF2ΔTR
TRF1F/F TRF2F/F Ku80-/- SV40-LT +4-OHT
C
TRF1F/F TRF2F/F Ku80-/- SV40-LT
+4-OHT
% T-loops
vec wt ΔTR
D
TRF1F/F TRF2F/F Ku80-/- SV40-LT
+4-OHT
Vec
TRF2
TRF2ΔTR
E
TRF1F/F TRF2F/F Ku80-/- SV40-LT
+4-OHT
% Telomere Fusions
vec wt ΔTR
F
TRF1F/F TRF2F/F Ku80-/- SV40-LT
+4-OHT
Vec
TRF2
TRF2ΔTR
G
TRF1F/F TRF2F/F Ku80-/- SV40-LT
+4-OHT
% of cells with 53BP1 TIFs
>10 BP1 TIFs
5-10 BP1 TIFs
vec wt ΔTR

However, one critical complication in this experiment is that TRF1 can bind TIN2 and therefore bring the rest of the shelterin complex to the telomere. In addition, TIN2, which has not been removed, has been implicated in repressing ATM signaling (*60*). This prevents us from determining whether TRF2 is truly sufficient to suppress ATM in this setting. We will return to this issue later in the chapter.

TRF2 and T-loops are Not Sufficient to Block Alternative Non-homologous End-Joining

We previously tested whether TRF2 is sufficient to repress two different pathways at telomeres: c-NHEJ and ATM signaling. In the settings we examined, TRF2 was sufficient to repress c-NHEJ but not ATM signaling. We next wanted to test whether TRF2 is sufficient for suppression of alt-NHEJ.

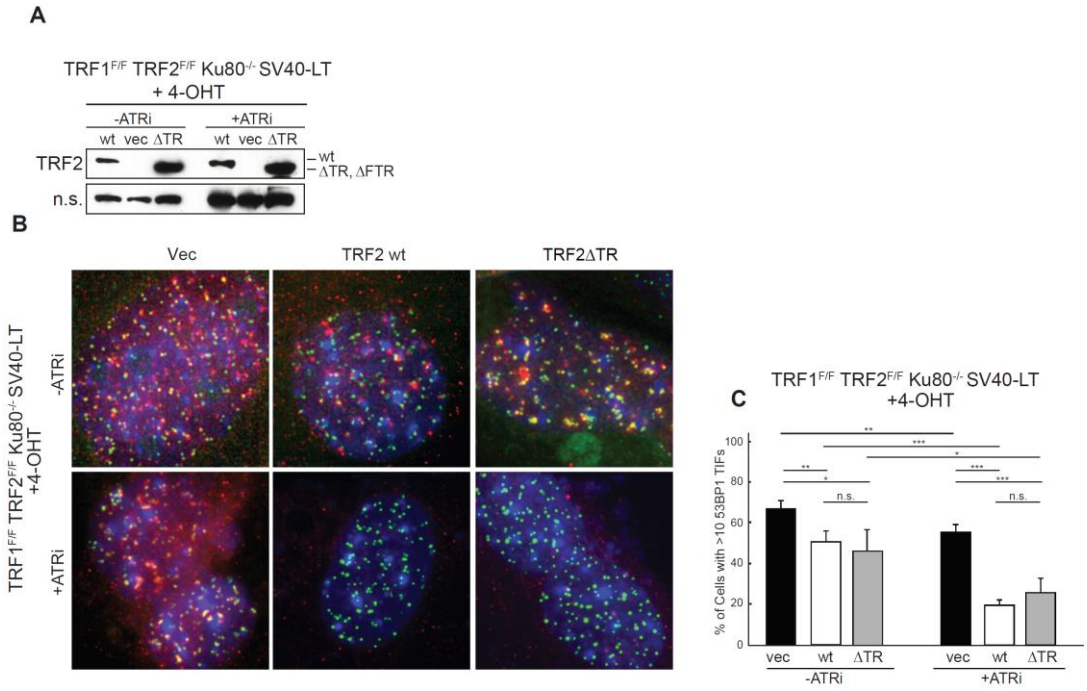
We used TRF1$^{F/F}$ TRF2$^{F/F}$ Ku80$^{-/-}$ MEFs and expressed an empty vector, TRF2, or TRF2ΔTR (Fig. 13A). When treated with Cre recombinase, these MEFs lose all shelterin proteins from their telomeres, and due to the loss of the c-NHEJ factor and alt-NHEJ suppressor Ku70/80, experience chromosome fusions that are dependent on Lig3, Pol θ and PARP1 (*82*, *109*, *122*, *124*, *125*, *156*). We first determined whether t-loops were present in these cells. After performing DNA spreading, we observed that both TRF2 and TRF2ΔTR had an equivalent level of t-loops (Fig. 13B,C). Interestingly, the overall level of t-loops was somewhat lower in this cell line (up to 30%), compared to the frequency found in other cells (up to 60%), however the significance of this difference is unclear. This may suggest that

Ku70/80 can play a role in t-loop formation and would be consistent with its interaction with TRF2, though how Ku70/80 would assist in t-loop formation is unclear.

In agreement with previous studies, there was a high level of telomere fusions in the empty vector control, with a partial repression of these alt-NHEJ fusions by wild type TRF2 (Fig. 13D, E). This suggests that TRF2 is not sufficient to repress alt-NHEJ. Furthermore, we observed that TRF2△TR was less capable of suppressing alt-NHEJ than wild type TRF2, suggesting that TIN2, TPP1, POT1 and/or Rap1 block alt-NHEJ. We also examined TIF formation, and unsurprisingly, neither TRF2 construct was able to fully repress the ATR and ATM signaling in this context (Fig. 13F,G). In the context of the fusions we observed, we would expect a reduced level of t-loops in TRF2△TR compared to wild type TRF2, but this was not the case, which suggests that fusions were occurring regardless of t-loop status.

Figure 14. TRF2 can suppress ATM signaling in the absence of other shelterin components.

(A) Western blot showing expression of wild type TRF2, and TRF2△TR (top panel) or a non-specific loading control (bottom panel) in $TRF1^{F/F}$ $TRF2^{F/F}$ $Ku80^{-/-}$ MEFs in the presence or absence of ATR inhibitor 120 hours after addition of tamoxifen. (B) Representative images of 53BP1 TIFs in $TRF1^{F/F}$ $TRF2^{F/F}$ $Ku80^{-/-}$ MEFs 120 hours after induction of Cre recombinase by tamoxifen and the addition of 2.5 μM ATR inhibitor. Blue – DAPI. Green – Telomeres. Red – 53BP1. (C) Quantification of B, showing cells with 5-10 or more than 10 53BP1 TIFs. n = 4 independent experiments, 100 cells per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$).


A
TRF1F/F TRF2F/F Ku80-/- SV40-LT
+ 4-OHT
-ATRi
+ATRi
wt vec ΔTR wt vec ΔTR
TRF2
n.s.
wt
ΔTR, ΔFTR
B
Vec
TRF2 wt
TRF2ΔTR
TRF1F/F TRF2F/F Ku80-/- SV40-LT
+4-OHT
-ATRi
+ATRi
C
TRF1F/F TRF2F/F Ku80-/- SV40-LT
+4-OHT
% of Cells with >10 53BP1 TIFs
0
20
40
60
80
100
**
***
*
**
*
n.s.
***
***
n.s.
vec
wt
ΔTR
vec
wt
ΔTR
-ATRi
+ATRi

TRF2 Is Not Sufficient to Repress ATM Activity at Telomere Ends

TRF2 is known to be the main repressor of ATM signaling at telomeres, since ATM activation occurs at all (or most) chromosome ends in its absence. In contrast, loss of Rap1, TRF1, POT1, or TPP1 does not lead to the activation of ATM at telomeres (*78*). Deletion of TIN2, however, shows a modest level of ATM activation at telomeres which cannot be explained from the mild loss of TRF2 since overexpression of TRF2 in TIN2-deficient cells does not fully repress this ATM response (*90*). In the setting of shelterin-free telomeres we can determine to what extent TRF2, in the absence of other shelterin components, has the ability to repress ATM signaling. Since this experiment needs to be performed under conditions where the telomeres do not become fused together, as fusions attenuate DDR signaling over time, we chose to express the TRF2△TR in TRF1$^{F/F}$ TRF2$^{F/F}$ Ku80$^{-/-}$ MEFs where fusions are less frequent.

However, this system has a complication in that loss of TRF1 induces ATR signaling, primarily through replication stress. Additionally, because TIN2 is absent, POT1 is removed from the telomeres, which exposes the telomeric overhang to ATR signaling. To counteract this, we treated cells with an ATR inhibitor, ETP-46464, concomitant with shelterin deletion by Cre-mediated recombination. As before, we used a TRF2 mutant, TRF2△TR, which does not recruit any other members of the shelterin complex to the telomere, allowing us to isolate the role of TRF2.

We expressed wild type TRF2 and TRF2ΔTR in TRF1$^{F/F}$ TRF2$^{F/F}$ Ku80$^{-/-}$ cells and then deleted the endogenous alleles of TRF1 and TRF2 by tamoxifen

induced Cre-mediated recombination. The cells were harvested 120 hours after gene deletion and were grown either with or without ATR inhibitor for 120 hours. We confirmed gene expression by western blotting and observed equal expression of TRF2 alleles in all conditions. (Fig. 14A). We then imaged the cells and scored the number of 53BP1 foci that colocalized with telomeres in each condition. In the absence of ATRi, there was strong DDR signaling in all conditions, consistent with the effect of TRF1 and POT1 loss at telomeres. However, DDR signaling was slightly attenuated in the wild type TRF2 and TRF2△TR conditions compared with the empty vector, potentially due to repression of ATM (Fig. 14B,C).

By comparing the level of TIFs in vector control with and without ATRi, it is possible to see that shelterin free telomeres are activating both ATM and ATR, through loss of TRF2 and recognition of the DSB structure, and exposure of the single-stranded overhang, respectively. In the ATR inhibitor treated conditions, there was a slight reduction in the TIF response in the vector control, however signaling remained high – an indication of ATM signaling at most telomeres. This is as we expected, and showed that there was no cross-inhibition of ATM by the ATRi.

When wild type TRF2 is expressed there is a modest effect of the ATRi – reducing DDR signaling by 60%. The residual TIFs are likely due to incomplete repression of ATR signaling by ATRi, because if the ATRi inhibitor was working fully, we would expect no TIFs in this condition.

Most importantly, in the cells expressing TRF2△TR the TIF response was substantially diminished by TRF2△TR, indicating that TRF2△TR has the ability to

repress ATM signaling. Further, although TRF2△TR had a slightly elevated level of signaling, there was no significant difference between the wild type TRF2 and TRF2△TR, suggesting that TRF2 has the ability to repress ATM largely on its own (Fig. 14B,C). The lack of a statistically significant difference between TRF2△TR and wild type TRF2 also suggests that TRF2△TR can fully complement wild type TRF2. This data supports a model in which TRF2 alone is capable of partially suppressing ATM signaling but requires TIN2 to fully block ATM signaling.

Discussion

We began this study by seeking to test whether TRF2 alone is sufficient to form t-loops. By utilizing shelterin free telomeres and expressing interaction deficient alleles of TRF2, we have shown that TRF2 is indeed both necessary and sufficient for t-loop formation. We have also reinforced the t-loop model of chromosome end protection by adding evidence that t-loops are the mechanism by which TRF2 acts to suppress c-NHEJ telomere fusions and attenuate ATM signaling.

The finding that TRF2 is the only member of the shelterin complex to form t-loops is confirmation of one of the central principles of the t-loop model. Previous data had suggested that TRF1 has the ability to loop DNA strands, but it was later shown that TRF1 is not required for t-loop formation. POT1, due to its ability to bind the ssDNA end was also thought to potentially play a role in stabilizing the t-loops, but work from this lab showed that loss of POT1 did not affect t-loop frequency (*138*). However, these experiments still left the possibility of redundancy

within the shelterin complex with multiple factors facilitating t-loop formation. We tested this and found that TRF2 was sufficient to form t-loops – as t-loops can be formed by TRF2 without the help of any other members of the shelterin complex. Although we did not directly demonstrate that no other proteins are present at telomeres, deletion of TRF1 and TRF2 concurrently has previously been shown to removal all of shelterin from the telomere (*82*).

This data also indicated that TRF2 alone is capable of repressing c-NHEJ. This result is fully consistent with what we have observed regarding t-loop formation, and with a model where TRF2 mediates this repression through forming t-loops. Other c-NHEJ fusions that have been seen in cells missing shelterin components are likely due to loss of TRF2 from telomeres: for instance, c-NHEJ occurs upon loss of TIN2 but has been shown to be due to loss of TRF2 because it can be suppressed by overexpression of TRF2 (*90*).

We were also able to show that TRF2 is required for repressing ATM activation at telomeres, and it is capable of partial repression of ATM signaling without other shelterin components. We have known that TRF2 represses ATM, presumably through t-loops, but in the past we have been unable to exclude the involvement of other shelterin factors (*78*, *86*, *128*). TRF2 alone was largely able to suppress the ATM response, but there was still a moderate level of signaling which may be the result of incomplete action of the ATR inhibitor and/or due to the loss of TIN2. Additionally, it is important to note that unlike the telomere fusion phenotype of TIN2, the TIF phenotype of TIN2 cannot be suppressed by

overexpression of TRF2, which lends more evidence towards the theory that TIN2 and TRF2 cooperate to suppress ATM signaling (*90*).

In addition, this work shows that TRF2 not only is unable to fully suppress alt-NHEJ, but that t-loops are not sufficient to suppress alt-NHEJ as well, as we observed alt-NHEJ fusions despite the presence of t-loops. TRF2△TR was only slightly better at repressing fusions than an empty vector, whereas wild type TRF2 was significantly better but could still not fully repress alt-NHEJ fusions. This suggests that there are multiple shelterin components acting in tandem to repress alt-NHEJ.

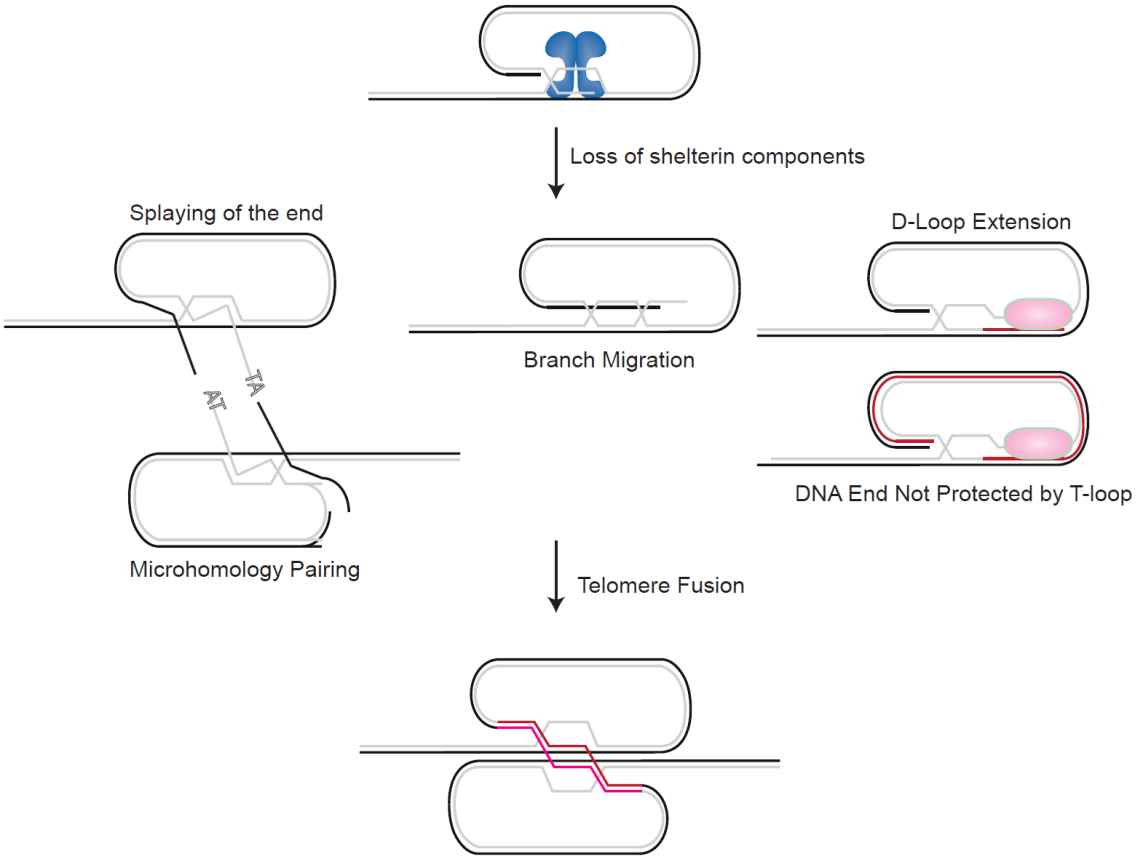
Ku70/80 is a general repressor of alt-NHEJ, and its loss gives rise to a low level of alt-NHEJ fusions (*157*). In addition to TRF2, we also know that TIN2 plays a role in repressing alt-NHEJ, partially through being necessary for full loading of TRF2 at the telomere, and partly through the repression of PARP1 (*60*). This indicates that there is a concerted effort at telomeres to repress both PARP1 and alt-NHEJ: TRF2 and TIN2 act with the general repressor Ku. It is still unclear whether TRF2 represses alt-NHEJ by forming t-loops or if TRF2 has a separate mechanism for repressing alt-NHEJ. Prior data, such as the presence of alt-NHEJ fusions in TIN2 and Ku70/80 double deficient cell lines, supports the hypothesis that neither t-loops nor TRF2 are sufficient to protect against alt-NHEJ. There is also another way to induce alt-NHEJ fusions: loss of TRF1 can lead to stalled replication forks at about 5% telomere fusions, and this number can increase to 14% when combined with the loss of Rap1 (*98*). It is possible that loss of TRF1 reduces the levels of TRF2 at the telomere, thereby derepressing alt-NHEJ (*98*).

Finally, it may also be the case that shelterin does not repress alt-NHEJ that occurs at telomere internal sites, such as may be the case with a stalled replication fork (*125*).

One explanation for the alt-NHEJ at telomeres with normal t-loop levels is that t-loops do not protect against alt-NHEJ. However, there are a few explanations that are consistent with the t-loop model of end-protection. Mechanistically, it is possible that the 5' end becomes accessible to PARP1, Lig3 or Pol θ without active repression by shelterin components. This can potentially occur due to t-loop branch migration, or due to thermodynamic splaying of the last 1 – 3 nt, which can then serve as a substrate for alt-NHEJ. Whatever the cause, we speculate that this can lead to Pol θ accessing the 3' end and extending it. If this extension continues around the entire t-loop, the telomeric end would no longer be protected by the D-loop at the base of the t-loop. This would lead to a telomeric end that could be fused or used as a substrate for recombination. Regardless of whether extension occurs or not, the splaying at the telomere can potentially lead to microhomology pairing between two different telomeres. If this occurs, alt-NHEJ can fuse the telomeres by fill in and Lig3 fusion. This can explain our finding that we observed both t-loops and telomere fusions: the end can be recognized by the alt-NHEJ machinery even if it is within a t-loop, and fused to another t-loop without unwinding it.

Figure 15. Model of t-loop fusions during alt-NHEJ

(A) Branch migration of the telomere end leads to accessibility of the end to alt-NHEJ factors, and consequently D-loop extension. This can generate and end that can be fused, and t-loop to t-loop fusions may occur.


Loss of shelterin components
Splaying of the end
D-Loop Extension
Branch Migration
TA
AT
DNA End Not Protected by T-loop
Microhomology Pairing
Telomere Fusion

A final possibility is that t-loops are protective, but loss of other shelterin components causes persistent damage or replication issues at the telomere, which may cause a fraction of t-loops to be unwound. It is in this state that alt-NHEJ can act. Other telomeres may have escaped persistent unwinding and returned to a t-loop confirmation, granting them a degree of protection from alt-NHEJ – which would explain why we see both looped and un-looped telomeres, as well as why alt-NHEJ fusions do not ever reach the absolute level of fusions seen with c-NHEJ.

It is interesting to ponder why TRF2 would be so extraordinarily self-sufficient – capable of making protecting telomeres from ATM, NHEJ, and making t-loops, yet share the responsibility of repressing alt-NHEJ. One possibility goes back to t-loops: if the primary mechanism of action of TRF2 is through the t-loop, and t-loops are not sufficient to repress alt-NHEJ, this would necessitate backup pathways within shelterin. Alternatively, alt-NHEJ may have evolved as a catch-all repair pathway which necessitates its ability to act on a wide variety of substrates, and therefore would require multiple repressors. Finally, alt-NHEJ may be a backup repair pathway, one that occurs so infrequently that shelterin has not evolved to repress it. Rather, shelterin may be geared towards repression of PARP1 signaling to prevent its activation at the chromosome end, and repression of alt-NHEJ is a byproduct of this. Regardless, we have shown that TRF2 is the only factor at telomeres that makes t-loops, and likely through t-loops, is the sole repressor of NHEJ, and shares the role of repressing ATM signaling with TIN2. However, neither TRF2 nor t-loops are capable of fully suppressing alt-NHEJ on their own.

Chapter 4

Investigating the Mechanism of T-loop Formation

TRF2 is a telomeric protein that is responsible for forming t-loops to protect the chromosome end from ATM signaling and the DNA repair machinery. How TRF2 creates and maintains a t-loop remains a mystery. Having established that TRF2 is sufficient to form t-loops, we test several hypotheses as to which features of TRF2 may be responsible for this action, such as the basic domain, the Top DNA binding feature, and TRF2 tetramerization. We go on to characterize alleles of TRF2 deficient for these functions and show that the Top and tetramerization features are defective in t-loop formation but are behaving as hypomorphic alleles, barring interpretation. Further, we wanted to test whether TRF2 fulfills its function by recruiting non-shelterin proteins for t-loop maintenance. We focused on RAD51 and RAD51D which have both been implicated in strand-invasion and show that neither is required for t-loop formation. In addition, we tested RTEL1, a TRF2-bound helicase implicated in telomere maintenance and show it does not affect t-loop frequency. Finally, we demonstrate that TRF2 is capable of maintaining t-loops throughout the cell cycle. Although we are still unable to definitively explain how TRF2 is capable of forming a t-loop, we have shown that both the Top and tetramerization features are essential to TRF2 function.

Introduction

TRF2 Features

In the past few years characterization of a number of TRF2 mutants has shed light on the individual functions of TRF2 domains. This has led to the hypothesis that the basic domain may play a role in t-loop maintenance. The basic domain, so called because of its high density of basic residues, is notable for its

ability to repress t-loop cleavage (*59*). TRF2's basic domain has been shown to bind branched DNA such as 3- and 4-way junctions, and can do so irrespective of the presence of a telomeric sequence (*65*). Recently, the basic domain has been shown to be a bona fide Holliday junction binding domain in vivo, in addition to being able to bind the aforementioned branched DNA structures with sub-micromolar affinity in vitro (*60*, *158*). Given that the base of the t-loop is a D-loop, which most closely resembles a 3-way junction, the binding activity of the basic domain may be relevant for t-loop formation. This DNA binding is functionally relevant because it acts to prevent recombination or aberrant cutting of the DNA by nucleases by preventing branch migration and recognition of the end by PARP1 (*60*, *159*). Given the higher order structure of the t-loop, and the fact that loss of the basic domain leads to cleavage of t-loop sized fragments, we hypothesized that the basic domain may play a role in t-loop formation.

Another domain hypothesized to be involved in t-loop formation is the TRFH domain, which has three reported functions: interaction with non-shelterin proteins, dimerization, and DNA wrapping. Crystal structures of the TRF2 TRFH domain have shown the interface for TRF2-TRF2 interactions (*68*). Dimerization is critical both for TRF2 localization to telomeres, and DNA binding in vitro (*73*, *160*). TRF2's TRFH domain has also been implicated in the ability of TRF2 to distort chromatin (*62*, *132*).

In addition to containing residues required for TRF2 homodimerization, the TRFH domain contains residues that may facilitate tetramerization. These interface residues, specifically in helix 2 and helix 3, can be observed in the

packing lattice crystal structure of the TRF2 TRFH domain but do not occur in TRF1. TRF2, but not TRF1 has been observed to form tetramers in a variety of conditions, such as on EM, AFM, with or without DNA, and with or without Rap1 (*62*, *130*, *141*, *142*, *161*, *162*). The oligomer size of TRF2 has been determined by analyzing the distribution of particle contour circumference on EM or AFM, and by gel filtration during protein purification (*62*, *161*, *162*). These studies have shown that TRF2 can be found as a dimer with Rap1, or as a tetramer with a tetramer of Rap1. We and others hypothesize that tetramerization may be important for t-loop formation (*62*, *130*). TRF2's tetramerization may act to loop and distort telomeric DNA and this can promote the unwinding needed for strand invasion. Additionally, tetramerization may occur between TRF2 at the telomere end and TRF2 at an internal distortion which would act to bring the telomere end into spatial proximity with a region that can be strand invaded. In this manner TRF2 would act similar to the tetramerization and looping observed with the lac repressor (*147*). Unpublished data suggests that TRF2 tetramerization may be mediated by helix 2/3 of the TRFH domain, and by a set of phenylalanine residues in helix 1 (F59, F69).

Recently, a series of basic residues were identified within the TRF2 TRFH domain that facilitate DNA wrapping around TRF2, and these were shown to be necessary for the protective role of TRF2 at telomeres (*61*). These residues were found based upon their protection from acetylation footprinting due to their contact with DNA. This series of residues (R69, R99, K158, K173, K176, K179, K241, K242, and K245) was termed the "Top" domain and has been proposed to be crucial for t-loop formation. It was proposed that this DNA binding feature wraps

90 base pairs of DNA around the TRFH domain in a right-hand manner, which induces torsional stress on the dsDNA via positive supercoiling. This torsion may melt the DNA, exposing a ssDNA region that can be used to strand invade and form a D-loop. In vitro, a Topless mutant where the critical basic residues were mutated to alanines lacks the ability to promote strand-invasion, which is a critical step in t-loop formation. This Topless mutant was unable to form t-loops in vivo and was also unable to repress ATM signaling in a setting where endogenous TRF2 was knocked down. However, Topless was able to localize to telomeres and complement endogenous TRF2 knockdown by repressing c-NHEJ. Repression of c-NHEJ was lost after knockdown of Rap1, which suggests that Rap1 acts as a back-up mechanism for TRF2. This suggests that Topless is a separation of function mutant that teases apart the role of t-loops in ATM suppression from TRF2 localization and c-NHEJ suppression.

To test the hypotheses that one or more of these features are required for t-loop formation, the function of these domains must be separated such that we ablate t-loop formation but otherwise do not disturb TRF2 localization. If TRF2 localization is disturbed, or if the mutant protein is not stable in cells, it will not be possible to interpret whether or not t-loop loss is due to a lack of functional TRF2 at the telomeres, or due to the specific mutation introduced.

Exogenous Factors

In vivo, it has not been determined whether TRF2 acts alone to form t-loops, or whether other exogenous, non-shelterin proteins may play a role. One class of

such proteins are recombinases, such as RAD51, its paralogs, and RAD52. RAD51 and RAD52 have already been proposed to play a role in t-loop formation (*163*). Homologous recombination was found to be required for D-loop formation in telomeric DNA, implicating a recombinase in a crucial step of t-loop formation.

RAD51 is known to be essential in mammalian cells (*164*, *165*). RAD51 is a filament forming bacterial RecA homolog and is known to be involved in homologous recombination and template search (*166–168*). Interestingly, RAD51 deletion is not lethal in yeasts, which so far have not been shown to harbor a t-loop, and one possible explanation is that the essential function of mammalian RAD51 is in t-loop formation and therefore telomere end-protection (*169*, *170*). Otherwise, RAD51 is known to act at telomeres in yeast as a backup extension pathway and RAD51 has been implicated in telomere replication and maintenance in human cells through BRCA2 (*171*, *172*).

RAD51 does not work alone, and there are other recombinases that function alongside it. Five RAD51 paralogs, named RAD51B-D and XRCC2,3 have been identified, and they play an important role in maintaining HR in cells (*173*, *174*). In particular, the RAD51D is essential in mice and this paralog has been implicated in a telomere overhang maintenance role (*175–177*). Additionally, RAD51D can bind branched DNA, and preferentially binds Holliday junctions, a structure that is believed to be found at the base of the t-loop (*178*). Finally, at least in *Arabidopsis*, RAD51D can function independently of RAD51, suggesting a mechanism by which it can aid t-loop formation, and therefore warrants further study (*179*).

Although any helicase in unlikely to play a role in the formation of t-loops, they may play a role in their proper regulation. One in particular, RTEL1, a DEAH-family helicase, has been singled out for its potential ability to unwind t-loops during S phase (*180*). RTEL1 was first identified as a telomere length regulating gene in mice and then was found to be the homolog of the yeast srs2 helicase dog-1 in *C. elegans* (*181*). Since then, it has been extensively implicated in both genomic stability and for maintaining telomeric integrity throughout replication. Furthermore, RTEL1 has been proposed to unwind t-loops, which led us to test whether or not RTEL1 had an effect on t-loops (*180*, *182–185*).

Cell Cycle

An outstanding question in the field has been whether t-loops are present throughout the cell cycle, or whether there are phases during which t-loops are unwound. If t-loops are specific to a phase of the cell cycle, such as S phase, it may lend mechanistic insights into how t-loops are formed and regulated. Studies have suggested that certain helicases, such as RTEL1, can unwind t-loops during S phase (*180*, *185*). Other work has shown that there is a transient telomere deprotection event during mitosis (*186*, *187*). Such an event would correspond to loss of TRF2 and would be expected to result in loss of t-loops. Finally, if t-loops require homologous recombination, their presence may be diminished in G_1.

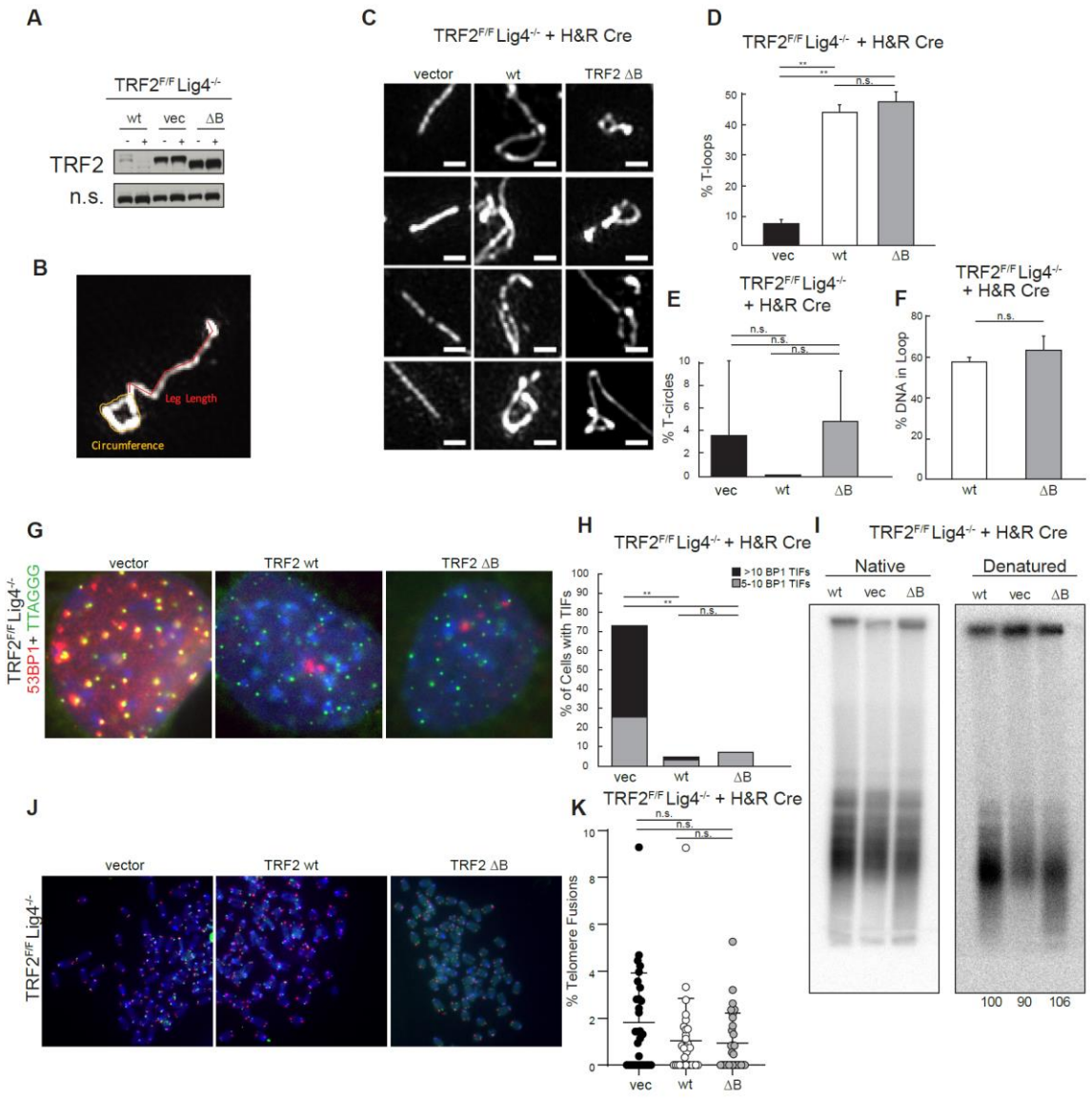
Results

The TRF2 Basic domain is not required for t-loop formation

We tested whether the TRF2 basic domain is required for t-loop formation. We first expressed TRF2$^{\Delta basic}$ in MEFs and verified expression by immunoblotting (Fig. 16A). We then performed OMX microscopy on DNA spreads and confirmed the presence of t-loops in TRF2$^{\Delta basic}$ (Fig. 16B-D). We did not see any significant difference in the formation of t-loops between wild type TRF2 and TRF2$^{\Delta basic}$. Further, we could detect t-circles by OMX imaging, but there was no significant difference between the samples, and no change in the loop size (Fig. 16B,E,F). TRF2$^{\Delta basic}$ is capable of functioning in telomere end protection in vivo, as there was no significant increase in 53BP1 foci at telomeres, consistent with previous findings (Fig. 16G,H). In addition, we also wanted to verify whether or not there were chromosome fusions in complemented TRF2$^{F/F}$ Lig4$^{-/-}$ cells, which could confound our t-loop analysis. We found that there was no increase in chromosome fusions (Fig. 16J,K). Finally, we also analyzed the telomeric overhang in these cells, and found no significant changes (Fig. 16I). Overall, this shows that although the basic domain of TRF2 may be important in protecting the t-loop, it is not required for t-loop formation.

Figure 16. TRF2's basic domain is not required for t-loop formation.

(A) Western blot showing expression of TRF2 and TRF2$^{\Delta Basic}$ in MEFs immortalized with p53-loss in TRF2$^{F/F}$ Lig4$^{-/-}$ cells, 120 hours after Cre-mediated deletion. TRF2 in the top panel, N.s. loading control in the bottom panel. (B) Schematic showing an analysis of t-loop length and the loop versus tail measurements. (C) Representative images of DNA spreads showing t-loops in TRF2$^{F/F}$ MEFs, imaged using super-resolution OMX microscopy, 120 hours after tamoxifen induced Cre deletion (D) Quantification of C, n = 3 independent experiments, 100 countable molecules per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * p > 0.05, ** p > 0.01, *** p > 0.001) (E) Quantification of C, n = 3 independent experiments, 100 countable molecules per replicate. T-circle is defined as a molecule where 85% or more of the total length of the molecule is within the loop portion. (F) Analysis of percentage of telomeric DNA in the loop versus the loop plus tail. n = 3 independent experiments, 100 countable molecules per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * p > 0.05, ** p > 0.01, *** p > 0.001) (G) Representative images of 53BP1 TIFs in TRF2$^{F/F}$ MEFs 120 hours after Cre-induction with tamoxifen. Blue – DAPI. Green – Telomeres. Red – 53BP1. (H) Quantification of F, showing cells with 5-10 or >10 53BP1 TIFs. (I) Representative images of metaphases in TRF2$^{F/F}$ MEFs, 120 hours after Cre-induction with tamoxifen. Blue – DAPI. Green – FITC-(TTAGGG)$_3$ Red – Cy3-(CCCTAA)$_3$ (J) Quantification of H, n = 2 independent experiments, 10 metaphases per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * p > 0.05, ** p > 0.01, *** p > 0.001) (K) Telomere blot in TRF2$^{F/F}$ Lig4$^{-/-}$ cells 120 hours after tamoxifen induced Cre-mediated deletion. Left panel, native gel. Right panel, denaturing gel. Relative quantification of the overhang (normalized to TRF2) are shown below the right panel.


A
TRF2F/F Lig4-/-
wt
vec
ΔB
TRF2
n.s.
B
Leg Length
Circumference
C
TRF2F/F Lig4-/- + H&R Cre
vector
wt
TRF2 ΔB
D
TRF2F/F Lig4-/- + H&R Cre
% T-loops
vec
wt
ΔB
E
TRF2F/F Lig4-/-
+ H&R Cre
% T-circles
F
TRF2F/F Lig4-/-
+ H&R Cre
% DNA in Loop
G
TRF2F/F Lig4-/-
53BP1+ TTAGGG
vector
TRF2 wt
TRF2 ΔB
H
TRF2F/F Lig4-/- + H&R Cre
>10 BP1 TIFs
5-10 BP1 TIFs
% of Cells with TIFs
I
TRF2F/F Lig4-/- + H&R Cre
Native
Denatured
100
90
106
J
TRF2F/F Lig4-/-
K
TRF2F/F Lig4-/- + H&R Cre
% Telomere Fusions

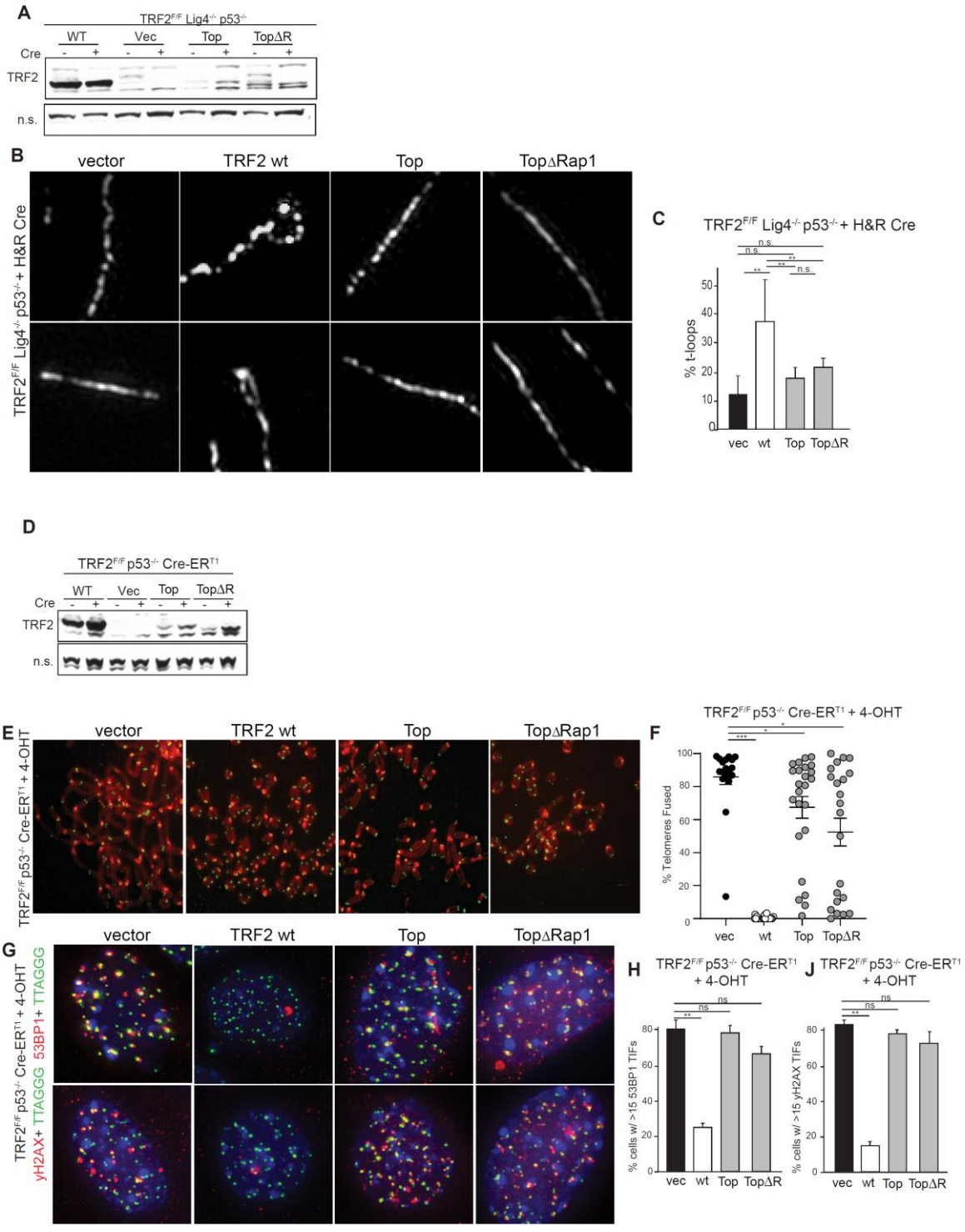
DNA Wrapping As a Mechanism for T-loop Formation

To test the DNA wrapping model, we generated the mouse equivalent of TRF2^{Topless}, mutating the 7 relevant lysines and 2 relevant arginines. In parallel we generated and tested TRF2^{ToplessΔRap1} to determine whether or not Rap1 can suppress telomere fusions. If the t-loop and wrapping models are correct, these mutants should be unable to form t-loops and therefore also unable to repress ATM signaling and c-NHEJ. Alternatively, we could observe the same results as previous studies on Topless and see that Topless fails to suppress ATM signaling but does repress c-NHEJ fusions.

After expressing the Topless mutants in TRF2$^{F/F}$ Lig4$^{-/-}$ p53$^{-/-}$ MEFs, we then tested whether either of the Topless mutants was competent for t-loop formation. We assayed for t-loop formation 120 hours after Cre induction and observed that the Topless mutants had a significantly reduced level of t-loops compared to the wild type TRF2 (Fig. 17A-C). It is important to note that TRF2^{Topless} it not expressed as readily in MEFs or 293T cells as wild type TRF2, which could be due to instability (Fig. 17A,D). However, it is still significantly overexpressed compared to endogenous levels of TRF2 and therefore its ability to form t-loops or protect telomere ends should not be affected by expression levels (Fig. 17A,D). The phenotypes of TRF2^{Topless} and TRF2^{ToplessΔRap1} were not significantly different, and both were comparable to the empty vector control, suggesting that these mutants were completely deficient in t-loop formation.

Figure 17. TRF2 Topless Mutants are Hypomorphs

(A) Western blot showing expression of TRF2, $TRF2^{Topless}$, and $TRF2^{ToplessΔRap1}$ in MEFs immortalized with p53-loss and floxed for TRF2 and deficient in Lig4. TRF2 in the top panel, N.s. loading control in the lower panel. (B) Representative images of DNA spreads showing t-loops in $TRF2^{F/F}$ $Lig4^{-/-}$MEFs, imaged using super-resolution OMX microscopy, 120 hours after tamoxifen induced Cre deletion. (C) Quantification of H, n = 3 independent experiments, 100 countable molecules per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * p > 0.05, ** p > 0.01, *** p > 0.001). (D) Western blot showing expression of TRF2, $TRF2^{Topless}$, and $TRF2^{ToplessΔRap1}$ in MEFs immortalized with p53-loss and floxed for TRF2. N.s. loading control in the top panel, TRF2 in the lower panel. (E) Representative images of metaphases in TRF2F/F MEFs, 120 hours after Cre-induction with tamoxifen. Red – DAPI. Green – Alexa647-(TTAGGG)3 (F) Quantification of E, n = 3 independent experiments, 10 metaphases per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * p > 0.05, ** p > 0.01, *** p > 0.001). (G) Representative images of TIFs in TRF2F/F MEFs 120 hours after Cre-induction with tamoxifen. Top panel: Blue – DAPI. Green – Telomeres. Red – 53BP1. Bottom panel: Blue – DAPI. Green – Alexa647-$(TTAGGG)_3$. Red – γH2AX. (H) Quantification of G, showing cells with 15 or more 53BP1 TIFs. n = 3 independent experiments, 100 cells per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * p > 0.05, ** p > 0.01, *** p > 0.001). (H) Quantification of G, showing cells with 15 or more γH2AX TIFs. n = 3 independent experiments, 100 cells per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * p > 0.05, ** p > 0.01, *** p > 0.001).

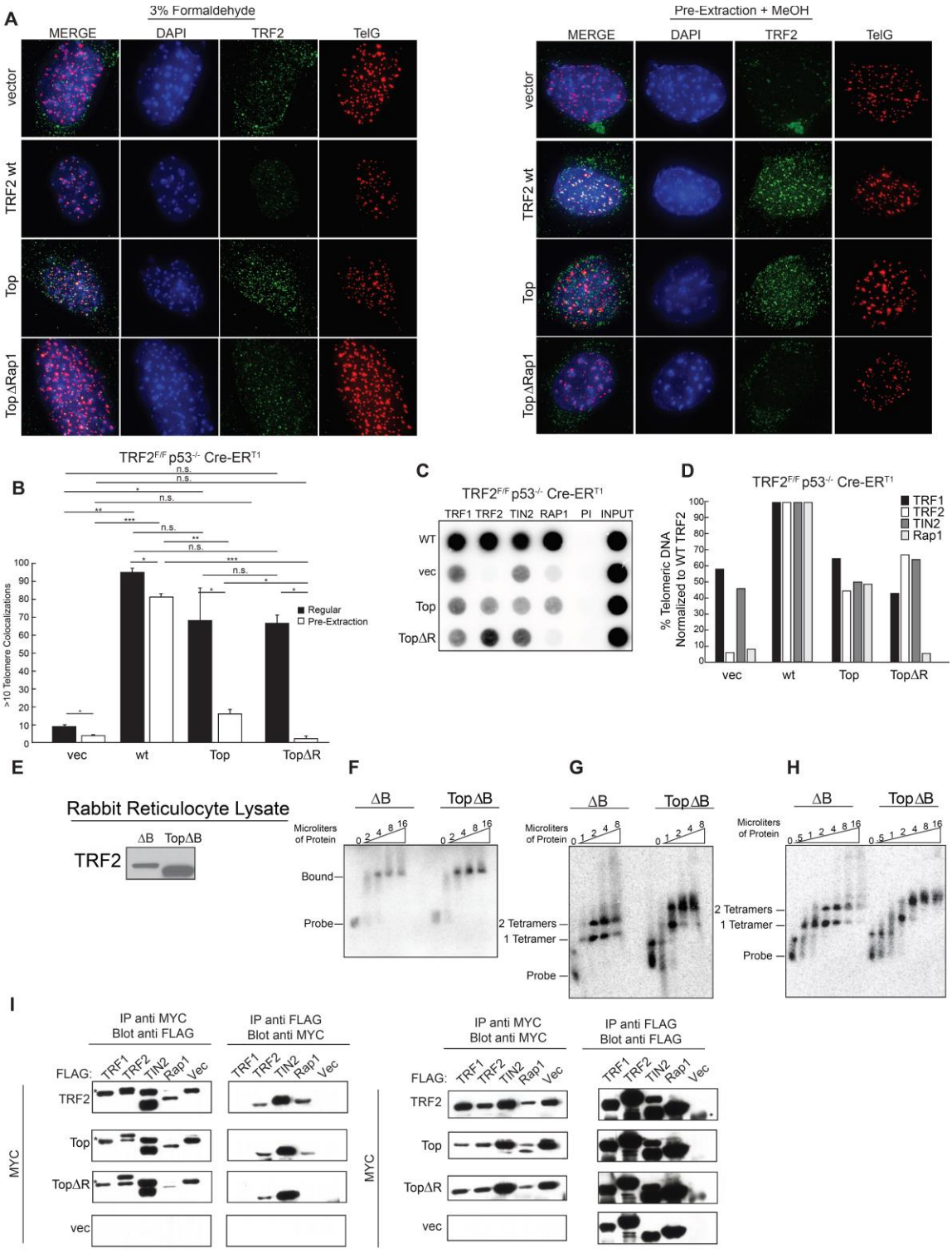

A
TRF2F/F Lig4-/- p53-/-
WT
Vec
Top
TopΔR
Cre
- + - + - + - +
TRF2
n.s.
B
vector
TRF2 wt
Top
TopΔRap1
TRF2F/F Lig4-/- p53-/- + H&R Cre
C
TRF2F/F Lig4-/- p53-/- + H&R Cre
n.s.
**
n.s.
% t-loops
vec wt Top TopΔR
D
TRF2F/F p53-/- Cre-ER T1
WT Vec Top TopΔR
Cre
- + - + - + - +
TRF2
n.s.
E
vector
TRF2 wt
Top
TopΔRap1
TRF2F/F p53-/- Cre-ER T1 + 4-OHT
F
TRF2F/F p53-/- Cre-ER T1 + 4-OHT
***
*
% Telomeres Fused
vec wt Top TopΔR
G
vector
TRF2 wt
Top
TopΔRap1
TRF2F/F p53-/- Cre-ER T1 + 4-OHT
yH2AX+ TTAGGG 53BP1+ TTAGGG
H
TRF2F/F p53-/- Cre-ER T1
+ 4-OHT
ns
**
% cells w/ >15 53BP1 TIFs
vec wt Top TopΔR
J
TRF2F/F p53-/- Cre-ER T1
+ 4-OHT
ns
**
% cells w/ >15 γH2AX TIFs
vec wt Top TopΔR

We next tested whether Topless was able to suppress fusions caused by the loss of TRF2. We analyzed metaphases spreads and assayed for end-to-end fusions 120 hours after Cre-mediated deletion of TRF2. We observed that while wild type TRF2 was capable of fully suppressing telomere fusions upon TRF2 deletion, the Topless mutants were not (Fig. 17E, F). However, they were able to partially suppress TRF2 deletion, compared to the empty vector. There was no significant difference between the mutants.

We continued by testing whether Topless and Topless$^{\Delta Rap1}$ can suppress either 53BP1 or γH2AX TIFs. At 120 hours after deleting endogenous TRF2, we assayed TIF formation by IF-FISH. While exogenous expression of wild type TRF2 suppressed TIF formation, neither Topless nor Topless$^{\Delta Rap1}$ were able to suppress TIF formation (Fig. 17G-J). This data is consistent with DNA wrapping being required for t-loop formation and t-loops being required for ATM and c-NHEJ suppression. However, the failure of the Topless mutants to behave as a separation of function between t-loop formation and ATM signaling and fusions led us to suspect that these mutants may be hypomorphs which are unable to fulfill any of the functions of TRF2.

Figure 18. TRF2 Topless maintains all shelterin interactions but binds DNA poorly.

(A) Immunofluorescence showing colocalization of TRF2, Empty Vector, or TRF2 Topless and TRF2 Topless$^{\Delta Rap1}$ in TRF2$^{F/F}$ MEFs infected with the aforementioned constructs and 96 hours after Cre-mediated deletion. Telomeres are visualized by Alexa647-(TTAGGG)3. Left panel – fixation with 3% formaldehyde for 10 minutes. Right panel – Pre-extraction followed by fixation in methanol for 10 minutes. Blue – DAPI, nuclear DNA. Green – TRF2 antibody. Red - telomeres. (B) Quantification of (A), n = 3 replicates and approximately 100 cells per replicate. (C) Chromatin Immunoprecipitation showing TRF2, Empty Vector, and various TRF2 mutants interacting with telomeric DNA and with TRF1, TIN2, or RAP1 on telomeric DNA in TRF2$^{F/F}$ MEFs infected with the aforementioned constructs and 120 hours after Cre-mediated deletion. Input is 20%, PI – preimmune. 2 replicates are shown. (D) Quantification of (A) normalized to wild type TRF2 set to 100%. (E) Western blot showing protein from a rabbit reticulocyte lysate used for in vitro translation of TRF2$^{\Delta basic}$ and TRF2 Topless (F) EMSA showing DNA binding activity of TRF2 Topless, 1 ng of TH12 probe was used. (G) EMSA that neither TRF2$^{\Delta basic}$ nor TRF2 Topless exhibit a laddering phenotype. 0.5ng of P-32 labeled TH12 probe was used and incubated for 20 minutes (H) EMSA that neither TRF2$^{\Delta basic}$ nor TRF2 Topless exhibit a laddering phenotype. 0.25ng of P-32 labeled TH12 probe was used and incubated for 20 minutes (I) Co-immunoprecipitation of Flag tagged shelterin components (TRF1, TRF2, TIN2, Rap1) and Myc-tagged TRF2, Empty Vector, TRF2 Topless and TRF2 Topless$^{\Delta Rap1}$.


A
3% Formaldehyde
Pre-Extraction + MeOH
MERGE
DAPI
TRF2
TelG
vector
TRF2 wt
Top
Top ΔRap1
B
TRF2F/F p53-/- Cre-ERT1
>10 Telomere Colocalizations
Regular
Pre-Extraction
vec
wt
Top
TopΔR
C
TRF2F/F p53-/- Cre-ERT1
TRF1 TRF2 TIN2 RAP1 PI INPUT
WT
vec
Top
TopΔR
D
TRF2F/F p53-/- Cre-ERT1
% Telomeric DNA Normalized to WT TRF2
TRF1
TRF2
TIN2
Rap1
E
Rabbit Reticulocyte Lysate
ΔB TopΔB
TRF2
F
ΔB
TopΔB
Microliters of Protein
Bound
Probe
G
2 Tetramers
1 Tetramer
H
I
IP anti MYC
Blot anti FLAG
IP anti FLAG
Blot anti MYC
FLAG:
TRF1
TRF2
TIN2
Rap1
Vec
MYC

We hypothesized that the Topless mutants may be unable to bind DNA, or lack stabilizing interactions within shelterin, which would explain why they are hypomorphs. We first tested their telomeric localization in vivo. We expressed TRF2^{Topless} in TRF2$^{F/F}$ MEFs and performed an immunofluorescence experiment where we varied the stringency of pre-extraction (methanol only or pre-extraction with 0.5% Triton X-100) and tracked the co-localization of TRF2^{Topless} with telomeres. We observed that TRF2^{Topless} not only localized to telomeres less efficiently than wild type TRF2 in standard IF conditions but was also almost entirely removed when the stringency of protein extraction before IF was increased (Fig. 18A,B). To confirm this result, we performed a ChIP and observed that TRF2^{Topless} mutants displayed poor telomeric localization (Fig. 18C, D). This suggests that the defect that TRF2^{Topless} has is related to DNA binding and telomere localization, which is a plausible explanation if it lacks the ability to properly wrap DNA around its Top domain and therefore has fewer contacts with the DNA that may stabilize it.

We then performed an in vitro experiment to test whether purified TRF2^{Topless} could bind telomeric DNA. We confirmed expression of these constructs in a rabbit reticulocyte lysate system by immunoblotting (Fig. 18E,F). Next, we tested the ability of these mutants to bind a telomeric probe with 12 telomeric repeats and the TRF2^{Topless} mutant displayed a defect in DNA binding (Fig. 18E,F). In three independent gel-shift experiments, the volume of Topless reticulocyte lysate required to fully shift the probe was twice the volume needed for the wild type TRF2 (compare 1 and 2 µl lanes) (Fig. 18F-H). This indicates a

deficiency in the Topless' ability to bind DNA because the reticulocyte lysate contained 2.3 times more translated protein (Fig. 18E).

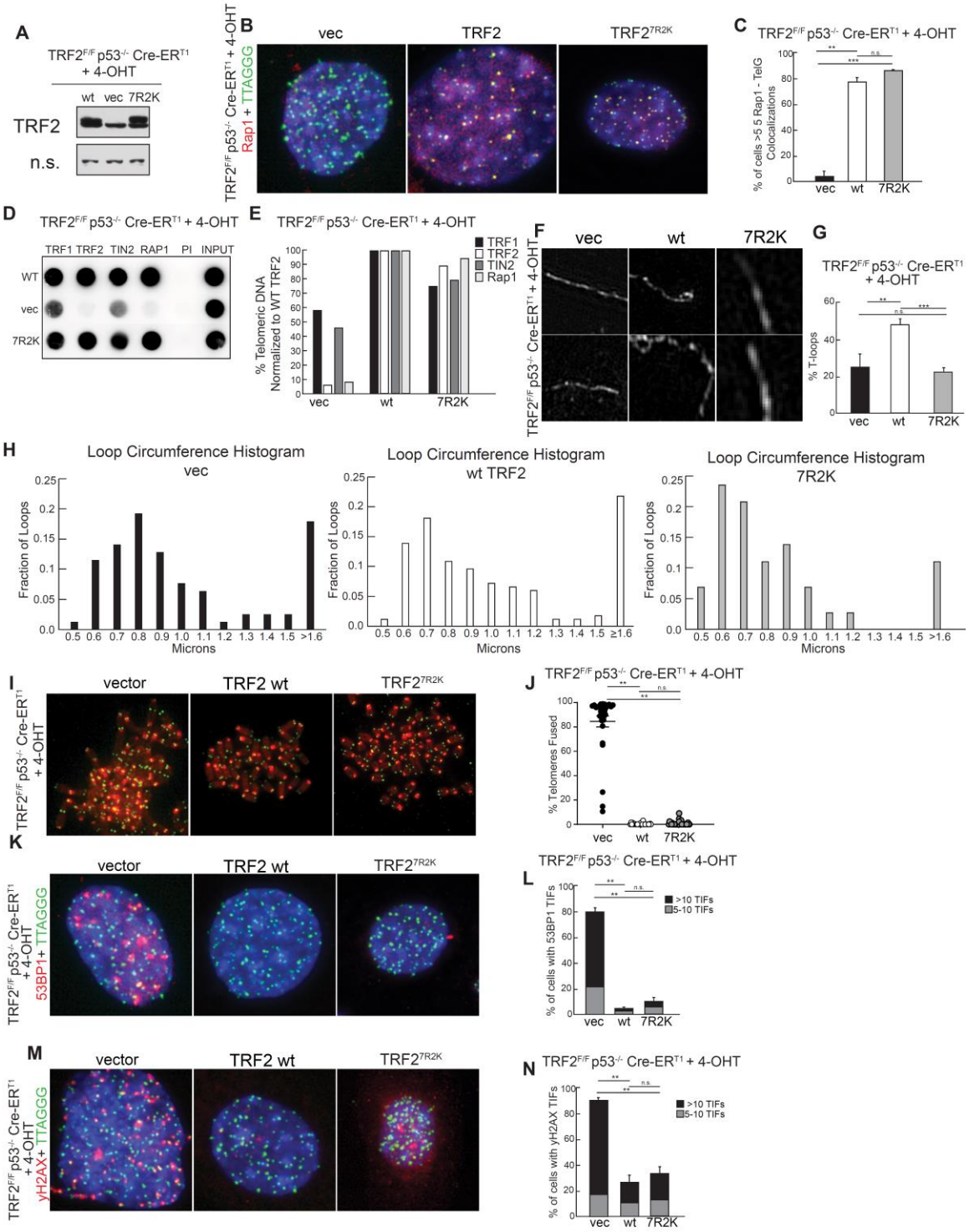
We also wanted to know whether TRF2^{Topless} maintained its interactions with other shelterin components. To test this, we performed a series of co-immunoprecipitations between TRF2 and other shelterin components. We observed that with the exception of TRF2^{ToplessΔRap1}, which predictably does not interact with Rap1, TRF2^{Topless} interacted with all components of the shelterin complex in a comparable way to wild type TRF2 (Fig. 18I). TRF2^{Topless} was also able to dimerize with wild type TRF2 (Fig. 18I). These results largely recapitulate the data seen from the ChIP experiments, which ultimately suggests that the defect that TRF2^{Topless} suffers from is an in vivo DNA binding and localization issue.

We also created TRF2^{7R2K}, a variant of the Topless mutant that maintained the charge of the critical Top domain residues but switched the lysines and arginines. We hypothesized that this mutant would either behave like wild type TRF2, or if side chain length was critical to Top domain function, would exhibit a milder phenotype than the Topless mutant and perhaps not be a hypomorph.

We began by asking whether TRF2^{7R2K} can bind telomeres in vivo. We found that TRF2^{7R2K} was able to localize to telomeres with comparable efficiency to wild type TRF2, as demonstrated by both IF and ChIP (Fig. 19A-E). We next asked whether TRF2^{7R2K} was able to form t-loops. Surprisingly, TRF2^{7R2K} mutant was unable to form t-loops: it displayed a significantly reduced level of t-loop formation compared to wild type TRF2 and was not significantly different from the empty vector control (Fig. 19F, G).

Figure 19. TRF2^{7R2K} confers a T-loop defect

(A) Western blot showing expression of TRF2 and TRF2^{7R2K} in MEFs immortalized with p53-loss and floxed for TRF2, 120 hours after Cre-mediated deletion. N.s. loading control in the top panel, TRF2 in the lower panel. (B) Immunofluorescence showing colocalization of Rap1 used as a proxy for TRF2, TRF2$^{F/F}$ MEFs infected with TRF2 or TRF2^{7R2K} and imaged 120 hours after Cre-mediated deletion. Telomeres are visualized by Alexa647-(TTAGGG)3. Blue – DAPI, nuclear DNA. Green – TRF2 antibody. Red - telomeres. (C) Quantification of (B), n = 2 replicates and approximately 100 cells per replicate. (D) Chromatin Immunoprecipitation showing TRF2, Empty Vector, and TRF2^{7R2K} mutants interacting with telomeric DNA and with TRF1, TIN2, or RAP1 on telomeric DNA in TRF2$^{F/F}$ MEFs infected with the aforementioned constructs and 120 hours after Cre-mediated deletion. Input is 20%, PI – preimmune. 2 replicates are shown. (E) Quantification of (A) normalized to wild type TRF2 set to 100%. (F) Representative images of DNA spreads showing t-loops in TRF2$^{F/F}$ MEFs, imaged using super-resolution OMX microscopy, 120 hours after tamoxifen induced Cre deletion. (G) Quantification of B, n = 3 independent experiments, 100 countable molecules per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$). (H) Histogram showing the distribution of the size of residual loops in wild type TRF2, vector, and 7R2K. (I) Representative images of metaphases in TRF2$^{F/F}$ MEFs, 120 hours after Cre-induction with tamoxifen. Red – DAPI. Green – Alexa647-(TTAGGG)$_3$ (J) Quantification of E, n = 3 independent experiments, 10 metaphases per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$). (K) Representative images of 53BP1 TIFs in TRF2$^{F/F}$ MEFs 120 hours after Cre-induction with tamoxifen. Blue – DAPI. Green – Telomeres. Red – 53BP1. (L) Quantification of F, showing cells with 5-10 or >10 53BP1 TIFs. n = 3 independent experiments, 100 cells per replicate. S Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$). (M) Representative images of γH2AX TIFs in TRF2$^{F/F}$ MEFs 120 hours after Cre-induction with tamoxifen. Blue – DAPI. Green – Telomeres. Red – γH2AX. (N) Quantification of H, showing cells with 5-10 or >10 γH2AX TIFs. n = 3 independent experiments, 100 cells per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$).


A
TRF2F/F p53-/- Cre-ERT1 + 4-OHT
wt vec 7R2K
TRF2
n.s.
B
vec
TRF2
TRF2 7R2K
TRF2F/F p53-/- Cre-ERT1 + 4-OHT
Rap1 + TTAGGG
C
TRF2F/F p53-/- Cre-ERT1 + 4-OHT
% of cells >5 5 Rap1 - TelG Colocalizations
vec wt 7R2K
D
TRF2F/F p53-/- Cre-ERT1 + 4-OHT
TRF1 TRF2 TIN2 RAP1 PI INPUT
WT
vec
7R2K
E
TRF2F/F p53-/- Cre-ERT1 + 4-OHT
TRF1
TRF2
TIN2
Rap1
% Telomeric DNA Normalized to WT TRF2
vec wt 7R2K
F
vec wt 7R2K
TRF2F/F p53-/- Cre-ERT1 + 4-OHT
G
TRF2F/F p53-/- Cre-ERT1 + 4-OHT
% T-loops
vec wt 7R2K
H
Loop Circumference Histogram
vec
wt TRF2
7R2K
Fraction of Loops
Microns
I
vector
TRF2 wt
TRF2 7R2K
TRF2F/F p53-/- Cre-ERT1 + 4-OHT
J
TRF2F/F p53-/- Cre-ERT1 + 4-OHT
% Telomeres Fused
vec wt 7R2K
K
vector
TRF2 wt
TRF2 7R2K
TRF2F/F p53-/- Cre-ERT1 + 4-OHT
53BP1 + TTAGGG
L
TRF2F/F p53-/- Cre-ERT1 + 4-OHT
% of cells with 53BP1 TIFs
>10 TIFs
5-10 TIFs
vec wt 7R2K
M
vector
TRF2 wt
TRF2 7R2K
TRF2F/F p53-/- Cre-ERT1 + 4-OHT
γH2AX + TTAGGG
N
TRF2F/F p53-/- Cre-ERT1 + 4-OHT
% of cells with γH2AX TIFs
>10 TIFs
5-10 TIFs
vec wt 7R2K

We measured the size of residual t-loops in the TRF2^{7R2K} samples, and we found the remaining loop size to be smaller than that of the vector (Fig. 19H). This suggests that we cannot exclude the possibility that TRF2^{7R2K} does make t-loops, but they are smaller than our limit of detection.

We next wondered how these properties are reflected in functional cellular assays. We stained metaphases by FISH and did not observed telomere fusion in TRF2^{7R2K} expressing cells, which behaved like wild type TRF2 expressing cells (Fig. 19I,J). We next investigated whether TRF2^{7R2K} is able to suppress the DNA damage response. After staining cell for 53BP1 and γH2AX TIFs, we observed that TRF2^{7R2K} was able to repress both forms of DNA damage signaling, though damage foci were still slightly, though not statistically significantly, elevated over wild type TRF2 (Fig. 19K-N).

We did not test TRF2^{7R2K}'s ability to wrap DNA, so we cannot rule out the possibility that this mutant does not have a defect in DNA wrapping. Therefore, we cannot conclude that the DNA wrapping is involved in any of the phenotypes that we have observed.

Tetramerization as a Mechanism for T-loop Formation

An alternative hypothesis for how t-loops are formed is that they require TRF2 to tetramerize. This tetramerization would then bring the DNA together in a conformation that promotes looping and strand invasion. To test for tetramerization deficiency, a laddering assay is used. The laddering assay is so named because tetramerization mutants "ladder", which is to say they exhibit addition binding

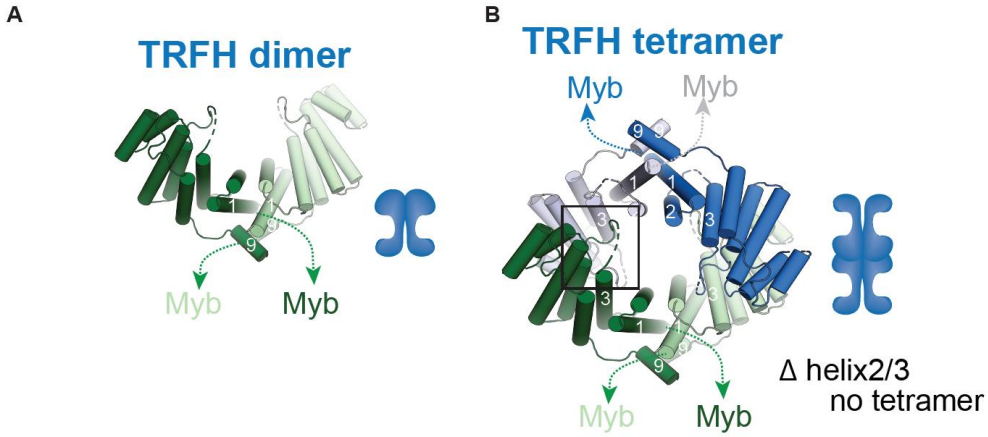
reactions that wild type TRF2 does not, creating the visual impression of laddering (Fig. 20A-C). This is due to TRF2 binding as dimers in tetramerization deficient mutants, which creates a series of small steps, in contrast to wild type TRF2 binding as a tetramer and creating a pattern of larger steps (Fig. 20C). To test the tetramerization hypothesis, we used two tetramerization mutants that had previously been described, $TRF2^{F59A}$ and $TRF2^{\Delta 70\text{-}114}$, in a biochemical laddering assay.

The laddering assay works most efficiently without TRF2's basic domain, so we began by cloning and expressing $TRF2^{F59A}$ and $TRF2^{\Delta 70\text{-}114}$ in a ΔBasic backbone in a rabbit reticulocyte system (Fig. 20D). In this assay, $TRF2^{F59A}$ and $TRF2^{\Delta 70\text{-}114}$ clearly exhibit intermediate binding states compared to wild type TRF2, which has two prominent binding bands that likely correspond to one and two tetramers (Fig. 20F,G). $TRF2^{F59A}$ and $TRF2^{\Delta 70\text{-}114}$ have intermediate bands that correspond to a single dimer, 3 dimers and 5 dimers of TRF2.

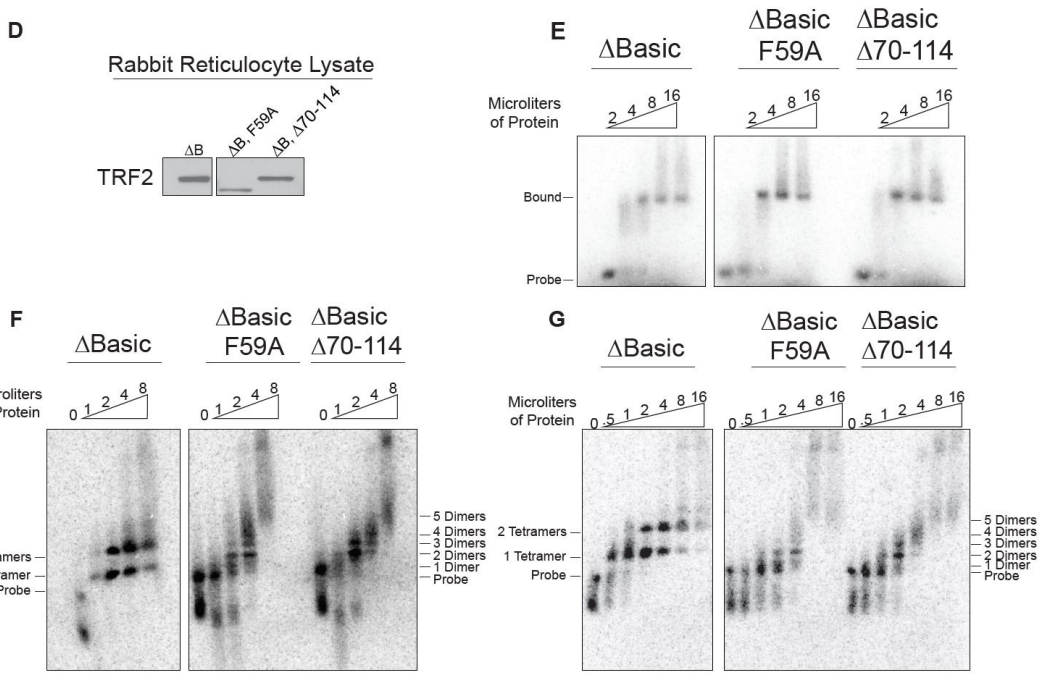
We hypothesized that perhaps the oligomerization of TRF2 is essential for t-loop formation and consequently protection of the telomere end. We re-cloned the laddering mutants into in vivo expression constructs, infected $TRF2^{F/F}$ MEFs with them by viral transduction and verified expression of the laddering mutants (Fig. 21A).

Figure 20. TRF2^{F59A} and TRF2$^{\Delta 70\text{-}114}$ exhibit the laddering phenotype.

(A) Crystal structure of TRF2 TRFH domains modeled in a proposed dimer structure. (B) Crystal structure of TRF2 TRFH domains modeled in a proposed tetramer structure. (C) Schematic of the laddering assay: wild type TRF2 will bind as tetramers and produce larger steps than a tetramerization deficient allele of TRF2, which will bind as dimer. (D) Western blot showing TRF2, TRF2^{F59A}, and TRF2$^{\Delta 70\text{-}114}$ proteins synthesized by in vitro translation in rabbit reticulocyte lysate. (E) EMSA showing DNA binding activity of all TRF2 mutants, 1 ng of TH12 probe was used. Irrelevant lanes have been cropped. (F) EMSA showing the laddering phenotype of TRF2^{F59A} and TRF2$^{\Delta 70\text{-}114}$ but not TRF2$^{\Delta basic}$. 0.5ng of P-32 labeled TH12 probe was used and incubated for 20 minutes. Irrelevant lanes have been cropped. (G) EMSA showing the laddering phenotype of TRF2^{F59A} and TRF2$^{\Delta 70\text{-}114}$ but not TRF2$^{\Delta basic}$. 0.25ng of P-32 labeled TH12 probe was used and incubated for 20 minutes. Irrelevant lanes have been cropped.


A
TRFH dimer
Myb
Myb
B
TRFH tetramer
Myb
Myb
Myb
Myb
Δ helix2/3
no tetramer


C
dsTTAGGG12 gel-shift
wt TRF2
Δ helix2/3
3 tetramers
2 tetramers
1 tetramer
probe
5 dimers
4 dimers
3 dimers
2 dimers
1 dimer
probe


D
Rabbit Reticulocyte Lysate
TRF2
ΔB
ΔB, F59A
ΔB, Δ70-114
E
ΔBasic
ΔBasic F59A
ΔBasic Δ70-114
Microliters of Protein
Bound
Probe
F
ΔBasic
ΔBasic F59A
ΔBasic Δ70-114
Microliters of Protein
2 Tetramers
1 Tetramer
Probe
5 Dimers
4 Dimers
3 Dimers
2 Dimers
1 Dimer
Probe
G
ΔBasic
ΔBasic F59A
ΔBasic Δ70-114
Microliters of Protein
2 Tetramers
1 Tetramer
Probe
5 Dimers
4 Dimers
3 Dimers
2 Dimers
1 Dimer
Probe

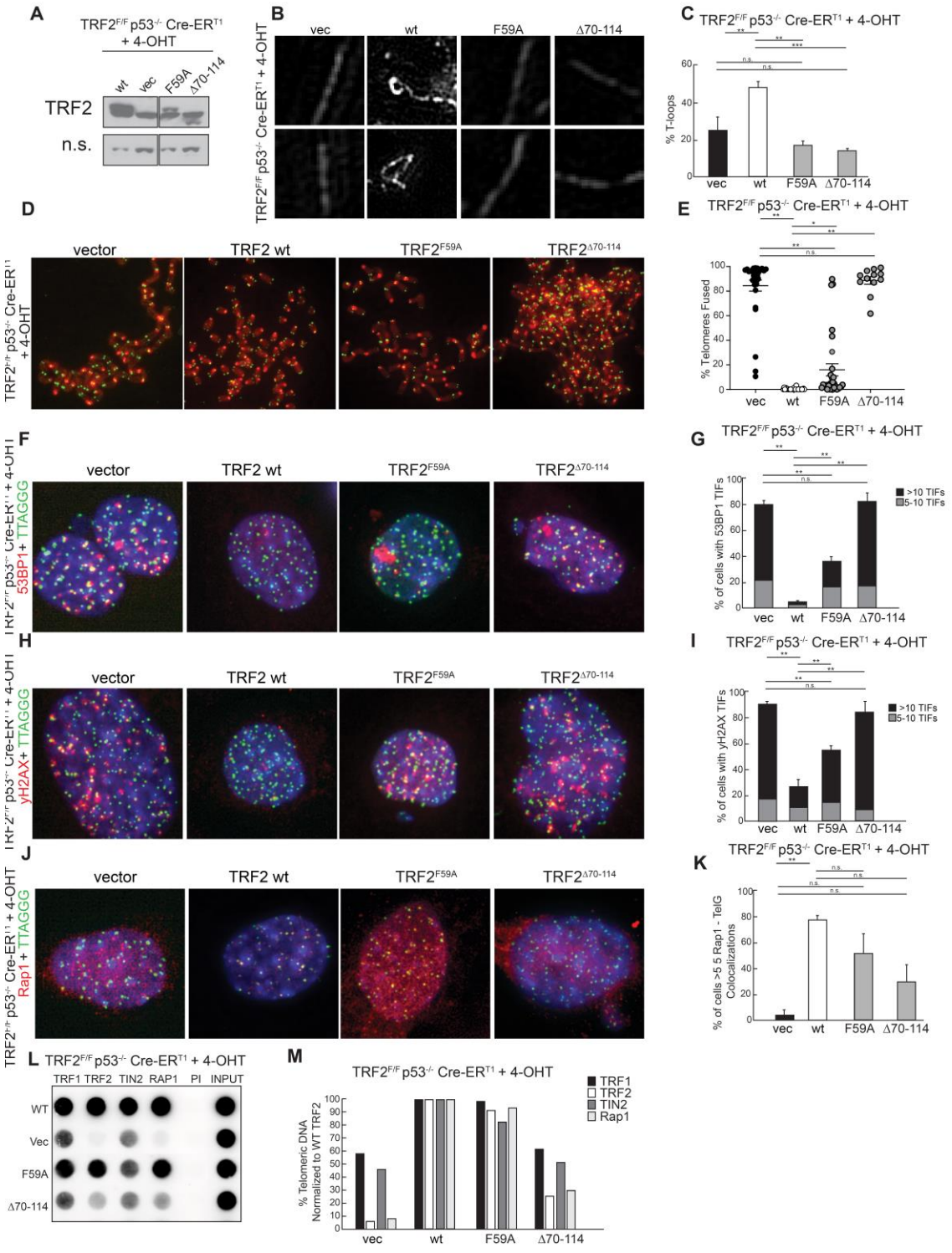
The first question we asked was whether or not these TRF2 mutants could form t-loops. None of the tested TRF2 mutants were able to form t-loops: all displayed a significantly reduced level of t-loop formation compared to wild type TRF2, and none were significantly different from the empty vector control (Fig. 21B, C).

We next determined how the tetramerization mutants behaved in cellular assays for TRF2 function. We stained metaphases by FISH and observed telomere fusion events in both tetramerization mutants (Fig. 21D,E). Interestingly, the phenotype of $TRF2^{\Delta 70\text{-}114}$ was significantly worse than that of $TRF2^{F59A}$, though both clearly were unable to fully suppress telomere fusions.

We next investigated whether these mutants were able to suppress the ATM response. After staining cell for 53BP1 and γH2AX TIFs, we observed a pattern of DNA damage that generally followed the telomere fusion phenotype (Fig. 21F-I). $TRF2^{F59A}$ and $TRF2^{\Delta 70\text{-}114}$ both displayed clear defects in repressing the ATM pathway, with $TRF2^{F59A}$ displaying an intermediate phenotype between empty vector and wild type TRF2 while $TRF2^{\Delta 70\text{-}114}$ was not statistically significantly different from the vector control.

Figure 21. TRF2 laddering mutants are hypomorphs.

(A) Western blot showing expression of TRF2, TRF2^{F59A}, and TRF2$^{\Delta 70\text{-}114}$ in MEFs immortalized with p53-loss and floxed for TRF2, 120 hours after Cre-mediated deletion. N.s. loading control in the top panel, TRF2 in the lower panel. Irrelevant lane has been cropped out. (B) Representative images of DNA spreads showing t-loops in TRF2$^{F/F}$ MEFs, imaged using super-resolution OMX microscopy, 120 hours after tamoxifen induced Cre deletion. (C) Quantification of B, n = 3 independent experiments, 100 countable molecules per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$) (D) Representative images of metaphases in TRF2$^{F/F}$ MEFs, 120 hours after Cre-induction with tamoxifen. Red – DAPI. Green – Alexa647-$(TTAGGG)_3$ (E) Quantification of E, n = 3 independent experiments, 2000 telomere ends scored per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$). (F) Representative images of 53BP1 TIFs in TRF2$^{F/F}$ MEFs 120 hours after Cre-induction with tamoxifen. Blue – DAPI. Green – Telomeres. Red – 53BP1. (G) Quantification of F, showing cells with 5-10 or >10 53BP1 TIFs. n = 3 independent experiments, 100 cells per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$). (H) Representative images of γH2AX TIFs in TRF2$^{F/F}$ MEFs 120 hours after Cre-induction with tamoxifen. Blue – DAPI. Green – Telomeres. Red – γH2AX. (I) Quantification of H, showing cells with 5-10 or >10 γH2AX TIFs. n = 3 independent experiments, 100 cells per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$). (J) Immunofluorescence showing colocalization of Rap1 used as a proxy for TRF2, TRF2$^{F/F}$ MEFs infected with TRF2, TRF2^{F59A}, and TRF2$^{\Delta 70\text{-}114}$ and imaged 120 hours after Cre-mediated deletion. Telomeres are visualized by Alexa647-(TTAGGG)3. Blue – DAPI, nuclear DNA. Green – TRF2 antibody. Red - telomeres. (K) Quantification of (J), n = 3 replicates and approximately 100 cells per replicate. (L) Chromatin Immunoprecipitation showing TRF2, Empty Vector, and TRF2^{7R2K} mutants interacting with telomeric DNA and with TRF1, TIN2, or RAP1 on telomeric DNA in TRF2$^{F/F}$ MEFs infected with the aforementioned constructs and 120 hours after Cre-mediated deletion. Input is 20%, PI – preimmune. 2 replicates are shown. (M) Quantification of (A) normalized to wild type TRF2 set to 100%


A
TRF2F/F p53-/- Cre-ER T1
+ 4-OHT
wt
vec
F59A
Δ70-114
TRF2
n.s.
B
vec
wt
F59A
Δ70-114
TRF2F/F p53-/- Cre-ER T1 + 4-OHT
C
TRF2F/F p53-/- Cre-ER T1 + 4-OHT
% T-loops
vec
wt
F59A
Δ70-114
D
vector
TRF2 wt
TRF2F59A
TRF2Δ70-114
TRF2F/F p53-/- Cre-ER T1 + 4-OHT
E
TRF2F/F p53-/- Cre-ER T1 + 4-OHT
% Telomeres Fused
F
vector
TRF2 wt
TRF2F59A
TRF2Δ70-114
TRF2F/F p53-/- Cre-ER T1 + 4-OHT
53BP1 + TTAGGG
G
TRF2F/F p53-/- Cre-ER T1 + 4-OHT
% of cells with 53BP1 TIFs
>10 TIFs
5-10 TIFs
H
vector
TRF2 wt
TRF2F59A
TRF2Δ70-114
TRF2F/F p53-/- Cre-ER T1 + 4-OHT
γH2AX + TTAGGG
I
TRF2F/F p53-/- Cre-ER T1 + 4-OHT
% of cells with γH2AX TIFs
>10 TIFs
5-10 TIFs
J
vector
TRF2 wt
TRF2F59A
TRF2Δ70-114
TRF2F/F p53-/- Cre-ER T1 + 4-OHT
Rap1 + TTAGGG
K
TRF2F/F p53-/- Cre-ER T1 + 4-OHT
% of cells >5 5 Rap1 - TelG Colocalizations
L
TRF2F/F p53-/- Cre-ER T1 + 4-OHT
TRF1 TRF2 TIN2 RAP1 PI INPUT
WT
Vec
F59A
Δ70-114
M
TRF2F/F p53-/- Cre-ER T1 + 4-OHT
% Telomeric DNA Normalized to WT TRF2
TRF1
TRF2
TIN2
Rap1
vec
wt
F59A
Δ70-114

To understand whether the laddering mutants are hypomorphs, we wanted to verify their localization to telomeres using IF. We found that both TRF2^{F59A} and TRF2$^{\Delta 70\text{-}114}$ did not localize well to telomeres in vivo, which may explain their inability to protect telomere ends (Fig. 21J,K). In addition, we performed a ChIP to more accurately quantify their DNA binding properties and found that TRF2$^{\Delta 70\text{-}114}$ exhibited a strong DNA binding defect, while TRF2^{F59A} had a mild DNA binding defect, consistent with the IF data (Fig.21L,M). Therefore, both TRF2$^{\Delta 70\text{-}114}$ and TRF2^{F59A} are hypomorphs.

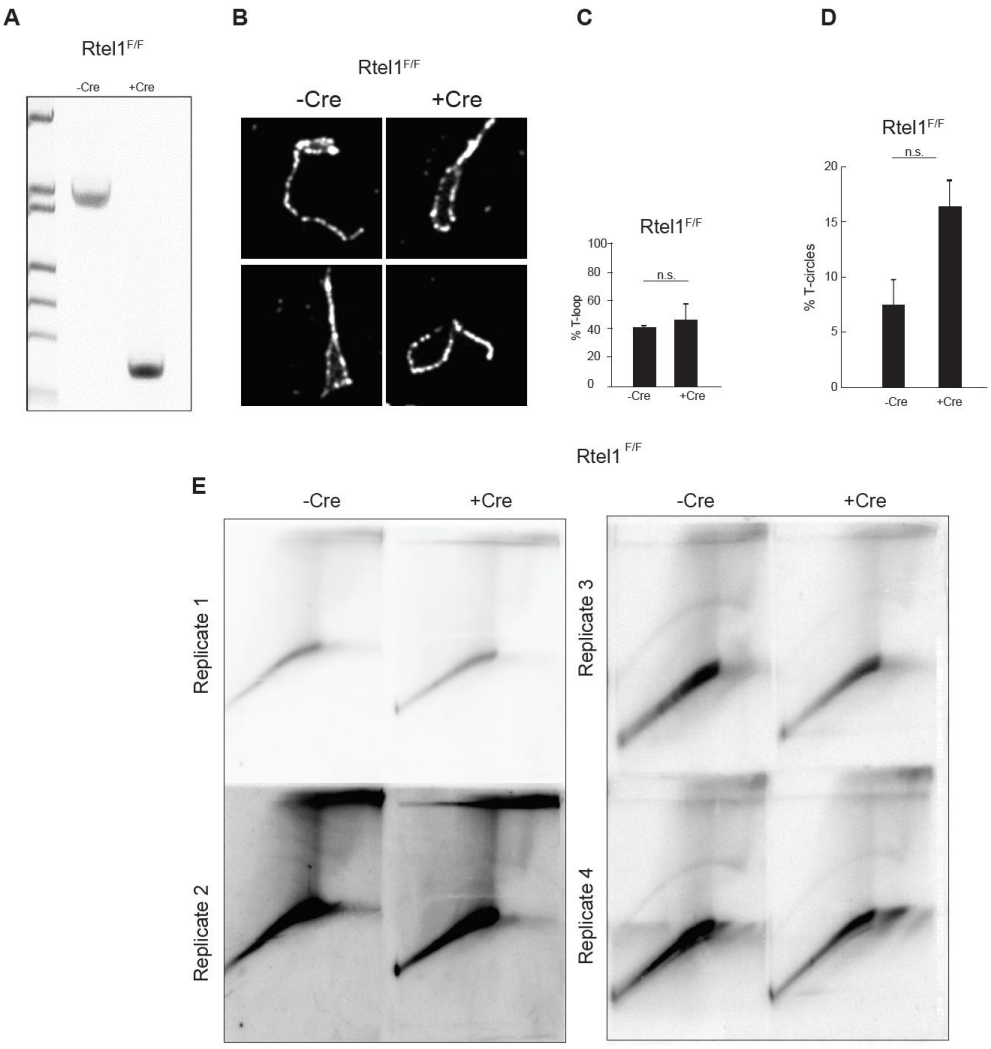
Loss of RTEL1 does not affect the frequency of t-loops

Having established that TRF2 is sufficient, we wanted to test whether TRF2 fulfills its function by recruiting non-shelterin components to telomeres for t-loop formation. To test what factors may be required for t-loop formation we began with RTEL1, a helicase strongly implicated in t-loop disassociation. Based on published work, we predicted that loss of RTEL1 should increase the frequency of t-loops due to the predicted role of RTEL1 in removing t-loops for replication or for telomerase action. For this study we used RTEL1$^{F/F}$ MEFs which were generously provided by the Boulton lab and had been generated Xiaoli Wu (*188*).

To examine the effects of RTEL1 deletion, we treated MEFs with Cre-virus and harvested them 120 hours after Cre-mediated deletion. Due to the lack of a functional commercial antibody, RTEL1 deletion was verified by a PCR reaction across the floxed exon 7. A smaller PCR fragment corresponds to successful excision of the exon from the genome (Fig. 22A).

Figure 22. RTEL1 is not required for maintenance of t-loops

(A) PCR across the floxed exon of RTEL1, showing successful excision of exon 7 120 hours after addition of tamoxifen. (B) Representative images of DNA spreads showing t-loops in RTEL1$^{F/F}$ MEFs, imaged using super-resolution OMX microscopy, 120 hours after tamoxifen induced Cre deletion. (C) Quantification of B, n = 3 independent experiments, 100 countable molecules per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$). (D) Quantification of the number of "t-circles" found, with any loop being defined as a t-circle if the loop exceeded 85% of the total length of the telomeric molecule. (E) 2-D Gels of RTEL1$^{F/F}$ cells, labeled with 32-P Sty11 telomeric probe. No telomeric abnormalities were detected by 2D gel.


A
Rtel1F/F
-Cre
+Cre
B
Rtel1F/F
-Cre
+Cre
C
Rtel1F/F
% T-loop
n.s.
-Cre
+Cre
D
Rtel1F/F
% T-circles
n.s.
-Cre
+Cre
E
Rtel1 F/F
-Cre
+Cre
-Cre
+Cre
Replicate 1
Replicate 2
Replicate 3
Replicate 4

Once deletion of RTEL1 was verified, we performed a t-loop assay and observed no statistically significant change in the frequency of t-loops in RTEL1 deleted compared to RTEL1 proficient cells (Fig. 22B,C). T-loops were present at high frequency in both samples, comparable to t-loop frequencies seen in other cell lines. It is possible that RTEL1 produces only a subtle effect, and therefore falls below the detection limit. We noted t-circles by OMX imaging, though we did not see any such arcs on 2D gels (Fig. 22D,E).

RAD51 is not required for t-loop formation

Given the fact that a t-loop contains a D-loop at the base, we hypothesized that RAD51 may play a role in t-loop formation. We first tested whether there was any interaction between shelterin and RAD51 and found that TRF2 interacts with RAD51 in co-immunoprecipitation assays (Fig. 23A,B). We wanted to test whether this interaction was phosphorylation dependent and found that the RAD51-TRF2 interaction was strengthened by the action of lambda phosphatase and weakened by the addition of PhosSTOP (Fig. 23C).

We next wanted to test whether RAD51 had an in vivo role in t-loop formation. Unfortunately, RAD51 is an essential gene, and no conditional knockout mice have been generated (*164*). Therefore, to delete RAD51, we chose to use a CRISPR knockout strategy that targeted the bulk population of cells. LentiCRISPRv2 allowed us to use a single plasmid that contained both the guide RNA to RAD51 as well as Cas9, and by producing a lentivirus we were able to

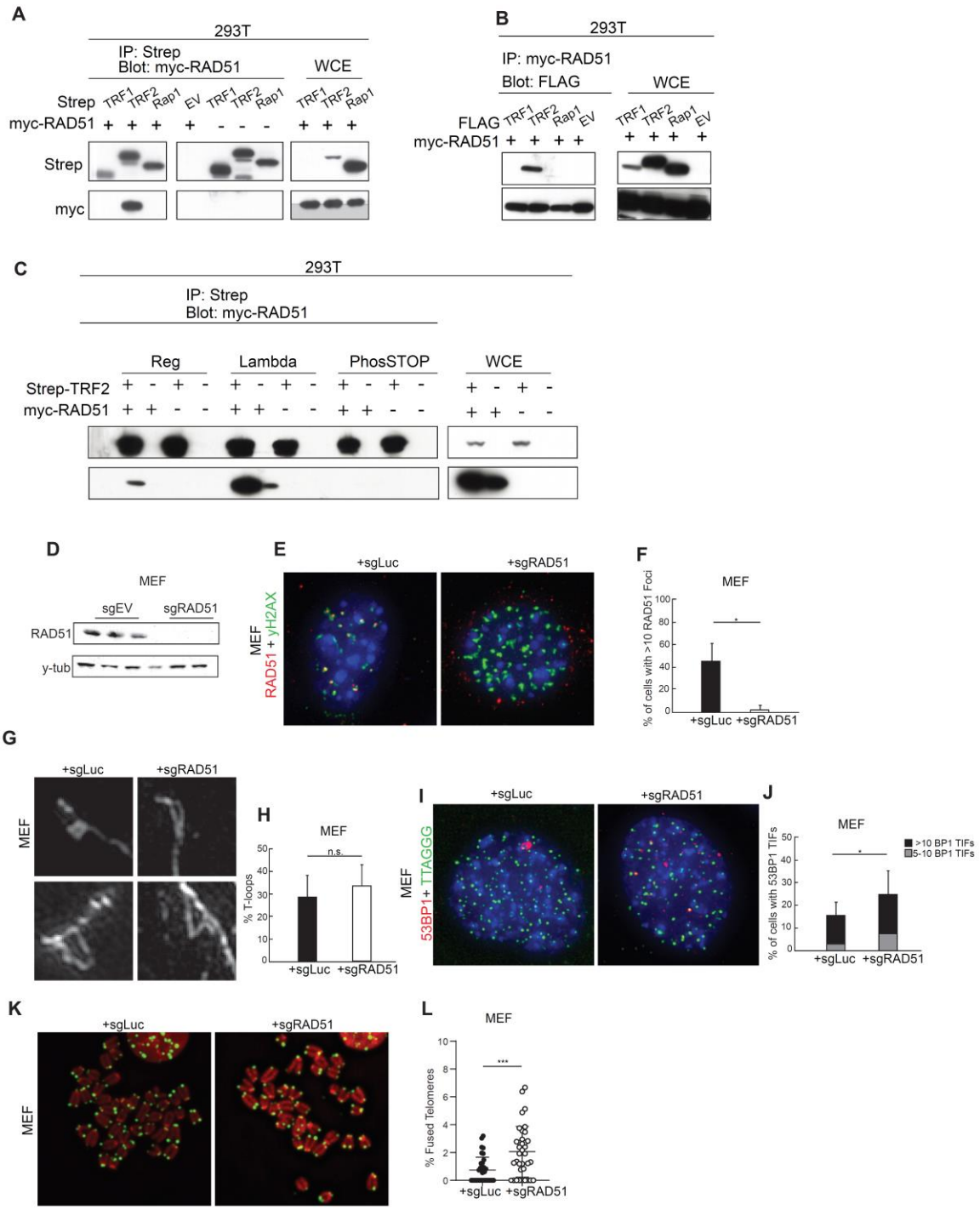
target cells in a way similar to using a Cre-virus to mediate deletion of a floxed gene (*189*).

We began by targeting wild type MEFs with the sgRAD51, and harvesting cells 120 hours after deletion, consistent with our protocol for other essential genes. Deletion of RAD51 was confirmed by immunoblotting (Fig. 23D). An additional functional assay for RAD51 was performed by irradiating cells with gamma rays and observing the colocalization of RAD51 with DDR γH2AX foci as cells attempted to repair the damage. We observed a significant reduction in RAD51 foci, consistent with deletion of RAD51 (Fig. 23E,F).

We next performed a t-loop assay to test whether or not RAD51 is required for t-loop formation. After super-resolution imaging, we observed a high frequency of t-loops both in cells that received a control luciferase sgRNA construct, and in cells that were targeted with an sgRAD51 construct (Fig. 23G,H). We did not note a statistically significant different between the two groups, which suggests that RAD51 does not play a role in t-loop formation, or if it does, it is not an essential role but rather redundant with another factor.

Figure 23. RAD51 is not required for t-loop formation.

(A) Co-immunoprecipitation of Strep tagged shelterin components (TRF1, TRF2, TIN2, Rap1) and Myc tagged RAD51 or vector. (B) Co-immunoprecipitation of myc tagged RAD51 and FLAG tagged shelterin components (TRF1, TRF2, TIN2, Rap1). (C) Co-immunoprecipitation of Strep tagged shelterin components (TRF1, TRF2, TIN2, Rap1) and Myc tagged RAD51 or vector in the presence of 400 U Lambda Phosphotase or 1X PhosSTOP (Roche). (D) Western blot of wild type MEFs showing expression of RAD51 (top) or gamma tubulin (bottom) after expression of an empty vector (left) or a CRISPR guide targeting RAD51 (right). (E) Representative images of wild type cells expressing either a CRISPR guide targeting luciferase or a CRISPR guide targeting RAD51 for 120 hours, stained for colocalization of γH2AX (Green) and RAD51 (red) after exposure to 4 Gy or ionizing radiation and allowed 4 hours to recover (F) Quantification of B, n = 3 independent experiments, 100 cells per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$). (G) Representative images of DNA spreads showing t-loops in wild type MEFs, imaged using super-resolution OMX microscopy, 120 hours after infection with either a luciferase targeting CRISPR guide or a RAD51 targeting CRISPR guide. (H) Quantification of D, n = 3 independent experiments, 100 countable molecules per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$). (I) Representative images of TIFs in wild type MEFs 120 hours after infection with either a luciferase targeting CRISPR guide or a RAD51 targeting CRISPR guide. Blue – DAPI. Green – Telomeres. Red – 53BP1. (J) Quantification of F, showing cells with 5-10 or more than 10 53BP1 TIFs. n = 4 independent experiments, 100 cells per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$) (K) Representative images of metaphases in in wild type MEFs 120 hours after infection with either a luciferase targeting CRISPR guide or a RAD51 targeting CRISPR guide. Red – DAPI. Green – Alexa647-(TTAGGG)3 (L) Quantification of H, n = 3 independent experiments, 2000 telomere ends per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$)


A
293T
IP: Strep
Blot: myc-RAD51
WCE
Strep
TRF1 TRF2 Rap1 EV TRF1 TRF2 Rap1 TRF1 TRF2 Rap1
myc-RAD51 + + + + - - - + + +
Strep
myc
B
293T
IP: myc-RAD51
Blot: FLAG
WCE
FLAG TRF1 TRF2 Rap1 EV TRF1 TRF2 Rap1 EV
myc-RAD51 + + + + + + + +
C
293T
IP: Strep
Blot: myc-RAD51
Reg
Lambda
PhosSTOP
WCE
Strep-TRF2 + - + - + - + - + - + - + - + -
myc-RAD51 + + - - + + - - + + - - + + - -
D
MEF
sgEV
sgRAD51
RAD51
y-tub
E
+sgLuc
+sgRAD51
MEF
RAD51 + γH2AX
F
MEF
% of cells with >10 RAD51 Foci
+sgLuc +sgRAD51
G
+sgLuc
+sgRAD51
MEF
H
MEF
n.s.
% T-loops
+sgLuc +sgRAD51
I
+sgLuc
+sgRAD51
MEF
53BP1 + TTAGGG
J
MEF
>10 BP1 TIFs
5-10 BP1 TIFs
% of cells with 53BP1 TIFs
+sgLuc +sgRAD51
K
+sgLuc
+sgRAD51
MEF
L
MEF
***
% Fused Telomeres
+sgLuc +sgRAD51

We also tested whether RAD51 is involved more generally in telomere end protection. We looked for the presence of 53BP1 TIFs in cells after RAD51 deletion, a sign of telomere dysfunction. Although we saw an increase in the number of 53BP1 TIFs, it was not statistically significant and it was consistent with a general increase in 53BP1 foci, likely due to the important role RAD51 has in repairing DNA damage (Fig. 23I,J). We suspect that this increase in TIFs is predominantly due to random colocalizations. We also looked directly at metaphases following RAD51 deletion, and also did not observe any striking phenotypes that would be suggestive of RAD51 being involved in telomere end-protection. We saw a very slight increase in telomere fusions which is unlikely to be physiologically relevant (Fig. 23K,L).

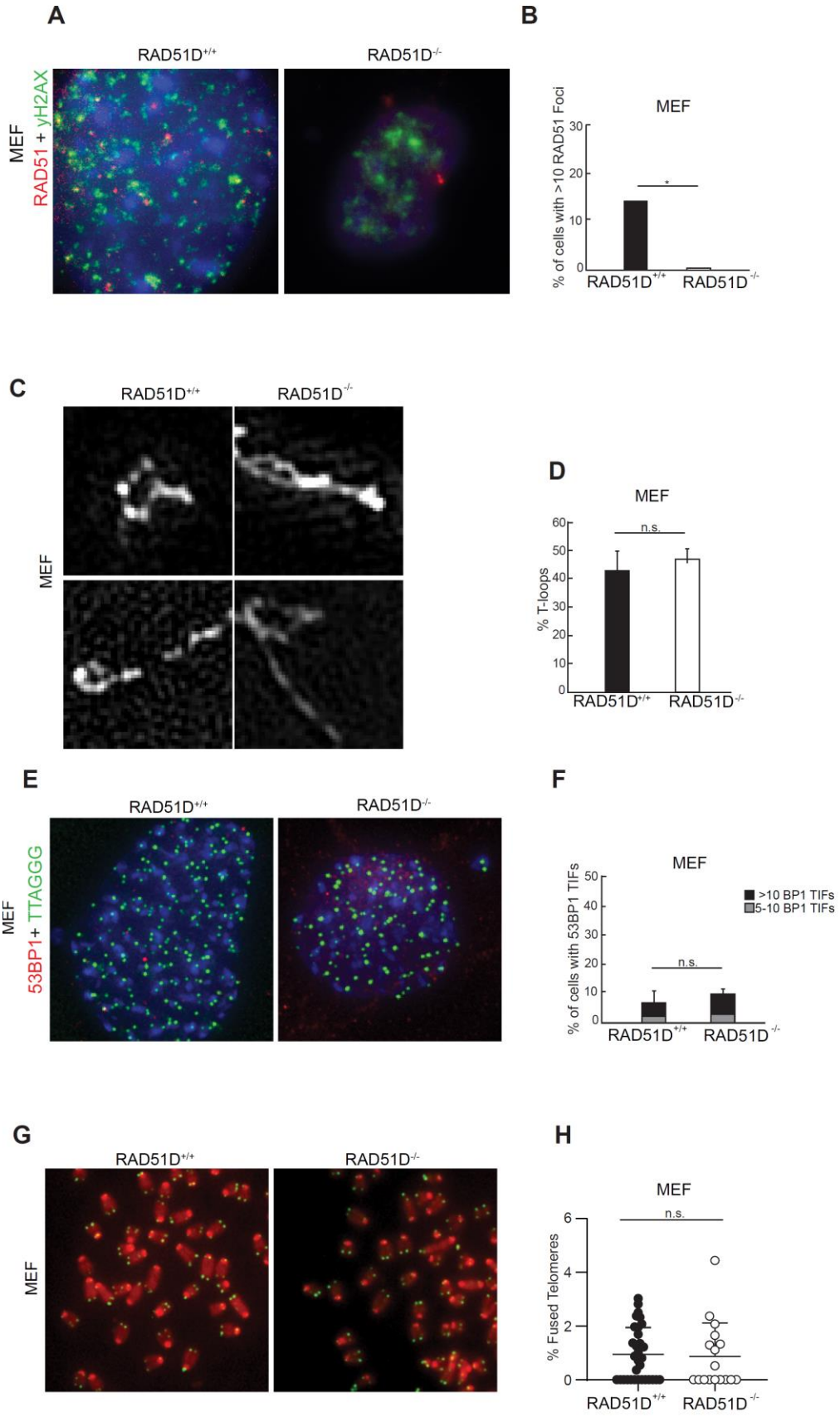
Overall, all our observations following RAD51 deletion are most consistent with RAD51 being a known DNA repair factor, rather than any novel function of RAD51 in telomere protection or t-loop formation.

RAD51D is not required for t-loop formation

RAD51 has a number of paralogs that assist in the filament formation and strand invasion functions of RAD51. Although we were able to show that RAD51 did not play a role in t-loop formation, the RAD51D paralog has been implicated in a telomere overhang maintenance role, and is also known to be essential (*175*, *176*). We were interested in testing whether RAD51D may be required for t-loop formation.

Figure 24. RAD51D is not required for t-loop formation.

(A) Representative images of RAD51D$^{+/+}$ or RAD51D$^{-/-}$ cells stained for colocalization of γH2AX (Green) and RAD51 (red) after exposure to 4 Gy or ionizing radiation and allowed 4 hours to recover (B) Quantification of B, n = 3 independent experiments, 100 cells per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$) (C) Representative images of DNA spreads showing t-loops in RAD51D$^{+/+}$ or RAD51D$^{-/-}$ MEFs, imaged using super-resolution OMX microscopy. (D) Quantification of D, n = 3 independent experiments, 100 countable molecules per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$) (E) Representative images of TIFs in RAD51D$^{+/+}$ or RAD51D$^{-/-}$ MEFs. Blue – DAPI. Green – Telomeres. Red – 53BP1. (F) Quantification of F, showing cells with 5-10 or more than 10 53BP1 TIFs. n = 4 independent experiments, 100 cells per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$) (G) Representative images of metaphases in in RAD51D$^{+/+}$ or RAD51D$^{-/-}$ MEFs. Red – DAPI. Green – Alexa647-(TTAGGG)3 (H) Quantification of G, n = 3 independent experiments, 2000 telomere ends per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$)


A
RAD51D+/+
RAD51D-/-
MEF
RAD51 + γH2AX
B
MEF
% of cells with >10 RAD51 Foci
30
20
10
0
*
RAD51D+/+
RAD51D-/-
C
RAD51D+/+
RAD51D-/-
MEF
D
MEF
% T-loops
60
50
40
30
20
10
0
n.s.
RAD51D+/+
RAD51D-/-
E
RAD51D+/+
RAD51D-/-
MEF
53BP1+ TTAGGG
F
MEF
% of cells with 53BP1 TIFs
50
40
30
20
10
0
>10 BP1 TIFs
5-10 BP1 TIFs
n.s.
RAD51D+/+
RAD51D-/-
G
RAD51D+/+
RAD51D-/-
MEF
H
MEF
% Fused Telomeres
6
4
2
0
n.s.
RAD51D+/+
RAD51D-/-

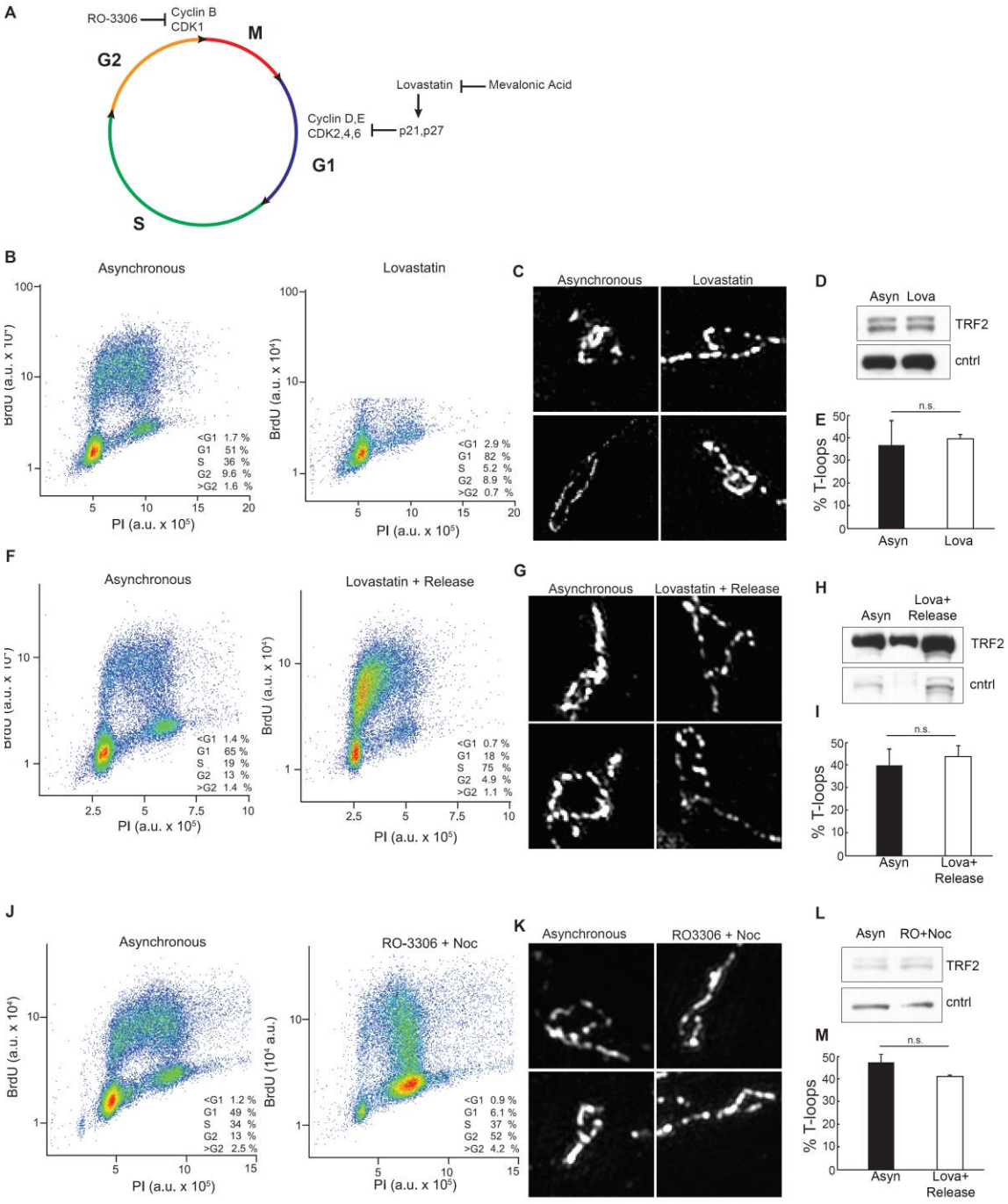
We received RAD51D knockout and control MEFs courtesy of the Pittman lab and verified that RAD51D is deleted by a functional assay, due to the lack of commercially available antibodies (*176*). RAD51D is required for the formation of RAD51 foci after gamma radiation induced damage (*175*, *190*, *191*). To that end, we irradiated RAD51D$^{-/-}$ p53$^{-/-}$ MEFs and their control counterparts and assayed for colocalizations between γH2AX and RAD51. We observed a significant decrease in RAD51 foci formation in the RAD51D-deficient cells, consistent with the cells having lost RAD51D (Fig. 24A,B).

We followed this up with a t-loop assay to test whether RAD51D is necessary for t-loop formation. As with RAD51, we saw a high level of t-loops in both cell lines, and no significant different between the two conditions (Fig. 24C,D). This points towards the fact that like RAD51, RAD51D does not play an important role in t-loop formation.

Finally, we wanted to see whether RAD51D contributed to general telomere maintenance. We assayed for TIF formation as well as chromosome fusions. We found no significant difference between RAD51D deficient or RAD51 proficient cells when assayed for 53BP1 TIFs or for chromosome fusions (Fig. 24E-H). This is consistent with previously published data that suggested either no phenotype, or a mild telomeric phenotype in cells lacking RAD51D that were otherwise untreated with any DNA damaging agents (*175*, *176*).

Figure 25. T-loops Are Present Throughout the Cell Cycle

(A) Schematic illustrating the action of Lovastatin and RO-3306 in the cell cycle. (B) FACS analysis of the cell cycle profile of asynchronous MEFs or MEFs treated with Lovastatin. (C) Representative images of DNA spreads showing t-loops in MEFs after 36 hours of 40μM Lovastatin treatment. (D) Western blot showing expression of TRF2 after treatment of MEFs with Lovastatin. (E) Quantification of DNA spreads showing t-loops in MEFs after 36 hours of 40μM Lovastatin treatment. n = 2 independent experiments, 100 countable molecules per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$) (F) FACS analysis of the cell cycle profile of asynchronous MEFs or MEFs treated with Lovastatin and released into Mevalonic Acid. (G) Representative images of DNA spreads showing t-loops in MEFs after 36 hours of 40μM Lovastatin treatment followed by 15 hours of release into 400 μM Mevalonic Acid(H) Western blot showing expression of TRF2 after treatment of MEFs with Lovastatin and subsequent release. (I) Quantification of DNA spreads showing t-loops in MEFs after 36 hours of 40μM Lovastatin treatment followed by 15 hours of release into 400 μM Mevalonic Acid. n = 2 independent experiments, 100 countable molecules per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$) (J) FACS analysis of the cell cycle profile of asynchronous MEFs or MEFs treated with RO-3306 and released into Nocodazole. (K) Representative images of DNA spreads showing t-loops in MEFs after 12 hours of 9μM RO-3306 treatment and 3 hours release into Nocodazole. (L) Western blot showing expression of TRF2 after treatment of MEFs with RO-3306 and Nocodazole. (M) Quantification of DNA spreads showing t-loops in MEFs after 12 hours of 9μM RO-3306 treatment and 3 hours release into Nocodazole. n = 2 independent experiments, 100 countable molecules per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$)


A
RO-3306
Cyclin B
CDK1
M
G2
Lovastatin
Mevalonic Acid
Cyclin D,E
CDK2,4,6
p21,p27
G1
S
B
Asynchronous
Lovastatin
BrdU (a.u. x 10⁴)
PI (a.u. x 10⁵)
<G1 1.7 %
G1 51 %
S 36 %
G2 9.6 %
>G2 1.6 %
<G1 2.9 %
G1 82 %
S 5.2 %
G2 8.9 %
>G2 0.7 %
C
Asynchronous
Lovastatin
D
Asyn Lova
TRF2
cntrl
E
n.s.
% T-loops
Asyn
Lova
F
Asynchronous
Lovastatin + Release
<G1 1.4 %
G1 65 %
S 19 %
G2 13 %
>G2 1.4 %
<G1 0.7 %
G1 18 %
S 75 %
G2 4.9 %
>G2 1.1 %
G
Asynchronous
Lovastatin + Release
H
Asyn
Lova+ Release
TRF2
cntrl
I
n.s.
% T-loops
Asyn
Lova+ Release
J
Asynchronous
RO-3306 + Noc
BrdU (10⁴ a.u.)
<G1 1.2 %
G1 49 %
S 34 %
G2 13 %
>G2 2.5 %
<G1 0.9 %
G1 6.1 %
S 37 %
G2 52 %
>G2 4.2 %
K
Asynchronous
RO3306 + Noc
L
Asyn
RO+Noc
TRF2
cntrl
M
n.s.
% T-loops
Asyn
Lova+ Release

T-loops are present in G1, S, and G2/M

We hypothesized that if t-loops are regulated in a cell cycle dependent manner, it may lend mechanistic insights into what factors are involved in t-loop formation. In addition, given the critical nature of t-loops in protecting telomeres from ATM and NHEJ, we wanted to investigate whether or not t-loops are present throughout the cell cycle, or whether they become unwound for extended periods of time in G_1, G_2/M, or S phase, which would conflict with the t-loop model. We used Lovastatin for G1 arrest and release into S phase, and RO-3306 along with Nocodazole for G2/M arrest (Fig. 25) (*192–194*).

We began by testing whether G_1 or S phase cells still maintained their t-loops. We arrested cells in G_1 with Lovastatin and saw a large increase in G_1 cells, as expected (Fig. 25B-E). TRF2 was still present in G_1 cells at normal levels, and consistent with the t-loop model, t-loops were unaffected (Fig. 25C-E). To synchronize cells in S phase, cells were released from Lovastatin arrest with mevalonic acid. TRF2 levels were unaffected by this treatment, and we observed a significant increase in S phase cells 16 hours after release (Fig. 25F-I). As before, t-loops were still present in S phase (Fig. 25G-I).

We next tested whether t-loops were present in G_2/M by arresting cells with RO-3306 and releasing them into nocodazole. We observed a significant enrichment of cells in G_2/M and TRF2 levels were unaffected (Fig. 25J-M). T-loops were also still present at levels comparable to asynchronous cells in this setting (Fig. 25K-M).

Combined with the previous data, this suggests that as long as TRF2 is present, it is capable of maintaining t-loops throughout the cell cycle. This is consistent with the fact that telomeres do not activate ATM at any specific point in the cell cycle.

Discussion

Here we tested a number of hypotheses regarding t-loop formation, such as the requirement of the basic domain, TRF2 DNA wrapping, and TRF2 tetramerization. We began by testing whether the branched DNA binding function of TRF2's basic domain of TRF2 is required for t-loop formation (*60*). Loss of the basic domain is known to produce t-loop sized cleavage products in certain cell lines, as well as lead to telomere shortening. However, loss of the basic domain does not lead to an ATM response nor to telomere fusions. We tested this and recapitulated previously published data, as well as noted that the basic domain is not required for t-loop formation. Evidently it is required to bind the base of the D-loop and protect it from cleavage, but it is not required to form it. This logically makes sense – one cannot produce t-loop cleavage products if one does not first produce t-loops and have the capacity to reform them. It is possible that the basic domain acts as a pin to hold the D-loop and t-loop in place and prevent slippage and consequently access to nucleases and branch-resolving helicases.

We next tested the hypothesis that DNA wrapping around the TRFH domain is essential for t-loop formation. Our data was generally consistent with previously published work suggesting that Topless mutants are unable to form t-loops and

are unable to suppress ATM (*61*). We found that when wild type TRF2 was replaced by a Topless mutant, that ATM was activated and no t-loops were present, but critically, we did see telomere fusions.

We noted that the Topless mutants were poorly expressed, though still greatly overexpressed compared to wild type TRF2. However, their localization to telomeres was poor, and their binding was not strong in vivo. This raises the possibility that the Top domain is required for proper DNA binding rather than t-loop formation. Such a result would not be surprising, given the extensive nature of the proposed contacts between the Top domain and telomeric DNA. Unfortunately, this causes Topless to behave as a hypomorph in our experiments, and it is unclear whether the phenotypes we observe are due to specific wrapping disruptions of the Top domain, or due to its poor DNA binding properties. If it is the latter – then we simply cannot distinguish whether the mutations in the Top domain are functionally relevant or just phenocopy low TRF2 levels.

We also tested a variant of TRF2^{Topless}, the 7R2K version of the Topless mutant. We had two competing hypotheses: first, if these charged residues were required for wrapping then switching them should not impact the function of the wild type protein because charge is maintained. A second hypothesis was that the side-chain length may be critical to proper wrapping and disrupting it would lead to a phenotype similar to Topless. However, we observed that TRF2^{7R2K} was able to localize to telomeres, protect it from ATM and NHEJ, but on first inspection did not form t-loops.

Such a result would challenge the t-loop model, as it would call into question whether t-loops are required for end-protection. Therefore, we decided to investigate whether the size of any residual loops in the sample were different. Unexpectedly, we found that these residual loops were statistically significantly smaller. Therefore, we cannot exclude the possibility that the 7R2K mutant makes more tightly wound t-loops which would appear to us as linear molecules or molecules we discard because we lack the resolution to see a loop. Unfortunately, until a separation of function mutant, or a system to generate an artificial t-loop is found, this question cannot be decisively answered.

Another key difference in our data is that it was previously reported that Topless was capable of suppressing telomere fusions, except in the absence of Rap1 (*61*). We did not observe this in our genetic setting since Topless was incapable of repressing c-NHEJ with or without Rap1.

Next, we tested whether TRF2 tetramerization is required for t-loop formation by using two tetramerization mutants. These mutants were first described in unpublished work from this lab, performed by Heidi Moss. We hypothesized that these mutants were defective in tetramerization of TRF2, and that this tetramerization is essential to t-loop formation. Of the two mutants we used, $TRF2^{F59A}$ had the smallest perturbation of the protein, while $TRF2^{\Delta 70\text{-}114}$ had the strongest laddering phenotype. We were able to successfully reproduce the laddering phenotype, but unfortunately in our experiments, the laddering mutants exhibited poor telomeric localization. $TRF2^{F59A}$ localized slightly better than $TRF2^{\Delta 70\text{-}114}$, but this nonetheless complicates our interpretation of the data. It is

clear that the laddering mutants produce a deprotection phenotype: both mutants have an increase in ATM signaling, and both mutants exhibit some level of fusions. The TRF2$^{\Delta 70\text{-}114}$ mutant is clearly a hypomorph – it performs no better than an empty vector at protecting telomeres, lacks t-loops entirely, has a full signaling phenotype, and localizes poorly to telomeres by both ChIP and IF. As an interesting aside – it is notable that although both laddering mutants do not localize perfectly to telomeres, they have no expression issues.

The TRF2^{F59A} mutant is more nuanced but most likely a hypomorph as well. By IF it localizes worse than wild type TRF2 to telomeres, despite having strong expression, but by ChIP it appears to be at telomeres at near wild type levels. Despite this, it only partially protects telomeres from ATM signaling, with 20-40% showing TIFs, and it only partially protects from fusions, with about 15% of telomeres being fused. However, the level of t-loops was not significantly different from the empty vector. This leaves two possible explanations for the data: TRF2^{F59A} is a hypomorph, or TRF2^{F59A} is important for t-loop formation. If TRF2^{F59A} is unable to tetramerize, this may affect its ability to bind telomeric DNA, and may explain its somewhat lower in vivo telomeric localization. The result would be that TRF2^{F59A} behaves like a hypomorph. However, one can speculate that oligomerization and t-loop formation itself may act to stabilize TRF2 at the telomere, and failure of TRF2^{F59A} to form a t-loop contributes to its poor localization. However, this result would challenge the t-loop model because it would suggest that t-loops are not sufficient for protection from ATM or c-NHEJ, as TRF2^{F59A} does not have a complete deprotection phenotype. In fact, it would suggest that t-loops

partially contribute to the repression by TRF2, the majority of which is likely provided by TRF2 itself, and would echo the results originally published on both the Topless mutants, as well as data from the Cesare lab (*145*). Mechanistically, we can imagine that TRF2 can bind at some location on the t-loop and at a site near the 3' overhang as a dimer, and subsequent tetramerization of TRF2 can promote strand invasion. The telomeric wrapping induced by the Top domain can melt the DNA and allow the 3' end access to form a D-loop.

Unfortunately, $TRF2^{F59A}$ is most likely a hypomorph. The localization data shows a defect in binding, and protection at telomeres is acutely sensitive to TRF2 levels. The loss of TRF2 binding somewhat correlates with the change level of fusions, and the mild fusion phenotype in turn correlates with the level of TIFs we observe. This means that we cannot rule out the possibility that the entire phenotype of $TRF2^{F59A}$ is due to its subpar telomeric localization.

In addition, we tested whether multiple different protein factors or cell cycle stages are important in maintaining telomere end protection through t-loop formation. We found that neither RAD51, RAD51D, nor RTEL1 contribute to telomere end-protection or t-loop formation in any meaningful way. Additionally, we found that the cell cycle stage does not impact t-loop formation.

We tested whether the RTEL1 helicase is involved in t-loop disassociation. We found that loss of RTEL1 did not significant decrease the frequency of t-loops, consistent with the hypothesis that RTEL1 is not required for t-loop disassociation. Rather, we believe that the replicative helicase CMG is more than capable of unwinding the t-loop.

The finding that RAD51 was not required for t-loop formation is interesting because RAD51 is known to be required for D-loop formation, and the base of a t-loop is in a D-loop conformation. This suggests that either the base of the t-loop is not a D-loop, or TRF2 has the ability to strand invade in a similar manner to a recombinase. The latter is consistent with our finding that TRF2 is sufficient for t-loop formation. This finding also rules out a possible explanation for why RAD51 is essential in mammals and not in yeast. Since t-loops are not formed by RAD51, this cannot be its essential function. This also suggests that any involvement of BRCA2 in telomere end protection is not likely to be due to BRCA2's role in RAD51 loading, but perhaps involve BRCA2's role in protection from end-resection (*172*). Although we showed an interaction between RAD51 and TRF2, the biological relevance of this is unclear.

Additionally, we showed that the essential RAD51 paralog RAD51D is not required for t-loop formation. This is another strike against a model that calls for a recombinase to mediate t-loop formation. One additional factor that we did not thoroughly test, but one that has been suggested to play a role, is RAD52.

We also showed that cell cycle is not a meaningful factor in t-loop formation. This suggests that t-loops are not synchronously unwound and deprotected in S phase. The data also is consistent with the idea that t-loops are able to be constantly reformed. Together this shows that t-loops are most likely unwound as needed for replication in S phase, and then quickly reformed. If t-loops were unwound during S phase in a coordinated manner, we would expect to see a decrease in the frequency of t-loops, and the same goes if they were unwound

when needed but only reformed in a specific moment. We cannot exclude that we simply do not have the sensitivity in our assay to detect such subtle changes, however, we can exclude the possibility that t-loops as a whole are unwound during S phase. The most likely explanation is that as each telomere is unwound from its t-loop conformation for replication, it is then processed for 3' overhang formation, and subsequently reforms into a t-loop by the action of TRF2. This would be consistent with data suggesting that telomeres do not replicate in a coordinated manner (*195*).

We have also demonstrated that t-loops are present in G1 and G2/M phase. First, this shows that the reason that historically observed t-loop levels have never been at 100% is not because the cells are in an asynchronous population that lacks t-loops in a certain cell cycle phase. Other reasons for the finding that t-loop frequency is always below 100% is that not every telomere ends in a t-loop, that the assay results either cannot detect very small loops, sample preparation results in broken telomeres, that Y-shapes are not scored, or that crosslinking is inefficient. However, we are able to exclude that a certain cell cycle phase lacks t-loops.

In addition, our data on G1 and G2/M cells suggests that t-loops can either reform during these stages of the cell cycle or are very stable. If for example cells arrest in G1 due to DNA damage, our data suggests that t-loops would still be protecting telomere ends. It also suggests that other arrested cells, such as quiescent or senescent cells, should also contain t-loops despite the fact that they are not cycling, as predicted by Griffith et al. in 1999 (*137*). This is consistent with

these cells not displaying a constitutively on DNA damage response at telomeres. It would be possible to test this by using a temperature sensitive allele of TRF2 to deprotect telomeres at specific cell cycle phases, and then test whether t-loops can be reformed in those phases.

In this study we have tested a number of different hypothetical mechanisms of t-loop formation: the basic domain, DNA wrapping, and TRF2 tetramerization. We also tested whether TRF2 requires exogenous factors. Although we were unable to determine how TRF2 forms t-loops, our data lends itself to the argument that TRF2 is the sole factor responsible for t-loop formation at telomeres and that it maintains t-loops throughout the cell cycle.

Chapter 5

The N-Terminally Extended Isoform of TRF2 is a Functionally Competent Member of the Shelterin Complex

Telomeric repeat binding factor 2 (TRF2) is a critical component of the telomere bound shelterin complex, responsible for suppressing ATM signaling as well as c-NHEJ mediated and other repair to prevent aberrant repair of the telomere. In metazoan cells, TRF2 is expressed and found in the shelterin complex in two forms, a short form and an N-terminally extended long isoform. This N-terminal tail is highly conserved, but its function remains unknown. Using a combination of cell biological and biochemical approaches, we show that $TRF2^{Long}$ can complement the deletion of both forms of TRF2 by suppressing telomere fusions by both classical and alternative non-homologous end-joining, forming t-loops, maintaining the telomeric overhang, and suppressing ATM. Furthermore, the N-terminal tail does not confer altered DNA binding or interactions with the rest of the shelterin complex. Our work demonstrates that both isoforms of TRF2 contribute the same functions to the shelterin complex, allowing it to solve the end protection problem. We also test a putative PIKK phosphorylation site on $TRF2^{Long}$. Unfortunately, the question of why the N-terminus is conserved remains unresolved.

Introduction

Many components of the shelterin complex have alternative splice forms or expression variants. For instance, TIN2 has a long and short isoform, of which the former strongly interacts with the nuclear matrix (*196*). In addition to TIN2, POT1 has alternative splice isoforms, as does TRF1 (*197*, *198*). The role of these isoforms in the shelterin complex is not clear, but some have been reported to be

differentially expressed across cell lines, and others, such as POT1, have different DNA binding affinities (*199*). Given presence of alternative splice variants of other shelterin components, it should come as no surprise that alternative splice variants of TRF2 have been identified before (*129*, *200–202*). In most mammalian cells, TRF2 is made in two forms: a dominant short version, and a somewhat less abundant N-terminally extended version which is the result of an alternative start site (Fig. 26A, B). The long form had been observed as second band in western blots, but was never fully characterized. This tail is highly conserved in metazoans, is similar in amino acid composition to the basic domain, and contains a putative S/T-Q PIKK phosphorylation site. The function of TRF2^{Long} function is unclear, but some reports suggest a non-telomeric role as it is expressed concurrent with neuronal axon differentiation where it has been proposed to regulate RNA transport (*203*). In this chapter, test whether both forms of TRF2 are functionally equivalent.

Results

TRF2^{Long} is proficient at blocking telomere fusions and can suppress ATM

Upon TRF2 deletion, c-NHEJ and ATM signaling are unleashed at telomere ends, which generates the most prominent phenotype of TRF2 deletion in MEFs: telomere end fusion and TIFs. To address the capabilities of the TRF2^{Long}, we designed a construct to express TRF2^{Long} but not TRF2^{Short}. This form of TRF2 has a 42 amino acid N-terminal extension, which has a 43% amino acid identity match to the basic domain as calculated by CLUSTAL (Fig. 26B) (*204*). As a control, we

first tested for endogenous expression of both isoforms in MEF cell lines and observed that TRF2^{Long} is expressed at comparable levels to TRF2^{Short} (Fig. 26C). Next, to test whether TRF2^{Long} can complement TRF2 deletion, we expressed it in a panel of cells: TRF2$^{F/F}$ to test for c-NHEJ suppression, TRF2$^{F/F}$ Lig4$^{-/-}$ to test for inhibition of ATM activity, and TRF2$^{F/F}$ Ku70$^{-/-}$ to test for suppression of alt-NHEJ fusions. (Fig. 26D-F). Cre deletion of TRF2 was done at 120 hours and removes both isoforms of TRF2. Both TRF2^{Short} and TRF2^{Long} are greatly overexpressed compared to endogenous levels of the proteins, which may occlude small defects in TRF2^{Long} function (Fig. 26D-F).

We first tested whether TRF2^{Long} could suppress TIF formation after deletion of TRF2 in TRF2$^{F/F}$ Lig4$^{-/-}$ MEFs. As discussed previously, one of TRF2's key functions is to repress ATM by blocking its ability to recognize a telomere end as a DSB (*205*). Because telomere fusions attenuate ATM signaling, TIFs were analyzed in Lig4$^{-/-}$ MEFs that do not fuse their telomeres. TIFs were scored on the basis of colocalization of 53BP1 foci with telomeric foci, 120 hours after deletion of endogenous TRF2. The analysis reveals that TRF2^{Long} can suppress the formation of 53BP1 foci (Fig. 26G, H). We observed both fewer TIFs on average per cell compared to the empty vector and fewer cells with many (>10) TIFs, indicating that TRF2^{Long} can suppress ATM. There was no significant difference between the efficacy of TIF suppression by TRF2^{Long} or TRF2^{Short}, indicating that TRF2^{Long} is fully competent at suppressing ATM signaling.

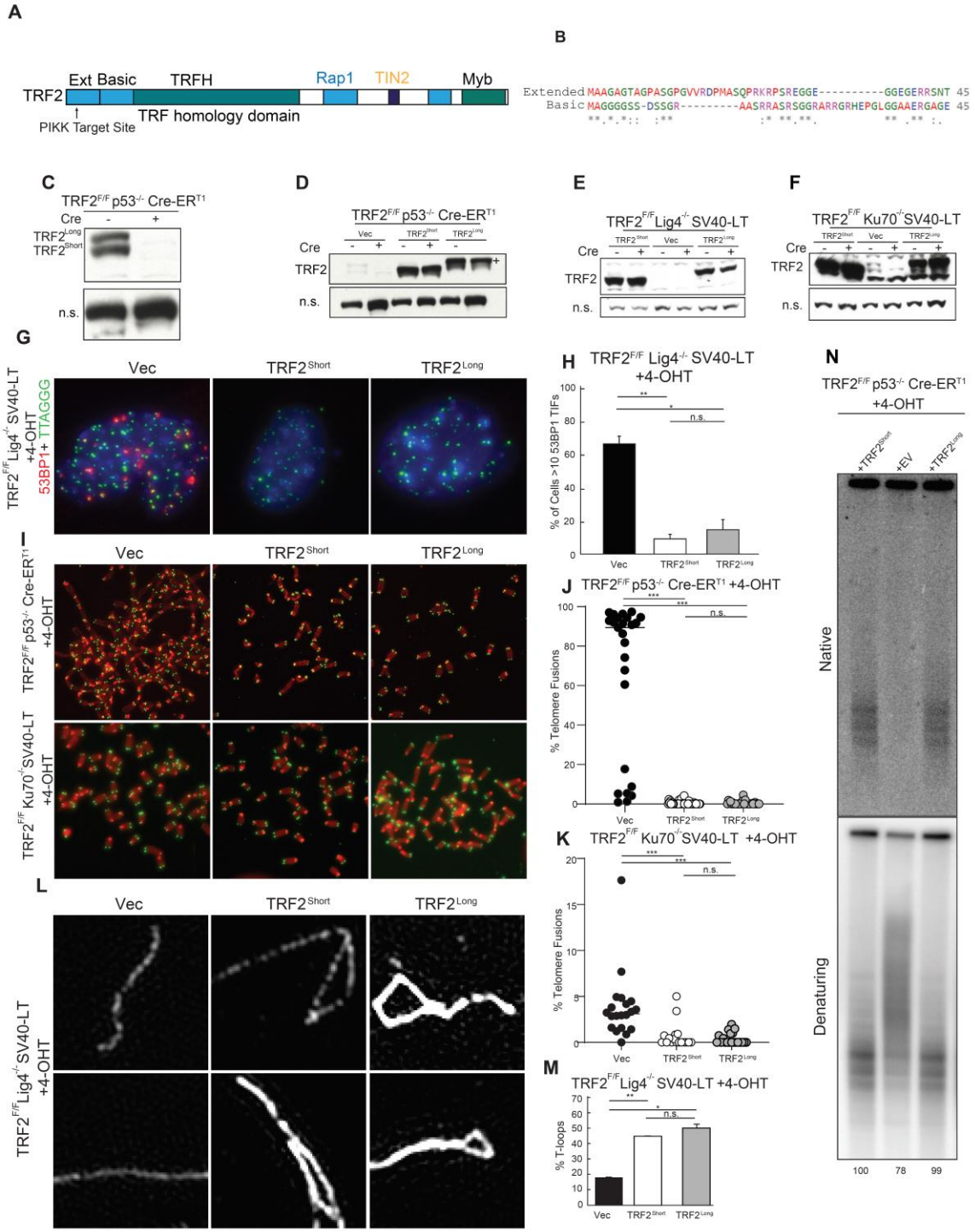
We next looked at whether TRF2^{Long} could suppress NHEJ fusions in a Ligase 4 proficient setting, or alt-NHEJ fusions in a Ku70-deficient setting. Ku70/80

is a general repressor of alt-NHEJ, and deletion of Ku70/80 allows us to observe alt-NHEJ repair phenotypes. 120 hours after TRF2 deletion, we observed that $TRF2^{Long}$ was capable of suppressing both c-NHEJ and alt-NHEJ with the same efficacy as $TRF2^{Short}$ (Fig. 26I-K). $TRF2^{Long}$ was slightly more effective at suppressing alt-NHEJ than $TRF2^{Short}$ but this difference was not statistically significant. In summary, there is no detectable difference in the ability of either isoform to repress c-NHEJ, alt-NHEJ, or ATM signaling.

One of the most striking features of TRF2 is its ability to catalyze the strand invasion of the telomeric 3' overhang to form a lariat structure known as a t-loop. To investigate whether $TRF2^{Long}$ can form t-loops, we expressed both isoforms of TRF2, deleted TRF2 by Cre-mediated recombination, 120 hours later UV/psoralen crosslinked purified nuclei, spread and hybridized the DNA with PNA probes and imaged the samples with OMX SIM. We observed that both $TRF2^{Long}$ and $TRF2^{Short}$ were able to form t-loops with high frequency (Fig. 26L, M). There was no statistically significant difference between t-loop frequency in the cells expressing the two isoforms, showing that they are both capable of catalyzing t-loop formation to an equal extent.

Figure 26. The N-Terminally Extended Isoform of TRF2 is a Functionally Competent Member of the Shelterin Complex.

(A) Schematic of the N-terminally extended isoform of TRF2. (B) CLUSTAL Omega protein sequence alignment between the N-terminal extended domain and the basic domain of TRF2. (C) Western blot from MEFs showing endogenous expression of both $TRF2^{Long}$ and $TRF2^{Short}$. TRF2 in the top panel, N.s. loading control in the bottom panel. (D) Western blot showing expression of $TRF2^{Short}$ and $TRF2^{Long}$ in MEFs immortalized with p53-loss and floxed for TRF2. TRF2 in the top panel, N.s. loading control in the bottom panel. (E) Western blot showing expression of $TRF2^{Short}$ and $TRF2^{Long}$ in MEFs immortalized with p53-loss and floxed for TRF2 and lacking Lig4. TRF2 in the top panel, N.s. loading control in the bottom panel. (F) Western blot showing expression of $TRF2^{Short}$ and $TRF2^{Long}$ in MEFs immortalized with p53-loss and floxed for TRF2 and lacking Ku70. TRF2 in the top panel, N.s. loading control in the bottom panel. (G) Representative images of TIFs in $TRF2^{F/F}$ Lig4-/- MEFs. Blue – DAPI. Green – Telomeres. Red – 53BP1. (H) Quantification of TIFs in $TRF2^{F/F}$ $Lig4^{-/-}$MEFs, showing cells with 10 or more TIFs 120 hours after tamoxifen induced Cre-deletion. n = 2 independent experiments, 100 cells per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$) (I) Representative images of metaphases. Top panel: $TRF2^{F/F}$ Bottom Panel: $TRF2^{F/F}$ $Ku70^{-/-}$ MEFs. Red – DAPI. Green – Alexa647-$(TTAGGG)_3$. (J) Quantification of metaphases in $TRF2^{F/F}$ MEFs, showing the percentage of telomeres fusions 120 hours after tamoxifen induced Cre-deletion. n = 3 independent experiments, 10 metaphases per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$). (K) Quantification of H, bottom panel, n = 2 independent experiments, 10 metaphases per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$). (L) Representative images of DNA spreads showing t-loops in $TRF2^{F/F}$ $Lig4^{-/-}$ MEFs, imaged using super-resolution OMX microscopy, 120 hours after tamoxifen induced Cre deletion. (M) Quantification of K, n = 2 independent experiments, 100 countable molecules per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$) (N) TRF analysis in $TRF2^{F/F}$ cells 120 hours after tamoxifen induced Cre-mediated deletion. Left panel, native gel. Right panel, denaturing gel. Relative quantification of the overhang (normalized to $TRF2^{Short}$) are shown below the right panel.


A
Ext Basic
TRFH
Rap1
TIN2
Myb
TRF2
PIKK Target Site
TRF homology domain
B
Extended MAAGAGTAGPASGPGVVRDPMASQPRKRPSREGGE---------GGEGERRSNT 45
Basic MAGGGGSS-DSSGR---------AASRRASRSGGRARRGRHEPGLGGAAERGAGE 45
C
TRF2F/F p53-/- Cre-ERT1
Cre
TRF2Long
TRF2Short
n.s.
D
TRF2F/F p53-/- Cre-ERT1
Vec
TRF2Short
TRF2Long
Cre
TRF2
n.s.
E
TRF2F/F Lig4-/- SV40-LT
F
TRF2F/F Ku70-/-SV40-LT
G
Vec
TRF2Short
TRF2Long
TRF2F/F Lig4-/- SV40-LT +4-OHT
53BP1+ TTAGGG
H
TRF2F/F Lig4-/- SV40-LT +4-OHT
% of Cells >10 53BP1 TIFs
n.s.
I
TRF2F/F p53-/- Cre-ERT1 +4-OHT
TRF2F/F Ku70-/-SV40-LT +4-OHT
J
TRF2F/F p53-/- Cre-ERT1 +4-OHT
% Telomere Fusions
K
TRF2F/F Ku70-/-SV40-LT +4-OHT
L
TRF2F/F Lig4-/- SV40-LT +4-OHT
M
TRF2F/F Lig4-/- SV40-LT +4-OHT
% T-loops
N
TRF2F/F p53-/- Cre-ERT1 +4-OHT
+TRF2Short
+EV
+TRF2Long
Native
Denaturing
100
78
99

To determine the status of the telomere end, we tested whether TRF2^{Long} is capable of properly maintaining the telomeric overhang length. To this end, performed CHEF gel electrophoresis on telomeric restriction fragments of TRF2$^{F/F}$ MEFs expressing either TRF2^{Long} or TRF2^{Short} on a pulse field gel and hybridized the DNA with a radiolabeled CCCTAA$_{(4)}$ probe under native conditions followed by denaturing conditions (Fig. 26N). After normalizing the total TTAGGG signal to the native DNA, we found that the telomeric overhang in TRF2^{Long} expressing cells the same as TRF2^{Short}. This finding is consistent with the ability of TRF2^{Long} to interact with Apollo and other shelterin members.

Mutations of a conserved ST/Q site do not affect the ability of TRF2^{Long} to repress ATM.

Sequence alignment across metazoans shows that the TRF2^{Long} is highly conserved (Fig. 27A). We identified an SQ/TQ site at Serine-23, which could potentially be phosphorylated by ATM or ATR, and computational predictions by NetPhos 3.1 strongly suggested that it would be phosphorylated by ATM (Fig. 27B) (*206*, *207*).

We therefore constructed an S23A mutant of TRF2^{Long} to block phosphorylation, and a phosphomimic S23D mutant. We expressed these in a MEFs and tested if these mutants would produce a phenotype different from TRF2^{Long} or TRF2^{Short} (Fig. 27C, D). The mutants were expressed at comparable levels to each other as well as to TRF2^{Short} and TRF2^{Long}. We first looked at the ability of these mutants to suppress TIFs 120 hours after TRF2 deletion. Both

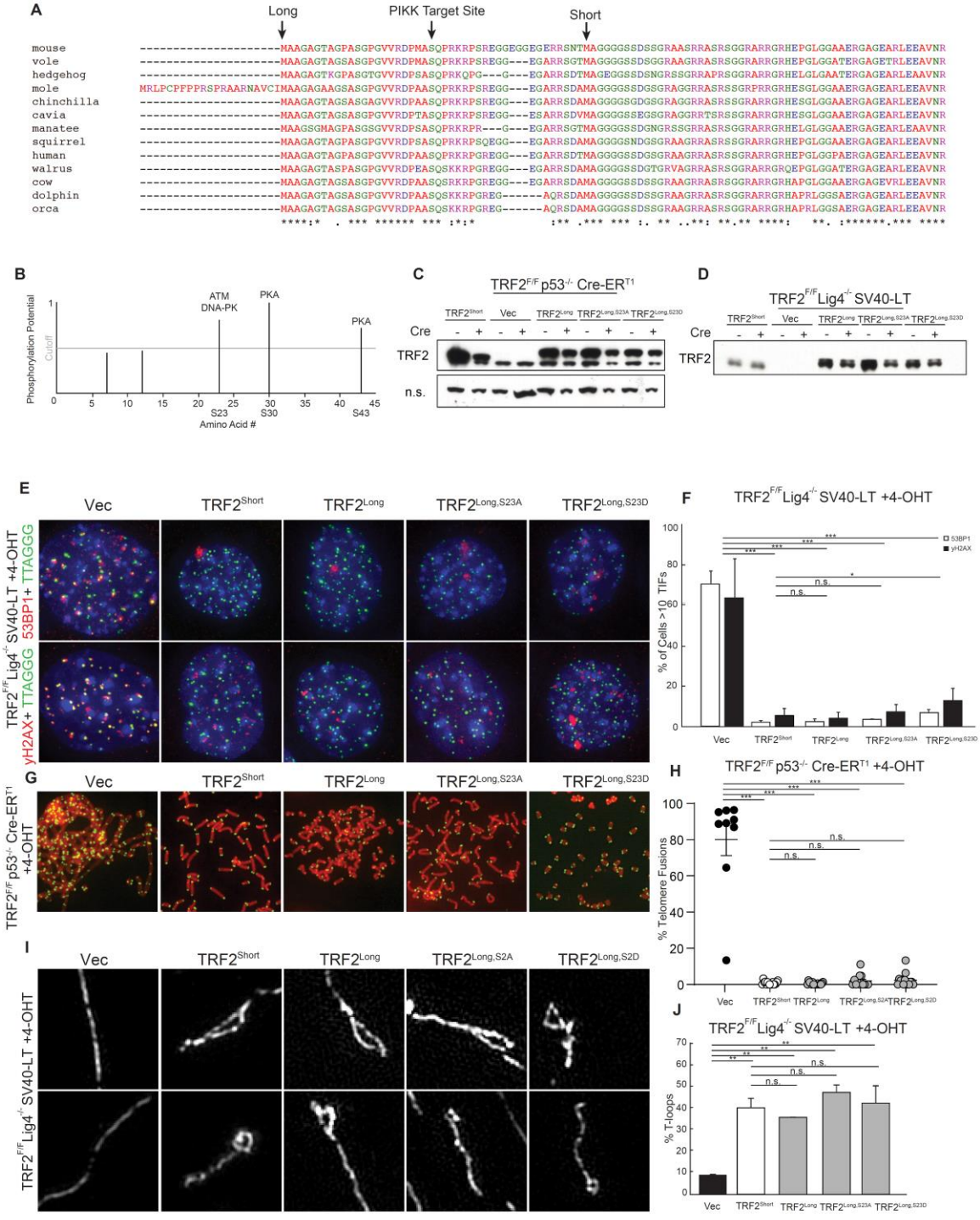
mutants were able to effectively rescue both 53BP1 and γH2AX TIFs (Fig. 27E, F). Although TRF2$^{Long, S23D}$ had a slightly elevated TIF response compare to TRF2^{Short}, it was not significantly different than any of the other mutants and still repressed TIFs efficiently.

Both mutants were also competent at preventing telomere fusions 120 hours after TRF2 deletion (Fig. 27G, H). None of the cells expressing either isoform of TRF2 had fusions that exceeded the vector control, and although once again, TRF2$^{Long, S23D}$ had a slightly elevated level of fusions compared to TRF2^{Long} and TRF2^{Short}, this difference was not statistically significant.

The phospho-mutants were also proficient at t-loop formation (Fig. 27I, J). Both were able to rescue t-loop formation after TRF2 deletion, and had comparable level of t-loops to TRF2^{Long} and TRF2^{Short}. The ability of these mutants to form t-loops is consistent with their ability to repress end-joining and ATM signaling.

Figure 27. Phosphomimetic Mutations in a Putative PIKK Site in the Extended Isoform of TRF2 do not Affect Function

(A) Sequence alignment demonstrating the highly conserved nature of the N-terminal extension of TRF2, as well as the location of the SQ site. (B) NetPhos 3.1 prediction of phosphorylated sites in the N-terminus of TRF2. (C) Western blot showing expression of $TRF2^{Short}$, $TRF2^{Long}$, $TRF2^{Long,S23A}$, $TRF2^{Long,S23D}$ in $TRF2^{F/F}$ MEFs immortalized with p53-loss and floxed for TRF2. TRF2 in the top panel, N.s. loading control in the bottom panel. (D) Western blot showing expression of $TRF2^{Short}$, $TRF2^{Long}$, $TRF2^{Long,S23A}$, $TRF2^{Long,S23D}$ in $TRF2^{F/F}$ $Lig4^{-/-}$ MEFs immortalized with SV40-LT and floxed for TRF2. TRF2 in the top panel, N.s. loading control in the bottom panel. (E) Representative images of TIFs in $TRF2^{F/F}$ $Lig4^{-/-}$ SV40-LT MEFs. Top panel: Blue – DAPI. Green – Telomeres. Red – 53BP1. Bottom panel: Blue – DAPI. Green – Telomeres. Red – γH2AX. (F) Quantification of TIFs in $TRF2^{F/F}$ $Lig4^{-/-}$MEFs, showing cells with 10 or more TIFs 120 hours after tamoxifen induced Cre-deletion. n = 3 independent experiments, 100 cells per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$) (G) Representative images of metaphases in $TRF2^{F/F}$ $p53^{-/-}$ MEFs. Red – DAPI. Green – Alexa647-$(TTAGGG)_3$. (H) Quantification of metaphases in TRF2F/F MEFs, showing the percentage of telomeres fusions 120 hours after tamoxifen induced Cre-deletion. n = 2 independent experiments, 2000 telomere ends scored per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$) (I) Representative images of DNA spreads showing t-loops in $TRF2^{F/F}$ $Lig4^{-/-}$MEFs, imaged using super-resolution OMX microscopy, 120 hours after tamoxifen induced Cre deletion. (J) Quantification of t-loops in $TRF2^{F/F}$ $Lig4^{-/-}$ SV40-LT MEFs n = 2 independent experiments, 100 countable molecules per replicate. Significance is shown using a two-tailed unpaired t-test where relevant (n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$)


A
Long
PIKK Target Site
Short
mouse
vole
hedgehog
mole
chinchilla
cavia
manatee
squirrel
human
walrus
cow
dolphin
orca
B
Phosphorylation Potential
Cutoff
ATM
DNA-PK
PKA
PKA
S23
S30
S43
Amino Acid #
C
TRF2F/F p53-/- Cre-ERT1
TRF2Short
Vec
TRF2Long
TRF2Long,S23A
TRF2Long,S23D
Cre
TRF2
n.s.
D
TRF2F/F Lig4-/- SV40-LT
E
Vec
TRF2Short
TRF2Long
TRF2Long,S23A
TRF2Long,S23D
TRF2F/F Lig4-/- SV40-LT +4-OHT
53BP1+ TTAGGG
γH2AX+ TTAGGG
F
TRF2F/F Lig4-/- SV40-LT +4-OHT
53BP1
γH2AX
% of Cells >10 TIFs
n.s.
G
TRF2F/F p53-/- Cre-ERT1 +4-OHT
H
TRF2F/F p53-/- Cre-ERT1 +4-OHT
% Telomere Fusions
I
Vec
TRF2Short
TRF2Long
TRF2Long,S2A
TRF2Long,S2D
TRF2F/F Lig4-/- SV40-LT +4-OHT
J
TRF2F/F Lig4-/- SV40-LT +4-OHT
% T-loops

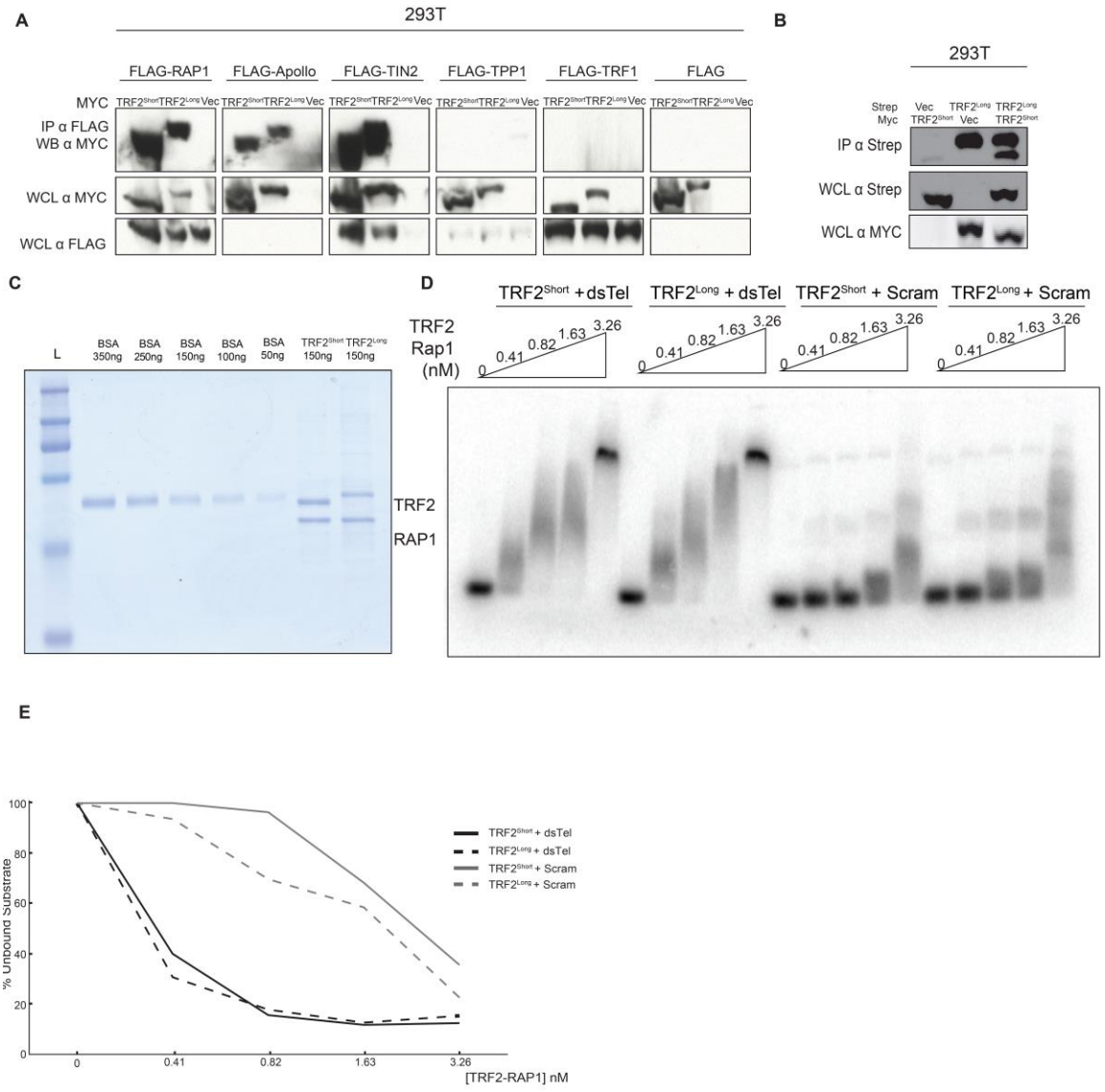
TRF2^{Long} Interacts with the Shelterin Complex and Binds DNA with High Affinity

To biochemically characterize TRF2^{Long} we performed co-immunoprecipitation assays with MYC-tagged TRF2^{Short} and TRF2^{Long}, and FLAG-tagged shelterin components. We found that all shelterin members that interact with TRF2^{Short} (RAP1, TIN2, TPP1), as well as Apollo, also interacted with TRF2^{Long} (Fig. 28A). There were no interactions with TRF2^{Long} that were not present with TRF2^{Short}. Unsurprisingly, we also found that TRF2 dimers can form between the two different isoforms (Fig. 28B).

To determine the DNA binding properties of TRF2^{Long}, we co-purified untagged TRF2 with its Strep-tagged RAP1 binding partner expressed in 293T cells by Streptavadin affinity purification (Fig. 28D). We then performed an electrophoresis mobility shift assay (EMSA) using either a short telomeric DNA substrate or a scrambled DNA substrate (Fig. 28C-E). TRF2^{Long} has comparable binding affinity to TRF2^{Short} when binding to telomeric DNA (0.33 nM and 0.341 nM, respectively).

Figure 28. The Extended Form of TRF2 Maintains Interactions with Shelterin and Binds DNA

(A) Co-immunoprecipitation of FLAG-tagged shelterin components (RAP1, Apollo, TIN2, TPP1, TRF1) and immunoblots against MYC-tagged $TRF2^{Short}$ or $TRF2^{Long}$ Whole cell lysate is also shown. Apollo cannot be detected because of low expression. (B) Co-immunoprecipitation of Strep-tagged $TRF2^{Long}$ and immunoblots against $TRF2^{Short}$ or $TRF2^{Long}$. Whole cell lysate is also shown. (C) One step purification of $TRF2^{Short}$ or $TRF2^{Long}$ from 293T cells by immunoprecipitation of its Strep-tagged interacting partner RAP1. (D) Electrophoresis mobility shift assay of $TRF2^{Short}$ or $TRF2^{Long}$ purified in (C). Increasing concentrations of protein were incubated for five minutes with .2 nM of either 142 base pairs of double-stranded telomeric DNA or 87 base pairs of scrambled DNA. (E) Quantification of the data in D. Unbound substrate as a percentage of total signal in the lane, normalized to 100% unbound in the first lane.


A
293T
FLAG-RAP1
FLAG-Apollo
FLAG-TIN2
FLAG-TPP1
FLAG-TRF1
FLAG
MYC
IP α FLAG
WB α MYC
WCL α MYC
WCL α FLAG
B
293T
Strep
Myc
Vec
TRF2Short
TRF2Long
Vec
TRF2Long
Vec
TRF2Short
IP α Strep
WCL α Strep
WCL α MYC
C
L
BSA 350ng
BSA 250ng
BSA 150ng
BSA 100ng
BSA 50ng
TRF2Short 150ng
TRF2Long 150ng
D
TRF2Short +dsTel
TRF2Long + dsTel
TRF2Short + Scram
TRF2Long + Scram
TRF2
Rap1
(nM)
0
0.41
0.82
1.63
3.26
TRF2
RAP1
E
% Unbound Substrate
100
80
60
40
20
0
0.41
0.82
1.63
3.26
[TRF2-RAP1] nM
TRF2Short + dsTel
TRF2Long + dsTel
TRF2Short + Scram
TRF2Long + Scram

Discussion

We examined the function of TRF2^{Long} to ask whether TRF2^{Long} can fully complement the function of TRF2. We found that TRF2^{Long} is able to complement TRF2 in protecting telomere ends from c-NHEJ and alt-NHEJ mediated fusion, as well as in suppressing ATM signaling. TRF2^{Long} was not significantly better at protecting telomeres than TRF2^{Short} in any assay we examined, they performed identically. TRF2^{Long} was also proficient in t-loop formation and telomeric overhang maintenance.

To a certain extent, the results found here are unsurprising. TRF2^{Short} was first cloned from a single cDNA in a library, and was fully able to complement the loss of a complete genetic knockout of TRF2 – suggesting that only one form is actually required in cells grown in vitro. Historically the N-terminus of TRF2 has been extraordinarily difficult to clone due to its GC rich nature and secondary structure formation, which may partly explain why TRF2^{Long} had been overlooked for so long. In the original cloning of TRF2 it may have been missing from the cDNA. Furthermore, TRF2^{Long} is slightly less abundant than TRF2^{Short} making it harder to detect and clone.

We also generated phospo-mutants of Serine-23 to attempt to address whether post-translational modifications may be an important regulatory aspect of TRF2^{Long}. These mutants targeted a conserved SQ/T site, which is a recognition and phosphorylation site of the ATM kinase, which TRF2 is known to repress. Neither mutant produced a striking phenotype and unfortunately, there was no clear indication that this site plays an important role in the regulation of TRF2. It is

possible that Serine-23 works is redundant with other ATM sites, such as T188 and T368, which could explain why we saw no significant effect from mutating this individual site.

We had also reasoned that the N-terminus of TRF2^{Long} might confer novel interactions within the shelterin complex or strengthen the DNA binding activity. However, this was not the case and there were no new interactions observed, and TRF2^{Long} did not form a tighter complex with any subunit of shelterin than does TRF2^{Short}. Similarly, there was no striking phenotype in DNA binding. TRF2^{Long}'s interactions within shelterin were consistent with the genetic data that show it can complement TRF2 function. It remains a possibility that TRF2^{Long} interacts with non-shelterin proteins in a way that TRF2^{Short} does not. For example, an alternative isoform of TIN2 preferentially interacts with the nuclear lamina (*208*). Similarly, TRF2 has been thought to interact with the nuclear lamina and nuclear pore complex as well (*209*). We did not test for such interactions, but differential binding partners can be identified by pulling down on a tagged version of each form of TRF2 and then performing mass spectrometry.

We cannot exclude the possibility that TRF2^{Long} has additional functions at telomeres that we did not observe. One such hint is that in human stem cells the relative expression of TRF2^{Long} and TRF2^{Short} is reversed: the TRF2^{Long} becomes somewhat more abundant than TRF2^{Short} (personal communication S. Boulton). Perhaps the long form is necessary during some step of development, and this could be tested by a specific knockout of the N-terminal extension in mice. TRF2

is known to be essential for embryonic development of mice, but there have been no studies to determine whether both the long and short form are required (*86*).

TRF2 has been described to play a role in neuronal function – a short isoform of TRF2 has been described to act as a differentiator of neurons (*210*). Intriguingly, the levels of $TRF2^{Long}$ and $TRF2^{Short}$ appears to be different in differentiated versus undifferentiated neurons (*210*). It is conceivable that the two forms of TRF2 may each play unique and regulated roles in neurons.

Ultimately, the question of the role of $TRF2^{Long}$ remains unresolved. It is possible that the N-terminal extension was preserved by chance – both transcripts are translated from the same RNA and there are three potential AUG start codons and Kozak consensus sequences in the N-terminus. However only the long and short isoforms are expressed, despite the central Kozak sequence being a comparable consensus match to $TRF2^{Short}$ (*211*). The N-terminal extension has remained highly conserved across 125 million years of evolution from dolphins to humans, and it would be very unusual for a sequence with no function to be retained and protected from genetic drift, especially since the sequence identity at the DNA level of the N-terminal extension is 85%. It is possible that at some point during evolution a codon was mutated to a start codon and it created a new start site – this could have led to the generation of either the long or the short variant. The creation of this variant TRF2 could be largely neutral – especially considering the small size change and the fact that both variants perform their essential functions. Compared to Myb domain proteins in more distantly related organisms, the N-terminus of TRF2 is poorly conserved, which makes it difficult to determine

whether TRF2^{Long} or TRF^{Short} is the actual ancestral sequence. Overall, the most likely explanation for our data is that we have not successfully identified the proper set of conditions under which we can observe the phenotype of TRF2^{Long}.

Chapter 6

Telomeres Across The Tree of Life

Telomeres are conserved nucleoprotein structures found at the end of linear chromosomes. Telomeres are required for the continued propagation of a linear DNA molecule in cells, and consequently are expected to be found in any organism that has linear chromosomes. In higher eukaryotes, where telomeres are most often studied, they typically consist of a conserved, tandemly arrayed DNA sequence, often $(TTAGGG)_n$. This sequence is usually bound by several proteins, which in mammals form a complex known as shelterin.

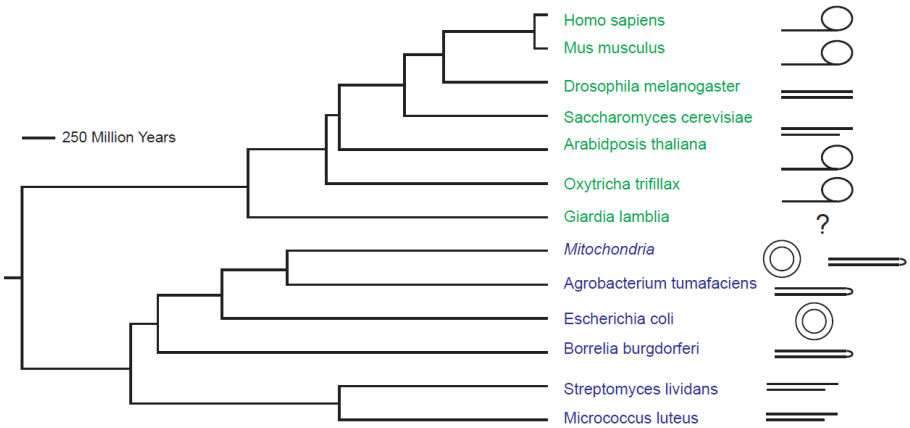
Telomeres serve several purposes. First, they solve what is known as the end-replication problem. In short, the end-replication problem arises due three reasons. First, RNA primase may not always place a primer at the exact end of a chromosome; second, removal of the RNA primer leaves a gap that cannot be filled-in; lastly, exonucleolytic processing of the telomere end. The resulting erosion of the telomere end threatens the integrity of the genome. Telomeres solve this problem by providing both a buffer of DNA sequence to be lost, but more importantly, by marking the end of the chromosome and recruiting proteins that allow the addition of new telomeric repeats to the end. In mammals, this telomere addition is mediated by telomerase, and it acts as a reverse transcriptase for de-novo telomere repeat addition at the ends of chromosomes (*1*–*3*). The telomerase protein is encoded by the TERT gene, while the RNA is encoded by TERC.

The second problem that telomeres solve is known as the end-protection problem. Cells have many methods of repairing breaks in DNA, but although the natural end of the chromosome resembles a break, it must not be repaired. Aberrant repair of a chromosome end can cause chromosome fusions, which

prevent proper cell division and may lead to cell death, or in some settings, aneuploidy and cancer. In eukaryotes, the shelterin complex acts to suppress both the DNA repair pathways as well as the DNA damage signaling pathways, which can induce cell cycle arrest and cell death (*217*).

Both the end-replication and end-protection problems will be discussed in greater detail in below. Higher eukaryotes have generally conserved both the telomeric sequence, telomeric proteins, and telomerase with few exceptions. However, most research on telomeres is done in a handful of model organisms, and their relevance to the rest of eukaryotes will be explored here. Although mouse telomeres are the model used for the data in this thesis, the relevance of the model to the rest of the evolutionary tree is of utmost importance when attempting to draw generalized conclusions. Despite a handful of interesting exceptions, such as fruit flies that lack a telomeric sequence, telomerase, and shelterin, and the newt *Pleurodeles waltl* that lacks telomerase and uses recombination to elongate its telomeres, eukaryotic telomeres are all structured in a very similar way.

Figure 29. Evolution of End-Protection. A) Rooted tree showing the evolutionary relationship between key model organisms from humans to bacteria. Eukaryotes are labeled in green, bacteria in blue. Schematics beside the species name illustrate whether the organism has a t-loop protection system, a blunt double-stranded system with end-binding proteins (Drosophila), an overhang with end-binding proteins (yeast), circular chromosomes, a hairpin, or an invertron system. Evolutionary tree is drawn to scale.


250 Million Years
Homo sapiens
Mus musculus
Drosophila melanogaster
Saccharomyces cerevisiae
Arabidposis thaliana
Oxytricha trifillax
Giardia lamblia
?
Mitochondria
Agrobacterium tumafaciens
Escherichia coli
Borrelia burgdorferi
Streptomyces lividans
Micrococcus luteus

Eukaryotes are not the only organisms that can have linear chromosomes (Fig. 29). Some bacteria have linear chromosomes, and current thinking would predict that they too must have some form of telomere at their chromosome end. Bacteria with linear chromosomes and telomeres deviate from the specific methods used by eukaryotes to solve the end-protection and end-replication problem, but nevertheless, they address these issues in their own unique way. Here we will compare telomeres across the entire tree of life and show that although there are exceptions to the generalized way telomere protection is controlled, many of the core principles have been conserved throughout the tree of life.

Chordate Telomeres

Homo sapiens

Human telomeres are a nucleoprotein structure comprised of the shelterin complex – a six subunit protein complex – and approximately 4 – 15 kilobases of the TTAGGG DNA repeat which ends in a 3’ overhang.

Human shelterin is composed of TRF1, TRF2, Rap1, TIN2, TPP1, and POT1. The individual components of the shelterin complex cooperate to repress ATM and ATR signaling, c-NHEJ, alt-NHEJ, PARP1, and HDR. The shelterin complex can be broken down into three parts: a set of myb domaining proteins, and a set of OB-fold containing proteins, as well as a linker protein.

The three myb domain containing proteins in shelterin are TRF1, TRF2 and Rap1. TRF1 and TRF2 both bind the double stranded telomeric DNA and although

Rap1 contains a myb domain, it does not bind DNA. TRF1 is a negative regulator of telomere length and is also important for semi-conservative replication of the telomere by recruiting helicases such as BLM (*98*, *246*). TRF2 functions to create and maintain the t-loop, which prevents ATM signaling and chromosome fusions through c-NHEJ (*56*). A t-loop is formed when the 3' telomeric overhang invades into the duplex region of the telomere and forms a D-loop. TRF2 also functions as a scaffold for other proteins, recruiting Apollo for 3' overhang processing and RTEL1 for telomere replication (*73*, *185*). In addition, TRF1 and TRF2 are both negative regulators of telomere length. RAP1, although a highly conserved member of the shelterin complex, has a relatively minor role, mostly in repressing homologous recombination at telomeres (*242*).

POT1 and TPP1 both contain OB-folds, though only POT1 binds the single stranded 3' overhang to prevent an ATR mediated DNA damage response. POT1 is also believed to play a role as a sensor of telomere length and regulates telomere elongation (*247*). Finally, POT1 also recruits the CST complex for Pol α/Primase mediated fill-in of the telomere end after telomere replication. TPP1 serves as a bridge to connect POT1 to the rest shelterin, and also recruits telomerase using the TEL patch in its OB fold (*248*, *249*).

TIN2 is the final member of the shelterin complex and acts as a bridge – connecting TRF1 to TRF2, as well as both to TPP1 (*77*, *90*, *250*). In addition, TIN2 is also responsible for repression of parsylation by inhibiting PARP1, and also contributes to inhibiting ATM independently of TRF2 (*60*, *90*).

Mammalian telomeres also typically contain a second complex called CST. CST consists of three members – CTC1, STN1, TEN1. CST is an RPA-like complex, itself a distant homolog of the ancient bacterial ssDNA binding proteins SSB (*251*). CST is believed to bind the 3' overhang and inhibit telomerase action, as well as recruit for Pol α/Primase to fill in the 3' overhang after it has been resected (*37*, *38*, *42*). CST is conserved in most higher eukaryotes including *Xenopus* and mouse (*11*, *252*).

Mus musculus

Rodent telomeres are highly conserved within the clade and with human telomeres. The telomeric sequence remains TTAGGG, and both rodents and humans share the key components of the shelterin complex. One difference lies in the duplication of the POT1 protein: rodents have two (POT1a and POT1b) while humans have only a single POT1 protein (*232*). Mouse POT1a is chiefly responsible for suppressing the ATR response, while POT1b interacts with and recruits CST for fill-in of the telomeric overhang (*11*, *232*).

A second critical difference relating to telomeres is that rodents constitutively express telomerase in all their tissues, in stark contrast to mammals where telomerase is expressed primarily in the germ line (*253–255*). In addition, it is known that while human cells will rarely spontaneously immortalize in culture, mouse cells do (*256*).

Diptera Telomeres

Diptera is an order of the class of Insecta which includes flies and mosquitos. The order of Diptera is also known as “true flies” because they do not include wasps, butterflies, or flying beetles. True flies are characterized by a single pair of primary wings, reduced hind wings, compound eyes, and an ability to undergo complete metamorphosis. Diptera such as flower flies, are critical to many ecosystems due to their role as pollinators and rival bees as the main plant pollinators. Some members of Diptera are either agricultural pests or disease vectors.

The last common ancestor between Diptera and Homo sapiens was approximately 794 million years ago, with a confidence interval of 678 – 916 million years (*257–264*). The main models used to study Diptera are the fruit fly *Drosophila melanogaster* and various species of mosquito. Diptera offer one of the most interesting exceptions to the standard model of telomere protection in eukaryotes: they have an entirely divergent system of telomeres that is based on transposons. Although these transposons share no similarities in sequence and bind different proteins, they accomplish the same fundamental functions as the TTAGGG and sheltered based telomeres of humans.

Table 1. Sequences of Diverse Telomeres. A) List of key organisms with different telomeric systems, indicating which telomeric proteins are at their telomere, the general structure of the end, and the sequence of the very 3' end of the chromosome.

Organism	Telomeric Sequence	Proteins	End Structure
H. sapien	TTAGGG	Yes Shelterin	T-loop
M. musculus	TTAGGG	Yes Shelterin	T-loop
D. melanogaster	Het-A & TAR-T Transposon	Yes Terminin	Blunt or 3' Overhang
S. cerevisae	TG_{1-3}	Yes	3' Overhang
A. thaliana	TTTAGGG	Yes Shelterin	T-loop
O. trifillax	TTTTGGGG	Yes, Shelterin homologs	T-loop
G. lamblia	TAGGG	No	Unknown
B. burgdorferi	TACTAATAAAAAATTATATATATAATTTTTTATTAGTA	Yes ResT	Hairpin
E. coli	None	No	Circular Chromosome
S. lividans	CAGCGATAGGGTACCCGCTCCGCGGG	Yes TapL TpgL	3' Overhang

Drosophila melanogaster

Drosophilia melanogaster is one of the most studied model organisms in biology, and yet it has one of the most unrepresentative telomeres amongst all eukaryotes. Drosophilia and the rest of the Diptera genome does not incorporate the canonical TTAGGG sequence at chromosome ends (*265*, *266*). Furthermore, Diptera lack telomerase (*267*). Drosophilia telomeres vary widely in size, ranging from 147 kb on the XL chromosome to 26 kb on the 3L chromosome, however this also varies depending on the strain of fly (*268*). In addition, there is no evidence that Drosophila telomeres end in t-loops, and there is no evidence of a 3' overhang at Diptera telomeres either.

Drosophila melanogaster telomeres have a unique structure in the animal kingdom. They consist of three repeating elements: a Het-A retrotransposon, a TAS interstitial sequence, and a TART retrotransposon (*269*, *270*). Het-A is a 6 kb sequence that contains a single ORF that codes for a Gag protein. Het-A usually also contains up to five 80 bp tandem repeats at the 3' end. TART is a 12 kb element with two ORFs, one encoding a reverse transcriptase, and one encoding a Gag protein. TAS, or telomere associated sequence, are the interstitial repeats between the Het-A and TART sequences. They vary in length considerably and contain modest homology to one another.

Due to the unique nature of Drosophila telomeres and their lack of telomerase, their method of elongating and solving the end-replication problem is also significantly different from that of other eukaryotes. Drosophila telomeres chromosomes shorten by 50 to 100 base pairs per generation, or 75 base pairs on

average (*271–274*). Telomeric sequence is added rarely, randomly, and in large bursts. Whenever a TART retrotransposon successfully targets and integrates into a Het-A element, 6 kb of DNA is added. This occurs at roughly 1% per chromosome per generation – but is highly variable (*275*). At this rate of addition, it would add approximately 60 base pairs per generation, which is enough to compensate for the experimentally observed rate of shortening.

As in other organisms, Diptera telomere length is also regulated by proteins which act at the telomere. Interestingly, many familiar players from the DDR have found a new a role in telomere length regulation. For example, the Ku70/80 complex, which is required for initiation of c-NHEJ, is a negative regulator of telomere length, and although the mechanism is unclear, it is believed to interfere with the addition of transposon repeats and thereby protect the telomere end (*276*). HP1, a chromatin remodeling factor, also plays an important role in Diptera telomere length maintenance. Loss of HP1 increases the frequency of addition events, likely because it leads to a loss of heterochromatin, and makes the chromatin structure more accessible to integration events by Het-A and TART (*277*). In addition, there are proteins that have been found to genetically affect telomere length in Diptera, but the mechanism of action is not known. One such protein is Tel, and mutations in tel increase telomere length by up to 7-fold (*278*).

Despite not having the shelterin complex at telomeres, Diptera have evolved their own set of telomeric proteins, known as terminin. Terminin consists of HHop, HOAP, HP1, Moi, and Ver (*267*). Ver is the only terminin complex member with homology to chordate telomeric proteins – it is a homolog of STN1,

which is a member of the CST complex (*279*). Unlike in the shelterin complex where only loss of TRF2 causes telomere fusions, loss of any of the Diptera terminin components results in telomere fusions (*267*, *280–284*).

The binding of terminin to Drosophila telomeres is not based on DNA sequence recognition, but rather on the heterochromatin state of telomeres. The telomeres are marked by HP1, which recruits HOAP. HOAP in turn recruits Moi and Ver (*285*). Loss of either ATM or Mre11 causes loss of HP1, and consistent with this, results in telomere fusions. Intriguingly, this is the opposite of what is observed with mammalian telomeres, where the signaling function of ATM and Mre11 is necessary for telomere fusions. How ATM or Mre11 are recruited to telomere ends is currently not known, but they regulate both localization of HP1 and the creation of heterochromatin at telomere ends.

Despite the advances in understanding Drosophila telomeres, Muller's observation that Drosophila telomeres are not readily made de-novo at chromosome breaks remains difficult to explain (*286*). Given that the telomeres are not demarcated in a sequence specific manner, they should be formed at double strand breaks. However, they are not, and this demonstrates that there is still more to be learned about Drosophila telomeres.

Coleoptera Telomeres

Coleoptera (beetles) also have a somewhat unique telomeric arrangement. They are related to Diptera, having diverged 310 Mya. Unlike Diptera, many Coleoptera have a telomeric repeat sequence, however 50% of the tested species

have lost the TTAGG repeat (*287–289*). Interestingly, there was no strong correlation between whether two species were related and whether they retained a telomeric sequence, suggesting that there were many random and independent events that lead to telomere sequence loss.

In addition, some beetles have an alternative telomere sequence, e.g. TCAGG, as is the case in at least three families of beetle (*290*). Further, in some species of Coleoptera, no telomeric sequence was present at the ends. The TERT sequence of some Coleoptera also harbors mutations in highly conserved regions, raising questions as to whether these beetle TERTs are functional (*291*). Despite some exceptions, most species of beetle, as well as insects, have been found to have an active telomerase (*292*).

Yeast Telomeres

Yeasts diverged from mammals approximately 1,100 Mya yet for the most part, their telomeres resemble those of humans and mice, lending them credence as a model system. One exception to this are the telomeres of *Saccharomyces cerevisiae*.

Saccharomyces cerevisiae

Although budding yeast has telomeric binding proteins, they bear no sequence nor structural similarity to chordate shelterin. This may be due to the whole genome duplication event that *Saccharomyces cerevisiae* suffered. However, budding yeast do contain a fully functional telomerase (*4*, *293*, *294*).

Budding yeast have very short telomeres, about 300 base pairs in length, and the telomeric sequence is heterogenous, often abbreviated as TG_{1-3} (*295*–*297*). In addition, there is no evidence for t-loops at budding yeast telomeres, consistent with the absence of TRF2. However, it may also be that their telomeres are not long enough to observe t-loops through conventional imaging means. Another closely related yeast species, *K. lactis*, was found to have observable t-loops in a mutant strain selected for longer telomeres (*298*).

The budding yeast telomeric proteins include Rap1, Rif1 and Rif2, and the CST complex which consists of Ten1, Stn1 and Cdc13. Rap1 was originally identified as a protein binding the mating type locus, but later it was found that temperature sensitive mutation in Rap1 affect telomere length and telomere stability (*299*–*303*). Interestingly, Rap1 is the only telomeric protein in budding yeast conserved across both yeasts and mammals, having two myb domains, a BRCT domain, and an RCT domain (*71*, *304*, *305*). Rap1 is believed to regulate telomere length through a counting mechanism whereby the number of Rap1 molecules bound to the DNA is inversely correlated with telomerase action (*297*). Rap1's other major function is to repress c-NHEJ at the telomere end, reminiscent of the action of TRF2 at mammalian telomeres (*306*). Additionally, Rap1 anchors Rif1 and Rif2 to the telomere.

Rif1 and Rif2 are critical for allowing higher order structurers to form between Rap1 molecules (*307*). Rif2 has been shown to negatively regulate the activity of the MRX complex and nucleolytic processing of the telomere end, as

well as block NHEJ (*306*, *308*). Like Rif2, Rif1 is important for repressing checkpoint signaling at telomere ends (*309*, *310*).

In budding yeast, CST performs many critical functions that otherwise would have been split between CST and shelterin because Saccharomyces cerevisiae lacks shelterin. Within CST, cdc13 is primarily responsible for DNA binding and also recruitment of telomerase to telomeres (*311–314*). In addition, CST is important for repressing Mec1 (ATR) but not Tel1 (ATM). Loss of CST in budding yeast results in telomere uncapping and DNA damage signaling (*315*).

Schizosaccharomyces pombe

The telomere structure of *S. pombe* is reminiscent of higher organisms: they have a somewhat homologous telomere sequence, and a set of telomeric binding proteins that functionally resemble the shelterin complex. The *S. pombe* telomeric consensus sequence is $TTACAGG_{(1-4)}$, although it can include rare insertions of additional Ts immediately following or preceding the G tract, and an insertion of an A following the C (*316*). Fission yeast has a functional telomerase (*317*, *318*). Fission yeast telomeres also have the capacity to form t-loops (*319*). Given that *S. pombe* telomeres are 1 kb long, and therefore longer than budding yeast telomeres, there is enough telomeric DNA to be able to form and resolve a t-loop (*320*).

The central telomeric protein of *Schizosaccharomyces pombe* is Taz1, which shares functional and structural homology to the mammalian TRF1 and TRF2 shelterin components (*321*). Taz1 contains a myb domain and is responsible

for telomere length regulation, as well as interacting with and anchoring the rest of the *S. pombe* telomeric proteins, Rif1, Rap1, Poz1, Tpz1, Pot1 and Ccq1 to the telomere. In addition Taz1 is responsible for preventing telomere fusions in fission yeast, similar to the role that TRF2 has in mammalian cells (*322*). Rif1 and Rap1 bind directly to Taz1 and may compete for the same site (*323*). Both Rap1 and Rif1 are essential for proper telomere length control in *S. pombe*.

S. pombe Pot1 was identified based on homology of the OB fold to the Oxytricha TEBPα telomeric ssDNA binding protein, and it was subsequently found to also be homologous to the human Pot1 (*63*). Loss of Pot1 in fission yeast leads to rapid telomere shortening and loss of the telomere sequence, as well as chromosome circularization. Fission yeast Pot1 binds to Tpz1 in a manner homologous to the Pot1-TPP1 interaction in mammalian cells, and like TPP1, Tpz1 has an OB fold (*324*).

Like the loss of Pot1, deletion of Tpz1 resulted in loss of viability and loss of the telomeric signal. *S. pombe* Poz1 functions similar to mammalian TIN2, bridging the Pot1-Tpz1 complex to the Taz1-Rap1 complex via its interaction with Rap1 (*324–326*). Poz1 is a negative regulator of telomerase, but loss of Poz1 does not lead to cell death. *S. pombe* shelterin contains another protein, Ccq1, which was first identified as containing an SMC domain (*327*). Loss of Ccq1 leads to telomere shortening, abnormalities during cell division, and persistent Mec1 activation (*328*). Ccq1 is also known to recruit telomerase to telomeres (*329*).

Remarkably, although no known eukaryotic organism has circular chromosomes, it is possible to circularize the telomeres of fission yeast. When

Tel1 and Rad3 were deleted, fission yeast loss their telomeric sequence and fused their chromosomes to create a single large circular chromosome (*330*). These cells are able to bypass both the end protection and end replication problem and are viable through mitosis, though they were unable to undergo productive meiosis.

Plant Telomeres

Plants diverged from humans approximately 1,600 million years ago, and despite this, the organization of their telomeres has remained surprisingly similar to that of other eukaryotes. Plants generally have a telomere sequence that is very closely related to the mammalian sequence, TTTAGGG, and plant telomeres were first isolated from *Arabidopsis thaliana* (*331*). Plants also have very variable telomere lengths, ranging from a few kb in *Arabidopsis* to over 150 kb in other species (*332*). T-loops have also been demonstrated in plants, highlighting the level of similarity of telomeric organization across over a billion years of evolution (*333*).

Despite these broad similarities, there are still specific exceptions to telomere organization in plants. Not all plants have the TTTAGGG telomeric sequence: interestingly, plants of the *Asparagales* genus have transitioned to the TTAGGG variant as their predominant repeat, with TTTAGGG and TTAGG repeats also being found (*334*). In addition, plants of the genus *Allium* hybridize poorly to all known telomeric repeats, and their telomeric sequence remains largely uncharacterized (*335–337*). Recently *Allium* telomeres have been tentatively

uncovered as having the unique repeat sequence of CTCGGTTATGGG, which is synthesized by an active telomerase (*338*).

One important yet under explored area is the plant shelterin proteins. Efforts to find shelterin homologs have been hindered by the large number of myb and OB fold domain containing proteins in plant genomes, and techniques such as pulldowns are complicated by the fact that many plant promoters contain the telomeric sequence (*332*). Despite this, six proteins have been discovered that are part of the TRF family that are also capable of forming homodimers or heterodimers with each other and binding telomeric DNA – TBP1, TRP1, TRFL1, TRFL2, TRFL4, TRFL9 (*339*). Homologs of these proteins have been found in many plants, such as rice, where one such protein was crystalized and shown to align well to TRF1 (*340*). Unsurprisingly, due to the six-fold redundancy, disruption of any single one of these TRF-like proteins did not produce a telomeric phenotype, with the exception of TBP1, whose loss lead to a mild telomere lengthening phenotype (*339*, *341*). However, in tobacco and rice plants where there are only two copies of the TRF proteins, strong telomeric phenotypes were observed, such as telomere length dysfunction, cell death, and anaphase bridges, which could be signs of telomere fusion (*342*, *343*).

Plant shelterin also contains up to three of the Pot1 proteins. Although the proteins share homology to the OB folds, neither Pot1 (Pot1A) nor Pot2 (Pot1B) appear to bind telomeric DNA in vitro (*344*). Despite this, Pot1 in *Arabidopsis* is known to bind telomerase and be required for positive telomere length regulation (*345*). Disruption of Pot2 leads to telomere shortening and anaphase bridges,

which appear to be the result of chromosome fusions at the telomere (*346*). Little is known about the third Pot1 protein. Finally, plants do contain a CST complex which closely resembles the CST of both yeast and mammalian origins, however, as in other higher eukaryotes, the CTC1 subunit does not share homology with the yeast CDC13 subunit (*38*, *347–349*).

Bikonts

Bikonts are a distant clade of eukaryotes, distinguished by the presence of multiple flagella rather than a single flagellum as in unikonts, which comprise all higher order eukaryotes. They are one of the most distant outgroups of Eukaryotes, having diverged approximately 1.7 billion years ago from chordates. Despite the evolutionary distance, bikonts have proven to be an important system for understanding telomere biology – both *Oxytricha* and *Tetrahymena* are members of the bikont group.

Oxytricha has a telomeric sequence of TTTTGGGG while *Tetrahymena* has TTGGGG, both are fairly similar to the TTAGGG sequence of higher eukaryotes (*350*, *351*). Consistent with the conserved nature of telomeres, t-loops have been observed at *Oxytricha* telomeres in the micronuclei (*144*, *352*). Telomerase is obviously present in both species (*1*, *353*). The key difference between bikont and unikont telomeres appears to be the presence of shelterin – the telomeric binding proteins that bikonts have contain OB folds that have sequence homology to Pot1 (*63*, *354*). *Oxytricha* has two telomeric binding proteins, TEBPα and TEBPβ (*355*). The two TEBP proteins are responsible for binding and protecting the G-rich

overhang of *Oxytricha*. Interestingly, only TEBPα has DNA binding activity and shows telomere protection, while TEBPβ, which is proposed to be more closely related to TPP1, is required for proper complex formation of TEBPα-TEBPβ, and for forming G quartets (*356–359*). Together, the complex acts as a negative regulator of telomerase by outcompeting its binding (*360*). Remarkably, recent genome sequencing has shown that there are multiple distinct paralogs of both TEBPα and TEBPβ, six of TEBPα and three of TEBPβ which can all potentially form complexes with one another (*361*).

Tetrahymena differs from *Oxytricha* and has a more extensive set of telomere end capping proteins, consisting of four proteins – Pot1a, Tpt1, Pat1, and Pat2 – but nonetheless does not have a full shelterin complex. Pot1a was originally identified based on sequence homology to TEBPα and is an essential protein in *Tetrahymena*. It binds the G-overhang and is responsible for negative regulation of telomerase, as well as blocking DNA damage signaling, though it has no effect on overhang structure (*362*). Tpt1 appears to be the equivalent of TEBPβ or TPP1 and interacts directly with Pot1a and deletion of Tpt1 produced similar phenotypes to that of Pot1a loss (*363*). In contrast, loss of Pat1 caused telomere shortening and it is thought that Pat1 is needed to recruit telomerase to the telomere end (*363*). Pat1 interacts with Tpt1 but not Pot1, lending further credence to the idea that Tpt1 is a TPP1-like bridge. The final telomeric protein in *Tetrahymena* is Pat2, which interacts with both Tpt1 and Pat1, and is responsible for telomere length regulation because its depletion results in shorter telomeres (*364*).

Tetrahymena is unique in that it has a CST- and RPA- like telomeric complex called the 7-4-1 complex, which consist of the subunits p75, p45 and p19 (*365*, *366*). Originally discovered in complex with Tetrahymena telomerase, it was identified as an being related to RPA, and later it was found that p45 and p19 are structural homologs of Stn1 and Ten1 (*367*). The 7-4-1 complex binds the ssDNA overhang of telomeres and its disruption causes an increase in the 3' overhang length, similar to loss of CST.

Excavata Telomeres

Excavata are the most distant eukaryote to chordates, being one of the first phyla to diverge from the rest of the eukaryotic lineage, somewhere between 1.7 and 2.3 billion years ago, and they are often regarded as the outgroup of a rooted eukaryotic tree (*368*). Excavata are single cellular and are often parasites, the best known of these are the disease-causing order *Trypanosomatida* and the genus *Giardia*. Given their basal position on the tree of life, it is somewhat remarkable how well conserved their telomeres are.

Trypanosomes have been used as a model system since the 1980s to study telomeres and have the ancestral telomeric sequence of TTAGGG and an active telomerase which adds repeats to telomere ends (*369*, *370*). Trypanosomes have also been found to harbor t-loops, though they are small compared to mammalian t-loops, especially given that the telomeres in trypanosomes and humans are approximately the same length at 10 – 20 kb (*371*). Rather, trypanosomes have a single protein, called tbTRF, which has strong sequence homology to the myb

domain of TRF1 and TRF2 (*372*). tbTRF is an essential gene that binds the telomeric sequence, can homodimerize, and loss of tbTRF causes a loss of the overhang signal, though unlike TRF2, no telomere fusions are observed. Recently, a homolog of RAP1, tbRAP1 has been identified. This protein contains up to two myb domains and a BRCT domain, is an interacting factor of tbTRF, and is essential for regulation of VSV-G expression (*373*).

The most distal eukaryote in the Excavata lineage is *Giardia lambis*, which is estimated to have diverged from *Homo sapiens* approximately 2.1 billion years ago. Despite the evolutionary distance, *Giardia* telomeres are also well conserved. The telomeric repeat is TAGGG and its telomeres are only about 500 base pairs in length (*374*, *375*). To date, no telomeric binding proteins have been found. Relatively weak telomerase activity has been found in *Giardia*, though interestingly the protein component lacks the T motif necessary for template binding and processivity (*376*, *377*). This fact, along with the discovery of active retrotransposons has led to the hypothesis that *Giardia* telomeres are maintained by retro transposition, similar to the way *Diptera* telomeres are (*378*). However, these two mechanisms do not need to be mutually exclusive and both can be used to actively maintain *Giardia*'s telomeres.

Archaea

Archaea comprise an entirely separate domain of life from eukaryotes and prokaryotes. The last common ancestor of archaea and eukaryotes existed between 2.1 and 4 billion years ago (*379*). Unlike eukaryotes, which have an end-

protection and end-replication problem due to their linear chromosomes, no Archaea studied to date have linear chromosomes, negating the need for telomeres.

Bacteria

Despite archaea and eukaryotes having exclusively circular or linear chromosomes, respectively, the genomic organization of prokaryotes is much more varied. The divergence between eukaryotes and bacteria occurred approximately 4.1 billion years ago (*380*–*382*), and the evolutionary distance between some of the prokaryotic species is immense. Although most prokaryotes have circular chromosomes and circular plasmids, some have linear plasmids and even linear chromosomes. The presence of linear chromosomes would suggest that these organisms must also find a mechanism to cope with the end protection and end replication problems.

Bacteria with linear chromosomes still face many of the problems that eukaryotes face with their linear chromosomes. First, the problem of linear chromosome replication does not change when using a bacterial DNA polymerase. Second, bacteria have an active ligation-based repair pathway in addition to the more common and better studied recombination-based repair pathway. *Bacillus subtilis* has an active Ku / LigD pathway that is important for resistance to ionizing radiation and is separate from the canonical RecA repair pathway (*383*). Similarly, *Mycobacterium tuberculosis* has a very active Ku/LigD pathway (*384*, *385*). *Escherichia coli* also has a LigA dependent repair pathway which becomes much

more prominent after its suppressor recBCD is impaired (*386*). The presence of an end-joining repair method would pose a threat to linear chromosomes due to the risk of inappropriate ligation of the ends together. Although end-joining repair pathways have not been studied in all the bacteria that have linear chromosomes, *Agrobacterium* has an active LigC and LigD NHEJ pathway and linear chromosomes which creates a true end-protection problem as in mammalian cells (*387*). Finally, bacteria can also respond to a DNA end and a DSB through the SOS signaling pathway, which detects ssDNA (*388*, *389*). Prolonged activation of the SOS response can lead to deleterious mutations as well as cell cycle arrest.

Mitochondria and Chloroplasts

Although not strictly prokaryotes, mitochondria and chloroplasts are of prokaryotic origin, as proposed in the endosymbiotic theory (*390–394*). What is most surprising however, is that many mitochondria and chloroplast have linear DNA molecules, rather than circular DNA as previously thought (*395*, *396*). This is true in a number of species, ranging from *Tetrahymena* to *Saccharomyces* to many plant species such as maize (*397*, *398*).

In species in which the mitochondrial or chloroplast telomere has been studied, the telomeric repeat often differs from the one found in the parental nucleus. For example, the mitochondrial telomere sequence of *Tetrahymena thermophila* is fourteen copies of a 53 nucleotide repeat of TCTTAGAGGTATGTTAGCTATTAGTGTTGTTTAGGCTTGTTATGGTATGTGTA (*397*, *399*). The mitochondrial telomeres of many yeast species are believed to end

in a closed hairpin (*400*). How these telomeres are replicated is not yet known, but it is hypothesized to occur via recombination or self-priming, similar to vaccinia virus (*396*, *401*, *402*). Reminiscent of Drosophila, *Fusarium oxysporum* mitochondrial telomeres have a retrotransposon system for maintaining telomeric integrity (*403*). These alternative telomere arrangements are not unique to mitochondria as chloroplast tend to share them as well, though it has been studied less. Surprisingly, t-loops have also been demonstrated in mitochondrial telomeres, suggesting a very ancient origin to the t-loop model of end protection (*404*). Unfortunately, the mechanism by which they are created remains elusive.

The presence of linear telomeres and t-loops would suggest the possibility of there being telomere binding proteins in these organelles as well. In fact, this is now known to be true, and at least two classes of telomere binding proteins have been found in mitochondria. The first, is a protein that binds the ends of the linear molecule and assists in replication by serving as a protein primer (*396*, *405*). This protein works similarly to the way adenovirus terminal protein acts, presumably by conjugating itself to a nucleotide to prime replication (*406*, *407*). The second protein encountered at some mitochondrial telomeres is known as mtTBP. This protein specifically binds the telomeric sequence (in the case of *C. parapsilosis*, a 110 nucleotide repeat), and GFP tagged mtTBP localizes to the mitochondria (*408*, *409*). The function of mtTBP remains unclear, though it is believed to be important for binding the 5' telomeric overhang and protecting the overhang from degradation, and it may play an unknown role in generating a protective t-loop.

Borrelia burgdorferi and Agrobacterium tumafaciens

Borrelia burgdorferi and *Agrobacterium tumafaciens* are two well studied examples of bacteria with linear telomeres. *Borrelia* is a spirochete that causes Lyme disease while *Agrobacterium* is an alphaproteobacterial that causes plant tumors, they both share a common method of solving the end protection and end-replication problems. Rather than leaving the ends of their chromosome free, both of these bacterial species have a linear covalently closed hairpin at the end of their chromosomes.

If *E. coli* is used as a standard dating marker, then *Borrelia* diverged 3 billion years ago from *E. coli*, and *Agrobacterium* about 2.4 billion. Two other bacteria also use this solution for their telomeres: *Cyanotheces 51142* (3.2 billion years from *E. coli*) and *Candidatus Phytoplasma* (also 3.2 billion years from *E. coli*) (*410*, *411*).

A linear covalently closed circle is a linear ssDNA where the ends have been connected with one another; it is actually a DNA circle but it is different from a double stranded circle (such as a plasmid) because there is only one molecule, and different from a single stranded circle because the interior DNA is complementary. Early evidence for the linear genomes of these two bacterial species comes from experiments based on DNA mobility in gels, as well as from electron microscopy, where *Borrelia* and *Agrobacterium* chromosomes exhibit snapback kinetics when partially denatured, do not form DNA bundles under electron microscopy when denatured unless treated with a nuclease, and appear as a linear circle when fully denatured under EM (*412–414*). The final piece of

evidence to suggest that these species have linear chromosomes came from whole genome sequencing and mapping overlapping reads (*415*). The end of the *Borrelia* chromosome is a 19 or 26 nucleotide hairpin with the sequence TACTAATAAAAAATTATATATATAATTTTTTATTAGTA (*416*).

A covalently closed end solves the end-protection problem by hiding the end from any factors that would recognize it as a broken molecule. Further, it also addresses the end replication problem by presenting a linear molecule as a circle that can be replicated without loss of DNA. However, this solution introduces a new problem during replication – the circle has to be resolved into two linear chromosomes. This is achieved by a prototelomerase protein, ResT in *Borrelia* and TelA in *Agrobacterium* (*417–419*). Although it is called a prototelomerase, it shares no homology with mammalian telomerase and does not have any reverse transcriptase function; instead it is both a ligase and a resolvase (*420*, *421*). This tyrosine recombinase recognizes the site-specific telomere sequence, binds as a dimer, and cuts at the cleavage site where the two telomeric palindromes meet (*422–425*). It then facilitates refolding of the intermediate and ligates the ends back together (*426*). This hairpin telomere with a tyrosine resolvase allows bacterial linear chromosomes to be protected from aberrant end-joining, replicated without loss of DNA, and properly resolved and segregated during replication and cell division.

Table 2. Comparison of Shared Protein Features. A) A comparison of proteins which share similar functions across the tree of life.

Feature	Human	Rodent	Diptera	Budding Yeast	Fission Yeast	Plants	Tetrahymena	Oxotrychia	Mitochondria	Bacteria
dsDNA Binding Protein	Myb domain: TRF1 / TRF2	Myb domain: TRF1 / TRF2	HOAP	Myb domain: Rap1	Myb domain: Taz1	Myb domain: TBP1 /TRP1 / TRFL1 / TRFL2 / TRFL4 / TRFL9				TpgL
ssDNA Binding Protein	OB fold: Pot1	OB fold: Pot1a/b	Ver		OB fold: Pot1	OB fold: Pot1 / Pot2 / Pot3	TEBPα	Pot1a	mtTBP	TapL
Bridging Proteins	TIN2 / TPP1	TIN2 / TPP1		Rif1 / Rif2	Rap1 / Poz1 / Ccq1 / Tpz1		TEBPβ	Tpt1		
Accessory Proteins	Rap1		Hhop, Moi		Rif1			Pat1 / Pat2		
RPA-like Complex	Ctc1 / Stn1 / Ten1	Ctc1 / Stn1 / Ten1	Ver	Ctc1 / Stn1 / cdc13	Ctc1 / Stn1	Ctc1 / Stn1 / Ten1		p75 / p45 / p19		SSB
Telomerase	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No	No

Streptomyces lividans

Streptomyces lividans and *Micrococcus luteus* have evolved a system to maintain their linear telomeres different from the hairpin system described above. Both of these bacterial species are approximately 3.2 billion years diverged from *E. coli* and about 1.6 billion years from each other, as they are both members of the Gram-positive bacteria. The telomeres of *Streptomyces* are 20 – 25 kb of inverted repeats that contain 11 13 – 26 bp palindromes, and these telomeres are bound by two terminal telomere maintenance proteins.

Like *Borrelia* and *Agrobacterium*, *Streptomyces* has a linear chromosome where both ends are identical in sequence and bound by a protein complex (*427*). This protein complex consists of two proteins, TpgL and TapL which bind specifically to the telomeric repeats and complete the “invertron” telomere cap system (*428*). These two proteins are absolutely required for proper maintenance of the *Streptomyces* chromosome, for both its replication and to prevent fusions. In fact, deletion of either factor lead to chromosome circularization, circularization of all *Streptomyces* plasmids, and the inability to maintain linear DNA (*428–430*). Interestingly, although the *Streptomyces* genome is comparatively unstable, both artificial circularization and removal of the telomeric repeats increases genomic instability and decreases growth and viability, pointing to the importance of the telomere in this species (*431–433*). In addition, the inverted repeats of the telomeric sequence create a complex secondary structure which serve to recruit and position the two telomeric proteins, as well as aid in replication (*434*).

TpgL and TapL function together for both telomere end protection and for proper priming of replication. They interact with one another in a two-hybrid experiment, and TapL preferentially binds ssDNA while TpgL prefers dsDNA (*429*). In addition to their role in preventing chromosome circularization, they are also needed to prime replication in a protein dependent manner (*435*). TpgL is covalently linked to the 3' end of the DNA and helps position both the DNA and TapL. TapL is a polymerase that adds nucleotides in a template dependent manner to TpgL. Once TapL has added a 13 nucleotide primer, Pol I extends this and takes over to replicate the rest of the chromosome (*436*). In addition, TapL is also capable of restoring truncated telomere sequences in vivo through a foldback mechanism (*437*).

The invertron system developed by *Streptomyces* is the closest bacteria have come to replicating the telomeres of eukaryotes. This is a system that has a specific telomere length and sequence, two binding proteins which are essential for chromosome end-protection, and together they are responsible for recruiting a polymerase and solving the end-replication problem.

Viruses

Viruses are typically thought of as outside the tree of life, yet they face similar replication problems as other organisms. The solutions they have evolved for either end-replication or end-protection are different from those used by eukaryotes or prokaryotes. Despite this, there are only so many ways that nature can solve a single problem, and the N15 phage, adenovirus, poxvirus, and

herpesvirus all answer the linear DNA question in familiar ways. Viruses also have the added caveat that they do not typically have any repair machinery to protect their genome from, and they are generally more concerned with proper replication than protecting their genetic information from aberrant host DNA repair.

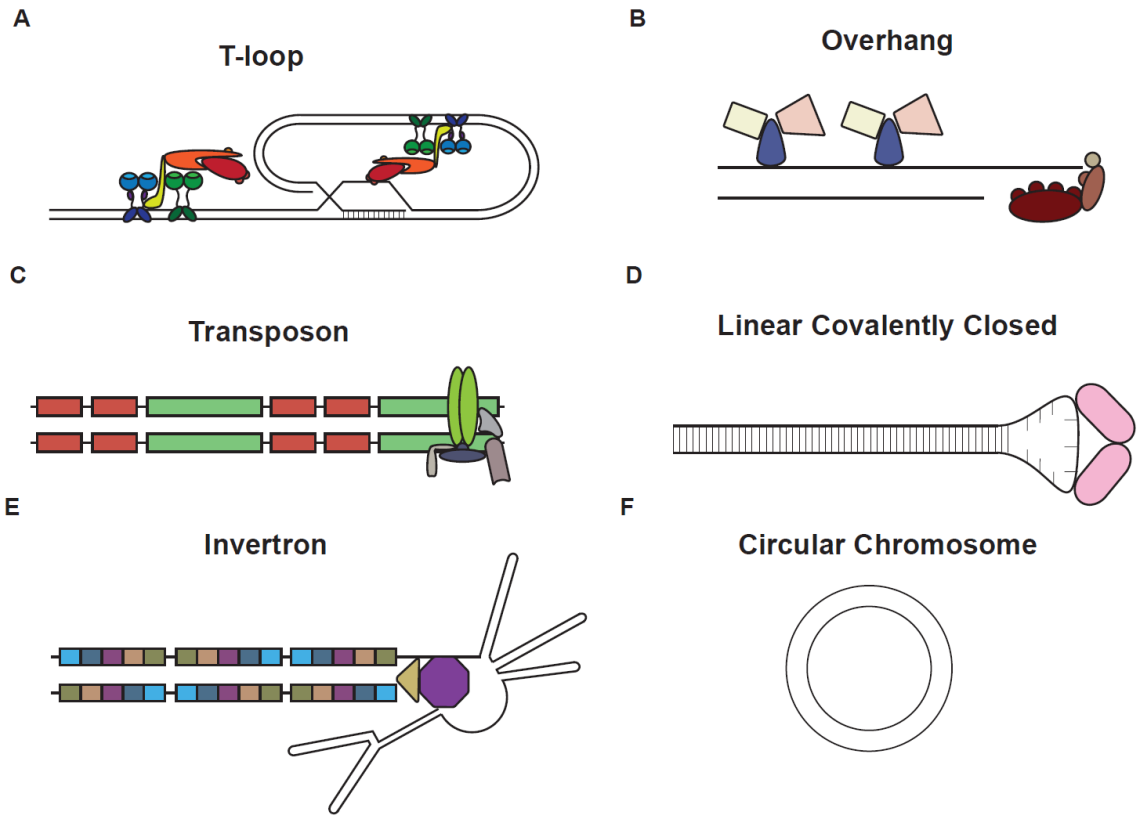
The N15 phage is a bacteriophage that infects *E. coli*, and it was the first example of linear DNA within prokaryotes (*438–440*). The ends of the phage are typically found as a closed linear hairpin, a structure very similar to *Borrelia*, and in fact, the phage also encodes a homologous prototelomerase enzyme named telN (*441*). This prototelomerase is responsible for ligating the ends of the linear plasmid as well as cutting them at the telRL sites after replication and for DNA segregation (*442–444*). In the eukaryotic realm, Poxviridae utilize a similar closed hairpin system like the N15 phage that includes a prototelomerase and inverted repeats for binding (*445–447*).

Adenovirus utilizes a telomere system reminiscent of *Streptomyces* which has a terminal protein that is essential for priming replication. Adenovirus terminal protein is covalently linked to the 5’ end and it serves to initiate replication of the adenovirus DNA by accepting the addition of a dCMP and protect the adenovirus genome from host c-NHEJ (*406*, *448–450*).

Finally, Herpesviridae has a simple trick up its sleeve to protect and replicate its linear genome. Shortly after a herpes virus infects its host, it circularizes its linear genome (*451*). Two factors are required for this: first, the viral protein ICP4 is necessary to circularize the genome (*452*). The second factor is the host’s ligase machinery, specifically Lig4 and XRCC4, without which viral

replication drops over 100-fold (*453*). Circularizing the genome allows for rolling circle replication and prevents the recognition of the Herpesvirus DNA as a broken DNA end and aberrant processing by cellular nucleases.

Figure 30. Six Methods of End-Protection. A) A t-loop based protection system, can be found in mammals and includes two double strand binding proteins (TRF1, TRF2), two bridging proteins (TIN2, TPP1), and a ssDNA binding protein (POT1), as well as an accessory factor (RAP1). B) A protein protected telomere, such as is found in yeast, contains a 3' overhang, a double stranded binding protein (Rap1), two accessory protein (Rif1, Rif2), and a single strand binding complex (CST). C) A transposon-based system found in Diptera with Het-A (red) and TAR-T (green), and the termini complex which consists of DNA binding dimer HP1, the DNA binders HOAP and Ver, and accessory factors HipHop and Moi. D) A linear covalently closed hairpin end such as is found in *Agrobacterium*, with the prototelomerase dimer bound at the end. E) An invertron telomere which can be found in *Streptomyces and* consists of a series of palindromic inverted repeats and two end-binding and replication priming proteins TapL and TpgL. F) A simple and common solution to end-protection, a circular chromosome which leaves no free ends.


A
T-loop
B
Overhang
C
Transposon
D
Linear Covalently Closed
E
Invertron
F
Circular Chromosome

Conclusion

The end-replication and end-protection problems are universal in evolution for all linear genomes. Despite the billions of years of life, only a few solutions have emerged to this puzzle. The most common solution is to forgo linear chromosomes in favor of circular ones, as most prokaryotes and all archaea have done. Otherwise, the most common solution to the challenges posed by linear chromosomes, is the system used primarily by eukaryotes: a short terminal repeat that provides a specific sequence to be recognized, a polymerase to add that sequence to chromosome ends and solve the end-replication problem, and a set of telomere binding proteins to prevent the end from being recognized as DNA break and aberrantly repaired. This system, which involves a t-loop, appears the be the ancestral form of end-protection in eukaryotes, based on parsimoniously drawing an evolutionary tree. However, even within eukaryotes, this system is not universal. An alternative, transposon-based system was also developed in some species. A third system, reminiscent of the eukaryotic telomere, is present in *Streptomyces:* a long terminal repeat and a set of two proteins which are responsible for both protecting the end and serving as a protein primer for replication. This system includes active transposons as the terminal repeats to solve the end-replication problem, and a series of DNA binding proteins that recognize the end and block the DNA damage response from activating. The fourth and final system discovered to date is the hairpin telomere system, one that was originally proposed in eukaryotes but discovered in prokaryotes and viruses: a specific terminal sequence both makes a hairpin and serves as a binding site for

a hairpin resolving protein. This protein can cut the hairpin to allow proper DNA replication and also re-ligate it to protect the end.

Even though these four systems evolved independently, and their components share no homology, they share many similarities. All these systems have a terminal repeat, whether it be a short 6 nucleotide repeat, or many repeats of a 6 kb transposon. In all the systems except the transposon telomeres of *Diptera*, these repeats are essential to mark the end. Finally, all these systems share a protein component who role is either to protect the end from end-joining, to recruit an end maintenance protein, or to process the end, whether by adding telomeric sequence or by serving as a primer to allow its replication.

Chapter 7

Discussion

Here we presented an extensive analysis of TRF2 and the telomere end. First, we started by testing an alternate model of end-protection and finding that it was not a likely explanation. Next, we analyzed TRF2 itself by testing whether or not TRF2 was sufficient to form t-loops and protect the telomere end from ATM signaling and c- as well as alt-NHEJ. After finding that it was, we hypothesized that t-loops were the primary effector by which TRF2 accomplishes end-protection and set out to test which parts of the protein are required to form t-loops. We tested three hypotheses regarding features of TRF2 that may be responsible for t-loop formation, but unfortunately were unable to find a pure separation of function mutant. We also hypothesized that exogenous, non-shelterin components may be required to form a t-loop, but out of a panel of likely candidates we were unable to find any that would be required. Finally, we also described the function of a long isoform of TRF2.

Models for End-Protection

A Compaction Based Model

Although the most commonly accepted model in the field is the t-loop model of end-protection, in recent years an alternative has come to light. This model posits that end-protection is based on compaction of the telomere. This compaction makes the telomere so dense that it is inaccessible to DNA damage sensors such as MRN, or the kinases ATM and ATR. We and others tested this model and found that we were unable to support this view (*244*). The compaction model already had a number of issues – why, for instance, does decompaction

cause a DDR but not lead to telomere fusions? Other evidence also cast doubt on this model. For instance, several DDR factors such as MRN are known to be at telomeres constitutively, which is not consistent with compaction denying access to DDR factors. There was also no clear correlation between decompaction and the severity of the response for instance, TRF1 was the subunit that produced the most striking decompaction phenotype but TRF1 does not produce the strongest DDR and does not activate ATM. Finally, other work from this lab showed that induction of a DSB specifically in the telomeric region was sufficient to induce a DDR, which stands in contrast to the proposed model whereby compaction should prevent a response (*125*).

Ultimately, we tested the model using cells that lacked shelterin at their telomeres and did not find evidence for decompaction that would be consistent with previously published data (Chapter 5, *244*, *454*). Rather, we made two striking discoveries. First – telomeres are not particularly compact to begin with, as compared to repressed heterochromatin regions in fruit fly (Chapter 2). Second – damaged telomeres cluster in a 53BP1-dependent manner, and this could explain the previously published data (*454*). This clustering phenotype correlates with DDR activity and increases the apparent size of the telomeric foci, which could be erroneously interpreted as decompaction in response to the DDR.

The finding that compaction is not the mediator of telomere end protection does not exclude that compaction may play some role. It is known that the recruitment of DDR factors is slowed in areas of heterochromatin, and this may be true of telomeres.

TRF2 is Sufficient for T-loop Formation

In contrast to a compaction-based protection model, our data supports the view that TRF2 is responsible for forming a t-loop, and that this t-loop is critical in preventing ATM activation and c-NHEJ. Our work has shown that TRF2 alone is capable of forming t-loops and protecting telomeres from c-NHEJ (Chapter 2). In addition, we have shown that TRF2 is also capable of repressing ATM signaling in the absence of other shelterin components. This is consistent both with the proposed role of a t-loop in carrying out these functions, and with TRF2 alone being sufficient to form a t-loop.

How Are T-loops Formed?

We next wanted to know what parts of TRF2 actually form t-loops. Being able to find a separation of function mutant in TRF2 that distinguishes t-loop formation from telomeric localization would be a major breakthrough – it would allow us to understand precisely how t-loops contribute to end-protection. We attempted to test three different models of how TRF2 may form t-loops: via its basic domain, through DNA wrapping around the TRFH domain, or by tetramerization. Unfortunately, we were unable to identify true separation of function mutants, which makes it difficult to interpret the data (Chapter 3).

Basic Domain

We had originally hypothesized that the basic domain may be required for t-loop formation, due to its role in t-loop protection (*59*, *60*). We envisioned a

mechanism where the basic domain could encourage formation of the t-loop by preferentially binding the 3-way junction that is formed, and therefore stabilizing the t-loop. However, deletion of the basic domain showed that it is not required for t-loop formation (Chapter 3, *60*). The basic domain may still be required to pin and hold the t-loop in place, and we cannot exclude a possibility where t-loops are rapidly being unwound and reformed in the absence of the basic domain.

A DNA Wrapping Model

Our second hypothesis for t-loop formation was based on the ability of TRF2 to wrap DNA. To test this, we recreated mutants that are unable to wrap DNA around their TRFH domain and tested whether or not they would protect telomeres in vivo. Unfortunately, not only were they unable to protect telomeres from ATM and c-NHEJ, they also had serious defects in telomeric localization, and are therefore hypomorphs (Chapter 3). This makes it impossible to draw any conclusions about whether DNA wrapping is required specifically for t-loop formation or more broadly for TRF2 function and localization.

Although our data regarding the 7R2K mutant was also inconclusive, it may still be a DNA wrapping mutant with a specific t-loop defect (Chapter 3). By changing the chain length of the amino acids required for melting the DNA, it is possible that the DNA cannot be properly melted and t-loops are not formed. Unlike the Top domain mutants, 7R2K still localizes to telomeres, presumably because it can still form DNA contacts. This mutant behaves almost entirely like wild type TRF2, with the exception of lacking t-loops. Once again, it is possible that TRF2 is

able to take over the role of ATM and c-NHEJ suppression from t-loops. It would be interesting to test these mutants in either a RAP1, or a shelterin-free context, to analyze the redundancies that may exist with respect to telomere protection. Similarly, 7R2K may wrap t-loops more tightly, causing them to be undetectable in our assay but still present. This could be addressed by electron microscopy imaging of spreads from 7R2K expressing cells.

A Tetramerization Model

The final hypothesis we tested was whether TRF2 tetramerization was responsible for t-loop formation. Unfortunately, as with our Topless mutants, the tetramerization mutants presented as hypomorphs. However, it is possible that at least one of the tetramerization mutants is a t-loop deficient allele of TRF2. The F59A mutant is an oligomerization mutant which expresses well, localizes moderately well by IF, very well by ChIP, but does not form t-loops. In addition, it only partially suppresses ATM and fusions. This partial suppression could be due to the loss of t-loops. Potentially it is not a complete deprotection because TRF2 is capable of repressing ATM and Ku70/80 through other means. This is purely speculative, and the simplest explanation is in fact that the localization of F59A is disrupted enough to cause it to be unable to fully protect telomere ends.

Hypomorphic TRF2 Mutants

To test both the tetramerization and wrapping models, we hoped to generate separation of function mutants that could divide t-loop formation from

either ATM or c-NHEJ suppression. However, the two sets of mutants that we used to the tetramerization and wrapping models of t-loop formation displayed a significant loss of function and were not separation of function mutants. Consequently, we classified these mutants as hypomorphs, in accordance with Muller's definitions of genetic mutants (*455*). A hypomorph is a mutation that leads to a partial loss of gene function, and this is what we observed with both the tetramerization and the wrapping mutants. These mutants had a consistent failure to perform all of the functions of TRF2: repression of TIFs, repression of c-NHEJ, and telomeric localization was all reduced. Furthermore, these mutants also had defects in telomeric localization and expression.

A second critical reason why we chose to call these mutants hypomorphs is because their defects were proportional across the functions of TRF2. For example, the Topless mutants had approximately 80% of telomeres with a TIF, 80% of chromosomes were fused, telomeric localization was reduced by 80%, as was the relative level of protein expression.

It may be impossible to obtain a simple separation of function mutation in TRF2 that divides t-loop formation from the other roles that TRF2 plays at telomeres. TRF2's action in t-loop formation may play a role in localizing and stabilizing TRF2 at the telomere. In addition, t-loop formation may be absolutely required for ATM and c-NHEJ suppression, in which case a separation of function mutant would be impossible.

Cooperation Across Domains of TRF2

Unfortunately, our work has not been able to identify any single portion of TRF2 that disrupts t-loop formation while maintaining other functions of TRF2. This raises the issue that if for example, the Myb domain is required for t-loop formation but is also necessary for proper localization, it would be impossible to parse out. It may be possible to tether TRF2 to the telomere or to shelterin in-lieu of the myb domain. Potentially domain swapping experiments between TRF1's and TRF2's myb could be done, or using the localization domain of TZAP (*456*). Ultimately, we only know that the basic domain, the extended domain, and the TIN2 and RAP1 binding features are not required. However, the dimerization, tetramerization, myb, and hinge domains may all be required.

In fact, it is very likely that multiple domains may be required. The current data suggests that the Top domain of TRF2 wraps the DNA, causing it to melt (*61*). This likely exposes an area of double-stranded DNA for the free 3' end to invade. Tetramerization of TRF2 may assist in distorting the DNA helix and also act to bring the 3' end to the newly melted DNA, after which it invades. This invasion is then bound and "locked" by the basic domain. Naturally, the myb domain, being the dsDNA binding domain, is required for this action.

No Exogenous Factors Are Required

We were also interested in whether or not non-shelterin factors may be necessary for t-loop formation. The first and easiest factor to test was Apollo, due to the use of a $TRF2^{F120A}$ mutant. Apollo is required for maintenance of the leading

strand overhang, and since loss of Apollo leads to chromatid type fusions, it is reasonable to believe that there are no t-loops formed on non-processed strands (*9*, *10*). We found that although there was a slight reduction in t-loops, it was not significant, and certainly not a 2-fold or more reduction that one would expect (Chapter 4). Therefore, we can conclude that a long 3' overhang is not a requirement for t-loop formation. We cannot rule out that a 3' overhang does not aid t-loop formation though – it may create favorable kinetics for TRF2 to invade the duplex region. DNA can thermodynamically splay, so it would make sense that the 3' overhang is not strictly required, especially if put in close proximity with other single stranded DNA. Slower t-loop formation kinetics may explain the fusions seen in Apollo, and also explain why the phenotype is weak.

We also considered whether recombinases may play a role in t-loop formation. We hypothesized that the essential nature of RAD51 could be explained by its role in t-loop formation, which would also explain why it is not essential in yeast but is essential in mammals. Unfortunately, this proved not to be true, and RAD51 is not required for t-loop formation (Chapter 4). Neither is its ortholog RAD51D. RAD52 remains an interesting and elusive recombinase that still may play a role: it has been known to localize to telomeres, and is an important player in ALT telomere maintenance, and may play a role in loop formation in some yeasts (*298*). Although we did not test whether RAD52 is required for t-loop formation, it would be an interesting experiment.

One other factor that we have considered is Ku70/80. Ku70/80 is present at telomeres, and is also important for repressing alt-NHEJ at telomeres, though it

likely does this as a general rather than a specific repressor (*109*). However, we observed that in the cells in which we removed Ku70/80, there was an overall decrease in the level of t-loops compared to other cell lines. To properly study this, we would need to utilize a floxed allele of Ku70/80. Unpublished data from the lab shows that these cells do not have a strong alt-NHEJ phenotype (likely due to the presence of other repressors). In mice, Ku70/80 is not essential, so either it does not function in t-loop formation or t-loops are not essential. At best, Ku70/80 may be an accessory factor that can aid TRF2, but it would come as a surprise if it were required. Similarly, loss of Ku70/80 does not produce a strong ATM response, which all point towards Ku70/80 not being necessary for t-loop formation. Why Ku70/80 is found at telomeres still remains unresolved.

What do T-loops do?

Protection from Alt-NHEJ and a 3-Factor Model

Our work also raises numerous questions regarding the role of TIN2. For one, we have shown that TRF2 is sufficient to repress both c-NHEJ and the ATM response. Previous work had shown that loss of TIN2 led to a slight increase in c-NHEJ and suggested that the reason that TIN2 has this phenotype is due to the loss of TRF2 from the telomere. This is consistent with data suggesting that overexpression of TRF2 can suppress the TIN2 phenotype (*90*). In addition, we also provide compelling evidence that TRF2 capable of protecting the telomere from the ATM response. Although our data strongly suggests this conclusion, there was still a base level of signaling remaining – this could be a minor independent

role of TIN2, or it could just be an experimental artifact. It is also possible that this low level of signaling comes from crosstalk between the PARP1 signaling pathway that TIN2 is known to suppress, and the ATM pathway (*60*).

We also learned that TRF2 is insufficient to repress alt-NHEJ, and that TIN2 plays an important role in this repression. Previous work has already implicated TIN2 in repressing PARP1 and alt-NHEJ so this does not come as a surprise. What is more surprising is that t-loops seem to provide little or no protection from alt-NHEJ fusions. Alt-NHEJ fusions are present even in TRF2 proficient cells: when TIN2 is stripped from the telomere, the alt-NHEJ fusion phenotype is exacerbated, despite the fact that t-loops are still present.

This leads to a potential model on how alt-NHEJ is repressed. We know that alt-NHEJ fusions increase upon loss of Ku70/80, TRF2, and TIN2. Therefore, we propose that each of these three factors is required to fully repress alt-NHEJ. Ku70/80 is a general repressor, but it is found constitutively at telomeres, and also interacts with TRF2 (*122*, *136*, *149*). Its role in repressing alt-NHEJ may explain why it localizes to telomeres. TIN2 is necessary for repression of PARP1 signaling. TRF2 is the third and final factor necessary for repression of alt-NHEJ. Whether it does so through t-loops or through its own action is unclear – but given that t-loops on their own are not protective, this suggests that TRF2 may be suppressing alt-NHEJ through another mechanism. Alternatively, it is possible that t-loops interfere with alt-NHEJ by presenting the chromosome end in an unfavorable conformation, and that the action of TRF2 is mediated by t-loops.

Loss of one factor increases alt-NHEJ slightly, in the range of 0-2%. This can be seen with single deletions of Ku70/80, TIN2, or TRF2 in a $Lig4^{-/-}$ setting. Loss of two factors result in fusions in the 2 – 10% range. Deletion of TRF2 and Ku70/80, or TIN2 and Ku70/80 are such examples. Loss of all three factors leads to a large increase in fusions: TRF2 TIN2 Ku70/80 or TRF1 TRF2 Ku70/80 cells exhibit 25 – 60% alt-NHEJ fusions. It is unclear why alt-NHEJ never reach 100% of telomeres - possibly there is a fourth factor that is contributes to repressing alt-NHEJ or possibly alt-NHEJ is not efficient and by the timepoints analyzed does not fuse all the telomeres. Since we have not been able to separate TRF2 from t-loops, it is possible t-loops are a fourth factor themselves, capable of repressing alt-NHEJ without TRF2.

It is also very possible that the basic domain, although not required for t-loop formation, provides an important stabilizing function to the 3-way junction. It may be the case that loss of the basic domain, in addition to loss of TIN2 and Ku70/80 would actually cause the same level of alt-NHEJ as loss of TRF2. If so, this would suggest that the t-loop and basic domain are the effectors of alt-NHEJ suppression.

The T-loop Model

T-loops are believed to protect telomeres from chromosome fusions as well as from activation of the ATM response. However, there are a few instances where t-loops appear to fail, but other considerations need to be taken into account. T-loops appear fail to protect from the ATM response in a prolonged mitosis, and

telomeres fuse if the fusion machinery is reactivated (*457*, *458*). However, this can be explained by the loss of TRF2 from telomeres and the loss of t-loops, rather than their inability to protect the chromosome end.

Both the stability of t-loops, and the kinetics of their formation could be monitored with a temperature sensitive allele of TRF2 (*226*). T-loops are also unable to protect telomeres from the ATR response – this is the role of the POT1 proteins, which occlude the 3' overhang (*78*). Another consideration is that loss of POT1 leads to a low level (~3.5%) of telomere fusions despite the fact that POT1 is not required for t-loop formation (*138*, *232*). It is possible that POT1, along with the basic domain, stabilize the t-loop, without which it becomes more vulnerable to fusions. It would be interesting to monitor t-loop frequencies in cells lacking POT1 and expressing TRF2$^{\Delta basic}$.

In other studies of TRF2 separation of function mutants, t-loops have been suggested to either protect from fusions but not signaling, or from signaling but not fusions (*61*, *148*, *458*). Other work has suggested that t-loops are necessary for both. It is possible that t-loops are only part of the mechanism by which TRF2 protects the telomere. One of the few ways to test this would be to generate an artificial t-loop – perhaps by bringing RAD51 to the telomere. It may also be possible to create t-loop structures in vitro with crosslinking, and then transfect them back into cells, though the risk of the crosslink being recognized as damage is present. Such structures would either need to have a loop on both ends, or an internal marker for sequencing so that it would be possible to tell which ends are preferentially fused. Finally, one could use an in vitro cell extract system and add

artificial t-loops or purified telomeric DNA gently stripped of protein to gauge if an ATM response or fusions are activated.

We do know that t-loops are not a silver bullet – they fail to protect telomeres from alternative non-homologous end-joining. We also know that they are not strictly required in other organisms – certain yeasts, and even fruit flies lack t-loops yet are still able to solve the end-protection problem. None-the-less, t-loops do an admirable job of protecting the chromosome from a litany of threats.

Other Work

Evolution of TRF2

We also examined the extended domain of TRF2. This domain clearly cannot be involved in t-loop formation because the short form of TRF2 does not contain this domain yet is capable of forming t-loops. However, what this domain does is still a mystery. We did not detect any differences in its interaction with DNA, or with known binding partners. However, we did not pursue whether or not it interacts with exogenous binding partners. It is possible, though unlikely, that the extended domain of TRF2 is just an evolutionary artifact with too little a burden to be selected against. What is more likely is that $TRF2^{Long}$ plays an important role in conditions that we have not tested, such as during differentiation of embryonic cells or in neurons, where $TRF2^{Long}$.

Conclusion

T-loops are complex structures – both in their arrangement at the chromosome end, and in their difficult to decipher role in end-protection. Lately the t-loop model has been called into question, either by chromatin compaction-based models, or by temporal models that suggest that t-loops are not entirely protective. Here we attempted to challenge the t-loop model, and we were unable to disprove it. In fact, we provided further evidence that TRF2 and t-loops are important players, demonstrating that they are sufficient to protect against ATM activation and protect from c-NHEJ alone.

Ultimately, one of the most convincing arguments for the function of t-loops is their evolutionary conservation. Why would nature preserve a structure for nearly a billion years if it did not provide any function? Even organisms that do not have t-loops find some way to hide the end of their chromosome, whether it be in a hairpin, or to occlude the 3’ end with a binding protein or complicated secondary structures. Masking the 3’ end must be important, even if it is not sufficient by itself to prevent deleterious downstream events.

Chapter 8

Telomere Imaging Appendix

Introduction

Traditionally, telomeres have been visualized using either negative stain, shadowing electron microscopy, or by fluorescence microscopy (*137*, *205*, *459*). Fluorescence microscopy can be used to visualize either the telomeric DNA using FISH probes, or telomeric proteins using immunofluorescence. Both EM and light microscopy techniques have their limitations: electron microscopy is time consuming and cannot be used to visualize telomeres inside cells, and fluorescence microscopy cannot resolve finer structural elements of the telomere.

Recent advances in microscopy have led to the development of super resolution technology, of which structured illumination microscopy (SIM), STORM, and photoactivatable localization microscopy (PALM), microscopy have been applied to the study of telomeres (*228*, *460–462*). STORM and PALM share a similar mechanism whereby dye molecules (or fluorescent proteins for PALM) are stimulated by a high powered laser to emit single photons at random intervals (as reviewed in (*463*)). A high-speed camera captures these emissions for many minutes and then each emission is precisely localized to within a 20 nm radius of the actual molecule to generate an image. SIM uses interfering wavelengths of light to generate grids on the image, which can then be shifted in phase and rotated. High frequency information is then extracted from the image sets using an algorithm, which generates an image with about a two-fold increase in resolution. STORM and SIM have been used to image t-loops, and PALM has been used to measure chromatin compaction at telomeres. Additionally, many of the studies on the interactions between shelterin and chromatin, as well as studies on DNA

wrapping around TRF2, have been done using AFM microscopy (*464*). Atomic force microscopy works by having a very fine detector physically brush over the sample which induces an electrical current and a force on the detector, which is subsequently conveyed into an image of the sample's surface (*465–467*).

An alternative technique to generate data beyond what a traditional light microscope can resolve in a cell is Expansion Microscopy (ExM) (*468*). This recently developed method works on the principle of uniformly physically enlarging a sample which increases the spatial resolution and the signal to noise ratio by diluting the background over the total expanded area. Such an approach allows a conventional microscope to resolve fine details that were previously obscured. Samples are typically embedded in a gel and the fluorescent tags detecting biological structures of interest are crosslinked to the gel matrix to mark their position. The sample is then dissolved and the gel osmotically swelled and imaged.

One of the original aims of the work presented here was to visualize a t-loop in a native conformation and in an intact nucleus. Prior work on t-loops has been done either in vitro, or on purified telomeric DNA by STORM or EM. Prior attempts to visualize a t-loop using immunogold EM were unable to do so (*469*). This raises concerns that t-loops may be artifacts of whatever method of preparation was used to visualize them. In addition, it also does not allow us to answer questions such as: are t-loops found at every telomere and are t-loops present continuously throughout the cell cycle.

We were also interested in using microscopy to study the telomeric overhang. Very little is known about the positioning of the overhang relative to the

rest of the telomere, and this is of particular interest because it is single-stranded rather than double-stranded DNA. Direct visualization of the overhang would add to our understanding of telomere replication and maintenance. Further, precise localization of the overhang would allow us to study the spatial organization of the shelterin complex; for instance, it may classify whether all the subunits are distributed evenly across the telomere or whether there is a positional preference (such as POT1 being enriched near the overhang).

ExM has not been used in combination with super-resolution imaging, but if it is, should generate images of higher resolution than either technique alone because they work in mechanistically different ways. ExM and STORM would also still retain the advantage of being able to use conventional dyes and antibodies. Therefore, the use of ExM in combination with a super-resolution technique such as STORM would be a significant advance in the field of light microscopy. If applied to the telomere, it would allow us to visualize a t-loop within an intact nucleus and provide visual proof for the relevance of the t-loop model in vivo.

Results

We began by focusing our efforts on visualizing the telomeric overhang in $TRF1^{F/F}$ $TRF2^{F/F}$ $53BP1^{-/-}$ cells, due to the fact that they have a very long overhang stemming from hyper-resection (*82*). A longer overhang should lead to a more prominent overhang signal that is easier to hybridize. In addition, removal of TRF2 prevents the formation of t-loops and would free the overhang from a t-loop conformation where it would be hidden and difficult to resolve.

We first attempted to use a dual hybridizing technique to label telomeres on DNA spreads (Fig. 31A). We reasoned that this would be a proving ground to test if we could visualize a longer overhang: without t-loops, we expect that the 3' end should extend further out, and we can visualize this by observing a non-overlapping portion of the signal. In addition, performing DNA spreading prior to visualization allows us to focus on a single telomere in a relaxed conformation, where we reasoned it would be easier to observe non-overlapping signal. The hybridization was performed in the same manner as a standard FISH, with the exception of using two probes added simultaneously in a 1:1 ratio. Unfortunately, we were never able to observe any non-overlapping signal or extension of the G-strand, and shown are the two most prominent examples of the 3' overhang (green) extending past the rest of the DNA molecule (Fig. 31A).

We believe that the failure to observe a phenotype is mainly due to a technical issue: the amount of labeling we were able to achieve was very poor. It is likely due to two factors: the crosslinking treatment damaging the DNA and the use of two probes. Using two probes that can hybridize to each other prevents binding of the probe with high affinity to the telomeric DNA, as they may preferentially bind each other. Consequently, we were never able to accurately assess where the double stranded region ends and where the single stranded region begins.

We next used the same cell lines but transitioned to imaging intact nuclei, where we hoped to see either a specific localization of the overhang, or an increase in overhang signal after we treated the cells with tamoxifen to induce Cre mediated

deletion (Fig. 31B). We believed that the nuclei may allow for better hybridization as there is no crosslinking step to interfere with probe binding, and there is also no chance of telomeres breaking and losing the 3' overhang. Unfortunately, we encountered a different technical issue: we were unable to properly align the telomeric signals. This makes it difficult to tell if there is an extension of an overhang signal beyond a double stranded telomeric signal (Fig. 31B). We attempted to validate our approach by comparing the signal of the probes before and after tamoxifen induced gene deletion, which should correspond to an increase in signal strength of the overhang probe. However, this analysis was confounded due to both the weak signal of the probes, as well as the high variability of signal intensity. This led to us being unable to conclude whether or not we could even detect overhang signal by STORM microscopy.

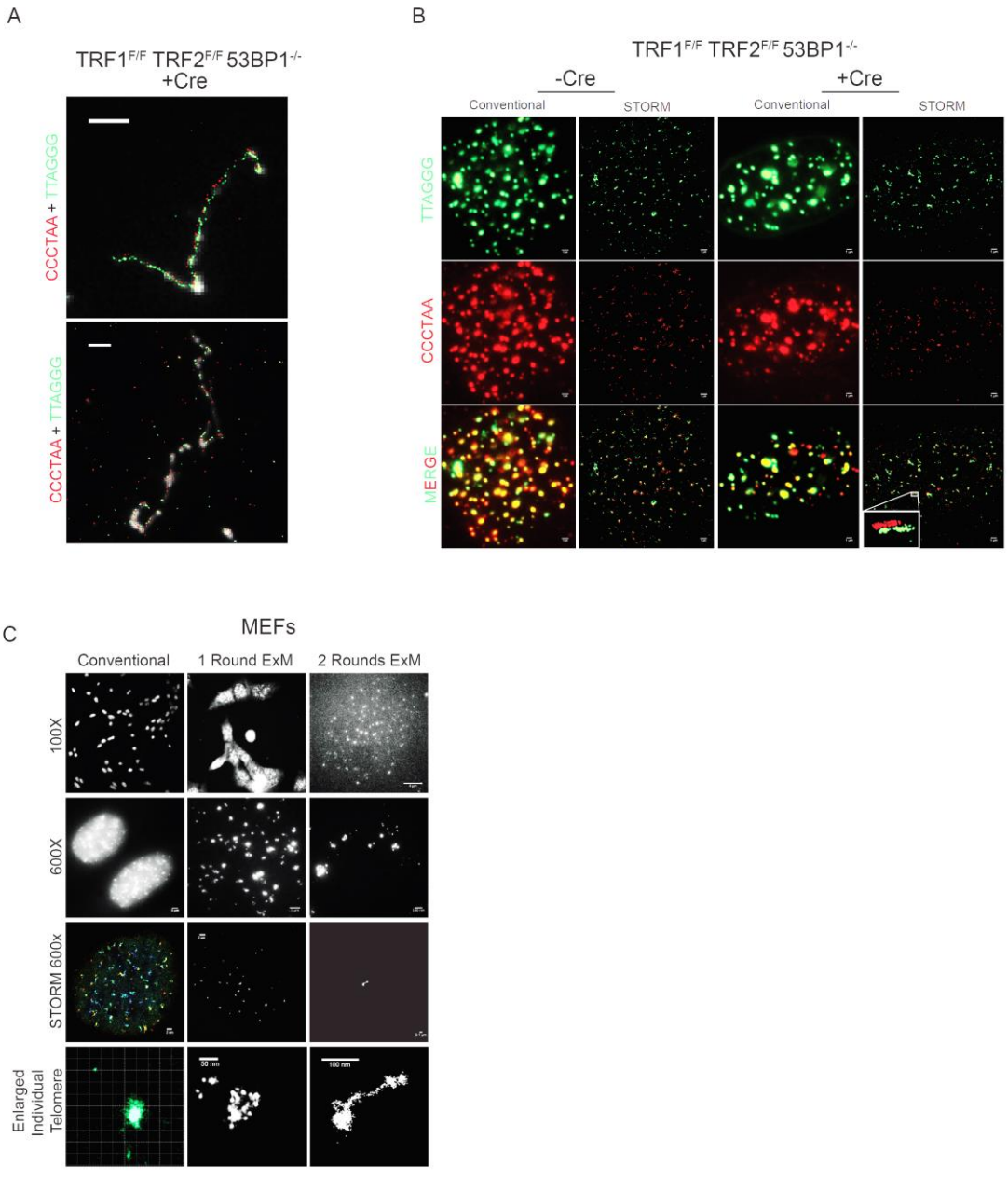
Finally, we combined ExM and STORM to try to visualize t-loops within individual nuclei in wild-type MEFs (Fig. 31C). We reasoned that the combined resolution increase of ExM (4 – 16-fold) and STORM (10-fold) would be enough to detect t-loops inside nuclei. We attempted to visualize telomeres using regular and iterative expansion microscopy, with mixed results. The greatest challenge was hybridizing telomeric probe to sufficient density for uniform labeling. We ultimately settled on a protocol of labeling the sample after gelation and expansion, because the free radicals in the gel polymerization step destroyed the probe. However, diffusion of the probe into the gel made uniform labeling challenging. We were able to verify that ExM was working by comparing the size of nuclei before and after no ExM, 1 round of ExM, or two rounds of ExM (left, middle, and right columns). After

individual nuclei or individual telomeric localizations, we observed three kinds of signal: a scattered diffuse pattern, balls of signal (bottom row, middle), or very rarely a tailed signal (bottom row, right) (Fig. 31C). Although we would like to suspect that the tailed signal pattern is evidence of t-loops, we are unable to know whether gaps in signal density are real or due to inefficient labeling. Furthermore, we did not perform this experiment in a setting where we could genetically remove t-loops as a control.

These results are preliminary but imaging the samples exposed two additional issues with this technique. First, stabilization of the gel is very challenging, and failure to have a secure sample leads to uninterpretable data. However, it is difficult to prevent the gel from wobbling, especially on the time scale that is required for STORM imaging. The second issue was finding the sample. Only the very top of the gel is usable for imaging due to its thickness. Additionally, due to the expansion, encountering a telomere – which may be oriented in the z plane rather than in the xy plane – is rare.

Figure 31. Super-resolution Imaging of Telomeres.

(A) Double labeled spread telomeres of TRF1$^{F/F}$ TRF2$^{F/F}$ 53BP1$^{-/-}$ cells 120 hours after tamoxifen induced gene deletion. Green – A647-TelC. Red – A750-TelG. (B) Double labeled telomeres within a nucleus of TRF1$^{F/F}$ TRF2$^{F/F}$ 53BP1$^{-/-}$ cells before tamoxifen induced gene deletion or 120 hours after. Green – A647-TelC. Red – A750-TelG. No differences were observed. (C) Panel of regular, ExM, STORM, iterative ExM, and ExM + STORM images taken in TRF1$^{F/F}$ TRF2$^{F/F}$ 53BP1$^{-/-}$ cells but untreated with tamoxifen. Bottom row: post-processing zoom on individual telomeric clusters, as determined by the density of localizations. Note the somewhat loop like appearance of the telomeres visualized by ExM.


A
TRF1F/F TRF2F/F 53BP1-/-
+Cre
CCCTAA + TTAGGG
CCCTAA + TTAGGG
B
TRF1F/F TRF2F/F 53BP1-/-
-Cre
+Cre
Conventional
STORM
Conventional
STORM
TTAGGG
CCCTAA
MERGE
C
MEFs
Conventional
1 Round ExM
2 Rounds ExM
100X
600X
STORM 600x
Enlarged Individual Telomere
50 nm
100 nm

Imaging t-loops was one of the original aims of this thesis, but we encountered significant technical challenges. We were particularly interested in in vivo visualization of a t-loop, within an intact nucleus. This would answer whether or not every telomere is in a t-loop conformation, provide further evidence that t-loops are not an artifact. It may also be possible to capture t-loops in distinct steps of the strand invasion process. We tested a number of different approaches, with varying results. Despite significant difficulties with labeling and stabilization our results gave a tantalizing glimpse into the future of telomere imaging.

We were also curious about the distribution of proteins along the telomere, and whether subunits of the shelterin complex were preferentially localized to different areas. For example, TRF2 and POT1 may be enriched at the base of the t-loop due to their role in protecting it. We attempted using STORM microscopy on intact telomeres but found that they were too dense to resolve. We tried to perform DNA spreads with an altered, more gentle lysis to preserve the protein-DNA interactions, or to crosslink proteins to DNA prior to spreading in the hopes of being able to detect a distribution of shelterin components along the telomeric DNA or t-loop. Unfortunately, we were unable to optimize and effectively use these techniques.

Ultimately, we decided to abandon this project due to the extensive technical difficulties. However, certain elements of these experiments show promise and could be expanded upon with additional optimization. First, we did observe loop structures after combined ExM and STORM. To verify if these are real t-loops, an experiment should be done where cells are imaged before and

after TRF2 deletion, and the number of these looped telomeres scored. This experiment could demonstrate the first t-loop inside an intact nucleus. Unfortunately, with the quality of imaging we achieved, this experiment would not be able to answer whether all t-loops are in a looped conformation. Second, we did not try to use ExM with an antibody, but it may prove more efficient than FISH.

The two-color imaging experiments could be improved upon by adding fiducial markers to help align the images. We tried to use a protein, Rap1, as a marker for the double-stranded telomeric region and this increased our image quality. However, because we were using cells lacking shelterin at their telomeres to generate a long telomeric overhang, we were unable to use this technique in the experimental conditions. As an alternative, one could use Rap1 as a marker for double-stranded DNA, use a tel-C probe for the overhang, and perform the experiment in $TPP1^{F/F}$ $53BP1^{-/-}$ cells, which should also have a dramatic overhang phenotype.

In the future advances in imaging techniques may help to solve these issues. In particular, Cryo-EM tomography of cellular structures in their native states may allow us to visualize a t-loop. Current proof of concept work has imaged the nuclear pore within a nucleus, ribosomes, and nucleosomes to 28 angstroms (*470*, *471*). It is possible that the “telosome” can be imaged in the near future as well, and the structure of the t-loop accurately determined. In addition, advances in mass spectrometry have made it possible to analyze the composition of subcellular structures down to 1 μM (*472–474*). If performed on spread telomeres

or isolated and purified telomeres, it would be possible to determine the distribution of the shelterin complex along the telomere.

Despite our inability to bring these experiments to a conclusion, they still demonstrate a tantalizing possibility: that imaging a t-loop and the telomeric overhang in a cell is nearly possible. By slightly changing the cellular conditions, optimizing the protocols further, and with the likely advent of better stabilization tools and other reagents, there is a real possibility that these experiments can be performed in the future.

Chapter 9

Automated Imaging Appendix

Introduction

One of the most time consuming and challenging activities in regards to imaging is the actual scoring and analysis of the images that have been acquired. To this end, we wrote two scripts to assist with this process that can help ensure greater accuracy of scoring, more efficient use of time, and the ability to compare results in an unbiased manner across different experiments. The program is compatible with all image types.

The script consists of two main parts: first, a macro in ImageJ that analyzes colocalizations, and second, a Python script that will export the data from ImageJ into organized Excel tables. The first file is meant to run internally in ImageJ or Fiji and is generally based on the Colocalization plugin developed by Pierre Bourdoncle (*475*, *476*). The last stable version of Fiji with which "The Incredible Focus Finder" is compatible with is 1.49. The macro can analyze images with up to four channels and comes with a value tester script. The premise of the macro is as follows: the user selects thresholding values to determine the cutoff for what they consider a focus, and then the program counts the foci it finds in all the automatically segmented nuclei, counts the overlap, and then returns the values. Nuclei segmentation and foci size can be customized in the actual script code, while foci values are requested from the user in prompts.

The second part of the program takes the exported ImageJ values, and converts them into a single Excel file per condition. This program is able to sort by the number of conditions or by image, and can provide the user with a readout of how many of each foci their images contain, how many colocalizations their images

contain, and also basic statistical analysis such as whether their images contain greater than 5, 10, 15, or 20 of each focus.

Script

Fiji

```
var filenameIndex = 1;


macro "The Incredible Focus Finder [F1]" {

    run("Set Measurements...", "integrated redirect=None decimal=3");

    dir1 = getDirectory("Choose Source Directory ");
    dir2 = getDirectory("Choose Destination Directory ");
    dir = dir2
    triple_coloc = false;

    File.makeDirectory(dir2 + "First Foci");
    File.makeDirectory(dir2 + "Second Foci");
    File.makeDirectory(dir2 + "Coloc");
    File.makeDirectory(dir2 + "Image");
    File.makeDirectory(dir2 + "Nuclei");
    File.makeDirectory(dir2 + "Third Foci");
    dirlist = getFileList(dir2);

    Telosave = dir2 + dirlist[1];
    BP1save = dir2 + dirlist[4];
    Colocsave = dir2 + dirlist[0];
    Imagesave = dir2 + dirlist[2];
    Nucleisave = dir2 + dirlist[3];
    Thirdsave = dir2 + dirlist[5];

    channel_number = getNumber("How Many Channels?", 3);
    if (channel_number == 4) {
        rows = 1;
        columns = 4;
        rc = rows * columns;
              labels = newArray("Channel 1", "Channel 2", "Channel 3",
      "Channel 4");
        defaults = newArray(rc);
        Dialog.create("Please Select Which Channels To Colocalize");
        Dialog.addCheckboxGroup(rows, columns, labels, defaults);
        Dialog.show();
        channels = newArray();
        C_1 = Dialog.getCheckbox();
        C_2 = Dialog.getCheckbox();
        C_3 = Dialog.getCheckbox();
        C_4 = Dialog.getCheckbox();
```

```
        if (channel_number == 4 && ((C_2 == 1 && C_3 == 1 && C_4 == 1)
|| (C_1 == 1 && C_3 == 1 && C_4 == 1) || (C_1 == 1 && C_2 == 1 && C_3
== 1))) {
            triple_coloc = true;

        } else {
            triple_coloc = false;
        }
    }


    first_autocontrastfull = getBoolean("Would you like use auto-
contrasting on the first foci?");
    if (first_autocontrastfull == 0) {

        first_autocontrastmin = getNumber("Enter Min Value", 0);
        first_autocontrastmax = getNumber("Enter Max Value", 255);

    }


    if (channel_number != 2) {
        second_autocontrastfull = getBoolean("Would you like use auto-
contrasting on the second foci?");
        if (second_autocontrastfull == 0) {

            second_autocontrastmin = getNumber("Enter Min Value", 0);
            second_autocontrastmax = getNumber("Enter Max Value", 255);

        }
    }

    if (channel_number == 4 && triple_coloc == 1) {
        third_autocontrastfull = getBoolean("Would you like use auto-
contrasting on the third foci?");
        if (third_autocontrastfull == 0) {

            third_autocontrastmin = getNumber("Enter Min Value", 0);
            third_autocontrastmax = getNumber("Enter Max Value", 255);

        }
    }

    first_thresh = getNumber("Threshold value for first foci", -180);

    if (channel_number != 2) {
        second_thresh = getNumber("Threshold value for second foci", -
180);
    }

    if (channel_number == 4 && triple_coloc == 1) {
        third_thresh = getNumber("Threshold value for third foci", -
180);
    }


    dapi_order = getNumber("Which channel is DAPI?", 1);
```

```
    setBatchMode(true);
    list = getFileList(dir1);
    for (n = 0; n < list.length; n++) {
        showProgress(n + 1, list.length);


        run("Collect Garbage");


        open(dir1 + list[n]);


        imagedir = File.directory;
        imagename = File.name;

        if (channel_number == 4) {
            run("Stack to Hyperstack...", "order=xyczt(default)
channels=4 slices=1 frames=1 display=Color");
            if (dapi_order == 1)
                run("Duplicate...", "title=dapi duplicate channels=1");
            else if (dapi_order == 2)
                run("Duplicate...", "title=dapi duplicate channels=2");
            else if (dapi_order == 3)
                run("Duplicate...", "title=dapi duplicate channels=3");
            else if (dapi_order == 4)
                run("Duplicate...", "title=dapi duplicate channels=4");


            if (triple_coloc == 1) {
                if (C_1 == 1 && C_2 == 1 && C_3 == 1) {
                    selectWindow(imagename);
                    run("Duplicate...", "title=firstfocichannel
duplicate channels=1");
                    selectWindow(imagename);
                    run("Duplicate...", "title=secondfocichannel
duplicate channels=2");
                    run("Merge Channels...", "c1=[firstfocichannel]
c2=[secondfocichannel] c3=[dapi] create");
                }
                if (C_2 == 1 && C_3 == 1 && C_4 == 1) {
                    selectWindow(imagename);
                    run("Duplicate...", "title=firstfocichannel
duplicate channels=2");
                    selectWindow(imagename);
                    run("Duplicate...", "title=secondfocichannel
duplicate channels=3");
                    run("Merge Channels...", "c1=[firstfocichannel]
c2=[secondfocichannel] c3=[dapi] create");

                }
                if (C_1 == 1 && C_3 == 1 && C_4 == 1) {
                    selectWindow(imagename);
                    run("Duplicate...", "title=firstfocichannel
duplicate channels=1");
                    selectWindow(imagename);
```

```
                    run("Duplicate...", "title=secondfocichannel
duplicate channels=3");
                    run("Merge Channels...", "c1=[firstfocichannel]
c2=[secondfocichannel] c3=[dapi] create");
                }
            }
            if (triple_coloc == 0) {
                if (C_1 == 1 && C_2 == 1) {
                    selectWindow(imagename);
                    run("Duplicate...", "title=firstfocichannel
duplicate channels=1");
                    selectWindow(imagename);
                    run("Duplicate...", "title=secondfocichannel
duplicate channels=2");
                    run("Merge Channels...", "c1=[firstfocichannel]
c2=[secondfocichannel] c3=[dapi] create");
                }
                if (C_1 == 1 && C_3 == 1) {
                    selectWindow(imagename);
                    run("Duplicate...", "title=firstfocichannel
duplicate channels=1");
                    selectWindow(imagename);
                    run("Duplicate...", "title=secondfocichannel
duplicate channels=3");
                    run("Merge Channels...", "c1=[firstfocichannel]
c2=[secondfocichannel] c3=[dapi] create");
                }
                if (C_1 == 1 && C_4 == 1) {
                    selectWindow(imagename);
                    run("Duplicate...", "title=firstfocichannel
duplicate channels=1");
                    selectWindow(imagename);
                    run("Duplicate...", "title=secondfocichannel
duplicate channels=4");
                    run("Merge Channels...", "c1=[firstfocichannel]
c2=[secondfocichannel] c3=[dapi] create");
                }
                if (C_2 == 1 && C_3 == 1) {
                    selectWindow(imagename);
                    run("Duplicate...", "title=firstfocichannel
duplicate channels=2");
                    selectWindow(imagename);
                    run("Duplicate...", "title=secondfocichannel
duplicate channels=3");
                    run("Merge Channels...", "c1=[firstfocichannel]
c2=[secondfocichannel] c3=[dapi] create");
                }
                if (C_2 == 1 && C_4 == 1) {
                    selectWindow(imagename);
                    run("Duplicate...", "title=firstfocichannel
duplicate channels=2");
                    selectWindow(imagename);
                    run("Duplicate...", "title=secondfocichannel
duplicate channels=4");
                    run("Merge Channels...", "c1=[firstfocichannel]
c2=[secondfocichannel] c3=[dapi] create");
                }
```

```
                if (C_3 == 1 && C_4 == 1) {
                    selectWindow(imagename);
                    run("Duplicate...", "title=firstfocichannel
duplicate channels=3");
                    selectWindow(imagename);
                    run("Duplicate...", "title=secondfocichannel
duplicate channels=4");
                    run("Merge Channels...", "c1=[firstfocichannel]
c2=[secondfocichannel] c3=[dapi] create");
                }
            }

            close(imagename);
            close("firstfocichannel");
            close("secondfocichannel");
            close("dapi");
            selectWindow("Composite");
            rename(imagename);
        }

        if (channel_number == 3 || channel_number == 4) {
            run("Stack to Hyperstack...", "order=xyczt(default)
channels=3 slices=1 frames=1 display=Color");
        }

        if (channel_number == 3) {

            if (dapi_order == 1)
                run("Arrange Channels...", "new=231");
            else if (dapi_order == 2)
                run("Arrange Channels...", "new=132");
            else if (dapi_order == 3)
                run("Arrange Channels...", "new=123");
        }

        if (channel_number == 2) {
            run("Stack to Hyperstack...", "order=xyczt(default)
channels=2 slices=1 frames=1 display=Color");
            if (dapi_order == 1)
                run("Arrange Channels...", "new=21");
            else if (dapi_order == 2)
                run("Arrange Channels...", "new=12");
        }

        run("32-bit");

        if (channel_number != 2) {
            run("Duplicate...", "title=dapi duplicate channels=3");
            selectWindow(imagename);
            run("Duplicate...", "title=bp1 duplicate channels=2");
            selectWindow(imagename);
            run("Duplicate...", "title=telo duplicate channels=1");
            run("Duplicate...", "title=telo1");
            run("Median...", "radius=10");
            imageCalculator("Subtract create", "telo", "telo1");
            selectWindow("bp1");
            run("Duplicate...", "title=bp2");
```

```
            run("Median...", "radius=10");
            imageCalculator("Subtract create", "bp1", "bp2");
            run("Merge Channels...", "c1=[Result of telo] c2=[Result of
bp1] c3=[dapi] create");
            close("telo");
            close("telo1");
            close("bp1");
            close("bp2");
            close(imagename);
            selectWindow("Composite");
            run("Stack to Hyperstack...", "order=xyczt(default)
channels=3 slices=1 frames=1 display=Color");
        }


        if (channel_number == 2) {
            run("Duplicate...", "title=dapi duplicate channels=2");
            selectWindow(imagename);
            run("Duplicate...", "title=telo duplicate channels=1");
            run("Duplicate...", "title=telo1");
            run("Median...", "radius=10");
            imageCalculator("Subtract create", "telo", "telo1");
            run("Merge Channels...", "c1=[Result of telo] c2=[dapi]
create");
            close("telo");
            close("telo1");
            selectWindow(imagename);
            selectWindow("Composite");
            run("Stack to Hyperstack...", "order=xyczt(default)
channels=2 slices=1 frames=1 display=Color");
            selectWindow("Composite");
        }

        run("Duplicate...", "title=Green duplicate channels=1");
        run("Grays");
        run("Subtract Background...", "rolling=14");
        selectWindow("Composite");

        //

        if (channel_number != 2) {
            run("Duplicate...", "title=Red duplicate channels=2");
            run("Grays");
            run("Subtract Background...", "rolling=14");
            selectWindow("Composite");
            run("Duplicate...", "title=Blue duplicate channels=3");
        } else {
            run("Duplicate...", "title=Blue duplicate channels=2");
        }


        run("Grays");
        run("Subtract Background...", "rolling=300");
        run("Smooth");
        run("Gaussian Blur...", "sigma=3");
        setAutoThreshold("Huang dark");
        //run("Threshold...");
```

```
        setOption("BlackBackground", false);
        run("Convert to Mask");
        selectWindow("Composite");
        close();

        if (channel_number != 2) {

            selectWindow("Red");
            run("8-bit");
            run("Despeckle");

            if (second_autocontrastfull == 1) {


                //autocontrast
                AUTO_THRESHOLD = 5000;
                getRawStatistics(pixcount);
                limit = pixcount / 10;
                threshold = pixcount / AUTO_THRESHOLD;
                nBins = 256;
                getHistogram(values, histA, nBins);
                i = -1;
                found = false;
                do {
                    counts = histA[++i];
                    if (counts > limit) counts = 0;
                    found = counts > threshold;
                } while ((!found) && (i < histA.length - 1))
                hmin = values[i];
                i = histA.length;
                do {

                    counts = histA[--i];
                    if (counts > limit) counts = 0;
                    found = counts > threshold;
                } while ((!found) && (i > 0))
                hmax = values[i];
                setMinAndMax(hmin, hmax);
                //print(hmin, hmax);
                run("Apply LUT");
            } else {
                setMinAndMax(second_autocontrastmin,
second_autocontrastmax);
                run("Apply LUT");
            }

            run("Auto Local Threshold", "method=Mean radius=500
parameter_1=second_thresh parameter_2=0 white"); //Thresholding for
second
            run("Convert to Mask");
            run("Make Binary");
        }

        selectWindow("Green");
        run("8-bit");
        run("Despeckle");
```

```
        if (first_autocontrastfull == 1) {


            //autocontrast
            AUTO_THRESHOLD = 5000;
            getRawStatistics(pixcount);
            limit = pixcount / 10;
            threshold = pixcount / AUTO_THRESHOLD;
            nBins = 256;
            getHistogram(values, histA, nBins);
            i = -1;
            found = false;
            do {
                counts = histA[++i];
                if (counts > limit) counts = 0;
                found = counts > threshold;
            } while ((!found) && (i < histA.length - 1))
            hmin = values[i];
            i = histA.length;
            do {

                counts = histA[--i];
                if (counts > limit) counts = 0;
                found = counts > threshold;
            } while ((!found) && (i > 0))
            hmax = values[i];
            setMinAndMax(hmin, hmax);
            //print(hmin, hmax);
            run("Apply LUT");
        } else {
            setMinAndMax(first_autocontrastmin, first_autocontrastmax);
            run("Apply LUT");
        }


        run("Auto Local Threshold", "method=Mean radius=500
parameter_1=first_thresh parameter_2=0 white"); //Thresholding for
first one
        run("Convert to Mask");
        run("Make Binary");


        if (channel_number != 2) {
            run("Merge Channels...", "red=Red green=Green blue=*None*
gray=*None* create");
            selectWindow("Composite");
            run("Stack to RGB");
            run("Split Channels");
            selectWindow("Composite (RGB) (blue)");
            close();
        } else {
            run("Merge Channels...", "red=Green green=Green blue=*None*
gray=*None* create");
            selectWindow("Composite");
            run("Stack to RGB");
            run("Split Channels");
            selectWindow("Composite (RGB) (blue)");
```

```
            close();
            selectWindow("Composite (RGB) (red)");
            close();
        }


        if (channel_number != 2) {
            run("Colocalization ", "channel_1=[Composite (RGB) (red)]
channel_2=[Composite (RGB) (green)] ratio=30 threshold_channel_1=15
threshold_channel_2=15 display=255  colocalizated");
        }
        if (channel_number == 4 && triple_coloc == 1) {

            selectWindow("Colocalizated points (RGB) ");
            run("Make Binary");
            rename("Colocalizated points Part 1");
            open(imagedir + imagename);


            if ((C_2 == 1 && C_3 == 1 && C_4 == 1) | (C_1 == 1 && C_3
== 1 && C_4 == 1)) {
                selectWindow(imagename);
                run("Duplicate...", "title=thirdfocichannel duplicate
range=4");
                close(imagename);
            }

            if ((C_1 == 1 && C_2 == 1 && C_3 == 1)) {
                selectWindow(imagename);
                run("Duplicate...", "title=thirdfocichannel duplicate
range=3");
                close(imagename);
            }

            selectWindow("thirdfocichannel");
            run("Duplicate...", "title=third_med");
            run("Median...", "radius=10");
            imageCalculator("Subtract create", "thirdfocichannel",
"third_med");
            close("thirdfocichannel");
            close("third_med");
            selectWindow("Result of thirdfocichannel");
            rename("third_proc");
            run("32-bit");
            run("Grays");
            run("Subtract Background...", "rolling=14");
            selectWindow("third_proc");
            run("8-bit");
            run("Despeckle");

            if (third_autocontrastfull == 1) {


                //autocontrast
                AUTO_THRESHOLD = 5000;
                getRawStatistics(pixcount);
                limit = pixcount / 10;
```

```
                threshold = pixcount / AUTO_THRESHOLD;
                nBins = 256;
                getHistogram(values, histA, nBins);
                i = -1;
                found = false;
                do {
                    counts = histA[++i];
                    if (counts > limit) counts = 0;
                    found = counts > threshold;
                } while ((!found) && (i < histA.length - 1))
                hmin = values[i];
                i = histA.length;
                do {

                    counts = histA[--i];
                    if (counts > limit) counts = 0;
                    found = counts > threshold;
                } while ((!found) && (i > 0))
                hmax = values[i];
                setMinAndMax(hmin, hmax);
                //print(hmin, hmax);
                run("Apply LUT");
            } else {
                setMinAndMax(third_autocontrastmin,
third_autocontrastmax);
                run("Apply LUT");
            }



            run("Auto Local Threshold", "method=Mean radius=500
parameter_1=third_thresh parameter_2=0 white"); //Thresholding for
third one
            run("Convert to Mask");
            run("Make Binary");

            run("Colocalization ", "channel_1=[Colocalizated points
Part 1] channel_2=[third_proc] ratio=30 threshold_channel_1=15
threshold_channel_2=15 display=255  colocalizated");

        }


        selectWindow("Blue");
        run("Convert to Mask");
        run("Make Binary");
        run("Adjustable Watershed", "tolerance=4");
        run("Analyze Particles...", "size=1000-infinity pixel
circularity=0.00-1.00 show=[Bare Outlines] summarize display add");


        if (channel_number != 2) {
            selectWindow("Colocalizated points (RGB) ");
            run("Make Binary");
            run("Analyze Particles...", "size=2-1000 pixel
circularity=0.00-1.00 show=[Masks] display"); //Size of Colocalizations
to Count
```

```
            run("Fill Holes");
            run("Find Maxima...", "noise=10 output=[Single Points]
light");
            roiManager("Show All with labels");
            roiManager("Show All");
            run("Clear Results");
            roiManager("Measure");
            IJ.renameResults("Foci Colocalization");


            selectWindow("Composite (RGB) (red)");
            run("Make Binary");
            run("Analyze Particles...", "size=2-1000 pixel
circularity=0.00-1.00 show=[Masks] display"); //Size of 53BP1 Foci to
Count
            run("Fill Holes");
            run("Find Maxima...", "noise=10 output=[Single Points]
light");
            roiManager("Show All with labels");
            roiManager("Show All");
            run("Clear Results");
            roiManager("Measure");
            IJ.renameResults("Red Foci");
        }


        selectWindow("Composite (RGB) (green)");
        run("Make Binary");
        run("Analyze Particles...", "size=2-1000 pixel
circularity=0.00-1.00 show=[Masks] display"); //Size of Telomeres to
Count
        run("Fill Holes");
        run("Find Maxima...", "noise=10 output=[Single Points] light");
        roiManager("Show All with labels");
        roiManager("Show All");
        run("Clear Results");
        roiManager("Measure");
        IJ.renameResults("Green Foci");

        if (channel_number == 4 && triple_coloc == 1) {
            selectWindow("third_proc");
            run("Make Binary");
            run("Analyze Particles...", "size=2-1000 pixel
circularity=0.00-1.00 show=[Masks] display"); //Size of Third Foci to
Count
            run("Fill Holes");
            run("Find Maxima...", "noise=10 output=[Single Points]
light");
            roiManager("Show All with labels");
            roiManager("Show All");
            run("Clear Results");
            roiManager("Measure");
            IJ.renameResults("Third Foci");

        }
```

```
        selectWindow("Green Foci");
        nameG = "First Foci";
        index = lastIndexOf(nameG, ".");
        if (index != -1) nameG = substring(nameG, 0, index);
        nameG = nameG + ".csv"; // Change here for put a different name
        saveAs("Measurements", Telosave + list[n] + " " + nameG);

        if (channel_number != 2) {
            selectWindow("Red Foci");
            nameR = "Second Foci";
            index = lastIndexOf(nameR, ".");
            if (index != -1) nameR = substring(nameR, 0, index);
            nameR = nameR + ".csv"; // Change here for put a different
name
            saveAs("Measurements", BP1save + list[n] + " " + nameR);

            selectWindow("Foci Colocalization");
            nameC = "Foci Colocalization Count";
            index = lastIndexOf(nameC, ".");
            if (index != -1) nameC = substring(nameC, 0, index);
            nameC = nameC + ".csv"; // Change here for put a different
name
            saveAs("Measurements", Colocsave + list[n] + " " + nameC);
            filenameIndex++;
        }

        if (channel_number == 4 && triple_coloc == 1) {
            selectWindow("Third Foci");
            nameT = "Third Foci";
            index = lastIndexOf(nameT, ".");
            if (index != -1) nameT = substring(nameT, 0, index);
            nameT = nameT + ".csv"; // Change here for put a different
name
            saveAs("Measurements", Thirdsave + list[n] + " " + nameT);

            selectWindow("third_proc");
            run("Invert LUT");
            run("Cyan");
            selectWindow("Composite");
            run("Split Channels");
            run("Merge Channels...", "c1=C1-Composite c2=C2-Composite
c5=third_proc create");

        }


        selectWindow("Composite");
        nameD = "Foci Colocalization Image";
        saveAs("Tiff", Imagesave + list[n] + " " + nameD);

        selectWindow("Drawing of Blue");
        nameB = "Nuclei Image";
        saveAs("Tiff", Nucleisave + list[n] + " " + nameB);


        array1 = newArray("0");;
        for (i = 1; i < roiManager("count"); i++) {
```

```
            array1 = Array.concat(array1, i);
        }
        roiManager("select", array1);
        roiManager("Delete");

        run("Close All");
        if (isOpen("ROI Manager")) {
            selectWindow("ROI Manager");
            run("Close");
        }


        call("java.lang.System.gc");

        closelist = getList("window.titles");
        for (x = 0; x < closelist.length; x++) {
            winame = closelist[x];
            selectWindow(winame);
            run("Close");
        }


    }

    if (first_autocontrastfull == 1) {
        print("First Foci Auto-contrasted");
    } else {
        print("First Foci Contrast Values" + " " +
first_autocontrastmin + " " + first_autocontrastmax);
    }

    if (channel_number != 2) {
        if (second_autocontrastfull == 1) {
            print("Second Foci Auto-contrasted");
        } else {
            print("Second Foci Contrast Values" + " " +
second_autocontrastmin + " " + second_autocontrastmax);
        }
    }

    if (channel_number == 4 && triple_coloc == 1) {
        if (third_autocontrastfull == 1) {
            print("Third Foci Auto-contrasted");
        } else {
            print("Third Foci Contrast Values" + " " +
third_autocontrastmin + " " + third_autocontrastmax);
        }
    }

    print("First Threshold Value is" + " " + first_thresh);


    if (channel_number != 2) {
        print("Second Threshold Value is" + " " + second_thresh);
    }

    if (channel_number == 4 && triple_coloc == 1) {
```

```
        print("Third Threshold Value is" + " " + third_thresh);
    }

    selectWindow("Log"); //select Log-window
    saveAs("Text", dir2 + "Value Log");

    if (isOpen("Log")) {
        selectWindow("Log");
        run("Close");
    }

}
exit



macro "Close All Windows [F4]" {
        while (nImages > 0) {
            selectImage(nImages);
            close();
            list = getList("window.titles");
            for (i = 0; i < list.length; i++) {
                winame = list[i];
                selectWindow(winame);
                run("Close");

            }
            exit


            macro "Value Tester [F6]" {
                runcycle = 1
                for (tryagainprogram = 0; tryagainprogram == 0;
runcycle++) {


                    waitForUser("Press OK and then please open a test
file");
                    open();

                    run("Set Measurements...", "integrated
redirect=None decimal=3");
                    test_image = File.directory;
                    test_name = File.name;
                    triple_coloc = false;

                    print("The test image is");
                    print(test_image + test_name);
                    print("Test cycle..." + runcycle);

                    channel_number = getNumber("How Many Channels?",
3);
                    dapi_order = getNumber("Which channel is DAPI?",
1);
                    if (channel_number == 4) {
                        rows = 2;
                        columns = 2;
```

```
                        rc = rows * columns;
                        labels = newArray("Channel 1", "Channel 2",
"Channel 3", "Channel 4");
                        defaults = newArray(rc);
                        Dialog.create("Please Select Which Channels To
Colocalize");
                        Dialog.addCheckboxGroup(rows, columns, labels,
defaults);
                        Dialog.show();
                        channels = newArray();
                        C_1 = Dialog.getCheckbox();
                        C_2 = Dialog.getCheckbox();
                        C_3 = Dialog.getCheckbox();
                        C_4 = Dialog.getCheckbox();


                        if (channel_number == 4 && ((C_2 == 1 && C_3 ==
1 && C_4 == 1) || (C_1 == 1 && C_3 == 1 && C_4 == 1) || (C_1 == 1 &&
C_2 == 1 && C_3 == 1))) {
                            triple_coloc = true;

                        } else {
                            triple_coloc = false;
                        }
                    }


                    if (channel_number == 4) {
                        run("Stack to Hyperstack...",
"order=xyczt(default) channels=4 slices=1 frames=1 display=Color");
                        if (dapi_order == 1)
                            run("Duplicate...", "title=dapi duplicate
channels=1");
                        else if (dapi_order == 2)
                            run("Duplicate...", "title=dapi duplicate
channels=2");
                        else if (dapi_order == 3)
                            run("Duplicate...", "title=dapi duplicate
channels=3");
                        else if (dapi_order == 4)
                            run("Duplicate...", "title=dapi duplicate
channels=4");

                        if (triple_coloc == 1) {
                            if (C_1 == 1 && C_2 == 1 && C_3 == 1) {
                                print("Colocalized Channels are 1 and 2
and 3");
                                selectWindow(test_name);
                                run("Duplicate...",
"title=firstfocichannel duplicate channels=1");
                                selectWindow(test_name);
                                run("Duplicate...",
"title=secondfocichannel duplicate channels=2");
                                run("Merge Channels...",
"c1=[firstfocichannel] c2=[secondfocichannel] c3=[dapi] create");
                            }
                            if (C_2 == 1 && C_3 == 1 && C_4 == 1) {
```

```
                                print("Colocalized Channels are 2 and 3
and 4");
                                selectWindow(test_name);
                                run("Duplicate...",
"title=firstfocichannel duplicate channels=2");
                                selectWindow(test_name);
                                run("Duplicate...",
"title=secondfocichannel duplicate channels=3");
                                run("Merge Channels...",
"c1=[firstfocichannel] c2=[secondfocichannel] c3=[dapi] create");

                            }
                            if (C_1 == 1 && C_3 == 1 && C_4 == 1) {
                                print("Colocalized Channels are 1 and 3
and 4");
                                selectWindow(test_name);
                                run("Duplicate...",
"title=firstfocichannel duplicate channels=1");
                                selectWindow(test_name);
                                run("Duplicate...",
"title=secondfocichannel duplicate channels=3");
                                run("Merge Channels...",
"c1=[firstfocichannel] c2=[secondfocichannel] c3=[dapi] create");
                            }
                        }

                        if (triple_coloc == 0) {


                            if (C_1 == 1 && C_2 == 1) {
                                print("Colocalized Channels are 1 and
2");
                                selectWindow(test_name);
                                run("Duplicate...",
"title=firstfocichannel duplicate channels=1");
                                selectWindow(test_name);
                                run("Duplicate...",
"title=secondfocichannel duplicate channels=2");
                                run("Merge Channels...",
"c1=[firstfocichannel] c2=[secondfocichannel] c3=[dapi] create");
                            }
                            if (C_1 == 1 && C_3 == 1) {
                                print("Colocalized Channels are 1 and
3");
                                selectWindow(test_name);
                                run("Duplicate...",
"title=firstfocichannel duplicate channels=1");
                                selectWindow(test_name);
                                run("Duplicate...",
"title=secondfocichannel duplicate channels=3");
                                run("Merge Channels...",
"c1=[firstfocichannel] c2=[secondfocichannel] c3=[dapi] create");
                            }
                            if (C_1 == 1 && C_4 == 1) {
                                print("Colocalized Channels are 1 and
4");
                                selectWindow(test_name);
```

```
                                run("Duplicate...",
"title=firstfocichannel duplicate channels=1");
                                selectWindow(test_name);
                                run("Duplicate...",
"title=secondfocichannel duplicate channels=4");
                                run("Merge Channels...",
"c1=[firstfocichannel] c2=[secondfocichannel] c3=[dapi] create");
                            }
                            if (C_2 == 1 && C_3 == 1) {
                                print("Colocalized Channels are 2 and
3");
                                selectWindow(test_name);
                                run("Duplicate...",
"title=firstfocichannel duplicate channels=2");
                                selectWindow(test_name);
                                run("Duplicate...",
"title=secondfocichannel duplicate channels=3");
                                run("Merge Channels...",
"c1=[firstfocichannel] c2=[secondfocichannel] c3=[dapi] create");
                            }
                            if (C_2 == 1 && C_4 == 1) {
                                print("Colocalized Channels are 2 and
4");
                                selectWindow(test_name);
                                run("Duplicate...",
"title=firstfocichannel duplicate channels=2");
                                selectWindow(test_name);
                                run("Duplicate...",
"title=secondfocichannel duplicate channels=4");
                                run("Merge Channels...",
"c1=[firstfocichannel] c2=[secondfocichannel] c3=[dapi] create");
                            }
                            if (C_3 == 1 && C_4 == 1) {
                                print("Colocalized Channels are 3 and
4");
                                selectWindow(test_name);
                                run("Duplicate...",
"title=firstfocichannel duplicate channels=3");
                                selectWindow(test_name);
                                run("Duplicate...",
"title=secondfocichannel duplicate channels=4");
                                run("Merge Channels...",
"c1=[firstfocichannel] c2=[secondfocichannel] c3=[dapi] create");
                            }
                        }
                        close(test_name);
                        close("firstfocichannel");
                        close("secondfocichannel");
                        close("dapi");
                        selectWindow("Composite");
                        rename(test_name);
                    }

                    if (channel_number == 3 || channel_number == 4) {
                        run("Stack to Hyperstack...",
"order=xyczt(default) channels=3 slices=1 frames=1 display=Color");
                    }
```

```
                    if (channel_number == 3) {

                        if (dapi_order == 1)
                            run("Arrange Channels...", "new=231");
                        else if (dapi_order == 2)
                            run("Arrange Channels...", "new=132");
                        else if (dapi_order == 3)
                            run("Arrange Channels...", "new=123");
                    }

                    if (channel_number == 2) {
                        run("Stack to Hyperstack...",
"order=xyczt(default) channels=2 slices=1 frames=1 display=Color");
                        if (dapi_order == 1)
                            run("Arrange Channels...", "new=21");
                        else if (dapi_order == 2)
                            run("Arrange Channels...", "new=12");
                    }

                    run("32-bit");

                    if (channel_number != 2) {
                        run("Duplicate...", "title=dapi duplicate
channels=3");
                        selectWindow(test_name);
                        run("Duplicate...", "title=bp1 duplicate
channels=2");
                        selectWindow(test_name);
                        run("Duplicate...", "title=telo duplicate
channels=1");
                        run("Duplicate...", "title=telo1");
                        run("Median...", "radius=10");
                        imageCalculator("Subtract create", "telo",
"telo1");
                        selectWindow("bp1");
                        run("Duplicate...", "title=bp2");
                        run("Median...", "radius=10");
                        imageCalculator("Subtract create", "bp1",
"bp2");
                        run("Merge Channels...", "c1=[Result of telo]
c2=[Result of bp1] c3=[dapi] create");
                        close("telo");
                        close("telo1");
                        close("bp1");
                        close("bp2");
                        close();
                        run("Stack to Hyperstack...",
"order=xyczt(default) channels=3 slices=1 frames=1 display=Color");
                    }

                    if (channel_number == 2) {
                        run("Duplicate...", "title=dapi duplicate
channels=2");
                        selectWindow(test_name);
                        run("Duplicate...", "title=telo duplicate
channels=1");
```

```
                        run("Duplicate...", "title=telo1");
                        run("Median...", "radius=10");
                        imageCalculator("Subtract create", "telo",
"telo1");
                        run("Merge Channels...", "c1=[Result of telo]
c2=[dapi] create");
                        close("telo");
                        close("telo1");
                        close(test_name);
                        selectWindow("Composite");
                        run("Stack to Hyperstack...",
"order=xyczt(default) channels=2 slices=1 frames=1 display=Color");
                        selectWindow("Composite");
                    }


                    run("Duplicate...", "title=Green duplicate
channels=1");
                    run("Grays");
                    run("Subtract Background...", "rolling=14");
                    selectWindow("Composite");

                    if (channel_number != 2) {
                        run("Duplicate...", "title=Red duplicate
channels=2");
                        run("Grays");
                        run("Subtract Background...", "rolling=14");
                        selectWindow("Composite");
                        run("Duplicate...", "title=Blue duplicate
channels=3");
                    } else {
                        run("Duplicate...", "title=Blue duplicate
channels=2");
                    }

                    run("Grays");
                    run("Subtract Background...", "rolling=300");
                    run("Smooth");
                    run("Gaussian Blur...", "sigma=3");
                    setAutoThreshold("Huang dark");
                    //run("Threshold...");
                    setOption("BlackBackground", false);
                    run("Convert to Mask");
                    selectWindow("Composite");
                    close();


                    selectWindow("Green");
                    run("8-bit");
                    run("Despeckle");

                    first_autocontrast = getBoolean("Would you like to
auto-contrast the first foci?");
                    if (first_autocontrast == 1) {


                        //autocontrast
```

```
                        AUTO_THRESHOLD = 5000;
                        getRawStatistics(pixcount);
                        limit = pixcount / 10;
                        threshold = pixcount / AUTO_THRESHOLD;
                        nBins = 256;
                        getHistogram(values, histA, nBins);
                        i = -1;
                        found = false;
                        do {
                            counts = histA[++i];
                            if (counts > limit) counts = 0;
                            found = counts > threshold;
                        } while ((!found) && (i < histA.length - 1))
                        hmin = values[i];
                        i = histA.length;
                        do {

                            counts = histA[--i];
                            if (counts > limit) counts = 0;
                            found = counts > threshold;
                        } while ((!found) && (i > 0))
                        hmax = values[i];
                        setMinAndMax(hmin, hmax);
                        //print(hmin, hmax);
                        run("Apply LUT");
                        print("1st Foci... Auto-contrasted");
                    } else {
                        run("Brightness/Contrast...");
                        waitForUser("Select your values AND DO NOT
PRESS APPLY, then press OK");
                        getMinAndMax(min, max);
                        print("First Foci Min value is...");
                        print(min);
                        print("First Foci Max value is...");
                        print(max);
                        setMinAndMax(min, max);
                        run("Apply LUT");
                    }

                    waitForUser("Zoom in on the foci to identify a
cutoff value. When you are done, press OK.");
                    first_thresh = getNumber("What thresholding value
would you like to test for the first foci?", -180);
                    print("1st Foci Thresholding Value is...");
                    print(first_thresh);

                    run("Auto Local Threshold", "method=Mean radius=500
parameter_1=first_thresh parameter_2=0 white"); //Thresholding for
first one
                    run("Convert to Mask");
                    run("Make Binary");

                    if (channel_number != 2) {

                        selectWindow("Red");

                        run("8-bit");
```

```
                        run("Despeckle");

                        second_autocontrast = getBoolean("Would you
like to auto-contrast the second foci?");
                        if (second_autocontrast == 1) {


                            //autocontrast
                            AUTO_THRESHOLD = 5000;
                            getRawStatistics(pixcount);
                            limit = pixcount / 10;
                            threshold = pixcount / AUTO_THRESHOLD;
                            nBins = 256;
                            getHistogram(values, histA, nBins);
                            i = -1;
                            found = false;
                            do {
                                counts = histA[++i];
                                if (counts > limit) counts = 0;
                                found = counts > threshold;
                            } while ((!found) && (i < histA.length -
1))
                            hmin = values[i];
                            i = histA.length;
                            do {

                                counts = histA[--i];
                                if (counts > limit) counts = 0;
                                found = counts > threshold;
                            } while ((!found) && (i > 0))
                            hmax = values[i];
                            setMinAndMax(hmin, hmax);
                            //print(hmin, hmax);
                            run("Apply LUT");
                            print("2nd Foci... Auto-contrasted");
                        } else {
                            run("Brightness/Contrast...");
                            waitForUser("Select your values AND DO NOT
PRESS APPLY, then press OK");
                            getMinAndMax(min, max);
                            print("Second Foci Min value is...");
                            print(min);
                            print("Second Foci Max value is...");
                            print(max);
                            setMinAndMax(min, max);
                            run("Apply LUT");
                        }

                        waitForUser("Zoom in on the foci to identify a
cutoff value. When you are done, press OK.");
                        second_thresh = getNumber("What thresholding
value would you like to test for the second foci?", -180);
                        print("2nd Foci Thresholding Value is...");
                        print(second_thresh);
                        run("Auto Local Threshold", "method=Mean
radius=500 parameter_1=second_thresh parameter_2=0 white");
//Thresholding for second
```

```
                        run("Convert to Mask");
                        run("Make Binary");
                    }


                    if (channel_number != 2) {
                        run("Merge Channels...", "red=Red green=Green
blue=*None* gray=*None* create");
                        selectWindow("Composite");
                        run("Stack to RGB");
                        run("Split Channels");
                        selectWindow("Composite (RGB) (blue)");
                        close();
                    } else {
                        run("Merge Channels...", "red=Green green=Green
blue=*None* gray=*None* create");
                        selectWindow("Composite");
                        run("Stack to RGB");
                        run("Split Channels");
                        selectWindow("Composite (RGB) (blue)");
                        close();
                        selectWindow("Composite (RGB) (red)");
                        close();
                    }

                    if (channel_number != 2) {
                        run("Colocalization ", "channel_1=[Composite
(RGB) (red)] channel_2=[Composite (RGB) (green)] ratio=30
threshold_channel_1=15 threshold_channel_2=15 display=255
colocalizated");
                    }
                    if (channel_number == 4 && triple_coloc == 1) {

                        selectWindow("Colocalizated points (RGB) ");
                        run("Make Binary");
                        rename("Colocalizated points Part 1");
                        open(test_image + test_name);


                        if ((C_2 == 1 && C_3 == 1 && C_4 == 1) | (C_1
== 1 && C_3 == 1 && C_4 == 1)) {
                            selectWindow(test_name);
                            run("Duplicate...", "title=thirdfocichannel
duplicate range=4");
                            close(test_name);
                        }

                        if ((C_1 == 1 && C_2 == 1 && C_3 == 1)) {
                            selectWindow(test_name);
                            run("Duplicate...", "title=thirdfocichannel
duplicate range=3");
                            close(test_name);
                        }

                        selectWindow("thirdfocichannel");
                        run("Duplicate...", "title=third_med");
                        run("Median...", "radius=10");
```

```
                        imageCalculator("Subtract create",
"thirdfocichannel", "third_med");
                        close("thirdfocichannel");
                        close("third_med");
                        selectWindow("Result of thirdfocichannel");
                        rename("third_proc");
                        run("32-bit");
                        run("Grays");
                        run("Subtract Background...", "rolling=14");
                        selectWindow("third_proc");
                        run("8-bit");
                        run("Despeckle");

                        third_autocontrast = getBoolean("Would you like
to auto-contrast the third foci?");
                        if (third_autocontrast == 1) {


                            //autocontrast
                            AUTO_THRESHOLD = 5000;
                            getRawStatistics(pixcount);
                            limit = pixcount / 10;
                            threshold = pixcount / AUTO_THRESHOLD;
                            nBins = 256;
                            getHistogram(values, histA, nBins);
                            i = -1;
                            found = false;
                            do {
                                counts = histA[++i];
                                if (counts > limit) counts = 0;
                                found = counts > threshold;
                            } while ((!found) && (i < histA.length -
1))
                            hmin = values[i];
                            i = histA.length;
                            do {

                                counts = histA[--i];
                                if (counts > limit) counts = 0;
                                found = counts > threshold;
                            } while ((!found) && (i > 0))
                            hmax = values[i];
                            setMinAndMax(hmin, hmax);
                            //print(hmin, hmax);
                            run("Apply LUT");
                            print("3rd Foci... Auto-contrasted");
                        } else {
                            run("Brightness/Contrast...");
                            waitForUser("Select your values AND DO NOT
PRESS APPLY, then press OK");
                            getMinAndMax(min, max);
                            print("Third Foci Min value is...");
                            print(min);
                            print("Third Foci Max value is...");
                            print(max);
                            setMinAndMax(min, max);
                            run("Apply LUT");
```

```
                        }

                        waitForUser("Zoom in on the foci to identify a
cutoff value. When you are done, press OK.");
                        third_thresh = getNumber("What thresholding
value would you like to test for the third foci?", -180);
                        print("3rd Foci Thresholding Value is...");
                        print(third_thresh);

                        run("Auto Local Threshold", "method=Mean
radius=500 parameter_1=third_thresh parameter_2=0 white");
//Thresholding for third one
                        run("Convert to Mask");
                        run("Make Binary");

                        run("Colocalization ",
"channel_1=[Colocalizated points Part 1] channel_2=[third_proc]
ratio=30 threshold_channel_1=15 threshold_channel_2=15 display=255
colocalizated");
                    }

                    selectWindow("Blue");
                    run("Convert to Mask");
                    run("Make Binary");
                    run("Adjustable Watershed", "tolerance=4");
                    run("Analyze Particles...", "size=1000-infinity
pixel circularity=0.00-1.00 show=[Bare Outlines] summarize display
add");

                    if (channel_number != 2) {
                        selectWindow("Colocalizated points (RGB) ");
                        run("Make Binary");
                        run("Analyze Particles...", "size=2-1000 pixel
circularity=0.00-1.00 show=[Masks] display"); //Size of Colocalizations
to Count
                        run("Fill Holes");
                        run("Find Maxima...", "noise=10 output=[Single
Points] light");
                        roiManager("Show All with labels");
                        roiManager("Show All");
                        run("Clear Results");
                        roiManager("Measure");
                        IJ.renameResults("Foci Colocalization");


                        selectWindow("Composite (RGB) (red)");
                        run("Make Binary");
                        run("Analyze Particles...", "size=2-1000 pixel
circularity=0.00-1.00 show=[Masks] display"); //Size of 53BP1 Foci to
Count
                        run("Fill Holes");
                        run("Find Maxima...", "noise=10 output=[Single
Points] light");
                        roiManager("Show All with labels");
                        roiManager("Show All");
                        run("Clear Results");
                        roiManager("Measure");
```

```
                        IJ.renameResults("Second Foci");
                    }

                    selectWindow("Composite (RGB) (green)");
                    run("Make Binary");
                    run("Analyze Particles...", "size=2-1000 pixel
circularity=0.00-1.00 show=[Masks] display"); //Size of Telomeres to
Count
                    run("Fill Holes");
                    run("Find Maxima...", "noise=10 output=[Single
Points] light");
                    roiManager("Show All with labels");
                    roiManager("Show All");
                    run("Clear Results");
                    roiManager("Measure");
                    IJ.renameResults("First Foci");

                    if (channel_number == 4 && triple_coloc == 1) {
                        selectWindow("third_proc");
                        run("Make Binary");
                        run("Analyze Particles...", "size=2-1000 pixel
circularity=0.00-1.00 show=[Masks] display"); //Size of Third Foci to
Count
                        run("Fill Holes");
                        run("Find Maxima...", "noise=10 output=[Single
Points] light");
                        roiManager("Show All with labels");
                        roiManager("Show All");
                        run("Clear Results");
                        roiManager("Measure");
                        IJ.renameResults("Third Foci");


                        selectWindow("Colocalizated points (8-bit) ");
                        close();
                        selectWindow("Mask of third_proc Maxima");
                        close();
                        selectWindow("Mask of third_proc");
                        close();
                        selectWindow("Colocalizated points Part 1");
                        close();
                        selectWindow("third_proc");
                        run("Invert LUT");
                        run("Cyan");
                        selectWindow("Composite");
                        run("Split Channels");
                        run("Merge Channels...", "c1=C1-Composite
c2=C2-Composite c5=third_proc create");
                    }

                    if (channel_number != 2) {
                        selectWindow("Composite (RGB) (red)");
                        close();
                        selectWindow("Composite (RGB) (green)");
                        close();
                        selectWindow("Colocalizated points (8-bit) ");
                        close();
```

```
                        selectWindow("Colocalizated points (RGB) ");
                        close();
                        selectWindow("Drawing of Blue");
                        close();
                        selectWindow("Mask of Colocalizated points
(RGB) ");
                        close();
                        selectWindow("Mask of Colocalizated points
(RGB)  Maxima");
                        close();
                        selectWindow("Mask of Composite (RGB) (red)");
                        close();
                        selectWindow("Mask of Composite (RGB) (red)
Maxima");
                        close();
                        selectWindow("Mask of Composite (RGB)
(green)");
                        close();
                        selectWindow("Mask of Composite (RGB) (green)
Maxima");
                        close();
                    } else {
                        selectWindow("Composite (RGB) (green)");
                        close();
                        selectWindow("Drawing of Blue");
                        close();
                        selectWindow("Mask of Composite (RGB)
(green)");
                        close();
                        selectWindow("Mask of Composite (RGB) (green)
Maxima");
                        close();
                    }
                    if (isOpen("ROI Manager")) {
                        selectWindow("ROI Manager");
                        run("Close");
                    }

                    open(test_image + test_name);
                    selectWindow(test_name);
                    rename("Original");

                    waitForUser("Review your results, then press OK");
                    tryagainuser = getString("Try Again?", "Y");
                    if (tryagainuser == "N") {
                        tryagainprogram = 1;
                    } else {
                        tryagainprogram = 0;
                        selectWindow("Original");
                        close();
                        selectWindow("Blue");
                        close();
                        selectWindow("Composite");
                        rename("Test" + runcycle);
                    }

                }
```

```
                exit

            }
```

Python

```
import os
import csv
import easygui

def telo(folder_name, i, starting_num):
    list_of_files = os.listdir(folder_name)# remove additional system
files
new_list_of_files = []
for file in list_of_files:
    if file.endswith(".csv"):
    new_list_of_files.append(file)
if i < len(starting_num) - 1:
    new_list_of_files = new_list_of_files[starting_num[i]:
starting_num[i + 1]]
num_of_files = len(new_list_of_files)
else :
    new_list_of_files = new_list_of_files[starting_num[i]::]
num_of_files = len(new_list_of_files)
foci = []
exclusions = []
for j in range(0, num_of_files):
    foci.append([])
exclusions.append([])
total_foci = 0
total_nuclei = 0

image_num = 0

for filename in new_list_of_files:
    f = open(os.path.join(folder_name, filename), "r")
csv_reader = csv.reader(f, delimiter = '\t')
row1 = next(csv_reader)
index = row1[0].split(",").index("RawIntDen")
nuclei_num = 0# exclude cells with < 20 telomeres

for row in csv_reader:
    value = int(float(row[0].split(",")[index]) / 255)

if value < 20:
    exclusions[image_num].append(nuclei_num)

if value >= 20:
    foci[image_num].append(value)
total_foci += value
total_nuclei += 1

nuclei_num += 1
image_num += 1
f.close()
return foci, num_of_files, total_foci, total_nuclei, exclusions
```

```
def foci(folder_name, exclusions, i, starting_num):
    list_of_files = os.listdir(folder_name)# remove additional system
files
new_list_of_files = []
for file in list_of_files:
    if file.endswith(".csv"):
    new_list_of_files.append(file)

if i < len(starting_num) - 1:
    new_list_of_files = new_list_of_files[starting_num[i]:
starting_num[i + 1]]
num_of_files = len(new_list_of_files)
else :
    new_list_of_files = new_list_of_files[starting_num[i]::]
num_of_files = len(new_list_of_files)

foci = []
for j in range(0, num_of_files):
    foci.append([])
total_foci = 0
total_nuclei = 0
a = 0#[0, 5)
b = 0#[5, 10)
c = 0#[10, 15)
d = 0#[15, 20)
e = 0#[20, inf)

image_num = 0
if exclusions == []:
    for filename in new_list_of_files:
    f = open(os.path.join(folder_name, filename), "r")
csv_reader = csv.reader(f, delimiter = '\t')
row1 = next(csv_reader)
index = row1[0].split(",").index("RawIntDen")
nuclei_num = 0

for row in csv_reader:
    value = int(float(row[0].split(",")[index]) / 255)

foci[image_num].append(value)
total_foci += value
total_nuclei += 1
if value < 5:
    a += 1
if value >= 5 and value < 10:
    b += 1
if value >= 10 and value < 15:
    c += 1
if value >= 15 and value < 20:
    d += 1
if value >= 20:
    e += 1
nuclei_num += 1
image_num += 1
f.close()
else :
```

```
    for filename in new_list_of_files:
    f = open(os.path.join(folder_name, filename), "r")
csv_reader = csv.reader(f, delimiter = '\t')
row1 = next(csv_reader)
index = row1[0].split(",").index("RawIntDen")
nuclei_num = 0

for row in csv_reader:
    value = int(float(row[0].split(",")[index]) / 255)

if not nuclei_num in exclusions[image_num]:
    foci[image_num].append(value)
total_foci += value
total_nuclei += 1
if value < 5:
    a += 1
if value >= 5 and value < 10:
    b += 1
if value >= 10 and value < 15:
    c += 1
if value >= 15 and value < 20:
    d += 1
if value >= 20:
    e += 1
nuclei_num += 1
image_num += 1
f.close()
return foci, num_of_files, total_foci, total_nuclei, a, b, c, d, e

def build_csv(foci_1, foci_2, coloc, option, i):
    csvfile = open('Condition ' + str(i + 1) + '.csv', 'w')
writer = csv.writer(csvfile)
writer.writerow(("Image No.", "Nuclei No.", "Telo Foci", "2nd Foci",
"Colocalization"))
for image_num in range(0, len(foci_1[0])):
    for nuclei_num in range(0, len(foci_1[0][image_num])):
    if nuclei_num == 0:
    writer.writerow((image_num + 1, nuclei_num + 1,
foci_1[0][image_num][nuclei_num], foci_2[0][image_num][nuclei_num],
coloc[0][image_num][nuclei_num]))
else :
    writer.writerow(("", nuclei_num + 1,
foci_1[0][image_num][nuclei_num], foci_2[0][image_num][nuclei_num],
coloc[0][image_num][nuclei_num]))
writer.writerow(("", "Total Nuclei", foci_1[3]))
writer.writerow(("", "", "Telo Foci", "2nd Foci", "Colocalization"))
writer.writerow(("", "Avg No. Foci", round(float(foci_1[2]) /
foci_1[3], ndigits = 2), round(float(foci_2[2]) / foci_2[3], ndigits =
2), round(float(coloc[2]) / coloc[3], ndigits = 2)))
if option == "Yes":
    writer.writerow(("", "[0,5)", "", 100 * round(float(foci_2[4]) /
foci_2[3], ndigits = 4), 100 * round(float(coloc[4]) / coloc[3],
ndigits = 4)))
writer.writerow(("", "[5,10)", "", 100 * round(float(foci_2[5]) /
foci_2[3], ndigits = 4), 100 * round(float(coloc[5]) / coloc[3],
ndigits = 4)))
```

```
writer.writerow(("", "[10,15)", "", 100 * round(float(foci_2[6]) /
foci_2[3], ndigits = 4), 100 * round(float(coloc[6]) / coloc[3],
ndigits = 4)))
writer.writerow(("", "[15,20)", "", 100 * round(float(foci_2[7]) /
foci_2[3], ndigits = 4), 100 * round(float(coloc[7]) / coloc[3],
ndigits = 4)))
writer.writerow(("", "[20, inf)", "", 100 * round(float(foci_2[8]) /
foci_2[3], ndigits = 4), 100 * round(float(coloc[8]) / coloc[3],
ndigits = 4)))

writer.writerow(("", ">=5", "", 100 * round(1 - float(foci_2[4]) /
foci_2[3], ndigits = 4), 100 * round(1 - float(coloc[4]) / coloc[3],
ndigits = 4)))
writer.writerow(("", ">=10", "", 100 * round(1 - float(foci_2[4] +
foci_2[5]) / foci_2[3], ndigits = 4), 100 * round(1 - float(coloc[4] +
coloc[5]) / coloc[3], ndigits = 4)))
writer.writerow(("", ">=15", "", 100 * round(1 - float(foci_2[4] +
foci_2[5] + foci_2[6]) / foci_2[3], ndigits = 4), 100 * round(1 -
float(coloc[4] + coloc[5] + coloc[6]) / coloc[3], ndigits = 4)))
writer.writerow(("", ">=20", "", 100 * round(float(foci_2[8]) /
foci_2[3], ndigits = 4), 100 * round(float(coloc[8]) / coloc[3],
ndigits = 4)))
else :
    writer.writerow(("", "[0,5)", 100 * round(float(foci_1[4]) /
foci_1[3], ndigits = 4), 100 * round(float(foci_2[4]) / foci_2[3],
ndigits = 4), 100 * round(float(coloc[4]) / coloc[3], ndigits = 4)))
writer.writerow(("", "[5,10)", 100 * round(float(foci_1[5]) /
foci_1[3], ndigits = 4), 100 * round(float(foci_2[5]) / foci_2[3],
ndigits = 4), 100 * round(float(coloc[5]) / coloc[3], ndigits = 4)))
writer.writerow(("", "[10,15)", 100 * round(float(foci_1[6]) /
foci_1[3], ndigits = 4), 100 * round(float(foci_2[6]) / foci_2[3],
ndigits = 4), 100 * round(float(coloc[6]) / coloc[3], ndigits = 4)))
writer.writerow(("", "[15,20)", 100 * round(float(foci_1[7]) /
foci_1[3], ndigits = 4), 100 * round(float(foci_2[7]) / foci_2[3],
ndigits = 4), 100 * round(float(coloc[7]) / coloc[3], ndigits = 4)))
writer.writerow(("", "[20, inf)", 100 * round(float(foci_1[8]) /
foci_1[3], ndigits = 4), 100 * round(float(foci_2[8]) / foci_2[3],
ndigits = 4), 100 * round(float(coloc[8]) / coloc[3], ndigits = 4)))

writer.writerow(("", ">=5", 100 * round(1 - float(foci_1[4]) /
foci_1[3], ndigits = 4), 100 * round(1 - float(foci_2[4]) / foci_2[3],
ndigits = 4), 100 * round(1 - float(coloc[4]) / coloc[3], ndigits =
4)))
writer.writerow(("", ">=10", 100 * round(1 - float(foci_1[4] +
foci_1[5]) / foci_1[3], ndigits = 4), 100 * round(1 - float(foci_2[4] +
foci_2[5]) / foci_2[3], ndigits = 4), 100 * round(1 - float(coloc[4] +
coloc[5]) / coloc[3], ndigits = 4)))
writer.writerow(("", ">=15", 100 * round(1 - float(foci_1[4] +
foci_1[5] + foci_1[6]) / foci_1[3], ndigits = 4), 100 * round(1 -
float(foci_2[4] + foci_2[5] + foci_2[6]) / foci_2[3], ndigits = 4), 100
* round(1 - float(coloc[4] + coloc[5] + coloc[6]) / coloc[3], ndigits =
4)))
writer.writerow(("", ">=20", 100 * round(float(foci_1[8]) / foci_1[3],
ndigits = 4), 100 * round(float(foci_2[8]) / foci_2[3], ndigits = 4),
100 * round(float(coloc[8]) / coloc[3], ndigits = 4)))
csvfile.close()
```

```
# test code
msg = "Do you want to exclude nuclei with less than 20 telomeric foci?"
choices = ["Yes", "No", "Exit"]
option = easygui.buttonbox(msg, choices = choices)

if option == "Exit":
    exit()

num = easygui.integerbox("How many different conditions?")

starting_num = [0]

if num == 2:
    starting_num.append(easygui.integerbox("What is the starting number
for the 2nd condition?", upperbound = 100000) - 1)

if num == 3:
    starting_num.append(easygui.integerbox("What is the starting number
for the 2nd condition?", upperbound = 100000) - 1)
starting_num.append(easygui.integerbox("What is the starting number for
the 3rd condition?", upperbound = 100000) - 1)

if num > 3:
    starting_num.append(easygui.integerbox("What is the starting number
for the 2nd condition?", upperbound = 100000) - 1)
starting_num.append(easygui.integerbox("What is the starting number for
the 3rd condition?", upperbound = 100000) - 1)
for i in range(4, num + 1):
    starting_num.append(easygui.integerbox("What is the starting number
for the " + str(i) + "th condition?", upperbound = 100000) - 1)

print(starting_num)

msg = "Which one is your Telo folder?"
title = "Telo Folder"
choices = os.listdir(os.curdir)
telo_foci = easygui.choicebox(msg, title, choices)

msg = "Which one is your 2nd foci folder?"
title = "2nd Foci Folder"
choices = os.listdir(os.curdir)
second_foci = easygui.choicebox(msg, title, choices)

msg = "Which one is your colocolization folder?"
title = "Colocalization Folder"
choices = os.listdir(os.curdir)
coloc_foci = easygui.choicebox(msg, title, choices)

for i in range(0, num):
    if option == "Yes":
    foci_1 = telo(telo_foci, i, starting_num)
foci_2 = foci(second_foci, foci_1[4], i, starting_num, )
coloc = foci(coloc_foci, foci_1[4], i, starting_num, )

else :
    foci_1 = foci(telo_foci, [], i, starting_num)
```

```
foci_2 = foci(second_foci, [], i, starting_num)
coloc = foci(coloc_foci, [], i, starting_num)

print(foci_1[1::])
print(foci_2[1::])
print(coloc[1::])

build_csv(foci_1, foci_2, coloc, option, i)
```

Instructions

To install the macros, users must have put the adaptiveThr_.class and adjustable_Watershed.class into FIJI Plugins, put the Colocalization_.class file into FIJI Plugins Analyze, and put the latest text file of the macro into FIJI Macros. Next, they should open FIJI, go to Plugins, Macros, Install, and select “The Incredible Focus Finder”. In addition, the computer they are using must have Python and the easygui plugin for Python. All of these files are kept together as part of the installation package.

From there, users should run the Value Tester and select a representative image. They will be prompted to enter the number of channels, and which channel contains the nuclei stain. They will then be asked if they would like to use auto-contrast or to select their own contrast values, and then what the cutoff value should be. To properly determine the cutoff value, zoom in on foci and identify their intensity; the cutoff value one should select would be the lowest intensity that the user still considers a true focus. Then, input the negative of that number. These steps need to be repeated for each channel analyzed. The value tester will then run and present to you a summary of the data analysis, as well as the values you selected for the data analysis (threshold, contrasting). If you are happy with the

results, you may record the values and use them to run the full program. If not, try another image or try different values.

To run the full analysis, simply select “The Incredible Focus Finder”. It will then prompt you to select which folder your input files are located in and which folder you would like your files to be saved to. Finally, input the values you obtained from value tester. The program will run in the background. When it completes, place the TIFF Excel Merger and easygui.py into the same folder as your exported FIJI data and run the TIFF Excel Merger. The prompts will ask the user for how many data sets they have, as well as how many images are in each data set. Upon completion, it will export the data into Excel and the process will be complete.

Chapter 10

Materials and Methods

Cell culture and gene deletion

TRF1$^{F/F}$ TRF2$^{F/F}$ Lig4$^{-/-}$ p53$^{-/-}$ RosaCre-ERT2 MEFs and SV40LT immortalized TRF1$^{F/F}$ TRF2$^{F/F}$ 53BP1$^{-/-}$, TRF1$^{F/F}$, TRF2$^{F/F}$, TRF2 Ku70$^{-/-}$ and TRF2$^{F/F}$ Lig4$^{-/-}$ RosaCre-ERT2 MEFs were cultured in DMEM supplemented with 10% fetal bovine serum, 10 mM HEPES (pH 7.3), 1 mM sodium pyruvate, 100 U/mL penicillin, 1 µg/mL streptomycin, 0.2 mM glutamine, and 0.1 mM nonessential amino acids and used as described previously (*82*, *86*, *138*, *242*). All MEF lines contained Rosa-CreERT2, allowing induction of Cre with 0.5 µM tamoxifen for 12 h. Time 0 was set at the time the medium was replaced (12 h after addition of tamoxifen). Cells were harvested at 96 or 120 h after medium replacement. For all experiments, the +Cre and −Cre samples were generated and processed in parallel. For experiments that required ATR inhibition, 2.5 µM of inhibitor ETP-46464 was incubated with cells from the point of Cre induction through harvest.

Viral Gene Transfer

For retroviral gene transfer, 2.5 million 293T Phoenix cells were plated in 10% BCS DMEM media and transfected 24 hours later with 10 µg of plasmid DNA using a calcium phosphate technique. The media was changed 12 hours later to correspond to the target cells. Virus was harvested 24 hours, 36 hours, and 48 hours post transfection and filtered through a 0.45 µm filter. Prior to infection, 4 µg/mL polybrene was added to aid in infection of target cells. Cells were infected in three rounds, each 12 hours apart, and selected with 2.5 µg puromycin.

CRISPR gene deletion was performed using the lentiCRISPR v2 system (*189*). Virus was produced in 293FT cells using a standard calcium phosphate transfection protocol. 293FTs were transfected with 37.5 µg of lentiCRISPR v2, 25 µg of VSV-G and 37.5 µg of pPAX2. The supernatant was harvested 48 hours post transfection. Target MEF cells were infected once or twice as needed, 12 hours apart, and selected for viral integration using 2.5 µg of puromycin.

Immunoblotting

Cells (10^6) were lysed in 2× Laemmli buffer and treated for 5 min at 98°C, and the DNA was sheared using an insulin syringe. One-tenth was loaded on an 8-16% gradient SDS-PAGE gel and transferred in 20% methanol transfer buffer for immunoblotting. Antibodies used were mTRF2 (no. 1254), mTRF1 (no. 1249), myc (9B11, Cell Signaling), FLAG (M2, Sigma), Strep (Quiagen). The chemiluminescent signals were detected using ECL Western blotting detection reagents (GE Healthcare) and BioMax MR film or XAR film (Kodak) according to the manufacturer's protocol.

Immunofluorescence and FISH

IF-FISH was performed as described previously (*205*, *477*). Briefly, cells were fixed on coverslips with 3% formaldehyde and subsequently permeabilized with 0.1% Triton and blocked with 3% goat serum and 1 mg/mL BSA. Primary antibodies were Abcam ab175933 for 53BP1 and Millipore 05-636 for γ-H2AX. Secondary antibodies were donkey anti-rabbit conjugated to FITC. For FISH,

slides were dehydrated sequentially in 70%, 90%, and 100% ethanol for 5 minutes each and then air-dried. Slides were hybridized with a $[TTAGGG]_3$-Alexa647 (TelG-A647) PNA or a $[CCCTAA]_3$ probe for 10 min at 70°C–80°C before being placed in a humidity chamber overnight. Slides were then washed twice in 70% formamide with 10 mM Tris (pH 7.2) for 30 minutes and twice in PBS-T for 5 minutes before a second serial dehydration prior to mounting in Prolong Gold and imaging. Images were acquired on a Zeiss Axioplan II with a 63× objective lens and a Hamamatsu C4742-95 and processed with Volocity or on a GE DeltaVision with a 60× objective and processed with FIJI. For higher stringency permeabilization, cells were incubated with pre-extraction buffer (20 mM HepesKOH, pH 7.9, 50 mM NaCl, 3 mM MgCl2, 0.5% Triton X-100, 300 mM sucrose) for 30 seconds on ice, washed twice with PBS, and then fixed with methanol. For irradiation, cells were irradiated with a cesium gamma ray source at 4 Gy and fixed 4 hours later.

CO-FISH Metaphase Spreads

Chromosome orientation FISH was performed as previously described (*478*). Briefly, cells were treated with 200ng/mL colcemid for two hours to enrich for metaphases. Cells were swelled in 75mM KCl for 30 minutes and fixed in 3:1 Methanol : Acetic Acid (added drop by drop while gently vortexing) prior to being dropped on microscope slides. Cells were then treated with RNase A for 15 minutes, Hoescht 33342, UV crosslinked for 5400 µJ, and digested with ExoIII for 30 minutes. Metaphases were then treated following a standard FISH protocol as

described below, except omitting a heat denaturing step. DNA was hybridized with either a 3x CCCTAA – Cy3 (TelC-Cy3) PNA probe or a 3x TTAGGG – Cy5 (TelG-Cy5) PNA probe. Images were acquired on a Zeiss Axioplan II with a 63X objective lens and a Hamamatsu C4742-95 and processed with Volocity or on a GE Deltavision with a 60X objective and processed with FIJI. Regular metaphase spreads were performed the same way, omitting the RNAse, UV crosslinking, and ExoIII digestion steps.

Co-immunoprecipitation

MYC-tagged or Strep-tagged short or long TRF2 was co-expressed with FLAG-tagged Apollo, TIN2, RAP1, TRF1 or TPP1 in 293T cells via transient transfection. Cells were harvested 28 hours after transfection and lysed in Lysis Buffer (50mM NaH_2PO_4 pH 8, 50mM NaCl, 0.5% Tween-20, Roche Protease Inhibitor, 0.5mM PMSF, 1mM DTT). 1% of lysate was used for the whole cell lysate. Cell lysates were incubated with MYC, FLAG, or Strep conjugated beads (Sigma) for two hours at 4°C in the presence of 250 units of Benzonase. The beads were washed twice with Lysis buffer before being resuspended in Laemmli, boiled at 98°C for five minutes, and run on an 8-16% gradient SDS-PAGE for immunoblot analysis. Antibodies used were detailed above.

If PhosSTOP (Roche) was used, it was added to the Lysis Buffer according to manufacturer instructions, and was used in the washes as well. If Lambda phosphatase was used, 400 Units were added to the Lysis Buffer along with 1mM $MnCl_2$, and the same concentration of both was maintained through the wash steps.

Protein Purification

293T cells were co-transfected with Strep tagged RAP1 and short or long TRF2 and harvested 28 hours after transfection. Cells were lysed with Lysis Buffer (50mM NaH_2PO_4 pH 8, 300mM NaCl, 0.5% Tween-20, Roche Protease Inhibitor, 0.5mM PMSF, 1mM DTT), treated with 10U/mL benzonase, and then incubated 2 hours with Strep conjugated beads at 4°C. Beads were washed in lysis buffer three times, and then eluted with Elution Buffer (50mM NaH_2PO_4 pH 8, 50mM NaCl, 0.1% Tween-20, 10mM Biotin). 1% and 5% of the isolated elution was loaded onto an SDS-PAGE gel and stained with Coomassie Brilliant Blue, and BSA along with quantification of band intensity was used determine the protein concentration.

Electrophoresis Mobility Shift Assay

A 147 nucleotide double stranded telomeric substrate or an 87 nucleotide double stranded scramble substrate was labeled 3’ labeled with 32-P using Klenow. 0nM, .41nM, .82nM, 1.63nM, or 3.26nM of protein was incubated with .2nM substrate for 10 minutes on ice in EMSA buffer (250mM HEPES-KOH pH 8.0, 0.5mM DTT, 250ng/µl β-Casein, 40mM NaCl). Reactions were then loaded onto a 0.7% Agarose Gel in TBE and run at 100 volts for 45 minutes. Gels were fixed in 20% Methanol 10% Acetic Acid, dried, and exposed on a phosphoimager cassette which was then visualized on a Typhoon scanner.

In vitro Transcription and Translation

In vitro transcription and translation were done in a one-step reaction using Promega's TnT® Quick Coupled Transcription/Translation System (L1170). Briefly, a master mix containing rabbit reticulocyte lysate with RNA polymerase, amino acids and RNAse inhibitor were incubated with 1µg of plasmid DNA at 30°C for 90 minutes. Luciferase was used as a control transcript. Results were analyzed on an SDS-Page gel and protein expression was confirmed by Western Blotting for TRF2.

Laddering Assay

The TRF2 laddering assay was performed as follows: 0.5 – 16 µl of in vitro translated protein was incubated at room temperature for 20 minutes with 0.25 – 1 ng of pTH12 probe. pTH12 contains 12 repeats of the telomeric TTAGGG template and was labeled with P32. The reactions were performed in a 20 µl volume with 4% glycerol, 500 ng E. coli DNA, 15mM Tris-HCl pH 7.5, 50ng β-casein, and brought up to a final volume using control in vitro translation mix to maintain protein concentration. The reactions were run on a 0.75% agarose gel in 0.1X TBE and run at 200V for 30 minutes. The gel was then dried onto Whatmann paper, and imaged on a Phosphoimager cassette and scanned on a GE Typhoon.

STORM Imaging

2D STORM images were acquired on a custom setup based on a Nikon TiU inverted microscope similar to that used in a previous study (Doksani et al. 2013).

The Alexa647 labels on the PNA probes were excited with a 647-nm laser (MPB Photonics) with a power of 60 mW at the back port of the microscope. In addition, a 405-nm coherent cube laser with a maximum power of 1 mW at the back port of the microscope was used, and the power of the 405-nm laser was ramped during the experiment to maintain an approximately constant density of activated dye molecules (Folling et al. 2008; Heilemann et al. 2008; Dempsey et al. 2011). To improve reproducibility, the same 405 laser power ramp was used for each pair of experiment and control (+Cre and −Cre, respectively) samples. The fluorescence from the activated Alexa647 dye molecules was imaged onto an Andor 897 EMCCD camera using a 100× 1.45 NA Nikon plan apo lambda objective. Using this objective, the image pixel size was 160 nm, and the field of view was 41 μm× 41 μm. The microscope was controlled with custom software written in Python (https://github.com/ZhuangLab/storm-control). Imaging was performed with an imaging buffer containing 100 mM mercaptoethylamine (MEA) as the thiol group to promote photo-switching. The imaging buffer also contained 100 mM Tris (pH 8) and 50 mM NaCl with an oxygen-scavenging system consisting of 5% (w/v) glucose, 300 μg/mL glucose oxidase, and 40 μg/mL catalase.

Localizations were identified in STORM movies with the 3DDAOSTORM software package (“STORM analysis,” n.d.) (Babcock et al. 2012) and were rendered as Gaussian peaks with σ = 20 nm in the STORM images presented in the figures. We used image correlation between images taken at different times during STORM image acquisition to correct for sample drift (Bates et al. 2007). To test the drift correction quality by the image correlation approach, we also added

fiducial markers (fluorescent beads) to the samples to correct for drift and found that the two drift correction approaches generated identical results.

Clusters of localizations that represent telomeres were identified using the DBSCAN algorithm (Ester et al. 1996). Eps values from 20 to 80 nm and a fixed Minpts of 10 were used for DBSCAN analysis. In addition, only clusters with the localization number above a threshold value was used for further analysis. The eps value and the localization number threshold value were adjusted based on the telomere label density and the density of background localizations to produce optimal cluster identification as judged by visual inspection, but the same settings for these two values were used for the paired experiment and control (+Cre and −Cre, respectively) samples. We obtained an accompanying conventional image for each STORM image, and the majority of the clusters identified in the STORM images using the above approach has a corresponding spot in the conventional image, and the majority of spots identified in the conventional image has a corresponding identified STORM cluster. For each localization cluster identified this way in the STORM image, we computed the Rg and convex hull volume values to quantify the size of the telomere in 3D. For convex hull volume calculation, we first determined the convex hull area of each telomere imaged in 2D and then determined the convex hull volume by raising the convex hull area to the 3/2 power.

As an alternative approach, we additionally screened the clusters of localizations identified in the STORM image using the accompanying conventional image by selecting only STORM clusters that had a corresponding spot in the

conventional image identified by a spot-finding program. The results did not change substantially after such screening: The changes in the average Rg values upon TRF1 deletion measured for the seven independent experiments became −7%, 9%, 12%, 11%, 30%, 4%, and 1% after screening; the changes in the average Rg values upon TRF2 deletion (in Lig4-negative cells) measured for the three independent experiments became 14%, 18%, and 45% after screening; the changes in the average Rg values upon TRF1/TRF2 codeletion measured for the five independent experiments became 16%, 78%, 7%, 28%, and 47% after screening. We note that, although this alternative screening approach could help remove nontelomere clusters from the STORM images, it could also remove some real telomere foci that did not appear as sufficiently bright spots in the conventional image and hence may not necessarily provide a more accurate measurement compared with the approach of using STORM images alone to identify telomere clusters.

Dual color STORM imaging was performed as described above. Briefly, both [TTAGGG]3 and $[CCCTAA]_3$ probes were mixed together and added to coverslips for hybridization at 80 °C for 10 minutes in equal concentrations, ranging from 1:500 to 1:10,000.

Simulations of telomeres with predefined volume expansion

Simulations were performed with the OpenMM library (Eastman et al. 2013). The DNA was modeled using freely jointed chains with a harmonic potential to maintain the segment length and a short-range repulsive force to prevent chain

intersection. After initial energy minimization, each simulation was run for 20,000 steps to generate a random chain conformation. The simulations were repeated 1000 times for each chain length to generate a library of random chain conformations.

The above generated chain conformations were used to generate a simulated list of localizations similar to the output from the real STORM experiments. This was done by first placing the centroid positions of the chains at uniform random X, Y, and Z locations in a 41-μm × 41-μm× 2-μm box. Next, a list of localizations was created from these chains by using the position of each segment in the chain as the location of a localization. Localizations that had a Z-value that was too large (>500 nm) or too small (less than −500 nm) were discarded to roughly simulate the experimentally probed Z range. To simulate the effect of a change in polymer size arising from decompaction, the chain segment lengths were multiplied by a fixed value. For example, the chain segment lengths were multiplied by 1.2, 1.4, 1.7, and 2.0 to model the effect of a 20%, 40%, 70%, and 100% increase in Rg (representing a 1.7-fold, 2.7-fold, fivefold, and eightfold increase in volume, respectively). The predecompaction Rg was set to match the experimentally measured Rg values under the −Cre condition. An average of 180 chain segments (corresponding to 180 localizations per telomeric cluster) was used for the simulation. To simulate background localizations, 35,000 localizations were added at random uniformly distributed locations in the STORM image. This corresponds to an average density of background localizations of 21 per square micrometer. The average values of 180 localizations per cluster and 21

background localizations per square micrometer are comparable with the experimentally measured values in the vast majority of our experiments. The simulated STORM images were then processed with the same DBSCAN analysis pipeline that was used to analyze the experimental STORM data to determine telomeric signal clusters.

Statistical Analysis

All statistical analysis was performed with GraphPad Prism software or Excel. The significance between means was determined by one-way or two-way ANOVA with Tukey post test for multiple comparisons and two tailed unpaired Student's t test when the means of two experimental conditions were compared. n.s. not significant, * $p > 0.05$, ** $p > 0.01$, *** $p > 0.001$.

OMX Imaging

Cells were spread for t-loop imaging as previously described (*138*). Briefly, cells were lysed in fibroblast lysis buffer (12.5 mM Tris, pH 7.4, 5 mM KCl, 0.1 mM spermine, 0.25 mM spermidine and 175 mM sucrose, supplemented with protease-inhibitor cocktail (Roche)), and nuclei were spun down at 1,000g for 5 min and washed in nuclei wash buffer (10 mM Tris-HCl, pH 7.4, 15 mM NaCl, 60 mM KCl, 5 mM EDTA and 300 mM sucrose). Nuclei were cross-linked three times with 100 µg/ml trioxsalen under 365-nm light, then washed in nuclei wash buffer. Subsequently, nuclei were lysed with spreading buffer (10 mM Tris-HCl, pH 7.4, 10 mM EDTA, 0.05% SDS and 1 M NaCl), and the DNA was cytospun onto

coverslips. The standard FISH procedure was used to hybridize a triple-TTAGGG-Cy3 (TelG-Cy3) PNA probe, and DNA spreads were then imaged on a GE OMX V4 microscope. At least 100 countable molecules were scored per sample per experiment. Countable molecules were split into two categories: linear (continuous straight molecules that did not have any branches and were uniformly dense) and loops (molecules with a continuous, hollow loop at one end). Molecules that could not be classified as either linear or loops were discarded from the analysis.

ATAC-seq

ATAC-seq (Buenrostro et al. 2013) was performed by the Memorial Sloan Kettering Cancer Center Epigenetics Center Facility using standard protocols. Briefly, 50,000 cells were harvested and resuspended in PBS. Next, cells were lysed to prepare nuclei. Next, the transposition reaction was carried out for 30 min at 37°C using Tn5 transposon and TD buffer. Finally, the PCR reaction was performed using barcoded primers and the library was Illumina sequenced.

Cell Cycle Analysis

MEFs were treated with either 9μM RO-3306 for 12 hours and released into Nocadozole for 2 hours, with 40μM Lovastatin for 36 hours, with 40μM Lovastatin for 36 hours and released into 400μM Mevalonic Acid for 16 hours, or a DMSO control. Flow cytometry was conducted on an Accuri C6 and cells were prepared as follows: 1 hour prior to harvest cells were treated with 10μM BrdU. 500,000 cells were then harvested and fixed overnight in -20°C 70% EtOH. Cells were denatured

and permeabilized with 2N HCl / 0.5% Triton X-100 for 30 minutes, then neutralized with 0.1M Sodium Borate. Cells were blocked with 0.5% BSA / 0.5% Tween-20 in PBS and incubated with a FITC-conjugated anti-BrdU antibody (BD Biosciences #347583), then resuspended in 2mM EDTA, 0.5 mg / mL RNAse A, and 5µg / mL Propidium Iodine. Accuri C6 software and FlowJo were used for analysis.

Expansion Microscopy

Expansion microscopy was performed as described (*468*, *479*). Briefly, cells were grown on 8-well glass dishes and fixed with 2% PFA and dehydrated with ethanol. DNA was then FISH labeled with probes containing an acrylate group by either the Zhuang laboratory or the Boyden laboratory. Labeling with fluorescent PNA probes was done as described in the regular FISH protocol, either before expansion, during gelation, or at the conclusion of expansion. Optimal conditions were to label the sample after the conclusion of expansion microscopy. Next, the sample was gel cast inside the well with 0.1% PAA, protease treated, and expanded using distilled water. If iterative expansion microscopy was performed, the gel was removed and re-embeded and expanded again. The gel was then cut into 100 µm sections using a cryo-microtome and hybridized with fluorescent probe overnight. Next it was set into a stabilizing gel to prepare it for regular fluorescence or STORM imaging. Single expansion lead to a 4x expansion, and iterative expansion to 16x expansion. Due to the mesh structure, we estimate resolution at 5 nm.

Chromatin Immunoprecipitation

ChIP was performed as previously described (*247*). Cells were grown in 15 cm dishes to 80% confluence and fixed with 1% Formaldehyde in PBS for 30 minutes at room temperature, which was then quenched with 1.5M Glycine. Cells were then scraped into PBS for harvest, lysed for 15 minutes ChIP lysis (1% SDS, 10 mM EDTA pH 8.0, 50 mM Tris-HCl pH 8.0 with Roche protease inhbitor) and sonicated for 15 minutes to shear the chromatin. 1 million cells were used per IP with 20 µl TRF1 (1448), TRF2 (1254), Rap1 (1252), or TIN2 (1446) antibodies at 4 °C overnight. Magnetic beads (Cell signaling, #9006) were added for 30 minutes at 4 °C. Beads were then sequentially washed with Buffer A (0.1% SDS, 1% Triton X-100, 2 mM EDTA pH 8.0, 20 mM Tris-HCl pH 8.0, 150mM NaCl), Buffer B (0.1% SDS, 1% Triton X-100, 2 mM EDTA pH 8.0, 20 mM Tris-HCl pH 8.0, 500mM NaCl), Buffer C (0.25 M LiCl, 1% NP-40, 1% Na-Deoxycholate, 1 mM EDTA pH 8.0, 10 mM Tris-HCl pH 8.0) and TE. Beads were then resuspended in 1% SDS/0.1 M NaHCO3 to elute, and 20 µl of 5M NaCl was added to the supernatant and incubated at 65 °C overnight to reverse the crosslink. The beads were then subsequently treated with 20 µg RNAse A for 30 minutes and Proteinase K for 1 hour at 37 °C. Samples were then phenol chloroform extracted, run through a dot blotting apparatus, and hybridized to a 32-P labeled $[CCCTAA]_3$ probe.

Telomeric Restriction Fragment Analysis

1 million cells were harvested per condition, resuspended in PBS, and mixed 1:1 with 2% agarose to create plugs. The plugs were then digested overnight at 50°C with 1 mg/ml Proteinase K (Roche) in 10 mM Tris HCl pH 8.0,

250 mM EDTA, 0.2% sodium deoxycholate and 1% sodium lauryl sarcosine. Next, plugs were washed 4 times with TE for 1 hour each at 37 °C, washed for 1 hour with NEB Cutsmart Buffer at 37 °C, and then incubated with 60 units of MboI and AluI overnight at 37 °C in fresh Cutsmart Buffer. Plugs were then washed with TE and equilibrated with 0.5X TBE and loaded into a 1% Agarose 0.5X TBE gel and run on a CHEF pulse field gel electrophoresis machine. The gel was then dried and prehybridized with Church Mix for 1 hour at 50 °C, and then hybridized overnight in Church Mix with 32-P labeled $[CCCTAA]_4$ probe. The membrane was then washed 2x with 4X SSC 0.1% SDS at 55 °C and 2x with 4X SSC at 55 °C and exposed on a phosphocassette overnight. For denaturing analysis, the membrane was denatured in 0.5 M NaOH, 1.5 M NaCl for 30 min, neutralized with 0.5 M Tris HCl pH 7.5, 3 M NaCl. Prehybridization and hybridization steps were then repeated as described above.

2D Gel Analysis

1 million cells were harvested per condition, and genomic DNA was extracted by lysing cells in tail lysis buffer (10mM Tris pH 8.0, 100mM NaCl, 10mM EDTA, 0.5% SDS, 1 mg/mL Proteinase K) overnight at 55 °C, and then phenol-chloroform and ethanol precipitating the DNA. The DNA was then digested with MboI and AluI as described above. 7–12 mg of MboI/AluI-digested genomic DNA was resolved on a 0.4% agarose/1× TBE gel at <1 V/cm for 24 h. The first-dimension gel was stained with 0.3 mg/ml ethidium bromide, and bands were excised and placed 90° relative to the direction of electrophoresis in a 1.1%

agarose/1× TBE second dimension gel supplemented with 0.3 mg/ml ethidium bromide. Electrophoresis was performed for 4 h at 5 V/cm in a cold room. The gel was then hybridized and imaged as described above.

Chapter 11

References

1. C. W. Greider, E. H. Blackburn, Identification of a specific telomere terminal transferase activity in tetrahymena extracts. *Cell*. **43**, 405–413 (1985).
2. C. W. Greider, E. H. Blackburn, The telomere terminal transferase of tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity. *Cell*. **51**, 887–898 (1987).
3. C. W. Greider, E. H. Blackburn, A telomeric sequence in the RNA of Tetrahymena telomerase required for telomere repeat synthesis. *Nature*. **337**, 331–7 (1989).
4. V. Lundblad, J. W. Szostak, A mutant with a defect in telomere elongation leads to senescence in yeast. *Cell*. **57**, 633–643 (1989).
5. M. Z. Levy, R. C. Allsopp, A. B. Futcher, C. W. Greider, C. B. Harley, Telomere end-replication problem and cell aging. *J. Mol. Biol.* (1992), doi:10.1016/0022-2836(92)90096-3.
6. R. C. Allsopp *et al.*, Telomere Shortening Is Associated with Cell Division in Vitro and in Vivo. *Exp. Cell Res.* (1995), doi:10.1006/excr.1995.1306.
7. M. van Overbeek, T. de Lange, Apollo, an Artemis-related nuclease, interacts with TRF2 and protects human telomeres in S phase. *Curr. Biol.* **16**, 1295–302 (2006).
8. C. Lenain *et al.*, The Apollo 5’ exonuclease functions together with TRF2 to protect telomeres from DNA repair. *Curr. Biol.* **16**, 1303–10 (2006).
9. P. Wu, M. van Overbeek, S. Rooney, T. de Lange, Apollo Contributes to G Overhang Maintenance and Protects Leading-End Telomeres. *Mol. Cell*.

39, 606–617 (2010).

10. Y. C. Lam *et al.*, SNMIB/Apollo protects leading-strand telomeres against NHEJ-mediated repair. *EMBO J.* (2010), doi:10.1038/emboj.2010.58.
11. P. Wu, H. Takai, T. De Lange, Telomeric 3′ overhangs derive from resection by Exo1 and apollo and fill-in by POT1b-associated CST. *Cell*. **150**, 39–52 (2012).
12. V. L. Makarov, Y. Hirose, J. P. Langmore, Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. *Cell* (1997), doi:10.1016/S0092-8674(00)81908-X.
13. L. Hayflick, The limited in vitro lifetime of human diploid cell strains. *Exp. Cell Res.* (1965), doi:10.1016/0014-4827(65)90211-9.
14. T. H. D. Nguyen *et al.*, Cryo-EM structure of substrate-bound human telomerase holoenzyme. *Nature* (2018), doi:10.1038/s41586-018-0062-x.
15. J. Jiang *et al.*, Structure of Telomerase with Telomeric DNA. *Cell* (2018), doi:10.1016/j.cell.2018.04.038.
16. N. W. Kim *et al.*, Specific association of human telomerase activity with immortal cells and cancer. *Science (80-.).* (1994), doi:10.1126/science.7605428.
17. A. G. Bodnar *et al.*, Extension of life-span by introduction of telomerase into normal human cells. *Science (80-.).* (1998), doi:10.1126/science.279.5349.349.
18. W. C. Hahn *et al.*, Inhibition of telomerase limits the growth of human cancer cells. *Nat. Med.* (1999), doi:10.1038/13495.

19. W. C. Hahn *et al.*, Creation of human tumour cells with defined genetic elements. *Nature* (1999), doi:10.1038/22780.

20. B. Elenbaas *et al.*, Human breast cancer cells generated by oncogenic transformation of primary mammary epithelial cells. *Genes Dev.* (2001), doi:10.1101/gad.828901.

21. K. Damm *et al.*, A highly selective telomerase inhibitor limiting human cancer cell proliferation. *EMBO J.* (2001), doi:10.1093/emboj/20.24.6958.

22. J. W. Shay, W. E. Wright, Role of telomeres and telomerase in cancer. *Semin. Cancer Biol.* (2011), , doi:10.1016/j.semcancer.2011.10.001.

23. J. D. Henson, A. A. Neumann, T. R. Yeager, R. R. Reddel, Alternative lengthening of telomeres in mammalian cells. *Oncogene* (2002), , doi:10.1038/sj/onc/1205058.

24. T. M. Bryan, A. Englezou, J. Gupta, S. Bacchetti, R. R. Reddel, Telomere elongation in immortal human cells without detectable telomerase activity. *EMBO J.* (1995).

25. A. J. Cesare, R. R. Reddel, Alternative lengthening of telomeres: Models, mechanisms and implications. *Nat. Rev. Genet.* (2010), , doi:10.1038/nrg2763.

26. M. A. Dunham, A. A. Neumann, C. L. Fasching, R. R. Reddel, Telomere maintenance by recombination in human cells. *Nat. Genet.* (2000), doi:10.1038/82586.

27. V. Lundblad, Telomere maintenance without telomerase. *Oncogene* (2002), , doi:10.1038/sj/onc/1205079.

28. V. Lundblad, E. H. Blackburn, An alternative pathway for yeast telomere maintenance rescues est1- senescence. *Cell* (1993), doi:10.1016/0092-8674(93)90234-H.

29. C. M. Heaphy *et al.*, Altered telomeres in tumors with ATRX and DAXX mutations. *Science (80-.).* (2011), , doi:10.1126/science.1207313.

30. Y. Jiao *et al.*, DAXX/ATRX, MEN1, and mTOR pathway genes are frequently altered in pancreatic neuroendocrine tumors. *Science (80-.).* (2011), doi:10.1126/science.1200609.

31. C. A. Lovejoy *et al.*, Loss of ATRX, genome instability, and an altered DNA damage response are hallmarks of the alternative lengthening of Telomeres pathway. *PLoS Genet.* (2012), doi:10.1371/journal.pgen.1002772.

32. J. D. Henson *et al.*, DNA C-circles are specific and quantifiable markers of alternative- lengthening-of-telomeres activity. *Nat. Biotechnol.* (2009), doi:10.1038/nbt.1587.

33. T. R. Yeager *et al.*, Telomerase-negative immortalized human cells contain a novel type of promyelocytic leukemia (PML) body. *Cancer Res.* (1999).

34. R. L. Dilley, R. A. Greenberg, ALTernative Telomere Maintenance and Cancer. *Trends in Cancer* (2015), , doi:10.1016/j.trecan.2015.07.007.

35. X. Feng, S. J. Hsu, C. Kasbek, M. Chaiken, C. M. Price, CTC1-mediated C-strand fill-in is an essential step in telomere length maintenance. *Nucleic Acids Res.* (2017), doi:10.1093/nar/gkx125.

36. P. Gu *et al.*, CTC1 deletion results in defective telomere replication, leading

to catastrophic telomere loss and stem cell exhaustion. *EMBO J.* (2012), doi:10.1038/emboj.2012.96.

37. Y. Miyake *et al.*, RPA-like Mammalian Ctc1-Stn1-Ten1 Complex Binds to Single-Stranded DNA and Protects Telomeres Independently of the Pot1 Pathway. *Mol. Cell* (2009), doi:10.1016/j.molcel.2009.08.009.

38. Y. V. Surovtseva *et al.*, Conserved Telomere Maintenance Component 1 Interacts with STN1 and Maintains Chromosome Ends in Higher Eukaryotes. *Mol. Cell* (2009), doi:10.1016/j.molcel.2009.09.017.

39. F. Wang *et al.*, Human CST Has Independent Functions during Telomere Duplex Replication and C-Strand Fill-In. *Cell Rep.* (2012), doi:10.1016/j.celrep.2012.10.007.

40. D. E. Casteel *et al.*, A DNA polymerase-α·primase cofactor with homology to replication protein A-32 regulates DNA replication in mammalian cells. *J. Biol. Chem.* (2009), doi:10.1074/jbc.M807593200.

41. X. Dai *et al.*, Molecular steps of G-overhang generation at human telomeres and its function in chromosome end protection. *EMBO J.* (2010), doi:10.1038/emboj.2010.156.

42. L. Y. Chen, S. Redon, J. Lingner, The human CST complex is a terminator of telomerase activity. *Nature* (2012), doi:10.1038/nature11269.

43. M. Nakamura, A. Nabetani, T. Mizuno, F. Hanaoka, F. Ishikawa, Alterations of DNA and Chromatin Structures at Telomeres and Genetic Instability in Mouse Cells Defective in DNA Polymerase . *Mol. Cell. Biol.* (2005), doi:10.1128/mcb.25.24.11073-11088.2005.

44. J. A. Stewart *et al.*, Human CST promotes telomere duplex replication and general replication restart after fork stalling. *EMBO J.* (2012), doi:10.1038/emboj.2012.215.

45. L. Y. Chen, J. Lingner, CST for the grand finale of telomere replication. *Nucl. (United States)* (2013), doi:10.4161/nucl.25701.

46. R. Gupta *et al.*, DNA Repair Network Analysis Reveals Shieldin as a Key Regulator of NHEJ and PARP Inhibitor Sensitivity. *Cell* (2018), doi:10.1016/j.cell.2018.03.050.

47. H. Dev *et al.*, Shieldin complex promotes DNA end-joining and counters homologous recombination in BRCA1-null cells. *Nat. Cell Biol.* (2018), doi:10.1038/s41556-018-0140-1.

48. S. M. Noordermeer *et al.*, The shieldin complex mediates 53BP1-dependent DNA repair. *Nature* (2018), , doi:10.1038/s41586-018-0340-7.

49. Z. Mirman *et al.*, 53BP1–RIF1–shieldin counteracts DSB resection through CST- and Polα-dependent fill-in. *Nature* (2018), , doi:10.1038/s41586-018-0324-7.

50. D. Liu *et al.*, PTOP interacts with POT1 and regulates its localization to telomeres. *Nat. Cell Biol.* (2004), doi:10.1038/ncb1142.

51. B. R. Houghtaling, L. Cuttonaro, W. Chang, S. Smith, A dynamic molecular link between the telomere length regulator TRF1 and the chromosome end protector TRF2. *Curr. Biol.* (2004), doi:10.1016/j.cub.2004.08.052.

52. J. Z. S. Ye *et al.*, POT1-interaction protein PIP1: A telomere length regulator that recruits POT1 to the TIN2/TRF1 complex. *Genes Dev.*

(2004), doi:10.1101/gad.1215404.

53. F. Erdel *et al.*, Telomere Recognition and Assembly Mechanism of Mammalian Shelterin. *Cell Rep.* (2017), doi:10.1016/j.celrep.2016.12.005.

54. A. Bianchi, S. Smith, L. Chong, P. Elias, T. De Lange, TRF1 is a dimer and bends telomeric DNA. *EMBO J.* (1997), doi:10.1093/emboj/16.7.1785.

55. A. Bianchi *et al.*, TRF1 binds a bipartite telomeric site with extreme spatial flexibility. *EMBO J.* (1999), doi:10.1093/emboj/18.20.5735.

56. D. Broccoli, A. Smogorzewska, L. Chong, T. de Lange, Human telomeres contain two distinct Myb-related proteins, TRF1 and TRF2. *Nat. Genet.* **17**, 231–5 (1997).

57. L. Chong *et al.*, A human telomeric protein. *Science (80-.).* (1995), doi:10.1126/science.270.5242.1663.

58. Z. Zhong, L. Shiue, S. Kaplan, T. de Lange, A mammalian factor that binds telomeric TTAGGG repeats in vitro. *Mol. Cell. Biol.* (1992).

59. R. C. Wang, A. Smogorzewska, T. de Lange, Homologous recombination generates T-loop-sized deletions at human telomeres. *Cell*. **119**, 355–68 (2004).

60. I. Schmutz, L. Timashev, W. Xie, D. J. Patel, T. De Lange, TRF2 binds branched DNA to safeguard telomere integrity. *Nat. Struct. Mol. Biol.* **24**, 734–742 (2017).

61. D. Benarroch-Popivker *et al.*, TRF2-Mediated Control of Telomere DNA Topology as a Mechanism for Chromosome-End Protection. *Mol. Cell*, 274–286 (2015).

62. S. Amiard *et al.*, A topological mechanism for TRF2-enhanced strand invasion. *Nat. Struct. Mol. Biol.* (2007), doi:10.1038/nsmb1192.

63. P. Baumann, T. R. Cech, Pot1, the putative telomere end-binding protein in fission yeast and humans. *Science*. **292**, 1171–5 (2001).

64. M. Lei, E. R. Podell, T. R. Cech, Structure of human POT1 bound to telomeric single-stranded DNA provides a model for chromosome end-protection. *Nat. Struct. Mol. Biol.* (2004), doi:10.1038/nsmb867.

65. N. Fouché *et al.*, The basic domain of TRF2 directs binding to DNA junctions irrespective of the presence of TTAGGG repeats. *J. Biol. Chem.* **281**, 37486–37495 (2006).

66. D. Loayza, H. Parsons, J. Donigian, K. Hoke, T. De Lange, DNA binding features of human POT1: A nonamer 5′-TAGGGTTAG-3′ minimal binding site, sequence specificity, and internal binding to multimeric sites. *J. Biol. Chem.* (2004), doi:10.1074/jbc.M312309200.

67. S. Hanaoka, A. Nagadoi, Y. Nishimura, Comparison between TRF2 and TRF1 of their telomeric DNA-bound structures and DNA-binding activities. *Protein Sci.* (2009), doi:10.1110/ps.04983705.

68. L. Fairall, L. Chapman, H. Moss, T. De Lange, D. Rhodes, Structure of the TRFH dimerization domain of the human telomeric proteins TRF1 and TRF2. *Mol. Cell* (2001), doi:10.1016/S1097-2765(01)00321-5.

69. C. Chen *et al.*, Structural insights into POT1-TPP1 interaction and POT1 C-terminal mutations in human cancer. *Nat. Commun.* (2017), doi:10.1038/ncomms14929.

70. C. Rice *et al.*, Structural and functional analysis of the human POT1-TPP1 telomeric complex. *Nat. Commun.* (2017), doi:10.1038/ncomms14928.

71. Y. Chen *et al.*, A conserved motif within RAP1 has diversified roles in telomere protection and regulation in different organisms. *Nat. Struct. Mol. Biol.* (2011), doi:10.1038/nsmb.1974.

72. C. Hu *et al.*, Structural and functional analyses of the mammalian TIN2-TPP1-TRF2 telomeric complex. *Cell Res.* (2017), doi:10.1038/cr.2017.144.

73. Y. Chen *et al.*, A shared docking motif in TRF1 and TRF2 used for differential recruitment of telomeric proteins. *Science*. **319**, 1092–6 (2008).

74. K. K. Takai, S. Hooper, S. Blackwood, R. Gandhi, T. de Lange, In vivo stoichiometry of shelterin components. *J. Biol. Chem.* (2010), doi:10.1074/jbc.M109.038026.

75. C. J. Lim, A. J. Zaug, H. J. Kim, T. R. Cech, Reconstitution of human shelterin complexes reveals unexpected stoichiometry and dual pathways to enhance telomerase processivity. *Nat. Commun.* (2017), doi:10.1038/s41467-017-01313-w.

76. D. Liu, M. S. O'Connor, J. Qin, Z. Songyang, Telosome, a mammalian telomere-associated complex formed by multiple telomeric proteins. *J. Biol. Chem.* (2004), doi:10.1074/jbc.M409293200.

77. J. Z. S. Ye *et al.*, TIN2 binds TRF1 and TRF2 simultaneously and stabilizes the TRF2 complex on telomeres. *J. Biol. Chem.* (2004), doi:10.1074/jbc.M409047200.

78. E. L. Denchi, T. de Lange, Protection of telomeres through independent

control of ATM and ATR by TRF2 and POT1. *Nature*. **448**, 1068–71 (2007).

79. P. Wu, T. de Lange, No overt nucleosome eviction at deprotected telomeres. *Mol. Cell. Biol.* **28**, 5724–35 (2008).

80. A. M. Baker, Q. Fu, W. Hayward, S. M. Lindsay, T. M. Fletcher, The Myb/SANT domain of the telomere-binding protein TRF2 alters chromatin structure. *Nucleic Acids Res.* **37**, 5019–5031 (2009).

81. A. M. Baker *et al.*, The telomere binding protein trf2 induces chromatin compaction. *PLoS One* (2011), doi:10.1371/journal.pone.0019124.

82. A. Sfeir, T. de Lange, Removal of shelterin reveals the telomere end-protection problem. *Science (80-.).* **336**, 593–7 (2012).

83. T. Uziel *et al.*, Requirement of the MRN complex for ATM activation by DNA damage. *EMBO J.* (2003), doi:10.1093/emboj/cdg541.

84. S. Matsuoka, M. Huang, S. J. Elledge, Linkage of ATM to cell cycle regulation by the Chk2 protein kinase. *Science (80-.).* (1998), doi:10.1126/science.282.5395.1893.

85. M. B. Kastan, J. Bartek, Cell-cycle checkpoints and cancer. *Nature* (2004), , doi:10.1038/nature03097.

86. G. B. Celli, T. de Lange, DNA processing is not required for ATM-mediated telomere damage response after TRF2 deletion. *Nat. Cell Biol.* **7**, 712–8 (2005).

87. F. A. Mallette *et al.*, RNF8- and RNF168-dependent degradation of KDM4A/JMJD2A triggers 53BP1 recruitment to DNA damage sites. *EMBO*

J. (2012), doi:10.1038/emboj.2012.47.

88. S. Burma, B. P. Chen, M. Murphy, A. Kurimasa, D. J. Chen, ATM Phosphorylates Histone H2AX in Response to DNA Double-strand Breaks. *J. Biol. Chem.* (2001), doi:10.1074/jbc.C100466200.

89. S. Bekker-Jensen, C. Lukas, F. Melander, J. Bartek, J. Lukas, Dynamic assembly and sustained retention of 53BP1 at the sites of DNA damage are controlled by Mdc1/NFBD1. *J. Cell Biol.* (2005), doi:10.1083/jcb.200503043.

90. K. K. Takai, T. Kibe, J. R. Donigian, D. Frescas, T. de Lange, Telomere protection by TPP1/POT1 requires tethering to TIN2. *Mol. Cell*. **44**, 647–59 (2011).

91. T. Davoli, E. L. Denchi, T. de Lange, Persistent Telomere Damage Induces Bypass of Mitosis and Tetraploidy. *Cell* (2010), doi:10.1016/j.cell.2010.01.031.

92. E. L. Denchi, G. Celli, T. De Lange, Hepatocytes with extensive telomere deprotection and fusion remain viable and regenerate liver mass through endoreduplication. *Genes Dev.* (2006), doi:10.1101/gad.1453606.

93. L. Zou, S. J. Elledge, Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. *Science (80-.).* (2003), doi:10.1126/science.1083430.

94. L. Zou, D. Cortez, S. J. Elledge, Regulation of ATR substrate selection by Rad17-dependent loading of Rad9 complexes onto chromatin. *Genes Dev.* (2002), doi:10.1101/gad.950302.

95. E. J. Brown, D. Baltimore, Essential and dispensable roles of ATR in cell cycle arrest and genome maintenance. *Genes Dev.* (2003), doi:10.1101/gad.1067403.

96. G. N. Parkinson, M. P. H. Lee, S. Neidle, Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature* (2002), doi:10.1038/nature755.

97. G. Biffi, D. Tannahill, J. McCafferty, S. Balasubramanian, Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.* (2013), doi:10.1038/nchem.1548.

98. M. Zimmermann, T. Kibe, S. Kabir, T. de Lange, TRF1 negotiates TTAGGG repeat associated replication problems by recruiting the BLM helicase and the TPP1/POT1 repressor of ATR signaling. *Genes Dev.* **28**, 2477–2491 (2014).

99. A. Rizzo *et al.*, Stabilization of quadruplex DNA perturbs telomere replication leading to the activation of an ATR-dependent ATM signaling pathway. *Nucleic Acids Res.* (2009), doi:10.1093/nar/gkp582.

100. A. Sfeir *et al.*, Mammalian Telomeres Resemble Fragile Sites and Require TRF1 for Efficient Replication. *Cell*. **138**, 90–103 (2009).

101. K. W. Caldecott, S. Aoufouchi, P. Johnson, S. Shall, XRCC1 polypeptide interacts with DNA polymerase β and possibly poly (ADP-ribose) polymerase, and DNA ligase III is a novel molecular “nick-sensor” in vitro. *Nucleic Acids Res.* (1996), doi:10.1093/nar/24.22.4387.

102. V. Schreiber *et al.*, Poly(ADP-ribose) polymerase-2 (PARP-2) is required

for efficient base excision DNA repair in association with PARP-1 and XRCC1. *J. Biol. Chem.* (2002), doi:10.1074/jbc.M202390200.

103. A. A. E. Ali *et al.*, The zinc-finger domains of PARP1 cooperate to recognize DNA strand breaks. *Nat. Struct. Mol. Biol.* (2012), doi:10.1038/nsmb.2335.

104. M. F. Langelier, J. L. Planck, S. Roy, J. M. Pascal, Structural basis for DNA damage-dependent poly(ADP-ribosyl)ation by human PARP-1. *Science (80-.).* (2012), doi:10.1126/science.1216338.

105. M. Y. Kim, S. Mauro, N. Gévry, J. T. Lis, W. L. Kraus, NAD+-dependent modulation of chromatin structure and transcription by nucleosome binding properties of PARP-1. *Cell* (2004), doi:10.1016/j.cell.2004.11.002.

106. D. Ahel *et al.*, Poly(ADP-ribose)-dependent regulation of DNA repair by the chromatin remodeling enzyme ALC1. *Science (80-.).* (2009), doi:10.1126/science.1177321.

107. J. F. Haince *et al.*, PARP1-dependent kinetics of recruitment of MRE11 and NBS1 proteins to multiple DNA damage sites. *J. Biol. Chem.* (2008), doi:10.1074/jbc.M706734200.

108. H. Shirai *et al.*, PARG dysfunction enhances DNA double strand break formation in S-phase after alkylation DNA damage and augments different cell death pathways. *Cell Death Dis.* (2013), doi:10.1038/cddis.2013.133.

109. M. Wang *et al.*, PARP-1 and Ku compete for repair of DNA double strand breaks by distinct NHEJ pathways. *Nucleic Acids Res.* (2006), doi:10.1093/nar/gkl840.

110. S. Smith, I. Giriat, A. Schmitt, T. De Lange, Tankyrase, a poly(ADP-ribose) polymerase at human telomeres. *Science (80-.).* (1998), doi:10.1126/science.282.5393.1484.

111. A. Nabetani, F. Ishikawa, Unusual Telomeric DNAs in Human Telomerase-Negative Immortalized Cells. *Mol. Cell. Biol.* (2009), doi:10.1128/mcb.00603-08.

112. S. J. Boulton, S. P. Jackson, Components of the Ku-dependent non-homologous end-joining pathway are involved in telomeric length maintenance and telomeric silencing. *EMBO J.* (1998), doi:10.1093/emboj/17.6.1819.

113. Y. Ma, K. Schwarz, M. R. Lieber, The Artemis:DNA-PKcs endonuclease cleaves DNA loops, flaps, and gaps. *DNA Repair (Amst).* (2005), doi:10.1016/j.dnarep.2005.04.013.

114. S. Burma, B. P. C. Chen, D. J. Chen, Role of non-homologous end joining (NHEJ) in maintaining genomic integrity. *DNA Repair (Amst).* (2006), , doi:10.1016/j.dnarep.2006.05.026.

115. B. Van Steensel, A. Smogorzewska, T. De Lange, TRF2 protects human telomeres from end-to-end fusions. *Cell*. **92**, 401–413 (1998).

116. J. Maciejowski, Y. Li, N. Bosco, P. J. Campbell, T. De Lange, Chromothripsis and Kataegis Induced by Telomere Crisis. *Cell*. **163**, 1641–1654 (2015).

117. N. Arnoult *et al.*, Regulation of DNA repair pathway choice in S and G2 phases by the NHEJ inhibitor CYREN. *Nature* (2017),

doi:10.1038/nature24023.

118. N. Dimitrova, T. de Lange, Cell cycle-dependent role of MRN at dysfunctional telomeres: ATM signaling-dependent induction of nonhomologous end joining (NHEJ) in G1 and resection-mediated inhibition of NHEJ in G2. *Mol. Cell. Biol.* **29**, 5552–63 (2009).

119. X. D. Zhu *et al.*, ERCC1/XPF Removes the 3′ Overhang from Uncapped Telomeres and Represses Formation of Telomeric DNA-Containing Double Minute Chromosomes. *Mol. Cell* (2003), doi:10.1016/S1097-2765(03)00478-7.

120. F. Lottersberger, R. A. Karssemeijer, N. Dimitrova, T. De Lange, 53BP1 and the LINC Complex Promote Microtubule-Dependent DSB Mobility and DNA Repair. *Cell* (2015), doi:10.1016/j.cell.2015.09.057.

121. A. Sfeir, L. S. Symington, Microhomology-Mediated End Joining: A Back-up Survival Mechanism or Dedicated Pathway? *Trends Biochem. Sci.* (2015), , doi:10.1016/j.tibs.2015.08.006.

122. N. Bennardo, A. Cheng, N. Huang, J. M. Stark, Alternative-NHEJ is a mechanistically distinct pathway of mammalian chromosome break repair. *PLoS Genet.* (2008), doi:10.1371/journal.pgen.1000110.

123. M. Audebert, B. Salles, M. Weinfeld, P. Calsou, Involvement of polynucleotide kinase in a poly(ADP-ribose) polymerase-1-dependent DNA double-strand breaks rejoining pathway. *J. Mol. Biol.* (2006), doi:10.1016/j.jmb.2005.11.028.

124. P. A. Mateos-Gomez *et al.*, Mammalian polymerase θ promotes alternative

NHEJ and suppresses recombination. *Nature* (2015), doi:10.1038/nature14157.

125. Y. Doksani, T. de Lange, Telomere-Internal Double-Strand Breaks Are Repaired by Homologous Recombination and PARP1/Lig3-Dependent End-Joining. *Cell Rep.* (2016), doi:10.1016/j.celrep.2016.10.008.
126. R. B. Jensen, A. Carreira, S. C. Kowalczykowski, Purified human BRCA2 stimulates RAD51-mediated recombination. *Nature*. **467**, 678–683 (2010).
127. G. B. Celli, E. L. Denchi, T. de Lange, Ku70 stimulates fusion of dysfunctional telomeres yet protects chromosome ends from homologous recombination. *Nat. Cell Biol.* **8**, 885–90 (2006).
128. J. Karlseder, D. Broccoli, Y. Dai, S. Hardy, T. de Lange, p53- and ATM-dependent apoptosis induced by telomeres lacking TRF2. *Science*. **283**, 1321–1325 (1999).
129. T. Bilaud *et al.*, Telomeric localization of TRF2, a novel human telobox protein. *Nat. Genet.* (1997), , doi:10.1038/ng1097-236.
130. S. H. Yoshimura, H. Maruyama, F. Ishikawa, R. Ohki, K. Takeyasu, Molecular mechanisms of DNA end-loop formation by TRF2. *Genes to Cells* (2004), doi:10.1111/j.1356-9597.2004.00719.x.
131. K. Okamoto *et al.*, A two-step mechanism for TRF2-mediated chromosome-end protection. *Nature*. **494**, 502–5 (2013).
132. A. A. Poulet *et al.*, The N-terminal domains of TRF1 and TRF2 regulate their ability to condense telomeric DNA. *Nucleic Acids Res.* **40**, 2566–2576 (2012).

133. T. Gudjonsson *et al.*, TRIP12 and UBR5 suppress spreading of chromatin ubiquitylation at damaged chromosomes. *Cell* (2012), doi:10.1016/j.cell.2012.06.039.

134. R. Court, L. Chapman, L. Fairall, D. Rhodes, How the human telomeric proteins TRF1 and TRF2 recognize telomeric DNA: A view from high-resolution crystal structures. *EMBO Rep.* (2005), doi:10.1038/sj.embor.7400314.

135. A. Smogorzewska, J. Karlseder, H. Holtgreve-Grez, A. Jauch, T. De Lange, DNA ligase IV-dependent NHEJ of deprotected mammalian telomeres in G1 and G2. *Curr. Biol.* **12**, 1635–1644 (2002).

136. A. Ribes-Zamora, S. M. Indiviglio, I. Mihalek, C. L. Williams, A. A. Bertuch, TRF2 Interaction with Ku Heterotetramerization Interface Gives Insight into c-NHEJ Prevention at Human Telomeres. *Cell Rep.* (2013), doi:10.1016/j.celrep.2013.08.040.

137. J. D. Griffith *et al.*, Mammalian telomeres end in a large duplex loop. *Cell*. **97**, 503–14 (1999).

138. Y. Doksani, J. Y. Wu, T. de Lange, X. Zhuang, Super-resolution fluorescence imaging of telomeres reveals TRF2-dependent T-loop formation. *Cell*. **155**, 345–56 (2013).

139. Y. Doksani, T. de Lange, The role of double-strand break repair pathways at functional and dysfunctional telomeres. *Cold Spring Harb. Perspect. Biol.* (2014), doi:10.1101/cshperspect.a016576.

140. M. Zeraati *et al.*, I-motif DNA structures are formed in the nuclei of human

cells. *Nat. Chem.* (2018), doi:10.1038/s41557-018-0046-3.

141. R. M. Stansel, T. De Lange, J. D. Griffith, T-loop assembly in vitro involves binding of TRF2 near the 3′ telomeric overhang. *EMBO J.* **20**, 5532–5540 (2001).

142. J. Griffith, A. Bianchi, T. de Lange, TRF1 promotes parallel pairing of telomeric tracts in vitro. *J. Mol. Biol.* **278**, 79–88 (1998).

143. T. Nikitina, C. L. Woodcock, Closed chromatin loops at the ends of chromosomes. *J. Cell Biol.* (2004), doi:10.1083/jcb.200403118.

144. K. G. Murti, D. M. Prescott, Telomeres of polytene chromosomes in a ciliated protozoan terminate in duplex DNA loops. *Proc. Natl. Acad. Sci.* (2002), doi:10.1073/pnas.96.25.14436.

145. D. Van Ly *et al.*, Telomere Loop Dynamics in Chromosome End Protection. *Mol. Cell* (2018), doi:10.1016/j.molcel.2018.06.025.

146. M. Fried, D. M. Crothers, Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Res.* (1981), doi:10.1093/nar/9.23.6505.

147. A. M. Friedman, T. O. Fischmann, T. A. Steitz, Crystal structure of lac repressor core tetramer and its implications for DNA looping. *Science (80-.).* (1995), doi:10.1126/science.7792597.

148. J. N. Bandaria *et al.*, Shelterin Protects Chromosome Ends by Article Shelterin Protects Chromosome Ends by Compacting Telomeric Chromatin. *Cell*. **164**, 735–746 (2016).

149. H. L. Hsu *et al.*, Ku acts in a unique way at the mammalian telomere to

prevent end joining. *Genes Dev.* (2000), doi:10.1101/gad.844000.

150. H.-L. Hsu, D. Gilley, E. H. Blackburn, D. J. Chen, Ku is associated with the telomere in mammals. *Proc. Natl. Acad. Sci.* (2002), doi:10.1073/pnas.96.22.12454.

151. X. D. Zhu, B. Küster, M. Mann, J. H. J. Petrini, T. De Lange, Cell-cycle-regulated association of RAD50/MRE11/NBS1 with TRF2 and human telomeres. *Nat. Genet.* (2000), doi:10.1038/77139.

152. Y. Deng, X. Guo, D. O. Ferguson, S. Chang, Multiple roles for MRE11 at uncapped telomeres. *Nature* (2009), doi:10.1038/nature08196.

153. C. T. Carson *et al.*, The Mre11 complex is required for ATM activation and the G2/M checkpoint. *EMBO J.* (2003), doi:10.1093/emboj/cdg630.

154. J. H. Lee, T. T. Paull, Direct Activation of the ATM Protein Kinase by the Mre11/Rad50/Nbs1 Complex. *Science (80-.).* (2004), doi:10.1126/science.1091496.

155. J. H. Lee, T. T. Paull, ATM activation by DNA double-strand breaks through the Mre11-Rad50-Nbs1 complex. *Science (80-.).* (2005), doi:10.1126/science.1108297.

156. D. Simsek *et al.*, DNA ligase III promotes alternative nonhomologous end-joining during chromosomal translocation formation. *PLoS Genet.* (2011), doi:10.1371/journal.pgen.1002080.

157. W. Palm, D. Hockemeyer, T. Kibe, T. de Lange, Functional Dissection of Human and Mouse POT1 Proteins. *Mol. Cell. Biol.* (2009), doi:10.1128/mcb.01352-08.

158. A. Poulet *et al.*, TRF2 promotes, remodels and protects telomeric Holliday junctions. *EMBO J.* (2009), doi:10.1038/emboj.2009.11.

159. A. Saint-Léger *et al.*, The basic N-terminal domain of TRF2 limits recombination endonuclease action at human telomeres. *Cell Cycle* (2014), doi:10.4161/cc.29422.

160. H. Kim *et al.*, TRF2 functions as a protein hub and regulates telomere maintenance by recognizing specific peptide motifs. *Nat. Struct. Mol. Biol.* (2009), doi:10.1038/nsmb.1575.

161. P. Kaur *et al.*, Enhanced electrostatic force microscopy reveals higher-order DNA looping mediated by the telomeric protein TRF2. *Sci. Rep.* (2016), doi:10.1038/srep20513.

162. N. Ö. Arat, J. D. Griffith, Human Rap1 interacts directly with telomeric DNA and regulates TRF2 localization at the telomere. *J. Biol. Chem.* (2012), doi:10.1074/jbc.M112.415984.

163. R. E. Verdun, J. Karlseder, The DNA Damage Machinery and Homologous Recombination Pathway Act Consecutively to Protect Human Telomeres. *Cell* (2006), doi:10.1016/j.cell.2006.09.034.

164. T. Tsuzuki *et al.*, Targeted disruption of the Rad51 gene leads to lethality in embryonic mice. *Proc. Natl. Acad. Sci.* (1996), doi:10.1073/pnas.93.13.6236.

165. S. K. Sharan *et al.*, Embryonic lethality and radiation hypersensitivity mediated by Rad51 in mice lacking Brca2. *Nature* (1997), doi:10.1038/386804a0.

166. A. Shinohara, H. Ogawa, T. Ogawa, Rad51 protein involved in repair and recombination in S. cerevisiae is a RecA-like protein. *Cell*. **69**, 457–470 (1992).

167. P. Sung, Catalysis of ATP-dependent homologous DNA pairing and strand exchange by yeast RAD51 protein. *Science (80-.).* (1994), doi:10.1126/science.8066464.

168. P. Sung, L. Krejci, S. Van Komen, M. G. Sehorn, Rad51 Recombinase and Recombination Mediators. *J. Biol. Chem.* (2003), , doi:10.1074/jbc.R300027200.

169. F. Pâques, J. E. Haber, Multiple pathways of recombination induced by double-strand breaks in Saccharomyces cerevisiae. *Microbiol. Mol. Biol. Rev.* (1999).

170. J. C. Game, R. K. Mortimer, A genetic study of X-ray sensitive mutants in yeast. *Mutat. Res. - Fundam. Mol. Mech. Mutagen.* (1974), doi:10.1016/0027-5107(74)90176-6.

171. S. Le, J. K. Moore, J. E. Haber, C. W. Greider, RAD50 and RAD51 define two pathways that collaborate to maintain telomeres in the absence of telomerase. *Genetics* (1999).

172. S. Badie *et al.*, BRCA2 acts as a RAD51 loader to facilitate telomere replication and capping. *Nat. Struct. Mol. Biol.* (2010), doi:10.1038/nsmb.1943.

173. M. Takata *et al.*, Chromosome Instability and Defective Recombinational Repair in Knockout Mutants of the Five Rad51 Paralogs. *Mol. Cell. Biol.*

(2001), doi:10.1128/mcb.21.8.2858-2866.2001.

174. J. Y. Masson *et al.*, Identification and purification of two distinct complexes containing the five RAD51 paralogs. *Genes Dev.* (2001), doi:10.1101/gad.947001.

175. M. Tarsounas *et al.*, Telomere maintenance requires the RAD51D recombination/repair protein. *Cell* (2004), doi:10.1016/S0092-8674(04)00337-X.

176. M. D. Wyatt *et al.*, Thiopurine-induced mitotic catastrophe in Rad51d-deficient mammalian cells. *Environ. Mol. Mutagen.* (2018), doi:10.1002/em.22138.

177. D. L. Pittman, J. C. Schimenti, Midgestation lethality in mice deficient for the RecA-related gene, Rad51d/Rad51l3. *Genesis* (2000), doi:10.1002/(SICI)1526-968X(200003)26:3<167::AID-GENE1>3.0.CO;2-M.

178. H. Yokoyama *et al.*, Preferential binding to branched DNA strands and strand-annealing activity of the human Rad51B, Rad51C, Rad51D and Xrcc2 protein complex. *Nucleic Acids Res.* (2004), doi:10.1093/nar/gkh578.

179. H. Serra, O. Da Ines, F. Degroote, M. E. Gallego, C. I. White, Roles of XRCC2, RAD51B and RAD51D in RAD51-Independent SSA Recombination. *PLoS Genet.* (2013), doi:10.1371/journal.pgen.1003971.

180. J.-B. B. Vannier, V. Pavicic-Kaltenbrunner, M. I. R. Petalcorin, H. Ding, S. J. Boulton, RTEL1 dismantles T loops and counteracts telomeric G4-DNA to maintain telomere integrity. *Cell*. **149**, 795–806 (2012).

181. L. J. Barber *et al.*, RTEL1 Maintains Genomic Stability by Suppressing

Homologous Recombination. *Cell* (2008), doi:10.1016/j.cell.2008.08.016.

182. J. B. Vannier *et al.*, RTEL1 is a replisome-associated helicase that promotes telomere and genome-wide replication. *Science (80-.).* (2013), doi:10.1126/science.1241779.

183. Z. Deng *et al.*, Inherited mutations in the helicase RTEL1 cause telomere dysfunction and Hoyeraal–Hreidarsson syndrome. *Proc. Natl. Acad. Sci.* (2013), doi:10.1073/pnas.1300600110.

184. E. J. Uringa, J. L. Youds, K. Lisaingo, P. M. Lansdorp, S. J. Boulton, RTEL1: An essential helicase for telomere maintenance and the regulation of homologous recombination. *Nucleic Acids Res.* **39**, 1647–1655 (2011).

185. G. Sarek, J. B. Vannier, S. Panier, H. J. Petrini, S. J. Boulton, TRF2 recruits RTEL1 to telomeres in s phase to promote t-loop unwinding. *Mol. Cell*. **57**, 622–635 (2015).

186. M. T. Hayashi, A. J. Cesare, J. A. J. Fitzpatrick, E. Lazzerini-Denchi, J. Karlseder, A telomere-dependent DNA damage checkpoint induced by prolonged mitotic arrest. *Nat. Struct. Mol. Biol.* **19**, 387–394 (2012).

187. M. T. Hayashi, A. J. Cesare, T. Rivera, J. Karlseder, Cell death during crisis is mediated by mitotic telomere deprotection. *Nature*. **522**, 492–496 (2015).

188. X. Wu, S. Sandhu, H. Ding, Establishment of conditional knockout alleles for the gene encoding the regulator of telomere length (RTEL). *Genesis* (2007), doi:10.1002/dvg.20359.

189. N. E. Sanjana, O. Shalem, F. Zhang, Improved vectors and genome-wide

libraries for CRISPR screening. *Nat. Methods* (2014), , doi:10.1038/nmeth.3047.

190. P. G. Smiraldo, A. M. Gruver, J. C. Osborn, D. L. Pittman, Extensive chromosomal instability in Rad51d-deficient mouse cells. *Cancer Res.* (2005), doi:10.1158/0008-5472.CAN-04-2079.

191. M. Tarsounas, A. A. Davies, S. C. West, in *Philosophical Transactions of the Royal Society B: Biological Sciences* (2004).

192. S. Rao *et al.*, Lovastatin-mediated G 1 arrest is through inhibition of the proteasome, independent of hydroxymethyl glutaryl-CoA reductase. *Biochemistry*. **96**, 7797–7802 (1999).

193. L. T. Vassilev *et al.*, Selective small-molecule inhibitor reveals critical mitotic functions of human CDK1. *Proc. Natl. Acad. Sci.* **103**, 10660–10665 (2006).

194. T. Miura, W. F. Blakely, Optimization of calyculin A-induced premature chromosome condensation assay for chromosome aberration studies. *Cytom. Part A*. **79 A**, 1016–1022 (2011).

195. W. E. Wright, V. M. Tesmer, M. L. Liao, J. W. Shay, Normal human telomeres are not late replicating. *Exp. Cell Res.* (1999), doi:10.1006/excr.1999.4602.

196. P. G. Kaminker, S. H. Kim, P. Y. Desprez, J. Campisi, A novel form of the telomere-associated protein TIN2 localizes to the nuclear matrix. *Cell Cycle* (2009), doi:10.4161/cc.8.6.7941.

197. P. Baumann, E. Podell, T. R. Cech, Human Pot1 (Protection of Telomeres)

Protein: Cytolocalization, Gene Structure, and Alternative Splicing. *Mol. Cell. Biol.* (2002), doi:10.1128/mcb.22.22.8079-8087.2002.

198. C. S. Lages *et al.*, Identification of alternative transcripts of the TRF1/Pin2 gene. *J. Cell. Biochem.* (2004), doi:10.1002/jcb.20235.
199. Q. Yang *et al.*, Functional diversity of human protection of telomeres 1 isoforms in telomere protection and cellular senescence. *Cancer Res.* (2007), doi:10.1158/0008-5472.CAN-07-1390.
200. I. Grammatikakis, P. Zhang, M. P. Mattson, M. Gorospe, The long and the short of TRF2 in neurogenesis. *Cell Cycle* (2016), , doi:10.1080/15384101.2016.1222339.
201. I. Grammatikakis *et al.*, Alternative Splicing of Neuronal Differentiation Factor TRF2 Regulated by HNRNPH1/H2. *Cell Rep.* (2016), doi:10.1016/j.celrep.2016.03.080.
202. P. Zhang *et al.*, Nontelomeric splice variant of telomere repeat-binding factor 2 maintains neuronal traits by sequestering repressor element 1-silencing transcription factor. *Proc. Natl. Acad. Sci.* (2011), doi:10.1073/pnas.1106906108.
203. P. Zhang *et al.*, Novel RNA-and FMRP-binding protein TRF2-S regulates axonal mRNA transport and presynaptic plasticity. *Nat. Commun.* (2015), doi:10.1038/ncomms9888.
204. F. Sievers *et al.*, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* (2011), doi:10.1038/msb.2011.75.

205. H. Takai, A. Smogorzewska, T. De Lange, DNA damage foci at dysfunctional telomeres. *Curr. Biol.* **13**, 1549–1556 (2003).

206. N. Blom, T. Sicheritz-Pontén, R. Gupta, S. Gammeltoft, S. Brunak, Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics* (2004), , doi:10.1002/pmic.200300771.

207. N. Blom, S. Gammeltoft, S. Brunak, Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J. Mol. Biol.* (1999), doi:10.1006/jmbi.1999.3310.

208. S. Smith, The long and short of it: A new isoform of TIN2 in the nuclear matrix. *Cell Cycle* (2009), , doi:10.4161/cc.8.6.8337.

209. A. M. Wood *et al.*, TRF2 and lamin A/C interact to facilitate the functional organization of chromosome ends. *Nat. Commun.* (2014), doi:10.1038/ncomms6467.

210. P. Zhang *et al.*, Nontelomeric TRF2-REST Interaction Modulates Neuronal Gene Silencing and Fate of Tumor and Stem Cells. *Curr. Biol.* (2008), doi:10.1016/j.cub.2008.08.048.

211. A. A. Salamov, T. Nishikawa, M. B. Swindells, Assessing protein coding region integrity in cDNA sequencing projects. *Bioinformatics* (1998), doi:10.1093/bioinformatics/14.5.384.

212. T. de Lange, How telomeres solve the end-protection problem. *Science (80-.).* **326**, 948–52 (2009).

213. M. Armanios, Syndromes of Telomere Shortening. *Annu. Rev. Genomics*

Hum. Genet. **10**, 45–61 (2009).

214. S. E. Artandi, R. A. DePinho, Telomeres and telomerase in cancer. *Carcinogenesis*. **31** (2009), pp. 9–18.

215. J. Maciejowski, T. De Lange, Telomeres in cancer: Tumour suppression and genome instability. *Nat. Rev. Mol. Cell Biol.* (2017), , doi:10.1038/nrm.2016.171.

216. T. De Lange, Shelterin: The protein complex that shapes and safeguards human telomeres. *Genes Dev.* (2005), , doi:10.1101/gad.1346005.

217. W. Palm, T. de Lange, How shelterin protects mammalian telomeres. *Annu. Rev. Genet.* **42**, 301–34 (2008).

218. P. Martinez *et al.*, Increased telomere fragility and fusions resulting from TRF1 deficiency lead to degenerative pathologies and increased cancer in mice. *Genes Dev.* **23**, 2060–2075 (2009).

219. Y. Gong, T. de Lange, A Shld1-Controlled POT1a Provides Support for Repression of ATR Signaling at Telomeres through RPA Exclusion. *Mol. Cell*. **40**, 377–387 (2010).

220. R. L. Flynn *et al.*, TERRA and hnRNPA1 orchestrate an RPA-to-POT1 switch on telomeric single-stranded DNA. *Nature* (2011), doi:10.1038/nature09772.

221. W. A. Bickmore, B. Van Steensel, Genome architecture: Domain organization of interphase chromosomes. *Cell* (2013), , doi:10.1016/j.cell.2013.02.001.

222. M. Murga *et al.*, Global chromatin compaction limits the strength of the

DNA damage response. *J. Cell Biol.* **178**, 1101–1108 (2007).

223. Y. Ziv *et al.*, Chromatin relaxation in response to DNA double-strand breaks is modulated by a novel ATM- and KAP-1 dependent pathway. *Nat. Cell Biol.* **8**, 870–876 (2006).

224. A. A. Goodarzi *et al.*, ATM Signaling Facilitates Repair of DNA Double-Strand Breaks Associated with Heterochromatin. *Mol. Cell*. **31**, 167–177 (2008).

225. R. C. Burgess, B. Burman, M. J. Kruhlak, T. Misteli, Activation of DNA Damage Response Signaling by Condensed Chromatin. *Cell Rep.* **9**, 1703–1718 (2014).

226. A. Konishi, T. Izumi, S. Shimizu, TRF2 protein interacts with core histones to stabilize chromosome ends. *J. Biol. Chem.* (2016), doi:10.1074/jbc.M116.719021.

227. M. Zimmermann, T. de Lange, 53BP1: pro choice in DNA repair. *Trends Cell Biol.* **24**, 108–17 (2014).

228. M. M. J. Rust, M. Bates, X. W. Zhuang, Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat. Methods*. **3**, 793–795 (2006).

229. M. Bates, B. Huang, G. T. Dempsey, X. Zhuang, Multicolor super-resolution imaging with photo-switchable fluorescent probes. *Science (80-.).* (2007), doi:10.1126/science.1146598.

230. B. Huang, H. Babcock, X. Zhuang, Breaking the diffraction barrier: Super-resolution imaging of cells. *Cell* (2010), , doi:10.1016/j.cell.2010.12.002.

231. M. Ester, H. P. Kriegel, J. Sander, X. Xu, A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *Second Int. Conf. Knowl. Discov. Data Min.*, 226–231 (1996).

232. D. Hockemeyer, J. P. Daniels, H. Takai, T. de Lange, Recent Expansion of the Telomeric Complex in Rodents: Two Distinct POT1 Proteins Protect Mouse Telomeres. *Cell*. **126**, 63–77 (2006).

233. J. A. Aten *et al.*, Dynamics of DNA Double-Strand Breaks Revealed by Clustering of Damaged Chromosome Domains. *Science (80-.).* (2004), doi:10.1126/science.1088845.

234. P. M. Krawczyk, C. Stap, C. van Oven, R. Hoebe, J. A. Aten, Clustering of double strand break-containing chromosome domains is not inhibited by inactivation of major repair proteins. *Radiat. Prot. Dosimetry* (2006), doi:10.1093/rpd/ncl479.

235. T. Neumaier *et al.*, Evidence for formation of DNA repair centers and dose-response nonlinearity in human cells. *Proc. Natl. Acad. Sci.* (2012), doi:10.1073/pnas.1117849108.

236. N. W. Cho, R. L. Dilley, M. A. Lampson, R. A. Greenberg, Interchromosomal homology searches drive directional ALT telomere movement and synapsis. *Cell* (2014), doi:10.1016/j.cell.2014.08.030.

237. P. Mao *et al.*, Homologous recombination-dependent repair of telomeric DSBs in proliferating human cells. *Nat. Commun.* (2016), doi:10.1038/ncomms12154.

238. S. Difilippantonio *et al.*, 53BP1 facilitates long-range DNA end-joining

during V(D)J recombination. *Nature* (2008), doi:10.1038/nature07476.

239. N. Dimitrova, Y.-C. M. Chen, D. L. Spector, T. de Lange, 53BP1 promotes non-homologous end joining of telomeres by increasing chromatin mobility. *Nature*. **456**, 524–528 (2008).

240. V. L. Makarov, S. Lejnine, J. Bedoyan, J. P. Langmore, Nucleosomal organization of telomere-specific chromatin in rat. *Cell* (1993), doi:10.1016/0092-8674(93)90256-P.

241. H. Tommerup, A. Dousmanis, T. de Lange, Unusual chromatin in human telomeres. *Mol. Cell. Biol.* (2012), doi:10.1128/mcb.14.9.5777.

242. A. Sfeir, S. Kabir, M. van Overbeek, G. B. Celli, T. de Lange, Loss of Rap1 induces telomere recombination in the absence of NHEJ or a DNA damage signal. *Science*. **327**, 1657–61 (2010).

243. J. D. Buenrostro, P. G. Giresi, L. C. Zaba, H. Y. Chang, W. J. Greenleaf, Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods*. **10**, 1213–8 (2013).

244. A. Vancevska, K. M. Douglass, V. Pfeiffer, S. Manley, J. Lingner, The telomeric DNA damage response occurs in the absence of chromatin decompaction. *Genes Dev.* (2017), doi:10.1101/gad.294082.116.

245. J. Tang *et al.*, Acetylation limits 53BP1 association with damaged chromatin to promote homologous recombination. *Nat. Struct. Mol. Biol.* (2013), doi:10.1038/nsmb.2499.

246. B. Van Steensel, T. de Lange, Control of telomere length by the human

telomeric protein TRF1. *Nature*. **385**, 740–3 (1997).

247. D. Loayza, T. De Lange, POT1 as a terminal transducer of TRF1 telomere length control. *Nature* (2003), doi:10.1038/nature01688.

248. F. L. Zhong *et al.*, TPP1 OB-fold domain controls telomere maintenance by recruiting telomerase to chromosome ends. *Cell* (2012), doi:10.1016/j.cell.2012.07.012.

249. J. Nandakumar *et al.*, The TEL patch of telomere protein TPP1 mediates telomerase recruitment and processivity. *Nature* (2012), doi:10.1038/nature11648.

250. S. H. Kim *et al.*, TIN2 mediates functions of TRF2 at human telomeres. *J. Biol. Chem.* (2004), doi:10.1074/jbc.M408650200.

251. C. Iftode, Y. Daniely, J. A. Borowiec, Replication protein A (RPA): The eukaryotic SSB. *Crit. Rev. Biochem. Mol. Biol.* (1999), , doi:10.1080/10409239991209255.

252. H. Nakaoka, A. Nishiyama, M. Saito, F. Ishikawa, Xenopus laevis Ctc1-Stn1-Ten1 (xCST) protein complex is involved in priming DNA synthesis on single-stranded DNA template in Xenopus egg extract. *J. Biol. Chem.* (2012), doi:10.1074/jbc.M111.263723.

253. M. A. Blasco, W. Funk, B. Villeponteau, C. W. Greider, Functional characterization and developmental regulation of mouse telomerase RNA. *Science (80-.).* (1995), doi:10.1126/science.7544492.

254. R. A. Greenberg, R. C. Allsopp, L. Chin, G. B. Morin, R. A. DePinho, Expression of mouse telomerase reverse transcriptase during

development, differentiation and proliferation. *Oncogene* (1998), doi:10.1038/sj.onc.1201933.

255. N. M. V. Gomes *et al.*, Comparative biology of mammalian telomeres: Hypotheses on ancestral states and the roles of telomeres in longevity determination. *Aging Cell* (2011), doi:10.1111/j.1474-9726.2011.00718.x.

256. K. R. Prowse, C. W. Greider, Developmental and tissue-specific regulation of mouse telomerase and telomere length. *Proc. Natl. Acad. Sci.* (1995), doi:10.1073/pnas.92.11.4818.

257. S. B. Hedges, J. E. Blair, M. L. Venturi, J. L. Shoe, A molecular timescale of eukaryote evolution and the rise of complex multicellular life. *BMC Evol. Biol.* (2004), doi:10.1186/1471-2148-4-2.

258. K. J. Peterson *et al.*, Estimating metazoan divergence times with a molecular clock. *Proc. Natl. Acad. Sci.* (2004), doi:10.1073/pnas.0401670101.

259. L. D. Bromham, M. D. Hendy, Can fast early rates reconcile molecular dates with the Cambrian explosion? *Proc. R. Soc. B Biol. Sci.* (2000), doi:10.1098/rspb.2000.1108.

260. M. Lynch, The Age and Relationships of the Major Animal Phyla. *Evolution (N. Y).* (1999), doi:10.2307/2640769.

261. X. Gu, Early metazoan divergence was about 830 million years ago. *J. Mol. Evol.* (1998), doi:10.1007/PL00013150.

262. F. J. Ayala, A. Rzhetsky, F. J. Ayala, Origin of the metazoan phyla: Molecular clocks confirm paleontological estimates. *Proc. Natl. Acad. Sci.*

(1998), doi:10.1073/pnas.95.2.606.

263. L. W. Parfrey, D. J. G. Lahr, A. H. Knoll, L. A. Katz, Estimating the timing of early eukaryotic diversification with multigene molecular clocks. *Proc. Natl. Acad. Sci.* (2011), doi:10.1073/pnas.1110633108.

264. M. Dos Reis *et al.*, Uncertainty in the Timing of Origin of Animals and the Limits of Precision in Molecular Timescales. *Curr. Biol.* (2015), doi:10.1016/j.cub.2015.09.066.

265. R. W. Levis, R. Ganesan, K. Houtchens, L. A. Tolar, F. miin Sheen, Transposons in place of telomeric repeats at a Drosophila telomere. *Cell* (1993), doi:10.1016/0092-8674(93)90318-K.

266. C. W. Roth, F. Kobeski, M. F. Walter, H. Biessmann, Chromosome end elongation by recombination in the mosquito Anopheles gambiae. *Mol. Cell. Biol.* (1997), doi:10.1128/mcb.17.9.5176.

267. G. D. Raffa, L. Ciapponi, G. Cenci, M. Gatti, Terminin: A protein complex that mediates epigenetic maintenance of Drosophila telomeres. *Nucleus* (2011), , doi:10.4161/nucl.2.5.17873.

268. J. P. Abad *et al.*, Genomic analysis of Drosophila melanogaster telomeres: Full-length copies of HeT-A and TART elements at telomeres. *Mol. Biol. Evol.* (2004), doi:10.1093/molbev/msh174.

269. J. M. Mason, A. Y. Konev, H. Biessmann, Telomeric position effect in Drosophila melanogaster reflects a telomere length control mechanism. *Genetica* (2003), , doi:10.1023/A:1022925003172.

270. E. Casacuberta, M. L. Pardue, HeT-A and TART, two Drosophila

retrotransposons with a bona fide role in chromosome structure for more than 60 million years. *Cytogenet. Genome Res.* (2005), , doi:10.1159/000084947.

271. J. M. Mason, H. Biessmann, The unusual telomeres of Drosophila. *Trends Genet.* (1995), , doi:10.1016/S0168-9525(00)88998-2.
272. H. Biessmann, S. B. Carter, J. M. Mason, Chromosome ends in Drosophila without telomeric DNA sequences. *Proc. Natl. Acad. Sci.* (2006), doi:10.1073/pnas.87.5.1758.
273. H. Biessmann, J. M. Mason, Progressive loss of DNA sequences from terminal chromosome deficiencies in Drosophila melanogaster. *EMBO J.* (2018), doi:10.1002/j.1460-2075.1988.tb02916.x.
274. H. Biessmann *et al.*, Frequent transpositions of Drosophila melanogaster HeT-A transposable elements to receding chromosome ends. *EMBO J.* (2018), doi:10.1002/j.1460-2075.1992.tb05547.x.
275. R. W. Levis, Viable deletions of a telomere from a Drosophila chromosome. *Cell* (1989), doi:10.1016/0092-8674(89)90112-8.
276. L. Melnikova, H. Biessmann, P. Georgiev, The Ku protein complex is involved in length regulation of Drosophila telomeres. *Genetics* (2005), doi:10.1534/genetics.104.034538.
277. M. Savitsky, O. Kravchuk, L. Melnikova, P. Georgiev, Heterochromatin Protein 1 Is Involved in Control of Telomere Elongation in Drosophila melanogaster. *Mol. Cell. Biol.* (2002), doi:10.1128/mcb.22.9.3204-3218.2002.

278. G. M. Siriaco *et al.*, Telomere elongation (Tel), a new mutation in Drosophila melanogaster that produces long telomeres. *Genetics* (2002).

279. G. D. Raffa *et al.*, Verrocchio, a Drosophila OB fold-containing protein, is a component of the terminin telomere-capping complex. *Genes Dev.* (2010), doi:10.1101/gad.574810.

280. L. Fanti, G. Giovinazzo, M. Berloco, S. Pimpinelli, The heterochromatin protein 1 prevents telomere fusions in Drosophila. *Mol. Cell* (1998), doi:10.1016/S1097-2765(00)80152-5.

281. G. Cenci, G. Siriaco, G. D. Raffa, R. Kellum, M. Gatti, The drosophila HOAP protein in required for telomere capping. *Nat. Cell Biol.* (2003), doi:10.1038/ncb902.

282. G. D. Raffa *et al.*, The Drosophila modigliani (moi) gene encodes a HOAP-interacting protein required for telomere protection. *Proc. Natl. Acad. Sci.* (2009), doi:10.1073/pnas.0812702106.

283. L. Ciapponi *et al.*, The Drosophila Mre11/Rad50 complex is required to prevent both telomeric fusion and chromosome breakage. *Curr. Biol.* (2004), doi:10.1016/j.cub.2004.07.019.

284. S. R. Oikemus *et al.*, Drosophila atm/telomere fusion is required for telomeric localization of HP1 and telomere position effect. *Genes Dev.* (2004), doi:10.1101/gad.1202504.

285. X. Bi, S. C. D. Wei, Y. S. Rong, Telomere protection without a telomerase: The role of ATM and Mre11 in Drosophila telomere maintenance. *Curr. Biol.* (2004), doi:10.1016/j.cub.2004.06.063.

286. H. J. Muller, INDUCED MUTATIONS IN DROSOPHILA. *Cold Spring Harb. Symp. Quant. Biol.* (1941), doi:10.1101/sqb.1941.009.01.019.

287. K. Sahara, F. Marec, W. Traut, TTAGG telomeric repeats in chromosomes of some insects and other arthropods. *Chromosom. Res.* (1999), doi:10.1023/A:1009297729547.

288. R. Frydrychová, F. Marec, Repeated losses of TTAGG telomere repeats in evolution of beetles (Coleoptera). *Genetica* (2002), doi:10.1023/A:1020175912128.

289. T. De Lange, T-loops and the origin of telomeres. *Nat. Rev. Mol. Cell Biol.* (2004), , doi:10.1038/nrm1359.

290. B. Mravinac, N. Meštrović, V. Vanja Čavrak, M. Plohl, TCAGG, an alternative telomeric sequence in insects. *Chromosoma* (2011), doi:10.1007/s00412-011-0317-x.

291. J. M. Mason, T. A. Randall, R. Capkova Frydrychova, Telomerase lost? *Chromosoma* (2016), , doi:10.1007/s00412-015-0528-7.

292. M. Korandová, T. Krůček, K. Vrbová, R. Č. Frydrychová, Distribution of TTAGG-specific telomerase activity in insects. *Chromosom. Res.* (2014), doi:10.1007/s10577-014-9436-6.

293. C. M. Counter, M. Meyerson, E. N. Eaton, R. A. Weinberg, The catalytic subunit of yeast telomerase. *Proc. Natl. Acad. Sci.* (1997), doi:10.1073/pnas.94.17.9202.

294. J. Lingner *et al.*, Reverse transcriptase motifs in the catalytic subunit of telomerase. *Science (80-.).* (1997), doi:10.1126/science.276.5312.561.

295. J. W. Szostak, E. H. Blackburn, Cloning yeast telomeres on linear plasmid vectors. *Cell* (1982), doi:10.1016/0092-8674(82)90109-X.

296. J. Shampay, J. W. Szostak, E. H. Blackburn, DNA sequences of telomeres maintained in yeast. *Nature* (1984), doi:10.1038/310154a0.

297. S. Marcand, E. Gilson, D. Shore, A protein-counting mechanism for telomere length regulation in yeast. *Science (80-.).* (1997), doi:10.1126/science.275.5302.986.

298. A. J. Cesare, C. Groff-Vindman, S. A. Compton, M. J. McEachern, J. D. Griffith, Telomere Loops and Homologous Recombination-Dependent Telomeric Circles in a Kluyveromyces lactis Telomere Mutant Strain. *Mol. Cell. Biol.* (2007), doi:10.1128/mcb.01122-07.

299. D. Shore, K. Nasmyth, Purification and cloning of a DNA binding protein from yeast that binds to both silencer and activator elements. *Cell* (1987), doi:10.1016/0092-8674(87)90095-X.

300. A. Krauskopf, E. H. Blackburn, Rap1 protein regulates telomere turnover in yeast. *Proc. Natl. Acad. Sci.* (2002), doi:10.1073/pnas.95.21.12486.

301. A. Krauskopf, E. H. Blackburn, Control of telomere growth by interactions of RAP1 with the most distal telomeric repeats. *Nature* (1996), doi:10.1038/383354a0.

302. G. Kyrion, K. A. Boakye, A. J. Lustig, C-terminal truncation of RAP1 results in the deregulation of telomere size, stability, and function in Saccharomyces cerevisiae. *Mol. Cell. Biol.* (2015), doi:10.1128/mcb.12.11.5159.

303. G. Kyrion, K. Liu, C. Liu, A. J. Lustig, RAP1 and telomere structure regulate telomere position effects in Saccharomyces cerevisiae. *Genes Dev.* (1993), doi:10.1101/gad.7.7a.1146.

304. P. König, R. Giraldo, L. Chapman, D. Rhodes, The crystal structure of the DNA-binding domain of yeast RAP1 in complex with telomeric DNA. *Cell* (1996), doi:10.1016/S0092-8674(00)81088-0.

305. D. Laporte, F. Courtout, S. Tollis, I. Sagot, Quiescent Saccharomyces cerevisiae forms telomere hyperclusters at the nuclear membrane vicinity through a multifaceted mechanism involving Esc1, the Sir complex, and chromatin condensation. *Mol. Biol. Cell* (2016), doi:10.1091/mbc.e16-01-0069.

306. S. Marcand, B. Pardo, A. Gratias, S. Cahun, I. Callebaut, Multiple pathways inhibit NHEJ at telomeres. *Genes Dev.* (2008), doi:10.1101/gad.455108.

307. T. Shi *et al.*, Rif1 and Rif2 shape telomere function and architecture through multivalent Rap1 interactions. *Cell* (2013), doi:10.1016/j.cell.2013.05.007.

308. M. Martina *et al.*, A Balance between Tel1 and Rif2 Activities Regulates Nucleolytic Processing and Elongation at Telomeres. *Mol. Cell. Biol.* (2012), doi:10.1128/mcb.06547-11.

309. Y. Hirano, K. Fukunaga, K. Sugimoto, Rif1 and Rif2 Inhibit Localization of Tel1 to DNA Ends. *Mol. Cell* (2009), doi:10.1016/j.molcel.2008.12.027.

310. S. Anbalagan, D. Bonetti, G. Lucchini, M. P. Longhese, Rif1 supports the

function of the CST complex in yeast telomere capping. *PLoS Genet.* (2011), doi:10.1371/journal.pgen.1002024.

311. N. Grandin, C. Damon, M. Charbonneau, Cdc13 Cooperates with the Yeast Ku Proteins and Stn1 To Regulate Telomerase Recruitment. *Mol. Cell. Biol.* (2002), doi:10.1128/mcb.20.22.8397-8408.2000.

312. N. Grandin, S. I. Reed, M. Charbonneau, Stn1, a new Saccharomyces cerevisiae protein, is implicated in telomere size regulation in association with CDc13. *Genes Dev.* (1997), doi:10.1101/gad.11.4.512.

313. N. Grandin, C. Damon, M. Charbonneau, Ten1 functions in telomere end protection and length regulation in association with Stn1 and Cdc13. *EMBO J.* (2001), doi:10.1093/emboj/20.5.1173.

314. S. Li *et al.*, Cdk1-Dependent Phosphorylation of Cdc13 Coordinates Telomere Elongation during Cell-Cycle Progression. *Cell* (2009), doi:10.1016/j.cell.2008.11.027.

315. B. Garvik, M. Carson, L. Hartwell, Single-stranded DNA arising at telomeres in cdc13 mutants may constitute a specific signal for the RAD9 checkpoint [published erratum appears in Mol Cell Biol 1996 Jan;16(1):457]. *Mol Cell Biol* (1995).

316. Y. Hiraoka, E. Henderson, E. H. Blackburn, Not so peculiar: Fission yeast telomere repeats. *Trends Biochem. Sci.* (1998), doi:10.1016/S0968-0004(98)01176-1.

317. N. F. Lue, Y. Peng, Identification and characterization of a telomerase activity from Schizosaccharomyces pombe. *Nucleic Acids Res.* (1997),

doi:10.1093/nar/25.21.4331.

318. C. J. Webb, V. A. Zakian, Identification and characterization of the Schizosaccharomyces pombe TER1 telomerase RNA. *Nat. Struct. Mol. Biol.* (2008), doi:10.1038/nsmb1354.

319. L. Tomaska, S. Willcox, J. Slezakova, J. Nosek, J. D. Griffith, Taz1 binding to a fission yeast model telomere: Formation of telomeric loops and higher order structures. *J. Biol. Chem.* (2004), doi:10.1074/jbc.M409790200.

320. J. P. Cooper, E. R. Nimmo, R. C. Allshire, T. R. Cech, Regulation of telomere length and function by a Myb-domain protein in fission yeast. *Nature* (1997), doi:10.1038/385744a0.

321. W. Deng *et al.*, Fission yeast telomere-binding protein Taz1 is a functional but not a structural counterpart of human TRF1 and TRF2. *Cell Res.* (2015), doi:10.1038/cr.2015.76.

322. M. G. Ferreira, J. P. Cooper, The fission yeast Taz1 protein protects chromosomes from Ku-dependent end-to-end fusions. *Mol. Cell* (2001), doi:10.1016/S1097-2765(01)00154-X.

323. J. Kanoh, F. Ishikawa, spRap1 and spRif1, recruited to telomeres by Taz1, are essential for telomere function in fission yeast. *Curr. Biol.* (2001), doi:10.1016/S0960-9822(01)00503-6.

324. T. Miyoshi, J. Kanoh, M. Saito, F. Ishikawa, Fission yeast pot1-Tpp1 protects telomeres and regulates telomere length. *Science (80-.).* (2008), doi:10.1126/science.1154819.

325. H. Scott *et al.*, Spatial Organization and Molecular Interactions of the

Schizosaccharomyces pombe Ccq1–Tpz1–Poz1 Shelterin Complex. *J. Mol. Biol.* (2017), doi:10.1016/j.jmb.2017.08.002.

326. J. Xue *et al.*, Structure of the fission yeast S. pombe telomeric Tpz1-Poz1-Rap1 complex. *Cell Res.* (2017), doi:10.1038/cr.2017.145.
327. M. R. Flory, A. R. Carson, E. G. Muller, R. Aebersold, An SMC-domain protein in fission yeast links telomeres to the meiotic centrosome. *Mol. Cell* (2004), doi:10.1016/j.molcel.2004.10.027.
328. K. Tomita, J. P. Cooper, Fission yeast Ccq1 is telomerase recruiter and local checkpoint controller. *Genes Dev.* (2008), doi:10.1101/gad.498608.
329. T. Sugiyama *et al.*, SHREC, an Effector Complex for Heterochromatic Transcriptional Silencing (DOI:10.1016/j.cell.2006.12.035). *Cell* (2007), , doi:10.1016/j.cell.2007.04.014.
330. T. Naito, A. Matsuura, F. Ishikawa, Circular chromosome formation in a fission yeast mutant defective in two ATM homologues. *Nat. Genet.* (1998), doi:10.1038/2517.
331. E. J. Richards, F. M. Ausubel, Isolation of a higher eukaryotic telomere from Arabidopsis thaliana. *Cell* (1988), doi:10.1016/0092-8674(88)90494-1.
332. J. M. Watson, K. Riha, Comparative biology of telomeres: Where plants stand. *FEBS Lett.* (2010), , doi:10.1016/j.febslet.2010.06.017.
333. A. J. Cesare, N. Quinney, S. Willcox, D. Subramanian, J. D. Griffith, Telomere looping in P. sativum (common garden pea). *Plant J.* (2003), doi:10.1046/j.1365-313X.2003.01882.x.
334. E. Sýkorová *et al.*, Telomere variability in the monocotyledonous plant

order Asparagales. *Proc. R. Soc. B Biol. Sci.* (2003), doi:10.1098/rspb.2003.2446.

335. A. V. Cox *et al.*, Comparison of plant telomere locations using a pcr-generated synthetic probe. *Ann. Bot.* (1993), doi:10.1006/anbo.1993.1104.
336. U. Pich, I. Schubert, Terminal heterochromatin and alternative telomeric sequences in Allium cepa. *Chromosom. Res.* (1998), doi:10.1023/A:1009227009121.
337. E. Sỳkorová *et al.*, Minisatellite telomeres occur in the family Alliaceae but are lost in Allium. *Am. J. Bot.* (2006), doi:10.3732/ajb.93.6.814.
338. P. Fajkus *et al.*, Allium telomeres unmasked: The unusual telomeric sequence (CTCGGTTATGGG)n is synthesized by telomerase. *Plant J.* (2016), doi:10.1111/tpj.13115.
339. Z. N. Karamysheva, Y. V. Surovtseva, L. Vespa, E. V. Shakirov, D. E. Shippen, A C-terminal Myb extension domain defines a novel family of double-strand telomeric DNA-binding proteins in arabidopsis. *J. Biol. Chem.* (2004), doi:10.1074/jbc.M407938200.
340. S. Ko *et al.*, Structure of the DNA-binding domain of NgTRF1 reveals unique features of plant telomere-binding proteins. *Nucleic Acids Res.* (2008), doi:10.1093/nar/gkn030.
341. M. G. Hwang, M. H. Cho, Arabidopsis thaliana telomeric DNA-binding protein 1 is required for telomere length homeostasis and its Myb-extension domain stabilizes plant telomeric DNA binding. *Nucleic Acids Res.* (2007), doi:10.1093/nar/gkm043.

342. S. W. Yang, Perturbation of NgTRF1 Expression Induces Apoptosis-Like Cell Death in Tobacco BY-2 Cells and Implicates NgTRF1 in the Control of Telomere Length and Stability. *PLANT CELL ONLINE* (2004), doi:10.1105/tpc.104.026278.

343. J.-P. Hong *et al.*, Suppression of RICE TELOMERE BINDING PROTEIN1 Results in Severe and Gradual Developmental Defects Accompanied by Genome Instability in Rice . *Plant Cell* (2007), doi:10.1105/tpc.107.051953.

344. E. V. Shakirov, T. D. McKnight, D. E. Shippen, POT1-independent single-strand telomeric DNA binding activities in Brassicaceae. *Plant J.* (2009), doi:10.1111/j.1365-313X.2009.03837.x.

345. Y. V. Surovtseva *et al.*, Arabidopsis POT1 associates with the telomerase RNP and is required for telomere maintenance. *EMBO J.* (2007), doi:10.1038/sj.emboj.7601792.

346. E. V. Shakirov, Y. V. Surovtseva, N. Osbun, D. E. Shippen, The Arabidopsis Pot1 and Pot2 Proteins Function in Telomere Length Homeostasis and Chromosome End Protection. *Mol. Cell. Biol.* (2005), doi:10.1128/mcb.25.17.7725-7733.2005.

347. X. Song *et al.*, STN1 protects chromosome ends in Arabidopsis thaliana. *Proc. Natl. Acad. Sci.* (2008), doi:10.1073/pnas.0807867105.

348. C. M. Price *et al.*, Evolution of CST function in telomere maintenance. *Cell Cycle* (2010), , doi:10.4161/cc.9.16.12547.

349. C. W. Greider, Regulating telomere length from the inside out: The replication fork model. *Genes Dev.* (2016), doi:10.1101/gad.280578.116.

350. L. A. Klobutcher, M. T. Swanton, P. Donini, D. M. Prescott, All gene-sized DNA molecules in four species of hypotrichs have the same terminal sequence and an unusual 3’ terminus. *Proc. Natl. Acad. Sci.* (2006), doi:10.1073/pnas.78.5.3015.

351. E. H. Blackburn, J. G. Gall, A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in Tetrahymena. *J. Mol. Biol.* (1978), doi:10.1016/0022-2836(78)90294-2.

352. T. De Lange, A loopy view of telomere evolution. *Front. Genet.* (2015), doi:10.3389/fgene.2015.00321.

353. A. M. Zahler, D. M. Prescott, Telomere terminal transferase activity in the hypotrichous ciliate Oxytricha nova and a model for replication of the ends of linear DNA molecules. *Nucleic Acids Res.* (1988), doi:10.1093/nar/16.14.6953.

354. R. M. Mitton-Fry, E. M. Anderson, T. R. Hughes, V. Lundblad, D. S. Wuttke, Conserved structure for single-stranded telomeric DNA recognition. *Science (80-.).* (2002), doi:10.1126/science.1068799.

355. D. E. Gottschling, V. A. Zakian, Telomere proteins: Specific recognition and protection of the natural termini of Oxytricha macronuclear DNA. *Cell* (1986), doi:10.1016/0092-8674(86)90442-3.

356. J. T. Gray, D. W. Celander, C. M. Price, T. R. Cech, Cloning and expression of genes for the Oxytricha telomere-binding protein: Specific subunit interactions in the telomeric complex. *Cell* (1991), doi:10.1016/0092-8674(91)90075-A.

357. C. M. Price, T. R. Cech, Properties of the Telomeric DNA-Binding Protein from Oxytricha nova. *Biochemistry* (1989), doi:10.1021/bi00428a053.

358. G. Fang, J. T. Gray, T. R. Cech, Oxytricha telomere-binding protein: Separable DNA-binding and dimerization domains of the α-subunit. *Genes Dev.* (1993), doi:10.1101/gad.7.5.870.

359. G. Fang, T. R. Cech, Characterization of a G-Quartet Formation Reaction Promoted by the β-Subunit of the Oxytricha Telomere-Binding Protein. *Biochemistry* (1993), doi:10.1021/bi00094a022.

360. S. J. Froelich-Ammon, B. A. Dickinson, J. M. Bevilacqua, S. C. Schultz, T. R. Cech, Modulation of telomerase activity by telomere DNA-binding proteins in Oxytricha. *Genes Dev.* (1998), doi:10.1101/gad.12.10.1504.

361. E. C. Swart *et al.*, The Oxytricha trifallax Macronuclear Genome: A Complex Eukaryotic Genome with 16,000 Tiny Chromosomes. *PLoS Biol.* (2013), doi:10.1371/journal.pbio.1001473.

362. N. K. Jacob, R. Lescasse, B. R. Linger, C. M. Price, Tetrahymena POT1a Regulates Telomere Length and Prevents Activation of a Cell Cycle Checkpoint. *Mol. Cell. Biol.* (2006), doi:10.1128/mcb.01975-06.

363. B. R. Linger, G. B. Morin, C. M. Price, The Pot1a-associated proteins Tpt1 and Pat1 coordinate telomere protection and length regulation in Tetrahymena. *Mol. Biol. Cell* (2011), doi:10.1091/mbc.e11-06-0551.

364. V. L. Premkumar *et al.*, The 3’ overhangs at Tetrahymena thermophila telomeres are packaged by four proteins, Pot1a, Tpt1, Pat1, and Pat2. *Eukaryot. Cell* (2014), doi:10.1128/EC.00275-13.

365. J. Jiang *et al.*, Structure of Tetrahymena telomerase reveals previously unknown subunits, functions, and interactions. *Science (80-.).* (2015), doi:10.1126/science.aab4070.

366. B. Wan *et al.*, The Tetrahymena telomerase p75-p45-p19 subcomplex is a unique CST complex. *Nat. Struct. Mol. Biol.* (2015), doi:10.1038/nsmb.3126.

367. B. Min, K. Collins, An RPA-Related Sequence-Specific DNA-Binding Subunit of Telomerase Holoenzyme Is Required for Elongation Processivity and Telomere Maintenance. *Mol. Cell* (2009), doi:10.1016/j.molcel.2009.09.041.

368. T. Cavalier-Smith, The phagotrophic origin of eukaryotes and phylogenetic classification on protozoa. *Int. J. Syst. Evol. Microbiol.* (2002), , doi:10.1099/00207713-52-2-297.

369. L. H. T. Van der Ploeg, A. Y. C. Liu, P. Borst, Structure of the growing telomeres of trypanosomes. *Cell* (1984), doi:10.1016/0092-8674(84)90239-3.

370. M. I. N. Cano, J. M. Dungan, N. Agabian, E. H. Blackburn, Telomerase in kinetoplastid parasitic protozoa. *Proc. Natl. Acad. Sci.* (2002), doi:10.1073/pnas.96.7.3616.

371. J. L. Muñoz-Jordán, G. A. M. Cross, T. De Lange, J. D. Griffith, t-loops at trypanosome telomeres. *EMBO J.* (2001), doi:10.1093/emboj/20.3.579.

372. B. Li, A. Espinal, G. A. M. Cross, Trypanosome Telomeres Are Protected by a Homologue of Mammalian TRF2. *Mol. Cell. Biol.* (2005),

doi:10.1128/mcb.25.12.5011-5021.2005.

373. X. Yang, L. M. Figueiredo, A. Espinal, E. Okubo, B. Li, RAP1 Is Essential for Silencing Telomeric Variant Surface Glycoprotein Genes in Trypanosoma brucei. *Cell* (2009), doi:10.1016/j.cell.2009.01.037.

374. S. M. Le blancq, R. S. Kase, L. H. t. Van De Ploeg, Analysis of a giardia lamblia rRNA encoding telomere with [TAGGG]n. *Nucleic Acids Res.* (1991), doi:10.1093/nar/19.20.5790.

375. R. D. Adam, T. E. Nash, T. E. Wellems, Telomeric location of Giardia rDNA genes. *Mol. Cell. Biol.* (1991), doi:10.1128/mcb.11.6.3326.

376. H. S. Malik, W. D. Burke, T. H. Eickbush, Putative telomerase catalytic subunits from Giardia lamblia and Caenorhabditis elegans. *Gene* (2000), doi:10.1016/S0378-1119(00)00207-9.

377. M. Uzlíková *et al.*, Characterization of telomeres and telomerase from the single-celled eukaryote Giardia intestinalis. *Mol. Biochem. Parasitol.* (2017), doi:10.1016/j.molbiopara.2016.09.003.

378. I. R. Arkhipova, H. G. Morrison, Three retrotransposon families in the genome of Giardia lamblia: Two telomeric, one dead. *Proc. Natl. Acad. Sci.* (2002), doi:10.1073/pnas.231494798.

379. S. B. Hedges *et al.*, A genomic timescale for the origin of eukaryotes. *BMC Evol. Biol.* (2001), doi:10.1186/1471-2148-1-4.

380. R. F. D. Doolittle, D. F. Feng, G. Cho, R. F. Doolittle, Determining divergence times with a protein clock: update and reevaluation. *Proc. Natl. Acad. Sci. U. S. A.* (1997).

381. J. Marin, F. U. Battistuzzi, A. C. Brown, S. B. Hedges, The Timetree of Prokaryotes: New Insights into Their Evolution and Speciation. *Mol. Biol. Evol.* (2017), doi:10.1093/molbev/msw245.

382. P. P. Sheridan, K. H. Freeman, J. E. Brenchley, Estimated minimal divergence times of the major bacterial and archaeal phyla. *Geomicrobiol. J.* (2003), doi:10.1080/01490450303891.

383. G. R. Weller *et al.*, Identification of a DNA nonhomologous end-joining complex in bacteria. *Science (80-.).* (2002), doi:10.1126/science.1074584.

384. M. Della *et al.*, Mycobacterial Ku and ligase proteins constitute a two-component NHEJ repair machine. *Science (80-.).* (2004), doi:10.1126/science.1099824.

385. R. S. Pitcher, L. M. Tonkin, A. J. Green, A. J. Doherty, Domain structure of a NHEJ DNA repair ligase from Mycobacterium tuberculosis. *J. Mol. Biol.* (2005), doi:10.1016/j.jmb.2005.06.038.

386. R. Chayot, B. Montagne, D. Mazel, M. Ricchetti, An end-joining repair mechanism in Escherichia coli. *Proc. Natl. Acad. Sci.* (2010), doi:10.1073/pnas.0906355107.

387. H. Zhu, S. Shuman, Characterization of Agrobacterium tumefaciens DNA ligases C and D. *Nucleic Acids Res.* (2007), doi:10.1093/nar/gkm145.

388. B. Michel, After 30 years of study, the bacterial SOS response still surprises us. *PLoS Biol.* (2005), , doi:10.1371/journal.pbio.0030255.

389. D. G. Anderson, S. C. Kowalczykowski, Reconstitution of an SOS response pathway: Derepression of transcription in response to DNA

breaks. *Cell* (1998), doi:10.1016/S0092-8674(00)81721-3.

390. L. Sagan, On the origin of mitosing cells. *J. Theor. Biol.* (1967).

391. R. M. Schwartz, M. O. Dayhoff, Origins of prokaryotes, eukaryotes, mitochondria, and chloroplasts. *Science (80-.).* (1978), , doi:10.1126/science.202030.

392. L. Bonen, R. S. Cunningham, M. W. Gray, W. F. Doolittle, Wheat embryo mitochondrial 18S ribosomal RNA: Evidence for its prokaryotic nature. *Nucleic Acids Res.* (1977), doi:10.1093/nar/4.3.663.

393. L. Bonen, W. F. Doolittle, On the prokaryotic nature of red algal chloroplasts. *Proc. Natl. Acad. Sci.* (1975), doi:10.1073/pnas.72.6.2310.

394. P. H. Raven, A multiple origin for plastids and mitochondria. *Science (80-.).* (1970), , doi:10.1126/science.169.3946.641.

395. R. Kolodner, K. K. Tewari, Molecular size and conformation of chloroplast deoxyribonucleic acid from pea leaves. *J. Biol. Chem.* (1972).

396. J. Nosek, L. Tomáška, H. Fukuhara, Y. Suyama, L. Kováč, Linear mitochondrial genomes: 30 years down the line. *Trends Genet.* (1998), , doi:10.1016/S0168-9525(98)01443-7.

397. G. B. Morin, T. R. Cech, The telomeres of the linear mitochondrial DNA of tetrahymena thermophila consist of 53 bp tandem repeats. *Cell* (1986), doi:10.1016/0092-8674(86)90069-3.

398. D. J. Oldenburg, A. J. Bendich, Most Chloroplast DNA of Maize Seedlings in Linear Molecules with Defined Ends and Branched Forms. *J. Mol. Biol.* (2004), doi:10.1016/j.jmb.2003.11.020.

399. G. B. Morin, T. R. Cech, Mitochondrial telomeres: Surprising diversity of repeated telomeric DNA sequences among six species of Tetrahymena. *Cell* (1988), doi:10.1016/S0092-8674(88)80029-1.

400. H. Fukuhara *et al.*, Linear mitochondrial DNAs of yeasts: frequency of occurrence and general features. *Mol. Cell. Biol.* (1993).

401. B. M. Baroudy, S. Venkatesan, B. Moss, Structure and replication of vaccinia virus telomeres. *Cold Spring Harb. Symp. Quant. Biol.* (1982).

402. J. M. Gerhold, A. Aun, T. Sedman, P. Jõers, J. Sedman, Strand invasion structures in the inverted repeat of candida albicans mitochondrial DNA reveal a role for homologous recombination in replication. *Mol. Cell* (2010), doi:10.1016/j.molcel.2010.09.002.

403. T. C. Walthert, J. C. Kennell, Linear mitochondrial plasmids of F. oxysporum are novel, telomere-like retroelements. *Mol. Cell* (1999), doi:10.1016/S1097-2765(00)80370-6.

404. L. Tomaska, A. M. Makhov, J. D. Griffith, J. Nosek, t-loops in yeast mitochondria. *Mitochondrion* (2002), doi:10.1016/S1567-7249(02)00009-0.

405. L. Erickson, W. D. Beversdorf, K. P. Pauls, Linear mitochondrial plasmid in Brassica has terminal protein. *Curr. Genet.* (1985), doi:10.1007/BF00449821.

406. B. W. Stillman, The replication of adenovirus DNA with purified proteins. *Cell* (1983), doi:10.1016/0092-8674(83)90201-5.

407. J. Nosek, Ľ. Tomáška, B. Kucejová, The chromosome end replication: lessons from mitochondrial genetics. *J. Appl. Biomed.* (2019),

doi:10.32725/jab.2004.008.

408. L. Tomáška, J. Nosek, H. Fukuhara, Identification of a putative mitochondrial telomere-binding protein of the yeast Candida parapsilosis. *J. Biol. Chem.* (1997), doi:10.1074/jbc.272.5.3049.

409. J. Nosek, L. Tomáška, B. Pagáčová, H. Fukuhara, Mitochondrial telomere-binding protein from Candida parapsilosis suggests an evolutionary adaptation of a nonspecific single-stranded DNA- binding protein. *J. Biol. Chem.* (1999), doi:10.1074/jbc.274.13.8850.

410. F. U. Battistuzzi, A. Feijao, S. B. Hedges, A genomic timescale of prokaryote evolution: Insights into the origin of methanogenesis, phototrophy, and the colonization of land. *BMC Evol. Biol.* (2004), doi:10.1186/1471-2148-4-44.

411. F. U. Battistuzzi, S. B. Hedges, A major clade of prokaryotes with ancient adaptations to life on land. *Mol. Biol. Evol.* (2009), doi:10.1093/molbev/msn247.

412. A. G. Barbour, C. F. Garon, Linear plasmids of the bacterium Borrelia burgdorferi have covalently closed ends. *Science (80-.).* (1987), doi:10.1126/science.3603026.

413. A. Allardet-Servent, S. Michaux-Charachon, E. Jumas-Bilak, L. Karayan, M. Ramuz, Presence of one linear and one circular chromosome in the Agrobacterium tumefaciens C58 genome. *J. Bacteriol.* (1993).

414. M. S. Ferdows, A. G. Barbour, Megabase-sized linear DNA in the bacterium Borrelia burgdorferi, the Lyme disease agent. *Trends Genet.*

(1989), doi:10.1016/0168-9525(89)90169-8.

415. C. M. Fraser *et al.*, Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi. *Nature* (1997), doi:10.1038/37551.

416. J. Hinnebusch, A. G. Barbour, Linear plasmids of Borrelia burgdorferi have a telomeric structure and sequence similar to those of a eukaryotic virus. *J. Bacteriol.* (1991), doi:10.1128/jb.173.22.7233-7239.1991.

417. K. Kobryn, G. Chaconas, The circle is broken: Telomere resolution in linear replicons. *Curr. Opin. Microbiol.* (2001), , doi:10.1016/S1369-5274(00)00251-4.

418. K. Shi, W. M. Huang, H. Aihara, An Enzyme-Catalyzed Multistep DNA Refolding Mechanism in Hairpin Telomere Formation. *PLoS Biol.* (2013), doi:10.1371/journal.pbio.1001472.

419. W. M. Huang *et al.*, Linear chromosome-generating system of agrobacterium tumefaciens C58: Protelomerase generates and protects hairpin ends. *J. Biol. Chem.* (2012), doi:10.1074/jbc.M112.369488.

420. K. Kobryn, G. Chaconas, in *Mobile DNA III* (2015).

421. G. Chaconas, K. Kobryn, Structure, Function, and Evolution of Linear Replicons in Borrelia . *Annu. Rev. Microbiol.* (2010), doi:10.1146/annurev.micro.112408.134037.

422. S. H. Huang, K. Kobryn, The Borrelia burgdorferi telomere resolvase, ResT, anneals ssDNA complexed with its cognate ssDNA-binding protein. *Nucleic Acids Res.* (2016), doi:10.1093/nar/gkw344.

423. S. H. Huang, M. R. Cozart, M. A. Hart, K. Kobryn, The Borrelia

burgdorferitelomere resolvase, ResT, possesses ATP-dependent DNA unwinding activity. *Nucleic Acids Res.* (2017), doi:10.1093/nar/gkw1243.

424. K. Kobryn, G. Chaconas, ResT, a telomere resolvase encoded by the Lyme disease spirochete. *Mol. Cell* (2002), doi:10.1016/S1097-2765(01)00433-6.

425. Y. Tourand, K. Kobryn, G. Chaconas, Sequence-specific recognition but position-dependent cleavage of two distinct telomeres by the Borrelia burgdorferi telomere resolvase, ResT. *Mol. Microbiol.* (2003), doi:10.1046/j.1365-2958.2003.03485.x.

426. K. Kobryn, G. Chaconas, Fusion of hairpin telomeres by the B. burgdorferi telomere resolvase ResT: Implications for shaping a genome in flux. *Mol. Cell* (2005), doi:10.1016/j.molcel.2005.02.025.

427. Y. -S Lin, H. M. Kieser, D. A. Hopwood, C. W. Chen, The chromosomal DNA of Streptomyces lividans 66 is linear. *Mol. Microbiol.* (1993), doi:10.1111/j.1365-2958.1993.tb00964.x.

428. K. Bao, S. N. Cohen, Terminal proteins essential for the replication of linear plasmids and chromosomes in Streptomyces. *Genes Dev.* (2001), doi:10.1101/gad.896201.

429. K. Bao, S. N. Cohen, Recruitment of terminal protein to the ends of Streptomyces linear plasmids and chromosomes by a novel telomere-binding protein essential for linear DNA replication. *Genes Dev.* (2003), doi:10.1101/gad.1060303.

430. Y. Nindita *et al.*, The tap-tpg gene pair on the linear plasmid functions to

maintain a linear topology of the chromosome in Streptomyces rochei. *Mol. Microbiol.* (2015), doi:10.1111/mmi.12904.

431. J. N. Volff, J. Altenbuchner, Genetic instability of the streptomyces chromosome. *Mol. Microbiol.* (1998), , doi:10.1046/j.1365-2958.1998.00652.x.

432. Y. Lin, C. W. Chen, Instability of artificially circularized chromosomes of Streptomyces lividans . *Mol. Microbiol.* (2003), doi:10.1046/j.1365-2958.1997.5991975.x.

433. J. N. Volff, P. Viell, J. Altenbuchner, Artificial circularization of the chromosome with concomitant deletion of its terminal inverted repeats enhances genetic instability and genome rearrangement in Streptomyces lividans. *Mol. Gen. Genet.* (1997), doi:10.1007/s004380050380.

434. C. H. Huang, Y. S. Lin, Y. L. Yang, S. W. Huang, C. W. Chen, The telomeres of Streptomyces chromosomes contain conserved palindromic sequences with potential to form complex secondary structures. *Mol. Microbiol.* (1998), doi:10.1046/j.1365-2958.1998.00856.x.

435. Z. Qin, S. N. Cohen, Replication at the telomeres of the Streptomyces linear plasmid pSLA2. *Mol. Microbiol.* (1998), doi:10.1046/j.1365-2958.1998.00838.x.

436. C. C. Yang, S. M. Tseng, C. W. Chen, Telomere-associated proteins add deoxynucleotides to terminal proteins during replication of the telomeres of linear chromosomes and plasmids in Streptomyces. *Nucleic Acids Res.* (2015), doi:10.1093/nar/gkv302.

437. C. C. Yang, S. M. Tseng, H. Y. Pan, C. H. Huang, C. W. Chen, Telomere associated primase Tap repairs truncated telomeres of Streptomyces. *Nucleic Acids Res.* (2017), doi:10.1093/nar/gkx189.

438. A. Svarchevsky, V. Rybchin, Physical mapping of plasmid N15 DNA. *Mol. Gen. Mikrobiol. Virusol* (1984).

439. Ai. Malinin, A. A. Vostrov, V. N. Rybchin, A. N. Svarchevskii, [Structure of ends of linear plasmid N15]. *Mol Gen Mikrobiol Virusol* (1992).

440. Ai. Malinin *et al.*, [Structure of a region in the genome of bacteriophage N15, necessary for formation of hairpins at ends of the linear plasmid prophage]. *Mol Gen Mikrobiol Virusol* (1992).

441. V. Ravin *et al.*, Genomic sequence and analysis of the atypical temperate bacteriophage N15. *J. Mol. Biol.* (2000), doi:10.1006/jmbi.2000.3731.

442. J. Deneke, G. Ziegelin, R. Lurz, E. Lanka, Phage N15 telomere resolution. Target requirements for recognition and processing by the protelomerase. *J. Biol. Chem.* (2002), doi:10.1074/jbc.M111769200.

443. J. Deneke, G. Ziegelin, R. Lurz, E. Lanka, The protelomerase of temperate Escherichia coli phage N15 has cleaving-joining activity. *Proc. Natl. Acad. Sci. U. S. A.* (2000).

444. W. M. Huang, L. Joss, T. Hsieh, S. Casjens, Protelomerase Uses a Topoisomerase IB/Y-Recombinase Type Mechanism to Generate DNA Hairpin Ends. *J. Mol. Biol.* (2004), doi:10.1016/j.jmb.2004.01.012.

445. A. M. DeLange, M. Reddy, D. Scraba, C. Upton, G. McFadden, Replication and resolution of cloned poxvirus telomeres in vivo generates linear

minichromosomes with intact viral hairpin termini. *J. Virol.* (1986).

446. A. M. DeLange, G. McFadden, Efficient resolution of replicated poxvirus telomeres to native hairpin structures requires two inverted symmetrical copies of a core target DNA sequence. *J. Virol.* (1987).

447. D. Stuart, K. Ellison, K. Graham, G. McFadden, In vitro resolution of poxvirus replicative intermediates into linear minichromosomes with hairpin termini by a virally induced Holliday junction endonuclease. *J Virol* (1992).

448. F. Tamanoi, B. W. Stillman, Function of adenovirus terminal protein in the initiation of DNA replication. *Proc. Natl. Acad. Sci.* (1982), doi:10.1073/pnas.79.7.2221.

449. T. Enomoto, J. H. Lichy, J. E. Ikeda, J. Hurwitz, Adenovirus DNA replication in vitro: purification of the terminal protein in a functional form. *Proc. Natl. Acad. Sci.* (1981), doi:10.1073/pnas.78.11.6779.

450. A. Baker, K. J. Rohleder, L. A. Hanakahi, G. Ketner, Adenovirus E4 34k and E1b 55k Oncoproteins Target Host DNA Ligase IV for Proteasomal Degradation. *J. Virol.* (2007), doi:10.1128/jvi.00029-07.

451. D. A. Garber, S. M. Beverley, D. M. Coen, Demonstration of circularization of herpes simplex virus DNA following infection using pulsed field gel electrophoresis. *Virology* (1993), doi:10.1006/viro.1993.1612.

452. Y.-H. Su, X. Zhang, X. Wang, N. W. Fraser, T. M. Block, Evidence that the Immediate-Early Gene Product ICP4 Is Necessary for the Genome of the Herpes Simplex Virus Type 1 ICP4 Deletion Mutant Strain d120 To Circularize in Infected Cells. *J. Virol.* (2006), doi:10.1128/jvi.01869-06.

453. I. Muylaert, P. Elias, Knockdown of DNA ligase IV/XRCC4 by RNA interference inhibits herpes simplex virus type I DNA replication. *J. Biol. Chem.* (2007), doi:10.1074/jbc.M611834200.

454. L. A. Timashev, H. Babcock, X. Zhuang, T. de Lange, The DDR at telomeres lacking intact shelterin does not require substantial chromatin decompaction. *Genes Dev.* **31**, 578–589 (2017).

455. H. J. Muller, Further studies on the nature and causes of gene mutations. *Int. Congr. Genet.* (1932).

456. J. S. Z. Li *et al.*, TZAP: A telomere-associated protein involved in telomere length control. *Science (80-.).* (2017), doi:10.1126/science.aah6752.

457. A. Orthwein *et al.*, Mitosis inhibits DNA double-strand break repair to guard against telomere fusions. *Science (80-.).* (2014), doi:10.1126/science.1248024.

458. M. T. Hayashi, A. J. Cesare, J. A. J. Fitzpatrick, E. Lazzerini-Denchi, J. Karlseder, A telomere-dependent DNA damage checkpoint induced by prolonged mitotic arrest. *Nat. Struct. Mol. Biol.* (2012), doi:10.1038/nsmb.2245.

459. P. M. Lansdorp *et al.*, Heterogeneity in telomere length of human chromosomes. *Hum. Mol. Genet.* (1996), doi:10.1093/hmg/5.5.685.

460. M. G. L. Gustafsson, Nonlinear structured-illumination microscopy: wide-field fluorescence imaging with theoretically unlimited resolution. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 13081–13086 (2005).

461. E. Betzig *et al.*, Imaging intracellular fluorescent proteins at nanometer

resolution. *Science*. **313**, 1642–5 (2006).

462. S. T. Hess, T. P. K. Girirajan, M. D. Mason, Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophys. J.* **91**, 4258–72 (2006).

463. B. Huang, M. Bates, X. Zhuang, Super-Resolution Fluorescence Microscopy. *Annu. Rev. Biochem.* **78**, 993–1016 (2009).

464. H. G. Hansma, K. Kasuya, E. Oroudjev, Atomic force microscopy imaging and pulling of nucleic acids. *Curr. Opin. Struct. Biol.* (2004), , doi:10.1016/j.sbi.2004.05.005.

465. G. Binning, H. Rohrer, C. Gerber, E. Weibel, Surface studies by scanning tunneling microscopy. *Phys. Rev. Lett.* (1982), doi:10.1103/PhysRevLett.49.57.

466. F. J. Giessibl, Advances in atomic force microscopy. *Rev. Mod. Phys.* (2003), , doi:10.1103/RevModPhys.75.949.

467. G. Binnig, H. Rohrer, Scanning tunneling microscopy from birth to adolescence. *Rev. Mod. Phys.* (1987), doi:10.1103/RevModPhys.59.615.

468. F. Chen, P. W. Tillberg, E. S. Boyden, Expansion microscopy. *Science (80-.).* **347**, 543–548 (2015).

469. M. E. E. Ludérus *et al.*, Structure, subnuclear distribution, and nuclear matrix association of the mammalian telomeric complex. *J. Cell Biol.* (1996), doi:10.1083/jcb.135.4.867.

470. J. Mahamid *et al.*, Visualizing the molecular sociology at the HeLa cell nuclear periphery. *Science (80-.).* (2016), doi:10.1126/science.aad8857.

471. M. Eltsov *et al.*, Nucleosome conformational variability in solution and in interphase nuclei evidenced by cryo-electron microscopy of vitreous sections. *Nucleic Acids Res.* (2018), doi:10.1093/nar/gky670.

472. R. Van De Plas, J. Yang, J. Spraggins, R. M. Caprioli, Image fusion of mass spectrometry and microscopy: A multimodality paradigm for molecular tissue mapping. *Nat. Methods* (2015), doi:10.1038/nmeth.3296.

473. M. K. Passarelli *et al.*, The 3D OrbiSIMS - Label-free metabolic imaging with subcellular lateral resolution and high mass-resolving power. *Nat. Methods* (2017), doi:10.1038/nmeth.4504.

474. M. Kompauer, S. Heiles, B. Spengler, Atmospheric pressure MALDI mass spectrometry imaging of tissues and cells at 1.4-μm lateral resolution. *Nat. Methods* (2016), doi:10.1038/nmeth.4071.

475. J. Schindelin *et al.*, Fiji: An open-source platform for biological-image analysis. *Nat. Methods* (2012), , doi:10.1038/nmeth.2019.

476. C. A. Schneider, W. S. Rasband, K. W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* (2012), , doi:10.1038/nmeth.2089.

477. U. Herbig, W. A. Jobling, B. P. C. Chen, D. J. Chen, J. M. Sedivy, Telomere shortening triggers senescence of human cells through a pathway involving ATM, p53, and p21 CIP1, but not p16 INK4a. *Mol. Cell*. **14**, 501–513 (2004).

478. London *et al.*, Alternative Lengthening of Telomeres Is Characterized by High Rates of Telomeric Exchange. *Cancer Res.* **64**, 2324–2327 (2004).

479. J. B. Chang *et al.*, Iterative expansion microscopy. *Nat. Methods* (2017), doi:10.1038/nmeth.4261.