Published online 24 April 2013 Nucleic Acids Research, 2013, Vol. 41, No. 11 5757–5768
doi:10.1093/nar/gkt310

De novo piRNA cluster formation in the *Drosophila* germ line triggered by transgenes containing a transcribed transposon fragment

Ivan Olovnikov[1,2], Sergei Ryazansky[1], Sergey Shpiz[1], Sergey Lavrov[1], Yuri Abramov[1], Chantal Vaury[3,4,5], Silke Jensen[3,4,5,*] and Alla Kalmykova[1,*]

[1]Institute of Molecular Genetics, Russian Academy of Sciences, Moscow 123182, Russia, [2]Division of Biology, California Institute of Technology, Pasadena, CA 91125, USA, [3]Clermont Université, Université d'Auvergne, Laboratoire du GReD, BP 38, 63001 Clermont-Ferrand, France, [4]INSERM, U 1103, BP 38, 63001 Clermont-Ferrand, France and [5]CNRS, UMR 6293, BP 38, 63001 Clermont-Ferrand, France

Received December 23, 2012; Revised March 6, 2013; Accepted April 3, 2013

ABSTRACT

PIWI-interacting RNAs (piRNAs) provide defence against transposable element (TE) expansion in the germ line of metazoans. piRNAs are processed from the transcripts encoded by specialized heterochromatic clusters enriched in damaged copies of transposons. How these regions are recognized as a source of piRNAs is still elusive. The aim of this study is to determine how transgenes that contain a fragment of the Long Interspersed Nuclear Elements (LINE)-like *I* transposon lead to an acquired TE resistance in *Drosophila*. We show that such transgenes, being inserted in unique euchromatic regions that normally do not produce small RNAs, become *de novo* bidirectional piRNA clusters that silence *I*-element activity in the germ line. Strikingly, small RNAs of both polarities are generated from the entire transgene and flanking genomic sequences—not only from the transposon fragment. Chromatin immunoprecipitation analysis shows that in ovaries, the trimethylated histone 3 lysine 9 (H3K9me3) mark associates with transgenes producing piRNAs. We show that transgene-derived *hsp70* piRNAs stimulate *in trans* cleavage of cognate endogenous transcripts with subsequent processing of the non-homologous parts of these transcripts into piRNAs.

INTRODUCTION

Silencing of transposable elements (TEs) in the germ cells depends on a distinct class of 24–30-nt-long RNAs, PIWI-interacting RNAs (piRNAs), associated with PIWI clade Argonaute proteins (1,2). TE repression is provided by the two piRNA classes, the primary and the secondary piRNAs. In fly, most primary piRNAs match defective transposons and derive from the discrete pericentromeric and telomeric heterochromatic loci enriched in damaged repeated sequences, which are called piRNA clusters or piRNA master loci. Primary piRNAs are processed from the putative long single-stranded transcripts encoded by these loci and demonstrate a strong 5′ terminal uridine bias (1U bias). Primary processing was recently shown to be dependent on the activities of an endoribonuclease Zucchini (3–5) and 3′ trimmer of yet unknown protein identity (6). A subsequent step of piRNA biogenesis, the 'ping-pong' mechanism, is a process resulting in amplification of antisense piRNAs (1,7). Secondary piRNAs are generated through piRNA-guided cleavage of transposon mRNA so that primary antisense piRNA and newly produced sense piRNA have 10-nt overlap. A sense piRNA that arises from this cleavage has adenine in position 10. In a similar process, sense piRNA can give rise to an antisense piRNA, by interacting with a piRNA cluster transcript. According to this model, active expression of TEs provokes amplification of antisense piRNAs reinforcing suppression of TE activity in the germ line. Indeed, the total population of piRNAs is strongly biased towards antisense piRNA in respect to most classes of TE found in *Drosophila*. piRNA clusters also produce significant levels of endo-siRNAs, which silence transposons both in the soma and in the germ line (8–10).

piRNA clusters are represented mainly by pericentromeric and subtelomeric regions enriched in remnants of TEs (1). Fragments of piRNA clusters inserted as transgenes into heterologous genomic positions in fly and mouse were able to produce piRNAs, indicating that

*To whom correspondence should be addressed. Tel: +7 499 1960019; Fax: +7 499 1960221; Email: allakalm@img.ras.ru
Correspondence may also be addressed to Silke Jensen. Tel: +33 4 7317 8182; Fax: +33 4 7327 6132; Email: silke.jensen@udamail.fr

The authors wish it to be known that, in their opinion, the first three authors should be regarded as joint First Authors.

specific location of endogenous piRNA clusters is not a prerequisite for piRNA processing (11). Artificial sequences introduced in piRNA clusters generate abundant piRNAs, suggesting that any sequence, if inserted into a piRNA cluster, will be a source of new piRNAs (11). Thus, transposon or transgene integrations into piRNA clusters cause repression of cognate sequences expressed in the germ line (11–13). However, the prerequisites to form a piRNA cluster have not yet been found.

piRNA cluster transcription in the germ line is thought to be regulated through the assembly of a distinct chromatin structure mediated by the methyltransferase dSETDB1, Cutoff protein and HP1 homologue Rhino (14–16). A special chromatin structure along the piRNA cluster is proposed to allow RNA polymerase to ignore splicing, termination and other signals providing generation of long piRNA precursor transcripts (11). The mechanism of piRNA cluster transcription still remains unclear.

Here, we address the questions concerning the structure and transcription of piRNA clusters by using *Drosophila* strains containing transgenes that silence *I*-element activity. The *I*-element is a non-long terminal repeat (LTR) retrotransposon involved in the phenomenon of I-R hybrid dysgenesis in *Drosophila melanogaster*. A cross of inducer ('I') males carrying active *I*-elements to reactive ('R') females devoid of functional *I*-elements yields dysgenic daughters (SF) with a sterility syndrome and elevated mutation rates owing to mobilization of *I*-element. These traits are not seen in the female progeny of a reciprocal cross. Despite the absence of functional *I*-elements in R strains, their genomes contain remnants of ancestral *I*-related elements in the pericentromeric heterochromatin, including the 42AB locus, which was described previously as one of the major piRNA loci (1). Natural populations of *D. melanogaster* have been recently reinvaded by the modern active *I*-element, resulting in the appearance of I strains containing damaged ancestral *I*-copies, as well as functional copies scattered along the chromosome arms. *I*-specific piRNAs from the ovaries of an I-strain exhibit the ping-pong signature (adenosine in position 10 of sense piRNAs and 5′-terminal uridine in antisense piRNAs). Notably, ~78% of the sense *I*-specific piRNAs are derived from modern copies, whereas 63% of antisense *I* piRNAs originate from ancestral heterochromatic *I*-copies (17). These data confirm previous observations that the productive ping-pong cycle takes place between the transcripts of active transposons and heterochromatic piRNA loci (1). A recent study has unveiled the role of endogenous piRNAs in the phenomenon of the I-R hybrid dysgenesis manifested in SF daughters (17). It was shown that the maternal inheritance of transposon silencing might be achieved through a direct transmission of maternal piRNAs to embryonic germ cells. Most of the uniquely mapping *I*-element piRNAs originate from the piRNA cluster 42AB. The reactive strain produces much less ovarian *I*-element piRNA, coming from ancestral *I*-related elements only, than the inducer strain containing both ancestral and modern *I*-elements (17). SF daughters inherit a low amount of *I*-specific piRNA from their R mothers. As a result, mobilization of *I*-elements coming from paternal chromosomes leads to the syndrome of hybrid dysgenesis. In terms of SF female sterility, the permissivity to *I*-element activity of R females, called reactivity, can be measured by a percentage of hatching eggs laid by their SF daughters. This score correlates with the *I*-element transposition frequency, providing a non-direct but convenient method for the estimation of the *I*-element's activity (18). According to this test, strong and weak R strains can be identified. Reactivity may be defined as a quantitative maternally inherited trait. Additionally, the reactivity level decreases on ageing or heat treatment. Under these experimental conditions, an increased expression of the *I*-related damaged copies correlated with decreased reactivity level in R females, indicating a significant role of these defective copies in the epigenetic mechanism of *I*-element suppression (19).

It was previously shown that the activity of functional *I*-elements introduced into R lines can be repressed by transgenes containing a fragment of the *I*-element in sense or antisense orientation (20–22). Here, we show that these transgenes produce *I*-specific small RNAs, which reduce reactivity of the transgenic lines. Moreover, *I*-transgenes become *de novo* piRNA clusters producing small RNAs throughout the construct and extending into flanking genomic sequences. R-lines containing such transgenic constructs represent a unique model to study the principles of piRNA cluster formation.

MATERIALS AND METHODS

Drosophila strains

The transgenic strains that carry insertions of the *I*-element fragment in sense (1.9, 2.1, 2.3, 2.6 and 2.10) or antisense (3.1, 3.6, 3.9 and 3.10) orientation are described in Jensen *et al*. (21). Briefly, the 167–2484-nt region of the GenBank sequence M14954, corresponding to the *I*-element, was inserted into the pW8-hsp-pA vector in sense (hsp[i1-2Δ/S]pA) or antisense (hsp[i1-2Δ/AS]pA) orientation and introduced into the genome of the *Drosophila* reactive w^K strain. The control strain 62.5.2 (T5) contains insertion of pW8-hsp-pA vector; strain 67.2.1 (7.1) carries a promoterless construct pA'[i1-2Δ]pA in which the *hsp70* polyadenylation sequence was inserted instead of the *hsp70* promoter upstream of the *I*-element fragment. These constructs were introduced into a *Drosophila* strain devoid of functional *I*-elements (the reactive w^K strain) (21). Reactivity was evaluated by measuring the percentage of dead embryos laid by the progeny resulting from the cross of transgenic females with w^{1118} males containing functional *I*-elements, according to a previously described procedure (22). Genomic sites of insertions were determined by inverse polymerase chain reaction (PCR) (Berkeley *Drosophila* Genome Project methods) and are shown in the Supplementary Table S1.

Small RNA library preparation and analysis

Small RNAs 19–29 nt in size from total ovarian RNA extracts were cloned as previously described in Muerdter *et al*. (11). Libraries were barcoded according to Illumina TrueSeq Small RNA sample prep kit instructions and submitted for sequencing using the Illumina HiSeq-2000 sequencing system. Bioinformatic analysis of small RNA libraries is described in Supplementary Materials and Methods. Published small RNA deep sequencing data from previously published data (23,24) were also analysed. Small RNA sequencing data are deposited at Gene Expression Omnibus (GEO), accession number GSE41780.

Northern analysis of short RNAs

Northern analysis of short RNAs was done as previously described (25). A chemical cross-linking step that enhances detection of small RNAs was used (26). For this, the damp membrane with RNA side facing up was placed onto 3 MM saturated in freshly prepared cross-linking EDC reagent [0.16 M 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) and 0.13 M 1-methylimidazole, pH 8) and incubated at 60°C for 2 h. After cross-linking, membrane was rinsed in excess RNase-free distilled water to remove any residual cross-linking solution. Enrichment for short RNA species was done using the miRNeasy Mini Kit (Qiagen). P^{32}-labelled riboprobes corresponding to the sense strand of the *I*-element were synthesized. The *I*-element probe contained a fragment of open reading frame (ORF) 2 corresponding to nucleotides 2109–2481 of the GenBank sequence M14954. Hybridization with P^{32} 5′-end-labelled oligonucleotide 5′-ACTCGTCAAAATGGCTGTGATA-3′ complementary to *miRNA-13b-1* was used as a loading control. The blots were visualized with a phosphor imager Storm-840 (Amersham).

Chromatin immunoprecipitation

About 200 pairs of ovaries were dissected for every Chromatin immunoprecipitation (ChIP) experiment. ChIP was performed according to the published procedure (27). Chromatin was immunoprecipitated with the following antibodies: anti-HP1a (Covance), anti-trimethyl-histone H3 Lys9 (Millipore), anti-H3K27me3 (Upstate) and anti-H3K4me2 (Upstate). Quantitative PCR was conducted on DT-96 machine from DNA Technology, Russia. Eight serial 3-fold dilutions of input DNA of corresponding strain were amplified in triplicates with each primer pair to build standard curves. Standard deviation of triplicate PCR measurements was calculated. *rp49* and histone H3 genes were used for normalization.

RESULTS

I-element-specific small RNAs are produced from transgenes carrying *I* fragments

It was previously shown that transgenes containing a fragment of the *I*-element repressed the *I*-element activity (20–22). In these constructs, a 2.3-kb *I*-element fragment (hereinafter referred to as I-TG) had been cloned into the pW8 transgenesis vector between the *hsp70* promoter and a sequence containing the *actin5C* polyadenylation signal (21). Constructs with the *I*-fragment in sense or antisense orientation were introduced into the genome of the *Drosophila* reactive w^K strain. We confirmed that at present, all transgenic strains used in this study are characterized by very low-reactivity levels (data not shown). To address the mechanism of the repressive effect of an *I*-containing transgene, we sequenced small RNAs from ovaries of w^K and transgenic strains (Supplementary Figure S1 and Supplementary Table S2). We analysed five strains with an *I*-sense construct (1.9, 2.1, 2.3, 2.6 and 2.10), four with an *I*-antisense (3.1, 3.6, 3.9 and 3.10), one strain with a construct containing the I-TG but no promoter, one control strain with a construct missing the I-TG region (strain 62.5.2) and the reactive w^K strain (21) (Supplementary Figure S2). For all of the transgenic strains, insertion sites were determined using inverse-PCR (Supplementary Table S1). In strain 3.1, the transgene was inserted into 3R telomere-associated sequences (TAS), which is a potent piRNA cluster; in the other strains, the insertions were located in euchromatic regions not adjacent to piRNA clusters. Insertion of TE *Tirant* in gene CG32486 present in the genome of the sequenced strain (insertion site indicated in Figure 4B) was not detected in the w^K and transgenic strains.

First, we focused on the *I*-fragment present in the transgene. Mapping of small RNAs to the canonical *I*-element revealed a 5- to 50-fold increase relative to w^K in *I*-specific small RNAs of both polarities corresponding to the I-TG portion of *I*-element for both *I*-sense and *I*-antisense transgenic strains (Figure 1A, Supplementary Figure S3 and Supplementary Table S3). These data show that transgenic *I* fragment transcripts generate additional small RNAs, which correlate with a decrease in reactivity (21). Interestingly, the pattern of piRNA distribution along the I-TG was almost identical irrespective of its orientation in the transgene (Figure 1A). The *I*-promoterless transgenic strain (67.2.1) and w^K generate comparable amounts of small RNAs complementary to the I-TG fragment, indicating that the *I*-promoterless construct does not produce *I*-specific small RNAs. Most of the I-TG small RNAs are 24–29 nt in size and show the characteristic nucleotide bias of piRNA species (1U/10A) (Figure 1B). We found sense/antisense (relative to canonical *I*-element) pairs corresponding to the I-TG region overlapped by 10 nt, which is a signature of the ping-pong amplification cycle (Supplementary Figure S4). Importantly, no ping-pong signal was detected in strain w^K before introduction of *I*-transgenes. Northern analysis of small RNAs confirmed the presence of *I*-element-specific piRNAs in ovaries of transgenic flies (Figure 1C).

Mapping of small RNAs from transgenic strains to the *I*-element with one to three mismatches reveals small RNAs that arise from the ancestral *I*-related elements residing in piRNA loci (17). Strains 2.1, 3.1 and 3.6 show moderate increase in such small RNA abundance across I-TG (Supplementary Table S3). We analysed the amount of *I*-element reads coming from the 42AB master

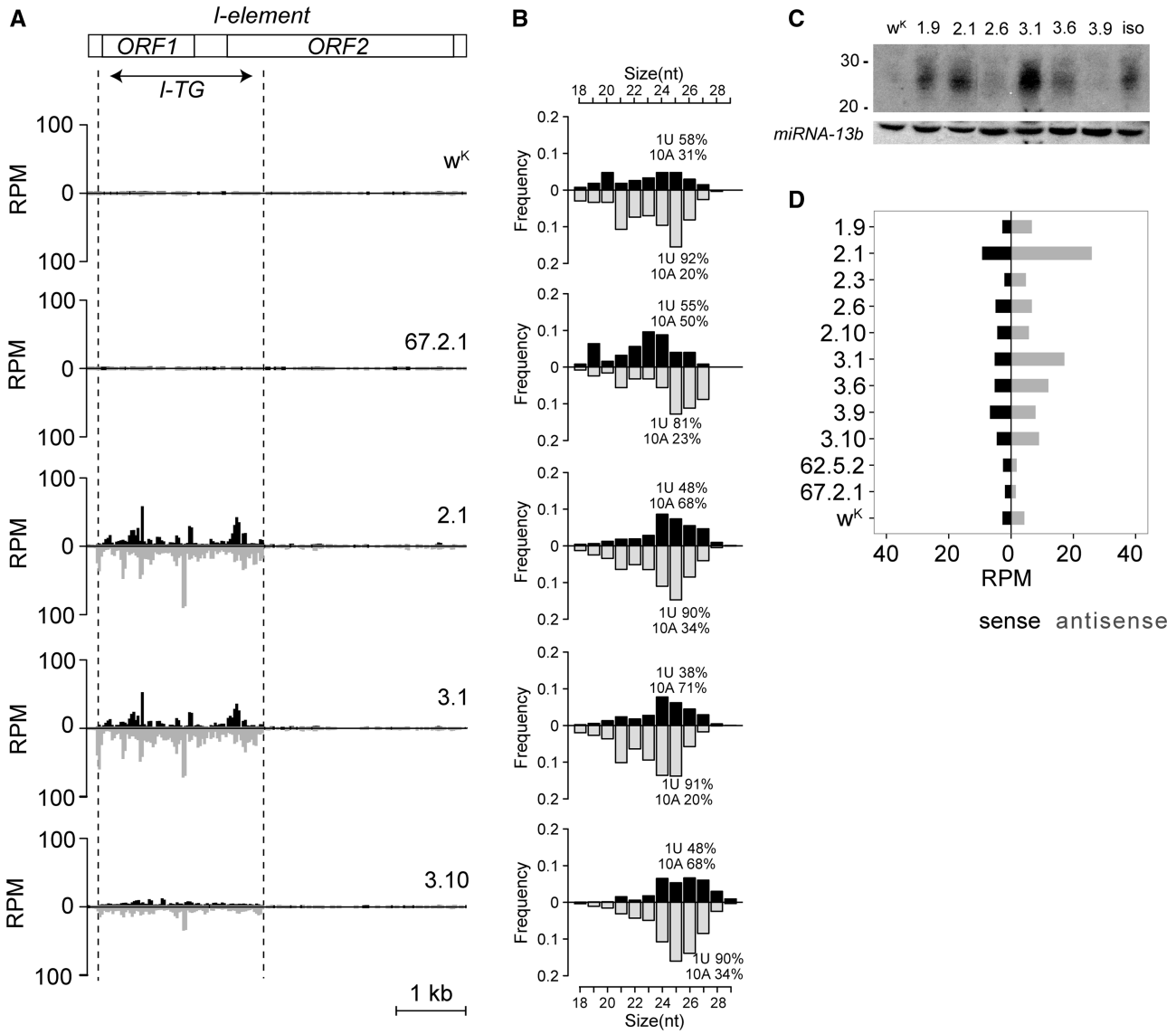

Figure 1. *I*-element-specific small RNAs in transgenic strains. (**A**) Normalized small RNA density in a 30-bp window along a canonical *I*-element sequence (shown above) in the transgenic fly ovaries (reads per million, rpm; no mismatches allowed). Reads mapped to the sense strand are shown in black, antisense in grey. *I*-element fragment present in the transgenic constructs is marked by a dashed line (I-TG). Note that I-TG is inverted relative to *hsp70* promoter in strains 3.1 and 3.10. (**B**) Length distribution of small RNA mapped to the *I*-element in the transgenic strains (the order of strains is the same as on Figure 1A). Percentages of reads having 1U and 10A biases are indicated for each strand (only 24–29-nt reads were considered). (**C**) Northern analysis of ovarian small RNAs in y^1; cn^1 bw^1 sp^1 (iso), w^K and transgenic strains using a sense *I*-element probe from the I-TG region to detect antisense RNAs. Lower panel: hybridization with oligonucleotide complementary to the *miRNA-13b-1* microRNA, loading control. Positions of RNA size markers are indicated. (**D**) Normalized numbers of small RNAs uniquely mapped to *I*-related element homologous to I-TG (chr2R: 2 148 773–2 149 491, 42AB locus) in different transgenic strains.

locus that contains an *I*-related element (chr2R: 2 148 773–2 149 491) homologous to I-TG. We observed a marked enrichment of piRNAs uniquely mapped to this fragment in transgenic strains (Figure 1D). Using reverse transcription (RT)–PCR, we have shown that the expression level of *I*-element from 42AB in the ovaries of transgenic strains is not higher than in w^K (Supplementary Figure S5). Thus, transgene insertions and an elevated level of *I*-specific piRNAs do not affect piRNA cluster expression. These data imply the involvement of transcripts encoded by master loci in the amplification of *I*-specific small RNAs in transgenic flies.

I-containing transgenes form *de novo* piRNA producing clusters

Mapping of small RNAs from transgenic strains revealed that small RNAs of both polarities are generated from the entire transgene, including I-TG, hs-mini-*white* gene (mini-*white* under *hsp70* promoter), P-element fragments, the *actin5C* poly(A) signal-containing sequence and *hsp70* promoters (Figure 2A, Supplementary Figure S6 and Supplementary Table S4). We analysed separately each part of the transgene for piRNA production. *hsp70* promoter and ancestral heterochromatic *I*-related

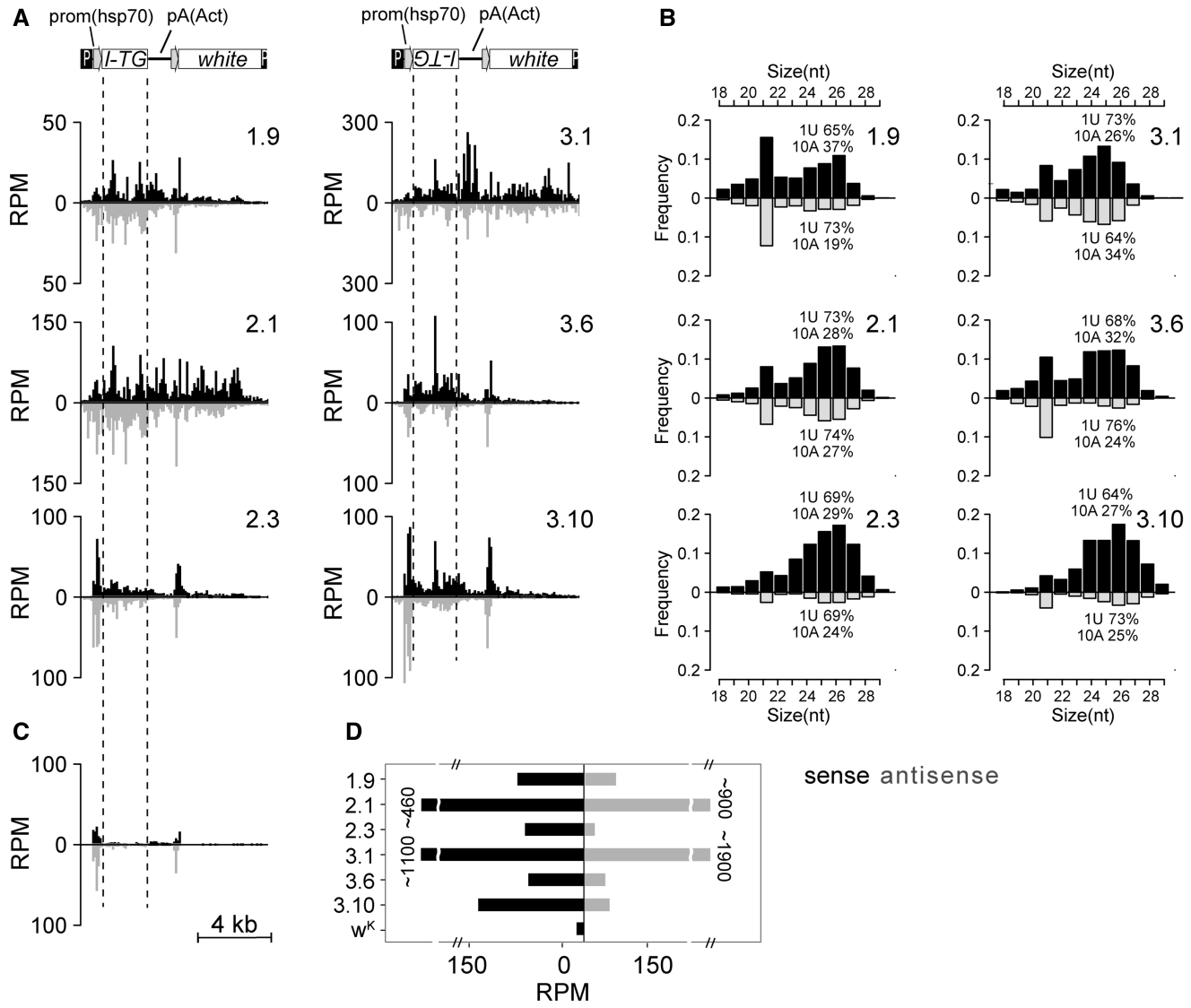

Figure 2. Generation of small RNAs by transgenes containing a fragment of the *I*-element. (**A**) Normalized numbers of small RNAs in a 100-bp window mapped to transgenic constructs (black: sense; grey: antisense; no mismatches allowed). *I*-sense and *I*-antisense transgenes are shown above the plots. (**B**) Length distribution of small RNAs mapping to all transgene sequences except for I-TG and *hsp70* promoters. Percentages of reads having 1U and 10A biases are indicated for each strand (only 24–29-nt reads were considered). Primers used in the ChIP analysis are indicated by arrows. (**C**) Normalized numbers of small RNAs corresponding to *I*-sense transgenic construct revealed in R strain w^K. (**D**) Normalized numbers of sense and antisense small RNAs mapping to all transgene sequences except for I-TG and *hsp70* promoters in different transgenic strains and w^K.

fragments produce piRNAs in w^K strain. *Actin5C*-specific piRNAs were also revealed in w^K (Figure 2C and Supplementary Table S4). We could not detect noticeable amount of *white* piRNAs. P-element and pW8 vector linkers are absent in the w^K genome. However, as parts of transgenic constructs containing transcribed *I*-fragment, all of these sequences produce sense and antisense piRNAs. It was shown previously that a transgene inserted in the TAS region behaves as part of this subtelomeric piRNA cluster and produces abundant piRNAs from transgenic sequences (11). In this study, however, only the 3.1 transgene inserted in a piRNA cluster, i.e. in the 3R TAS region. The transgenes of the other strains are inserted in unique euchromatic regions. Nevertheless, most of them behave as *de novo* double-stranded piRNA clusters (Figure 2 and Supplementary Table S4).

We analysed size and 1U/10A bias of small RNAs mapping to P-element, *actin5C* and hs-mini-*white* portions in transgenic flies (Figure 2B and Supplementary Table S5). Most of these small RNAs are 24–29 nt in size, whereas the 21-nt small RNA fraction is also present. piRNAs of both polarities show a strong 1U bias—a signature of primary piRNAs. The characteristic feature of secondary piRNA (10A-bias) is not observed for the sense piRNA population, suggesting that transgenic P-element, *actin5C* and *white* piRNAs are not involved in the ping-pong amplification loop.

Endogenous *hsp70* gene loci represent piRNA clusters producing noticeable amounts of piRNAs (Figure 3 and Supplementary Table S4). Given the fact that *hsp70* piRNAs are bound by Aub, PIWI and Ago3 proteins and amounts of *hsp70* piRNAs are considerably reduced

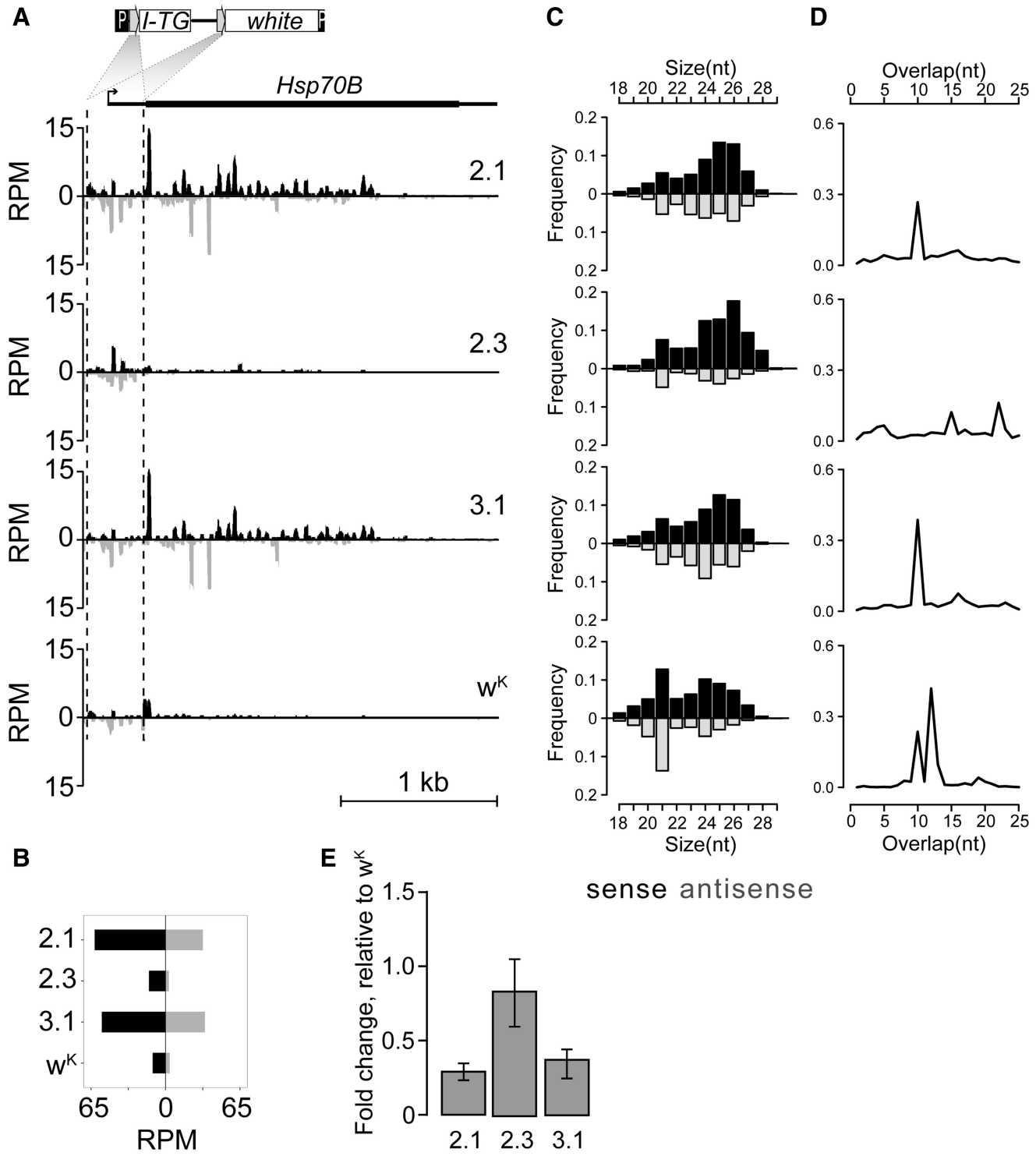

Figure 3. *Trans* effects caused by transgene-associated *hsp70* piRNAs. (**A**) Diagrams of transgene and endogenous *hsp70B* gene are shown above. Positions of transgenic *hsp70* fragments are indicated by dotted lines. Plots represent normalized abundance of small RNAs in a 30-bp window (in reads per million, rpm; black: sense; grey: antisense; no mismatches allowed) along the endogenous *hsp70B* gene in w^K and transgenic strains.(**B**) Normalized numbers of small RNAs mapped to *hsp70B*, excluding the promoter region in different transgenic strains. (**C**) Length distribution of small RNA mapped to *hsp70B*, excluding the promoter region that is present in the transgene (the order of strains is the same as on Figure 1A). (**D**) The relative frequencies (*Z*-score) of 5′-overlap for sense and antisense 24–29-nt *hsp70* piRNAs, excluding promoter region. (**E**) RT–qPCR analysis of the *hsp70B* transcript amount in the ovaries of transgenic strains. Bars of histograms represent normalized ratio of *hsp70B* transcript abundance in the ovaries of transgenic strains to that in w^K.

in mutants of the germ line piRNA pathway genes, we conclude that these small RNAs have a germinal origin [data from previous publications (23,24)] (Supplementary Figure S7). Transgenic small RNA libraries are moderately enriched in small RNAs corresponding to the *hsp70* promoter (Supplementary Table S4). However, it is impossible to distinguish endogenous piRNAs from transgenic ones.

Endogenous small RNAs, complimentary to *hsp70* and/or *I*-related elements, may serve as triggers for piRNA generation by the transgenic sequences. To differentiate between these alternatives, small RNA libraries made from the ovaries of control and *I*-promoterless strains have been analysed. In the control strain without *I*-fragment (62.5.2), piRNAs corresponding to the *actin5C* fragment and to the hs-mini-*white* gene are generated. In the *I*-promoterless strain (67.2.1), *white*-specific piRNAs are also produced (Supplementary Figure S6C and Supplementary Table S4). In both cases, small RNA production is observed downstream of *hsp70* promoters and only collinear with transcription, whereas in *I*-sense (1.9, 2.1, 2.3) and *I*-antisense (3.1, 3.6, 3.10) transgenic strains, small RNAs of both polarities homologous to *actin5C* and *white* are present (Supplementary Table S4). One may suggest that endogenous *hsp70*-specific piRNAs can induce generation of small RNAs from *hsp70*-driven transcription units. Ten independent *I*-promoterless strains analysed previously did not silence *I* activity (21), indicating that piRNA production does not spread upstream of the *hsp70* fragment into the *I*-fragment. Importantly, most of randomly chosen *I*-sense or *I*-antisense transgenic strains were low reactive and produced abundant transgene-specific small RNAs of both polarities (Supplementary Tables S3 and S4). These data indicate that the presence of a transcribed *I*-element fragment in the transgenic constructs is essential for the *de novo* dual-strand piRNA cluster formation.

Transgene insertions promote generation of small RNAs from genomic sequences flanking the transgenes

In all but strain 3.1, transgenes were inserted in unique euchromatic regions. In w^K, no noticeable amounts of small RNAs derived from these genomic regions were detected. Mapping of small RNAs from *I*-sense (1.9, 2.1) and *I*-antisense (3.6) strains to the reference genome showed the appearance of uniquely mapped sense and antisense small RNAs generated from the regions flanking the transgene (Figure 4). These small RNAs are represented predominantly by 21-nt RNAs that do not show obvious 1U bias (Figure 4D). The 21-nt RNAs detected in the transgene flanking regions can be rather attributed to the endo-siRNA population. We failed to detect phasing generated by Dicer among these RNAs; however, this observation may be explained by insufficient abundance of small RNAs detected in the transgene flanking regions. In strains 1.9 and 3.6, the regions producing small RNAs of both polarities spread as far as 10-kb upstream of the transgene insertion sites. In both cases, the transgene is inserted just upstream of a gene. In 2.1, production of small RNAs was observed upstream and downstream of the insertion site. This transgene was inserted in the 5′-untranslated region (UTR) of gene CG32486 in the opposite direction relative to the direction of gene transcription (Figure 4B). Production of small RNAs downstream of the transgene is restricted by the CG32486 transcription initiation site. All these data indicate that many genomic sequences could be

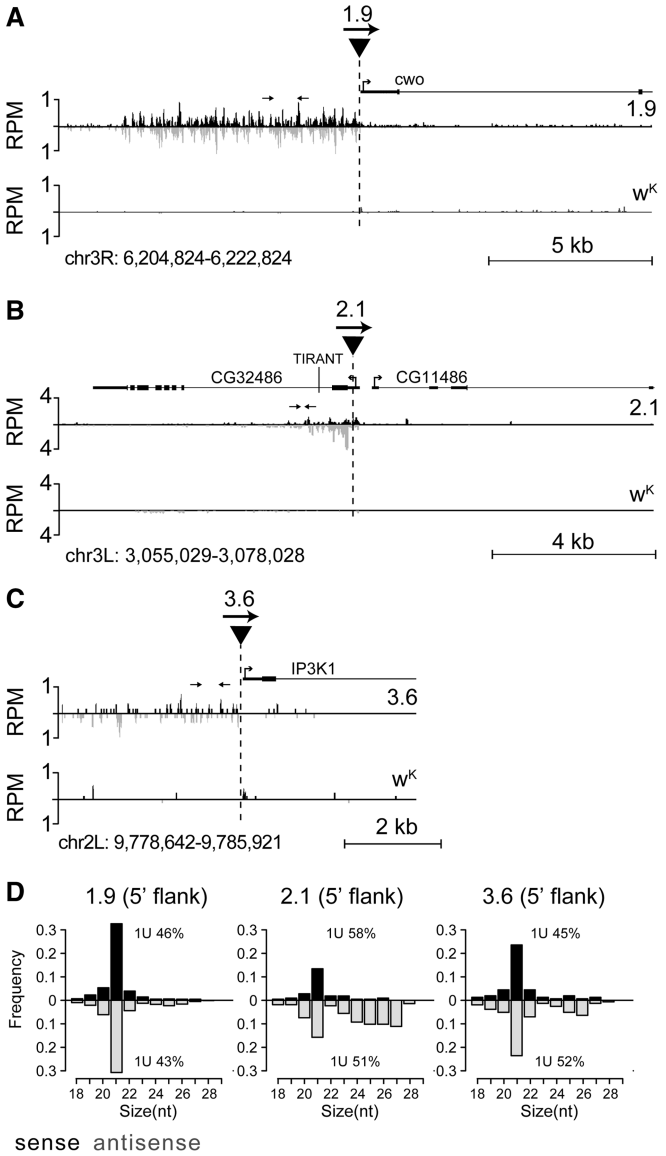

Figure 4. Transgene insertions induce generation of small RNAs from flanking genomic sequences. (**A–C**) Plots of unique small RNAs density, in a 30-bp window, around transgene insertion sites for genomic plus (black) and minus (grey) strand, in transgenic strains 1.9, 2.1, 3.6 and R-strain w^K (genomic positions according to dm3 assembly are indicated). Read numbers were normalized to sequencing depths of libraries (rpm). Position and orientation of transgenes are shown by arrowheads and arrows, respectively. The structures of the genomic regions are diagrammed above plots. Insertion of *Tirant* present in the genome of the sequenced strain (insertion site indicated in B) was not detected in the w^K and transgenic strains. (**D**) Length distribution of small RNAs mapping to flanking genomic regions. Percentages of reads having 1U bias are indicated for each strand. Primers used in the RT–PCR and ChIP are shown (not to scale).

involved in the transgene-mediated production of small RNAs (Supplementary Discussion).

piRNAs are believed to be generated from long precursor molecules encoded by piRNA clusters. We show that *I*-transgenes are bidirectionally transcribed, producing the substrate for biogenesis of piRNAs of both polarities from the entire transgene and flanking genomic regions (Supplementary Figures S8 and S9). We found small RNAs overlapping the border between transgene and neighbouring genomic sequences, which confirm the presence of read-through transcripts that start within or outside the transgene and span into flanking regions (Supplementary Figure S10). Formation of the transgene-associated piRNA clusters may change the expression of target genomic regions (Supplementary Results and Discussion and Supplementary Figure S9).

Transgene-associated *hsp70* piRNAs cause *trans*-effect on the endogenous *hsp70* expression

Libraries from transgenic lines are enriched in piRNAs homologous to the *hsp70* promoter fragments as compared with the wild-type levels of *hsp70*-derived piRNA found in w^K ovaries (Supplementary Table S4). Small RNAs complementary to *hsp70* promoter show strong 1U bias, which is a signature of primary piRNA populations (Supplementary Table S5). Moreover, in transgenic lines, the amount of piRNAs complementary to *hsp70* downstream of the fragment present in transgenes was estimated to be considerably higher than in the w^K strain (Figure 3A–C). The 10-nt overlap between 5′-ends of *hsp70* piRNAs that map to opposite strands indicates that they participate in the ping-pong amplification cycle (Figure 3D). RT–PCR analysis showed that the expression of the *hsp70B* in the ovaries of transgenic flies decreased in comparison with w^K in non–heat-shocked flies (Figure 3E). Under heat-shock conditions, no differences in *hsp70* expression were detected between w^K and transgenic ovaries (data not shown). At the same time, the appearance of *actin5C*-specific piRNAs of both polarities (Figure 2) in the transgenic flies does not lead to expansion of piRNA density beyond the homology region within *actin5C* mRNA or changes in the expression level of endogenous *actin5C* (data not shown). It is likely that the transgene-derived *hsp70* piRNAs facilitate cleavage of homologous endogenous *hsp70* transcripts with subsequent processing of heterologous parts of transcripts to small RNAs. We believe that potent bidirectional transcription of the endogenous *hsp70* loci (28) provides substrates for efficient piRNA amplification stimulated by the transgenic *hsp70* piRNAs.

Chromatin status of the *I*-element containing transgenes

Heterochromatin formation is required for piRNA cluster transcription (14–16). Trimethylated histone H3 lysine 9 (H3K9me3) generated by the methyltransferase dSETDB1 serves as a mark of piRNA clusters in the germ line (15). Therefore, we investigated H3K9me3 distribution along transgene-generated piRNA clusters in the ovaries by ChIP with anti-H3K9me3 antibody. By using transgene-specific primers, we compared H3K9me3 accumulation in strains that produce either high or low levels of transgene-specific piRNA. We show that in *I*-sense (1.9, 2.1) and *I*-antisense (3.1, 3.6) strains, the level of H3K9me3 at *I*-TG portion of transgenes is 4- to 6-fold higher relative to transgenic strain, producing the lowest amount of piRNAs (3.9) (Figure 5A). Using primers specific to the 5′-P-element fragment of transgene, we observe high level of H3K9me3 marks at the corresponding region of *I*-sense and *I*-antisense constructs compared with I-promoterless transgene (Figure 5B). Knowing that small RNA production spans into unique genomic regions bordering the transgenes (see earlier in the text), we decided to look whether this correlated with the expansion of H3K9me3 occupancy into these regions. To do this, we compared H3K9me3 deposition at genomic regions flanking the transgene with the same genomic regions in strains that do not contain transgenes in these sites, i.e. w^K or strains with a transgene in a different insertion site. We found that genomic regions adjacent to the 1.9, 2.1 and 3.6 transgenic sequences were highly enriched in H3K9me3 in comparison with the same sequences in the absence of transgene (Figure 5C). Regions flanking control (62.5.2) and I-promoterless (67.2.1) constructs display ~1- to 2-fold enrichment in H3K9me3 over 'empty' sites. We do not observe differences in the deposition of H3K27me3 and H3K4me3 histone marks or HP1a recruitment to the transgenic insertion sites in comparison with 'empty' sites (data not shown). Thus, H3K9me3 occupancy precisely correlates with the level of piRNA production by transgene-induced clusters.

Analysis of 21-nt RNAs mapping to *I*-transgenes and 42AB piRNA cluster

As mentioned earlier in the text, *I*-transgenes generate abundant 21-nt RNAs. During the analysis of ping-pong signals in the piRNA populations, we noticed that sense/antisense pairs that overlap by 10 nt are also formed between 21- and 24–29-nt RNAs mapped to I-TG, and, to a smaller extent, within 21-nt RNA population (Supplementary Figure S4). Notably, 21mers were enriched on the antisense strand of I-TG (relative to canonical *I*-element); hence, a prominent ping-pong signature was more obvious when sense piRNAs were compared with antisense 21-nt RNAs (also relative to canonical *I*-element). The 21-nt-long RNAs demonstrate a strong enrichment of uridine at the first position comparable with that of genuine piRNA populations (Supplementary Table S5). These observations held true for the rest of the transgene, at least in lines with big numbers of mappers to non-I-TG regions, such as 2.1 and 3.1 (data not shown). The 1U bias cannot be explained by the presence of a small fraction of authentic 21-nt piRNAs, as the length distribution of 1U-only mappers showed a clear peak at 21 nt (Supplementary Figure S11), which is absent from small RNA libraries made from immunoprecipitated PIWI subfamily (8,24). Genuine siRNAs do not demonstrate a tendency to begin with uridine (9).

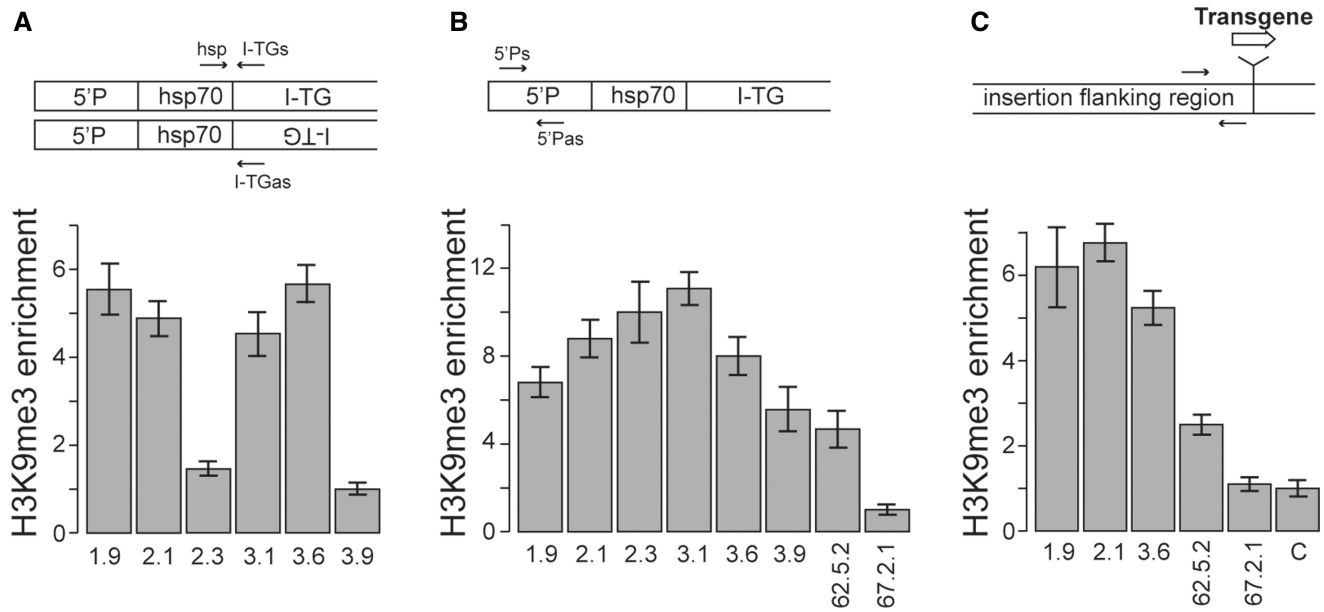

Figure 5. Chromatin status of the transgene-associated piRNA producing loci. ChIP was done using antibody specific to H3K9me3. (**A**) The ratio of normalized quantities of precipitated transgenic construct regions in *I*-transgenic strains to that in *I*-transgenic strain 3.9 (this strain produces the lowest amount of small RNAs mapping to the transgene). Locations of primers that amplify hsp70-I-TG transgenic regions in *I*-sense and *I*-antisense constructs are indicated above the graphs. (**B**) The ratio of normalized quantities of precipitated 5′-P-element transgenic construct regions (primers are indicated above) in *I*-transgenic strains to that in the *I*-promoterless strain 67.2.1 (**C**) The ratio of normalized amounts of precipitated transgene flanking regions in transgenic strains 62.5.2, 67.2.1, 1.9, 2.1 and 3.6 to that in w^K or in strains with different transgene insertion sites (designated as C, for control). Error bars indicate standard deviation of triplicate PCR measurements. Primers specific to the flanking genomic regions adjacent to the 5′-P fragment of transgenes were used in this analysis (shown above the graph). Primers to 1.9, 2.1 and 3.6 flanking regions are shown in Figure 4.

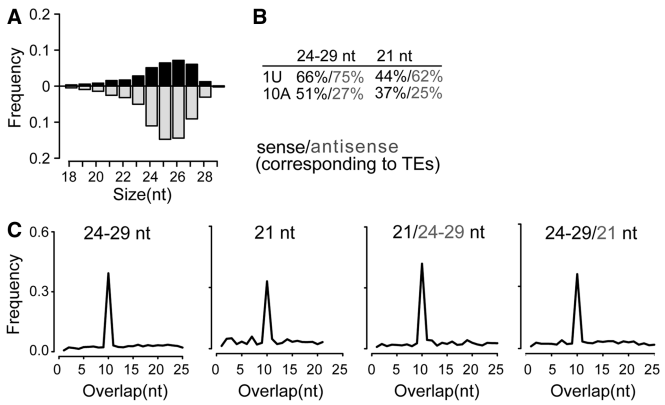

Figure 6. Characteristics of 21-nt-long RNAs coming from piRNA cluster 42AB. (**A**) Length distribution of small RNAs mapped uniquely to piRNA cluster 42AB. (**B**) Percentages of reads having 1U and 10A are indicated for each strand. (**C**) The relative frequencies (*Z*-score) of 5′-overlap for sense and antisense small RNAs within 21-nt RNA population and between 21-nt-long RNAs and piRNAs uniquely mapped to 42AB. Reads mapped to the sense strand of TEs are shown in black, antisense in grey.

We set out to test the universality of this phenomenon by looking at uniquely mapping small RNAs of major piRNA cluster 42AB. Previously, endo-siRNAs and piRNAs were shown to have a similar distribution at this locus (8). Similar to the I-TG region in transgenic strains, unique mappers to 42AB sequence in w^K were represented by genuine 24–29-nt-long piRNAs and shorter populations of RNAs (Figure 6). The 1U bias, ping-pong within 21-nt RNA population and between 21-nt-long RNAs and piRNAs was observed (Figure 6). The peculiar piRNA-like properties of 21-nt-long RNAs mapping to transgenic and endogenous piRNA clusters suggest that they belong to either a subclass of piRNAs or represent a distinct piRNA cluster-associated population of endo-siRNAs that might be involved in piRNA biogenesis.

DISCUSSION

We have studied the mechanism of the acquired TE resistance in the *Drosophila* strains carrying transgenes containing transcribed *I* retrotransposon fragments. We found that *I*-element silencing in the transgenic strains is linked to production of *I*-specific piRNAs by transgenes. Moreover, small RNAs of both polarities are generated from the entire transgene and from the flanking genomic sequences. These data indicate that *I*-containing transgenes, being inserted in unique euchromatic regions that do not generate piRNAs, seem to form *de novo* dual-strand piRNA clusters. Importantly, all tested transgenes, even with the lowest piRNA production level, were able to silence *I*-element activity (21). Several transgene insertions induce generation of small RNAs in the regions flanking the transgenes. It seems that almost any genomic sequence can be involved in the formation of a piRNA cluster. However, analysis of *I*-transgenic strains indicates that the genomic position of transgene insertions is an important factor, which determines the abundance of small RNAs generated by a given transgene (Supplementary

Results and Discussion). Antisense read-through transcription of transgenes, as in the case of strain 2.1, promotes generation of small RNAs, whereas intronic transgenes produce the lowest amount of small RNAs. At the same time, transgenes located in the intergenic regions (1.9, 3.6 and 2.10) also form double-stranded piRNA clusters, suggesting that this process itself could stimulate bidirectional transcription of the locus. Further investigation is needed to elucidate principles of transcriptional regulation of the double-stranded piRNA clusters.

Both transgene and flanking sequences generate two types of small RNAs, 24–29 nt and 21 nt. Most of the I-TG small RNAs were 24–29 nt in size and showed the signature of primary piRNA species (1U bias), as well as a secondary piRNA subpopulation (10A bias), which indicates that transgenic primary piRNAs generate additional small RNAs via the ping-pong amplification mechanism, and that this occurs in the germ line. Irrespective of the I-TG orientation in the transgene, we observe 1U bias for the I-TG-antisense piRNA and 10A bias for the I-TG-sense piRNA populations (sense/antisense according to the functional *I*-element). Such biases are expected for transposons that are actively expressed in the germ line (1). However, strains used in this study are devoid of functional *I*-elements. It was previously hypothesized that the bias originated a long time ago during the first invasion of the *Drosophila* genome by functional *I*-elements and is transmitted through generations via maternal piRNAs (29). Indeed, piRNA-mediated silencing initiated by maternally transmitted piRNAs remained strong—at least for 55 generations in case of the artificial piRNA cluster model (30). These facts strongly suggest that sense/antisense bias for *I*-specific transgene-associated piRNAs may be determined by ancestral *I*-element-derived piRNAs.

Large piRNA clusters produce abundant endo-siRNAs in the germ line acting alongside the piRNA pathway (8). Clusters of tandem repeated transgenes have been shown to produce not only piRNAs, but a significant fraction of 21mers (30). Transgene-associated piRNA clusters described in the present study also produce 21-nt RNAs, suggesting generality of this phenomenon. At the same time, the endo-siRNA pathway is not essential for piRNA biogenesis, at least on the stage when clusters are already established (30,31). Interestingly, 21-nt RNAs produced by transgenes demonstrate 1U bias unlike siRNAs. Furthermore, 21-nt and 24–29-nt RNAs mapped to transgenes form sense/antisense pairs that overlap by 10 nt. The 21-nt class of small RNAs, possibly represents a mix of endo-siRNAs and distinct subpopulation of short piRNAs. The 19–22-nt piRNA subpopulation was previously suggested to be involved in the silencing of telomeric retrotransposons in *Drosophila* (32). Thus, different population of small RNAs operate to silence transposon sequences in the germ line. Moreover, for certain transgenes, we observed spreading effect of small RNA production extended beyond the *I*-containing transgenes into flanking genomic regions. The presence of small RNAs overlapping the boundary between transgene and flanking sequences confirms the existence of read-through transcripts, which are processed into small RNAs. Genomic sequences surrounding transgenes (1.9, 2.1 and 3.6) produce predominantly 21-nt small RNAs, which may instead be attributed to the endo-siRNA population. Recently, a particular class of endo-siRNA was shown to be associated with double-strand breaks in plants, mammals and *Drosophila* (33,34). It is possible that the introduction of transgenes in germ cells also stimulates production of specific endo-siRNAs similar to those observed in tissue culture. In co-operation with piRNAs, these endo-siRNAs can potentially trigger chromatin silencing of the transgene (transposon). We cannot exclude the possibility that transgenes without any significant homology in the genome can also be efficiently silenced. It remains to be tested whether the *Drosophila* germ line contains a specific small RNA pathway for recognition and silencing of new, potentially detrimental, insertions. We noticed that in some cases, production of small RNA could spread into flanking genomic sequences in a non-symmetric manner. For example, in strains 1.9 and 3.6, the density of small RNA is observed only upstream of the transgene insertion. In strain 2.1, production of small RNAs occurs both upstream and downstream of the transgene but abruptly terminates at the target gene transcription start site. In this case, the endogenous promoter probably provides a transcription start site of the small RNA-producing transcript. Further experiments are needed to elucidate the role of genomic contexts in the genesis of piRNA clusters.

Particular chromatin marks are believed to provide a signal that distinguishes a locus as a source of piRNAs. Histone H3 lysine 9 tri-methylation, deposited by the methylase, dSETDB1, is essential for piRNA generation by dual-strand piRNA clusters in the germ line of *Drosophila* (15). Germ line-specific HP1 homologue Rhino potentially binds H3K9me3 marks at piRNA clusters. Consistent with their small RNA pattern reminiscent of native double-stranded piRNA clusters, *I*-transgenes are enriched in H3K9me3 chromatin mark. In *Caenorhabditis elegans*, siRNAs induce locus-specific accumulation of H3K9me3 extending up to several kilobases away from the site of dsRNA homology (35). We also observed spreading of loci producing small RNA and H3K9me3 chromatin modification footprints beyond the *I*-containing transgenes into adjacent genomic regions.

Despite the fact that R-strains, including w^K used in this study, are devoid of functional *I*-elements, they still contain *I*-specific piRNAs. These piRNAs are expressed on levels not sufficient for suppression of *I*-element transpositions, at least immediately after the introduction of *I*-element in the genome of SF females (17). We believe that these pre-existing *I*-specific piRNAs are responsible for *de novo* piRNA cluster formation at *I*-transgenes. This hypothetical process may be divided into several steps. On the first step, I-TG transcripts can be recognized by PIWI proteins loaded with *I*-specific piRNAs derived from endogenous clusters. Most likely, PIWI itself, being a nuclear protein (36) and previously shown to be responsible for TE transcriptional silencing (37,38), performs this function. This recognition causes recruitment of dSETDB1, accumulation of H3K9me3 chromatin mark and possibly other factors or epigenetic modifications at

transgene sequences. These modifications could stimulate transcription of both strands of the locus and provide recognition of transgenic sequences as a source of piRNAs. As piRNAs of either orientation should stimulate these processes with equal efficiency, one would expect transgenes expressing I-TG in either orientation to be equally potent suppressors of *I*-element hybrid dysgenesis. Indeed, all of the randomly chosen *I*-sense and *I*-antisense transgenic strains used in our work were of low reactivity and produced transgenic small RNAs. According to this scenario, insertion of a transposon in a euchromatic site would be targeted to undergo chromatin modifications and become a source of additional piRNAs that further enhance silencing. We found that this phenomenon occurs in *Drosophila* germ line (S.S., S.R. and A.K., unpublished data). Insertion-dependent piRNA cluster formation is a compelling mechanism that adds a new layer of protection against TEs in addition to the cytoplasmic piRNA-mediated silencing.

Endogenous *hsp70*-specific small RNAs may also be considered as potential triggers of piRNA production from transgenic transcripts. *hsp70* genes produce Ago2-associated small RNAs in somatic cells (28). In ovaries, *hsp70* loci also generate a considerable amount of small RNAs. Here, production of *hsp70* piRNAs depends on Ago3, Aub, Zuc and Spn-E proteins, indicating their germinal origin (23,24). Recently, it was shown that maternally inherited transgene-derived piRNAs could initiate production of piRNAs by other homologous transgenes (30). This effect was exerted by a distinct tandem repeat cluster of *P{lacW}* transgenes, *T-1*, inserted in the ectopic heterochromatic region of chromosome 2R and producing a significant amount of small RNAs. *P{lacW}* contains the *hsp70* promoter, which might be responsible for piRNA production in this case. However, failure to generate piRNAs by a *BX2* tandem *P{lacW}* cluster (30) argues against a role of *hsp70* small RNAs in piRNA cluster establishment. In *I*-promoterless, as well as control constructs lacking I-TG, all containing *hsp70* promoter in some part of the transgene, only transcripts initiated at *hsp70* promoters are processed into sense piRNAs, without production of antisense piRNAs. Thus, *hsp70*-induced piRNA production, in these cases, is somewhat reminiscent of native single-stranded piRNA clusters. In these transgenic lines, we do not observe spreading of piRNA density along the entire transgene or flanking regions. Therefore, *hsp70*-specific piRNAs do not seem to be sufficient for authentic piRNA cluster formation as we observe for I-TG-containing transgenes.

We did find an interesting property of *hsp70*-specific piRNAs. Transgene-derived piRNAs complementary to the *hsp70* promoter have the capacity to exert a *trans* effect on the production of piRNAs from the endogenous *hsp70* transcripts, resulting in repression of *hsp70* expression. The presence of abundant endogenous *hsp70*-derived piRNAs distinguishes it from most other genes expressed in the germ line of *Drosophila* and may serve endogenous functions that are not yet understood. Our data suggest that the piRNA-mediated cleavage of target transcripts facilitates subsequent processing of non-homologous parts of these transcripts into piRNAs. Previously, it was shown that *I*-sense and *I*-antisense transgenic strains could induce silencing of a non-homologous reporter gene, which contained a 100-bp promoter fragment of *I*-element (I-CAT) (39). We speculate that, in this case, increase in I-TG-specific piRNA abundance has stimulated small RNA processing from the non-homologous portions of the piRNA cluster transcripts, including the promoter region of *I*-element. However, owing to an insufficient number of *I*-promoter reads, we could not conclude about the enrichment of transgenic libraries in such piRNAs (Supplementary Table S3).

Transgenic strains that we examine in this study represent a unique model to elucidate the principles of double-stranded piRNA cluster genesis. Taking into account the diversified *cis* and *trans* homology-dependent effects of transgene-associated small RNAs on gene expression, we believe that transgenic constructs may serve as powerful tools in the modulation of gene expression in the germinal tissues.

ACCESSION NUMBERS

Small RNA sequencing data are deposited at Gene Expression Omnibus (GEO), accession number GSE 41780.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1–5, Supplementary Figures 1–11, Supplementary Materials and Methods, Results and Discussion and Supplementary References [40,41].

ACKNOWLEDGEMENTS

The authors thank Alexei Aravin for help with small RNA cloning, Igor Antoshechkin (Caltech) for help with small RNA sequencing, Pierre Pouchin and Yoan Renaud for help with bioinformatic analysis and Alexandre Webster for editing the English.

FUNDING

'Wildlife: Current State and Development' of the Presidium of RAS (to A.K.); Russian Foundation for Basic Researches [09-04-00996 to S.S.]. Funding for open access charge: 'Wildlife: Current State and Development' of the Presidium of RAS.

Conflict of interest statement. None declared.

REFERENCES

1. Brennecke,J., Aravin,A.A., Stark,A., Dus,M., Kellis,M., Sachidanandam,R. and Hannon,G.J. (2007) Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell*, **128**, 1089–1103.
2. Aravin,A.A., Sachidanandam,R., Bourc'his,D., Schaefer,C., Pezic,D., Toth,K.F., Bestor,T. and Hannon,G.J. (2008) A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. *Mol. Cell*, **31**, 785–799.

3. Ipsaro,J.J., Haase,A.D., Knott,S.R., Joshua-Tor,L. and Hannon,G.J. (2012) The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis. *Nature*, **491**, 279–283.
4. Nishimasu,H., Ishizu,H., Saito,K., Fukuhara,S., Kamatani,M.K., Bonnefond,L., Matsumoto,N., Nishizawa,T., Nakanaga,K., Aoki,J. *et al*. (2012) Structure and function of Zucchini endoribonuclease in piRNA biogenesis. *Nature*, **491**, 284–287.
5. Voigt,F., Reuter,M., Kasaruho,A., Schulz,E.C., Pillai,R.S. and Barabas,O. (2012) Crystal structure of the primary piRNA biogenesis factor Zucchini reveals similarity to the bacterial PLD endonuclease Nuc. *RNA*, **18**, 2128–2134.
6. Kawaoka,S., Izumi,N., Katsuma,S. and Tomari,Y. (2011) 3′ end formation of PIWI-interacting RNAs *in vitro*. *Mol. Cell*, **43**, 1015–1022.
7. Gunawardane,L.S., Saito,K., Nishida,K.M., Miyoshi,K., Kawamura,Y., Nagami,T., Siomi,H. and Siomi,M.C. (2007) A slicer-mediated mechanism for repeat-associated siRNA 5′ end formation in *Drosophila*. *Science*, **315**, 1587–1590.
8. Czech,B., Malone,C.D., Zhou,R., Stark,A., Schlingeheyde,C., Dus,M., Perrimon,N., Kellis,M., Wohlschlegel,J.A., Sachidanandam,R. *et al*. (2008) An endogenous small interfering RNA pathway in *Drosophila*. *Nature*, **453**, 798–802.
9. Ghildiyal,M., Seitz,H., Horwich,M.D., Li,C., Du,T., Lee,S., Xu,J., Kittler,E.L., Zapp,M.L., Weng,Z. *et al*. (2008) Endogenous siRNAs derived from transposons and mRNAs in *Drosophila* somatic cells. *Science*, **320**, 1077–1081.
10. Kawamura,Y., Saito,K., Kin,T., Ono,Y., Asai,K., Sunohara,T., Okada,T.N., Siomi,M.C. and Siomi,H. (2008) *Drosophila* endogenous small RNAs bind to Argonaute 2 in somatic cells. *Nature*, **453**, 793–797.
11. Muerdter,F., Olovnikov,I., Molaro,A., Rozhkov,N.V., Czech,B., Gordon,A., Hannon,G.J. and Aravin,A.A. (2012) Production of artificial piRNAs in flies and mice. *RNA*, **18**, 42–52.
12. Khurana,J.S., Wang,J., Xu,J., Koppetsch,B.S., Thomson,T.C., Nowosielska,A., Li,C., Zamore,P.D., Weng,Z. and Theurkauf,W.E. (2011) Adaptation to P element transposon invasion in *Drosophila melanogaster*. *Cell*, **147**, 1551–1563.
13. Ronsseray,S., Lehmann,M. and Anxolabehere,D. (1991) The maternally inherited regulation of P elements in *Drosophila melanogaster* can be elicited by two P copies at cytological site 1A on the X chromosome. *Genetics*, **129**, 501–512.
14. Klattenhoff,C., Xi,H., Li,C., Lee,S., Xu,J., Khurana,J.S., Zhang,F., Schultz,N., Koppetsch,B.S., Nowosielska,A. *et al*. (2009) The *Drosophila* HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. *Cell*, **138**, 1137–1149.
15. Rangan,P., Malone,C.D., Navarro,C., Newbold,S.P., Hayes,P.S., Sachidanandam,R., Hannon,G.J. and Lehmann,R. (2011) piRNA production requires heterochromatin formation in *Drosophila*. *Curr. Biol.*, **21**, 1373–1379.
16. Pane,A., Jiang,P., Zhao,D.Y., Singh,M. and Schupbach,T. (2011) The Cutoff protein regulates piRNA cluster expression and piRNA production in the *Drosophila* germline. *EMBO J.*, **30**, 4601–4615.
17. Brennecke,J., Malone,C.D., Aravin,A.A., Sachidanandam,R., Stark,A. and Hannon,G.J. (2008) An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science*, **322**, 1387–1392.
18. Picard,G., Bregliano,J.C., Bucheton,A., Lavige,J.M., Pelisson,A. and Kidwell,M.G. (1978) Non-mendelian female sterility and hybrid dysgenesis in *Drosophila melanogaster*. *Genet. Res.*, **32**, 275–287.
19. Dramard,X., Heidmann,T. and Jensen,S. (2007) Natural epigenetic protection against the I-factor, a *Drosophila* LINE retrotransposon, by remnants of ancestral invasions. *PLoS One*, **2**, e304.
20. Jensen,S., Cavarec,L., Gassama,M.P. and Heidmann,T. (1995) Defective I elements introduced into *Drosophila* as transgenes can regulate reactivity and prevent I-R hybrid dysgenesis. *Mol. Gen. Genet.*, **248**, 381–390.
21. Jensen,S., Gassama,M.P. and Heidmann,T. (1999) Cosuppression of I transposon activity in *Drosophila* by I-containing sense and antisense transgenes. *Genetics*, **153**, 1767–1774.
22. Jensen,S., Gassama,M.P. and Heidmann,T. (1999) Taming of transposable elements by homology-dependent gene silencing. *Nat. Genet.*, **21**, 209–212.
23. Li,C., Vagin,V.V., Lee,S., Xu,J., Ma,S., Xi,H., Seitz,H., Horwich,M.D., Syrzycka,M., Honda,B.M. *et al*. (2009) Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. *Cell*, **137**, 509–521.
24. Malone,C.D., Brennecke,J., Dus,M., Stark,A., McCombie,W.R., Sachidanandam,R. and Hannon,G.J. (2009) Specialized piRNA pathways act in germline and somatic tissues of the *Drosophila* ovary. *Cell*, **137**, 522–535.
25. Aravin,A.A., Naumova,N.M., Tulin,A.V., Vagin,V.V., Rozovsky,Y.M. and Gvozdev,V.A. (2001) Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the *D. melanogaster* germline. *Curr. Biol.*, **11**, 1017–1027.
26. Pall,G.S. and Hamilton,A.J. (2008) Improved northern blot method for enhanced detection of small RNA. *Nat. Protoc.*, **3**, 1077–1084.
27. Chanas,G., Lavrov,S., Iral,F., Cavalli,G. and Maschat,F. (2004) Engrailed and polyhomeotic maintain posterior cell identity through cubitus-interruptus regulation. *Dev. Biol.*, **272**, 522–535.
28. Cernilogar,F.M., Onorati,M.C., Kothe,G.O., Burroughs,A.M., Parsi,K.M., Breiling,A., Lo Sardo,F., Saxena,A., Miyoshi,K., Siomi,H. *et al*. (2011) Chromatin-associated RNA interference components contribute to transcriptional regulation in *Drosophila*. *Nature*, **480**, 391–395.
29. Grentzinger,T., Armenise,C., Brun,C., Mugat,B., Serrano,V., Pelisson,A. and Chambeyron,S. (2012) piRNA-mediated transgenerational inheritance of an acquired trait. *Genome Res.*, **22**, 1877–1888.
30. de Vanssay,A., Bouge,A.L., Boivin,A., Hermant,C., Teysset,L., Delmarre,V., Antoniewski,C. and Ronsseray,S. (2012) Paramutation in *Drosophila* linked to emergence of a piRNA-producing locus. *Nature*, **490**, 112–115.
31. Vagin,V.V., Sigova,A., Li,C., Seitz,H., Gvozdev,V. and Zamore,P.D. (2006) A distinct small RNA pathway silences selfish genetic elements in the germline. *Science*, **313**, 320–324.
32. Khurana,J.S., Xu,J., Weng,Z. and Theurkauf,W.E. (2010) Distinct functions for the *Drosophila* piRNA pathway in genome maintenance and telomere protection. *PLoS Genet.*, **6**, e1001246.
33. Michalik,K.M., Bottcher,R. and Forstemann,K. (2012) A small RNA response at DNA ends in *Drosophila*. *Nucleic Acids Res.*, **40**, 9596–9603.
34. Wei,W., Ba,Z., Gao,M., Wu,Y., Ma,Y., Amiard,S., White,C.I., Rendtlew Danielsen,J.M., Yang,Y.G. and Qi,Y. (2012) A role for small RNAs in DNA double-strand break repair. *Cell*, **149**, 101–112.
35. Gu,S.G., Pak,J., Guang,S., Maniar,J.M., Kennedy,S. and Fire,A. (2012) Amplification of siRNA in *Caenorhabditis elegans* generates a transgenerational sequence-targeted histone H3 lysine 9 methylation footprint. *Nat. Genet.*, **44**, 157–164.
36. Cox,D.N., Chao,A. and Lin,H. (2000) piwi encodes a nucleoplasmic factor whose activity modulates the number and division rate of germline stem cells. *Development*, **127**, 503–514.
37. Shpiz,S., Olovnikov,I., Sergeeva,A., Lavrov,S., Abramov,Y., Savitsky,M. and Kalmykova,A. (2011) Mechanism of the piRNA-mediated silencing of *Drosophila* telomeric retrotransposons. *Nucleic Acids Res.*, **39**, 8703–8711.
38. Sienski,G., Donertas,D. and Brennecke,J. (2012) Transcriptional silencing of transposons by piwi and maelstrom and its impact on chromatin state and gene expression. *Cell*, **151**, 964–980.
39. Jensen,S., Gassama,M.P., Dramard,X. and Heidmann,T. (2002) Regulation of I-transposon activity in *Drosophila*: evidence for cosuppression of nonhomologous transgenes and possible role of ancestral I-related pericentromeric elements. *Genetics*, **162**, 1197–1209.
40. Dufourt,J., Pouchin,P., Peyret,P., Brasset,E. and Vaury,C. (2013) NucBase, an easy to use read mapper for small RNAs. *Mob. DNA*, **4**, 1.
41. Langmead,B., Trapnell,C., Pop,M. and Salzberg,S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.*, **10**, R25.