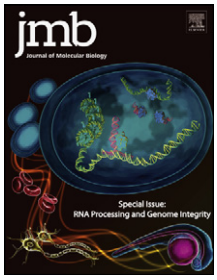

Telomeric Retrotransposon *HeT-A* Contains a Bidirectional Promoter that Initiates Divergent Transcription of piRNA Precursors in *Drosophila* Germline

Elizaveta Radion†, Sergei Ryazansky†, Natalia Akulenko, Yakov Rozovsky, Dmitry Kwon, Valeriya Morgunova, Ivan Olovnikov and Alla Kalmykova

Institute of Molecular Genetics, *Russian Academy of Sciences, Moscow 123182, Russia*

Correspondence to Alla Kalmykova: *Institute of Molecular Genetics RAS, Kurchatov square 2, Moscow 123182, Russia.*
allakalm@img.ras.ru
http://dx.doi.org/10.1016/j.jmb.2016.12.002
Edited by Fonseca Maria Carmo

Abstract

PIWI-interacting RNAs (piRNAs) provide the silencing of transposable elements in the germline. *Drosophila* telomeres are maintained by transpositions of specialized telomeric retroelements. piRNAs generated from sense and antisense transcripts of telomeric elements provide telomere length control in the germline. Previously, we have found that antisense transcription of the major telomeric retroelement *HeT-A* is initiated upstream of the *HeT-A* sense transcription start site. Here, we performed a deletion analysis of the *HeT-A* promoter and show that common regulatory elements are shared by sense and antisense promoters of *HeT-A*. Therefore, the *HeT-A* promoter is a bidirectional promoter capable of processive sense and antisense transcription. Ovarian small RNA data show that a solo *HeT-A* promoter within an euchromatic transgene initiates the divergent transcription of transgenic reporter genes and subsequent processing of these transcripts into piRNAs. These events lead to the formation of a divergent unistrand piRNA cluster at solo *HeT-A* promoters, in contrast to endogenous telomeres that represent strong dual-strand piRNA clusters. Solo *HeT-A* promoters are not immunoprecipitated with heterochromatin protein 1 (HP1) homolog Rhino, a marker of the dual-strand piRNA clusters, but are associated with HP1 itself, which provides piRNA-mediated transcriptional repression of the reporter genes. Unlike endogenous dual-strand piRNA clusters, the solo *HeT-A* promoter does not produce overlapping transcripts. In a telomeric context, however, bidirectional promoters of tandem *HeT-A* repeats provide a read-through transcription of both genomic strands, followed by Rhi binding. These data indicate that *Drosophila* telomeres share properties of unistrand and dual-strand piRNA clusters.

Introduction

PIWI-interacting RNAs (piRNAs) mediate biochemical pathways responsible for the silencing of transposable elements (TEs) in the germline [1]. piRNAs are generated from long single-stranded precursor RNAs that are transcribed from distinct genomic loci called piRNA clusters [2]. Dual-strand piRNA clusters produce piRNAs from both genomic strands; in unistrand clusters, piRNAs are mapped to one genomic strand.

Distinct types of piRNA clusters are characterized by transcriptional peculiarities such as convergent or divergent bidirectional transcription. In both cases, transcription is initiated at opposite genomic strands. Transcription is termed divergent when RNA polymerases produce non-overlapping transcripts, whereas in the case of convergent transcription, RNA polymerases move toward each other along opposite genomic strands producing complementary RNAs. Dual-strand clusters enriched in damaged TE fragments are typical for the *Drosophila* germline. Promoters driving the convergent transcription of dual-strand clusters have yet to be identified.

Unistrand piRNA clusters are the main source of piRNAs in *Drosophila* follicular cells [3,4] and at the pachytene stage during mammalian spermatogenesis [5]. In contrast to *Drosophila* piRNA clusters, many mammalian clusters are transcribed divergently from bidirectional promoters. A transcription

factor of the Myb family binds to promoters of both divergently and unidirectionally transcribed pachytene piRNA clusters in mouse testes [6].

Notably, divergent transcription is a general feature of not only piRNA clusters but also many mammalian genic promoters [7–10] and TEs. The regulatory 5′ region of the human non-long terminal repeat (LTR) retrotransposon LINE-1 contains sense and antisense transcription start sites (TSSs) that give rise to overlapping transcripts [11–15]. At most genic promoters, however, productive elongation occurs unidirectionally, while initiation in the antisense direction produces short, unstable transcripts [9]. The majority of *Drosophila* coding gene promoters are unidirectional [16]. Bidirectional transcription has been observed only for a few *Drosophila* TEs, such as the non-LTR retrotransposon *F*-element and the telomeric retroelements [17–19]. Antisense transcription of the *F*-element was initiated near the sense TSS, and it was therefore proposed that discrete sense and antisense promoters of the *F*-element are located close to each other [18].

Drosophila melanogaster telomeres consist of tandem head-to-tail arrays of the specialized telomeric retrotransposons *HeT-A*, *TART*, and *TAHRE* represented both by complete and 5′-truncated copies [20,21]. Analysis of ovarian small RNA deep sequencing data revealed abundant telomere-specific piRNAs corresponding to both sense and antisense strands of telomeric repeats [2], suggesting that both genomic strands of telomeres are transcribed. It has previously been shown that transcription of the antisense strand of telomeric retroelements is initiated within their regulatory regions [17,19]. The *TART* element was shown to be transcribed bidirectionally from the region that is located within its non-terminal direct repeats [17,22]. Promoters of the telomeric retroelements *HeT-A* and *TAHRE* are located in the 3′ untranslated region (3′ UTR) and drive the transcription of a downstream element [23,24]. We have previously demonstrated that *HeT-A* antisense transcription in ovaries is also initiated within the 3′ UTR, approximately 150 nt upstream of the sense TSS [19]. Here, we present a study of *HeT-A* promoter by using dual reporter constructs and transfection of *Drosophila* cultured cells. We show that both sense and antisense transcription of *HeT-A* are controlled by common regulatory elements. Our results suggest that the *HeT-A* 3′ regulatory region can be considered as a single bidirectional promoter. The *HeT-A* promoter, being inserted in euchromatic sites within a transgenic construct, initiates the divergent transcription of transgenic reporter genes in ovaries. Subsequent processing of their transcripts into piRNAs results in the formation of divergent unistrand piRNA clusters reminiscent of the pachytene piRNA clusters in mouse testes [5,6]. *HeT-A* promoter is a unique example of *Drosophila* regulatory element that can drive divergent transcription. In the context of telomeres, divergent *HeT-A* promoters described here can provide the transcription of both genomic strands of tandemly repeated telomeric transposons. In the germline, such transcripts are processed into piRNAs that regulate telomeric element expression and their transposition rate onto chromosome ends [25]. We suggest that *HeT-A* bidirectional promoter activity may have evolved to provide a piRNA-mediated control of telomere length.

Results

Bidirectional promoter activity of *HeT-A* 3′ region: constructs and cell culture transfection assay

A dual reporter system was created to simultaneously monitor sense and antisense promoter activity of the *HeT-A* 3′ region. *lucF* and *lucR* reporter genes encoding firefly and *Renilla* luciferase, respectively, were inserted in a head-to-head orientation in the *2luc* vector designed for this experiment (see Supplementary Materials and Methods). The *HeT-A* promoter region and its deletion variants were derived from the *HeT-A/z2* clone [24], which contains a fragment of an active, recently transposed *HeT-A* element. The *HeT-A* promoter sequence was incorporated into a polylinker and inserted between *lucF* and *lucR*. Deletion constructs were designated as H(#), where (#) corresponds to the distance (in bp) from the *HeT-A* polyadenylation site. Schneider 2 cells were co-transfected with the *HeT-A* constructs and vector pCaSpeR-AUG-β-gal expressing *lacZ* under the *ovo* gene promoter (*ovo/lacZ* construct; Supplementary Materials and Methods). Reverse transcription followed by quantitative PCR (RT-qPCR) was considered a more reliable method of measuring promoter activity than dual luciferase assay, because cryptic AUG codons could affect reporter protein expression levels. Expression levels of *lucF* and *lucR* reporter genes measured by RT-qPCR were normalized to *lacZ* transcript abundance. Results of the deletion analysis are shown in Fig. 1. Promoter activities of 675 bp and 434 bp of the 3′-terminal *HeT-A* sequences were approximately equal in both directions. Further deletions resulted in decreased sense and antisense expression levels; the most prominent decrease was observed for the H347 construct. In this construct, the *HeT-A* fragment from –434 to –347 bp, relative to the polyadenylation site, was deleted. Another group has previously shown that the deletion of the *HeT-A* 3′ region (–367 to –208 bp) considerably decreased reporter gene expression [23]. Here, we have constructed other deletion variants in order to test both sense and antisense transcription. Notably, we also observed the importance of the *HeT-A* promoter region located ~350–430 bp upstream of the polyadenylation site for the *HeT-A* promoter activity.

To verify potential enhancer/activator activity that the –434 to –347-bp fragment (relative to the