Evolution of species-specific promoter-associated mechanisms for protecting chromosome ends by Drosophila Het-A telomeric transposons

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1073/pnas.1000612107">http://dx.doi.org/10.1073/pnas.1000612107</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>National Academy of Sciences (U.S.)</td>
</tr>
<tr>
<td>Version</td>
<td>Final published version</td>
</tr>
<tr>
<td>Accessed</td>
<td>Fri Jun 23 06:06:16 EDT 2017</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/61409">http://hdl.handle.net/1721.1/61409</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Article is made available in accordance with the publisher’s policy and may be subject to US copyright law. Please refer to the publisher’s site for terms of use.</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td></td>
</tr>
</tbody>
</table>
Evolution of species-specific promoter-associated mechanisms for protecting chromosome ends by Drosophila Het-A telomeric transposons

Karen L. Traverse¹, Janet A. George³, P. G. DeBaryshe, and Mary-Lou Pardue²

Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139

Contributed by Mary-Lou Pardue, January 19, 2010 (sent for review December 3, 2009)

The non-LTR retrotransposons forming Drosophila telomeres constitute a robust mechanism for telomere maintenance, one which has persisted since before separation of the extant Drosophila species. These elements in D. melanogaster differ from nontelomeric retrotransposons in ways that give insight into general telomere biology. Here, we analyze telomere-specific retrotransposons from D. virilis, separated from D. melanogaster by 40 to 60 million years, to evaluate the evolutionary divergence of their telomeric traits. The telomeric retrotransposon Het-A from D. melanogaster has an unusual promoter near its 3′-terminus that drives not the element in which it resides, but the adjacent downstream element in a head-to-tail array. An obvious benefit of this promoter is that it adds nonessential sequence to the 5′ end of each transcript, which is reverse transcribed and added to the chromosome. Because the 5′ end of each newly transposed element forms the end of the chromosome until another element transposes onto it, this nonessential sequence can buffer erosion of sequence essential for the telomere. Thus, every 5′ end of the element persists in telomere arrays, necessitating a more precise processing of the extreme end of the telomere in D. virilis.

telomeres | retrotransposons | end replication problem | promoter evolution

Het-A, TART, and TAHRE are non-LTR retrotransposons that maintain the telomeres in Drosophila melanogaster. They transpose only to chromosome ends, where they form long arrays of head-to-tail repeats that make up the telomeres. Despite their differences, this retrotransposon mechanism for extending chromosome ends is functionally equivalent to the telomerase mechanism. In each case, an RNA template is reverse transcribed to add DNA repeats to the chromosome. Telomerase repeats are short sequences copied from part of the enzyme’s RNA component. In Drosophila, each repeat is a copy of one of the three telomere-specific retrotransposons. These retrotransposons have unusual features, some in common, some not, but all presumably related to their role at the telomere (reviewed in refs. 1, 2). This study of species differences in the Het-A promoters lends insight into the mechanisms by which Het-A elements protect chromosome ends.

A notable feature of D. melanogaster Het-A is its promoter, which has unexpected similarities to promoters of retroviruses and LTR retrotransposons (3). Het-A promoter sequences are found at the 5′ end of each element, in the 3′ untranslated region (UTR). They direct transcription of the neighbor immediately downstream, rather than the element in which they reside (see Fig. 1). To an outside observer, if the downstream element is another Het-A, the combined sequence (consisting of the 3′ sequence of the upstream element plus the entire downstream element) appears to be, and is formally equivalent to, an LTR retrotransposon or a retrovirus. Like the Het-A 3′ promoter, 5′ LTRs contain the promoter and transcription start sites and are identical to the 3′ sequence of the transcribed element. An important difference is that, unlike a true 5′LTR, the 3′ sequences carrying the Het-A promoter region are not attached to the element they regulate and do not move to its new site. Thus, when Het-A transposes to a new site, it cannot be replicated unless it acquires another Het-A element immediately upstream. Although many Het-A elements have appropriate neighbors in their long head-to-tail arrays in telomeres, this feature exacts an evolutionary cost on any Het-A not immediately preceded by a viable Het-A promoter, raising questions about the evolutionary tradeoffs involved.

Because Het-A transcription starts in the upstream element, each transcript carries on its 5′ end a short segment of sequence from its neighbor. We refer to the sequence copied from the neighboring element as a “tag” because it is not an essential part of the element (4). A tag almost certainly never has enough upstream sequence to serve as a promoter. Instead the 5′ tags provide nucleotides that can be lost without diminishing the element, thus serving as protection from terminal erosion. This protection is crucial because the 5′ end of the most distal element in each telomere forms the end of the chromosome and is at risk for under-replication (the well known end replication problem) or degradation.

The protective effect of the 5′ tags can be seen in sequences of Het-A elements in telomere arrays. Although Het-A elements can be variably 5′-truncated, many are complete and almost all complete elements have two or more 5′ tags. The D. melanogaster Het-A promoter has three transcription starts, which would yield tags of 31, 62, or 93 nucleotides (3, 5). Tags in the telomere arrays are variably 5′-truncated, presumably because they undergo some sequence loss while protecting the end of the element. Because a new tag is added each time an element is transcribed, the number of 5′ tags on each element should increase continuously, reflecting the number of times that element has transposed, much as rings indicate the age of a tree. Of course, tags could also be completely lost while buffering the end. In fact, however, there must be a mechanism preventing accumulation of long strings of tags because six tags are the most that have been seen at a single junction, yet our analyses of BACs from XL and 4R (4; updated in Flybase, http://Flybase.org/cgi-bin/gbrowse/dmel/) show that most of the complete elements available for transcription have four to six tags (median, 6.0), ranging in length from six to 68 nt (median, 17 nt).

The fact that Drosophila uses retrotransposable elements, rather than telomerase, to maintain its telomeres raises interesting questions about the evolution of both telomeres and retrotransposable elements. Het-A and TART maintain telomeres in all Drosophila species that have been studied, including D. virilis, which is separated from D. melanogaster by 40 to 60 million years (6). Although their sequence evolves faster than most Drosophila genes (7), Het-A elements have the same basic structure and


The authors declare no conflict of interest.

¹K.L.T. and J.A.G. contributed equally to this work.

²To whom correspondence should be addressed. E-mail: mlpardue@mit.edu.
share many unusual features in *D. melanogaster* and *D. virilis* (8, 9). Surprisingly, none of the complete HeT-A elements in *D. virilis* have 5’ tags, suggesting that *D. virilis* and *D. melanogaster* differ in their mechanisms for protecting the extreme end of the telomere.

In the studies reported here we have characterized the *D. virilis* HeT-A promoter in order to define the 5’ end of this element and understand how it is protected from degradation. We find that the *D. virilis* promoter has no mechanism for adding expendable nucleotides and thus must have more stringent end-protection than *D. melanogaster*. Instead, HeT-A′′ RNA shows extremely high sequence conservation at the 5’ end. We speculate that these conserved sequences may position protein and/or RNA components that prevent loss of sequence from the 5’ end of *D. virilis* HeT-A.

**Results**

**Comparison of the 3′ and 5′ UTR Sequences of HeT-A′′ and HeT-Amel.** Non-LTR retrotransposons transpose by target-primed reverse transcription (10). Normally, the 5’ end of the retrotransposon RNA associates with a nick in chromosomal DNA and reverse transcription is primed from the 3’OH of the nicked DNA. In contrast, telomere elements appear to be primed from the end of the chromosome, rather than a nick (4). Because reverse transcription begins within the 5’ poly(A) tail, the 3’ ends transposed elements are easy to identify. However elements in the telomeric arrays can be variably 5’-truncated. To determine the complete 5’ sequence of *D. virilis* HeT-A, we analyzed the *D. virilis* λ-phage clones from which we initially identified HeT-A′′ (9) and also searched the *D. virilis* genome database, although few of these variable repeated sequences are complete and assembled in this database. The longest HeT-A elements that were found all joined their upstream elements at almost the same nucleotide; we considered these to be complete elements, a decision validated by this promoter study. In Fig. 2 we show sequence surrounding the junction of five of these elements with their upstream HeT-A neighbors.

Looking for sequence similarities between the *D. melanogaster* HeT-A promoter and *D. virilis* HeT-A sequences, we examined junctions consisting of upstream 3’UTRs linked to the adjacent downstream, 5’UTR. These regions of the RNA were not similar enough to give meaningful nucleotide alignment between species. Within each species, UTRs of different HeT-A elements show numerous indels and nucleotide changes and differ slightly in length. In *D. melanogaster* both UTRs are variable: differences are scattered throughout the length of the UTRs, although the last 29 nucleotides of the 3’UTR are absolutely conserved. In *D. virilis* the first 1.4 kb of the 5’UTR is extremely conserved and the 3’-most 69 nucleotides of the 3’UTR are absolutely conserved in each complete element (Fig. 3).

**D. virilis HeT-A Promoter Sequences Are in the 5’UTR.** The significant differences between the sequences of the two species led us to characterize the promoter sequences of *D. virilis* HeT-A: we tested the ability of segments of HeT-A′′ 3′ and/or 5′UTR sequence to drive expression of a bacterial lacZ reporter gene. Reporter constructs were transiently transfected into cultured *D. virilis* cells, which endogenously express HeT-A RNA, thus providing a biologically relevant test system. Promoter activity was evaluated from the activity of the lacZ enzyme.

**5′UTR Activity.** All of the constructs with promoter activity included nucleotides +1 to +89 of the 5’UTR (Fig. 4). Thus this region appears to contain the minimal promoter, a conclusion supported by its sequence where we found good matches to the

---

*Fig. 1.* Diagram of *D. melanogaster* HeT-A transcription and subsequent transposition. Top diagram shows adjacent elements from the interior of a telomere array. [Dark gray: 5′ and 3′ UTR. Light gray: Gag coding regions. White arrowheads: 3′ oligo(A). Bent arrows: transcription start sites.] The promoter of the central element (star at transcription start) directs transcription of the element on the right. Steps 1 through 4 show the resulting RNA transcript with a 5′ tag of sequence from the element supplying the promoter. This RNA is reverse transcribed onto a chromosome end, with some or all of the tag remaining on the 5′ end of the newly added element. When this new end is extended by reverse transcription of another RNA, the promoter on the new terminal element (star) directs transcription of an RNA with two tags, sequence from its own 3′ end plus the tag remaining on the element that is transcribed. The chain of tags can grow if this RNA transposes to continue the cycle; however all or part of the 5′-most sequence can also be lost. For simplicity only a minimum of tags are shown here.

*Fig. 2.* Sequence at five junctions between head-to-tail *D. virilis* HeT-A elements, taken from a multiple alignment of tandem HeT-A elements with apparently complete 5′UTRs on the downstream element. This section surrounds the junctions (arrow) between the 3′ ends of upstream elements and the downstream 5′ ends of the neighboring elements. [The 3′ end of each upstream element is indicated by an oligo(A) sequence, the length of which is determined by the site at which reverse transcriptase initiated synthesis on the poly(A) tail of the transcription intermediate and is presumably much shorter than that poly(A) tail.] We have arbitrarily marked the junction on this alignment at the end of the shortest oligo(A) in the set. Nucleotides in lowercase bold immediately after this arrow are the only nucleotides not conserved in all five elements. For their origins, see Discussion. Bent arrows mark the two nucleotides where transcripts started in our experiments (results). Boxed sequences next to these arrowheads are a match to Inr. Boxed sequences on right are a match to DPE. DPE sequence begins at position +28 nt from the A (+1) at the transcription start site. Sequence 1 was used for the constructs in this study.

*Fig. 3.* Diagram of 3′ and 5′UTR sequences separating ORFs of tandem *D. virilis* HeT-A elements. The line linking the two ORFs summarizes results of a multiple alignment of full length elements, a small section of which is seen in Fig. 2. The 3′ and 5′UTRs make up approximately 3.8 kb of sequence but vary in length slightly from element to element. Thin lines denote regions where elements differ by multiple in-dels and base changes. The dark gray box on this line marks the region where all five elements are identical or nearly so. Vertical line marks the junction of the two elements (arrow in Fig. 2). As indicated above the line, the region of identity consists of the final 69 bp in the 3′UTR and 1.4 kb in the 5′UTR.
lacZ′−′ Have Evolved Different Ways to Protect the 5′+1,731) had little effect, whereas

Adding the 49 3′ tags before they reach internal End of the Element.

D. virilis sequences have two roles at the telomere. First, the end while transposing and later serving as the cells. Diagram shows junction between two ele-

quences necessary to get reasonable promoter activity. 

Retrotransposable elements must regulate their own constructs (Fig. 4) with signi-

end of the 5′ construct were 

show that complete 

were tested for antisense promoters but produced no activity in our 

this end of the 

www.pnas.org/cgi/doi/10.1073/pnas.1000612107

HeT-A′′ and HeT-A′′′′′′ Non-LTR Elements with Very Different Promoter Architectures. Retrotransposable elements must regulate their own transcription after they transpose; otherwise they will not spread and multiply. Typically, non-LTR retrotransposons do so by locating promoter sequences in the 5′UTR, downstream of the transcription start. The regulatory sequences are included in the RNA transcript and reverse-transcribed into the new site, enabling the new element to continue transposing. Promoters of several Drosophila non telomere elements, jockey, Doc, G, F, and I element, which never appear in the telomeres, have been charac-

terized. All have the typical 5′UTR promoter, with an Inr sequence defining the transcription start site and a DPE sequence precisely spaced downstream (12). As noted, the D. virilis HeT-A promoter uses this same structure. 

D. melanogaster HeT-A, in contrast, is the only non-LTR element found to have a promoter resembling that typical of LTR-containing retroelements (3). Because the HeT-A promoter is not attached to the element it regulates, we consider it a “pseudo LTR.”

HeT-A′′ and HeT-A′′′′′′ Have Evolved Different Ways to Protect the 5′ End of the Chromosome. This promoter study has identified the transcription start sites that define the true 5′ end of HeT-A in D. virilis. The 5′ end of each newly added element becomes the chromosome terminus until the next element transposes onto the telomere. Therefore these promoter studies give a glimpse of the terminus of the D. virilis HeT-A telomere. The 5′ ends of these newly transcribed elements can be compared with the ends of elements that are no longer at risk for terminal loss because they have become interior elements in the telomere array. These comparisons show that a large fraction of D. virilis elements have kept an intact 5′ end while transposing and later serving as the chromosome end. In contrast, D. melanogaster HeT-A elements undergo some erosion of their 5′ tags before they reach internal positions. Thus the two species appear to have different mech-

anisms for 5′ end protection.

Although different, these protective mechanisms appear to be equally effective for the two species. Our analyses of the available assembled telomere sequences from D. melanogaster and D. virilis show that complete HeT-A elements make up more than half of the total HeT-A sequence in telomeres in both species. The remainder of the HeT-A sequence is in elements that are 5′-truncated to varying degrees, many rather extensively. We note that HeT-A sequences have two roles at the telomere. First, the sequences form telomere-specific chromatin, analogous to chromatin formed by short telomerase repeats. For this, 5′-truncated elements are probably as good as intact elements because these truncated elements are enriched in the conserved sequences of the 3′UTR (15). Second, HeT-A sequences must

Fig. 4. Relative promoter activity of sequences from D. virilis HeT-A ele-

ments tested in D. virilis cells. Diagram shows junction between two ele-

ments showing the 3′UTR of upstream element and the 5′UTR promoters of the other non-LTR retrotransposons that have been studied in Drosophila (11, 12). There is a four-nucleotide match to the Inr, beginning with A at position +1, and a six-nucleotide match to the DPE, running from +28 to +33 (see Fig. 2).

Although the 5′-most 89 nucleotides appear to contain the sequences necessary for initiating transcription, the activity is also influenced by downstream flanking sequences (Fig. 4). Extending the sequence to +429 approximately doubled the activity and adding the entire 5′UTR approximately tripled the activity. Nevertheless the 5′-most sequences were necessary: the fragment +172 → +1,731 was inactive.

3′UTR Activity. Adding the 49 3′-most nucleotides of the 3′UTR to the complete 5′UTR (−49 → +1,731) had little effect, whereas adding 110 nucleotides diminished activity by one third to one half (Fig. 4). The apparent repression of activity by some 3′UTR sequence could indicate that transcription is modulated by 3′UTR promoters of the other non-LTR retrotransposons that have been studied in Drosophila (11, 12). Because the 5′UTR promoters of the other non-LTR retrotransposons do so by locating promoter sequences in the 5′UTR, downstream of the transcription start. The regulatory sequences are included in the RNA transcript and reverse-transcribed into the new site, enabling the new element to continue transposing.

Antisense Promotion. Constructs spanning the entire 3′/5′ region were tested for antisense promoters but produced no activity in our reporter assay. This was unexpected because the same reporter assay has detected an antisense promoter at the 3′ end of D. melanogaster HeT-A (14).

HeT-A′′ Transcription Starts at the 5′ End of the Element. All of the D. virilis constructs (Fig. 4) with significant promoter activity contain the 5′ end of the 5′UTR, which has good matches to the typical non-LTR element Inr and DPE. This strongly suggests that this end of the D. virilis 5′UTR contains the transcription start site. We determined the start site by using 5′RACE to map the 5′ end

of lacZ transcripts from two of our D. virilis constructs, selected to test the effects of the two UTRs on location on the transcription start. One construct (−49 → +1,731) contained the 5′UTR plus some 3′ sequence to test for a possible upstream start. The other (−2,778 → +429) contained the 3′UTR plus the smallest frag-

ment of 5′ sequence necessary to get reasonable promoter activity. 

The 5′RACE analyses showed that both constructs used the same transcription start sites (arrowheads in Fig. 2). Four clones of RNA expressed from the −49 → +1,731 construct were sequenced. One clone began with the A in the Inr and three began with the G immediately 5′ of this A. Both sequenced clones from RNA expressed by the −2,778 → +429 construct began with the G used by the −49 → +1,731 construct. Others studying different retrotransposons have found that A was the preferred start site (12) but G was the predominant start in our experiments. Although the 3′UTR sequences differ from 5′UTR sequences in their effect on the transcriptional activity, they did not move the start of transcription away from the Inr sequence.

Discussion

Inr (initiator) and DPE (downstream promoter element) motifs. The Inr and DPE characterize the 5′UTR promoters of the other
preserve a breeding stock of complete elements to maintain the appropriate rate of transposition onto ends. Therefore, at least some newly transposed elements must not lose 5’ sequence. In the following sections we consider what is known about the ways HeT-A elements could become 5’-truncated and why many HeT-A elements in both species survive without truncation.

Telomere Elements Might Lose 5’ Sequence While Transposing onto the Chromosome End. Non-LTR transposons show a tendency to be 5’-truncated. This truncation has generally been assumed to be caused by premature dissociation of reverse transcriptase from the RNA. There is now reason to question this assumption, which was based largely on studies of retroviral reverse transcriptase. A recent study on the reverse transcriptase of a non-LTR element, B. mori R2, found that this enzyme differed in important ways from the retroviral enzyme. For example, the R2 enzyme was more processive; therefore premature dissociation is less likely to explain 5’-truncation of R2 and probably other non-LTR elements (16).

In addition, studies on B. mori R2 (16) and two other non-LTR elements, mouse and human L1 elements (17), suggest that truncation of these elements is a result of premature initiation of second-strand synthesis. This premature initiation occurs where the RNA template has a region of microhomology with DNA at the other end of the nicked target site. Telomere elements should escape such premature second strand synthesis because they attach to an end and therefore do not have a nicked target site to provide microhomology. Thus it seems unlikely that HeT-A elements are subject to truncation during reverse transcription.

If HeT-A RNA is completely reverse transcribed, the next step in which truncation might occur is synthesis of the second strand of DNA. Nothing is known about how this is accomplished; however, the analogy to telomerase is strong enough to suggest that the HeT-A second strand is synthesized by DNA polymerases δ and α/primase, which synthesize the second strand on telomerase telomeres (18). Loss of terminal sequence could occur if second strand synthesis did not start at the extreme end of the reverse transcribed DNA or if the synthesis were primed by sequence that was removed and not replaced after replication. In Tetrahymena (19), yeast (20), and humans (21, 22), the ends of telomere DNA are known to be shaped by precise postreplicative processing. Much of this processing is accomplished by as yet unidentified proteins, but one protein identified in Saccharomyces cerevisiae (20) and humans (23) is Mre11. Although the processing required for retrotransposition ends may be different, the studies show that telomere ends in other species are accessible for postreplicative shaping. The terminal structure in Drosophila telomeres has not been defined; however, there is genetic evidence that Mre11 interacts with these telomeres (24), suggesting that at least one agent for DNA processing is available at the Drosophila telomere.

Telomere Elements Might Lose 5’ Sequence While They Form the Extreme End of the Chromosome. Conventional DNA synthesis of linear DNA would yield a blunt end on the leading strand telomere and a 3’ single strand overhang where the terminal primer is removed from the lagging strand telomere, causing sequence loss with each round of replication. This model is strongly supported by studies of terminally deleted D. melanogaster chromosomes completely lacking telomeric sequence. These chromosomes continuously shorten by approximately 70 nucleotides per fly generation (25, 26). Taking into account the size of the RNA primer, the number of germline replications between generations, and the fact that underreplication occurs only on the lagging strand, it has been estimated that there is an average loss of two to three nucleotides at each cell division (25). However, the deleted chromosomes of these studies (25, 26) are made and retained in stock only by the “magic” of Drosophila genetics. Their broken ends bind some telomere proteins (27) and perform some, but not all, of the functions of telomeres: they protect ends from checkpoints and terminal fusions. Importantly, they do not protect from end erosion. Unless they are healed by transposition of retrotransposons (28–30), they will eventually shorten enough to lose essential genetic material and be lost. Because they do not prevent shortening, these chromosomes cannot tell us what happens at the ends of chromosomes terminated by telomeric DNA.

It has not been possible to study what happens to Drosophila telomeres when new sequences cannot be added. However, in other organisms, deletion or inactivation of telomerase produces ever-shortening telomeres, but the specific results of loss of telomerase activity differ from species to species. For example, after telomerase deletion, telomeres in Kluyveromyces lactis (31) and Trypanosoma brucei (32) shorten at three to five nucleotides per cell cycle, much like the broken Drosophila chromosomes. Surprisingly, when T. brucei telomeres become critically short, they can stabilize and retain short TTAGGG tracts over long periods despite the loss of telomerase (32). The mechanism for this is not understood. Human chromosomes shorten much more rapidly than predicted from primer nonreplacement, with different cell lines having different loss rates ranging from 50 nucleotides to several hundred nucleotides per division (33). These results suggest that loss can be modulated or even reversed by species-specific mechanisms determined by as yet unidentified components of the telomere complex. The conservation of telomere ends in Drosophila must also be the result of species-specific regulation.

Finally, terminal sequences can be lost by breakage and/or recombination, both between telomere arrays and within an array (see ref. 34). These mechanisms can result in rapid sequence loss and may well be responsible for the severely truncated HeT-A elements seen in telomere arrays.

D. melanogaster HeT-A Promoter Provides Protection for the 5’ End of the Element. The pseudo-LTR mechanism offers a significant advantage for elements that transpose only to chromosome ends because it adds redundant sequence to prevent loss of the true sequence of the terminal element. This buffering effect could have driven, or more likely evolved in concert with, evolution of the D. melanogaster HeT-A promoter. Also, 5’-truncated HeT-A promoter should be capable of promoting an adjacent downstream HeT-A element, thus partially alleviating the evolutionary cost of using a different element for promotion.

Despite the potential for loss from the 5’ terminus, many elements in telomere arrays still retain some of this extra 3′ sequence on their 5′ end. This observation suggests that these tags are sufficient to protect the true ends of a reasonable number of the elements. The variable truncation of sequence within each tag and the fact that no more than six tags have been seen on a single element suggests that nucleotides, as well as complete tags, are lost frequently enough to indicate that they are useful as buffers.

D. virilis HeT-A Promoter Does Not Add Sequence to Protect the 5′ End of the Element. In contrast to the pseudo-LTR mechanism of the D. melanogaster promoter, the typical non-LTR mechanism of the D. virilis promoter provides no obvious way of replacing lost terminal sequence. Nevertheless, much of the HeT-A sequence in D. virilis telomere arrays is in complete elements with apparently intact 5′ ends. Three of the HeT-A junctions in Fig. 2 have both of the nucleotides that were identified as 5′ start sites in our experiments. The other two junctions have an A, the preferred start in other Drosophila non-LTR elements, moved one position to the left. As the G in this position was used most frequently in our experiments, we believe that this A is still appropriately positioned with respect to the DPE and that these two elements also retain their complete 5′ ends. Thus D. virilis HeT-A elements appear to have a mechanism for maintaining an intact 5′ end without resorting to the addition of expendable sequence.
The only positions in the alignment shown in Fig. 2 where nucleotides are not identical in all five elements are the seven positions (lowercase and bolded) immediately following the arrow marking the end of the upstream element. [This end was arbitrarily chosen because it is the end of the shortest oligo(A) in the aligned elements.] As has been reported for the R2 reverse transcriptase (13), the C, T, and G residues in this region are probably unphosphorylated nucleotides added by the enzyme before it engaged the RNA template; whereas the A residues seen in the variable region in Fig. 2 can be explained by initiation at different sites on the poly(A) tail of the RNA template.

As just discussed, we believe both GA and AC are bona fide start sites. Variation in the other five positions probably reflects the mechanism by which non-LTR RNA is reverse transcribed onto the chromosome: because the RNA template does not base pair with the DNA strand extended by reverse transcriptase, the enzyme is somewhat imprecise in starting the reverse transcript.

The evidence that D. virilis can retain intact 5′ ends on a significant fraction of its HeT-A elements suggests that it has a well-coordinated set of regulatory proteins that shape its telomeres. Telomerase-maintained telomeres are known to be associated with a large and growing number of proteins. Drosophila telomerase-associated proteins have been much less studied. Therefore we can only speculate that some of these proteins may have the functions required to precisely replicate the D. virilis telomere end. Mammalian telomeres have a telomere-specific complex, shelterin, composed of six proteins, as well as a number of proteins that are transiently associated with telomeres but have additional nontelomere functions (34). Their roles at the telomere are still being clarified, even for such familiar proteins as Rap1 (35). A particularly relevant study recently identified a protein, CTC1c, in plants and mammals and demonstrated that it had characteristics suggesting a role in telomere replication (36). At least some of the proteins associated with telomeres in other species have roles at telomeres in Drosophila. One of these is Mre11, mentioned earlier. None of the shelterin proteins have been found in Drosophila; however, two telomere-specific proteins, HOAP (37) and Moi (38), are thought to be the founding members of an analogous telomere-specific complex for Drosophila (38).

The striking conservation of the D. virilis HeT-A sequence at both the 3′ end of the upstream element and the 5′ end of the downstream element (thick gray line in Fig. 3) provides a likely explanation for the unexpectedly low level of end erosion seen in D. virilis HeT-A. Sequence conservation at the end of the element differs markedly from that in D. melanogaster. We suggest that this strong sequence conservation in D. virilis may be driven by a requirement for precise associations with multiple proteins and/or RNAs that act to process the extreme end of the telomere, prevent its erosion, and align reverse transcription of the next element. In contrast, D. melanogaster uses nonessential sequence to buffer less precarious processing of the telomere end. Both D. melanogaster and D. virilis have many intact HeT-A elements in the genome, showing that both mechanisms of 5′ end protection are effective.

Methods

Cell Line. The D. virilis cell line WR Dv-1 was obtained from the Drosophila Genomics Resource Center and maintained at room temperature in Schneider Drosophila media (Gibco) supplemented with 10% heat-inactivated FBS (HyClone).

Constructs to Assay for Promoter Activity. Constructs were made in the pCasper-AUG-β-gal vector. HeT-A sequences were inserted into the polylinker and drove expression of the lacZ reporter gene. Constructs are named by the nucleotide at each end of the construct. Nucleotides in the 5′ UTR are numbered consecutively 3′ → 5′ from the junction with the downstream element counting down from −1 [omitting nucleotides in the oligo(A) at the junction, which collectively serve as nucleotide 0]. Nucleotides in the 5′ UTR are numbered 5′ → 3′ from the junction, counting up from +1. All constructs were PCR-amplified from the 3′ UTR of element V7b joined to the 5′ UTR of element V7c (GenBank accession no. AY369260) and verified by sequencing.

Transient Transfection. Promoter strength was measured by the activity of β-gal expressed from each construct. To normalize the transfection, pCMV-Luc (from N. Kamoshita, Tokyo University, Tokyo, Japan), which constitutively expresses luciferase under the control of the CMV promoter, was cotransfected with each experimental promoter construct. Cells were seeded at 5 × 10^6 cells/mL in six-well plates (2 mL/well) and grown overnight at room temperature, then transfected with 20 μg of promoter construct DNA and 0.2 μg of control luciferase plasmid DNA, using a calcium phosphate–DNA coprecipitation method (39). Calcium phosphate precipitate was removed 16 to 20 h after addition of DNA and the cells were treated with 15% glycerol in cell culture medium for 1.5 min. The glycerol shock solution was removed, cells were washed two times with media, and grown overnight before the assay.

Expression Assays. Cells were harvested into 1× reporter lysis buffer (Promega) 48 h after addition of DNA. Luciferase activity was measured by the Luciferase Assay System (Promega) and β-gal activity was measured by the Beta-Glo Assay System (Promega). β-Gal activity was normalized to luciferase activity. Data for analysis came from at least three independent experiments. Data from each experiment was normalized by the activity of one construct used for this purpose in every experiment.

5′ RACE Determination of Transcription Start Sites. 5′ RACE analysis of RNA from transfected cells was with the FirstChoice RLM-RACE kit (Ambion) as described by Maxwell et al. (5). Two rounds of nested PCR were carried out with forward primers for the 5′ UTR adaptase and reverse primers for the lacZ element of the vector. The lacZ primers were as follows: D60263, GCTTTCAGC CACTTGTGCTGGTCCAGCAGGAAAAGAC; and 753R, 5′-GTTGGCCACGTCTGAGGAAAGGGG-3′. PCR products were cloned into the Stratagene PCR cloning vector pSC-A (Stratagene) and sequenced.

Acknowledgments. We thank E. Casacuberta and N. Kamoshita for clones; T. Rajbhandary for use of the luminometer; C. Koehrer for instruction; and K. Lowenhaupt, E. Casacuberta, and members of the Pardue laboratory for helpful discussions and critical reading of the manuscript. This work was supported by National Institutes of Health Grant GM50315 to M.-L.P.


