

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/108191/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Baird, Duncan M. 2018. Telomeres and genomic evolution. *Philosophical Transactions B: Biological Sciences* 373 (1741) , 20160437. 10.1098/rstb.2016.0437

Publishers page: <http://dx.doi.org/10.1098/rstb.2016.0437>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



Telomeres and genomic evolution

Duncan M. Baird

Division of Cancer & Genetics, School of Medicine, Cardiff University, Heath Park, Cardiff,
CF14 4XN, UK. ORCID ID: 0000-0001-8408-5467

Corresponding author: bairddm@cardiff.ac.uk

Keywords: Telomere, DNA repair, genome, evolution

Abstract

The terminal regions of eukaryotic chromosomes, composed of telomere repeat sequences and sub-telomeric sequences, represent some of the most variable and rapidly evolving regions of the genome. The sub-telomeric regions are characterised by segmentally duplicated repetitive DNA elements, interstitial telomere repeat sequences and families of variable genes. Sub-telomeric repeat sequence families are shared amongst multiple chromosome-ends, often rendering detailed sequence characterisation difficult. These regions are composed of constitutive heterochromatin and are subjected to high-levels of meiotic recombination. Dysfunction within telomere repeat arrays, either due to disruption in the chromatin structure, or because of telomere shortening, can lead to chromosomal fusion and the generation of large-scale genomic rearrangements across the genome. The dynamic nature of telomeric regions therefore provides functionally useful variation to create genetic diversity, but also provides a mechanism for rapid genomic evolution that can lead to reproductive isolation and speciation.

Telomeres and the evolution of the cancer genome

It is well known that cancer arises because of acquired somatic mutation. This creates the genetic diversity upon which Darwinian selection can operate to drive clonal evolution within the cellular ecosystem of multicellular organisms. Multiple mutational mechanisms generate cellular genetic diversity, including the action of APOBEC cytidine deaminases, exposure to mutagens, DNA replication and age, all creating distinctive mutational signatures [1]. In addition to this mutational burden, large-scale genomic rearrangements are frequently detected and represent a hallmark of cancer cells [2]. Cancers exhibiting karyotypic complexity display a more aggressive and rapidly evolving phenotype that confers a poorer prognosis [3]. Telomere dysfunction is one key mechanism that is considered to drive karyotypic complexity in cancer. Short telomeres, relative to the specific normal somatic tissues from which tumours are derived, have been identified in the majority of tumour types, both solid and haematological [4]. Mutations leading to the upregulation of telomerase activity are required for malignant progression in at least 85% of tumours [5], with the remaining tumours adopting the alternative lengthening of telomeres phenotype [6]. Murine models with a combined lack of telomerase and Tp53 demonstrate that telomere dysfunction can drive the types of genomic rearrangements, such as non-reciprocal translocations, that are frequently observed in human cancer [7]. Cancers with both short telomeres and telomere fusion exhibit higher levels of genomic complexity [8-10]. Importantly also, patients with tumours that display telomeres within the length ranges at which fusion occurs, display a much poorer prognosis [11-14] and an impaired response to genotoxic chemotherapeutics [15]. Taken together this body of evidence is consistent with the view that telomere-driven mutation can be one contributor to the overall genetic heterogeneity of cancer, that facilitates clonal progression and the evolution of the cancer genome. In this article, I consider how our increasingly detailed understanding of telomere biology, primarily in the context of cancer but also human genetic disease, may inform about the role that telomeres may play in modulating the evolution of the genome across species.

Telomere erosion and senescence.

The DNA sequence that constitutes the terminus of linear eukaryotic chromosomes was first characterised in *Tetrahymena* to be the hexameric nucleotide repeat sequence TTGGGG, tandemly repeated into variable lengthened arrays of between 20-70 repeats [16]. In the

subsequent years, it has become apparent that the chromosomes in the majority of eukaryotes terminate with short G-rich repetitive sequences orientated 5' to 3' toward the terminus. The short repeat structure containing a group of 2-4 guanines appears to be largely conserved between taxonomic groupings, for example all vertebrates utilise the sequence TTAGGG [17]. Whereas within the majority of plant species TTTAGGG is utilised, with some exceptions for example, *Nicotiana tabacum* (TTAGGG) and *Alliums* (CTCGGTTATGGG). The chromosomal terminus is not blunt ended, but is instead a single-stranded 3' overhang, composed in human cells of 200-300 nt of TTAGGG repeats [18-20]. It is considered that this overhang is capable of folding back and annealing within proximal regions of the telomere repeat array, to form the T-loop structure, that effectively sequesters the natural end of the chromosome and prevents its recognition as a double-stranded DNA break (DSB)[21]. The only exceptions to the small repeat unit paradigm for telomeric structure are the dipterans, which utilise retro-transposable elements that specifically transpose to the chromosomal termini [22].

Telomere repeat sequences are synthesised *de novo* by the specialised reverse transcriptase, telomerase, that utilises an RNA template to catalyse the addition of telomere repeats at the chromosomal terminus [23]. This counteracts the ongoing loss of terminal sequences that occurs due the end-replication problem as cells divide [24], thereby defining one of the key functions of telomeres (Figure 1). A further function of telomeres is to protect the natural chromosome end from recognition as a DSB and to prevent aberrant DNA repair activities. This 'end-capping' function is mediated via the specialised chromatin structure of the telomere, which is composed of telomere repeat specific binding proteins and associated proteins, collectively referred to as the 'Shelterin complex' [21]. Abrogation of TRF2 function, a key component of the Shelterin complex, leads to an immediate and catastrophic chromosomal fusion phenotype, with chromosomes joined telomere-to-telomere [25, 26](Figure 2A).

Telomerase expression is not consistent across all tissues within the same organism and has species-specific expression profiles [27]. In general, longer lived, larger species, such as humans, have a tendency towards telomerase repression in the majority of somatic tissues and relatively short mean telomere lengths that are typically less than 20 kb. Which contrasts

with shorter lived, smaller organisms, where telomerase is active and the telomeres are longer [28, 29]. Thus in longer lived species, ongoing cell division throughout life, results in progressive telomere shortening as a function of age (Figure 1). *In vitro* cell culture allows cells to be passaged to the end of their natural replicative lifespan, at this point telomeres can lose their end-capping function. This can lead to various outcomes depending on the cellular context and the nature by which end-capping is lost. In normal human fibroblast cells, gradual telomere erosion ultimately results in the partial loss of end-capping, whereby cells undergo a Tp53-dependent G1/S cell cycle arrest, referred to as replicative senescence. The repression of telomerase, together with a telomere-dependent limit to replicative lifespan is considered to provide a stringent tumour suppressive mechanism in long-lived species [30-32]. The corollary of which, is that whilst the telomeres of senescent cells illicit a cell-cycle arrest, it does not result in the repair of telomeres. Thus replicative senescence is permanent and cells can remain in a metabolically active, but non-dividing, state for many years. The phenotype of senescent cells can be distinct from younger cells, potentially becoming more catabolic and actively degrading the tissue matrix in which they reside [33] and they can acquire the pro-inflammatory senescence associated secretory phenotype that can promote tumour progression [34]. The accumulation of senescent cells as a function of age, may therefore lead to a loss of tissue homeostasis and frailty, that may underpin the ageing processes (discussed elsewhere in this issue, [35]). Thus, telomere erosion may provide a tumour suppressive mechanism during the reproductive years, yet the same process may underpin the ageing process in long lived species.

Telomere-driven chromosomal instability

Whilst normal cells undergo replicative senescence in response to short telomeres, cells that are compromised in their ability to respond to DNA-damage, for example via mutations within components of the Tp53 and RB pathways, can continue to divide and undergo further telomere erosion [36]. In this situation telomeres can erode to a length at which they are almost completely denuded of telomere repeats and will no longer be capped by the Shelterin complex [37]. This complete uncapping results in chromosome-ends that are recognised and processed by the cellular DSB repair mechanisms leading to the fusion of telomeres (Figure 2B and 2C) [36-38]. In human cells the characterisation of the DNA sequences across telomere fusion breakpoints, reveals a mean of just 5 TTAGGG repeats at the break point and that

fusion displays a distinctive mutational profile [37]. This profile includes micro-homology at the fusion point and that one, or both, of the participating telomeres is subjected to extensive processing, resulting in deletions extending several kilobases into the sub-telomeric DNA [37-39]; a similar profile was also identified in *Arabidopsis* mutants lacking telomerase and Ku70 [40]. This characteristic mutational signature, of extensive deletion and micro-homology at the fusion points, implicates the more error-prone, alternative form of non-homologous end joining (A-NHEJ) as the predominate DNA repair mechanism catalysing the fusion of short dysfunctional telomeres. Indeed, it is apparent that if the classical-NHEJ pathway is abrogated, for example by the deletion of LIG4, then short telomeres are still capable of undergoing fusion, consistent with the utilisation of A-NHEJ in this process [41-44]. The removal of specific Shelterin components also reveals the relative roles of C- and A-NHEJ in telomere fusion. Following the loss of TRF2 function telomere fusions are mediated by C-NHEJ [25, 45]; whereas following the loss of POT1 function, fusion is mediated by A-NHEJ [45]. Telomeres are subject to both C- and A-NHEJ mediated repair following the removal of both TRF1 and TRF2 [46]; thus, the involvement of each pathway appears to be dependent on how telomeres are rendered dysfunctional. The error-prone nature of A-NHEJ is consistent with the sub-telomeric deletion events, presumably generated by nucleolytic resection, that accompanies telomere fusion. The distribution of fusion points, with respect to the start of the telomere repeat arrays, indicates that deletion may be much more extensive, with the potential to extend into the coding regions of the chromosome [37, 39]. Thus, the processing of short dysfunctional telomeres, via the A-NHEJ pathway, may result in sub-telomeric resection that is sufficient to result in the deletion of distal genes (Figure 2D).

Telomere fusion analysis shows that telomeres can be subjected to fusion with other shortened telomeres within the cell [37]; leading to the formation of dicentric chromosomes and the initiation of cycles of breakage, fusion and anaphase-bridging (BFB) [47], that can drive large-scale genome rearrangements, of the kind frequently observed in human cancer [7]. Telomere fusion preferentially occurs between newly replicated sister chromatids; in this situation, the dicentric chromatid is broken during anaphase, yielding one chromatid with a potentially large deletion and one with a large inverted repeat. The daughter cell with the inverted repeat now contains duplicated copies of genes within the repeat region, with a gene copy number of 3; further cycles of BFB will result in further gene amplifications (Figure 2B).

These BFB cycles can only be prevented by chromosome healing via the synthesis of a new telomere on the broken end via the action of telomerase, or by recombination with a pre-existing telomere [48-50](Figure 2D). Inter-chromosomal telomere-telomere fusion can occur at any stage of the cell cycle. Following replication, these structures can form a double anaphase-bridge that, depending on how the centromeres are resolved and the position of the subsequent breakage, will result in daughter cells with a non-reciprocal translocation (NRT), or a deletion (Figure 2C), additional cycles of BFB will further perpetuate the state of genome instability. BFBs initiated by either inter-chromosomal or sister-chromatid telomere fusion, will result in the internalisation of telomere repeat sequences to form new interstitial telomere repeat sequences (ITS; Figure 2B and 2C).

The state of wide-spread genome instability induced following telomere dysfunction is referred to as 'crisis' during which telomere fusions initiate mitotic arrest that leads to apoptotic cell death [51]. Cell death progressively increases to a point at which it exceeds cell growth and the culture crashes. Rarely cells can escape crisis, but only following the reestablishment of a telomere maintenance mechanism, such as the upregulation of telomerase activity, that allows the genome to be stabilised and the outgrowth a clonal derivative with a highly rearranged genome emerges [43]. There are good data from human cancers demonstrating telomere erosion to the lengths observed in cells undergoing crisis in culture, as well as telomere fusion and evidence of anaphase-bridging [8, 9, 52-55]. Whilst a telomere driven crisis observed in *in vitro* cell culture experiments represents an extreme situation, the observations in cancers indicate that the malignant progression requires a stage akin to a telomeric crisis. The BFB cycles that may arise, following telomere fusion, will lead to cells containing chromosomes with large deletions, non-reciprocal translocations, inversions, duplications and the generation of interstitial telomeric sequences (ITS).

In addition to fusion with other telomeres and the initiation of BFB cycles, it is also apparent that short dysfunctional telomeres can be subjected to fusion with non-telomeric loci [39, 42]. Interestingly these fusion events can be detected, captured between sister chromatids, from just a single dysfunctional telomere, in the context of an otherwise functional telomere component. These often-complex events, appear to involve the ligase 4 dependent classical-NHEJ pathway and can involve multiple loci across the genome, providing a direct mechanism

for mutation. If these mutational events result in translocation of a functional telomere, then the resulting chromosome will be stable and thus provides a mechanism for direct mutation without the initiation of BFB cycles.

Chromosome healing and recombination

Meiotic recombination rates in humans, are not linear throughout the chromosome; instead rates are at their lowest across centromeres and increase consistently to the sub-telomeric regions of the chromosomes, leading to a linkage map expansion at the telomeres whereby the genetic distance between markers increases relative to the physical distance [56]. These regions of the genome are characterised by variable satellite and mini-satellite repeat sequences that are often associated with recombination hotspots [57, 58]. These sub-telomeric repetitive sequences comprise 80% of the 100 kb of DNA adjacent to telomeres [59]. They are considered to arise via NHEJ mediated translocations, followed by ectopic meiotic recombination and gene conversion between non-homologous chromosome ends. This allows for the distribution and homogenisation of sub-telomeric repeat sequences shared amongst multiple chromosome ends in the human genome [60]. This sub-telomeric recombination activity renders these regions of the genome highly dynamic and variable; a property that may have been selected for to generate genetic diversity in the human population. A sub-class of the Wiscott-Aldrich Syndrome Protein family (WASH) are the most distal protein coding RNA transcripts identified in the human genome, with transcripts terminating just 5 kb from the telomere in some individuals [59, 61]. Due to their sub-telomeric location gene dosage and distribution of the WASH family varies widely both within and between species [59]. This gene family appears to be involved the organisation of the actin cytoskeleton in response to various extracellular stimuli; why such diversity in this gene family is maintained is not clear, but they may contribute to host responses to pathogen infection and phenotypic diversity [61]. Other sub-telomeric gene families include a sub-set of the large olfactory receptor gene family, which also display a polymorphic distribution in the human population [62]. The phenotypic consequences are not entirely clear, but diversity in both sequence and copy number of these receptors could alter specificity and sensitivity to particular odorants [63].

The mechanisms that provide potentially adaptive genotypic and phenotypic diversity in the population, can also result in deleterious genetic rearrangements. Chromosome healing is a process that can result in *de novo* telomere formation at DSBs, this can occur by the action of telomerase or the capturing of a pre-existing telomere by recombination [64]. Telomerase activity is controlled by Pif1, a helicase that negatively regulates telomerase processivity and prevents aberrant telomerase activity at non-telomeric DSBs [65]. Chromosome healing has been observed in yeast and *Tetrahymena* [66], as well in several genetic conditions in humans that result from chromosomal terminal deletions 'healed' by *de novo* telomere formation, either by the action of telomerase [67, 68], or by acquiring a pre-existing telomere by recombination with sub-telomeric sequences or break induced replication [69, 70](Figure 2D). Numerous sub-telomeric deletions and translocations have been documented that lead to genetic disease, including mental retardation [71], holoprosencephaly [67] and facioscapulohumeral muscular dystrophy [72]. Many of these conditions arise from deletion and genetic rearrangements, but also changes to the chromosomal location with respect to telomeres, that may cause telomere position effects on gene expression [73].

Therefore, whilst the sub-telomeric telomere regions of human chromosomes are amongst the most genetically dynamic regions of the genome, providing potentially useful adaptive variation, their dynamic nature can in turn have detrimental mutational effects.

Stochastic telomeric deletion

Superimposed on gradual end-replication losses as a function of cell division, are mechanisms that generate large telomeric deletion events. Observed initially in *Trypanosoma brucei* [74], human cancer cell lines, then later in yeast models [75, 76] and *C.elegans* [77], these deletion events are also detected in normal human somatic cells and tissues [78, 79] as well as in the human male germline [80]. These events lead to a single chromosome end with a telomere almost completely denuded of telomere repeats [79], or the complete loss of the telomere and the generation of sub-telomeric double stranded DNA break (DMB unpublished observations; Figure 1). These apparently stochastic telomere deletion events, can result in short dysfunctional telomeres, in otherwise normal cells, that can then be subjected to DNA repair activity and fusion [37]. The mechanisms by which large-scale telomeric deletion occurs are not clear, however they may be related to difficulties in replication in these regions.

Telomere repeat sequences can form stable higher-order structures such as G-quadruplexes [81] and T-loops [21], these are likely to be difficult to replicate and require specialised helicase activities such as WRN and RTEL1 to unwind these structures prior to replication [82, 83]. The resolution of stalled replication forks within telomeres may result in a telomeric deletion event and fusion; consistent with this, telomeres have been identified as potential fragile sites arising as a consequence of replication stress [84]. The deleted telomeres may then be subjected to further DNA repair activity, leading to the same mutational mechanisms described above (Figure 2), including BFB cycles, sub-telomeric deletion and de novo telomere formation. Thus stochastic telomere deletion provides a telomere-driven mutational mechanism that can bypass the various DNA damage response apparatus in normal cells, even in the presence of telomerase.

Can telomere-driven mutation facilitate genomic evolution?

We have generated an increasingly detailed understanding of the mechanisms underlying telomere dysfunction and the role that this plays in driving mutation both within the cancer genome and genetic disease in humans. Can we apply this knowledge to understand how telomere biology may facilitate the evolution of the genome within and between species?

Sub-telomeric sequences – rapid evolution

The repetitive and variable nature of sub-telomeric sequences appears to be consistent across the majority of eukaryote species, however the specific sequences involved are less conserved. Comparisons of the karyotypes of the great ape species show distinct additional G-bands at the terminal regions of approximately half of the telomeres in chimpanzee karyotypes and nearly all the telomeres in the gorilla karyotypes, but are absent in the human and orangutan karyotypes [85]. The use of a sub-telomeric cloning and sequencing strategy in chimpanzees identified sub-telomeric sequences that showed no similarity with human sequences characterised using the same approaches [86], but instead identified a 32bp A-T rich repetitive sequence (5'-GATATTTCCATGTTTATACAGATAGCGGTGTA-3')[87]. Southern hybridisation and PCR analysis showed that this sequence was also present, but at higher levels, in the gorilla genome, but was absent from the human and orangutan genomes. These repetitive arrays may therefore account for the additional G-bands at the end of the chimpanzee and gorilla karyotypes. A recent 'Blast' search using this sequence against the

completed genomes of the great apes, shows that both the chimpanzee and gorilla have abundant copies of this repeat sequence, but it is entirely absent from the genome databases of the human and orangutan genomes (DMB unpublished observation). Based on our current understanding of the great ape phylogeny, it is apparent that over a relatively short period of evolutionary time, this 32bp repeat sequence has been generated and propagated in the ancestral genome to chimpanzee, gorillas and humans, but has been entirely eradicated from the human genome. Thus the location of the start of the telomere repeat arrays in the great ape species are likely to be unique to each species. Consistent with this, sequence analysis of the telomere-adjacent DNA of a single chromosome end in the great apes, revealed that the more ancestral telomere-adjacent sequence was observed in orangutans, which had been truncated in chimpanzee and replaced by the 32bp repeat sequences and in human by a new telomere repeat array, potentially via a telomerase mediated chromosome healing event [88].

The virulence genes of *Plasmodium falciparum*, *Trypanosoma brucei* and *cruzi* and *Pneumocystis carinii* are found in the sub-telomeric regions [89, 90], frequent ectopic recombination in these regions is considered to increase variability and diversity. Moreover as the telomeres in *T. brucei* are subjected to large sporadic deletion events [74], it has been considered that if these deletion events occur in telomeres that are already short, then the resulting DSBs may occur within the sub-telomeric genes. The resulting repair via break induced replication, may result in conversion between virulence genes creating an antigenic switch [91].

Sub-telomeric variation has been exploited in many other species to facilitate genetic diversity and adaptation; this includes the genes involved in sugar metabolism in yeast species [92], as well as the avirulence genes of the rice blast fungus [93].

The occurrence of variable repeat sequence families is a generalised property of sub-telomeric regions across a broad range of species [94-98]. It appears that these regions of the genome have a propensity to accumulate repetitive DNA, with ectopic recombination being invoked as the key driver of this variation. However, it is also apparent that telomeric deletion, both within the telomere repeat array, but also extending into the sub-telomeric regions,

followed by DNA repair activity to create translocations, gene conservations or to seed new telomeres, contributes further to this diversity. Importantly also, these mechanisms lead to rapid changes over short periods of evolutionary time, in the organisation and sequence composition of chromosome ends, resulting in species specific structures and positioning of telomere repeat arrays.

Interstitial telomeres

Most vertebrate species studied have chromosomes in which there are non-telomeric loci that contain interstitial telomere repeats. Using *in situ* hybridisation Meyne *et al* undertook a comprehensive analysis of telomere repeat distributions across a wide range of species [99]. They could detect ITS in 55 of 100 species analysed, with 44 of these containing 3 or more ITS, and some contained chromosomes with extensive telomere repeat hybridisation patterns across specific chromosomes, for the example, the Y chromosomes in Bennett's wallaby and African elephant [99]. The remaining 45 species displayed a telomere only hybridisation signal, however *in situ* hybridisation is limited in resolution and it is difficult to detect blocks of telomere repeats less than 1 kb and thus the full extent of ITS will be underestimated; this will become more apparent with the completion of the whole genome sequences from a broader range of species. Many subsequent studies have extended our knowledge of the distribution of ITS across a broad range of species including fish, amphibians and reptiles [100-102]. The chromosomal distribution of ITS provide markers to document the evolutionary history of the genome in which they reside. In some situations ITS are clearly derived from telomere fusion events, an example of which is the human chromosome 2 which is derived from a telomere fusion between two ancestral chromosomes, creating a dicentric chromosome that was stabilised following the inactivation one of the two centromeres (Figure 2E)[103]. The sequence of the fusion event that gave rise to human chromosome 2, is indistinguishable to the types of fusion observed in experimental systems, or directly in human cancers and is consistent with a NHEJ mediated fusion [39, 103].

An even more dramatic example of telomere fusion mediated events creating ITS, is observed in Muntjac deer chromosomes. Muntjac exhibit considerable variation in chromosomal complement, at the extremes of the distribution is the Indian muntjac whose genome is comprised of 6-7 chromosomes, whereas the genome of the closely related Chinese muntjac

contains 46 chromosomes, which is more typical of mammalian genomes [104]. Despite this large karyotypic difference, both genomes are of a similar size and the species can hybridise [105]; other Muntjac species display intermediate chromosome numbers. Cytogenetic analysis of the Indian muntjacs revealed multiple ITS adjacent to blocks of centromeric heterochromatin [106]. Long-range mapping and sequencing across these ITS sites reveals pattern of telomere fusion between telocentric chromosomes, in the same head-to-tail orientation [107]. It is proposed that the ancestral genome contained 70 chromosomes and 29 fusion events were required to reduce the genome to the complement of chromosomes observed in the extant population [107], over very a short period of evolutionary time [108]. The mechanisms underlying such a large-scale series of telomere fusions are not clear, but the dramatic scale of the events that led to this genome might be consistent with a single catastrophic chromosomal fusion event, that is reminiscent of the end-to-end fusion of chromosomes observed following the loss of TRF2 function [25] (Figure 2A). These types of large-scale genomic rearrangements, have been long associated with chromosomal speciation [109]. The subsequent fixation of these events in the population may be driven by a combination of; immediate selective advantage because of change in gene expression associated with a specific chromosomal rearrangement; meiotic drive, creating an imbalance in chromosomal segregation; suppression of recombination to counter act the detrimental effects of heterozygosity [110, 111].

ITS sequences are frequently identified in centromeric regions of metacentric and sub-metacentric chromosomes, for example in several lizard and snake species [102]. These are likely consistent with a Robertsonian fusion of telocentric chromosomes, resulting the internalisation of telomere sequences adjacent to the centromere, as documented in Brazilian geckos [112]. Centromere adjacent ITS sequences observed on acrocentric chromosomes are often large [102] and this might arise because of repeat expansion in these regions. Pericentromeric regions can be transcriptionally active creating non-coding RNA species that contribute to the chromatin structure of these regions. Reverse transcription of these non-coding RNAs to DNA (RNA-derived DNA, rdDNA) can lead to repeat expansions [113]. The heterogametic sex chromosomes in some species contain extensive ITS sequences [99] and these can be extreme, for example as observed in the W chromosome of Sand lizards, where they appear to constitute the bulk of the chromosome [114]. Whilst the mechanism of these

repeat expansions is not clear, it is tempting to speculate that a that an rdDNA mediated repeat expansion, associated centromeric function, could be tolerated in gene poor degenerate chromosomes.

Other ITS loci are consistent with simple insertions of telomere sequences, these have the appearance of a chromosomal healing event that was not complete and thus at this telomere sequences may have been captured, or generated *de novo* via telomerase activity, prior to the repair of the locus [115].

ITS represent a signature of the molecular events underlying genomic evolution, their continued presence in the genome may therefore be selectively neutral; however there is evidence to indicate that ITS may contribute to chromosomal function and thus their maintenance may be adaptive. Specific subsets of ITS provide binding sites for the shelterin components RAP1 and TRF2, but this distribution might be controlled by the relative expression levels of TRF2 [116]. The binding of these proteins at ITS may regulate the expression of proximal genes [116]. Moreover, evidence has been presented that indicates that the interaction of TRF2 and A-type lamins with ITS, may result in the formation of interstitial T-loop structures [117], thereby potential linking biological ageing and telomeres, with chromosomal architecture and gene expression profiles.

Summary

Telomeric and sub-telomeric regions of eukaryotic chromosomes, display extraordinary levels of variation both within and between species. The biology of telomeres, their dysfunction and repair, provides mechanisms by which rapid large-scale changes in genomes can occur over short periods of evolutionary time and has the potential to facilitate chromosomal speciation.

Funding

Work in the Baird lab is funded by Cancer Research UK (C17199/A18246).

Competing interests

I have no competing interests.

Figure Legends

Figure 1.

Illustrating telomere dynamics; gradual telomere erosion as a consequence of end-replication losses, stochastic telomere deletion and sub-telomeric DSB formation. Centromeres are depicted by a blue filled oval and telomeres by black and white filled rectangles. Sub-telomeric sequences, or genes, are illustrated by the coloured squares with capital letters.

Figure 2.

Illustrating telomeric DNA-damage processing. A, Catastrophic telomere fusion because of dysfunction within the Shelterin complex. B, Sister chromatid telomere fusion and the initiation of cycles of fusion, anaphase-bridging and breakage gradual telomere erosion. C, Inter-chromosomal fusion. D, Telomeric resection, creating sub-telomeric deletion that can lead to fusion, or the chromosomal healing and the acquisition of a new telomere. E, Robertsonian chromosomal fusion, creating a dicentric chromosome that is stabilised following the inactivation of one of the two centromeres. Active centromeres are depicted by a blue filled oval, inactive centromeres by an unfilled oval and telomeres by black and white filled rectangles. Sub-telomeric sequences, or genes, are illustrated by the coloured squares with capital letters. Interstitial telomere sequences (ITS) are highlighted

References

- [1] Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Aparicio, S. A., Behjati, S., Biankin, A. V., Bignell, G. R., Bolli, N., Borg, A., Borresen-Dale, A. L., et al. 2013 Signatures of mutational processes in human cancer. *Nature* **500**, 415-421. (DOI:10.1038/nature12477).
- [2] Hanahan, D. & Weinberg, R. A. 2011 Hallmarks of cancer: the next generation. *Cell* **144**, 646-674. (DOI:10.1016/j.cell.2011.02.013).
- [3] Papaemmanuil, E., Gerstung, M., Bullinger, L., Gaidzik, V. I., Paschka, P., Roberts, N. D., Potter, N. E., Heuser, M., Thol, F., Bolli, N., et al. 2016 Genomic Classification and Prognosis in Acute Myeloid Leukemia. *N Engl J Med* **374**, 2209-2221. (DOI:10.1056/NEJMoa1516192).
- [4] Jones, C. H., Pepper, C. & Baird, D. M. 2012 Telomere dysfunction and its role in haematological cancer. *Br J Haematol* **156**, 573-587. (DOI:10.1111/j.1365-2141.2011.09022.x).
- [5] Kim, N. W., Piatyszek, M. A., Prowse, K. R., Harley, C. B., West, M. D., Ho, P. L., Coviello, G. M., Wright, W. E., Weinrich, S. L. & Shay, J. W. 1994 Specific association of human telomerase activity with immortal cells and cancer. *Science* **266**, 2011-2015.
- [6] Bryan, T. M., Englezou, A., Dalla-Pozza, L., Dunham, M. A. & Reddel, R. R. 1997 Evidence for an alternative mechanism for maintaining telomere length in human tumors and tumor-derived cell lines. *Nat Med* **3**, 1271-1274.
- [7] Artandi, S. E., Chang, S., Lee, S. L., Alson, S., Gottlieb, G. J., Chin, L. & DePinho, R. A. 2000 Telomere dysfunction promotes non-reciprocal translocations and epithelial cancers in mice. *Nature* **406**, 641-645.
- [8] Lin, T. T., Letsolo, B. T., Jones, R. E., Rowson, J., Pratt, G., Hewamana, S., Fegan, C., Pepper, C. & Baird, D. M. 2010 Telomere dysfunction and fusion during the progression of chronic lymphocytic leukaemia: evidence for a telomere crisis. *Blood* **116**, 1899-1907.
- [9] Roger, L., Jones, R. E., Heppel, N. H., Williams, G. T., Sampson, J. R. & Baird, D. M. 2013 Extensive telomere erosion in the initiation of colorectal adenomas and its association with chromosomal instability. *J Natl Cancer Inst* **105**, 1202-1211. (DOI:10.1093/jnci/djt191).
- [10] Jones, R. E., Grimstead, J. W., Sedani, A., Baird, D. & Upadhyaya, M. 2017 Telomere erosion in NF1 tumorigenesis. *Oncotarget*. (DOI:10.18632/oncotarget.16981).
- [11] Hyatt, S., Jones, R. E., Heppel, N. H., Grimstead, J. W., Fegan, C., Jackson, G. H., Hills, R., Allan, J. M., Pratt, G., Pepper, C., et al. 2017 Telomere length is a critical determinant for survival in multiple myeloma. *Br J Haematol*. (DOI:10.1111/bjh.14643).
- [12] Williams, J., Heppel, N. H., Britt-Compton, B., Grimstead, J. W., Jones, R. E., Tauro, S., Bowen, D. T., Knapper, S., Groves, M., Hills, R. K., et al. 2017 Telomere length is an independent prognostic marker in MDS but not in de novo AML. *Br J Haematol*. (DOI:10.1111/bjh.14666).
- [13] Lin, T. T., Norris, K., Heppel, N. H., Pratt, G., Allan, J. M., Allsup, D. J., Bailey, J., Cawkwell, L., Hills, R., Grimstead, J. W., et al. 2014 Telomere dysfunction accurately predicts clinical outcome in chronic lymphocytic leukaemia, even in patients with early stage disease. *Br J Haematol* **167**, 214-223. (DOI:10.1111/bjh.13023).
- [14] Simpson, K., Jones, R. E., Grimstead, J. W., Hills, R., Pepper, C. & Baird, D. M. 2015 Telomere fusion threshold identifies a poor prognostic subset of breast cancer patients. *Mol Oncol* **9**, 1186-1193. (DOI:10.1016/j.molonc.2015.02.003).
- [15] Strefford, J. C., Kadalayil, L., Forster, J., Rose-Zerilli, M. J., Parker, A., Lin, T. T., Heppel, N., Norris, K., Gardiner, A., Davies, Z., et al. 2015 Telomere length predicts progression and

overall survival in chronic lymphocytic leukemia: data from the UK LRF CLL4 trial. *Leukemia* **29**, 2411-2414. (DOI:10.1038/leu.2015.217).

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

[17] Podlevsky, J. D., Bley, C. J., Omana, R. V., Qi, X. & Chen, J. J. 2008 The telomerase database. *Nucleic Acids Res* **36**, D339-343. (DOI:10.1093/nar/gkm700).

[18] Wright, W. E., Tesmer, V. M., Huffman, K. E., Levene, S. D. & Shay, J. W. 1997 Normal human chromosomes have long G-rich telomeric overhangs at one end. *Genes Dev* **11**, 2801-2809.

[19] Makarov, V. L., Hirose, Y. & Langmore, J. P. 1997 Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. *Cell* **88**, 657-666.

[20] McElligott, R. & Wellinger, R. J. 1997 The terminal DNA structure of mammalian chromosomes. *EMBO J* **16**, 3705-3714. (DOI:10.1093/emboj/16.12.3705).

[21] de Lange, T. 2005 Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes Dev* **19**, 2100-2110.

[22] Pardue, M. L. & DeBaryshe, P. G. 2008 *Drosophila* telomeres: A variation on the telomerase theme. *Fly (Austin)* **2**, 101-110.

[23] Greider, C. W. & Blackburn, E. H. 1985 Identification of a specific telomere terminal transferase activity in *Tetrahymena* extracts. *Cell* **43**, 405-413.

[24] Harley, C. B., Vaziri, H., Counter, C. M. & Allsopp, R. C. 1992 The telomere hypothesis of cellular aging. *Exp Gerontol* **27**, 375-382.

[25] Smogorzewska, A., Karlseder, J., Holtgreve-Grez, H., Jauch, A. & de Lange, T. 2002 DNA ligase IV-dependent NHEJ of deprotected mammalian telomeres in G1 and G2. *Curr Biol* **12**, 1635-1644.

[26] van Steensel, B., Smogorzewska, A. & de Lange, T. 1998 TRF2 protects human telomeres from end-to-end fusions. *Cell* **92**, 401-413.

[27] Kolquist, K. A., Ellisen, L. W., Counter, C. M., Meyerson, M., Tan, L. K., Weinberg, R. A., Haber, D. A. & Gerald, W. L. 1998 Expression of TERT in early premalignant lesions and a subset of cells in normal tissues. *Nat Genet* **19**, 182-186.

[28] Gomes, N. M., Ryder, O. A., Houck, M. L., Charter, S. J., Walker, W., Forsyth, N. R., Austad, S. N., Venditti, C., Pagel, M., Shay, J. W., et al. 2011 Comparative biology of mammalian telomeres: hypotheses on ancestral states and the roles of telomeres in longevity determination. *Aging Cell* **10**, 761-768. (DOI:10.1111/j.1474-9726.2011.00718.x).

[29] Seluanov, A., Chen, Z., Hine, C., Sasahara, T. H., Ribeiro, A. A., Catania, K. C., Presgraves, D. C. & Gorbunova, V. 2007 Telomerase activity coevolves with body mass not lifespan. *Aging Cell* **6**, 45-52. (DOI:10.1111/j.1474-9726.2006.00262.x).

[30] Tian, T., Doerig, K., Park, R., Can Ran Qin, A., Hwang, C., Neary, A., Gilbert, M. R., Seluanov, A. & Gorbunova, V. 2017 Evolution of telomere maintenance and tumor suppressor mechanisms across mammals. *Philosophical Transactions B*.

[31] Risques, R. A. & Promislow, D. 2017 All's well that ends well: Why large species have short telomeres. *Philosophical Transactions B*.

[32] Young, A. 2017 The role of telomeres in the mechanisms and evolution of life-history trade-offs and ageing. *Philosophical Transactions B*.

[33] Campisi, J. 2013 Aging, cellular senescence, and cancer. *Annu Rev Physiol* **75**, 685-705. (DOI:10.1146/annurev-physiol-030212-183653).

- [34] Coppe, J. P., Desprez, P. Y., Krtolica, A. & Campisi, J. 2010 The senescence-associated secretory phenotype: the dark side of tumor suppression. *Annu Rev Pathol* **5**, 99-118. (DOI:10.1146/annurev-pathol-121808-102144).
- [35] Aviv, A. & Shay, J. W. 2017 Reflections on Telomere Dynamics and Aging-Related Diseases in Humans *Philosophical Transactions B*.
- [36] Counter, C. M., Avilion, A. A., LeFeuvre, C. E., Stewart, N. G., Greider, C. W., Harley, C. B. & Bacchetti, S. 1992 Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity. *Embo J* **11**, 1921-1929.
- [37] Capper, R., Britt-Compton, B., Tankimanova, M., Rowson, J., Letsolo, B., Man, S., Haughton, M. & Baird, D. M. 2007 The nature of telomere fusion and a definition of the critical telomere length in human cells. *Genes Dev* **21**, 2495-2508.
- [38] Tankimanova, M., Capper, R., Letsolo, B. T., Rowson, J., Jones, R. E., Britt-Compton, B., Taylor, A. M. & Baird, D. M. 2012 Mre11 modulates the fidelity of fusion between short telomeres in human cells. *Nucleic Acids Res* **40**, 2518-2526. (DOI:10.1093/nar/gkr1117).
- [39] Letsolo, B. T., Rowson, J. & Baird, D. M. 2010 Fusion of short telomeres in human cells is characterised by extensive deletion and microhomology and can result in complex rearrangements. *Nucleic Acids Res* **38**, 1841-1852.
- [40] Heacock, M., Spangler, E., Riha, K., Puizina, J. & Shippen, D. E. 2004 Molecular analysis of telomere fusions in Arabidopsis: multiple pathways for chromosome end-joining. *Embo J* **23**, 2304-2313.
- [41] Maser, R. S., Wong, K. K., Sahin, E., Xia, H., Naylor, M., Hedberg, H. M., Artandi, S. E. & DePinho, R. A. 2007 DNA-dependent protein kinase catalytic subunit is not required for dysfunctional telomere fusion and checkpoint response in the telomerase-deficient mouse. *Mol Cell Biol* **27**, 2253-2265.
- [42] Liddiard, K., Ruis, B., Takasugi, T., Harvey, A., Ashelford, K. E., Hendrickson, E. A. & Baird, D. M. 2016 Sister chromatid telomere fusions, but not NHEJ-mediated inter-chromosomal telomere fusions, occur independently of DNA ligases 3 and 4. *Genome Res* **26**, 588-600. (DOI:10.1101/gr.200840.115).
- [43] Jones, R. E., Oh, S., Grimstead, J. W., Zimbric, J., Roger, L., Heppel, N. H., Ashelford, K. E., Liddiard, K., Hendrickson, E. A. & Baird, D. M. 2014 Escape from telomere-driven crisis is DNA ligase III dependent. *Cell Rep* **8**, 1063-1076. (DOI:10.1016/j.celrep.2014.07.007).
- [44] Heacock, M. L., Idol, R. A., Friesner, J. D., Britt, A. B. & Shippen, D. E. 2007 Telomere dynamics and fusion of critically shortened telomeres in plants lacking DNA ligase IV. *Nucleic Acids Res* **35**, 6490-6500.
- [45] Rai, R., Zheng, H., He, H., Luo, Y., Multani, A., Carpenter, P. B. & Chang, S. 2010 The function of classical and alternative non-homologous end-joining pathways in the fusion of dysfunctional telomeres. *Embo J* **29**, 2598-2610. (DOI:10.1038/emboj.2010.142).
- [46] Sfeir, A. & de Lange, T. 2012 Removal of shelterin reveals the telomere end-protection problem. *Science* **336**, 593-597. (DOI:10.1126/science.1218498).
- [47] McClintock, B. 1939 The behaviour in successive nuclear divisions of a chromosome broken at meiosis. *Proc Natl Acad Sci U S A* **25**, 405-416.
- [48] Murnane, J. P. 2006 Telomeres and chromosome instability. *DNA Repair (Amst)* **5**, 1082-1092.
- [49] Sabatier, L., Ricoul, M., Pottier, G. & Murnane, J. P. 2005 The loss of a single telomere can result in instability of multiple chromosomes in a human tumor cell line. *Mol Cancer Res* **3**, 139-150.

- [50] Lo, A. W., Sabatier, L., Fouladi, B., Pottier, G., Ricoul, M. & Murnane, J. P. 2002 DNA amplification by breakage/fusion/bridge cycles initiated by spontaneous telomere loss in a human cancer cell line. *Neoplasia* **4**, 531-538.
- [51] Hayashi, M. T., Cesare, A. J., Rivera, T. & Karlseder, J. 2015 Cell death during crisis is mediated by mitotic telomere deprotection. *Nature* **522**, 492-496. (DOI:10.1038/nature14513).
- [52] Chin, K., de Solorzano, C. O., Knowles, D., Jones, A., Chou, W., Rodriguez, E. G., Kuo, W. L., Ljung, B. M., Chew, K., Myambo, K., et al. 2004 In situ analyses of genome instability in breast cancer. *Nat Genet* **36**, 984-988.
- [53] Meeker, A. K., Hicks, J. L., Platz, E. A., March, G. E., Bennett, C. J., Delannoy, M. J. & De Marzo, A. M. 2002 Telomere shortening is an early somatic DNA alteration in human prostate tumorigenesis. *Cancer Res* **62**, 6405-6409.
- [54] van Heek, N. T., Meeker, A. K., Kern, S. E., Yeo, C. J., Lillemoe, K. D., Cameron, J. L., Offerhaus, G. J., Hicks, J. L., Wilentz, R. E., Goggins, M. G., et al. 2002 Telomere shortening is nearly universal in pancreatic intraepithelial neoplasia. *Am J Pathol* **161**, 1541-1547.
- [55] Meeker, A. K., Hicks, J. L., Iacobuzio-Donahue, C. A., Montgomery, E. A., Westra, W. H., Chan, T. Y., Ronnett, B. M. & De Marzo, A. M. 2004 Telomere length abnormalities occur early in the initiation of epithelial carcinogenesis. *Clin Cancer Res* **10**, 3317-3326.
- [56] Coop, G. & Przeworski, M. 2007 An evolutionary view of human recombination. *Nat Rev Genet* **8**, 23-34. (DOI:10.1038/nrg1947).
- [57] Royle, N. J., Clarkson, R. E., Wong, Z. & Jeffreys, A. J. 1988 Clustering of hypervariable minisatellites in the proterminal regions of human autosomes. *Genomics* **3**, 352-360.
- [58] Jeffreys, A. J., Murray, J. & Neumann, R. 1998 High-resolution mapping of crossovers in human sperm defines a minisatellite-associated recombination hotspot. *Mol Cell* **2**, 267-273.
- [59] Riethman, H. 2008 Human subtelomeric copy number variations. *Cytogenet Genome Res* **123**, 244-252. (DOI:10.1159/000184714).
- [60] Linardopoulou, E. V., Williams, E. M., Fan, Y., Friedman, C., Young, J. M. & Trask, B. J. 2005 Human subtelomeres are hot spots of interchromosomal recombination and segmental duplication. *Nature* **437**, 94-100.
- [61] Linardopoulou, E. V., Parghi, S. S., Friedman, C., Osborn, G. E., Parkhurst, S. M. & Trask, B. J. 2007 Human subtelomeric WASH genes encode a new subclass of the WASP family. *PLoS Genet* **3**, e237. (DOI:10.1371/journal.pgen.0030237).
- [62] Trask, B. J., Friedman, C., Martin-Gallardo, A., Rowen, L., Akinbami, C., Blankenship, J., Collins, C., Giorgi, D., Iadonato, S., Johnson, F., et al. 1998 Members of the olfactory receptor gene family are contained in large blocks of DNA duplicated polymorphically near the ends of human chromosomes. *Hum Mol Genet* **7**, 13-26.
- [63] Mefford, H. C. & Trask, B. J. 2002 The complex structure and dynamic evolution of human subtelomeres. *Nat Rev Genet* **3**, 91-102.
- [64] Gao, Q., Reynolds, G. E., Wilcox, A., Miller, D., Cheung, P., Artandi, S. E. & Murnane, J. P. 2008 Telomerase-dependent and -independent chromosome healing in mouse embryonic stem cells. *DNA Repair (Amst)* **7**, 1233-1249. (DOI:10.1016/j.dnarep.2008.04.004).
- [65] Boule, J. B., Vega, L. R. & Zakian, V. A. 2005 The yeast Pif1p helicase removes telomerase from telomeric DNA. *Nature*.
- [66] Yu, G. L. & Blackburn, E. H. 1991 Developmentally programmed healing of chromosomes by telomerase in Tetrahymena. *Cell* **67**, 823-832.

- [67] Varley, H., Di, S., Scherer, S. W. & Royle, N. J. 2000 Characterization of terminal deletions at 7q32 and 22q13.3 healed by De novo telomere addition. *Am J Hum Genet* **67**, 610-622.
- [68] Flint, J., Craddock, C. F., Villegas, A., Bentley, D. P., Williams, H. J., Galanello, R., Cao, A., Wood, W. G., Ayyub, H. & Higgs, D. R. 1994 Healing of broken human chromosomes by the addition of telomeric repeats. *Am J Hum Genet* **55**, 505-512.
- [69] Flint, J., Rochette, J., Craddock, C. F., Dode, C., Vignes, B., Horsley, S. W., Kearney, L., Buckle, V. J., Ayyub, H. & Higgs, D. R. 1996 Chromosomal stabilisation by a subtelomeric rearrangement involving two closely related Alu elements. *Hum Mol Genet* **5**, 1163-1169.
- [70] Meltzer, P. S., Guan, X. Y. & Trent, J. M. 1993 Telomere capture stabilizes chromosome breakage. *Nat Genet* **4**, 252-255. (DOI:10.1038/ng0793-252).
- [71] Flint, J., Wilkie, A. O., Buckle, V. J., Winter, R. M., Holland, A. J. & McDermid, H. E. 1995 The detection of subtelomeric chromosomal rearrangements in idiopathic mental retardation. *Nat Genet* **9**, 132-140.
- [72] van Deutekom, J. C., Bakker, E., Lemmers, R. J., van der Wielen, M. J., Bik, E., Hofker, M. H., Padberg, G. W. & Frants, R. R. 1996 Evidence for subtelomeric exchange of 3.3 kb tandemly repeated units between chromosomes 4q35 and 10q26: implications for genetic counselling and etiology of FSHD1. *Hum Mol Genet* **5**, 1997-2003.
- [73] Stadler, G., Rahimov, F., King, O. D., Chen, J. C., Robin, J. D., Wagner, K. R., Shay, J. W., Emerson, C. P., Jr. & Wright, W. E. 2013 Telomere position effect regulates DUX4 in human facioscapulohumeral muscular dystrophy. *Nat Struct Mol Biol* **20**, 671-678. (DOI:10.1038/nsmb.2571).
- [74] Pays, E., Laurent, M., Delinte, K., Van Meirvenne, N. & Steinert, M. 1983 Differential size variations between transcriptionally active and inactive telomeres of *Trypanosoma brucei*. *Nucleic Acids Res* **11**, 8137-8147.
- [75] Lustig, A. J. 2003 Clues to catastrophic telomere loss in mammals from yeast telomere rapid deletion. *Nat Rev Genet* **4**, 916-923.
- [76] Murnane, J. P., Sabatier, L., Marder, B. A. & Morgan, W. F. 1994 Telomere dynamics in an immortal human cell line. *Embo J* **13**, 4953-4962.
- [77] Cheung, I., Schertzer, M., Baross, A., Rose, A. M., Lansdorp, P. M. & Baird, D. M. 2004 Strain-specific telomere length revealed by single telomere length analysis in *Caenorhabditis elegans*. *Nucleic Acids Res* **32**, 3383-3391.
- [78] Enoch, S., Wall, I., Peake, M., Davies, L., Farrier, J., Giles, P., Baird, D., Kipling, D., Price, P., Moseley, R., et al. 2009 Increased oral fibroblast lifespan is telomerase-independent. *J Dent Res* **88**, 916-921. (DOI:88/10/916 [pii] 10.1177/0022034509342979).
- [79] Baird, D. M., Rowson, J., Wynford-Thomas, D. & Kipling, D. 2003 Extensive allelic variation and ultrashort telomeres in senescent human cells. *Nat Genet* **33**, 203-207.
- [80] Baird, D. M., Britt-Compton, B., Rowson, J., Amso, N. N., Gregory, L. & Kipling, D. 2006 Telomere instability in the male germline. *Hum Mol Genet* **15**, 45-51.
- [81] Lipps, H. J. & Rhodes, D. 2009 G-quadruplex structures: in vivo evidence and function. *Trends Cell Biol* **19**, 414-422. (DOI:10.1016/j.tcb.2009.05.002).
- [82] Ding, H., Schertzer, M., Wu, X., Gertsenstein, M., Selig, S., Kammori, M., Pourvali, R., Poon, S., Vulto, I., Chavez, E., et al. 2004 Regulation of murine telomere length by Rtel: an essential gene encoding a helicase-like protein. *Cell* **117**, 873-886.

- [83] Vannier, J. B., Sandhu, S., Petalcorin, M. I., Wu, X., Nabi, Z., Ding, H. & Boulton, S. J. 2013 RTEL1 is a replisome-associated helicase that promotes telomere and genome-wide replication. *Science* **342**, 239-242. (DOI:10.1126/science.1241779).
- [84] Sfeir, A., Kosiyatrakul, S. T., Hockemeyer, D., MacRae, S. L., Karlseder, J., Schildkraut, C. L. & de Lange, T. 2009 Mammalian telomeres resemble fragile sites and require TRF1 for efficient replication. *Cell* **138**, 90-103. (DOI:S0092-8674(09)00721-1 [pii] 10.1016/j.cell.2009.06.021).
- [85] Yunis, J. J. & Prakash, O. 1982 The origin of man: a chromosomal pictorial legacy. *Science* **215**, 1525-1530.
- [86] Royle, N. J., Hill, M. C. & Jeffreys, A. J. 1992 Isolation of telomere junction fragments by anchored polymerase chain reaction. *Proc Biol Sci* **247**, 57-67.
- [87] Royle, N. J., Baird, D. M. & Jeffreys, A. J. 1994 A subterminal satellite located adjacent to telomeres in chimpanzees is absent from the human genome. *Nat Genet* **6**, 52-56.
- [88] Baird, D. M. & Royle, N. J. 1997 Sequences from higher primates orthologous to the human Xp/Yp telomere junction region reveal gross rearrangements and high levels of divergence. *Hum Mol Genet* **6**, 2291-2299.
- [89] Keely, S. P., Renauld, H., Wakefield, A. E., Cushion, M. T., Smulian, A. G., Fosker, N., Fraser, A., Harris, D., Murphy, L., Price, C., et al. 2005 Gene arrays at *Pneumocystis carinii* telomeres. *Genetics* **170**, 1589-1600. (DOI:10.1534/genetics.105.040733).
- [90] Moraes Barros, R. R., Marini, M. M., Antonio, C. R., Cortez, D. R., Miyake, A. M., Lima, F. M., Ruiz, J. C., Bartholomeu, D. C., Chiurillo, M. A., Ramirez, J. L., et al. 2012 Anatomy and evolution of telomeric and subtelomeric regions in the human protozoan parasite *Trypanosoma cruzi*. *BMC Genomics* **13**, 229. (DOI:10.1186/1471-2164-13-229).
- [91] Dreesen, O., Li, B. & Cross, G. A. 2007 Telomere structure and function in trypanosomes: a proposal. *Nat Rev Microbiol* **5**, 70-75. (DOI:10.1038/nrmicro1577).
- [92] Brown, C. A., Murray, A. W. & Verstrepen, K. J. 2010 Rapid expansion and functional divergence of subtelomeric gene families in yeasts. *Curr Biol* **20**, 895-903. (DOI:10.1016/j.cub.2010.04.027).
- [93] Farman, M. L. 2007 Telomeres in the rice blast fungus *Magnaporthe oryzae*: the world of the end as we know it. *FEMS Microbiol Lett* **273**, 125-132. (DOI:10.1111/j.1574-6968.2007.00812.x).
- [94] Vijay, N., Weissensteiner, M., Burri, R., Kawakami, T., Ellegren, H. & Wolf, J. B. W. 2017 Genomewide patterns of variation in genetic diversity are shared among populations, species and higher-order taxa. *Mol Ecol* **26**, 4284-4295. (DOI:10.1111/mec.14195).
- [95] Iwata-Otsubo, A., Radke, B., Findley, S., Abernathy, B., Vallejos, C. E. & Jackson, S. A. 2016 Fluorescence In Situ Hybridization (FISH)-Based Karyotyping Reveals Rapid Evolution of Centromeric and Subtelomeric Repeats in Common Bean (*Phaseolus vulgaris*) and Relatives. *G3 (Bethesda)* **6**, 1013-1022. (DOI:10.1534/g3.115.024984).
- [96] Mlinarec, J., Chester, M., Siljak-Yakovlev, S., Papes, D., Leitch, A. R. & Besendorfer, V. 2009 Molecular structure and chromosome distribution of three repetitive DNA families in *Anemone hortensis* L. (Ranunculaceae). *Chromosome Res* **17**, 331-346. (DOI:10.1007/s10577-009-9025-2).
- [97] Riethman, H. 2008 Human telomere structure and biology. *Annu Rev Genomics Hum Genet* **9**, 1-19.
- [98] Hernandez-Ibarra, N. K., Leitch, A. R., Cruz, P. & Ibarra, A. M. 2008 Fluorescent in situ hybridization and characterization of the Sall family of satellite repeats in the *Haliotis* L. species (abalone) of the Northeast Pacific. *Genome* **51**, 570-579. (DOI:10.1139/G08-041).

- [99] Meyne, J., Baker, R. J., Hobart, H. H., Hsu, T. C., Ryder, O. A., Ward, O. G., Wiley, J. E., Wurster-Hill, D. H., Yates, T. L. & Moyzis, R. K. 1990 Distribution of non-telomeric sites of the (TTAGGG)_n telomeric sequence in vertebrate chromosomes. *Chromosoma* **99**, 3-10.
- [100] Wiley, J. E., Meyne, J., Little, M. L. & Stout, J. C. 1992 Interstitial hybridization sites of the (TTAGGG)_n telomeric sequence on the chromosomes of some North American hyliid frogs. *Cytogenet Cell Genet* **61**, 55-57.
- [101] Ocalewicz, K. 2013 Telomeres in fishes. *Cytogenet Genome Res* **141**, 114-125. (DOI:10.1159/000354278).
- [102] Rovatsos, M., Kratochvil, L., Altmanova, M. & Johnson Pokorna, M. 2015 Interstitial Telomeric Motifs in Squamate Reptiles: When the Exceptions Outnumber the Rule. *PLoS One* **10**, e0134985. (DOI:10.1371/journal.pone.0134985).
- [103] Ijdo, J. W., Baldini, A., Ward, D. C., Reeders, S. T. & Wells, R. A. 1991 Origin of human chromosome 2: an ancestral telomere-telomere fusion. *Proc Natl Acad Sci U S A* **88**, 9051-9055.
- [104] Wurster, D. H. & Benirschke, K. 1970 Indian muntjac, *Muntiacus muntjak*: a deer with a low diploid chromosome number. *Science* **168**, 1364-1366.
- [105] Johnston, F. P., Church, R. B. & Lin, C. C. 1982 Chromosome rearrangement between the Indian muntjac and Chinese muntjac is accompanied by a deletion of middle repetitive DNA. *Can J Biochem* **60**, 497-506.
- [106] Lee, C., Sasi, R. & Lin, C. C. 1993 Interstitial localization of telomeric DNA sequences in the Indian muntjac chromosomes: further evidence for tandem chromosome fusions in the karyotypic evolution of the Asian muntjacs. *Cytogenet Cell Genet* **63**, 156-159.
- [107] Tsipouri, V., Schueler, M. G., Hu, S., Program, N. C. S., Dutra, A., Pak, E., Riethman, H. & Green, E. D. 2008 Comparative sequence analyses reveal sites of ancestral chromosomal fusions in the Indian muntjac genome. *Genome Biol* **9**, R155. (DOI:10.1186/gb-2008-9-10-r155).
- [108] Wang, W. & Lan, H. 2000 Rapid and parallel chromosomal number reductions in muntjac deer inferred from mitochondrial DNA phylogeny. *Mol Biol Evol* **17**, 1326-1333.
- [109] Dobzhansky, T. 1950 Genetics of natural populations. XIX. Origin of heterosis through natural selection in populations of *Drosophila pseudoobscura*. *Genetics* **35**, 288-302.
- [110] Nachman, M. W. & Searle, J. B. 1995 Why is the house mouse karyotype so variable? *Trends Ecol Evol* **10**, 397-402.
- [111] Leache, A. D., Banbury, B. L., Linkem, C. W. & de Oca, A. N. 2016 Phylogenomics of a rapid radiation: is chromosomal evolution linked to increased diversification in north american spiny lizards (Genus *Sceloporus*)? *BMC evolutionary biology* **16**, 63. (DOI:10.1186/s12862-016-0628-x).
- [112] Pellegrino, K. C., dos Santos, R. M., Rodrigues, M. T., Laguna, M. M., Amaro, R. C. & Yonenaga-Yassuda, Y. 2009 Chromosomal evolution in the Brazilian geckos of the genus *Gymnodactylus* (Squamata, Phyllodactylidae) from the biomes of Cerrado, Caatinga and Atlantic rain forest: evidence of Robertsonian fusion events and supernumerary chromosomes. *Cytogenet Genome Res* **127**, 191-203. (DOI:10.1159/000295175).
- [113] Bersani, F., Lee, E., Kharchenko, P. V., Xu, A. W., Liu, M., Xega, K., MacKenzie, O. C., Brannigan, B. W., Wittner, B. S., Jung, H., et al. 2015 Pericentromeric satellite repeat expansions through RNA-derived DNA intermediates in cancer. *Proc Natl Acad Sci U S A* **112**, 15148-15153. (DOI:10.1073/pnas.1518008112).
- [114] Matsubara, K., Uno, Y., Srikulnath, K., Matsuda, Y., Miller, E. & Olsson, M. 2015 No Interstitial Telomeres on Autosomes but Remarkable Amplification of Telomeric Repeats on

the W Sex Chromosome in the Sand Lizard (*Lacerta agilis*). *J Hered* **106**, 753-757.
(DOI:10.1093/jhered/esv083).

[115] Nergadze, S. G., Rocchi, M., Azzalin, C. M., Mondello, C. & Giulotto, E. 2004 Insertion of telomeric repeats at intrachromosomal break sites during primate evolution. *Genome Res* **14**, 1704-1710. (DOI:10.1101/gr.2778904).

[116] Yang, D., Xiong, Y., Kim, H., He, Q., Li, Y., Chen, R. & Songyang, Z. 2011 Human telomeric proteins occupy selective interstitial sites. *Cell Res* **21**, 1013-1027.
(DOI:10.1038/cr.2011.39).

[117] Wood, A. M., Rendtlew Danielsen, J. M., Lucas, C. A., Rice, E. L., Scalzo, D., Shimi, T., Goldman, R. D., Smith, E. D., Le Beau, M. M. & Kosak, S. T. 2014 TRF2 and lamin A/C interact to facilitate the functional organization of chromosome ends. *Nat Commun* **5**, 5467.
(DOI:10.1038/ncomms6467).