The role of recombination in telomere length maintenance

Nicola J. Royle1, Aarón Méndez-Bermúdez, Athanasia Gravani, Clara Novo, Jenny Foxon, Jonathan Williams, Victoria Cotton and Alberto Hidalgo

Department of Genetics, University of Leicester, University Road, Leicester LE1 7RH, U.K.

Introduction to human telomeres

Human telomeric DNA comprises tandem arrays of the consensus sequence TTAGGG that is oriented 5’→3’ towards the terminus. The majority of the telomeric DNA is double-stranded, but at the terminus, the G-rich strand is longer than the C-strand, producing a 3’ single-stranded overhang of 100–300 nt. The double-stranded portion of the telomere varies in length (2–20 kb) between chromosome ends within a cell, between cells and between tissues (Figure 1a). A major contributor to the variation in telomere length is the inability of the cell to fully replicate a linear chromosome during lagging strand synthesis. In addition, other factors that damage telomeric DNA (such as oxidative damage) contribute to a telomere length dynamics, but these events tend to be stochastic in nature. At the start of human telomere repeat arrays, there is often a region that contains DNA-sequence-variant repeats (such as TGAGGG,

TCAGGG and TTGGGG) interspersed with consensus TTAGGG repeats [1–4]. The portion of the telomere that includes divergent repeats is variable, but, when present, these variant repeats are not usually found beyond 3 kb into the repeat array. The distribution of variant and TTAGGG repeats within these regions is highly variable between alleles (at a single telomere) within a population, which implies that there is a high underlying mutation rate. Consequently, these variable regions have been used as a tool to study mutation processes that operate on telomeric DNA [5,6].

The origin of the degenerate repeats is unknown, but it is likely that they arise from single base mutation events and, once present, the variant repeats can be propagated within the telomere primarily by intra-allelic processes such as slippage during replication or unequal sister-chromatid exchange. Population and pedigree analysis support the hypothesis that germine turnover of repeats within the proximal portion of the telomere is driven by intra-allelic processes, although rare recombination or conversion events cannot be excluded [2–4].

DNA damage and repair

The normal cellular response to damage that creates a DSB (double-strand break) in the DNA is cell-cycle arrest that can lead to repair of the DSB via NHEJ (non-homologous end-joining) or HRR (homologous recombination repair). The decision as to which repair pathway is selected is complex, but is partly dependent on the availability of a homologous
sequence that can act as a template. Thus NHEJ is favoured in G1- and early S-phase, whereas HRR is favoured during S- or G2-phases of the cell cycle when sister chromatids are available [7,8]. When damage is detected at an internal location in the chromosome, a DNA-damage-response pathway is activated and orchestrated primarily through two serine/threonine protein kinases, ATM (ataxia telangiectasia mutated) and ATR (ATM- and Rad3-related). The current consensus is that ATM responds to damage that creates DNA DSBs (such as the damage caused by ionizing radiation), whereas ATR responds to the presence of single-stranded DNA (as might be generated if a replication fork stalls) [9]. Following recognition of these DNA lesions, ATM or ATR are activated by autophosphorylation and this results in the phosphorylation of histone H2AX on Ser139 in a large region of chromatin around the DNA damage site (Figure 1b). Subsequently, many other proteins are recruited to the damaged site and this facilitates the repair of the DNA damage. Coincident with the formation of DNA-damage foci, the cell cycle is arrested via ATM- or ATR-mediated activation of checkpoint kinases (Chk1 or Chk2), leading to activation of p53 (encoded by \( TP53 \)) and expression of CDKN1A (cyclin-dependent kinase inhibitor 1A) (also known as p21) [8,9].

**Telomeres and capping**

The ends of human chromosomes must not be recognized as DNA DSBs, because inappropriate repair via NHEJ or HRR would cause chromosome and genome instability. Therefore a primary function of telomeres is to cap chromosomes, and the shelterin protein complex is essential for this capping function. Shelterin is composed of six proteins that only function at telomeres, but it can recruit a wide variety of other proteins to the telomere. Three of the shelterin components bind directly to telomeric DNA. The \( TRF1 \) and \( TRF2 \) genes (encoding telomeric-repeat-binding factor 1 and 2 respectively) are evolutionarily related and encode proteins that bind to double-stranded TTAGGG repeats as homodimers via their Myb DNA-binding domains [8]. The \( TRF2 \) protein preferentially binds at the end of the duplex telomeric repeats, adjacent to the single-stranded G-overhang. The single human \( POT1 \) (protection of telomeres 1) gene [10] encodes a protein that binds to the TTAGGG single-stranded DNA via its two OB (oligonucleotide/oligosaccharide-binding)-fold domains. The three other protein components of shelterin [encoded by \( RAP1 \) (repressor activator protein 1), \( TPP1 \) (a POT1-binding partner) and \( TIN2 \) (TRF1-interacting nuclear factor 2)] interact with one or several of the telomere DNA-binding proteins. Thus human Rap1 binds to \( TRF2 \), TPP1 interacts with \( POT1 \), and TIN2 is able to connect with all three DNA-binding proteins as it interacts directly with \( TRF1 \) and \( TRF2 \) and with \( POT1 \) via its interaction with TPP1 [8]. All of the shelterin components are essential for correct capping and telomere length regulation, as disruption of any of them results in altered telomere length dynamics.

Besides binding to duplex telomeric repeats, \( TRF1 \) and \( TRF2 \) can bend the DNA, thus forming loop structures [11]. The structures observed indicate that the G-rich single-stranded overhang inserts into the duplex DNA and pairs with the C-strand at the point of insertion,
The telomere hypothesis of cellular senescence and immortalization

Figure 2 | The telomere hypothesis of cellular senescence and immortalization

Forming a D-loop. Moreover, TRF2 is required for D-loop formation. The factors that dictate the point of insertion are not known, but presumably this can be accomplished anywhere along the telomere length, including the stretches of TTAGGG repeats found among the degenerate repeats. These terminal t-loop structures or other structures, bound to the shelterin complex, therefore sequester the terminus of the chromosome and so suppress activation of ATM and ATR and activation of a DNA-damage response (Figure 1c).

Telomeres, replication and cellular senescence

Replication of telomeric DNA is problematic for several reasons. First, there is evidence in yeast that the initiation of replication occurs at an origin that lies internal to the telomere. As the replication fork hits the repetitive DNA, it may slow or even pause, possibly because it is necessary to remove the t-loop, or deal with secondary structures that form on the G-rich or C-rich strands. If the single replication fork stalls while travelling through the repetitive DNA it must be restarted to avoid loss of telomeric DNA [12,13]. Furthermore, it is not possible to replicate a linear DNA molecule fully via lagging strand synthesis, as removal of the most terminal RNA primer leaves a gap that cannot be filled (the ‘end replication problem’ [14]), so causing a gradual replication-dependent erosion of telomeric DNA. Following replication, it is necessary to process the terminus to generate G-strand overhangs required for capping and this resectioning also contributes to the erosion of telomeric DNA.

Expression of the enzyme telomerase, a reverse transcriptase that carries its own RNA template, enables a cell to overcome telomere attrition [12]. However, telomerase activity is tightly regulated, mainly via expression of the human TERT (telomerase reverse transcriptase) gene that encodes the reverse transcriptase component. In humans, telomerase activity is only sufficient to maintain telomere length fully in the germline (Figure 2); although telomerase activity can be detected in the progenitor cell compartment of tissues with a high cell turnover, the activity is insufficient to maintain telomere length over a human lifespan. In differentiated human cells, telomerase is inactive, and the gradual erosion of telomere length can be measured as a function of number of cell divisions and age [15].

In 1961, Hayflick and Moorehead [16] first described the limited replicative potential of normal somatic cells in culture, showing that they arrest after a defined number of PDs (population doublings), at the mortality 1 (M1) checkpoint (also known as the Hayflick limit). At this checkpoint, most cells enter a senescent state. They stop dividing and show morphological and gene expression changes, including the expression of high levels of β-galactosidase in lysosomes (SA-β-galactosidase (senescence-associated β-galactosidase) is a marker of senescent cells). It has been shown that there are several routes to the senescent phenotype [17,18], including telomere dysfunction, oxidative stress, DNA damage, oncogene activation and suboptimal growth conditions of cells in culture. These various stimuli seem to act through one of two signalling pathways: activation of tumour-suppressor protein p53 and/or the pRb (retinoblastoma protein) and their downstream cyclin-dependent kinase inhibitors p21 (CDKN1A) or p16 (CDKN2A). Thus various types of cellular stress can trigger senescence.

Recent advances have shown that telomere-dependent senescence arises when one or a few short telomeres that can no longer cap the chromosomes are present in the nucleus [19]. These uncapped or dysfunctional telomeres precipitate the formation of DNA-damage foci [20]. The content of these senescence-associated DNA-damage foci substantially overlaps with other DNA-damage foci, so including phosphorylated ATM or ATR, γ-H2AX (phosphorylated histone H2AX) (bound to subtelomeric DNA) leading to phosphorylation of p53 and expression of p21 [18,20]. Following the onset of senescence, the telomere-induced DNA-damage foci persist, suggesting that senescent cells are in a state of continuous DNA damage response and therefore cell-cycle arrest.

Bypass of the M1 checkpoint can be achieved by disruption of the signalling pathway (for example, loss or silencing of p53), this allows the cell to continue to replicate, but, in the absence of a TMM (telomere maintenance mechanism), telomeres continue to shorten, chromosomes fuse (via NHEJ) [21] and enter a cycle of breakage–fusion–anaphase bridge formation, and the genome becomes increasingly unstable. In the absence of a TMM, the majority of cells will undergo apoptosis at the stage known as M2 or crisis (Figure 2).

Reactivation of telomerase in differentiated cells leads to the avoidance of senescence and therefore to immortality of cells in culture [22], but if accompanied by somatic mutations (oncogene activation and loss of tumour suppressor genes) in tissues, it leads to tumour formation. Furthermore, there is some evidence that genome instability, driven by dysfunctional telomeres as the cells approach crisis, can contribute to tumour formation, although eventually the genome must be stabilized by the activation of a TMM [23]. Telomerase is active in the majority of tumours (~85%), leading to the
stabilization of mean telomere length. However, approx. 15\% of tumours either have no known TMM or they activate the ALT (alternative lengthening of telomeres) mechanism [24].

Features of the ALT mechanism

ALT was first described in immortal cell lines that did not express telomerase [25]. The telomeres in these cell lines are very heterogeneous in length (<1 to >50 kb [24,26]) and contain single-stranded regions on the C-rich or G-rich strand [27]. The telomeres in ALT+ cell lines show gradual shortening due to incomplete replication, but they are also subjected to random events that suddenly shorten or elongate the repeat array. These features of telomeres in ALT+ cells resemble the telomere length dynamics in type II survivors of yeast that lack telomerase (see [28]). It is known that recombination-like processes underlie the telomere length maintenance in type II survivors as they are dependent on RAD32 and other genes that are required for recombination. The first evidence that ALT in human cells is recombination-based came from telomere-tagging experiments in which a plasmid inserted into a telomere repeat array moved to other telomeres after a number of PDs, whereas a plasmid internal to the telomere did not [29]. Furthermore, analysis of the degenerate repeats in human telomeres revealed a class of complex telomere mutations only seen in ALT+ cells [5]. These complex mutations result in the truncation of the progenitor telomere (defined by the interspersion pattern of the degenerate repeats) and addition of a novel telomere repeat array (with a different interspersion pattern) distal to the breakpoint. Such complex mutations cannot be explained by intramolecular events such as slippage during replication, deletions or t-loop replication. Furthermore, they cannot be explained by equal or unequal sister-chromatid exchange and therefore arise by other intramolecular recombination-like processes (see below).

A subset of the nuclear bodies that contain the PML (promyelocytic leukaemia) protein are specific to ALT+ cells [APBs (ALT-associated PML bodies)], as they contain telomeric DNA and components of the shelterin complex in addition to PML and proteins involved in DNA repair, replication and recombination [24]. APBs are only present in a subset of ALT+ cells, and they are more abundant in the S/G2-phases of the cell cycle, when it is thought that telomeres can be elongated. Moreover, some evidence shows that telomeres are closely associated with APBs at certain times [30], but, in contrast with this, there are cell lines and tumours that lack APBs, yet maintain telomeres by a recombination-based process [24]. Thus it is currently not clear what role APBs play in ALT+ cells, but there are three overlapping views: (i) they may be repositories for proteins that are required for telomere length maintenance; (ii) they may be the sites where telomere elongation occurs; and (iii) they may be the repository for by-products of telomere elongation that would otherwise cause cell-cycle arrest. Nonetheless, characterization of the content of APBs has given some indication as to the proteins that may be involved in telomere elongation in ALT+ cells.

Another feature of ALT+ cells is the presence of an ECTR (extrachromosomal telomeric repeat) that is present in multiple forms, including short linear molecules, a proportion of which are located in APBs [31], and relaxed circular molecules that can be single-stranded, double-stranded or indeed partially double-stranded [27,32]. It is thought that the ECTR is generated as a by-product of the recombination processes operating at telomeres in ALT+ cells, but it has also been proposed that the circular forms may sometimes act as substrates for telomere elongation.

A further marker of ALT+ cell lines and indeed some tumours that show telomere recombination is instability at a GC-rich minisatellite called MS32 (D15S8), which is located at an internal site 1q42 [26,33]. The mutation rate at this minisatellite varies between ALT+ cell lines, but is on average >50-fold higher than in telomerase+ or normal cell lines. It is unknown why this minisatellite is destabilized by activation of the ALT mechanism, but it appears to be a localized effect as none of the other minisatellites that have been investigated shows somatic instability in ALT+ cell lines ([33] and C. Novo and N.J. Royle, unpublished work). One hypothesis for the observed MS32 instability is that factors required for the correct replication of MS32 are recruited to telomeres in ALT+ cells and this causes incomplete replication or other errors at MS32 that are processed into DSBs. The subsequent repair by error-prone processes, that may involve the sister chromatid, then result in the loss, gain or other more complex rearrangement of repeats along the minisatellite.

Telomeres and recombination in ALT+ cells

With the exception of instability at MS32, there is no evidence of a general increase in recombination-like activity across the genome in ALT+ cells [34]. However, use of the CO-FISH (chromosome-orientation fluorescent in situ hybridization) technique with strand-specific probes [35] has shown that telomeres in ALT+ cells undergo an elevated level of post-replicative exchange, some of which appears to be unequal T-SCEs (telomere sister-chromatid exchanges) [36]). It has been proposed that unequal T-SCE could be the basis of telomere elongation in ALT+ cells, but a confounding problem is that unequal T-SCE does not result in a net gain of telomeric DNA as one sister chromatid is elongated at the expense of the other. Therefore T-SCE seem to contribute to the erratic telomere length dynamics in ALT+ cells, but other types of recombination-based event must also play roles (Figure 3).

Overexpression of a mutant form of TRF2 (TRF2ΔB) in telomerase-+ cells led to the formation of extrachromosomal circles of telomeric DNA (t-circles). The mutant TRF2ΔB was able to suppress NHEJ between telomeres, but was not able to suppress misprocessing of t-loops [37]. This stochastic misprocessing resulted in t-loop excision, by a XRCC3 (X-ray repair complementing defective repair in Chinesehamster cells 3)- and NBS1 (Nijmegen breakage syndrome 1)-dependent process, so releasing a t-circle and causing dramatic telomere shortening (Figure 3) [37,38]. Moreover, deletion of one of the two mouse POT1 genes (Pot1a) also led to increased T-SCE and to t-circle formation [39]. As T-SCE and t-circles are abundant in ALT+ cell lines, it seems probable...
that the capping function of TRF2, POT1 and possibly other shelterin components is dysfunctional in ALT+ cells.

As indicated above, the deregulation of recombination-like processes in ALT+ cells results in movement of telomeric sequence between chromosome ends [29]. The details of the mechanism that underlies these events are still unclear and several models have been proposed. One of the most attractive models is a type of BIR (break-induced replication), whereby the G-strand of a short ‘recipient’ telomere invades a longer ‘donor’ telomere and uses the C-strand as a template to extend the short telomere (Figure 3). The C-strand is then generated by lagging strand synthesis either concurrently or following G-strand synthesis [40]. The attraction of this model is the net gain of telomeric sequences, i.e. the short recipient telomere is extended without loss of sequences from the donor telomere. In this model, DNA synthesis would be initiated at the 3′-end of the G-strand within the D-loop. Replication would then proceed round the t-loop, extending the telomere by an intramolecular mechanism [24]. This process would not, however, move sequences between telomeres and so it cannot on its own explain all of the known features of telomeres in ALT+ cells.

In summary, ALT+ cells show: (i) elevated T-SCE, but no elevation of SCE across the genome; (ii) t-circles, which seem to arise via recombination-dependent misprocessing of t-loops; and (iii) movement of sequences between telomeres, either directly or via a t-circle intermediate. These features indicate that recombination-like mechanisms are deregulated at telomeres in ALT+ cells, but it is not yet clear whether telomere extension is dependent on one mechanism or whether several mechanisms can achieve telomere elongation within ALT+ cells.

**DNA damage response in ALT+ cells**

The wide range of telomere lengths within ALT+ cells plus the possible presence of stretches of single-stranded DNA along the repeat arrays indicate that some chromosomes in ALT+ cells have dysfunctional telomeres that will not be able to cap the chromosome. Indeed, TIFs (telomere-induced DNA-damage foci) can be detected in ALT+ cell lines throughout the cell cycle [41,43], and ATM appears to be permanently activated in ALT+ cell lines [44]. However, ALT+ cells survive and continue to divide, and so they must avoid a full-blown DNA-damage response, and corroborations of this comes from the observation that most ALT+ cell lines are deficient in p53. Nevertheless, there may be other routes to the management of the DNA-damage response in ALT+ cells. Disruption of TRF2 expression in telomerase+ p53-deficient cells leads to activation of ATM and downstream substrates, leading to loss of G-strand overhangs from telomeres and multiple telomere–telomere fusions via NHEJ. This precipitates an abrupt cell-cycle arrest and apoptosis. Similar disruption of TRF2 expression in an ALT+ cell line, which retains wild-type p53, showed a different response. Stable clones in which TRF2 expression was disrupted were generated, but at a reduced frequency; apoptosis was not induced, but, after a delay, there was an increase in p53/p21-mediated senescence. Finally, there was a significant loss of telomeric DNA from the TRF2-deficient ALT+ clones [44], although it is not known whether this loss was from the ECTR or from telomeres.

It is known that TRF1 and TRF2 can be post-translationally modified by SUMO (small ubiquitin-related modifier) proteins. Moreover, disruption of the SUMOylation pathway in ALT+ cells leads to telomere shortening, the loss of APBs and an increase in cellular senescence [45]. Thus the mutated response to TRF2 disruption in ALT+ cells compared with normal or telomerase+ cells may not be via alterations to the SUMOylation of telomere-binding proteins in ALT+ cells.

**Genes involved in ALT**

The MRN [MRE11 (meiotic recombination 11)–Rad50–NBS1] complex is required for HRR of DSBs at interstitial sites in the genome. Each component contributes different
activities to the complex: MRE11 has exonuclease (3’-5’ \textit{in vitro}) and endonuclease activities, and Rad50 has helicase activity. The third component NBS1 only associates with the MRE11 and Rad50 components at certain times of the cell cycle when it is thought to phosphorylate and thus activate the complex. Interestingly, the MRN complex associates with telomeres in normal cells at specific times during the cell cycle and it may play a role in the formation and disassociation of t-loops. Disruption of the MRN complex in ALT+ cells leads to telomere shortening and loss of APBs, so showing that it is required for telomere length maintenance in these cells [38,46,47].

The \textit{RAD51D} gene, a member of a family of genes homologous with yeast RAD51, which is required for filament formation and strand invasion during HRR, has also been implicated in the ALT mechanism, although its role has not been fully determined [48]. More recently, the protein complex encoded by the \textit{BLM} (Bloom’s syndrome protein), \textit{TOP3A} (topoisomerase III\alpha) and \textit{BLAP75} (BLM-associated polypeptide 75) genes [49,50] has been implicated in the ALT mechanism. Topoisomerase III\alpha localizes with TRF2 in ALT+ cells and siRNA (small interfering RNA) disruption of topoisomerase III\alpha resulted in reduction of TRF2 levels, loss of G-strand overhangs and a reduction of ALT+ cell viability, but without induction of apoptosis.

Following the demonstration that the human ALT mechanism is recombination-based, a role for DNA helicase genes could be envisaged. Two members of the RecQ helicase gene family have been considered, the \textit{BLM} ( Werner’s syndrome protein) gene. Werner’s syndrome patients (\textit{WRN}\textsuperscript{−/−}) show genome instability, leading to increased cell loss resulting in a premature aging phenotype associated with predisposition to cancers, particularly sarcomas. Loss of \textit{WRN} from mouse cells that lack telomerase led to genome instability, increased T-SCE and resulted in cells that were more readily immortalized in culture [51]. This indicates that \textit{WRN} acts as a barrier to T-SCE in mouse cells, in addition to its role in telomere lagging strand replication. Bloom’s syndrome patients (\textit{BLM}\textsuperscript{−/−}) have a cancer-predisposition syndrome and show elevated levels of T-SCE, indicating that BLM is involved in regulating such exchanges. The fact that ALT+ cells show elevated T-SCE raises a question as to what, if anything, BLM does in the ALT mechanism. The evidence that disruption of topoisomerase III\alpha has an impact on ALT+ cell survival also suggests that BLM plays a role in the ALT mechanism. Recently, we have been investigating the role of the \textit{BLM} and \textit{WRN} genes in ALT+ cells by examining the effect they have on minisatellite instability and on mutations in the telomere repeat array.

\textbf{Note added in proof (received 6 April 2009)}

Since submission of the present article, new data have been published which show that intratelomeric copying also contributes to telomere dynamics in ALT+ cells [52].

\textbf{Funding}

Research in N.J.R.’s group is supported by grants from the Medical Research Council and Cancer Research UK. J.W. is funded by a studentship from the Biotechnology and Biological Sciences Research Council, and A.K. is funded by Consejo Nacional de Ciencia y Tecnologia (Mexico).

\textbf{References}


43 Nabetani, A., Yokoyama, O. and Ishikawa, F. (2004) Localization of hRad9, hHus1, hRad1, and hRad17 and caffeine-sensitive DNA replication at the alternative lengthening of telomeres-associated promyelocytic leukemia body. J. Biol. Chem. 279, 25849-25857.


