Available online at www.sciencedirect.com
ScienceDirect
Biochemical and Biophysical Research Communications 362 (2007) 126–131
BBRC
www.elsevier.com/locate/ybbrc

Identification and characterization of Piwi subfamily in insects

Xue Zhou [a,b,1], Zhen Liao [a,1], Qidong Jia [a], Luogen Cheng [b,*], Fei Li [a,*]

[a] *Department of Entomology, Nanjing Agricultural University, Nanjing 210095, China*
[b] *College of Life Sciences, Nanjing Normal University, Nanjing 210097, China*

Received 26 July 2007
Available online 10 August 2007

Abstract

As a subfamily of Argonaute proteins, Piwi is poorly understood compared with Ago subfamily until recent discovery of Piwi protein interacting with piRNA. We did a large scale screening of insect genomes to identify piwi-like genes. Full or partial cDNA sequences were obtained by EST elongation and GENSCAN. We found that the exon numbers were totally different between vertebrates and invertebrates, approximately 20 exons in mammals but only 6–9 exons in insects. This infers either intron insertion or loss occurred during evolution. Characterized PAZ, c-terminal PIWI domains exist in almost all predicted Piwi-like proteins. We found six conserved motifs, which contain active catalytic triad "Asp-Asp-His/Lys" required for slicer activity. The expression of siwi1 and siwi2 in *Bombyx mori* were verified with RT-PCR. Phylogenetic tree inferred by Bayesian algorithm indicates invertebrate Piwi-like proteins are classified into three clades, of which Ago3 clade is closer to mammalian Piwi proteins.

Keywords: Piwi; Insect genome; Domain; Motif; EST; Evolution

Argonaute proteins have important functions in gene silencing through RNA interfering mechanism [1–3]. These proteins are essential for a wide range of regulation machinery, such as induction of histone and DNA methylation, mRNA breakdown and inhibition of translation, retrotransposons silencing, germ cell development and stem cell self-renewal [4]. Argonaute proteins are characterized by the presence of two domains: PAZ and PIWI. PAZ domain is of 110 aa in the middle region, named after proteins Piwi, Argonaut, and Zwille. It functions as a nucleic acid binding domain with a strong preference for single-stranded nucleic acids or RNA duplexes with single-stranded 3′overhangs [5]. PIWI domain is of 300 aa in the C-terminal region and structurally similar to RNaseH catalytic domain, which performs a crucial role in slicer activity [6–8].

In animals, Argonaute protein family can be divided into two classes: Ago subfamily and Piwi subfamily. Ago subfamily is ubiquitously expressed and participates in siRNAs and miRNAs pathways, whereas Piwi subfamily is expressed specifically in germline and early development of embryo [1,6,9]. The first protein in Piwi subfamily was found in *Drosophila* germline stem cells (GSCs) [10]. It has been reported that Piwi subfamily plays an essential role in germ cell development, retrotransposons silencing and germ stem cell self-renewal. Recent deep sequencing of small RNA profile in mouse and zebrafish testis revealed that Piwi proteins interact with rasiRNA or piRNA [11–13]. This highlights the functional importance of Piwi subfamily, though the detailed mechanism remains unclear.

Though the first Piwi protein has been found ten years ago, Piwi lag behind Ago subfamily in function and structure analysis. This is partially due to only Ago subfamily involves in siRNA and miRNA induced gene silencing pathway. Recently, Piwi subfamily also becomes a hotspot as an important participator in piRNA biogenesis. Present understanding implies that the organism expressing Piwi protein might also express piRNAs. Unfortunately, studies

* Corresponding authors. Fax: +86 25 84399025.
E-mail addresses: chengluogen@njnu.edu.cn (L. Cheng), lifei@njau.edu.cn (F. Li).
[1] The authors wish it to known that, in their opinion, the first two authors should be regarded as joint first authors.

0006-291X/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2007.07.179

of Piwi proteins are limited to some model organisms [14–17]. Four Piwi-like proteins have been identified in *Homo sapiens* [18], three in *Mus musculus* [17,19] and one in *Rattus norvegicus* [15]. Two Piwi homologues, Ziwi and Zili, have been reported in *Danio rerio*. Ziwi is expressed in both male and female gonad and is a component of a germline-specifying structure called nuage [20]. In *Drosophila melanogaster*, Piwi subfamily consists of three members: Piwi, Aubergine, and Ago3 [12,21,22], while two Piwi homologues, *prg-1* and *prg-2*, have been found in *C. elegans* [22]. Nevertheless, there is no Piwi homologous protein found in yeast and plants. It raises a question of how many species containing Piwi subfamily and its interacting partner piRNA or rasiRNA.

Insect species is the largest group of animals. Nevertheless, Piwi proteins have been only investigated in *D. melanogaster*. The increasing insect genome and expression data enable us to carry out a large scale survey of piwi-like sequences. Moreover, identification of Piwi subfamily in invertebrate is not only important for elucidating function and evolution of Piwi proteins but also helpful for piRNA or rasiRNA research. Here, we did a large scale screen of insect genomes to identify Piwi-like proteins in *Bombyx mori*, *Anopheles gambiae*, *Apis mellifera* and ten *Drosophila* species. Thirty five novel Piwi-like proteins were found. The full cDNA sequences of six putative Piwi proteins were obtained by e-PCR, while major part of cDNA sequences of other Piwi-like proteins were clustered by EST elongation. A rich amount of ESTs were perfectly aligned to predicted piwi-like genes. Characterized PAZ and PIWI domain and six motifs were found, indicating high conservation of Piwi subfamily. Almost all insect Piwi proteins have active catalytic residues "Asp-Asp-His/Lys", implying possessing of slicer activity. We also found that exon numbers of piwi genes were different between vertebrates and invertebrates. Phylogenetic tree shows that piwi is a high conserved family across diverse organisms including mammal, insect, etc.

Material and methods

Data collection. The *Bombyx mori* genome sequence were downloaded from SilkDB [23]. The genome sequences of ten *Drosophila* species and *An. gambiae*, *A. mellifera* were accessed via UCSC Genome Bioinformatics database. The sequences of 13 known Piwi proteins of *D. melanogaster*, *H. sapiens*, *M. musculus*, *R. norvegicus* and *D. rerio* were downloaded from Genbank according to gene Accession numbers reported in references (Table S1).

Homology searching. We used BioEdit to create local blast databases of insect genomic sequences. Thirteen Piwi protein sequences were used as queries to search local databases using TBLASTN algorithm with default parameters. All blasted sequences with *E*-value lower than 1E-20 were manually examined and remained as candidate seeds for further analysis. We wrote Perl scripts to extract 10 kb upstream and downstream genome sequences flanking candidate seeds for *An. gambiae*, *A. mellifera* and *Bombyx mori*. For those species with abundant ESTs, we did *in silico* elongation by aligning ESTs with genome sequences to get cDNA sequences as long as possible (as detailed below). We also used GENSCAN to predict transcriptional starting sites (TSS) and exon/intron boundary.

EST elongation and expression analysis. The EST sequences of ten *Drosophila* species and *An. gambiae*, *A. mellifera* and *Bombyx mori* were downloaded from UCSC Genome Bioinformatics database and dbEST. In total, there are 153,165 ESTs for *An. gambiae*, 78,085 for *A. mellifera*, 184,509 for *Bombyx mori*, 22,662 for *Drosophila erecta*, 5013 for *Drosophila simulans*, 11,135 for *Drosophila yakuba*, 541,595 for *D. melanogaster*, 21,262 for *Drosophila grimshawi*, 19,844 for *Drosophila mojavensis*, 36,014 for *Drosophila pseudoobscura*, 23,379 for *Drosophila ananassae* and 22,152 for *Drosophila virilis*. We aligned ESTs with extracted genome sequences to determine exon/intron boundaries. For those species with abundant ESTs, we did *in silico* elongation to get cDNA sequences as long as possible. We then blasted local EST databases with putative cDNA sequences as queries for expression analysis. The ESTs meeting following criteria were remained: (1) sequence identity is more than 95%; (2) the length of matched sequences is larger than 160 bp.

Domain and motif prediction. The deduced amino acid sequences were aligned with ClustalX and viewed with Genedoc. Sequence Identity Matrices were obtained from BioEdit. Protein domains were analyzed by searching CDS Conserved Domains databases. Motif finding was carried out with MEME program [24] and motif sequence logos were drawn by WebLogo [25].

RT-PCR verification and sequencing. Total RNA was isolated from egg and second instar larva of silkworm using the TRIzol reagent (Invitrogen). The first strand of cDNA was synthesized with M-MLV reverse transcriptase (Promega). PCRs for amplifying siwi1 genes were performed with standard procedure at an annealing temperature of 55 °C. Touchdown PCR was carried out to amplify the fragments of siwi2. The annealing temperature was designed to decrease from 65 to 50 °C with intervals of 3 °C, three cycles at each annealing temperature. Additional 25 cycles were appended at 50 °C. The PCR products were purified with AxyPrep DNA Gel Extraction Kit (AXYGEN) and then cloned into pGEM®–T Easy vector (Promega). The ligation reactions were used for transformations with *Escherichia coli* DH5α competent cells. The primers for siwi1 were: forward primer: 5′AGCACTGGAGTCCGAAAGGC3′, reverse primer: 5′GCGACCACGTTGAACGAAGA3′. The primers for siwi2 were: forward primer: 5′TCAGTTGAATTGCAAACTGGG3′, reverse primer: 5′CACATCTTGTACGTCAAACGCT3′. PCR products were separated on a 2.0% agarose gel and visualized with Goldview staining. Recombinant plasmids were sequenced by Shanghai Sangon Biological Engineering Technology & Services Co., Ltd.

Phylogenetic analysis. An evolutional tree of Piwi subfamily was constructed with conserved amino acid blocks of 13 known Piwis and 35 predicted Piwi-like proteins. All protein sequences were aligned using ClustalX. Phylogenetic tree was constructed using MrBayes [26] and viewed with TreeView [27]. The consensus tree was generated using the following parameters: Minimum Evolution method, Dayhoff matrix model for protein distance calculation.

Results

Predicted insect piwi-like genes

Thirteen well-studied Piwi proteins from human, mouse, rat, zebrafish and fruitfly were downloaded from Genbank. The protein sequences are used as queries to blast a local insect genome database for searching putative piwi-like sequence in mosquito, silkworm, honeybee and ten related fruitfly species. In total, 35 novel piwi-like genes were obtained, of which 30 were from ten *Drosophila* species, two from *An. gambiae* and one from *A. mellifera*. Two piwi-like genes, siwi1 (EU034629) and siwi2 (EU045577), were also identified from domesticated *Bombyx mori* unassembled genomic sequence. Siwi1 is located in Scaffold000329 plus strand, while siwi2 located in

Scaffold000138 minus strand (Table S1 and Supplementary materials).

Since only part of gene sequences can be obtained by homology searching, we extracted genome sequences of about 20 kb flanking predicted piwi-like fragments. We then tried to get the cDNA sequence as long as possible by EST alignments. We also used new version of GENSCAN software to scan genome sequence for transcriptional starting site (TSS) and exon/intron boundary information. Six predicted piwi-like genes have intact open reading frames (Table S1). By blasting Genbank database with predicted piwi-like genes, we found that some Piwi-like proteins from *An. gambiae* (XP_564296 and XP_310187) and *A. mellifera* (XP_395884) have already been deposited in Genbank. These sequences were annotated to be predicted by automated computational analysis using GNOMON method. Sequence alignments indicated that piwi-like genes predicted by our strategy provide more detailed information, as we used ESTs to increase reliability.

RT-PCR and EST abundance analysis

To ensure the reliability of our predictions, we performed RT-PCR verification and EST abundance analysis. Siwi1 and siwi2 in silkworm were selected for experimental validation. Specific primers were designed based on predicted cDNA sequences. As shown in Fig. 1, fragments of siwi1 and siwi2 were obtained with expected size (1800 and 600 bp). The expression of siwi1 and siwi2 were verified by sequencing.

ESTs were also used to estimate expression of predicted piwi-like genes in diverse insects. Multiple piwi-like genes were predicted for some insects such as *Bombyx mori* and *An. gambiae*. Their expressions are quite different according to EST abundance analysis. In silkworm, there are 13 ESTs supporting siwi1 but only one matched with siwi2. This is also the case for the mosquito. Ten ESTs are matched to AgPIWI2 but only 5 supporting AgPIWI1. This is possible due to developmental or tissue-specific expression of diverse piwi-like genes ([2] and unpublished data). It also remains possible that present available ESTs in insects are not abundant enough to cover all genes.

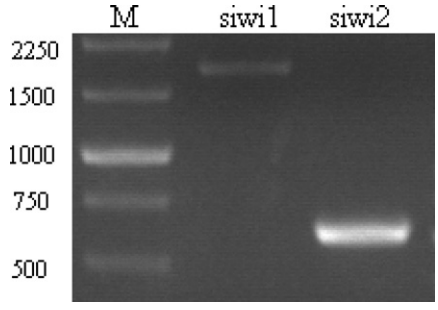

Fig. 1. PCR results of siwi1 and siwi2 in *Bombyx mori* viewed with 2% agarose gel stained with Goldview. M: 250 bp DNA ladder.

Domain, motif and catalytic residues in insect Piwi subfamily

All predicted 35 novel Piwi-like and 13 known Piwi proteins were analyzed using CDS-Conserved-Domains-prediction-sever. Two conserved domains, PAZ and PIWI, were identified in all Piwi proteins (Table S2). Averagely, the C-terminal PIWI domain is 392 aa and PAZ domain in the middle is 113 aa residues. This characterized feature also proves that predicted genes are piwi-like. We also used MEME server to predict conserved motifs in insect Piwi-like proteins (Fig. 2). Six highly conserved motifs were found, which contain active catalytic residues. We interestingly found that sequences flanking active catalytic triad are not the most conserved region of Piwi proteins. This infers there are other functional residues remained to be elucidated.

The catalytic triad is important for slicer activity of Piwi protein. The Piwi proteins containing degenerate active catalytic motifs as "Asp-Asp-Asp/His/Glu/Lys" are believed to have slicer activity or slicer-like. Those without an intact catalytic motif are non-slicer. Only one member, HsAgo2, of human Argonaute family demonstrates slicer activity. We found that all insect Piwi proteins except DvPiwi have active catalytic motifs as "Asp-Asp-His/Lys". The DvPiwi has an incomplete catalytic motif as "Asp-Asp-Asn", implying it is a non-slicer.

Genomic structure of Piwi family members

Genomic structures of Piwi subfamily members were determined by sequence alignment of putative cDNA and genomic sequences. Exon number and the length of each exon in insect Piwi proteins are in general similar except AgPIWI2 (Fig. 3). Though there is no apparent difference for the full length of cDNA sequence about 800–900 aa between diverse organisms, less exon is found in invertebrate Piwi proteins than their counterparts in vertebrates. There are 20–23 exons in mammal Piwi proteins but only 6–9 exons exist in insects. Either intron insertion or loss during evolutional procedures may cause the exon number difference. Which is the case needs further investigations. We also found that the first intron of insect Piwi subfamily is much longer than other introns in same gene.

Phylogenetic analysis

A total of 48 Piwi proteins were collected for phylogenetic analysis, including 35 predicted insect piwi-like genes and 13 known piwi genes mining from published references. Bayesian algorithm was used to construct phylogenetic tree using the minimum Evolution method with a multiple alignment of 48 Piwi protein sequences. As shown by evolutional tree, Piwi subfamily proteins are classified into three major clades (Fig. 4). The largest clade is insect specific, comprising 27 insect Piwi subfamily genes such as DmePIWI and DmeAUB. Insect Ago3 protein is in a separate clade along with vertebrate Piwi proteins, Hili, zili

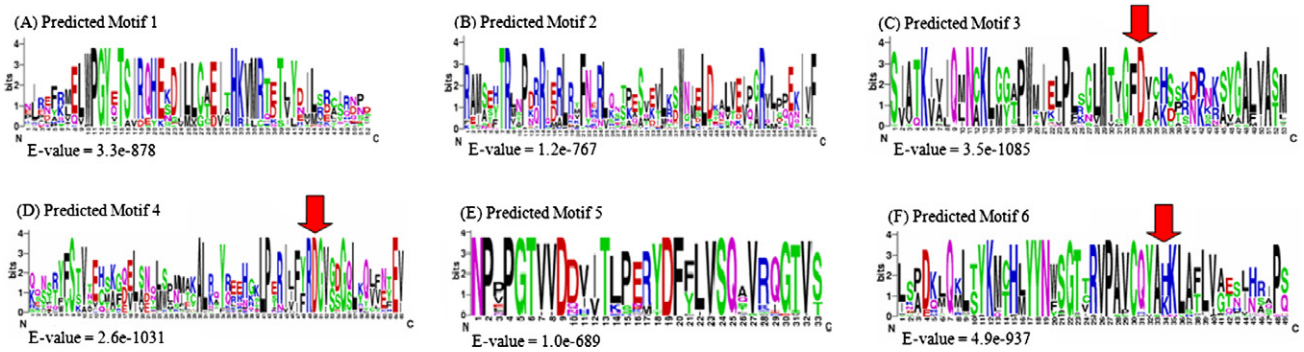

Fig. 2. Six motifs of Piwi subfamily predicted by MEME sever. The motif1 is in PAZ domain and other motifs are in PIWI domain. *E*-value of each motif is given. Three catalytic residues of Piwi subfamily are indicated by arrows.

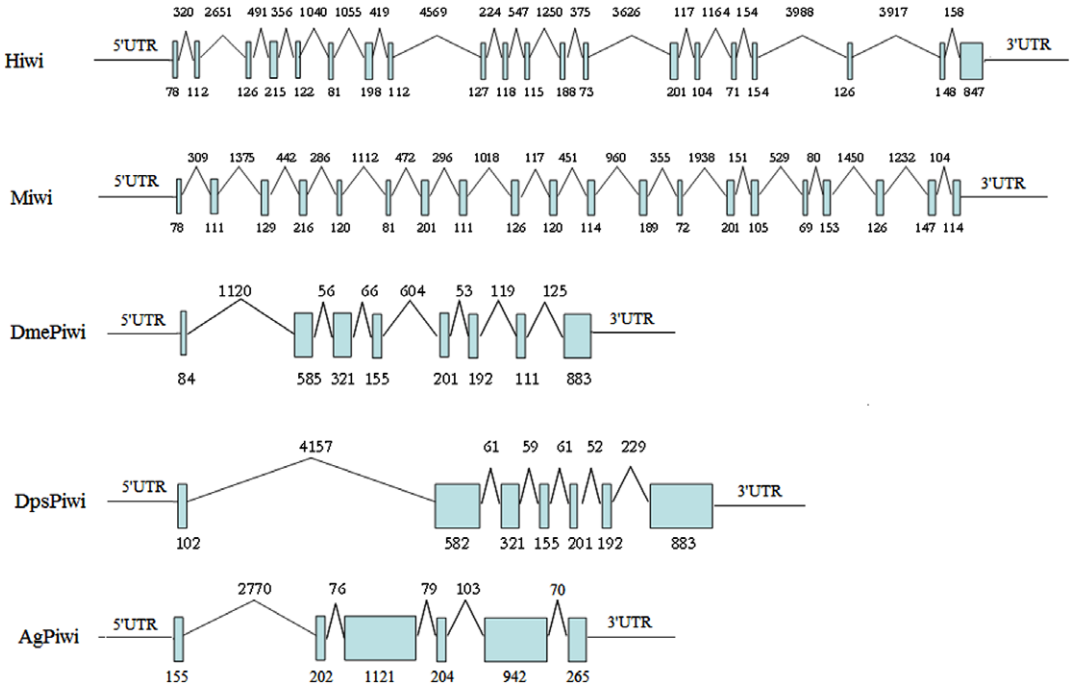

Fig. 3. Genomic structure of piwi-like genes from *Homo sapiens* (HIWI), *Mus musculus* (MIWI), *Drosophila melanogaster* (DmePIWI), *Drosophila pseudoobscura* (DpsPIWI), *Anopheles gambiae* (AgPIWI1). The number above the black line (intron) or below blue box (exon) represents length of corresponding intron or exon. Genomic structures are conserved within insects or mammals. However, exon numbers in human and mouse are more than that of insects, suggesting intron insertion or loss occurs during evolution. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this paper.)

and mili. All vertebrate Piwi proteins except Hili, zili and mili are clustered as an independent clade. Consistent with present understanding, evolutional tree infers different function of Mili (and its homologue) and Miwi (and its homologue) [6].

Insect Piwi subfamily can be divided into three clusters, Piwi, Aubergine and Ago3 [22]. Interestingly, Piwi-cluster and Aubergine-cluster are not only in same clade of evolutional tree, but also located adjacently within genome (Table S1). For example, both piwi and Aubergine gene of *D. melanogaster* are in chromosome 2L and the distance between them is only 10.6 kb. It is likely that appearance of either Piwi or Aubergine might due to an ancient gene duplication event.

Discussion

New gene discovery remains an important work in genome annotation, especially for non-model organisms. Much effort is required for better understanding of those sequenced genomes. Unfortunately, it is still a difficult task in insect species considering poor annotations and relative few ESTs available. Some insect genomes are still not assembled. Moreover, there is no gene-finding algorithm specifically designed for insect genome annotation at present. In this work, we adopted a strategy to identify Piwi subfamily in insects, exploiting available ESTs to increase reliability. Comparing with automated computational analysis of insect genome, this strategy provides more information including

Fig. 4. Evolutionary tree of piwi subfamily constructed by MrBayes and viewed with TreeView. Forty-eight piwi proteins are included, including 35 predicted genes in this work. The gene Accession numbers of 13 known piwi genes are given in Table S1.

intron/exon boundary, intact ORF, and expression level. Though this method had been applied in vertebrate species, it was seldom reported in insects except *D. melanogaster* and other few species due to lacking of abundant ESTs.

The investigations of predicted insect piwi-like genes provided some new insights into Piwi proteins. An interesting finding is that exon difference between vertebrate and invertebrate. Since the full length of cDNA sequence remains similar, an increasing number of exon in mammals might due to intron insertion. Another possibility is ancient introns were removed during evolution in insect species, resulting in decreasing of exon number. Existence of many exons in mammal implies complex regulation of Piwi expression in human and mouse. For example, it increases the possibilities of producing alternative spliced isoform by combining different exons [28,29]. It can be speculated that mammal spermatogenesis or early embryo development require fine regulation. However, it has been reported that there is a negative association between expression and exon numbers or gene length [29,30]. This means that mammal Piwi protein are possibly finely regulated but at cost of transcriptional efficiency. In other words, insect Piwi proteins could be transcribed efficiently. The selection of few exons in insect may be drove by requirements of fast response to environmental stimuli in germ cell development as insects are oviparous animals.

As expected, Piwi subfamily is highly conserved between diverse organisms, as evidenced by domain and motif analysis. In Argonaute proteins, only those with active catalytic motifs "Asp-Asp-Asp/His/Glu/Lys" have slicer activity. However, possessing of active catalytic motif is not sufficient for slicer activity. Piwi proteins demonstrated to have slicer activity are called Slicer, whereas those only with active catalytic motifs but no slicer activity are Slicer-like. The other members without complete catalytic motif are named as non-Slicer [6]. In human, only one member, HsAgo2, is Slicer and HsAgo3 is Slicer-like. The other members are surprisingly non-Slicer. It seems this is not the case for insect species, as most predicted insect Piwis have active catalytic motifs "Asp-Asp-His/Lys" except DvPIWI. Most insect Piwi proteins are at least Slicer-like. It is likely that slicer activity has been eliminated in mammal Piwi proteins but the reason requires to be elucidated.

Evolutionary analysis inferred by Bayesian algorithm reveals that insect Ago3 cluster is closer to vertebrate Piwi proteins, hili, mili and zili. These proteins are in separate clade of evolutional tree. This is in partial consistent with previous expression analysis of mili and miwi in mouse. It has been reported that mili expressed in both spermatogonia and spermatocytes in adult testis, whereas miwi only expressed in spermatocytes [13]. The expression pattern infers functional difference of mili and miwi, as also

indicated by evolutionary tree. Phylogenetic analysis indicated that insect Ago3-cluster may have similar function as mili.

Finally, we noticed that the nomenclature of Argonaute proteins is a little confusable, especially when refers to piwi (Piwi, PIWI). All Argonaute proteins have characterized PAZ and PIWI domain. According to their functions, Argonaute proteins can be classified into two sub-families, Ago and Piwi. Piwi sub-family was named after *Drosophila* Piwi, the first protein in this class. In this work, multiple genes are predicted to be members of Piwi sub-family in some organisms. We named one of them as piwi, also after *Drosophila* piwi. The others are named as Ago3 or Aub according to their sequence similarities.

Acknowledgments

We thank professor Guoqing Li in Nanjing Agricultural University for critic comments on the manuscript. Peng Li and Jinjun Qian kindly provided assistance in constructing phylogenetic tree. This work was in part supported by the National Science Foundation of China (60405001) and National Science Foundation of Jiangsu Province (BK2007524).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bbrc.2007.07.179.

References

[1] A.G. Seto, R.E. Kingston, N.C. Lau, The coming of age for Piwi proteins, Mol. Cell 26 (2007) 603–609.
[2] L. Peters, G. Meister, Argonaute proteins: mediators of RNA silencing, Mol. Cell 26 (2007) 611–623.
[3] M.A. Carmell, Z. Xuan, M.Q. Zhang, G.J. Hannon, The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis, Genes Dev. 16 (2002) 2733–2742.
[4] P.D. Zamore, B. Haley, Ribo-gnome: the big world of small RNAs, Science 309 (2005) 1519–1524.
[5] Y.R. Yuan, Y. Pei, H.Y. Chen, T. Tuschl, D.J. Patel, A potential protein-RNA recognition event along the RISC-loading pathway from the structure of *A. aeolicus* Argonaute with externally bound siRNA, Structure 14 (2006) 1557–1565.
[6] N.H. Tolia, L. Joshua-Tor, Slicer and the argonautes, Nat. Chem. Biol. 3 (2007) 36–43.
[7] Y.R. Yuan, Y. Pei, J.B. Ma, V. Kuryavyi, M. Zhadina, G. Meister, H.Y. Chen, Z. Dauter, T. Tuschl, D.J. Patel, Crystal structure of A. aeolicus argonaute, a site-specific DNA-guided endoribonuclease, provides insights into RISC-mediated mRNA cleavage, Mol. Cell 19 (2005) 405–419.
[8] L. Cerutti, N. Mian, A. Bateman, Domains in gene silencing and cell differentiation proteins: the novel PAZ domain and redefinition of the Piwi domain, Trends Biochem. Sci. 25 (2000) 481–482.
[9] P.D. Zamore, RNA silencing: genomic defence with a slice of pi, Nature 446 (2007) 864–865.
[10] H. Lin, A.C. Spradling, A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the *Drosophila* ovary, Development 124 (1997) 2463–2476.
[11] S. Kuramochi-Miyagawa, T. Kimura, T.W. Ijiri, T. Isobe, N. Asada, Y. Fujita, M. Ikawa, N. Iwai, M. Okabe, W. Deng, H. Lin, Y. Matsuda, T. Nakano, Mili, a mammalian member of piwi family gene, is essential for spermatogenesis, Development 131 (2004) 839–889.
[12] M. Pal-Bhadra, U. Bhadra, J.A. Birchler, RNAi related mechanisms affect both transcriptional and posttranscriptional transgene silencing in *Drosophila*, Mol. Cell 9 (2002) 315–327.
[13] W. Deng, H. Lin, miwi, a murine homolog of piwi, encodes a cytoplasmic protein essential for spermatogenesis, Dev. Cell 2 (2002) 819–830.
[14] J.V. Hartig, Y. Tomari, K. Forstemann, piRNAs—the ancient hunters of genome invaders, Genes Dev. 21 (2007) 1707–1713.
[15] N.C. Lau, A.G. Seto, J. Kim, S. Kuramochi-Miyagawa, T. Nakano, D.P. Bartel, R.E. Kingston, Characterization of the piRNA complex from rat testes, Science 313 (2006) 363–367.
[16] S.T. Grivna, E. Beyret, Z. Wang, H. Lin, A novel class of small RNAs in mouse spermatogenic cells, Genes Dev. 20 (2006) 1709–1714.
[17] A. Aravin, D. Gaidatzis, S. Pfeffer, M. Lagos-Quintana, P. Landgraf, N. Iovino, P. Morris, M.J. Brownstein, S. Kuramochi-Miyagawa, T. Nakano, M. Chien, J.J. Russo, J. Ju, R. Sheridan, C. Sander, M. Zavolan, T. Tuschl, A novel class of small RNAs bind to MILI protein in mouse testes, Nature 442 (2006) 203–207.
[18] A.K. Sharma, M.C. Nelson, J.E. Brandt, M. Wessman, N. Mahmud, K.P. Weller, R. Hoffman, Human CD34(+) stem cells express the hiwi gene, a human homologue of the *Drosophila* gene piwi, Blood 97 (2001) 426–434.
[19] A. Girard, R. Sachidanandam, G.J. Hannon, M.A. Carmell, A germline-specific class of small RNAs binds mammalian Piwi proteins, Nature 442 (2006) 199–202.
[20] S. Houwing, L.M. Kamminga, E. Berezikov, D. Cronembold, A. Girard, H. van den Elst, D.V. Filippov, H. Blaser, E. Raz, C.B. Moens, R.H. Plasterk, G.J. Hannon, B.W. Draper, R.F. Ketting, A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish, Cell 129 (2007) 69–82.
[21] L.S. Gunawardane, K. Saito, K.M. Nishida, K. Miyoshi, Y. Kawamura, T. Nagami, H. Siomi, M.C. Siomi, A slicer-mediated mechanism for repeat-associated siRNA 5′ end formation in *Drosophila*, Science 315 (2007) 1587–1590.
[22] D.N. Cox, A. Chao, J. Baker, L. Chang, D. Qiao, H. Lin, A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal, Genes Dev. 12 (1998) 3715–3727.
[23] J. Wang, Q. Xia, X. He, M. Dai, J. Ruan, J. Chen, G. Yu, H. Yuan, Y. Hu, R. Li, T. Feng, C. Ye, C. Lu, J. Wang, S. Li, G.K. Wong, H. Yang, J. Wang, Z. Xiang, Z. Zhou, J. Yu, SilkDB: a knowledgebase for silkworm biology and genomics, Nucleic Acids Res. 33 (2005) D399–D402.
[24] T.L. Bailey, N. Williams, C. Misleh, W.W. Li, MEME: discovering and analyzing DNA and protein sequence motifs, Nucleic Acids Res. 34 (2006) W369–W373.
[25] G.E. Crooks, G. Hon, J.M. Chandonia, S.E. Brenner, WebLogo: a sequence logo generator, Genome Res. 14 (2004) 1188–1190.
[26] F. Ronquist, J.P. Huelsenbeck, MrBayes 3: Bayesian phylogenetic inference under mixed models, Bioinformatics 19 (2003) 1572–1574.
[27] R.D. Page, TreeView: an application to display phylogenetic trees on personal computers, Comput. Appl. Biosci. 12 (1996) 357–358.
[28] B. Budagyan, A. Loraine, Gene length and alternative transcription in fruit fly. Computational Systems Bioinformatics Conference (2004) 515–516.
[29] F. Chiaromonte, W. Miller, E.E. Bouhassira, Gene length and proximity to neighbors affect genome-wide expression levels, Genome Res. 13 (2003) 2602–2608.
[30] C.I. Castillo-Davis, S.L. Mekhedov, D.L. Hartl, E.V. Koonin, F.A. Kondrashov, Selection for short introns in highly expressed genes, Nat. Genet. 31 (2002) 415–418.