Dynamics and plasticity of chromosome ends: consequences in human pathologies

Caroline Schluth-Bolard, Alexandre Ottaviani, Amadou Bah, Amina Boussouar, Eric Gilson, Frédérique Magdinier

Laboratoire de Biologie Moléculaire de la Cellule, CNRS UMR5239, Ecole Normale Supérieure de Lyon, UCBL1, IFR128. 46 allée d'Italie, 69364 Lyon Cedex 07, France (CSB, AO, AB, AB, EG, FM)

Abstract

In eukaryotic cells, chromosome ends of linear chromosomes are particular regions of the genome formed by telomeres at the very end and subtelomeres, complex sequences that separate telomeres from chromosome-specific regions. These two regions are highly dynamic and contribute to the stability and integrity of the human genome. Furthermore, in human cells, dysregulation of these regions are implicated in a wide range of physiological events and pathological manifestations. Due to the amount of information available on the biology of telomeres, we will give an overview and discuss here what is currently known of the regulation of telomere length and homeostasis and describe the complex organization of subtelomeric regions and their implication in multiple pathologies.

Introduction

In eukaryotic cells, the genetic information is carried by linear chromosomes, which start and end with a particular structure named telomere. Telomeres are constituted by non-coding repeated DNA, usually considered of heterochromatic nature. They are bound by a specific multi-proteins complex. In particular, mammalian telomeres are constituted by TTAGGG repeats, bound to shelterin or telosome complexes containing six different polypeptides (Gilson and Geli, 2007; Palm and de Lange, 2008).

Telomeres are maintained by a ribonucleoprotein complex named telomerase. This cellular reverse transcriptase counteracts telomere shortening resulting from the incomplete telomere elongation after each round of DNA replication (Gilson and Geli, 2007) by using an RNA component complementary to the telomeric repeats. In the absence of telomerase in somatic cells, telomeres shorten at each round of replication because of the asymmetrical replication process that does not allow the entire replication of the 5' to 3' end. Over the years, accumulating evidence has strengthened the view that excessive telomere erosion is a tumor-suppressor mechanism and contributes to acquired and inherited aging processes. Thus, the balance between telomere shortening and telomerase activity and the regulation of telomere length, is a signal controlling the fate of the cell: proliferation, senescence or apoptosis.

Subtelomeres are complex sequences that separate the chromosome ends from gene specific regions. These regions are composed of repetitive sequences usually shared by different chromosome, degenerated telomeric repeats and numerous families of genes or individual ones. These regions are highly polymorphic and various haplotypes are found in the human population. The frequent rearrangements involving these regions are associated with evolution but also several unrelated human pathologies (Linardopoulou et al., 2005; Ottaviani et al., 2008; Riethman et al., 2001).

This review mainly focuses on the regulation of human telomeres and subtelomeres with a particular emphasis on their importance in pathologies.

I. Telomeres

I-A. General information

Telomeres are specialized chromatin structures at the end of linear chromosomes in eukaryotic species that prevent the chromosome ends from being recognized and processed as double strand breaks. The first clear
recognition of chromosome ends and of their importance in chromosome biology came from the cytogenetic observations of Barbara McClintock in maize in 1931 (McClintock, 1931). Upon X-ray irradiation, she observed an increased rate of chromosomal translocations and the appearance of ring chromosomes and envisaged that such rearrangements were caused by fusion of broken chromosome ends giving for the first time a particular role to chromosome termini in the maintenance of the genome. In the 70s, Blackburn and Gall cloned the first telomeric DNA in the ciliate Tetrahymena thermophila and described the presence of the repetitive TTGGGG sequence (Blackburn and Gall, 1978).

In many organisms, including vertebrates, plants and many protistes, the telomeres are composed of tandem repeats of simple G-rich sequences. In these cases, the G-rich strand always corresponds to the 3' end, which protrudes as a 3' overhang. Of note, the C. elegans telomeres exhibit both 3' G-rich and 5' C-rich overhangs (Raices et al., 2008). In Drosophila, the telomeres are formed by the insertion of specific retroposons. In mammals, telomeres consist in a G-rich TTAGGG hexanucleotidic sequence repeated thousands of times. The size of the telomeres varies among species from a few base pairs in ciliates to thousands of base pairs in higher eukaryotes (Fajkus et al., 1995; Kipling and Cooke, 1990; Klobutcher et al., 1981) and the size of the 3' overhang in mammals varies between 50 and 500 nucleotides (Figure 1).

Telomeres are organized in a large duplex structure than can be seen by electron microscopy, called the t-loop (Telomere loop), which presumably forms through strand invasion of the duplex telomeric repeat by the 3' overhang. Since this structure hides the 3' overhang region, it has been proposed that such a conformation would protect the telomere terminus from being recognized as a damaged DNA sequence and subsequently processed by the DNA damage machinery.

Due to their particular nature, telomeres are processed and maintained by specialized pathways and protein complexes that we will describe below (Figure 1).

Figure 1. Schematic representation of chromosome ends.
Chromosome termini ends with an array of hexanucleotides repeated in tandem. The telomeric sequence found in higher eukaryotes is shown with the protruding 5' overhang ending with a G nucleotide (see text for details). Telomeric sequences are bound by a complex formed by 6 different proteins with specificity for either the single or the double strand DNA. These proteins form the shelterin or telosome complex. TRF1 and Tin2 regulate telomere length and TRF2 protects ends from fusion and prevent activation of the DNA damage response. Additional proteins able to interact with telomeric proteins and involved in DNA damage response and double-strand break repair have also been implicated in telomere length regulation and chromosome end protection (De Boeck et al., 2009; Gilson and Gelli, 2007; Palm and de Lange, 2008). Some of these proteins are indicated here.
I-B. Telomere maintenance and telomere binding proteins

Telomeres are maintained by a specialized ribonucleoprotein complex named telomerase. The telomerase comprises a reverse transcriptase (TERT), a RNA component (TERC) and associates with Dyskerin. This enzymatic complex extends the 3’ end of chromosomes by reverse transcription of the template region of its tightly associated RNA moiety. Telomerase expression is required for unlimited proliferation of yeast, protozoa and immortal human tumor cells, as well as for the extended proliferation in germinal, embryonic and some stem cells where it contributes to the maintenance of telomere length but is absent in most somatic cells. In the absence of telomerase, telomeres shorten at each round of replication in cells ongoing division partly because of the asymmetrical replication process that cannot allow the entire replication of parental DNA (end of replication problem, see below) (Gilson and Geli, 2007) (Figure 2).

For the TERT subunit, distantly related organisms share regions of conserved sequence motifs such as the amino-terminal, the reverse transcriptase and the carboxy-terminal domains. Each domain has a specialized function including catalysis, nucleolar localization, RNA binding, dimerization and recruitment to the telomere. In mammals, telomeres are bound by a specialized complex formed by 6 proteins (TRF1, TRF2, POT1, TIN2, TPP1 and RAP1) and called shelterin (Figure 1). These components either bind to the double stranded DNA regions (TRF1, TRF2 and their interacting factors RAP1 (repressor activator protein 1) and TIN2 (TRF1-interacting nuclear protein 2) or the to the G-strand overhang such as the POT1-TPP1 heterodimer (Blasco, 2007a; Gilson and Geli, 2007; Palm and de Lange, 2008). This complex is implicated in the formation of the t-loop, affects the structure of telomere terminus and controls the synthesis of telomeric DNA by telomerase. The assembly of this complex relies on two bridging proteins that interact with each other, TRF2 that bridges TRF1 to TRF2 and TPP1 that bridges the TRF1-TRF2-Tin2 complex to POT1. However, the shelterin components can be found in cells as separate subcomplexes (Chen et al., 2008; Liu et al., 2004a).

Moreover, photobleaching experiments revealed pools of TRF1 and TRF2 proteins with different dynamics on telomeric DNA in vivo (Mattern et al., 2004). TRF1 was the first human telomeric binding protein isolated by biochemical methods (Bilaud et al., 1996; Chong et al., 1995). TRF1 and TRF2 share a common domain structure consisting of the TRF homology domain (TRFH) allowing homodimerization and heterodimerization and a C-terminal SANT/Myb DNA binding domain, which are connected through a flexible hinge domain. TRF2 N-terminal domain contains basic amino acids while TRF1 contains acidic residues. The TRFH domain of TRF1 and TRF2 contain a docking site through which they recruit other proteins to telomeres (Chen et al., 2008). In particular, TRF1 interacts with the Ku70/80 heterodimer, the BLM helicase (Hsu et al., 2000; Lillard-Wetherell et al., 2004; Opresko et al., 2004), the ATM kinase (Kishi et al., 2001), the nucleotide diphosphate kinase mm23-H2 (Nosaka et al., 1998), components of the mitotic spindle (Nakamura et al., 2001) and the transcriptional repressor SALL1 (Netzer et al., 2001).

TRF2 interacts with components of the Mre11 complex involved in non homologous end joining (NHEJ) and homologous recombination (HR) (Zhu et al., 2000), the Apollo nuclease (Lenain et al., 2004; van Overbeek and de Lange, 2006), the MDC1 DNA signaling factor (Dimitrova and de Lange, 2006), the nucleotide excision repair XPF-ERCC1 nuclease (Zhu et al., 2003) and ATM, canceling thereby the DNA damage response activated by ATM (Figure 1). In addition to binding telomeric sequences, TRF2 can fold DNA by creating higher-order structures such as the t-loop, a lasso-like structure where the end of the telomeric tract is joined to more internal telomeric sequences, or topologically-constraint DNA-protein complexes (Amiard et al., 2007). TRF2 specifically recognize the junction between double and single stranded DNA allowing the G-strand overhang to be sequestered into a protective structure (Griffith et al., 1999; Khan et al., 2007; Stansel et al., 2001). Moreover, TRF2 greatly increases the rate of Holliday junction (HJ) formation and blocks the cleavage by various types of HJ resolving activities (Poulet et al., 2009). Therefore, TRF2 could favor t-loop also by stimulating HJ formation and by preventing resolvase cleavage. Abrogation of TRF2, or expression of a TRF2 dominant negative allele, result in loss of the G-strand overhang leading to end-to-end chromosome fusions and p53 mediated apoptosis (Denchi and de Lange, 2007; Karlsgard et al., 1999; van Steensel et al., 1998). TRF1 and TRF2 are also involved in the protein counting model for telomere length regulation that controls in cis the elongation by telomerase (Ancelin et al., 2002; Loayza et al., 2004).

RAP1 is a 49 kDa protein composed of three domains: a Myb domain, an N-terminal BRCT motif and a C-terminal domain that mediates interaction with TRF2 (Li and de Lange, 2003; Li et al., 2000). Mammalian RAP1 is not able to bind directly to telomeric DNA but associates to telomere through interaction with TRF2. RAP1 has a role in telomere length regulation (Ye et al., 2004) and interacts with several components of the DNA damage machinery.

TIN2 is a 40 kDa protein that binds both TRF1 and TRF2 (Ye et al., 2004) by its central region and its amino-terminal end, respectively. It also binds to TPP1 and depletion or mutation of TIN2 has profound effect on the stability of the shelterin complex (Kim et al., 2004; Ye et al., 2004).
TPP1 (formerly PTOP, PIP1, TINT1) connects POT1 and TIN2 (Hockemeyer et al., 2005; Houghtaling et al., 2004; Liu et al., 2004b; Ye and de Lange, 2004). Impaired TPP1 function is associated with telomere deprotection and telomere loss. TPP1 binds to the amino-terminal half of TIN2 and to the carboxy-terminal region of POT1 and is important for the recruitment of POT1 to telomeres. A splice defect in the TPP1 gene results in adrenocortical dysplasia that arose spontaneously in laboratory mouse (Keegan et al., 2005).

POT1 (Protection of Telomeres 1) contains two OB folds in its N-terminus mediating the interaction with the G-strand telomeric sequence with a high affinity for the 5'-TGGGTTAG-3' sequence (Baumann and Cech, 2001; Kelleher et al., 2005; Lei et al., 2005; Loayza et al., 2004). The crystal structure of POT1 suggests a role in the protection of the 3' end (Lei et al., 2005) and depletion in POT1 after siRNA transfection results in a DNA damage response at telomeres, a reduction in the single strand and telomere fusions (Hockemeyer et al., 2005; Veldman et al., 2004; Yang et al., 2005).

Telomeres exert an effect on the different DNA transactions such as replication (Figure 2), recombination and repair. Short telomeres can replicate earlier than long telomeres (Bianchi and Shore, 2007; Gilson and Geli, 2007) and in yeast, the Sir-mediated silent chromatin emanating from telomeres might block replication initiation in the subtelomeric regions. Semi-conservative DNA replication represents a problem for the full replication of linear DNA molecules and results in the loss of terminal DNA (Gilson and Geli, 2007) (Figure 2). In organisms expressing the telomerase, this loss is rapidly compensated. However, the expression of telomerase is not universal and several processing steps limit the erosion of chromosome ends in human cells. Indeed, in the absence of any counteracting mechanism, cell division will result in a loss of telomeric sequence and loss of capping eventually leading to cellular senescence or apoptosis. The replication timing differs between different species. In human cells, telomeres replicate throughout the S phase and homologous ends seem to be highly coordinated while p and q arms of a single telomere show different timing (Zou et al., 2004). The G-tail results from the processing of the 5' strand rather than from the elongation of the 3' strand. For the leading strand, the existence of a 5' resection activity converts it into a 3' overhang. Another telomere specific problem occurring on the lagging strand is the removal of the last RNA primer. During normal replication, this removal is compensated by the downstream of Okasaki fragment. In the absence of a downstream Okasaki fragment at chromosome ends, a telomere-specific event is required for the removal of the last RNA primer. In humans, the completion of telomere replication is an ATM-dependent damage signal required for the generation of the G-tail. Finally in humans and ciliates, a last step involved in the determination of the last nucleotide (either G or C) is required. The POT1 protein might modulate this activity of a nuclease responsible for this processing since POT1 deficient cells harbor random ends (Hockemeyer et al., 2005).

Genome-wide telomere length exhibits considerable variations in the human population. Several genetic and environmental factors might be involved and a few loci containing genetic determinants of telomere length have been identified by linkage analysis (Andrew et al., 2006; Baird, 2008; Gilson and Londono-Vallejo, 2007; Graakjaer et al., 2004; Vasa-Nicotera et al., 2005). Moreover, this mechanism of allele specific telomere length might be influenced by the nature or the epigenetic status of subtelomeric regions but no telomere sequences adjacent to a telomere that can modulate its length have been identified so far.

Subtelomeric domains are cold-spots for meiotic recombination in a variety of organisms (Miklos and Nankivell, 1976). However, meiotic recombination can occur at an elevated rate near some human telomeres and can have both advantageous and pathological consequences in human biology (Kipling et al., 1996). In addition, chromosome instability through the loss of telomeres has been observed in a wide range of pathologies including cancer. Loss of telomere function leading to chromosome fusion occurs through different mechanisms. Fusion can result from the loss of capping in cells deficient for some telomeric proteins or proteins involved in telomere maintenance. Chromosome instability can occur through breakage/fusion/bridge cycles (B/F/B), which are generated when a chromosome without a telomere replicates and the sister chromatids fuse at their ends. The fused chromatids form a bridge during anaphase that breaks when the two centromeres are pulled in opposite directions. Since none of these two chromatids has a telomere, the same cycle occurs at the next division and can continue for numerous generations leading to increasing rearrangements until the broken chromosome acquire a new telomere and becomes more stable (Murnane, 2006).
Dynamics and plasticity of chromosome ends: consequences in human pathologies

Schluth-Bolard C, et al.

Atlas Genet Cytogenet Oncol Haematol. 2010; 14(5)

Figure 2. The "end of replication problem" and consequences on cell proliferation.
Schematic representation of the replication of the telomeric lagging and leading strand. During the replication of telomeres, the gap left by the degradation of the last Okasaki fragment or the stalling of the fork cannot be overcome by the replication complex that progresses for 5' to 3' leaving a single strand. The action of nucleases and the C strand degradation, regulated in part by the telomere binding proteins increase the size of this 3' overhang.

I-C. Epigenetic regulation of chromosome ends

I-C-1. Nucleosomal structures
Unlike yeast telomeres, mammalian ones contain nucleosomes. Indeed, DNA in higher eukaryotes is packaged in nucleosomes characterized by an unusual repeat length about 20-40 bp shorter than bulk nucleosome spacing. In vitro reconstitution of nucleosome formation revealed that the energy required for the wrapping of telomeric DNA and the formation of nucleosome is the highest among other sequences, suggesting that the repetitive and G-rich nature of telomeric sequence together with the periodicity limits the folding of these regions (Cacchione et al., 1997; Filesi et al., 2000; Rossetti et al., 1998; Widom, 2001). Furthermore, the reduced H1 content at telomeres is in agreement with the short nucleosome spacing (Parseghian et al., 2001; Woodcock et al., 2006) and telomeric sequences have only minor interactions with the histone tails (Cacchione et al., 2003). Human TRF1 is able to recognize its binding site on the nucleosomal surface (Galati et al., 2006; Rossetti et al., 2001). TRF2 overexpression is associated with aberrant nucleosomal organization of distal telomeric sequences and impinges on nucleosomal density at telomeric regions (Benetti et al., 2008).

I-C-2. Telomere maintenance and chromatin
Much of what is currently known on the distribution and regulation of chromatin marks at telomeres and their influence on subtelomeres come from recent studies in the laboratory mouse. By chromatin immunoprecipitation experiments, it has been shown that mouse telomeres are enriched in dimethylated and trimethylated Lysine 9 on the amino terminal tail of histone H3 (H3K9) (Garcia-Cao et al., 2002; Garcia-Cao et al., 2004) and trimethylated Lysine 20 on the amino terminal tail of histone H4 (H4K20) (Gonzalo et al., 2005).

Furthermore, mammalian telomeres are enriched in all three HP1 paralogs (Koering et al., 2002; Sharma et al., 2003), which interacts with the component of the telosome, TIN2 (Kaminsker et al., 2005). When HP1 isoforms are overexpressed in human cells, perturbed telomere structures lead to an increase in end-to-end fusions, an increased sensitivity to ionizing radiations, the suppression of tumorigenicity in a xenograft model and the association of hTERT to telomeres (Sharma et
methylation precedes H3K9 modification (Gonzalo et al., 2005). Interstingly, the level of trimethylated H3K9 and HP1 binding are not affected. However, the level of trimethylated H4K20 and the level of the Suvs4-20 histone methyltransferase, which interacts with the Rb proteins are decreased suggesting that the assembly of chromatin at telomeres is Rb-dependent since H4K20 methylation precedes H3K9 modification (Gonzalo et al., 2005). Despite the absence of CpG dinucleotides within the telomeric track, telomeres are also sensitive to the level of DNA methylation of the subtelomeric regions. A lack of DNA methyltransferase (DNMTs) in mouse embryonic stem cells alters the DNA methylation status of subtelomeric regions, leading to telomere elongation and increased telomeric sister chromatid exchanges (Gonzalo et al., 2006). Consistent with this hypothesis, loss of heterochromatin marks at telomeres increases recombination at these regions and activation of telomerase-independent telomere elongation mechanism (ALT) (Benetti et al., 2007a; Benetti et al., 2007b; Garcia-Cao et al., 2004) suggesting a narrow link between subtelomere chromatin and telomere regulation.

1-C-3. Telomeric Position Effect (TPE)

TPE has been extensively studied in baker's yeast, although it was revealed in this model organism five years after its discovery in Drosophila melanogaster (Gehring et al., 1984; Hazlerigg et al., 1984; Levis et al., 1985). Unlike Drosophila, telomeres in Saccharomyces cerevisiae are constituted of stretches of highly repetitive telomerase-added repeats and thus resemble most of the eukaryotic telomeres therefore constituting a powerful genetic system for the study of TPE. TPE in yeast was first demonstrated by insertion of a construct containing a URA3 at the subtelomeric ADH4 locus, 1.1 kb from the VII-L telomere. Expression of the URA3 gene allows growth of the cells on plates lacking uracil. However, on plates containing a drug toxic for cells expressing URA3 (5-fluoro-orotic acid or 5-FOA), 20 to 60% of the cells were still able to grow, suggesting that the URA3 was silenced in the vicinity of the telomere (Gottschling et al., 1990). Some of the features of TPE were concomitantly described, such as the stochastic reversibility, promoter independence and expression variation and the inward Sir-dependent heterochromatin spreading (Ottaviani et al., 2008). Increasing the length of telomeres improves TPE while telomere shortening limits the silencing of subtelomeric genes (Eugster et al., 2006; Kyriou et al., 1993; Renauld et al., 1993) but the influence of telomere length on TPE is not merely due to the length of the telomere itself but rather to the changes in the recruitment of silencing factors. Thus, the interplay between TPE and length regulation might be less direct (Ottaviani et al., 2008).

The heterochromatin nature of mammalian telomeres and their capacity to induce position effect have been controversial for many years. The first example of telomeric position effect in vivo came from the analysis of replication timing of human chromosome 22 carrying chromosomal abnormality frequently observed in pathologies such as cancer or genetic diseases. Various processes that result in the addition of a new telomere can stabilize these broken chromosome ends. One of these pathways is the process in which the broken chromosome acquires a telomeric sequence from another chromosome, homolog or sister chromatid called "telomere capture". An alternative is de novo telomere addition where the end of a broken chromosome is stabilized by telomerase-dependent addition of telomeric repeats named "telomeric healing". Telomere healing following the deletions of subtelomeric elements delays the replication timing of chromosome 22 (Ofir et al., 1999). This delayed replication is not associated with differences in DNA methylation status, condensation of the chromatin structure of the region or silencing of some subtelomeric genes located 50 kb from the telomere suggesting that the large distance between the telomere and the genes may protect from the spreading of telomeric silencing (Ofir et al., 1999). However, other studies implied that human telomeres neither modulate the expression of nearby genes nor affect the homeostasis of telomeres (Bayne et al., 1994; Sprung et al., 1996).
Comprehensive evidence for transcriptional silencing in the vicinity of human telomeres was provided experimentally by using transgenes inserted adjacent to telomeres, similar to the approach used with yeast after telomere fragmentation (Baur et al., 2001; Koering et al., 2002) and was fueled by additional observations in cell culture and clinical samples (Ottaviani et al., 2008). Reporter genes in the vicinity of telomeric repeats were found to be expressed on average ten-fold lower than reporter at non-telomeric sites. Overexpression of the human telomerase reverse transcriptase (hTERT) in the telomeric clones resulted in telomere extension and decrease in transgene expression (Baur et al., 2001) while overexpression of TRF1, involved in telomere length regulation, lead to the re-expression of the transgene (Koering et al., 2002) indicating the involvement of both the telomere length and architecture in TPE as observed in yeast. In addition, the treatment of cells with Trichostatin A, an inhibitor of class I and II histone deacetylases antagonizes TPE. In human cells, TPE is not sensitive to DNA methylation (Koering et al., 2002) while hypermethylation of the transgene appears as a secondary effect in TPE in mouse ES cells (Pedram et al., 2006).

In human cells, there is a correlation between HP1 delocalization and TPE alleviation by TSA treatment (Koering et al., 2002). By comparison to position effect variegation, TPE might thus be an alternative and specialized silencing process acting for instance through the interaction between the chromatin remodeling factor SALL1 and TRF1 (Netzer et al., 2001) or the telomeric shelterin component TIN2 and HP1 (Kaminker et al., 2005). Thus, in mammals, like in other simpler eukaryotic organisms, classical heterochromatin factors cooperate with telomere-associated proteins in the remodeling of the telomeric and subtelomeric regions and the propagation of the silencing at chromosome ends (Blasco, 2007b).

**I-D. TERRA**

Because they are constituted of highly repeated regions, are enriched with different heterochromatic marks and exert a general repressing effect on the expression of neighbouring genes, telomeres have always been considered as transcriptionally silent. Recently, Northern-Blots with telomeric probes revealed the existence of telomeric transcripts called TERRA or TelRNAs consisting of UUAGGG repeats implying that they are transcribed from the C-rich strand while antisense transcripts consisting of CCCUAA repeats are present at low to undetectable levels (Azzalin et al., 2007; Schoeftner and Blasco, 2008). TERRA has been identified in human, mouse, hamster and zebrafish, and also recently, in yeast (Luke et al., 2008) suggesting that telomere transcription is a common feature in eukaryotes. In mammals, TERRA transcripts are present in all adult tissues and cell lines but absent in the mouse embryos. These transcripts are heterogeneous in size, ranging from 100 bp up to more than 9 kb. Their transcription starts in subtelomeric regions (Azzalin et al., 2007), is processed by the RNA Polymerase II and transcripts are polyadenylated (Schoeftner and Blasco, 2008). The great majority of transcripts are found in nuclear fractions of the cells where they colocalize with various telomeric components and are also found at the telomere tips in metaphases.

TERRA levels are increased in cells deficient in the Suv39h1, h2 and Suv4-20h histone methyltransferases and decreased in cells lacking the Dnmt1 DNA methyltransferase or Dicer activity suggesting that heterochromatin represses TERRA formation. At this step, the involvement of DNA methylation in TERRA synthesis is not clear (Figure 3). Indeed, the Dnmt1 effect observed by Schoeßner et al. in the laboratory mouse, differs from the recent investigation of human telomeres in cells from ICF (Immunodeficiency, centromeric region instability, facial anomalies) patients (Ng et al., 2009). This rare disorder is linked to a mutation in the de novo DNA methyltransferase DNMT3B gene and correlates with a hypomethylation of several repeats and subtelomeric regions (Kondo et al., 2000). Hypomethylation of subtelomeres was associated with an elevated level of TERRA transcript suggesting that the hypomethylation of subtelomere and potentially TERRA promoter regions might facilitate the transcription of these repeats. Interestingly, telomerase-positive cells harbor elevated DNA methylation level at the proximal subtelomere and a reduction in telomeric transcription and TERRA production suggesting that telomere elongation negatively regulates TERRA production possibly through the methylation of the subtelomeres (Ng et al., 2009). On the opposite, telomerase-negative cancer cells maintaining their telomere via the ALT pathway, show heterogeneous methylation pattern.

Although many questions remain on the role of this transcript in telomere regulation, evidences from mammals and yeast revealed a key role in the regulation of the telomerase (Azzalin et al., 2007; Luke et al., 2008; Schoeftner and Blasco, 2008). Previous results from human cells suggested that longer telomeres produce more TERRA, which may then feed back and negatively regulate telomerase access (Azzalin et al., 2007; Schoeftner and Blasco, 2008), while in yeast, the accumulation of TERRA may inhibit telomerase action through the formation of DNA/RNA hybrids hindering thereby telomerase access (Horard and Gilson, 2008; Luke et al., 2008). This new partner in telomere biology opens a wide field of investigation.
Dynamics and plasticity of chromosome ends: consequences in human pathologies

Schluth-Bolard C, et al.

Atlas Genet Cytogenet Oncol Haematol. 2010; 14(5)

Figure 3. Distribution of the epigenetic marks in subtelomeric and telomeric regions controls the transcription of telomeres. Telomeric DNA is packaged into nucleosomes and telomeric chromatin contains marks of heterochromatin such as HP1, trimethylated H3K9 and trimethylated H4K20. At chromosome ends, the shortening of telomeres leads to changes in chromatin condensation. In the mouse, telomere shortening is associated with loss of heterochromatin marks at telomeres and subtelomeres and increase in sister chromatid exchange. Telomeres are transcribed into non-coding transcript named TERRA or TelRNA whose production is initiated in the subtelomeric region. TERRA molecules are produced by the RNA polymerase II from the G-strand and TERRA transcripts are polyadenylated. Several factors regulate TERRA production and stability such as the mRNA decay pathways (NMD) and epigenetic mechanisms such as DNA methylation and methylation of histone tails regulated the Suv39h and Suv4-20h histone methyltransferase. Concerning DNA methylation, contradictory results have been obtained in mice or human cells. In mice, hypomethylation seems to repress TERRA while in human cells, hypermthylation blocks TERRA production possibly through hypermethylation of TERRA promoters.

for understanding the homeostasis of chromosome ends in normal and pathological human samples.

I-E. Localisation of telomeres within the nuclear space

The localization of sequences within the nuclear space is of paramount importance for proper genome functions. The localization of chromosome regions at the periphery of the nucleus enriched in heterochromatin marks seems to play a central role in gene regulation, especially silencing and DNA repair (Finlan et al., 2008; Kumaran and Spector, 2008; Reddy et al., 2008; Therizols et al., 2006). The telomeres are not randomly localized within the nucleoplasm and can be found at the nuclear periphery (Gilson et al., 1993). This positioning varies greatly among organisms, cell types, cell cycle stages and individual telomeres. At the bouquet stage for instance, the clustering of all telomeres at the edge of the nucleus, is a nearly universal feature of meiosis (Scherthan, 2007). In budding yeast, the 32 telomeres are clustered into 4-6 foci which are primarily associated with the nuclear envelope (Gotta et al., 1996). This peripheral localization of telomeres is dependent on redundant pathways (Maillet et al., 2001). One acts through Ku and the second through Sir4-Esc1 (Hediger and Gasser, 2002; Taddei and Gasser, 2004). Relocation to this peripheral nuclear compartment probably does not cause repression per se (Tham et al., 2001) and silencing can be maintained without perinuclear anchoring (Gartenberg et al., 2004). All of the data on silencing and anchoring at the nuclear periphery converge toward a reservoir model where telomere clusters act as a subnuclear compartment concentrating key heterochromatin factors like the Sir proteins (Maillet et al., 1996).

A perinuclear positioning of telomeres is also observed in Plasmodium, where it favors subtelomeric gene conversion (Freitas-Junior et al., 2000) while in plants, telomeres are observed either close to the nuclear periphery (Rawlins and Shaw, 1990) or around the nucleolus (Fransz et al., 2002). In mammalian nuclei, telomeres adopt different locations (Luderus et al., 1996). While human telomeres are clustered at the nuclear periphery in sperm (Gilson et al., 1993; Zalenskaya et al., 2000), most telomeres in lymphocyte nuclei are located in the interior of the nucleolus (Weierich et al., 2003). Thus, it seems that by default, human telomeres are localized internally in most cell types. However, some subtelomeric elements might antagonize this internal localization and target their proximal telomere to the nuclear envelope as suggested by the presence of LADs at different subtelomeres (Guelen et al., 2008). Such an example of localization at the periphery of the nucleus is the positioning of the 4q35 subtelomeric locus, involved in the Facio-Scapulo-Humeral Dystrophy (FSHD) (Masny et al., 2004; Tam et al., 2004).

I-F. Telomeres, senescence and organismal aging

Human primary fibroblasts maintained in culture undergo many divisions before reaching the "Hayflick limit" and arresting growth (Hayflick, 1965). Later, it was shown that this limit corresponds to a critical telomere shortening and dysfunction in the absence of telomere maintenance and accumulating evidence
strengthened the view that an excessive telomere erosion and subsequent activation of the senescence pathway is a tumor-suppressor mechanism (Figure 4).

In mammalian somatic cells, the presence of a minimal set of very short telomeres is sufficient to trigger replicative senescence. Senescence depends on the essential phosphoinositide (PI)-3-kinase-related protein kinase ATM and ATR involved in DNA damage checkpoint. If these checkpoints fail, cells resume division, develop genomic instability and ultimately die during crisis. A few cells escape from crisis and their telomeres are maintained by recombination (Alternative telomere lengthening, ALT, see below) or telomerase reactivation. Viewing replicative senescence as one of the protective mechanisms against tumor formation, it is plausible that senescence-associated genes play significant roles in tumorgenesis repression. At this point, further studies are needed to elucidate the respective biological function of genes differentially expressed in senescent cells and cells suffering from telomere dysfunction.

A common feature among eukaryotes is the progressive decline in vitality over the time. Aging encompasses a wide spectrum of degenerative processes such as the accumulation of carbonylated proteins, damaged enzymes, protein misfolding, lipid peroxidation or activation of inflammatory response pathways. Increasing evidence supports the hypothesis of a link between modification of telomeres and senescence or aging. In S. cerevisiae, a key regulator in the aging process is the Sir2 NAD+ Histone Deacetylase (HDAC) that prevents these regions from recombination and fusion throughout cell division (Kaeberlein et al., 1999; Sinclair and Guarente, 1997).

Telomere length correlates with longevity and disease resistance in Human (Cawthon et al., 2003) whereas loss of telomerase causes accelerated aging in mice and several human syndromes (Blasco, 2007a; Chang et al., 2004). For instance, mutations reducing telomerase activity give rise to premature-ageing syndromes, which may be caused by telomere exhaustion and a reduced replicative potential of stem cells (see below).

In addition, telomere shortening is observed in the Hutchinson-Gilford progeroid syndrome (HGPS) linked to a mutation in the gene encoding A-type lamins (De Sandre-Giovannoli et al., 2003; Eriksson et al., 2003) indicating a close relation between telomere erosion and aging phenotypes (Decker et al., 2009; Ding and Shen, 2008).

Deficiency in the mammalian SIRT6 protein, a member of the Sir2 family (Dali-Youcef et al., 2007; Haigis and Guarente, 2006) increases chromosomal aberrations such as fragmented chromosomes, detached centromeres and gaps and leads to the development of a degenerative aging-like phenotype (Mostoslavsky et al., 2006).

SIRT6 is H3K9 deacetylase specific for telomeres that is critical for maintaining functional telomeres, allowing WRN interaction and preventing end-to-end fusion (Michishita et al., 2008).

Also several environmental factors that affect telomere length such as stress, smoking, obesity and socio-economic status might also accelerate aging (Canela et al., 2007; Cherkas et al., 2006; Epel et al., 2004; Valdes et al., 2005).

Figure 4. Telomere attrition controls the proliferative capacity of somatic cells.

Telomere length is central to the process of aging and tumorigenesis in human cells. As result of the “end of replication problem”, telomeres shorten after each cell division. When telomeres reach a critical length, the recruitment of telomere binding proteins and telomere capping are insufficient. Uncapped chromosome end resembles a double stranded DNA break that is highly unstable and can give rise to chromosome rearrangements. Since telomere erosion limits the proliferative capacity of cells, premalignant transformed clones cease to expend when their telomeres shorten. Critically shortened telomeres elicit a potent DNA damage response and accumulate several DNA damage markers such as phosphorylated gH2AX, 53 BP1, CHK2 for instance (d’Adda di Fagagna et al., 2003; Takai et al., 2003). Many of these proteins localize directly to dysfunctional telomeres to form dysfunctional telomere-induced foci (TIFs). However, entry into senescence or apoptosis can be bypassed by dysfunctional cell cycle controls and lead to the accumulation of chromosomal rearrangement and a global instability of the genome. Telomeres are tightly linked to tumorigenesis and the reactivation of telomerase circumvents telomere shortening and enhances the proliferative capacity of a subset of cells.
I-G. Telomeres and pathologies

I-G-1. Telomeres and cancer

I-G-1a. Reactivation of telomerase

Telomere dysfunction has a dual role in neoplasia either by initiating or suppressing tumorogenesis. As mentioned earlier, telomeres shorten after each division until they reach a critical length incompatible with normal capping (Figure 4). The induction of a DNA damage pathway is triggered, followed by growth arrest. In the presence of a functional p53 pathway, telomere shortening and inherent genomic instability promote cellular senescence, a potent tumor suppressor mechanism. However, in a few cells with disabled checkpoints, short dysfunctional telomeres escape this control step and accumulate rearranged DNA perpetuated through recurrent breakage-fusion-bridge cycles. Generalized chromosomal instability leads to cell death, however a few cells may escape and be maintained after reactivation of a telomere maintenance mechanism.

Human cancers and some human somatic cells are able to maintain or extend their telomere length. Indeed, in approximately 90-95% of human cancers and 60-70% of immortalized cell lines, telomerase is up-regulated. Furthermore, hTERT together with the large T and H-ras oncogenes result in the tumorigenic conversion of normal epithelial and fibroblast cells (Hahn et al., 1999). Chromosomal aberrations resulting from telomere-driven chromosomal instability contributes to the acquisition of the tumoral phenotype and early cancerous lesions seem to accumulate more breaks toward chromosome ends whereas more advanced lesions accumulate breaks along chromosomes (Gisselsson et al., 2000). At later stage, further telomere shortening and interstitial breaks amplify the global instability and heterogeneity observed in tumoral samples. In cellular models, the absence of TRF2 contributes to tumor development in cells expressing the telomerase and an ER-SV40 antigen suggesting that telomeric proteins also contribute to cancer (Brunori et al., 2006). However, in another cellular model, the dominant negative form of TRF2 introduced in melanoma cell lines is associated with a loss of tumorigenicity in at least one cell line (Biroccio et al., 2006). In human tumors, altered levels of telomeric proteins have been observed and might also be associated with tumor phenotype (Bellon et al., 2006; Lin et al., 2006; Ning et al., 2006; Oh et al., 2005; Poncelet et al., 2008; Yamada et al., 2002; Yamada et al., 2000). The targeting of telomere may thus be a promising target in the cure of a wide range of cancer types. However, only a few potent small molecules have been discovered so far and the main targets of these bioactive compounds are either the components of the telomerase or the ligands of G-quadruplexes formed at chromosome ends (Gomez et al., 2006a; Gomez et al., 2006b; Salvati et al., 2007; Tahara et al., 2006).

I-G-1b. The ALT pathway

Approximately 5% of human cancer cells maintain their telomere by ALT, Alternative Lengthening of Telomeres. ALT cells are characterized by the presence of long and heterogeneous telomeres. ALT appears more common in tumors derived from tissues of mesenchymal and neuroepithelial origin. The onset of the ALT pathway might be associated with loss of p53 or p16^INKA or activation of the Cyclin D1. The ALT mechanism likely involves recombination-mediated DNA replication, however, the exact mechanism is still unclear. A mechanism of replication/recombination using a circular telomeric DNA generated by homologous recombination might be implicated, allowing the extension of a telomere end using a sister telomere as a template (Dunham et al., 2000; Yeager et al., 1999). Cellular characteristics of the ALT pathway include the heterogeneous and rapid change in individual telomere length and the presence of intranuclear aggregates defined as ALT-associated promyelocytic leukemia (PML) protein nuclear bodies (APBs) (Henson et al., 2002; Yeager et al., 1999). These APBs contain extra-chromosomal telomeric DNA, TRF1 and TRF2 and several proteins involved in DNA recombination and replication. Decreased DNA methylation of the subtelomeric region in DNA methyltransferase deficient mice is associated with features of ALT including long and heterogeneously sized telomeres, increased telomeric recombination and the presence of APBs (Gonzalo et al., 2006). However, in a subset of cells that resembles typical ALT cells, the absence of APBs in telomerase-negative immortalized cells has been reported (Marciniak et al., 2005; Wu et al., 2000) suggesting the existence of alternative pathways for ALT but the mechanisms underlying either the onset of ALT or telomerase reactivation remain elusive.

I-G-2. Telomere dysfunction in genetic diseases

Recent experimental evidences revealed the existence of different genetic diseases linked to mutation in the telomerase or telomerase associated proteins (Garcia et al., 2007; Kirwan and Dokal, 2009).

Dyskeratosis congenital (DKC) or Zinsser-Engman-Cole syndrome is a rare inherited pathology with an incidence of 1/1 000 000 individuals. The pathology is characterized by the mucocutaneous triad of abnormal skin pigmentation, nail dystrophy and mucosal leukoplakia. Additional symptoms such as dental, gastrointestinal, pulmonary, immunological or neurological abnormalities have also been reported. Usually, patients appear healthy at birth and develop the mucocutaneous features later in life (around the age of 10) with death occurring around the median age of 16. For some patients, malignancy usually occurs in the third decade and includes carcinomas, leukemias and lymphomas.
Three genetic forms have been described for the disease. The first form is recessive, linked to the X chromosome and mainly occurs in male (MIM 305000) but autosomal dominant (MIM 127550) and autosomal recessive forms (MIM 224230) have also been identified. In these genetic diseases, mutations in the gene encoding components of the telomerase result in telomerase deficiency, telomere shortening, end-to-end fusions, increased chromosomal instability and the subsequent development of multisymptomatic syndromes.

The X-linked form of DKC was identified by linkage analysis and positional cloning and involves mutations spread across the DKC1 gene (encoding Dyskerin) on Xq28 (Heiss et al., 1998; Knight et al., 1998). Dyskerin has been predicted to have a pseudouridylation activity and DC pathogenesis might be in part linked to defective ribosome biogenesis. Furthermore, Dyskerin and other associated proteins (GAR1, NH2, NPO10) are part of the telomerase complex and interact with the telomerase RNA component (TERC). Dyskerin mutations lead to reduction in TERC and experimental evidences converge toward abnormalities in telomere maintenance in patients suffering DC.

The autosomal dominant form of DKC is associated with mutations in the TERC gene on chromosome 3q. The pathology results in reduced telomerase activity through either impaired RNA accumulation/stability or catalytic defect of the telomerase complex. The genetic alteration underlying the autosomal recessive form of DKC is currently unknown (Mitchell et al., 1999; Vulliamy et al., 2001; Vulliamy et al., 2004). In around 11% of cases of AD-DC, mutations in the TINF2 gene have been reported (Savage et al., 2008; Walne et al., 2008). This gene encodes the TIN2 shelterin component and TIN2 mutation might lead to a more direct degradation of telomere by preventing TRF1 binding and possibly by leaving telomere ends unprotected although the mechanism has not been fully determined yet (Heiss et al., 1998; Knight et al., 1998). Despite the heterogeneity of symptoms in patients with DKC, some features seem to overlap with the Hoyerhaal-Hreidarsson syndrome (HH, OMIM 300240), a severe multisystem disorder occurring in the neonatal period and infancy. As observed in the X-linked form of DKC, several DKC1 mutations have now been identified in patients affected with this syndrome (Marrone et al., 2007).

Like in patients with DKC, individuals with aplastic anemia and myelodysplasia have short telomeres and increased chromosomal instability compared to age-matched controls. These patients are believed to suffer from a defect at the level of stem cells, especially affecting the renewal and proliferative capacity of haematopoietic and epithelial tissues (Vulliamy et al., 2002; Yamaguchi et al., 2005). Idiopathic pulmonary fibrosis (IPF) is a progressive fatal lung disease characterized by lung scarring and abnormal gas exchange that usually arise around the age of 50. However, a familial form of IPF with early onset has been described in occidental countries. In a significant percentage of cases, this pathology has been associated with mutations in TERC or HTERT (Tsakiri et al., 2007), which are different from those usually observed in patients with DC.

In general, the mutations of the telomerase complex or proteins involved in telomere maintenance might affect the recruitment of telomerase, the activity of the enzyme and in turn lead to a dysfunctional telomere maintenance, especially in stem cells whose renewal potential become limited.

II. Subtelomeres

II-A. General information

Subtelomeres are DNA sequences placed between chromosome-specific regions and chromosome ends with features that distinguish them from the rest of the genome (Mefford and Trask, 2002; Riethman, 2008; Riethman et al., 2001). Human subtelomeres vary in size from 10 to up to 500 kb in human cells. They contain repetitive sequences of different types and numerous genes but very little are known on their function in the regulation of cellular homeostasis (Mefford and Trask, 2002; Riethman, 2008; Riethman et al., 2001). However, these regions, prone to recombination and rearrangements, are associated with genome evolution, human disorders but also aging possibly through TPE (Ottaviani et al., 2008). In the human population the subtelomeric regions are highly polymorphic and the rate of recombination at chromosome ends is higher than in the rest of the genome. Such rearrangements participate in the genome variability and the length of variation may be up to hundreds of kilobases among different haplotypes (Figure 5). Although the coverage of chromosome ends has not been fully achieved, available sequences allowed the representation of a
In eukaryotes, the subtelomeres are patchworks of genes (pink rectangles) interspersed within repeated elements (blue rectangles) and degenerated telomeric repeats (ITS, internal Telomeric Sequences). In human, large polymorphic blocks of repeated sequences are distributed between the different chromosomes and subtelomeres contain genes.

A detailed paralogy map showing that several blocks of sequences are shared by different human subtelomeres (Flory et al., 2004; Linardopoulou et al., 2005; Riethman et al., 2001). Various tandemly repeated units called Telomere-Associated Repeats (TAR1), short native telomeric arrays and numerous degenerate telomere-like repeats are also located at variable distance from the telomere and subtelomeres contain members of 25 small families of genes encoding potentially functional proteins (Flory et al., 2004; Linardopoulou et al., 2005; Riethman et al., 2001). Interestingly, many of them are involved in the adaptation to the environmental changes like in other species suggesting that the plasticity of chromosome ends is likely to play a key role in genome evolution and that abnormalities or dysregulation of these genes may have phenotypical consequences (see below for examples).

II-B. Subtelomeric sequences

II-B-1. Families of genes

Human subtelomeric genes vary in copy number and chromosomal distribution. Large families of genes are present at subtelomeres such as odorant and cytokine receptors, homeodomain proteins, secretoglobins together with several genes of unknown function.

II-B-1a. The WASH genes

The most terminally located human subtelomeric genes were recently identified and encode a third class of the Wiskott-Aldrich syndrome protein (WASP) family (Linardopoulou et al., 2007). Five WASP family members are known in mammals and are involved in cell motility, phagocytosis, cytokinesis and in processes such as angiogenesis, embryogenesis, inflammatory immune response, microbial infection and cancer metastasis (Millard et al., 2004; Takenawa and Suetsugu, 2007; Yamaguchi and Condeelis, 2007). Thus, the recent characterization of the human subtelomeric MGC52000 genes allowed the identification of these new members of the WASP family, named WASH for Wiskott Aldrich Syndrome protein and Scar Homolog. Human genome harbors multiple functional WASH paralogs at subtelomeres with the coding sequence ending within 5 kb of the telomere. Several functional WASH variants and multiple pseudogenes have been identified per genome and subtelomeric dynamic might contribute to the variation and diversification of the WASH family. The WASH genes are found in numerous species suggesting conservation during evolution with extensive duplication and dispersal to chromosome ends in primates.
Their colocalization with actin in vivo suggests a role in actin polymerization and cytoskeleton reorganization. The drosophila WASP protein was recently identified as a component of a nuclear complex containing various transcriptional factors and chromatin modifiers and is essential in development. Interestingly, the WASP gene defective in the Wiskott-Aldrich syndrome causes eczema, thrombocytopenia and immunodeficiency (Ochs and Thrasher, 2006) and by analogy, loss of the WASH genes in human through subtelomeric rearrangements for instance might have consequences in pathologies.

II-B-1b. The α and β Defensin loci

The human genome is rich in genomic regions repeated several times. These regions named CNVs (Copy Number Variations) are highly polymorphic but also involved in predisposition to diseases. Among these CNVs, the subtelomeric α and β Defensin loci located on chromosome 20 and at the 8p23.1 subtelomere and containing at least 8 Defensin genes, are associated with different pathologies mainly characterized by an altered immune response. Defensin are small cationic secreted peptides (3-5 kDa) with antimicrobial activity against gram-positive and gram-negative bacteria as well as fungi and enveloped viruses acting by disrupting membrane integrity and function (Ganz, 2003). Some Defensin are produced constitutively while some are synthesized in response to microbial products or pro-inflammatory cytokines amplifying subsequent innate and adaptive immune response (Ganz, 2003; Lehrer and Ganz, 2002). α Defensin are mainly expressed in neutrophils and paneth cells of the intestine while β Defensin are expressed by epithelial tissues.

This locus is the fastest changing CNV currently known (Abu Bakar et al., 2009) and probably the most clinically relevant CNV involved in lupus for the α Defensin (Bennett et al., 2003; Ishii et al., 2005), the Crohn's disease (Fellerman et al., 2006), psoriasis (Hollox et al., 2008), but also potentially to other inflammatory disorders for the β Defensin genes.

II-B-1c. Olfactory Receptors

Olfactory receptors (ORs) comprise one of the largest gene families in the genome of mammals with over 400 olfactory receptor genes and pseudogenes in human (Glusman et al., 2001; Niimura and Nei, 2007). Each neuron in the olfactory epithelium expresses a single allele of a single OR gene and axon neurons expressing the same gene converge in the olfactory bulb of the brain. A single odorant can be recognized by a number of different receptor types and the differential affinity of these different receptors together with the combinatorial use of other OR allow the detection of millions of chemicals by the olfactory system. Human ORs are often clustered and might represent > 0.1% of the human genome. The numerous OR clusters arose through tandem duplications and several ORs clusters are localized in subtelomeric or centromeric regions. Over 60% of human OR genes bear one or several sequence disruptions likely resulting in the inactivation of the corresponding protein. Interestingly, the human OR repertoire is interspersed with numerous repeated elements suggesting a high capacity to recombine. In some cases, recurrent 8p rearrangements might occur as a consequence of an inversion polymorphism mediated by two ORs genes between the subtelomeric 8p23.1 and 4p16 loci (Wieczorek et al., 2000a; Wieczorek et al., 2000b). In addition, unequal crossovers between two OR genes in 8p clusters are responsible for the formation of three recurrent chromosome rearrangements (inv dup(8p), +der(8p) and inv(8p)) associated with distinct phenotypes (Ciccone et al., 2006; Giglio et al., 2001). Furthermore, a much faster functional deterioration of the large OR gene superfamily occurred in the human lineage compared to the great apes and old world monkeys.

It is thus tempting to speculate that a lesser need for the sense of smell in humans compared to other species involved a rapid evolution of the OR gene repertoire, possibly through subtelomeric rearrangements (Gilad et al., 2003).

II-C. Regulation of subtelomeric sequences

II-C-1. Does telomere shortening modulate expression of subtelomeric genes?

In the human population, subtelomeric regions are highly polymorphic and length variation can be up to hundreds of kilobases among the different haplotypes. As described above, telomere shortening affects the epigenetic regulation of subtelomeres and might also impact on their recombination rate. Thus, transcriptional regulation of natural subtelomeric genes in human cells likely depends on telomere length, the structure of the telomeric chromatin but also on the composition of the subtelomeric regions and the spatial organization of chromosome ends.

The effect of telomere shortening on the expression of subtelomeric genes was recently investigated during senescence in human fibroblasts maintained in culture for an extended period of time (Ning et al., 2003). A total of 34 subtelomeric genes and the length of the corresponding telomeres were analyzed in young and senescent cells. Despite a differential expression for 17 out of these 34 genes, telomere length alone is not sufficient to determine the expression status of telomeric genes (Ning et al., 2003) suggesting a complex interplay between telomere and subtelomere regulation.

Age-dependent telomere erosion might also be a key player in the regulation of subtelomeric genes in elders as it was observed experimentally in artificial systems.
(Baur et al., 2001; Koering et al., 2002). In mammals, aging is associated with a multitude of gene expression changes and increasing evidence supports the hypothesis of a link between senescence or aging and modification of chromatin since the architecture of the telomeric and subtelomeric regions is also remodelled during these two processes and a number of factors that can influence directly or indirectly telomere structure may alter the expression of subtelomeric genes by changing telomere conformation and maintenance and vice versa although clear demonstration in higher eukaryotes are still lacking.

II-C-2. Telomeres and subtelomeres: adaptation to environment?

Different mechanisms could be designed to accommodate the evolution of environmental conditions throughout life. One way might be the rapid regulation of subtelomeric genes by Telomeric Position Effect. Indeed, a subtelomeric enrichment of genes related to stress response and metabolism in non-optimal growth conditions appears to be a conserved feature in many yeast species (Robyr et al., 2002) and clustering stress response genes at subtelomeres might be an evolutionary conserved strategy allowing their reversible silencing and a fast response to changes in environmental conditions (Barry et al., 2003; Borst and Ulbert, 2001; Dreesen et al., 2007; Ottaviani et al., 2008). Aging is characterized by an increasing susceptibility to environmental stress and a wide range of diseases. Interestingly, older subjects are more susceptible than younger ones to pathogenic stimuli (Krabbe et al., 2004) and the IgH genes cluster localized at subtelomeres is up-regulated in aging hematopoietic stem cells. Also, in human hepatic stellate cells undergoing senescence, increased expression of genes mediating inflammatory response has been observed (Schnabl et al., 2003). Interestingly, numerous genes encoding cytokines are located at subtelomeric loci and might be also influenced by telomere length. Impressively, sensory perception may also affect life span in higher animals (Libert et al., 2007; Lindemann, 2001) and some gustatory and olfactory neurons either promote or inhibit longevity (Aldco and Kenyon, 2004). Olfactory genes are preferentially positioned at subtelomeric positions, it would be interesting to correlate changes in olfactory stimuli in aged individuals and control of expression of subtelomeric clusters of genes by telomeres.

II-D. Subtelomeres and pathologies

II-D-1. General information

As described above, subtelomeres are highly polymorphic and a broad range of expression level of natural subtelomeric genes is likely to be found from individual to individual as described in yeast (Pryde and Louis, 1999). Thus, telomere length-mediated transcriptional regulation of natural subtelomeric genes in human cells is likely to operate through the telomeric heterochromatin structure, involving long and variable stretches of subtelomeric sequences and renders analysis of telomeres and subtelomeres challenging in human pathologies. The only naturally occurring situations wherein telomeric repeats are adjacent to unique sequences are those that occur in patients with truncated chromosomes ends that have been repaired by the process of telomere healing or that lead to the formation of ring chromosomes. However, the molecular pathogeneses associated with these rearrangements have never been investigated. Moreover, TPE may play a direct role in human diseases as a result of repositioning of active genes near telomeres or subtelomeric sequences following such chromosome rearrangements and subtelomeric element may either participate in the spreading of silencing in the vicinity of a telomere or shelter genes from this silencing.

II-D-2. Idiopathic mental retardation

Mental retardation (MR) affects 3% of the general population (Tyson et al., 2004) and approximately 15% of cases are explained by chromosome aberrations (Stankiewicz and Beaudet, 2007). Since subtelomeric regions are among the most gene rich regions of the genome and are particularly prone to recombination, it was logical to think that subtelomeric imbalances would account for MR. However, subtelomeric regions are difficult to explore with standard cytogenetic techniques and new strategies had to be developed in order to confirm their role in MR. In 1995, Flint et al. found 6% of cryptic subtelomeric imbalances in MR patients using polymorphic microsatellite markers (Flint et al., 1995). Since then, faster techniques have been developed such as fluorescent in situ hybridization (FISH) (Knight and Flint, 2000), Multiplex Ligation dependent Probe Amplification (Rooms et al., 2004) or array comparative genomic hybridization (CGH) (Ballif et al., 2007a; Ballif et al., 2007b; Veltman et al., 2002). They allowed the rapid screening of children with unexplained MR. The largest study on 12,000 MR patients investigated by FISH demonstrated that 2.5% of the clinical cases with severe to mild mental retardation display relatively small subtelomeric abnormalities of all the chromosomes arms with exception of the p-arm of the acrocentric chromosomes. Subtelomeric imbalances include deletions, duplications, unbalanced translocations and complex rearrangements (Shao et al., 2008). They are terminal as well as interstitial. Their size is extremely variable. However, these rearrangements are quite large since 40% of them are over 5 Mb in size (Ballif et al., 2007a; Ballif et al., 2007b). Fifty percent of the subtelomeric imbalances are familial cases. In a series of 56 families, Adeyinka and colleagues demonstrated that 65% of the derivative chromosomes were inherited from a parent carrier of a
balanced translocation, and that 32% of the subtelomeric deletions were inherited from a parent presented with normal clinical features or a milder phenotype than the affected children (Adeyinka et al., 2005). Among the benign subtelomeric copy number variations, it can be distinguished common telomeric polymorphisms and transmitted subtelomeric imbalances without phenotypic effect (Ledbetter and Martin, 2007). Common telomeric polymorphisms are present in at least 1% of the population. They are well known for telomeres 2q, 4q, 7q, 9p, 10q, Xp and Yq. Diagnostic assays now avoid the detection of these clinical insignificant variations. Transmitted subtelomeric imbalances without phenotypic effect are less common (Barber, 2008). They have a wide range of sizes from 150 kb to 10 Mb. They have now been detected at 24 of the 41 telomeres (Balikova et al., 2007). Several mechanisms may explain the absence of abnormal phenotype in carriers of the subtelomeric imbalances: variable expressivity, unmasking of recessive allele, somatic mosaicism in the normal parent and epigenetic modifications.

During the recent years, the development of high resolution genetic analysis techniques allowed a better characterization of the genotype of patients affected with such developmental delays and lead to the identification of different genes disrupted by these subtle terminal deletions (Kleefstra et al., 2006; Lamb et al., 1993; Walter et al., 2004). Nevertheless, gene deletion does not always explain the pathological manifestations. Among the hundreds of patients analyzed, the size of the subtelomeric region disrupted may be accompanied by variable degrees of chromatin condensation and explain the penetrance of the clinical manifestations in patients (Walter et al., 2004). Surprisingly, rather few subtelomeric imbalances are associated with a distinct, recognizable phenotype. Genotype-phenotype correlations for these syndromes are not well established. If one or few genes could have been identified for some syndromes, molecular mechanisms are still unclear for others. It appears that gene dosage alone cannot account for the phenotype in many cases and those alternative mechanisms, in particular epigenetic modifications, may be involved. We will take three examples to illustrate this point. Thus, it is conceivable that genes residing in close proximity to healed telomeres become epigenetically inactivated contributing to the phenotype. However, characterization of the rearrangement's effect on gene expression is still needed to prove that mental retardation is caused by modification of the chromatin architecture at telomeric and subtelomeric loci.

II-D-3. The 22qter syndrome
The 22qter deletion syndrome is characterized by severe neonatal hypotonia, global development delay, autistic-like behavior, normal to accelerated growth, absent to severely delayed speech and minor dysmorphic features (Balikova et al., 2007). The deletion is usually terminal although cases of interstitial deletions sparing the telomere have been described. Deletions range from 130 kb to 9 Mb in size (Wilson et al., 2003). However, there is little correlation between the size of the deletion and the severity of phenotype (Koolen et al., 2005). It has been demonstrated that all the patients shared a deletion of SHANK3, a gene encoding a scaffolding protein found in excitatory synapses, and that a recurrent breakpoint was in this gene (Bonaglia et al., 2006; Wilson et al., 2003). Moreover, patients with SHANK3 disruption or point mutations present the same phenotype than the 22qter deletion syndrome (Bonaglia et al., 2001; Durand et al., 2007). So, it appears that SHANK3 is the major gene responsible for at least the neuro-behavioral phenotype of 22qter deletion syndrome.

II-D-4. The cri-du-chat syndrome
Cri-du-chat syndrome is due to the terminal deletion of the short arm of chromosome 5 (5pter). The phenotype is characterized by microcephaly, facial dysmorphism, high-pitched cat-like cry, severe mental retardation and speech delay (Cerruti Mainardi, 2006). The range of deletion size is also very wide, since deletion can be visible on standard karyogram. Genotype-phenotype correlation studies delineated three critical regions corresponding to cry (5p15.31), speech delay (5p15.32-15.33) and facial dysmorphism (5p15.31-15.2) (Zhang et al., 2005) but no gene has been identified for each specific feature yet. The severity of mental retardation seems to be correlated with the deletion size although some patient presented with a disproportionately severe retardation regarding to the deletion size.

II-D-5. The 1p36 mental retardation syndrome
The 1p36 monosomy syndrome is the most frequent subtelomeric microdeletion syndrome with a frequency of 1/5,000. It is characterized by mental retardation, developmental delay, hearing impairment, seizures, growth impairment, hypotonia, heart defect and distinctive dysmorphic features (Gajecka et al., 2007). Two thirds of de novo rearrangements are apparently simple terminal truncation. The remaining third corresponds to complex structures including deletions with inverted duplication, large duplications and triplications with small deletions and interstitial deletions. Attempts have been made to demonstrate that monosomy 1p36 was a contiguous gene syndrome. Candidate genes for seizures and facial features have been proposed (Heilstedt et al., 2003a; Heilstedt et al., 2003b). However, arguments are in favor of alternative mechanisms. No correlation between the deletion size and the number of clinical features could be observed (Gajecka et al., 2007). There is neither common breakpoint nor common deletion interval in monosomy 1p36 patients. Redon et al., studied 6 patients using tiling path array CGH, two of the six patients presented with very similar features.
but had non-overlapping 1p36 deletions (Redon et al., 2005). Thus it was proposed that the 1p36 monosomy syndrome might be due to a positional effect rather than haploinsufficiency of contiguous genes.

**II-D-6. Facio-Scapulo-Humeral Dystrophy**

One of the best-characterized human genetic diseases potentially linked to TPE is the Facio-Scapulo-Humeral Dystrophy (FHS). This puzzling pathology is associated with the deletion of repeated elements at the 4q35 locus. Normal 4q35 chromosome termini carry from 11 up to 150 copies of a 3.3 kb repeated element named D4Z4 while in FSHD patients the pathogenic allele has only 1-10 repeats. This autosomal dominant disorder is the first most common myopathy clinically described by a progressive and asymmetric weakening of the muscles of the face, scapular girdle and upper limb. The pathogenic alteration does not reside within the gene responsible for the disease but is rather related to an epigenetic mechanism. Several hypotheses have been proposed to explain this enigmatic pathology (Gabellini et al., 2004; van der Maarel and Frants, 2005). Evidence for the binding of a repressor complex to D4Z4 that might regulate the expression of the nearby genes was provided (Gabellini et al., 2002) but remains controversial (Jiang et al., 2003; Winokur et al., 2003). However, the most popular hypothesis to explain this dystrophy is the involvement of PEV or TPE (van Deutekom et al., 1996). D4Z4 shares some of the properties of heterochromatic sequences such as DNA methylation and it was postulated that D4Z4 and surrounding sequences would be packed as heterochromatin leading to the silencing of nearby genes. In patients, the partial loss of the D4Z4 repeat would lead to local chromatin relaxation and to the transcriptional upregulation of genes (Hewitt et al., 1994; Winokur et al., 1994). However, the analysis of the chromatin structure of this locus either in normal individuals or in FSHD patients does not fully support this hypothesis (Jiang et al., 2003). Alternatively D4Z4 may act as an insulator, separating heterochromatic subtelomeric sequences distal to D4Z4 from euchromatic sequences upstream (van Deutekom et al., 1996). We recently showed that FSHD might be associated with a gain of function of CTCF (Ottaviani et al., 2009). Interestingly, different allelic variants might also be linked to the pathology and the 4q35 subtelomeric region appears as a mosaic of regulatory elements. The understanding of the cross talks between D4Z4, the 4q35 subtelomere and the telomere would provide insights in the deciphering of this complex epigenetic disease and the involvement of the D4Z4 subtelomeric element in transcriptional activity or replication timing of the 4q35 chromosome end.

**II-D-7. Ring chromosomes**

Hundreds of patients have been reported with various combinations of malformations, minor abnormalities and growth retardation usually associated with mental retardation linked to the formation of a ring chromosome (Cote et al., 1981; Kosztolanyi, 1987). Ring chromosomes are thought to be formed by deletion near the end(s) of chromosomes followed by fusion at breakage points and have been described for all human chromosomes. The resulting phenotypes vary greatly depending on the size and the nature of the deleted segments. Most ring chromosomes are formed by fusion of the deleted ends of both chromosome arms coupled with the loss of genetic material. However, in a few cases, the rings are formed by telomere-telomere fusion with little or no loss of chromosomal material and have intact subtelomeric and telomeric sequences suggesting that the "ring syndrome" might be associated with the silencing of genes in the vicinity of a longer telomere. The formation of intact ring caused by telomere-telomere fusion and associated with putative telomeric position effect effects has been reported for different autosomes (Pezzolo et al., 1993; Sigurdardottir et al., 1999; Vermeesch et al., 2002). For instance, a severe seizure disorder with features of non-convulsive epilepsy is a characteristic of ring chromosome 20. In this pathology, the formation of ring chromosome is generally associated with a breakage in each chromosome arm and the subsequent fusion of the broken ends with the loss of the telomere, subtelomeric regions or CHRNA4 and KCNQ2, two well-known epilepsy genes. In a patient with a typical severe epilepsy, classical cytogenetic methods, chromosome and quantitative FISH showed that the ring had a longer telomere than either of the 20p or 20q telomere ends suggesting that telomeric position effect silences the CHRNA4 and KCNQ2 genes (Zou et al., 2006).

Strikingly, most of these genetic diseases associated with mental retardation and different malformations are either linked to terminal deletion or fusion raising the hypothesis of a major contribution for telomere and subtelomere integrity in development. However, attempts to make genotype-phenotype correlations with specific anomalies have been difficult because of the paucity of reported cases and the variability in the size of the terminal deletion. In addition these chromosomal abnormalities are often mosaic and the occurrence of sister chromatid exchange complicates the description of these heterogeneous developmental disorders and the precise classification of the genetic alterations.

**Conclusions**

Telomeres on natural chromosomes are dynamic regions involved in numerous cellular pathways, which in turn controls cell fate. Next to telomeres, the nature and structure of subtelomeric regions might directly act on the specific topology of chromatin at chromosome ends. Consequently, a number of factors that can influence directly or indirectly the telomere length would likely affect the expression of subtelomeric sequences by changing telomere conformation and
maintenance and vice versa. Furthermore, chromosome ends are associated with multiple pathologies and a more complete knowledge of telomere and subtelomere regulation, especially those involving epigenetic mechanisms would likely provide important insights into the role of chromosome ends in cancer, age-related diseases or response to environmental stress and infection but also numerous pathologies such as developmental disorders, mental retardation, infertility and spontaneous recurrent miscarriages. In addition the deciphering of the molecular mechanism sustaining these pathologies would provide a new avenue for the development of therapeutic approaches aimed at correcting the molecular defects caused by inappropriate modification of telomeric silencing and rearrangements of chromosome ends.

Acknowledgements

The work in Gilson lab is supported by the Ligue Nationale contre le Cancer (Equipe labellisée) and by the Association Française contre les Myopathies (AFM).

References


Klobutcher LA, Swanton MT, Donini P, Prescott DM. All gene-sized DNA molecules in four species of hypotrichs have the same terminal sequence and an unusual 3' terminus. Proc Natl Acad Sci U S A. 1981;78(5):3015-9


Renaud H, Aparicio OM, Zierath PD, Billington BL, Chhablani SK, Gottschling DE. Silent domains are assembled continuously from the telomere and are defined by promoter distance and strength, and by Sir3 dosage. Genes Dev. 1993 Jul;7(7A):1133-45


Luderus ME, van Steensel B, Chong L, Sibon OC, Cremers FF, de Lange T. Structure, subnuclear distribution, and nuclear


Cachieone S, Cerone MA, Savino M. In vitro low propensity to form nucleosomes of four telomeric sequences. FEBs Lett. 1997 Jan 2;400(1):37-41


Pryde FE, Louis EJ. Limitations of silencing at native yeast telomeres. EMBO J. 1999 May 4;18(9):2538-50


Filesi I, Cachieone S, De Santis P, Rossetti L, Savino M. The main role of the sequence-dependent DNA elasticity in determining the free energy of nucleosome formation on telomeric DNAs. Biophys Chem. 2000 Jan 24;83(3):223-37


Dynamics and plasticity of chromosome ends: consequences in human pathologies

Schluth-Bolard C, et al.


Henson JD, Neumann AA, Yeager TR, Reddel RR. Alternative lengthening of telomeres in mammalian cells. Oncogene. 2002 Jan 21;21(4):598-610


Selicorni A, Zuffardi O. Disruption of the ProSAP2 gene in a t(12;22)(q24.1;q13.3) is associated with the 22q13.3 deletion syndrome. Am J Hum Genet. 2001 Aug;69(2):261-8


Vulliamy T, Marrone A, Dokal I, Mason PJ. Association between aplastic anaemia and mutations in telomerase RNA. Lancet. 2002 Jun 22;359(9324):2168-70


Bennett L, Palucka AK, Arce E, Cantrell V, Borvák J, Banchereau J, Pascual V. Interferon and granulopoiesis

Atlas Genet Cytogenet Oncol Haematol. 2010; 14(5)


Li B, de Lange T. Rap1 affects the length and heterogeneity of human telomeres. Mol Biol Cell. 2003 Dec;14(12):5060-8


Sharma GG, Hwang KK, Pandita RK, Gupta A, Dhar S, Parenteau J, Agarwal M, Worman HJ, Wellinger RJ, Pandita TK. Human heterochromatin protein 1 isoforms HP1(Hsalpha) and HP1(Hsbeta) interfere with hTERT-telomere interactions and correlate with changes in cell growth and response to ionizing radiation. Mol Cell Biol. 2003 Nov;23(22):8363-76


Veldman T, Etheridge KT, Counter CM. Loss of hPOT1 function leads to telomere instability and a cut-like phenotype. Curr Biol. 2004 Dec 29;14(24):2264-70


Ye JZ, de Lange T. TIN2 protects telomeres from a transient DNA damage response and determines how human chromosomes end. EMBO J. 2005 Jul 20;24(14):2667-78

Zou Y, Gryzavov SM, Shay JW, Wright WE, de Lange T. POT1 protects telomeres from a transient DNA damage response and determines how human chromosomes end. EMBO J. 2005 Jul 20;24(14):2667-78

Atlas Genet Cytogenet Oncol Haematol. 2010; 14(5)


Cerruti Mainardi P, Ciuti di Chat syndrome. Orphanet J Rare Dis. 2006 Sep 5;1:33


Dimitrova N, de Lange T. MDC1 accelerates nonhomologous end-joining of dysfunctional telomeres. Genes Dev. 2006 Dec 1;20(23):3238-43


Haigis MC, Guarente LP. Mammalian siruins—emerging roles in physiology, aging, and calorie restriction. Genes Dev. 2006 Nov 20;21(29):2913-21


Murnane JP. Telomeres and chromosome instability. DNA Repair (Amst). 2006 Sep 8;5(9-10):1082-92


Denchi EL, de Lange T. Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1. Nature. 2007 Aug 30;448(7157):1068-71


Telomere dysfunction in B-chronic lymphocytic leukemia. 


