



OPEN ACCESS

Edited by:

Robert Joseph Shmookler Reis,
Central Arkansas Veterans
Healthcare System Eugene J.
Towbin Healthcare Center,
United States

Reviewed by:

Bibo Li,
Cleveland State University,
United States
Mark A. McCormick,
University of New Mexico,
United States

***Correspondence:**

Jack Griffith
jdg@med.unc.edu

[†]Present address:

Anirban Kar,
National Institutes of Environmental
Health Sciences, Genome Integrity
and Structural Biology Laboratory,
Research Triangle Park, NC,
United States

Specialty section:

This article was submitted to
Genetics of Aging,
a section of the journal
Frontiers in Genetics

Received: 07 May 2019

Accepted: 26 July 2019

Published: 14 August 2019

Citation:

Tomaska L, Nosek J, Kar A,
Willcox S and Griffith JD (2019)
A New View of the T-Loop Junction:
Implications for Self-Primed Telomere
Extension, Expansion of Disease-
Related Nucleotide Repeat Blocks,
and Telomere Evolution.
Front. Genet. 10:792.
doi: 10.3389/fgene.2019.00792

A New View of the T-Loop Junction: Implications for Self-Primed Telomere Extension, Expansion of Disease-Related Nucleotide Repeat Blocks, and Telomere Evolution

Lubomir Tomaska¹, Jozef Nosek¹, Anirban Kar^{2†}, Smaranda Willcox² and Jack D. Griffith^{2*}

¹ Departments of Genetics and Biochemistry, Faculty of Natural Sciences, Comenius University in Bratislava, Bratislava, Slovakia, ² Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

Telomere loops (t-loops) are formed at the ends of chromosomes in species ranging from humans to worms, plants, and with genetic manipulation, some yeast. Recent *in vitro* studies demonstrated that transcription of telomeric DNA leads to highly efficient t-loop formation. It was also shown that both DNA termini are inserted into the preceding DNA to generate a highly stable t-loop junction. Furthermore, some telomeric RNA remains present at the junction, potentially acting as a plug to further protect and stabilize the t-loop. Modeling the loop junction reveals two mechanisms by which the canonical chromosomal replication factors could extend the telomere in the absence of telomerase. One mechanism would utilize the annealed 3' terminus as a *de novo* replication origin. *In vitro* evidence for the ability of the t-loop to prime telomere extension using the T7 replication factors is presented. A second mechanism would involve resolution of the Holliday junction present in the t-loop bubble by factors such as GEN1 to generate a rolling circle template at the extreme terminus of the telomere. This could lead to large expansions of the telomeric tract. Here, we propose that telomeres evolved as terminal elements containing long arrays of short nucleotide repeats due to the ability of such arrays to fold back into loops and self-prime their replicative extension. In this view, telomerase may have evolved later to provide a more precise mechanism of telomere maintenance. Both pathways have direct relevance to the alternative lengthening of telomeres (ALT) pathway. This view also provides a possible mechanism for the very large repeat expansions observed in nucleotide repeat diseases such as Fragile X syndrome, myotonic dystrophy, familial amyotrophic lateral sclerosis (ALS), and frontotemporal dementia (FTD). The evolution of telomeres is discussed in the framework of these models.

Keywords: telomere, t-loop, telomerase, replication, evolution, intron, Holliday junction

FORMATION OF T-LOOPS AT THE TELOMERE: A LINK BETWEEN LOOPING AND TRANSCRIPTION

Looping the end of the telomere back into the preceding $(5'-TTAGGG-3')_n$ repeat sequence to generate a lasso-like structure (t-loop) was first described in 1999 (Griffith et al., 1999) in human and murine cells, and in following studies extended by us and others to trypanosomes, *Oxytricha*, chickens, *Caenorhabditis elegans*, plants, and with some genetic manipulation, to yeasts [reviewed in Doksani et al. (2013) and Griffith (2013)]. These studies employed electron microscopy (EM) to directly visualize the DNA loops. A major stride was accomplished in 2013 when Doksani et al. (2013) visualized t-loops using stochastic optical reconstruction microscopy (STORM) imaging and employed this technique to confirm the presence of t-loops in murine cells and demonstrate their dependence on functional TRF2 protein. The rationale which drove the original proposal of t-loops was based on our knowledge of the mechanism of homologous recombination (HR) and the fact that a double-stranded (ds) DNA containing a long single-stranded (ss) overhang with a 3' terminus provides a highly efficient template for strand invasion of the ssDNA into a homologous dsDNA in reactions driven by proteins such as recA, uvsX, or Rad51. The telomere represents a structure in which both pairing partners are present in the same molecule. The demonstration of t-loops and this thinking led to the classic diagram of a t-loop (Figure 1A) in which the bubble formed at the t-loop junction contains just the 3' ssDNA overhang inserted. However, even in early papers, it was proposed that the 5' strand (Tomaska et al., 2002) or both strands may also insert, leading to a more stable t-loop junction (Stansel et al., 2001).

The discovery that telomeres are transcribed (Azzalin et al., 2007; Schoeftner and Blasco, 2008) and that the G-rich RNA termed TERRA is increased in cells with elevated HR (Arora et al., 2014) led us to ask if t-loop formation might be related to TERRA

formation. In Kar et al. (2016), we developed a simple *in vitro* system in which a terminal block of 576 bp of TTAGGG repeats in a linearized plasmid can be transcribed through the repeats by T7 RNA polymerase to yield TERRA. Following transcription, EM was used to probe for the presence of t-loops at the end of the DNA containing the mini-telomere. It was found that transcription of the telomeric block led to highly efficient t-loop formation with up to 60% of the input DNA containing a 576-bp or smaller loop at the telomeric end. Surprisingly, t-loops were formed at high levels irrespective of the presence of a long 3' ss overhang. Telomeres that were blunt ended, contained a 4-nt 5' overhang or 3' overhangs of 54–96 nt presented equally good templates for looping. Examination of the t-loop junction showed the frequent presence of a small nucleic acid bead, which, when labeled with a biotin-tagged RNA precursor, revealed that some TERRA remained at the junction in spite of RNase treatment. The results forced a revision of the t-loop junction model.

CONSEQUENCES OF A T-LOOP JUNCTION WITH BOTH TERMINI INSERTED

In the new t-loop model (Figure 1B), both terminal strands are annealed to their complementary strands in the bubble. This structure has features of both a replication fork (right side) and a Holliday junction (left side). With both strands base paired, the t-loops would be much more stable than with just the ss overhang paired. Although comparison of the stability of single- versus double-stranded paired t-loops will need to be quantified by assessment of their thermodynamic properties (such as free Gibbs energy), we have observed that t-loops generated by transcription *in vitro* were highly stable to deproteinization and remained present for days at 4°C (Kar et al., 2016). The suggestion that the 3' G-rich overhang is not essential for t-loop formation is in line with the

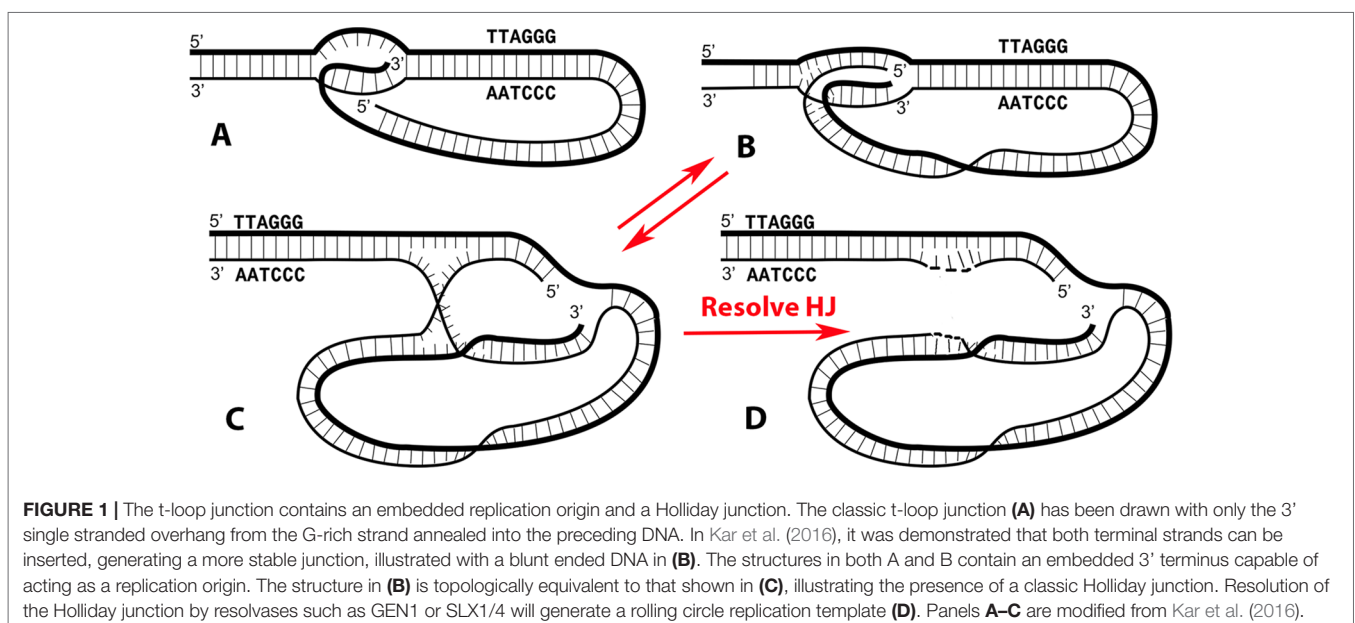


FIGURE 1 | The t-loop junction contains an embedded replication origin and a Holliday junction. The classic t-loop junction (A) has been drawn with only the 3' single stranded overhang from the G-rich strand annealed into the preceding DNA. In Kar et al. (2016), it was demonstrated that both terminal strands can be inserted, generating a more stable junction, illustrated with a blunt ended DNA in (B). The structures in both A and B contain an embedded 3' terminus capable of acting as a replication origin. The structure in (B) is topologically equivalent to that shown in (C), illustrating the presence of a classic Holliday junction. Resolution of the Holliday junction by resolvases such as GEN1 or SLX1/4 will generate a rolling circle replication template (D). Panels A–C are modified from Kar et al. (2016).

observation that while t-loops have been seen in plants (Cesare et al., 2003) and can be very large, it remains unclear whether significant ss overhangs are present at plant telomeres. Furthermore, in worms, some telomeric termini contain 5' ss termini (Raices et al., 2008). Based on these new findings, a model of t-loop formation was proposed by Kar et al. (2016) in which transcription of the telomere opens the DNA helix, frequently leaving R-loops behind. These would provide preferential sites for shelterin complexes (Palm and de Lange, 2008; Erdel et al., 2017) loaded at the telomeric terminus to loop back and establish a stable t-loop. The central function of the long 3' ss overhang in mammalian telomeres may be loading Pot1 and the shelterins at the end of the telomere enabling them to interact with other shelterin complexes bound to internal R-loops along the telomere. Indeed, TERRA has been shown to bind TRF2 (Deng et al., 2009) and thus would recruit a shelterin complex to an internal R-loop. Such interactions would generate a broad distribution of sizes for the circular portion of the t-loop, from large loops formed near the subtelomeric repeats to smaller loops formed at sites closer to the telomere end. This is, in fact, what is seen in naturally isolated t-loops (Griffith et al., 1999; Doksan et al., 2013).

In early t-loop studies (Griffith et al., 1999; Stansel et al., 2001), it was pointed out that a 3' ss telomeric overhang base paired in the t-loop bubble presents a *de novo* replication origin. This remains so with the newer model (Figure 1B). Self-primed telomere extension from the t-loop would employ the normal cadre of chromosomal replication factors including a DNA polymerase, sliding clamp, helicase, and primase among others. The kind of telomeric extension exemplified by this experiment has features of the alternative lengthening of telomeres (ALT) phenomenon. Telomeres in ALT cells (which lack active telomerase) show a much broader range of sizes and changes in size than normal cells (Bryan et al., 1995), reviewed in (Cesare and Reddel, 2010). In self-primed replication, the degree of extension will vary from short (small terminal t-loop) to long (large t-loop with the bubble close to the sub-telomeric sequences). Extension from the annealed 3' terminus will generate a ssDNA loop as replication proceeds toward the end of the telomere followed by its release as a longer ss tail upon encountering the telomere end. Conversion to duplex DNA would be carried out by the chromosomal replication factors.

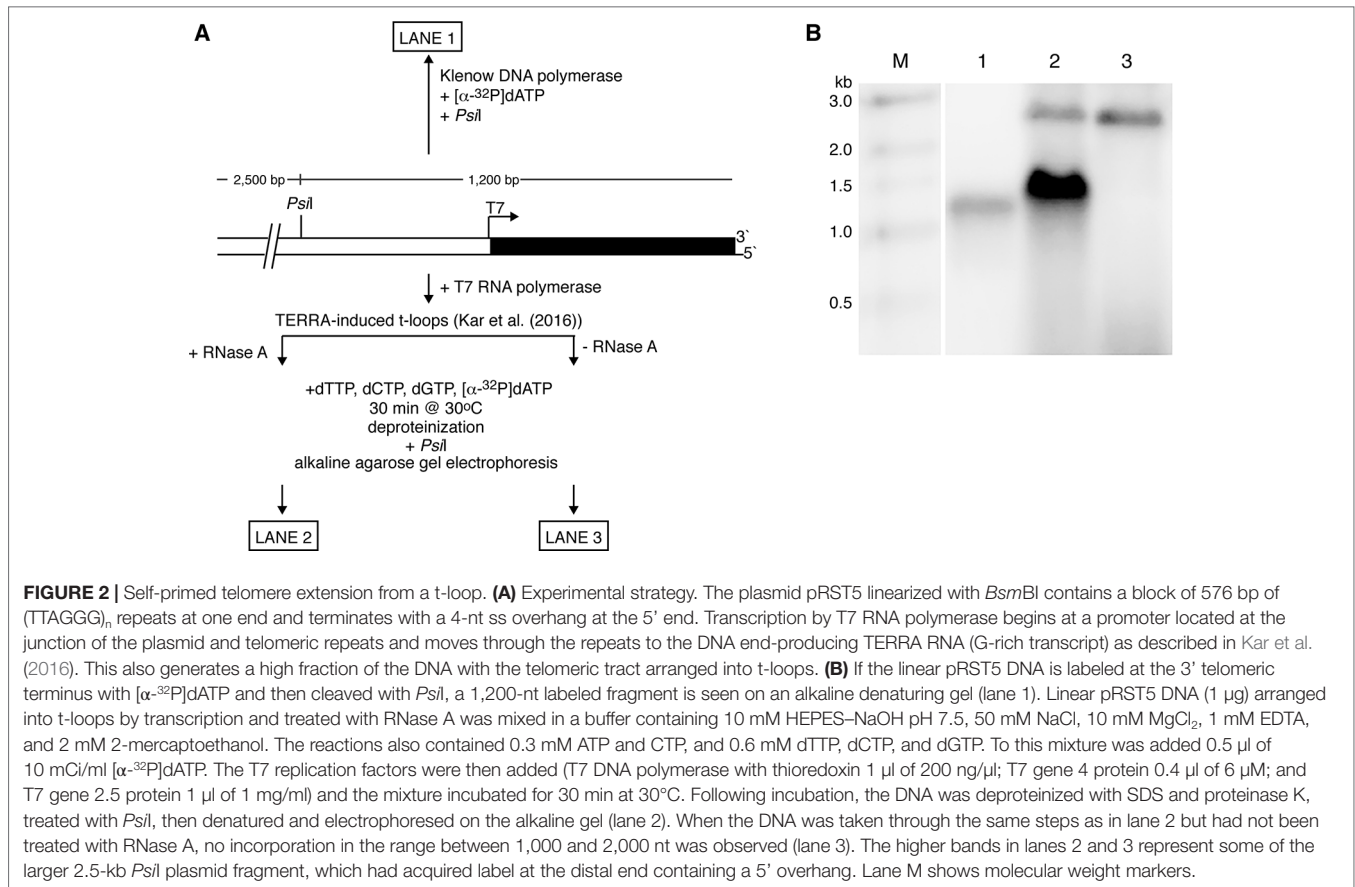
In vivo, access of the chromosomal replication factors to the embedded 3' telomeric terminus in the t-loop must be tightly controlled. Otherwise, telomeric extension *via* self-priming could go on unchecked in a cellular environment where telomerase would be expected to provide a more precise control over telomere length. Access could be blocked by assembly of the shelterin complexes at the t-loop junction. In addition, in Kar et al. (2016), we showed that a visible particle containing TERRA was often present at the t-loop junction. This RNA very likely was present in a G-quadruplex-stabilized structure formed between TERRA and the G-rich telomeric strand. It was relatively resistant to several RNases including RNase H and H1. While the detailed organization of the t-loop bubble containing TERRA remains to be determined, the presence of an "RNA plug" may further stabilize the t-loop structure and sequester it from unwanted activation. Further analysis using the *in vitro* system may shed light on these possibilities.

To conduct a "proof of principle" test of the ability of a t-loop to self-prime telomeric replication, we utilized the T7 phage

replication factors and linear pRST5 plasmid DNA containing t-loops generated by T7 RNA polymerase transcription through the TTAGGG repeats as described in Kar et al. (2016). The T7 factors include T7 gene 5 DNA polymerase with thioredoxin, gene 4 helicase-primase, and gene 2.5 single-strand binding protein (Kulczyk and Richardson, 2016). Treatment of the linear DNA with *PsiI* endonuclease generates a 1.2-kb fragment containing the telomeric block and a larger 2.5-kb plasmid fragment. The DNA (not arranged into t-loops) can be labeled at the telomeric 3' terminus with Klenow DNA polymerase and [α - 32 P]dATP in an exchange reaction. Subsequent cleavage with *PsiI* and electrophoresis on a denaturing alkaline gel results in a ssDNA band at the position of a 1,200-nt fragment (Figure 2, lane 1). To ask if the 3' terminus embedded in a t-loop can self-prime extension by the T7 replication factors, the 3.5-kb plasmid arranged into t-loops (including deproteinization and RNase A treatment to remove RNA at the t-loop) was incubated with the T7 replication proteins for 30 min at 30°C (see Figure 2 legend for details) in a reaction containing [α - 32 P]dATP. This was followed by cleavage with *PsiI* and electrophoresis on an alkaline gel. A strong band of label incorporation (lane 2) into DNA was observed with the median position shifted to a molecular weight higher than the control band in lane 1. Depending on the size of the t-loop formed at the end of the pRST5 DNA, self-primed extension could increase the size of the *PsiI* fragment by an amount between ~100 (the smallest t-loop observed) and ~550 nucleotides (the largest) resulting in a distribution of fragments on the gel between 1,300 and 1,750 nucleotides. This is what was observed. We also found that removal of the RNA by stringent RNase A treatment was absolutely required to observe self-primed replication. Shown in lane 3 is the result of an experiment identical to that in lane 2 except that the DNA was not treated with RNase A following transcription and deproteinization. This results in an abundance of RNA remaining along the telomeric tract [see Figure 3A in Kar et al. (2016)]. In this case, no incorporation of radiolabel was observed following incubation with the T7 proteins (lane 3) showing that the presence of TERRA strongly inhibited self-primed extension. The higher bands in lanes 2 and 3 represent the larger *PsiI* fragment, which has acquired some label. Of note, DNAs not arranged into t-loops would not be expected to incorporate label as they would lack the embedded 3' self-priming terminus and in agreement, little or no label was seen in lane 3 at the position of the nonextended fragment (lane 1). Clearly, further work would employ a fully reconstituted eukaryotic replication system, such as those that have been developed in *Saccharomyces cerevisiae* (Yeeles et al., 2015), and larger t-loop substrates.

CONVERSION OF THE T-LOOP TO A ROLLING CIRCLE TEMPLATE

Inspection of the new model of the t-loop bubble (Figure 1C) reveals that resolution of the Holliday junction (HJ) at the point where the two ssDNAs cross over each other will generate a covalently closed ssDNA circle annealed to the strand containing the free 3' terminus (Figure 1D). HJ resolution at the t-loop thus



creates a classic rolling circle template capable of extending the preceding DNA by large amounts. This would employ the canonical chromosomal replication factors. In human cells, resolution could be accomplished by GEN1 or the SLXC1/4-MUS81 complex (Rass et al., 2010; Garner et al., 2013; Wyatt et al., 2013; Wyatt and West, 2014; Chan and West, 2015). This presents a second pathway by which t-loops can function to extend the telomere.

Probably, the greatest challenge will be to show that t-loops corresponding to the structures depicted by **Figures 1B, C** are formed *in vivo*. One possible experimental approach would be based on the *in vitro* observation that t-loops are induced by transcription and stabilized by TERRA. So far, telomeric fragments have been isolated from cells as either deproteinized DNA or as DNA–protein complexes (disregarding RNA). It would be worth trying to purify telomeric fragments associated with TERRA from cell nuclei. This could be done either by immunoprecipitating one of the shelterin components under RNase-free conditions or by directly purifying TERRA in complex with telomeric DNA. The latter approach could be based on the CHIRT protocol developed to investigate the genomic binding sites of TERRA transcripts (Chu et al., 2017). The topology of the DNA–TERRA complexes would then be assessed by means of electron microscopy combined with mapping the double- and single-stranded regions.

NUCLEOTIDE REPEAT EXPANSION: CAN THIS MECHANISM EXPLAIN THE VERY LONG EXPANSIONS SEEN IN SOME OF THE NUCLEOTIDE REPEAT DISEASES?

In the classic nucleotide repeat diseases, short blocks of repeating tri- to hexa-nucleotides which are normally stable in the genome expand in size with consequences resulting in severe disease. In Huntington's disease, (CTG)_n expansions in the coding sequences for the huntingtin protein are relatively modest, growing from ~10 to 20 repeats in the unaffected genome, to ~100 repeats in severely afflicted individuals [reviewed in Bates et al. (2015)]. However, very large expansions are observed in other diseases. In Fragile X syndrome, the normal repeat block of 10–25 (CCG)_n triplets may expand to over 1,000 bp in afflicted male patients [reviewed in Ciaccio et al. (2017)]. The myotonic dystrophy type 2 locus begins with a (CCTG)_n repeat block a few tens of repeats in size but can grow to thousands of base pairs in patients presenting with the disease (Liquori et al., 2001). More recently, in familial amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD), a hexanucleotide block (CCGGGG)_n in the C9orf72 locus which is normally a few tens of repeats in size has been shown to expand to thousands of base pairs (DeJesus-Hernandez et al., 2011; Renton et al., 2011).

The molecular mechanisms by which expansions are initiated and grow have been the topic of many studies but to date are not fully understood. The predominant model poses that replication through the repeat block results in backwards slippage of the polymerase machinery and looping out of one of the strands (Wells and Ashizawa, 2006). If the ssDNA in the loop can pair on itself as in the (CTG)_n or (CCG)_n repeats, this will produce a partially base paired stable hairpin at the fork. Nicking and filling in steps would then add the looped-out nucleotides to the repeat block. However, ssDNA with the (GAA)_n repeat that expands in Friedreich's ataxia cannot form partially paired hairpins. More importantly, the very large expansions seem difficult to explain by simple slippage models since this would generate extremely large ssDNA loops at the replication fork if expansion were to occur in one event and these would be targets for nucleases. Hence, it seems possible that the very large expansions involve a two-step mechanism possibly beginning with replication slippage but then employing a different mechanism in the second stage potentially involving the formation of a t-loop like structure.

The models described above offer possible mechanisms for the very large nucleotide repeat expansions. Replication through long (5'-CCGGG-3')_n repeats can result in replication fork stalling (Thys and Wang, 2015) producing ds breaks. A break within the nucleotide repeats would create two DNA ends, each with similarities to telomeres and with the capability of forming "t-loops" either spontaneously, following transcription through the DNA to the ends, or by action of HR factors including Rad51. Once formed, the loops would protect the broken ends from recognition as ds breaks. However, they could also generate cycles of self-primed replicative extension either by initiation from the embedded 3' terminus or *via* HJ cleavage and production of a rolling circle template. These expansions would eventually be followed by opening the loops and fusion of the repeat block, now expanded greatly by the cycles of self-primed replication, back into the chromosome.

EVOLUTIONARY IMPLICATIONS

The involvement of t-loops in DNA transactions described above supports the participation of these structures in the evolution of telomeres. The origin of linear chromosomes and telomeres providing solutions to end-protection and end-replication problems are enigmatic and still a matter of debate (e.g., Nakamura and Cech, 1998; Ishikawa and Naito, 1999; de Lange, 2004; Fajkus et al., 2005; Nosek et al., 2006; Linger and Price, 2009; Lue, 2010; Forterre, 2013; Garavís et al., 2013; Fulcher et al., 2014; de Lange, 2015; Slijepcevic, 2016; Cervenak et al., 2017; Lue, 2018). Even the basic question regarding the advantage of linear versus circular genomes in eukaryotes remains unanswered. The early speculation supported by the results from mutants of fission yeast *Schizosaccharomyces pombe* possessing circularized chromosomes, i.e., that circular chromosomes would compromise meiotic segregation (Ishikawa and Naito, 1999), was questioned (de Lange, 2015) by the existence of the bacterial XerD/C resolution machinery efficiently cleaving circular chromosomes at specific *dif* sites (Barre et al., 2001). Furthermore, many eukaryotes lack

sexual reproduction and do not undergo meiosis, yet they retain linear chromosomes and canonical telomeres. These arguments support the conclusion that telomeres are structures exhibiting very potent means for escaping recircularization through ligation (de Lange, 2015), and t-loops provide a simple, powerful means of blocking circularization. In addition, it has been proposed that centromeres were derived from telomeres during the evolution of eukaryotic chromosomes implying that linear chromosomes arose even before the emergence of a mitotic segregation machinery (Villasante et al., 2007). Another line of speculation would be that large eukaryotic genomes are better off being fragmented into a number of linear chromosomes whose ends would provide the possibility of their specific spatial organization/localization and/or means for transcriptional regulation of genes. However, the reduction in the number of chromosomes in *S. cerevisiae* from 16 to 1 or 2 did not yield any substantial changes in the transcriptomes of such strains compared to the wild type (Luo et al., 2018; Shao et al., 2018). In addition, the existence of an ant species possessing a single chromosome demonstrates that even a complex eukaryote does not need to have a fragmented genome (Crosland and Crozier, 1986). All in all, these results suggest that linearity of eukaryotic chromosomes might have been forced by a selfish element(s) without necessarily providing selective advantage for the host.

In line with this hypothesis is the scenario that takes into account early events in eukaryogenesis (de Lange, 2015). It was indicated that accommodation of a α -proteobacterial endosymbiont that gave rise to mitochondria was accompanied by a massive invasion of group II introns (Martin and Koonin, 2006). According to de Lange (2015), the spreading of introns across the genome resulted in the inevitability of linear chromosomes. The occurrence of a double-strand DNA break would be either repaired by nonhomologous end joining pathways, leading to recircularization of the genome, or through HR resulting in the formation of a loop *via* invasion of the broken end into a copy of the intron. When the number of introns was low, the HR pathway would result in repair of the double-strand break followed by recircularization of the chromosome. However, as the number of introns increased, the chance of recircularization of a chromosome decreased, since it would require the two loops present at the ends of a single DNA molecule to be passed by replication forks synchronously. Under these circumstances, the t-loops would provide solutions for both end-protection and end-replication problems serving as primordial telomeres, later replaced by a more robust telomerase-dependent telomere replication (de Lange, 2004; de Lange, 2015).

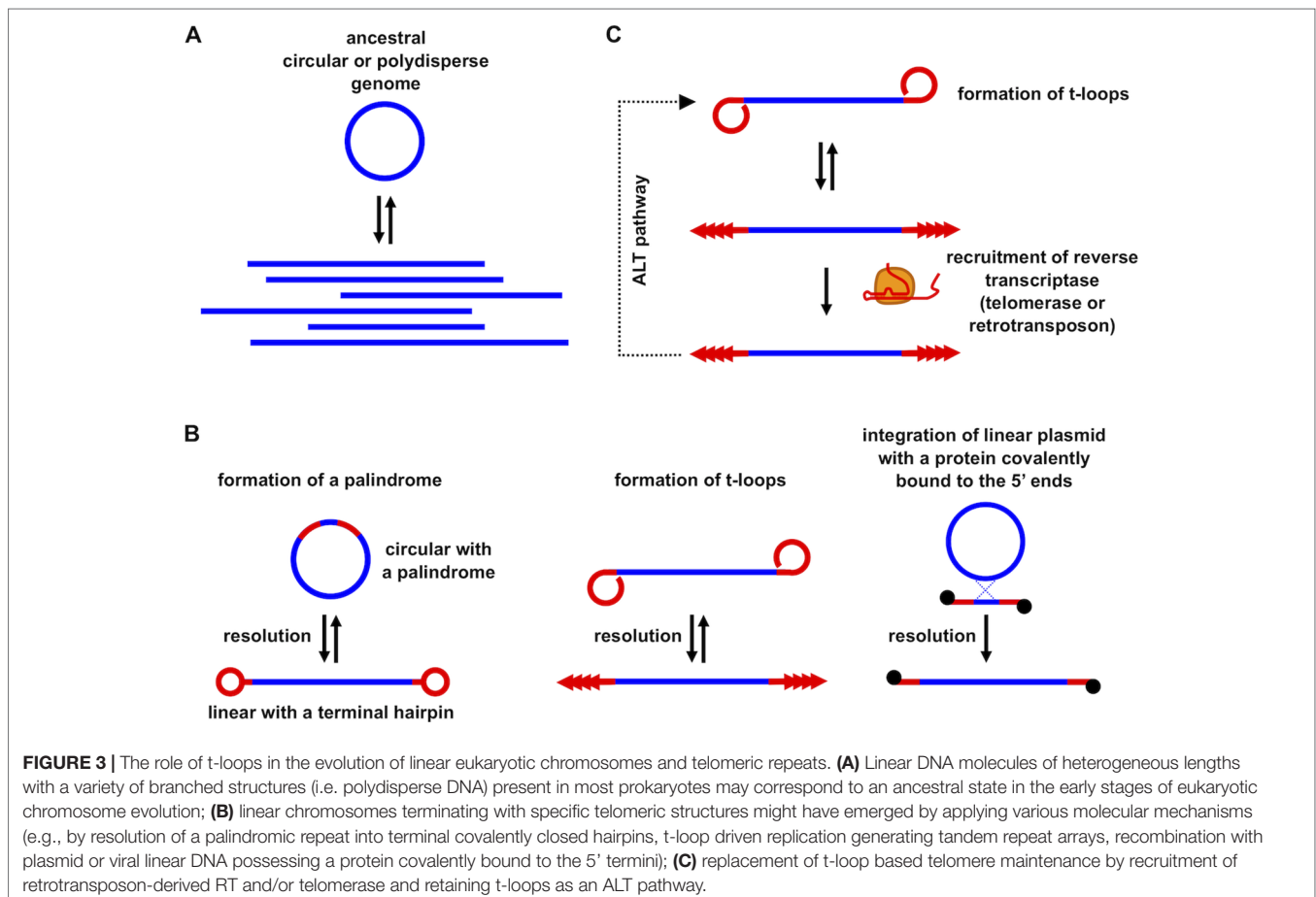
This appealing scenario also implies that the t-loop-based telomere formed on intronic repeats was replaced by modern telomeres by accommodation of a telomerase reverse transcriptase derived from a reverse transcriptase (RT) encoded by group II introns that gained the ability to use the 3' end of a chromosome as a primer for reverse transcription of its associated RNA, derived from another group II intron. This would be followed by the development of specific interactions between RT and RNA, the reduction in the RNA template region to a short sequence that will eventually become a terminal repeat and by recognition of all chromosomal ends by the same RT/RNA pair.

Although this is a plausible chain of events, the hypothesis of eukaryogenesis based on a massive invasion of group II introns has its opponents. For example, Poole (2006) argues that whereas selfish genetic elements (i.e., group II introns) are tolerated by sexual organisms, this is not the case for asexual species. Moreover, group II introns were repeatedly introduced into Archaea (i.e., Euryarcheota), apparently without massive expansion accompanied by linearization of their genomes (Simon et al., 2008). Hence, massive expansion of these introns might have occurred later in evolution and could be associated with increased host fitness, as it has been shown that spliceosomal introns play a role in survival of yeast cells during starvation (Morgan et al., 2019; Parenteau et al., 2019).

In addition to these observations that do not comply with the role of group II introns in the evolutionary emergence of linear chromosomes, the scenario proposed by de Lange (2015) implies that the evolution of telomeres must have proceeded in multiple steps and thus would require a substantial time. Why then are telomeres so conserved in such diverse eukaryotes as protists, fungi, plants, and animals [assuming that the exemptions like retrotransposons in *Drosophila* (Pardue and DeBaryshe, 2008), large tandem repeats in *Chironomus*, or extremely variable telomeric repeats in many yeast species (Lue, 2010; Cervenak et al., 2017) are secondary derivatives of the canonical TTAGGG-like repeats maintained by telomerase]?

If we disregard the possibility of a convergent evolution of modern telomeres that in independent phylogenetic branches yielded the same result, there are several possible scenarios. First, the early evolution of eukaryotes could have taken place in a distinct ecological niche for a period of time sufficient to slowly replace the ancestral telomeres [e.g., those depicted by de Lange (2015)] with their modern counterparts in a small population of cells representing the Last Universal Eukaryotic Ancestor (LECA) that subsequently spread around the globe. Second, telomerase could have been present in the pre-eukaryotic host cell (e.g., Nakamura and Cech, 1998) before acquisition of the α -proteobacterial symbiont and thus was preadapted for rapid complementation of the t-loop system of telomere maintenance. This would be in line with the hypothesis of Cavalier-Smith (2010) stating that linear chromosomes with telomeres existed before the endosymbiotic event leading to the origin of the eukaryotic cell. However, there is no report of linear bacterial or archaeal chromosomes containing telomerase-derived arrays at the ends, so this hypothesis is not supported by solid evidence.

A third scenario assumes that it might have been the other way around (Figure 3). As argued for a long time by Arnold Bendich, the genomes of prokaryotes, even though they exhibit circular physical maps (e.g., by restriction enzyme mapping or sequencing), are in fact linear and polydisperse *in vivo* (Bendich, 2001). In contrast to “true” linear genomes such as chromosomes



found in eukaryotic nuclei, they do not possess telomeres and thus are easy targets of DNA-damage response machinery and other DNA processing enzymes. The polyploid nature of prokaryotic genomes (Soppa, 2014) makes them suitable for evolutionary experimentation including tinkering with their topology. Indeed, there are numerous examples of Archaeal linear extragenomic genetic elements employing various strategies of maintaining their ends (Pina et al., 2014; Wang et al., 2015). It is conceivable that the pre-eukaryotic ancestor stabilized the ends through a protective structure similar to that present in several species of bacteria [e.g., *Agrobacterium*, *Borrelia*, *Coxiella*, *Streptomyces* (Volf and Altenbuchner, 2000)] or mitochondria in numerous modern eukaryotes (Nosek and Tomaska, 2003; Smith and Keeling, 2015). These examples illustrate that formation of a telomere was not a eukaryotic innovation but could occur quite frequently in prokaryotes. Although linear prokaryotic and organellar genomes employ several types of strategies of how to stabilize their DNA ends, one of the most similar to eukaryotic telomeres is based on very diverse sequences whose only common feature is the ability to replicate autonomously via a rolling circle mechanism generating a long array of tandem repeats (Nosek et al., 2005; Nosek et al., 2006; Tomaska et al., 2009). Autonomously replicating circular DNA molecules undergoing frequent excision–integration cycles enabling their expansion were observed in a wide range of cells (Cohen et al., 2007; Møller et al., 2015; Shoura et al., 2017), and may lead to a formation of tandem repeat arrays. If such arrays are integrated at the unprotected end of a DNA molecule, it would immediately provide a substrate for the formation of a t-loop, thus solving the end-protection and end-replication problem in a single step [for more details see, Nosek et al. (2006)]. Moreover, the tandem repeat sequences could undergo transcription, generating a population of RNA molecules similar to TERRA. These, as outlined above, could even pronounce the propensity of the sequence to form t-loops (Kar et al., 2016). After such a host cell engulfed the future mitochondrion, the invasion of group II introns would initiate eukaryogenesis (Martin and Koonin, 2006) and at the same time provide candidate RTs for future telomerase development. The only step required to transform an RT to a modern-type telomerase would be acquiring the ability to associate with RNAs derived from preexisting telomeres and employ a small sequence as a template and the rest of the molecule as a structural scaffold. The template sequence might have been short and conserved because it provided some particular structural features (e.g., formation of a protective

secondary structure such as the G-quadruplex), whereas the rest of the RNA molecule, as well as the proteins associated with telomeres and telomerase were much more prone to evolutionary changes (e.g., Fajkus et al., 2005; Linger and Price, 2009; Lue, 2010; Fulcher et al., 2014; Cervenak et al., 2017). Retainment of t-loops as a backup or complementary means for solution of end-protection and end-replication problems was dependent on the evolutionary trajectory of a particular phylogenetic line. For example, in macronuclei of ciliates with very short telomeric tracts, or in yeast species with short (300–500 bp) telomeric tracts containing heterogeneous repeats (e.g., *S. cerevisiae*, *S. pombe*), t-loops are difficult to form even from a mechanistic point of view (e.g., limitations associated with DNA bending), and their absence is compensated by other mechanisms such as amplification of subtelomeric regions, gene conversion, or HR [reviewed in Wellinger and Zakian (2012)]. However, even these organisms retained their capacity to form t-loops as exemplified by their observation in micronuclei of *Oxytricha* (Murti and Prescott, 1999) or the ability of *S. pombe* telomeric protein Taz1 to promote formation of t-loops *in vitro* (Tomaska et al., 2004).

In summary, the recent results indicating the participation of t-loops in various forms of DNA transactions shed additional light not only on the maintenance of telomeres but also has much broader implications ranging from the molecular basis of nucleotide repeat expansion to the evolution of telomeres.

AUTHOR CONTRIBUTIONS

AK carried out the replication experiment. JDG, JN, and LT developed the hypotheses and wrote the original draft. JDG, LT, JN, and SW prepared the later and final drafts of the work. JDG, AK, LT, and JN prepared the figures. All authors approved of the final work.

FUNDING

This work was supported by grants from the NIH (GM31819, ES013773) to JDG and APVV-15-0022 and VEGA 1/0052/16 (to LT) and APVV-14-0253 and VEGA 1/0027/19 (to JN).

ACKNOWLEDGMENTS

We wish to thank Charles C. Richardson and his laboratory for the T7 replication factors.

REFERENCES

- Arora, R., Lee, Y., Wischnewski, H., Brun, C. M., Schwarz, T., and Azzalin, C. M. (2014). RNaseH1 regulates TERRA-telomeric DNA hybrids and telomere maintenance in ALT tumour cells. *Nat. Commun.* 5, 5220. doi: 10.1038/ncomms6220
- Azzalin, C. M., Reichenbach, P., Khoraiuli, L., Giulotto, E., and Lingner, J. (2007). Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science* 318, 798–801. doi: 10.1126/science.1147182
- Barre, F.-X., Soballe, B., Michel, B., Aroyo, M., Robertson, M., and Sherratt, D. (2001). Circles: the replication–recombination–chromosome segregation connection. *Proc. Natl. Acad. Sci. U.S.A.* 98, 8189–8195. doi: 10.1073/pnas.111008998
- Bates, G. P., Dorsey, R., Gusella, J. F., Hayden, M. R., Kay, C., Leavitt, B. R., et al. (2015). Huntington disease. *Nat. Rev. Dis. Prim.* 1, 15005. doi: 10.1038/nrdp.2015.5
- Bendich, A. J. (2001). The form of chromosomal DNA molecules in bacterial cells. *Biochimie* 83, 177–186. doi: 10.1016/S0300-9084(00)01209-8
- Bryan, T. M., Englezou, A., Gupta, J., Bacchetti, S., and Reddel, R. R. (1995). Telomere elongation in immortal human cells without detectable telomerase activity. *EMBO J.* 14, 4240–4248. doi: 10.1002/j.1460-2075.1995.tb00098.x
- Cavalier-Smith, T. (2010). Origin of the cell nucleus, mitosis and sex: roles of intracellular coevolution. *Biol. Direct* 5, 7. doi: 10.1186/1745-6150-5-7

- Cervenak, F., Jurikova, K., Sepsiova, R., Nebohacova, M., Nosek, J., and Tomaska, L. (2017). Double-stranded telomeric DNA binding proteins: diversity matters. *Cell Cycle* 16, 1568–1577. doi: 10.1080/15384101.2017.1356511
- Cesare, A. J., Quinney, N., Willcox, S., Subramanian, D., and Griffith, J. D. (2003). Telomere looping in *P. sativum* (common garden pea). *Plant J.* 36, 271–279. doi: 10.1046/j.1365-313X.2003.01882.x
- Cesare, A. J., and Reddel, R. R. (2010). Alternative lengthening of telomeres: models, mechanisms and implications. *Nat. Rev. Genet.* 11, 319–330. doi: 10.1038/nrg2763
- Chan, Y. W., and West, S. (2015). GEN1 promotes Holliday junction resolution by a coordinated nick and counter-nick mechanism. *Nucleic Acids Res.* 43, 10882–10892. doi: 10.1093/nar/gkv1207
- Chu, H. P., Cifuentes-Rojas, C., Kesner, B., Aeby, E., Lee, H. G., Wei, C., et al. (2017). TERRA RNA antagonizes ATRX and protects telomeres. *Cell* 170, 86–101. doi: 10.1016/j.cell.2017.06.017
- Ciaccio, C., Fontana, L., Milani, D., Tabano, S., Miozzo, M., and Esposito, S. (2017). Fragile X syndrome: a review of clinical and molecular diagnoses. *Ital. J. Pediatr.* 43, 39. doi: 10.1186/s13052-017-0355-y
- Cohen, S., Houben, A., and Segal, D. (2007). Extrachromosomal circular DNA derived from tandemly repeated genomic sequences in plants. *Plant J.* 53, 1027–1034. doi: 10.1111/j.1365-313X.2007.03394.x
- Crosland, M. W., and Crozier, R. H. (1986). *Myrmecia pilosula*, an ant with only one pair of chromosomes. *Science* 231, 1278. doi: 10.1126/science.231.4743.1278
- de Lange, T. (2004). T-loops and the origin of telomeres. *Nat. Rev. Mol. Cell Biol.* 5, 323–329. doi: 10.1038/nrm1359
- de Lange, T. (2015). A loopy view of telomere evolution. *Front. Genet.* 6, 321. doi: 10.3389/fgene.2015.00321
- DeJesus-Hernandez, M., Mackenzie, I. R., Boeve, B. F., Boxer, A. L., Baker, M., Rutherford, N. J., et al. (2011). Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* 72, 245–256. doi: 10.1016/j.neuron.2011.09.011
- Deng, Z., Norseen, J., Wiedmer, A., Riethman, H., and Lieberman, P. M. (2009). TERRA RNA binding to TRF2 facilitates heterochromatin formation and ORC recruitment at telomeres. *Mol. Cell* 35, 403–413. doi: 10.1016/j.molcel.2009.06.025
- Doksani, Y., Wu, J. Y., de Lange, T., and Zhuang, X. (2013). Super-resolution fluorescence imaging of telomeres reveals TRF2-dependent T-loop formation. *Cell* 155, 345–356. doi: 10.1016/j.cell.2013.09.048
- Erdel, F., Kratz, K., Willcox, S., Griffith, J. D., Greene, E. C., and de Lange, T. (2017). Telomere recognition and assembly mechanism of mammalian shelterin. *Cell Rep.* 18, 41–53. doi: 10.1016/j.celrep.2016.12.005
- Fajkus, J., Sykorová, E., and Leitch, A. R. (2005). Telomeres in evolution and evolution of telomeres. *Chromosom. Res.* 13, 469–479. doi: 10.1007/s10577-005-0997-2
- Forster, P. (2013). Why are there so many diverse replication machineries? *J. Mol. Biol.* 425, 4714–4726. doi: 10.1016/j.jmb.2013.09.032
- Fulcher, N., Derboven, E., Valuchova, S., and Riha, K. (2014). If the cap fits, wear it: an overview of telomeric structures over evolution. *Cell. Mol. Life Sci.* 71, 847–865. doi: 10.1007/s00018-013-1469-z
- Garavis, M., González, C., and Villasante, A. (2013). On the origin of the eukaryotic chromosome: the role of noncanonical DNA structures in telomere evolution. *Genome Biol. Evol.* 5, 1142–1150. doi: 10.1093/gbe/evt079
- Garner, E., Kim, Y., Lach, F. P., Kottmann, M. C., and Smogorzewska, A. (2013). Human GEN1 and the SLX4-associated nucleases MUS81 and SLX1 are essential for the resolution of replication-induced Holliday junctions. *Cell Rep.* 5, 207–215. doi: 10.1016/j.celrep.2013.08.041
- Griffith, J. D. (2013). Many ways to loop DNA. *J. Biol. Chem.* 288, 29724–29735. doi: 10.1074/jbc.X113.515981
- Griffith, J. D., Comeau, L., Rosenfield, S., Stansel, R. M., Bianchi, A., Moss, H., et al. (1999). Mammalian telomeres end in a large duplex loop. *Cell* 97, 503–514. doi: 10.1016/S0092-8674(00)80760-6
- Ishikawa, F., and Naito, T. (1999). Why do we have linear chromosomes? A matter of Adam and Eve. *Mutat. Res.* 434, 99–107. doi: 10.1016/S0921-8777(99)00017-8
- Kar, A., Willcox, S., and Griffith, J. D. (2016). Transcription of telomeric DNA leads to high levels of homologous recombination and t-loops. *Nucleic Acids Res.* 44, 9369–9380. doi: 10.1093/nar/gkw779
- Kulczyk, A. W., and Richardson, C. C. (2016). The replication system of bacteriophage T7. *Enzymes* 39, 89–136. doi: 10.1016/bs.enz.2016.02.001
- Linger, B. R., and Price, C. M. (2009). Conservation of telomere protein complexes: shuffling through evolution. *Crit. Rev. Biochem. Mol. Biol.* 44, 434–446. doi: 10.3109/10409230903307329
- Liquori, C. L., Ricker, K., Moseley, M. L., Jacobsen, J. F., Kress, W., Naylor, S. L., et al. (2001). Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. *Science* 293, 864–867. doi: 10.1126/science.1062125
- Lue, N. F. (2010). Plasticity of telomere maintenance mechanisms in yeast. *Trends Biochem. Sci.* 35, 8–17. doi: 10.1016/j.tibs.2009.08.006
- Lue, N. F. (2018). Evolving linear chromosomes and telomeres: a C-strand-centric view. *Trends Biochem. Sci.* 43, 314–326. doi: 10.1016/j.tibs.2018.02.008
- Luo, J., Sun, X., Cormack, B. P., and Boeke, J. D. (2018). Karyotype engineering by chromosome fusion leads to reproductive isolation in yeast. *Nature* 560, 392–396. doi: 10.1038/s41586-018-0374-x
- Martin, W., and Koonin, E. V. (2006). Introns and the origin of nucleus-cytosol compartmentalization. *Nature* 440, 41–45. doi: 10.1038/nature04531
- Møller, H. D., Parsons, L., Jørgensen, T. S., Botstein, D., and Regenberg, B. (2015). Extrachromosomal circular DNA is common in yeast. *Proc. Natl. Acad. Sci. USA* 112, E3114–E3122. doi: 10.1073/pnas.1508825112
- Morgan, J. T., Fink, G. R., and Bartel, D. P. (2019). Excised linear introns regulate growth in yeast. *Nature* 565, 606–611. doi: 10.1038/s41586-018-0828-1
- Murti, K. G., and Prescott, D. M. (1999). Telomeres of polytene chromosomes in a ciliated protozoan terminate in duplex DNA loops. *Proc. Natl. Acad. Sci. USA* 96, 14436–14439. doi: 10.1073/pnas.96.25.14436
- Nakamura, T. M., and Cech, T. R. (1998). Reversing time: origin of telomerase. *Cell* 92, 587–590. doi: 10.1016/S0092-8674(00)81123-X
- Nosek, J., Kosa, P., and Tomaska, L. (2006). On the origin of telomeres: a glimpse at the pre-telomerase world. *BioEssays* 28, 182–190. doi: 10.1002/bies.20355
- Nosek, J., Rycovska, A., Makhov, A. M., Griffith, J. D., and Tomaska, L. (2005). Amplification of telomeric arrays via rolling-circle mechanism. *J. Biol. Chem.* 280, 10840–10845. doi: 10.1074/jbc.M409295200
- Nosek, J., and Tomaska, L. (2003). Mitochondrial genome diversity: evolution of the molecular architecture and replication strategy. *Curr. Genet.* 44, 73–84. doi: 10.1007/s00294-003-0426-z
- Palm, W., and de Lange, T. (2008). How shelterin protects mammalian telomeres. *Annu. Rev. Genet.* 42, 301–334. doi: 10.1146/annurev.genet.41.110306.130350
- Pardue, M.-L., and DeBaryshe, P. G. (2008). Drosophila telomeres: a variation on the telomerase theme. *Fly (Austin)* 2, 101–110. doi: 10.4161/fly.6393
- Parenteau, J., Maignon, L., Berthoumieux, M., Catala, M., Gagnon, V., and Abou Elela, S. (2019). Introns are mediators of cell response to starvation. *Nature* 565, 612–617. doi: 10.1038/s41586-018-0859-7
- Pina, M., Basta, T., Quax, T. E. F., Joubert, A., Baconnais, S., Cortez, D., et al. (2014). Unique genome replication mechanism of the archaeal virus AFV1. *Mol. Microbiol.* 92, 1313–1325. doi: 10.1111/mmi.12630
- Poole, A. M. (2006). Did group II intron proliferation in an endosymbiont-bearing archaeon create eukaryotes? *Biol. Direct* 1, 36. doi: 10.1186/1745-6150-1-36
- Raices, M., Verdun, R. E., Compton, S. A., Haggblom, C. I., Griffith, J. D., Dillin, A., et al. (2008). *C. elegans* telomeres contain G-strand and C-strand overhangs that are bound by distinct proteins. *Cell* 132, 745–757. doi: 10.1016/j.cell.2007.12.039
- Rass, U., Compton, S. A., Matos, J., Singleton, M. R., Ip, S. C. Y., Blanco, M. G., et al. (2010). Mechanism of Holliday junction resolution by the human GEN1 protein. *Genes Dev.* 24, 1559–1569. doi: 10.1101/gad.585310
- Renton, A. E., Majounie, E., Waite, A., Simón-Sánchez, J., Rollinson, S., Gibbs, J. R., et al. (2011). A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* 72, 257–268. doi: 10.1016/j.neuron.2011.09.010
- Schoeftner, S., and Blasco, M. A. (2008). Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. *Nat. Cell Biol.* 10, 228–236. doi: 10.1038/ncb1685
- Shao, Y., Lu, N., Wu, Z., Cai, C., Wang, S., Zhang, L.-L., et al. (2018). Creating a functional single-chromosome yeast. *Nature* 560, 331–335. doi: 10.1038/s41586-018-0382-x
- Shoura, M. J., Gabdank, I., Hansen, L., Merker, J., Gotlib, J., Levene, S. D., et al. (2017). Intricate and cell type-specific populations of endogenous circular DNA (eccDNA) in *Caenorhabditis elegans* and *Homo sapiens*. *G3-Genes Genom. Genet.* 7, 3295–3303. doi: 10.1534/g3.117.300141
- Simon, D. M., Clarke, N. A. C., McNeil, B. A., Johnson, I., Pantuso, D., Dai, L., et al. (2008). Group II introns in Eubacteria and Archaea: ORF-less introns and new varieties. *RNA* 14, 1704–1713. doi: 10.1261/rna.1056108
- Slijepcevic, P. (2016). Mechanisms of the evolutionary chromosome plasticity: integrating the ‘centromere-from-telomere’ hypothesis with telomere length regulation. *Cytogenet. Genome Res.* 148, 268–278. doi: 10.1159/000447415

- Smith, D. R., and Keeling, P. J. (2015). Mitochondrial and plastid genome architecture: reoccurring themes, but significant differences at the extremes. *Proc. Natl. Acad. Sci. USA* 112, 10177–10184. doi: 10.1073/pnas.1422049112
- Soppa, J. (2014). Polyploidy in Archaea and Bacteria: about desiccation resistance, giant cell size, long-term survival, enforcement by a eukaryotic host and additional aspects. *J. Mol. Microbiol. Biotechnol.* 24, 409–419. doi: 10.1159/000368855
- Stansel, R. M., de Lange, T., and Griffith, J. D. (2001). T-loop assembly *in vitro* involves binding of TRF2 near the 3' telomeric overhang. *EMBO J.* 20, 5532–5540. doi: 10.1093/emboj/20.19.5532
- Thys, R. G., and Wang, Y.-H. (2015). DNA replication dynamics of the GGGGCC repeat of the *C9orf72* gene. *J. Biol. Chem.* 290, 28953–28962. doi: 10.1074/jbc.M115.660324
- Tomaska, L., Makhov, A. M., Griffith, J. D., and Nosek, J. (2002). t-Loops in yeast mitochondria. *Mitochondrion* 1, 455–459. doi: 10.1016/S1567-7249(02)00009-0
- Tomaska, L., Nosek, J., Kramara, J., and Griffith, J. D. (2009). Telomeric circles: universal players in telomere maintenance? *Nat. Struct. Mol. Biol.* 16, 1010–1015. doi: 10.1038/nsmb.1660
- Tomaska, L., Willcox, S., Slezakova, J., Nosek, J., and Griffith, J. D. (2004). Taz1 binding to a fission yeast model telomere: formation of telomeric loops and higher order structures. *J. Biol. Chem.* 279, 50764–50772. doi: 10.1074/jbc.M409790200
- Villasante, A., Abad, J. P., and Mendez-Lago, M. (2007). Centromeres were derived from telomeres during the evolution of the eukaryotic chromosome. *Proc. Natl. Acad. Sci. U.S.A.* 104, 10542–10547. doi: 10.1073/pnas.0703808104
- Volf, J.-N., and Altenbuchner, J. (2000). A new beginning with new ends: linearisation of circular chromosomes during bacterial evolution. *FEMS Microbiol. Lett.* 186, 143–150. doi: 10.1111/j.1574-6968.2000.tb09095.x
- Wang, H., Peng, N., Shah, S. A., Huang, L., and She, Q. (2015). Archaeal extrachromosomal genetic elements. *Microbiol. Mol. Biol. Rev.* 79, 117–152. doi: 10.1128/MMBR.00042-14
- Wellinger, R. J., and Zakian, V. A. (2012). Everything you ever wanted to know about *Saccharomyces cerevisiae* telomeres: beginning to end. *Genetics* 191, 1073–1105. doi: 10.1534/genetics.111.137851
- Wells, R. D., and Ashizawa, T. (2006). *Genetic instabilities and neurological diseases*. San Diego, CA: Elsevier Academic Press.
- Wyatt, H. D. M., Sarbajna, S., Matos, J., and West, S. C. (2013). Coordinated actions of SLX1-SLX4 and MUS81-EME1 for Holliday junction resolution in human cells. *Mol. Cell* 52, 234–247. doi: 10.1016/j.molcel.2013.08.035
- Wyatt, H. D. M., and West, S. C. (2014). Holliday junction resolvases. *Cold Spring Harb. Perspect. Biol.* 6, a023192–a023192. doi: 10.1101/cshperspect.a023192
- Yeeles, J. T., Deegan, T. D., Janska, A., and Diffley, J. F. X. (2015). Regulated eukaryotic DNA replication origin firing with purified proteins. *Nature* 519, 431–435. doi: 10.1038/nature14285

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Tomaska, Nosek, Kar, Willcox and Griffith. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.