1. Introduction

Telomeres serve two vital functions to eukaryotes. They act as a protective chromosome cap to distinguish natural chromosome ends from double stranded DNA breaks and to avoid inappropriate fusions of telomeric sequences, and they maintain chromosome length by adding DNA to the ends of chromosomes. Telomeres thus balance the loss of terminal DNA due to the inability of the replication machinery to completely replicate linear DNA molecules (Olovnikov, 1973; Watson, 1972). In many cases the newly replicated chromosome ends are resected to allow for the formation of a t-loop that helps to hide the tip (Griffith et al., 1999; Wellinger et al., 1996). Most eukaryotes elongate chromosome ends with a special reverse transcriptase, telomerase, that carries a specific RNA template with telomeric sequence (Greider, 1996). The telomerase enzyme repeatedly adds copies of the short telomeric DNA sequence to the chromosome end. While there is strict conservation of telomeric sequence repeat in most species, the repeat unit has changed over evolutionary time. Holotrichous ciliates, e. g. Tetrahymena, use the sequence (TTGGGG)_n (Blackburn & Gall, 1978), while hypotrichous ciliates, e. g. Oxytricha, use (TTTTGGGG)_n (Klobutcher et al., 1981; Oka et al., 1980). The primary telomeric sequence in plants is (TTTAGGG)_n (Richards & Ausubel, 1988; Zellinger & Riha, 2007), although the alga Chlamydomonas uses (TTTTAGGG)_n. In the yeasts the telomeric sequence has the same general motif, but is not as tightly controlled. Saccharomyces for example uses (TG_t)_n (Shampay et al., 1984; Wang & Zakian, 1990), while Schizosaccharomyces has (TTACAG_t)_n (Matsumoto et al., 1987). The sequence found at the telomeres of most metazoa is (TTAGGG)_n (Meyne et al., 1989; Traut et al., 2007), although arthropods use (TTAGG)_n (Okazaki et al., 1993).

Lack of the predominant telomeric sequence in a species does not, however, signify that telomerase-generated terminal sequences are missing. For example, the metazoan-type telomeric sequence is found in place of the plant sequence in Aloe species (Weiss & Scherthan, 2002). In order to establish that a telomerase-independent, chromosome maintenance system exists it is also necessary to show a lack of a telomerase gene and telomerase activity, and to identify the nature of the DNA sequence at the chromosome termini. Establishing the negative is always difficult, and confirming that a specific sequence is at, not merely near, the chromosome tip is not trivial. Conversely, the presence of a
A completely different mechanism has been found in *Drosophila*, where tandem arrays of non-long terminal repeat (LTR) retrotransposons are found. Newly synthesized copies of these retrotransposons target chromosome ends and can even transpose to unique sequence chromosome ends. Similar telomere-specific retrotransposons have been found in *Drosophila* species that diverged as long as 40 million years ago, suggesting that this mechanism is reasonably stable. Three families of retrotransposons are found at *Drosophila* telomeres; these elements may cooperate with each other during transposition to maintain all three in the *Drosophila* genome. Mutations are known that either increase or decrease the rate of addition to the chromosome ends, leading to longer or shorter terminal retrotransposon array lengths. While these mutations have not been well characterized, they suggest that telomere maintenance by retrotransposition is genetically regulated by the host.

2. Plants without telomerase

The plant telomeric sequence (TTAGGG)$_n$ appears to be highly conserved in all phyla of the plant kingdom (Fuchs et al., 1995; Fuchs & Schubert, 1996; Richards & Ausubel, 1988). Nevertheless, in the order Asparagales the plant telomeric motif has been replaced with (TTAGGG)$_n$ but is still maintained by telomerase (Fajkus et al., 2005). In addition, three genera within the family Solanaceae appear to have lost both the canonical telomeric DNA motif as well as telomerase, which is required to maintain this motif.

2.1 The nightshade family

In the family Solanaceae the canonical plant telomeric repeat is replaced by a less conventional telomeric sequence that may be associated with a different compensation pathway. Detailed analysis of Solanaceae species revealed that although plant telomeric sequence is present in tobacco, tomato and other representatives of this family, the telomeric motif and telomerase activity are missing in the three closely related genera of *Cestrum*, *Vestia* and *Sessea*. The actual telomeric sequence and compensation mechanism in this group of plants, however, remain unknown (Fajkus et al., 1995; Peska et al., 2008; Sykorova et al., 2003; Watson & Riha, 2010).

2.2 The onion family

Chromosome termini of the onion, *Allium cepa*, and other Alliaceae species represent another known case of unusual telomeres lacking telomerase in plants. Telomeres of *A. cepa*
consist of two tandemly organized repeats – a 375-bp satellite sequence and rDNA repeats (Barnes et al., 1985; Pich et al., 1996; Pich & Schubert, 1998). Besides this, the telomeres in A. cepa are enriched with En/Spm transposable element-like sequence and Ty1-copia-like retrotransposons. The Ty1-copia retroelements have been reported not only at telomeres of A. cepa but dispersed throughout its genome (Pearce et al., 1996; Pich & Schubert, 1998). Based on these findings it has been proposed that the telomeres of Alliaceae species are maintained through transposition of the mobile elements or through homologous recombination between the satellite sequences (Pich et al., 1996; Pich & Schubert, 1998).

3. Animals without telomerase

In the case of animals, the lack of a telomerase system has been reported in a few insect species. The (TTAGG)_n sequence has been detected in most tested insect orders and is considered as the ancestral telomeric motif not only for insects but also for all arthropods (Vitkova et al., 2005). In some groups of arthropods, such as damselflies or spiders this telomeric motif was lost (Frydrychova et al., 2004; Vitkova et al., 2005), however in most cases it remains unknown if the sequence was replaced by another similar motif or a different type of sequence associated with a telomerase-independent elongation mechanism (Figure 1).

3.1 The silkworm

A highly interesting case of telomeres was revealed in another model organism, the silkworm, Bombyx mori (Lepidoptera). The telomeres of the silkworm consist of the insect telomeric repeats but harbor many types of non-LTR retrotransposons, designated TRAS and SART families (Fujiwara et al., 2005; Kubo et al., 2001; Okazaki et al., 1995). TRAS and SART are abundantly transcribed and actively transpose into TTAGG telomeric repeats in a highly sequence-specific manner. The silkworm genome contains a telomerase gene, but the telomerase itself displays little or no enzymatic activity. It is believed that compensation of telomeric loss in B. mori occurs almost exclusively by transposition of TRAS and SART elements to the chromosome ends (Fujiwara et al., 2005; Tatsuke et al., 2009). Nevertheless, in contrast to Drosophila, in which the telomerase system was completely lost and replaced by telomeric retrotransposition (see below), B. mori may be in transition from one telomere elongation pathway to another.

3.2 Lower diptera

Telomerase has not been found in any dipteran species (Figure 1). As fossils for this order date to the middle Triassic period, it is possible that telomerase may have been lost as much as 225 million years ago. Nevertheless, Diptera as a group are very successful, accounting for some 10% of known animal species. Thus, loss of telomerase does not seem to have been a major impediment to survival. Replacement of short telomerase-generated repeats with long satellite sequences is reported in lower dipteran species. Chromosome tips of non-biting midges (genus Chironomus) consist of large, 50-200 kb, blocks of complex, tandemly repeated sequences that are classified into subfamilies based on sequence similarities. Different telomeres display different sets of subfamilies, and the distribution of subfamilies differs between different individuals in a stock. The variation of the satellite sequences supports the proposal that telomeres in Chironomus are elongated by a homologous
Fig. 1. Distribution of telomere repeat motifs in Metazoa. (TTAGGG)$_n$ is the ancestral telomeric sequence of Metazoa and its sister group, Choanozoa. The ancestral motif was replaced with (TTAGG)$_n$ and (TTAGGC)$_n$ in Arthropoda and Nematoda, respectively. Tardigrada (green) do not display either of these motifs. Insect orders in red do not exhibit the arthropod sequence. Coleoptera (blue) is heterogeneous for the arthropod motif. With a few exceptions among Diptera, Tardigrada and the insect orders in color have unknown telomeric sequences. Arrows mark the replacement of the metazoan motif with other motifs, as shown. The cladogram is based on Frydrychova et al., 2004; Vitkova et al., 2005; Traut et al., 2007.
recombination mechanism involving these long blocks of complex repeat units (Biessmann & Mason, 1997; Cohn & Edstrom, 1992; Cohn & Edström, 1992; Nielsen & Edstrom, 1993). A situation has been observed in *Anopheles gambiae* with a plasmid insertion into the complex satellite telomeric sequences at the tip of chromosome 2L. The plasmid sequence was used as a marker to follow the specific telomere, which was found to engage in frequent recombination events to extend the array length (Biessmann et al., 1998; Roth et al., 1997). Recently, a similar case was reported in *Rhynchosciara americana* (Madalena et al., 2010). Tandem arrays of short repeats, 16 and 22 bp in length, were found to extend to chromosome ends. Although telomere elongation could not be assayed in this case, it seems likely that the mechanism is similar to that seen in other dipterans. In many respects, these complex arrays resemble subtelomeric sequences (Pryde et al., 1997).

### 3.3 Drosophila

Most of our information on the structure and maintenance of telomeres in *Drosophila* is based on *D. melanogaster*, although some recent studies have been performed on other species, especially *D. virilis*. As in other dipterans, *Drosophila* telomeres do not possess a canonical telomeric sequence and are not maintained by a telomerase-dependent system. Instead, chromosome ends in *Drosophila* carry an array of retrotransposons. This unusual telomere structure is common among all drosophilids that have been studied (Casacuberta & Pardue, 2002, 2005), although species within this genus may have diverged as much as 40 million years ago (Russo et al., 1995).

#### 3.3.1 Drosophila melanogaster

Three distinct telomeric regions have been identified in *Drosophila* (Andreyeva et al., 2005; Biessmann et al., 2005). At the extreme terminus is a proteinaceous chromosome cap that covers approximately 4 kb of terminal DNA sequence (Melnikova & Georgiev, 2005) and identifies the end as distinct from a chromosome break. The telomere-specific components of the cap in *Drosophila* are collectively termed ‘terminin’ by analogy to the shelterin protein complex at mammalian telomeres (Raffa et al., 2009). The terminin proteins differ from the shelterin proteins, in part because the TRF1 and TRF2 components of shelterin bind specifically to the canonical telomeric repeat, while the formation of the telomere cap in *Drosophila* is sequence independent, and in part because many of the terminin proteins are among the fastest evolving proteins in *Drosophila* (Gao et al., 2010; Raffa et al., 2010; Schmid & Tautz, 1997). There is no direct evidence that the cap in *Drosophila* plays a role in maintaining chromosome length. Most chromosome ends in *Drosophila* carry a tandem array of telomere-specific non-LTR retrotransposons (Mason & Biessmann, 1995; Pardue & DeBaryshe, 2003), although the length of this array can vary considerably. Located between the terminal retrotransposons and the unique sequence DNA of euchromatin is another repeat array. This array is often referred to as telomere associated sequences (TAS) or the subtelomere region (Karpen & Spradling, 1992; Walter et al., 1995). As in other eukaryotes TAS sequences in *Drosophila* include irregular arrays of relatively long repeat units that can vary from one chromosome end to another within the same organism (Pryde et al., 1997).

#### 3.3.1.1 Telomeric retrotransposons

Studies on *D. melanogaster* revealed three telomere-specific retrotransposable elements, *HeT-A*, *TART* and *TAHRE* (collectively abbreviated HTT) present in multiple copies on each
chromosome end. These retrotransposons are in the same family of elements as mammalian LINEs. Although the \textit{D. melanogaster} genome has some 60 families of known retrotransposable elements, only these three are found at chromosome ends. Further, these three elements are present only in the telomere arrays. HTT elements are not found in euchromatic regions, although tandem arrays of short segments of the 3' noncoding region of \textit{HeT-A} have been found in centric heterochromatin, especially in the \textit{Y} chromosome (Abad et al., 2004a; Agudo et al., 1999; Berloco et al., 2005). As a group the HTT elements have characteristics that distinguish them from other retrotransposons (Figure 2A). \textit{HeT-A} is about 6 kb in length, has only a single open reading frame (ORF), encoding a Gag-like nucleic acid binding protein, but lacks an ORF for a

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig2.png}
\caption{Structure of Drosophila telomeres. (A) There are three families of telomeric non-LTR retrotransposons. The GAG open reading frame encodes a nucleic acid binding protein that helps to target chromosome ends. The RT open reading frame encodes a reverse transcriptase needed to copy the RNA intermediate onto the chromosome end. \textit{HeT-A} does not carry a reverse transcriptase gene. All three elements carry relatively short 5' UTRs and very long 3' UTRs. Promoters are indicated by bent arrows. The 3' oligo(A) tail used to attach to chromosome ends is indicated by AAA. \textit{TART} has a strong antisense promoter. (B) The Drosophila terminal array is composed of a tandem mixed array of variably 5' truncated transposons. At the distal end the chromosome carries a protein complex that binds to the end independently of DNA sequence and stabilizes the terminus. The "A" at each junction indicates the 3' oligo(A) tail. Proximal to the retrotransposons is a complex subterminal, telomere associated sequence (TAS) followed by unique sequence chromosomal DNA. Adapted from Capkova Frydrychova et al., 2008).}
\end{figure}
reverse transcriptase. HeT-A also has a 2.4 kb 3' untranslated region (UTR), which includes two to five imprecise 80 bp tandem repeats that may help establish chromatin structure. This region has a strong G-rich strand bias, which resembles the strand bias found in canonical telomeric motifs (Abad & Villasante, 1999; Biessmann et al., 1992b; Danilevskaya et al., 1998a), suggesting selection for their presence. Despite the fact that the sequence of the 3' UTR of D. yakuba, a sister species of D. melanogaster, has diverged by about 50%, these repeat features have been conserved (Danilevskaya et al., 1998b). The TART element is about 10 kb in length with two ORFs, which encode a Gag protein and a Pol protein with a reverse transcriptase domain. TART also carries a pair of perfect non-terminal repeats that may be important for its replication (Sheen & Levis, 1994; George et al., 2010). TAHRE is about 11-13 kb in length and has extensive sequence similarity to HeT-A along its entire length, except that it carries a second ORF for a reverse transcriptase (Abad et al., 2004b; Shpiz et al., 2007). As they all carry unusually long 3' UTRs of about 2-3 kb, the HTT retrotransposons are exceptions to the pattern that transposable elements usually have very little sequence that does not code for polypeptides involved in their own transposition (Abad et al., 2004b; Biessmann et al., 1992b; Sheen & Levis, 1994). It seems likely that this non-coding DNA is related to their role at the telomere (Villasante et al., 2007).

3.3.1.2 The terminal retrotransposon array

The three retrotransposons present in Drosophila telomeres are arranged in head-to-tail arrays of mixed complete and 5'-truncated elements with their 3' oligo-A tails oriented toward the centromere (Figure 2B). HeT-A is the most abundant of the three families, accounting for 80-90% of the telomeric array. TART elements occupy about 10%, while TAHRE elements occupy only 1-2% of the telomeric array. The 5' ends of many of these elements are truncated to varying extents, as might be expected from terminal erosion due to the end replication problem or incomplete reverse transcription (Mason & Biessmann, 1995). In one stock the HTT array length varied from about 20 to 150 kb for individual chromosome ends (Abad et al., 2004a). The length and composition of the telomeric retrotransposon arrays can also vary significantly between chromosomes and among fly stocks (Walter et al., 1995). In some mutants the terminal array may be several fold longer than found in standard laboratory strains (Melnikova & Georgiev, 2002; Savitsky et al., 2002; Siriaco et al., 2002). Conversely, not all chromosome ends in Drosophila have terminal retrotransposon arrays. Broken chromosomes with ends far from the original telomere have been found in a number of different circumstances (Capkova Frydrychova et al., 2008). It is important to note that these broken chromosome ends lack both the retrotransposon array and TAS but do not induce cell cycle arrest and are not subject to DNA repair or telomere fusions. These broken chromosome ends can be maintained in vivo for hundreds of generations without gaining new HTT sequences (Biessmann et al., 1990a; Cenci et al., 2003; Fanti et al., 1998). Thus, these broken chromosome ends have 'healed' in the sense that McClintock (1941) described newly stabilized broken chromosome ends, and they are associated with a protein complex that includes a number of terminin proteins (Cenci et al., 2005; Ciapponi & Cenci, 2008). This suggests that chromosome caps in Drosophila are epigenetic and form independently of telomeric DNA sequence (Biessmann & Mason, 1988, 2003). These capped broken ends may eventually acquire retrotransposons by what appears to be a stochastic process (Biessmann et al., 1992a; Mikhailovsky et al., 1999). As expected from the end replication problem, the broken chromosome ends recede. Erosion at these terminally deficient chromosomes was estimated at a constant rate of about 75 bp
per sexual generation (Biessmann & Mason, 1988; Levis, 1989; Mikhailovsky et al., 1999). Considering the number of germline cell divisions, the rate of terminal erosion was estimated at 2-3 bp per chromosome end per cell cycle (Biessmann & Mason, 1988). This is formally equivalent to the degradation of an 8-12 nt RNA primer from the end of the lagging strand after each round of replication leaving a short 3' overhang (Biessmann et al., 1990a). In mammals telomere erosion is faster, in large part because chromosome ends are resected to produce relatively long 3' overhangs necessary for t-loop formation (Griffith et al., 1999; Wellinger et al., 1996). The slow rate of loss in Drosophila suggests that resection of the chromosome ends after replication is not extensive, and that t-loops are not required for telomere protection. It is possible that chromosome ends with telomeric retrotransposons behave differently from healed broken ends, but to date the evidence is lacking.

### 3.3.1.3 Transposition to elongate telomeres

To counter terminal erosion and maintain their length, telomeres must be elongated. The broken chromosome ends provide an entrée to study telomere elongation. Southern blots to monitor terminal fragment length at a broken chromosome end identified the addition of new sequence onto the terminal fragment at a frequency of about 1% per generation, with an average length for the added fragment of 6 kb (Biessmann et al., 1992a; Biessmann et al., 1990b). This averages out to an addition of 60 bp per generation, just enough to balance terminal erosion. It is important to note that the frequency of addition onto chromosome ends may be sensitive to different factors, including genetic background and possibly external conditions. Using genetic assays, two groups have identified stocks with much different frequencies of addition, possibly by as much as two orders of magnitude in either direction (Golubovsky et al., 2001; Savitsky et al., 2002; Savitsky et al., 2006). The new additions onto the receding chromosome ends were identified as HeT-A and TART elements, the same retrotransposons as found at natural telomeres. These elements were attached to the broken end by an oligo(A) tail, as would be expected from retrotransposition (Figure 3). Further, when broken chromosome ends that had gained a HeT-A element were used as a target they too acquired new HeT-A elements by transposition (Biessmann et al., 1992a). Thus, retrotransposition through target primed reverse transcription may be a mechanism for extending natural chromosome ends as well as broken ends.

#### 3.3.1.3.1 The transcription step

The first step in the process of retrotransposition is transcription of the transposable element (Figure 3). All three telomeric elements have unusual transcription patterns. TART has active promoters at both the 5' and 3' ends that initiate in both the sense and antisense directions, although the major product seems to be a nearly full length antisense RNA (Danilevskaya et al., 1999; Maxwell et al., 2006). TART has a single promoter in the 5' UTR that drives transcription of the transposition intermediate (Maxwell et al., 2006). HeT-A and TAHRE, on the other hand, do not have promoters at the 5' end. Instead, they have a promoter in the 3' end that drives transcription of the adjacent downstream element (Danilevskaya et al., 1997; Shpiz et al., 2007). This literally means that HeT-A promotes its neighbor. The placement of this promoter is important for the long-term integrity of the telomeric array, because a promoter in the standard 5' position would be subject to erosion due to the end replication problem and lost immediately after transposition. A 3' promoter resurrects the element downstream. HeT-A transcription is developmentally regulated and occurs only in diploid cells of ovaries, testes, imaginal discs, and embryos (Capkova Frydrychova et al., 2007).
Fig. 3. Transposition as a mechanism for telomere elongation. The model proposes that transcripts (colored arrows) are generated from telomeric retrotransposons using promoter activity located in the 3' UTR of an upstream HeT-A or TAHRE element. Transcripts leave the nucleus to serve as mRNA for translation of the encoded Gag protein and possibly reverse transcriptase (ovals). Gag proteins bind the RNA, facilitate re-entry into the nucleus and target the chromosome end. After docking to a telomere a reverse transcriptase uses the free 3' hydroxyl group at the chromosome end as primer to copy the RNA intermediate into the first DNA strand. Second strand synthesis completes the addition of a new retrotransposon. Sequence analyses of recently transposed HeT-A elements and several in native telomeric arrays suggest that there is a selection for the incorporation of elements with a functional Gag ORF.

One appealing mechanism for controlling the length of the terminal retrotransposon array is to regulate transcription of these elements. Two forms of this mechanism have been proposed, but both have problems. First, it was noted that transgenes inserted into subtelomere regions are repressed and variegate (Cryderman et al., 1999; Roseman et al., 1995). In addition, TAS arrays can silence *in cis* the activity of a neighboring transgene as well as a HeT-A element when they are distal (telomeric) of TAS (Boivin et al., 2003;
Capkova Frydrychova et al., 2007; Kurenova et al., 1998), suggesting that TAS can control terminal retrotransposon array length by regulating transcription (Mason et al., 2003a). The TAS silencing effect, however, only extends a short distance into the terminal array and has little or no effect on overall HeT-A transcript levels (Biessmann et al., 2005; Capkova Frydrychova et al., 2007). Thus, silencing orchestrated by TAS arrays is not sufficiently strong to regulate transcription of the terminal retrotransposons. Second, it was noted that HeT-A and TART transcript levels are under the control of an RNA interference pathway (Savitsky et al., 2006; Shpiz et al., 2009). Transposition frequency and terminal array length, however, did not increase with increasing retrotransposon transcript levels (Capkova Frydrychova et al., 2008). It thus appears that the transcript levels of these retrotransposons are not the limiting factor in their transposition.

### 3.3.1.3.2 Telomere targeting

After transcription the RNA is transported into the cytoplasm and translated. The HeT-A RNA produces only a Gag protein, which binds a transcript, enters the nucleus and attaches to chromosome ends. Evidence supports the hypothesis that the Gag protein binds preferentially to the transcript that encoded it, because while many HeT-A elements in the terminal array are 5’ truncated or otherwise lack an ORF, newly transposed HeT-A elements have a complete Gag ORF (Biessmann et al., 1994). Unlike Gag proteins for closely related parasitic retrotransposons, the HeT-A and TART Gag proteins are transported efficiently into the nucleus (Rashkova et al., 2002b). Unlike other non-LTR elements telomere specific elements do not require nicked DNA, because they are reverse transcribed directly onto the end of the chromosome. The HeT-A Gag can associate with telomeres on its own. The TART Gag, however, can only be seen to associate with telomeres in the presence of expressed HeT-A Gag protein (Rashkova et al., 2002a). Similarly, transport of the TAHRE Gag into the nucleus is facilitated by HeT-A and TART Gag proteins (Fuller et al., 2010). This presents a possible explanation for the presence of HeT-A, TAHRE and TART transposons in all Drosophila stocks. HeT-A does not encode a reverse transcriptase, which is required for retrotransposition, but may use the one encoded by either the TART or TAHRE elements. TAHRE and TART, on the other hand, cannot target chromosome ends without the aid of the HeT-A Gag protein.

### 3.3.1.3.3 Consequences of transposition on terminal array structure

Given the constant erosion of chromosome ends and the stochastic addition of transposon sequences to the same ends, one might expect that the terminal retrotransposon array would be very dynamic, constantly changing in length and composition. We have found this to be true using a genetic assay for the number of transposons at a specific telomere (Golubovsky et al., 2001; Mason et al., 2003b). One consequence of this turnover is that the transposon elements at the terminus are younger than those nearer to TAS. Virtually complete terminal arrays have been identified in overlapping BAC clones (Abad et al., 2004a). The age differential can be seen in the distribution of transposable elements that do not specifically target the chromosome end. These transposons are found primarily in the older, proximal portion of the terminal array (Pardue & DeBaryshe, 2008). Turnover in the younger, distal portion of array the removes evidence of these transposons. Newly transposed P elements have also been found inserted into the terminal array with reasonable frequency (Biessmann et al., 2005). Although the exact positions of these P elements in the terminal array could not be determined for most of the insertions, there is no evidence that any portion of the HTT array is refractory to insertion by non-telomere-specific elements.
HeT-A transcription start sites reside 31 and 62 bp upstream of the oligo-A tail (Danilevskaya et al., 1997). Thus, newly transposed transposons are slightly longer than the same elements before transposition, because they carry a tag at the 5' end identical to the 3' end of the previous upstream element (Traverse et al., 2010). Surprisingly, some of the elements carry multiple tags, suggesting that they have transposed several times without being subjected to terminal erosion. The simplest explanation is that multiple transposition events occur in rapid succession, possibly more than one per generation. Evidence of this has been found in measurements of transposition rate (Biessmann et al., 1992a), in which it was found that many of the new sequence additions were 12 kb or longer, and the 3' half of these long additions consisted of a (6 kb) HeT-A element. As these long addition events were shown to be the result of transposition, they could have resulted from either rapid multiple transposition events or transcriptional read-through to produce an RNA intermediate encompassing more than one element. The latter, however, have been found to be relatively rare (Capkova Frydrychova et al., 2007) and don't explain the presence of tandem tags. Rapid multiple transposition events may be the natural consequence of terminal transposition. When one retrotransposon attaches to the chromosome end, the old protective telomere cap must jump to the new terminus 6-12 kb away. If the cap is unstable during this transition, more transposons may have access to the new terminus, allowing for more transposition events to occur in rapid succession.

The 5' and 3' UTRs of TART carry perfect non-terminal repeats. These repeated regions vary among TART families and among individuals within a family but are identical at both ends of individual elements (Sheen & Levis, 1994). It has been proposed that the two repeated sequences evolve in concert by a mechanism of template switching during the reverse transcription step (George et al., 2010).

### 3.3.1.4 Recombination to elongate telomeres

Transposition is not the only mechanism for telomere elongation in *Drosophila*. Gene conversion allows genetic information to be transferred from one chromosome to another by homologous recombination (Figure 4). Georgiev and colleagues made use of broken chromosome ends with the *yellow* gene placed close to the terminus, such that the upstream controlling sequences were deleted, but the ORF was still present (Kahn et al., 2000; Mikhailovsky et al., 1999). Expression of the *yellow* gene was thus inactivated, but HeT-A transposition to the broken end could activate *yellow* expression via the promoter in its 3' UTR, while recombination with a wild type chromosome can reintroduce the *yellow* promoter and enhancers to their position on the broken end. Genetic assays were used to identify changes in *yellow* expression, then the length and sequence of the upstream region were characterized. Approximately 20-30% of the *yellow* reactivation events were the result of gene conversion. In one study the average length of the conversion track was estimated at 2.7 kb (Mikhailovsky et al., 1999), in another conversion tracks exceeding 20 kb were found (Kahn et al., 2000).

Although experiments using broken chromosome ends to monitor telomere elongation use an artificial system of telomere maintenance, it is assumed that the telomere elongation mechanisms identified in these experiments also work at the ends of long retrotransposon arrays. Extensions of long terminal arrays by individual transposition events or short gene conversion tracks cannot be monitored genetically or molecularly. If, however, genetic factors cause an imbalance between elongation and erosion, terminal retrotransposon arrays may grow or shrink. This can be measured cytologically by *in situ* hybridization on polytene
chromosomes as changes in the terminal array length, or molecularly by quantitative PCR as changes in the genomic copy number of the telomere-specific retrotransposons.

![Gene conversion as a mechanism for telomere elongation.](image)

Fig. 4. Gene conversion as a mechanism for telomere elongation. The model proposes that the 3' strand of a chromosome end invades another chromosome, possibly a sister or homologue. The invading strand is extended using the host sequence as a template then is used as a template in second strand synthesis. Ligation of the newly replicated fragment results in an extended chromosome.

Mutations have been identified in three genes that cause terminal transposon array length to increase. Telomere length is sensitive to HP1 concentration, as mutations in the gene encoding this protein lead to an increase in HeT-A and TART transcript levels and a 100 fold increase in the frequency of new HeT-A and TART attachments (Savitsky et al., 2002). The increased rate of elongation resulted from both transposition and gene conversion, and was associated with extremely long terminal array length after several generations. As HP1 protein is enriched in
the telomere cap, at least two possible hypotheses present themselves: (1) increased transcription increases transposition of the telomere-specific elements, and (2) disruption of the cap by decreasing one of its component proteins increases accessibility of the transposons. Our data (RCF and JMM, unpublished data), however, suggest that neither is true. Other mutations associated with increased HeT-A transcript levels are not associated with long telomeric arrays, and disruption of the cap by making heterozygous mutations in other genes encoding cap proteins does not increase telomeric array length. Two dominant mutations, E(tc) and Tel (Melnikova & Georgiev, 2002; Siriaco et al., 2002), exhibit abnormally long telomeres and are located in the same small genetic region in the middle of chromosome 3R. In the Tel mutant the copy number of HeT-A at telomeres is increased seven-fold, while TART and TAHRE copies are increased somewhat less (Siriaco et al., 2002; Walter et al., 2007). The mechanism of action of these mutations has not been elucidated, although one study indicated that the Tel mutation causes telomere elongation by transposition as well as gene conversion, while E(tc) causes mainly gene conversion (Proskuryakov & Melnikova, 2008).

3.3.2 Drosophila virilis

The DNA sequence of individual HeT-A and TART elements in D. melanogaster differ considerably throughout their lengths, but especially in the 3' UTR regions. Although it is possible to identify families of these elements, there is still some variation within each family. A comparison of telomeric retrotransposons between two sibling species of Drosophila, D. melanogaster and D. yakuba, shows substantial divergence in the HeT-A and TART UTRs but less divergence in the ORFs that encode the Gag-like polypeptides (Casacuberta & Pardue, 2002; Danilevskaya et al., 1998b). These two elements have the most amino acid sequence conservation around the zinc knuckle motif typical of Gag proteins, and there are conserved islands scattered throughout the coding region. The overall structure of the elements, however, is well conserved. HeT-A in D. yakuba also lacks a pol ORF and has a very long 3' UTR. Although the high sequence divergence of the telomeric elements makes it difficult to find these elements in new species, it also increases the probability that the conserved features are of biological importance.

Searching for telomeric retrotransposons in more distantly related species presents a problem because of the extensive sequence divergence. Only the most conserved part of the D. melanogaster TART pol gene can cross-hybridize, even at low stringency, with D. virilis DNA. This hybridization, however, allowed the isolation of DNA fragments that provided entry into the D. virilis telomere arrays (Casacuberta & Pardue, 2003a). The D. virilis TART resembles its D. melanogaster homolog in several respects. They are both found in tandem arrays, but not in the euchromatic arms, and they both produce an excess of antisense transcripts. The TART in D. virilis is different in that it has a relatively short 3' UTR without perfect non-terminal repeats and a pol gene (ORF2) that encodes an additional ‘X’ domain 3' to the reverse transcriptase domain. A HeT-A element was found in the terminal array next to a TART from D. virilis. As with its homolog in D. melanogaster, the HeT-A element carried only a single ORF for a Gag protein and had a long 3' UTR (Casacuberta & Pardue, 2003b). Experiments to localize GFP-tagged Gag proteins indicated that the Gag encoded by the D. virilis TART element requires HeT-A Gag to target the telomeres, similar to the situation found in D. melanogaster (Casacuberta et al., 2007).

There are significant differences between the telomere specific elements in D. melanogaster and D. virilis. HeT-A in D. virilis has its promoter in the 5' UTR, similar to nontelomeric
retrotransposons and thus produces transcripts that lack 5’ tags. Even so, full length HeT-A elements persist in the array. The D. virilis TART, on the other hand, has a 3’ promoter that generates 5’ tags on its transcripts (George et al., 2010; Traverse et al., 2010). This arrangement is the reverse of that found in D. melanogaster. Unlike in D. melanogaster, the HeT-A 5’ UTR in D. virilis is highly conserved. This suggests a different transposition mechanism for HeT-A in these two species. Unlike the situation in D. melanogaster, in which the retrotransposons attach to the chromosome terminus, it is possible that the HeT-A element in D. virilis inserts into the 5’ UTR of other elements already in the array by making a sequence-specific nick followed by target primed reverse transcription. This proposed mechanism resembles that used by canonical non-LTR retrotransposons. The ‘X’ domain specific to the D. virilis pol gene may play a role in this endonuclease activity. If it is true that HeT-A elements behave differently in these two distantly related Drosophila species and that HeT-A in D. virilis uses a mechanism similar to nontelomeric retrotransposons, then retrotransposon telomeres may have arisen near the dawn of Drosophila. HeT-A transposons in the Sophophora subgenus, which includes D. melanogaster, may have lost the endonuclease needed to nick chromosomal DNA to initiate insertion, while HeT-A transposons in the Drosophila subgenus, which includes D. virilis, may have retained the endonuclease but made it sequence specific. Further, if telomeric retrotransposons arose in the Drosophila genus, it follows that other Dipteran species may have other means of controlling telomere length. This is consistent with the finding of complex tandem sequence arrays at the extreme chromosome ends in Chironomus and Sciariidae species (Cohn & Edstrom, 1992; Cohn & Edström, 1992; Madalena et al., 2010; Nielsen & Edstrom, 1993).

4. Conclusion

Although rare, telomerase has been lost several times in plants and animals. In some cases, such as Diptera, telomerase was lost in the distant past, and the descendents of this event have thrived and diversified. This raises the possibility that, once established, organisms with noncanonical mechanisms of telomere maintenance may not be at a severe selective disadvantage. How, then, do we account for the paucity of organisms lacking telomerase? One possibility is that there is a strong selective barrier to the loss of telomerase-generated DNA motifs. Binding of the shelterin protein complex necessary for the protection of chromosome ends depends on sequence-specific binding. The shelterin components TRF1 and TRF2 in particular recognize the double stranded telomeric motif, while POT1 recognizes the single stranded form (Palm & de Lange, 2008). Thus, in most cases loss of telomerase results in loss of the telomeric motif, followed by loss of the chromosome cap, massive chromosome rearrangement and death. If, however, telomeric attachment of the cap complex in some lineages does not depend strongly on a specific DNA sequence, loss of the telomeric motif might not have the same catastrophic consequences. This might explain why insects have lost telomerase and the canonical arthropod-type telomeric sequence multiple times (Figure 1). This hypothesis assumes the existence of an effective backup mechanism that can replace the canonical telomerase system. Alternatively, it may be misleading to suggest that loss of telomerase occurred only in the small number of organisms already reported. There may be, for example, cases similar to B. mori, in which an unconventional telomere structure maintained without telomerase is camouflaged by the presence of canonical telomeric sequences. Loss of telomerase in Solanaceae, Alliaceae and
insects may thus represent the tip of the proverbial iceberg, and it is possible that many other groups will be found with unusual telomere structures.

Recombination seems like a ready backup mechanism for telomere maintenance by transferring information from one DNA strand to another, because it is widely used by eukaryotes both during meiotic recombination and as a means of repairing DNA damage (Heyer et al., 2010). Some organisms are known to use recombination to maintain the canonical telomeric DNA sequence in the absence of telomerase. In *Saccharomyces cerevisiae*, for example, loss of telomerase causes gradual loss of the terminal array and ultimately cell death. A few survivors, however, appear in a recombination-dependent manner (Lundblad & Blackburn, 1993). In mammals telomerase activity is reduced in most somatic tissues, thus limiting the growth of tumors. Most cancer cells overcome this obstacle by reactivating telomerase, but about 15% use an alternative lengthening of telomeres mechanism, which is recombination-dependent (Cesare & Reddel, 2010). Dysfunctional telomeres may actually stimulate telomeric recombination (Brault & Autexier, 2011). It is thus reasonable to hypothesize that a recombination-based gene conversion mechanism would be available if telomerase fails. As seen in yeast and human tumors, this pathway can maintain telomeric repeats. If the canonical motif is lost another DNA sequence, possibly related to the complex arrays of subtelomeric regions, could be maintained by the same mechanism. It is difficult to prove that organisms without telomerase use gene conversion to elongate their chromosome ends, in large part because many of these organisms lack the genetic tools to test the hypothesis. Recombination could be demonstrated in the malaria vector *Anopheles gambiae* because of the fortuitous transgene insertion into the telomeric array (Roth et al., 1997). In other cases it can be shown that a complex repeat array extends to, or close to, the chromosome end (Madalena et al., 2010; Nielsen & Edstrom, 1993). In these cases gene conversion was suggested as the default mechanism. Regardless of the mechanism, it is clear that a few well established lineages of both plants and animals lack telomerase and the canonical telomeric DNA motif it produces. Elucidation of these unusual telomeres will help us to understand what it means to be a telomere.

5. Acknowledgments

We thank František Marec, Jan Zrzavý and Sailesh Surapureddi for valuable comments on the manuscript. The authors were supported by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences, the MOBITAG project (Reg. No. 229518 of EU program FP7-REGPOT-2008-1) and the Institute of Entomology project Z50070508

6. References


The study of DNA advanced human knowledge in a way comparable to the major theories in physics, surpassed only by discoveries such as fire or the number zero. However, it also created conceptual shortcuts, beliefs and misunderstandings that obscure the natural phenomena, hindering its better understanding. The deep conviction that no human knowledge is perfect, but only perfectible, should function as a fair safeguard against scientific dogmatism and enable open discussion. With this aim, this book will offer to its readers 30 chapters on current trends in the field of DNA replication. As several contributions in this book show, the study of DNA will continue for a while to be a leading front of scientific activities.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:
