Opinion

Cell PRESS

How might flukes and tapeworms maintain genome integrity without a canonical piRNA pathway?

Danielle E. Skinner[1], Gabriel Rinaldi[1], Uriel Koziol[2,3], Klaus Brehm[2], and Paul J. Brindley[1]

[1] Department of Microbiology, Immunology, and Tropical Medicine, and Research Center for Neglected Tropical, Diseases School of Medicine and Health Sciences, George Washington University, Washington, DC 20037, USA
[2] Institute of Hygiene and Microbiology, University of Würzburg, Josef-Schneider-Strasse 2, Würzburg, D-97080, Germany
[3] Sección Bioquímica y Biología Molecular, Facultad de Ciencias, Universidad de la República, Iguá 4225, CP 11400, Montevideo, Uruguay

Surveillance by RNA interference is central to controlling the mobilization of transposable elements (TEs). In stem cells, Piwi argonaute (Ago) proteins and associated proteins repress mobilization of TEs to maintain genome integrity. This defense mechanism targeting TEs is termed the Piwi-interacting RNA (piRNA) pathway. In this opinion article, we draw attention to the situation that the genomes of cestodes and trematodes have lost the *piwi* and *vasa* genes that are hallmark characters of the germline multipotency program. This absence of Piwi-like Agos and Vasa helicases prompts the question: how does the germline of these flatworms withstand mobilization of TEs? Here, we present an interpretation of mechanisms likely to defend the germline integrity of parasitic flatworms.

Flukes and tapeworms lack conserved post-transcriptional regulators associated with multipotency in the Metazoa

During the past decade draft genomes of several species of the phylum Platyhelminthes (see Glossary) have been reported and/or made available in public databases: (i) the freshwater planarian *Schmidtea mediterranea* (turbellarian); (ii) the blood flukes *Schistosoma mansoni*, *Schistosoma haematobium*, and *Schistosoma japonicum* (trematodes); and (iii) the cyclophyllidean tapeworms *Echinococcus multilocularis*, *Echinococcus granulosus*, *Taenia solium*, and *Hymenolepis microstoma* (cestodes) [1–6]. Closer scrutiny of these genomes revealed that *piwi*, *vasa*, and genes encoding group 4 Tudor homologs were absent from the cestode and trematode classes of the Platyhelminthes [7,8]. These otherwise conserved, post-transcriptional regulators are associated with stem cell maintenance and germ cell development in diverse taxa – pre-bilaterians, the platyhelminth order Tricladida, the free-living platyhelminth relatives of the parasitic flatworms, and many others.

Phylogenetic analysis of the Argonaute (Ago) protein family with six platyhelminth genomes and other informative species revealed the loss of the Piwi subfamily of Argonautes in cestodes and trematodes [3,8]. This contrasts strikingly with planarians, which express Piwi-like Agos in the somatic and germ stem cells [3,8]. Additionally, it has become apparent that cestodes and trematodes evolved a clade of Argonaute proteins exclusive to the parasitic flatworms. This clade is termed the group 4 Argonautes; it is discrete from the Ago-like clade, the Piwi-like clade, and the *Caenorhabditis elegans* group 3 Argonautes clade that traditionally constitute the Ago protein family [3,8] (Figure 1). It is also noteworthy that group 4 Tudor proteins, the interactive partners of Piwi, appear to be absent from these parasitic flatworms (flukes, tapeworms) [3]. In a similar manner, phylograms of Vasa and related DEAD-box helicases further revealed the absence of orthologs of *vasa* in trematodes and cestodes [3,9]. Instead, these flatworms have evolved at least two *PL10*-like genes that cluster together and branch from the PL10 clade of DEAD-box helicases (Figure 2). Moreover, the flukes *S. mansoni* and *Clonorchis sinensis*, the cestodes noted above, and the monogenean, *Neobenedenia girellae*, encode a DEAD-box helicase that might be assignable to the PL10 family or the closely related p68 DEAD-box family [3,8–10]. Expression analysis of these latter DEAD-box helicases termed *S. mansoni vasa*-like gene 2 (*Smvlg2*) and *N. girellae vasa*-like gene 2 (*Ngvlg2*) revealed that they exhibit germline specific expression [9,10]. Silencing by RNA interference (RNAi) of *Ngvlg2* led to loss of germ cells, which indicated that the genes perform *vasa*-like roles during germline development [10]. *N. girellae* also has a third *N. girellae vasa*-like gene (*Ngvlg3*) that groups with *vasa* orthologs. This is noteworthy from a phylogenetic point of view because it has been proposed that monogeneans are basal to a clade composed of classes Trematoda and Cestoda [11] and suggests that *vasa* orthologs had been lost in the common ancestor of

Corresponding author: Brindley, P.J. (pbrindley@gwu.edu, pbrindley@email.gwu.edu).
Keywords: Platyhelminthes; Cestoda; Trematoda; *piwi*; *vasa*; *argonaute*; transposable elements; germline; piRNA pathway.

1471-4922/$ – see front matter
 http://dx.doi.org/10.1016/j.pt.2014.01.001

Glossary

Ago clade of Argonautes: the Ago subfamily of Argonaute proteins interacts with miRNAs and siRNAs. Ago proteins, which are expressed in all tissues, exhibit three structural features: PAZ (Piwi, Argonaute, and Zwille) domain, the MID (middle) domain, and the C-terminal PIWI domain.
Argonaute (Ago): catalytic component in RISCs responsible for coordinating downstream gene silencing events. Argonaute proteins bind to different classes of sncRNAs, that is, miRNAs, siRNAs, and piRNAs.
Argonaute protein family: includes the Ago and Piwi clades. *C. elegans* contain specific group 3 argonaute proteins, termed WAGO proteins, as well as the Ago and Piwi clade. Parasitic flatworms evolved a group 4 Ago clade.
DEAD-box RNA helicases: member of Superfamily 2 (SF2) of helicases. Characterized by presence of 12 conserved motifs: Q-motif, motif I, motif Ia, motif Ib, motif Ic, motif II, motif III, motif IV, motif IVa, motif V, motif Va, and motif VI. Motif II contains the amino acid sequence DEAD (Asp-Glu-Ala-Asp), from which the family name was derived. The motifs are involved in ATP binding, hydrolysis, ATP-dependent RNA binding, and helicase activities. DEAD-box helicases occur in all cells: central to metabolism of RNA, from transcription to decay.
Germline stem cells: germline stem cells have the capacity to generate all cells in the body. Under normal fate the germline stem cells are unipotent, producing gametes and self-renewing to maintain the generative population.
Group 4 Tudor proteins: Tudor domains were initially identified as a common structural motif, approximately 60 amino acids, in the *Drosophila* Tudor protein. The motif has now been identified in a wide variety of organisms. These Tudor domain-containing proteins bind to proteins with methylated arginine and lysine residues and play a role in epigenetics, gene expression, and regulation of small RNAs. There are four different functional groups of Tudor proteins distinguished by the sequence flanking of the Tudor domain. The group 4 Tudor proteins have been shown to interact with the methylated RG/RA box of Piwi proteins and are involved in piRNA biogenesis by promoting the formation of Piwi ribonucleoprotein (piRNP) complexes.
microRNAs (miRNAs): endogenous ncRNAs (~22 nt) that regulate gene expression at transcriptional and post-transcriptional levels; produced from genes or introns, processed from long primary transcripts. miRNAs target and bind complementary sequences in the 3′-untranslated region of mRNAs, leading to gene silencing via translation repression and/or mRNA degradation. Mature miRNAs associate with Ago clade of Argonautes.
Neighbor-joining method: a method for construction of phylograms (evolutionary trees) that is distanced based and uses bottom-up clustering.
Neoblasts: pluripotent adult stem cells with capacity for indefinite self-renewal, totipotency; drive long-term tissue homeostasis; endow planarians with ability to regenerate. Neoblast-like stem cells occur in tapeworms and schistosomes. Characteristics include roundish-to-ovoid cells with a high nuclear-to-cytoplasmic ratio, large nucleolus, exhibit chromatoid bodies (RNA-protein bodies near nuclei), and numerous free ribosomes.
Neodermata: the subphylum of the Platyhelminthes that includes the classes Monogenea, Trematoda, and Cestoda. All members of the Neodermata are parasites.
Nuage: term used to describe the electron dense, ribonucleoprotein-rich, perinuclear granule organelle-like structure in germ cells of *Drosophila* that plays a role in the piRNA pathway and piRNA-mediated silencing of transposons. The structure has been described in other animals and is termed chromatoid bodies, germinal granules, P granules, and mitochondrial cloud in other species. The nuage/chromatoid body has been described in the neoblasts of planarians.
Ping-Pong amplification cycle: a mechanism for piRNA biogenesis hypothesized from observations in *Drosophila* and later in other organisms. This mechanism obviates the need for RNA-dependent RNA polymerase. The mechanism is initiated by mature sense primary piRNAs, derived from TEs and transcripts from genomic loci termed piRNA clusters, bound to proteins of the Piwi clade which recognize complementary sequences, antisense transcripts from the same piRNA cluster. The antisense piRNAs are cleaved, generating secondary piRNAs with new 5′-ends targeted by another Piwi protein. The 3′-end is then trimmed and then the mature antisense secondary piRNAs are now able to target sense transcripts from the piRNA clusters generating new sense piRNAs.
Piwi clade of Argonautes: subfamily of Argonautes; typically expressed in germline and adult somatic stem cells of invertebrates; responsible for regeneration, for example, neoblasts in planarians; does not occur in parasitic flatworms. Piwi, similar to Ago, has three main structural motifs: PAZ (Piwi, Argonaute, and Zwille) domain, the MID (middle) domain, and the C-terminal PIWI domain. Piwis associate with sncRNAs termed piRNAs. This is essential for biogenesis and/or stability of piRNAs, the piRNA biogenesis pathway, and the Ping-Pong amplification cycle.
Piwi-interacting RNAs (piRNAs): small RNAs of 24–30 nt in length thought to arise from long single-stranded RNA precursors derived from TEs, intergenic regions, and protein-coding genes. piRNAs preserve genome integrity and stability, and play a role in antiviral immunity. piRNAs associate with specific Piwi proteins and together function in silencing TEs. Piwi proteins function in germline development, gametogenesis, germline stem cell maintenance, and meiosis.
PL10: an ATP-dependent DEAD-box RNA helicase, also termed DEAD-box helicase 3 (DDX3), closely related to Vasa. PL10 is not restricted to the germline or somatic stem cells. It is conserved in eukaryotes from yeast: plants to animals. It participates in RNA metabolism, tissue differentiation, embryogenesis, asexual reproduction, cell regeneration, tumorigenesis, apoptosis, co-option in viral pathology, and innate immunity.
Platyhelminthes: phylum Platyhelminthes includes parasitic clades Cestoda (tapeworms), Trematoda (flukes), and Monogenea (mostly ectoparasites of fish), and the mostly free-living planarian-like worms of order Tricladida.
Small interfering RNAs (siRNAs): class of sncRNAs (~20–25 nucleotides) notable for their role in the RNA interference pathway. They are derived from endogenous loci or foreign genetic material and are involved in silencing of invading viruses, TE activity, and in some cases protein-coding genes.
Transposable element (TE): DNA elements that can move from one location to another within the genome of a cell; these elements can move either by a cut-and-paste mechanism (DNA transposons) or indirectly through an RNA intermediate, that is, a copy-and-paste mechanism (retrotransposons).
Vasa: ATP-dependent DEAD-box RNA helicase; archetypal germ cell marker; termed DEAD-box helicase 4 (Ddx4).

flukes and tapeworms. However, transcripts encoding *Ngvlg3* mRNA were not detectable, whereas silencing *Ngvlg3* did not affect gonad anatomy in the same manner as silencing *Ngvlg1* (PL10-like cluster) and *Ngvlg2* [10].

Considered at large, the absence of Piwi, Vasa, and group 4 Tudor proteins indicates that parasitic flatworms (subphylum Neodermata, the cestodes, monogeneans and trematodes) have lost the ancestral components, classified traditionally as germline stem cell associated markers that are conserved in metazoans (Box 1). These enzymes and other factors establish and maintain *inter alia* the multipotency of progenitor germ cells (PGCs), germline stem cells (GSCs), and multipotent progenitor cells including the neoblasts (totipotent adult stem cells) of planarians, a hypothetical regulatory network collectively referred to as the germline multipotency program (GMP) [12,13]. Furthermore, it has been proposed that these factors are ancestrally linked in the GMP because together they participate in protection of the genome from the mobilization of endogenous and exogenous transposable elements (TEs), an ostensibly essential role in germline stem cells and somatic stem cells devoted to the production of progenitors for tissue renewal [14–16]. It is relevant to note here that approximately 30–50% of the genomes of the three major species of human schistosomes and other trematodes consists of TEs and other repetitive elements, although substantially less of the genome of (at least cyclophyllidean) cestodes is composed of these elements (~2–11%) [3–6,17] (Box 1).

This opinion article highlights the absence of *piwi*, the piRNA pathway, and *vasa* from flukes and tapeworms, and speculates on alternative mechanisms that might defend the integrity of the germline of the parasitic flatworms against TEs.

piRNA pathway – guardian of the germline

Piwi and associated proteins, including Vasa, are involved in the functions and synthesis of a novel class of small non-coding RNAs (sncRNAs) of 24–31 nucleotides in length termed Piwi-interacting RNAs (piRNAs). With model species, it became clear that the piRNA pathway was restricted to germ cells and, in some cases, gonadal somatic cells [13]. Deciphering the function of Piwi in the context of