bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

*article*

# Evolutionary implications of the first microRNA- and piRNA complement of *Lepidodermella squamata* (Gastrotricha)

**Bastian Fromm [1*], Juan Pablo Tosar [2,3] , Felipe Aguilera [4,5], Marc R. Friedländer [1], Lutz Bachmann [6], Andreas Hejnol [5]**

1 - Science for Life Laboratory, Department of Molecular Biosciences, The Wenner-Gren Institute, Stockholm University, Stockholm, Sweden

2 - Functional Genomics Unit, Institut Pasteur de Montevideo, Montevideo, Uruguay

3 - Nuclear Research Center, Faculty of Science, Universidad de la República, Montevideo, Uruguay

4 - Departamento de Bioquímica y Biología Molecular, Facultad de Ciencias Biológicas, Universidad de Concepción, Concepción, Chile

5 - Sars International Centre for Marine Molecular Biology, University of Bergen, Bergen, Norway

6 - Research group Frontiers in Evolutionary Zoology, Natural History Museum, University of Oslo, Oslo, Norway

* Correspondence: Bastian.Fromm@scilifelab.se

bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

**Abstract (183/200)**

Gastrotrichs - hairy bellies - are microscopic free-living animals inhabiting marine and freshwater habitats. Based on morphological and early molecular analyses, gastrotrichs were placed close to nematodes, but recent phylogenomic analyses have suggested their close relationship to flatworms (Platyhelminthes) within Spiralia. Small non-coding RNA data on e.g. microRNAs (miRNAs) and PIWI-interacting RNAs (piRNA) may help to resolve this long-standing question. MiRNAs are short post-transcriptional gene regulators that together with piRNAs play key roles in development. In a 'multi-omics' approach we here used small-RNA sequencing, available transcriptome and genomic data to unravel the miRNA- and piRNA complements along with the RNAi protein machinery of *Lepidodermella squamata* (Gastrotricha, Chaetonotida). We identified 52 miRNA genes representing 35 highly conserved miRNA families specific to Eumetazoa, Bilateria, Protostomia, and Spiralia, respectively, with overall high similarities to platyhelminth miRNA complements. In addition, we found four large piRNA clusters that also resemble flatworm piRNAs but not those earlier described for nematodes. Congruently, transcriptomic annotation revealed that the *Lepidodermella* protein machinery is highly similar to flatworms, too. Taken together, miRNA, piRNA and protein data support a close relationship of gastrotrichs and flatworms.

**Keywords:** microRNAs, piRNAs, RNAi protein machinery, Gastrotricha, *Lepidodermella squamata*, Platyhelminthes

bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

## 1. Introduction

The Gastrotricha is an animal taxon with hitherto around 700 described species that are divided into the two main clades Macrodasyida and Chaetonotida [1–3]. Gastrotrichs are microscopic animals that occur abundantly in marine and freshwater interstitial environments [1,4]. They show a direct development, and are bilaterally symmetric, with an acoelomate body, and worm-like body plan [5]. Most gastrotrich species have a commissural brain [6–9] and within the clade there is a large variety of reproductive strategies [5,10]. Interestingly, most gastrotrichs lack an ectodermal hindgut, and a specialized respiratory system has not been described [5]. Gastrotrichs share complex morphological characteristics, including a radial myoepithelial pharynx with terminal mouth, epidermal monociliation, and protonephridial structures, with several other lineages within Bilateria. This has made their phylogenetic position a matter of debate. [11]. Initially, gastrotrichs were considered to be close to rotifers ("Trochelminthes") [12]. Later, classic systematists placed gastrotrichs either as the sister group of Introverta, together forming Cycloneuralia (i.e., Kinorhyncha + Loricifera + Priapulida + Nematoda + Nematomorpha) [13,14], or in the Neotrochozoa group that comprises Gnathostomulida + Gastrotricha [15,16]. When nucleotide sequence data became available, gastrotrichs were placed within Spiralia with uncertain relatedness, within the debated group of 'Platyzoa' comprising Platyhelminthes, Gnathifera, and Gastrotricha [17–19]. Recent phylogenomic studies suggest a sister group relationship between Gastrotricha and Platyhelminthes (Rouphozoa) [20,21]. However, this Gastrotricha-Platyhelminthes clade is still controversial because there is no morphological apomorphy supporting the monophyly of the clade. Even though gastrotrichs and most flatworms ingest food with a simple pharynx [6,22,23], this common alimentary strategy is not necessarily an autapomorphy of Gastrotricha and Platyhelminthes. Thus, additional lines of evidence are required to further test the proposed close phylogenetic relationship of gastrotrichs and flatworms.

MicroRNAs (miRNAs) are a unique class of short non-coding RNAs with major roles as post-transcriptional gene regulators in animals and plants [24]. In contrast to short interfering RNAs (siRNAs), responsible for RNA interference (RNAi), and PIWI-interacting RNAs (piRNAs), responsible for genome integrity [24,25], miRNAs are deeply conserved in metazoans. Their continuous addition to genomes during evolution has often been correlated with increasing morphological complexity and number of cell-types in a given organism [26–29]. Accordingly, absence/presence of miRNAs has also been successfully applied as phylogenetic and taxonomic arguments [30–33], and were also used to ascertain clade specificity of highly derived organisms such as Myzostomida [34]. However, because miRNA annotation is a challenging task, many published miRNA complements have considerable issues with respect to quality and completeness [35,36]. For example, missing miRNA families have frequently been misinterpreted as secondary losses, and the conservation and the utility of miRNAs as phylogenetic

bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

marker has been questioned [37]. However, losses of miRNA families have rarely been observed in well-annotated complements, and the few confirmed cases were correlated to high degrees of genomic and phenotypic reductions, for instance in parasitic species [38–40]. Currently, information on the presence of the RNAi pathway and small RNAs in gastrotrichs is lacking, but several studies have reported on miRNAs, piRNA, and the RNAi protein machinery in flatworms [39,41–45]. Closer scrutiny of these studies revealed a dynamic evolution of RNAi proteins [43,44] and small RNAs, with the substantial loss of conserved and gain of novel miRNAs [39], and the loss of piRNAs in obligate parasitic flatworms and nematodes [42,43,46]. In free-living nematodes such as *Caenorhabditis elegans*, however, piRNAs have been detected and described in detail [47]. With respect to miRNAs there is substantial different between flatworms and nematodes. In addition to the miRNA families found in all protostome species, flatworms share two miRNA families MIR-1989 and MIR-1992 with spiralian taxa (we are following the nomenclature proposed by Fromm et al [35]). Nematodes are not part of Spiralia and, thus, do not share MIR-1989 and MIR-1992 but show instead additional four miRNA families specific for Ecdysozoa (MIR-305) and Nematoda (MIR-54, MIR-86, MIR-791), respectively. Thus at least six miRNA families differ between flatworms and nematodes. It is also noteworthy that, although free-living flatworms and nematodes have both piRNAs, the actual mature piRNAs and the protein machinery in nematodes are unique, and so far do not resemble what has been described for any other animal taxon. Thus, piRNAs of free-living flatworms and nematodes are clearly distinguishable. Using small-RNA sequencing and available transcriptome data, we here describe the first small RNA complement (miRNA and piRNA) as well as t the corresponding protein machinery for a gastrotrich, namely the freshwater species *L. squamata*. The analyses revealed that the genome of *L. squamata* encodes an extensive small RNA repertoire of *bona fide* miRNAs and piRNAs, as well as the full miRNA & piRNA biogenesis and RNAi protein machinery. The comparison of miRNAs, but also of piRNAs and RNAi protein machinery, to those of flatworms and nematodes strongly supports a close relationship of gastrotrichs to flatworms and not to nematodes.

bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

## 2. Results

*The microRNA complement of Lepidodermella squamata*

Small-RNA-sequencing yielded 22.4 millions raw and 217,000 unique high quality and trimmed miRNA reads (SRA Accession number: ). These reads were matched against 194.5 millions of high-quality genomic reads provided by the *L. squamata* genome sequencing project (Aguilera et al. in prep) using the miRCandRef pipeline [39]. In total 20,375 miRNA-candidate contigs were assembled from the genomic reads with perfect matches to small RNA reads, all representing potential miRNA loci. Subsequently, this miRNA-candidate reference assembly was screened with an adapted version of the MirMiner pipeline [30] for miRNA loci with longer hairpins [Fromm et al. in prep]. 46 *L. squamata* miRNAs that belonged to 35 conserved families and 6 miRNAs with yet undescribed seed sequences were identified, all of them fulfilling the latest criteria for miRNA annotation [35]. (see supplementary file S1). Given the undisputed protostome affinity of *L. squamata*, a set of conserved miRNA families was expected [30,31]. Indeed, we found the Eumetazoa-specific miRNA family (MIR-10) (Figure 1a), 25 of the 31 Bilateria-specific families and 8 of the 11 Protostome-specific miRNA families. We did not find any representatives of the four Nematoda-specific miRNA families, but there was a representative of the Spiralia-specific family MIR-1992 (Figure 1a and 1b), which is in line with the hypothesis of a close relationship of gastrotrichs to flatworms [20,21]. Interestingly, we found seven miRNAs of conserved families with particularly long hairpins (Let-7, Mir-750, Mir-22, Mir-31, Mir-1992, Mir-1) that exceeded commonly observed hairpin-sizes of 60-90 nucleotides (nt) with miRNA precursor lengths up to 386 nucleotides (Mir-750) (Figure 1b and 1c). Such long precursors are so far only described in flatworms [39,48–50].

The miRNA complement of *L. squamata* was submitted to MirGeneDB and will be available in one of the next releases.

bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

*Lepidodermella squamata has a conventional animal piRNA biogenesis mechanism*

Four putative piRNA clusters composed of 1,377 unique reads and with a total length of 18,789 bp were identified (supplementary figures S1-S4). These piRNAs were identified using stringent criteria for piRNA cluster definition to minimize false positives.

Clusters #1-#3 include piRNAs deriving from a single DNA strand, and the vast majority of the sequences had a length of 30-32 nt (Figure 2b). In contrast, cluster #4 comprised reads mapping to both strands (minus > plus). Minus strand-derived reads in cluster #4 had a similar length distribution as those in the unistrand clusters #1-#3. However, the length distribution of plus strand-derived reads was clearly bimodal, ranging from 23-26 nt and 30-32 nt, respectively (Figure 2b and 2c). Overlap with transposable elements (TEs) was partial in some of the clusters, and cluster #4 was remarkably devoid of identifiable TEs (supplementary figures S1-S4).

Main strand-derived sequences in the four clusters were strongly biased towards 5′ uridine (1T in our data), while 5′ ends in the antisense (plus strand) sequences in cluster #4 were virtually random (1T = 25%). Conversely, these sequences showed 10A bias (Figure 2d). The presence of sense sequences, with strong 1T bias and antisense sequences with evident 10A bias, strongly suggested the generation of initiator and responder piRNAs by a conventional ping-pong amplification cycle. This was confirmed by the observation that sense and antisense sequences in cluster #4 overlapped by exactly 10 nucleotides (Figure 2e).

A more detailed analysis of the genomic distribution of mapped reads in cluster #4 was also consistent with the generation of trailing piRNAs by a piRNA-independent endonuclease acting on large piRNA precursor transcripts (reviewed in [51]) (Figure 2f). Such piRNAs are phased by a length which comprises the footprint of a putative PIWI protein plus a few nucleotides to the next available uridine in the 5′-3′ direction. As observed in most minus strand-derived sequences, these putative trailing piRNAs were 30-32 nt in length. In contrast, responder piRNAs generated from the opposite strand (or from antisense transcripts present in *trans*) are either of this size or 24-25 nt long. This has been explained by the co-existence of at least two different PIWI-clade proteins (see below) acting on piRNAs derived from this cluster, with the proteins being responsible for the shorter piRNAs acting exclusively on transcripts that are antisense to the initial piRNA precursor. In summary, a conventional piRNA pathway involving all the characteristic features of the PIWI processing machinery was found in *L. squamata that*is is distinct from the situation in most nematodes [47].

bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

*The RNAi protein machinery of Lepidodermella squamata and Gastrotricha*

The small non-coding RNA pathways are best described for *C. elegans* [52]. A systematic survey for protein repertoires directly involved in RNA interference pathways was performed using a curated set of 65 RNAi proteins as queries (see supplementary table S1). We found putative orthologs of most *C. elegans* RNAi protein complexes, except for Secondary Argonautes (SAGOs) (Figure 3a).

Using a phylogenetic framework the distribution of RNAi protein repertoires within Spiralia was targeted. While the core of the RNAi pathways was generally conserved, several specific factors were absent. They may diverged to a degree that we could not identify the respective orthologs any longer in Gastrotricha(Figure 3a). Focusing on six gastrotrich species including *L. squamata*, we found that Gastrotricha possessed putative orthologs of at least one gene from each of the respective nematode RNAi complexes (Figure 3a), corroborating that gastrotrichs have the complete RNAi protein machinery involved in miRNA biogenesis. As expected, we found multiple hits of Argonaute proteins (*ALG-1* and *ALG-2*) for all gastrotrich species, except for *Mesodasys laticaudatus* (Figure 3a and supplementary table S2). We also confirmed the complete protein sets of the Microprocessor (*DRSH-1* and *PASH-1*) and Dicer complexes (*DCR-1* and *DRH-1*), except for *RDE-4* gene that is conserved only in some nematode species [53,54] and also absent in flatworms [44] (Figure 3a). Possible duplicates of *DCR-1* were found in the transcriptomes of *L. squamata* (Chaetonotida) and *Macrodasys sp.* (Macrodasyida) (supplementary table S2). Besides these core RNAi proteins, we found that *L. squamata* encoded the full canonical protein repertoire of the animal RNAi pathway, while in other gastrotrich species we could not find the complete miRNA protein sets (supplementary table S2). however, transcriptomes can only be used to identify RNAi genes, but not to conclude t that a particular gene is missing from the genome.

A screen for the presence of functional domains of some core proteins revealed that *L. squamata DCR-1* proteins had a helicase conserved C-terminal domain, a PAZ domain, a dicer dimerization domain, and two ribonuclease III domains. (Figure 3b). The predicted *L. squamata DRSH-1* and *PASH-1* proteins had a conserved organization, with two ribonuclease III domains and a double-stranded RNA binding motif in Drosha and a double-stranded RNA binding motif in Pasha (Figure 3b). The domain organization of theses proteins resembled largely to what has been described for similar proteins w in flatworms but only to a minor degree to those of nematodes [44]. The organization of functional domains in *XPO-1*, a transmembrane channel protein that transports pre-miRNAs into the cytoplasm in *C. elegans [55]*, ws also more conserved between gastrotrichs and platyhelminths rather than between gastrotrichs and nematodes (Figure 3b).

An analysis of PIWI genes targeted the proteins involved in the organization of the four piRNA clusters found in *L. squamata*. Systematic Blast searches of predicted proteins from animal genomes and

bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

gastrotrich transcriptomes identified 180 sequences corresponding to PIWI-like genes across metazoan species and confirmed the presence of seven PIWI-like proteins in Gastrotricha with PAZ and Piwi domains (supplementary table S3). Alignments of these functional domains across *L. squamata* and representatives of Platyhelminthes and Ecdysozoa revealed several conserved residues, that provided further support for considering these sequences PIWI-like proteins (supplementary figure S5).

Phylogenetic analysis (maximum likelihood) of PIWI-like proteins revealed good support for taxon-specific clades, but with only poor support for early diverging branches (Figure 4). Specifically, in *L. squamata*, we found that five PIWI-like proteins were grouped into three phylogenetic clades (Figure 4). Three PIWI-like paralogs (i.e., *Lsqu.Gene.78021, Lsqu.Gene.95003*, and *Lsqu.Gene.176266*) were grouped with *Drosophila melanogaster Piwi* and *Aubergine*, and *C. elegans PRG-1* and *PRG-2*, but with low statistical support. Similarly, the *Lsqu.Gene.203941*, along with a *Diuronotus aspetos* PIWI-like protein, was clustered with *D. melanogaster AGO3*, but also with low bootstrap support (Figure 4). Interestingly, the *Lsqu.Gene.20027* protein that grouped with strong support (i.e., 100% bootstrap) with proteins from platyhelminth species (i.e., *Schmidtea mediterranea, Echinococcus multilocularis*, and *Hymenolepis microstoma*) (Figure 4). While we did not find *L. squamata* representatives of the *Piwil1* and *Piwil2* clades, we found one PIWI-like protein from *Megadasys sp.,* that is part of the *Piwil2* clade (Figure 4). Finally, we found two clade-specific groups comprising PIWI-like proteins from *S. mediterranea* and *Macrostomum lignano,*

bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

## 3. Discussion

Gastrotrichs are an abundant group of organisms in marine habitats with important roles in aquatic ecosystems [1,4]. However, only little is known about their molecular biology [5]. In order to obtain a better molecular understanding of this, and to contribute to the discussion on their phylogenetic relatedness to other protostome species, we have studied the miRNA and piRNA gene-regulatory pathways in the chaetonotid gastrotrich *Lepidodermella squamata* using a 'multi-omics' approach.

The miRNA complement of *L. squamata* consisted of highly conserved and some novel miRNAs. There was an almost complete set of the expected miRNA families and the spiralian-specific MIR-1992, but no Ecdysozoa- or Nematoda-specific miRNAs were detected. Some of the few losses observed in *L. squamata* were shared with all other flatworms studied so far [39,48,56,57]. Interestingly, some of the gastrotrich miRNAs showed extended precursor lengths of up to 386 nucleotides, which, to this date, has only been described in flatworms [39,50,58]. Taken together, we found a near complete set of miRNAs that by the presence and absence of particular miRNA families, their high sequence similarities and the length of some precursors support a close relationship of gastrotrichs and flatworms.

The conservation of individual piRNA sequences is extremely low [59–61] and thus their potential for evolutionary or phylogenetic analyses is thought to be limited. However, we showed, that similarities in piRNA biogenesis patterns can be used to infer relatedness of animal groups such as the gastrotrichs to flatworms or nematodes. While both free-living flatworms and free-living nematodes possess piRNAs, many parasitic lineages have lost them altogether [47] or, in the case of nematodes, have retained a highly derived pathway, producing piRNAs that are clearly distinct from most other animals studied thus far [51]. For instance, piRNAs in *C. elegans* are not produced from long primary transcripts derived from genomic regions known as piRNA clusters and have a size of 21 nucleotides instead of the conventional 24-32 nt lengths typical of piRNAs in other animals [62].

Although, the four identified piRNA clusters do not constitute the full piRNA complement of *L. squamata*, they already allowed for some insights into the piRNA processing machinery in Gastrotricha. Three piRNA clusters with piRNAs mapping to only one DNA strand and a fourth cluster including piRNAs from both DNA strands were identified. The piRNAs of the latter cluster showed processing patterns consistent with the so-called ping-pong amplification mechanism [51], and the majority of the piRNAs were derived from the minus strand and showed 1T bias. In contrast, plus strand-derived piRNAs, which are antisense to the putative precursor transcript, showed strong bias towards adenine at the tenth position and overlapped minus strand-derived piRNAs by exactly ten nucleotides. It is known that the 10A bias is caused by the intrinsic affinity of PIWI-clade proteins for targets bearing adenine at position target-1 or t1 (i.e., in front of the piRNA 5′ end). When PIWI cleaves their targets in a

bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

piRNA-directed manner, it does so at position 10 of the target counting from the piRNA 5′ end. As a consequence, adenine at position t1 becomes adenine at position guide-10 or g10 [63]. piRNAs generated in this way were initially termed as secondary piRNAs, although the term *responder piRNA* has been recommended [51]. In *Drosophila melanogaster* (an ecdysozoan), the ping-pong cycle involves two of the three PIWI-clade proteins present in the genome. Aubergine (*Aub*) binds piRNAs produced from piRNA precursors while Argonaute 3 (*AGO3*) associate with complementary transcripts, usually corresponding to TEs. Aub-bound sequences tend to start with uridine, while AGO3-associated piRNAs show 10A, but no 5′ bias [64]. A similar heterotypic ping-pong amplification cycle was also recently described in mollusks, where two different PIWI proteins bind sense and antisense transcripts, respectively [65]. These two proteins bind piRNAs of 29-30 nt (with 1U bias) and 24-25 nt (with 10A bias), respectively. Interestingly, homotypic ping-pong amplification was also described in the oyster *Crassostrea gigas*, involving a single PIWI-clade protein acting on both sense and antisense transcripts. As a consequence of concomitant homotypic and heterotypic ping-pong amplification in the gonads of *C. gigas*, 1U-biased piRNAs showed a size distribution of 29-30 nt, while partially complementary, 10A-biased piRNAs were either 24-25 (heterotypic) or 29-30 (homotypic) nt long.

The piRNAs of cluster #4 piRNAs in *L. squamata* (Figure 2a-e) meet this pattern, with the only difference that the 1U-biased piRNAs derived from the putative piRNA precursor transcript were slightly longer in Gastrotricha (30-32 nt vs 29-30 nt, Figure 2a-b). Mammalian PIWI proteins are also associated with piRNAs of slightly different lengths. For instance, mouse piRNAs bound to MILI or MIWI2 show mean lengths of 26 and 28 nt, respectively [66]. 1U-biased piRNAs in *L. squamata* tend to be slightly longer than in the majority of animals analyzed to date [65–67]. They are very similar to what was previously reported in the planarian *S. mediterranea* [41], where bands of 31-32 nt were identified by both sequencing and Northern blot analyses. However, an exhaustive analysis of piRNA lengths across metazoan species is still lacking,

Although we cannot exclude the existence of 21U piRNAs (as not all 21 nt reads were assigned to other classes), the presence of canonical piRNAs and their genomic distribution in piRNA clusters suggest that piRNA processing patterns in *L. squamata* is different than in nematodes, but, instead, very similar to what was recently reported for mollusks [65] (a spiralian) and comparable in size to the piRNA complement in flatworms [41,44]. Accordingly, the piRNA data indicate that Gastrotricha (*L. squamata*) are more closely related to Platyhelminthes within Spiralia rather than to Nematoda within Ecdysozoa.

A strict homology approach to the protein repertoires of the RNAi and PIWI protein machinery across protostomes (i.e., spiralians and ecdysozoans) corroborated that nematodes (including *C. elegans*) have a more derived RNAi pathway [51] compared with other spiralians. Among spiralians, we found that RNAi proteins from gastrotrichs (including *L. squamata*) and flatworms are very similar in terms of

bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

RNAi protein phylogenetic distribution and domain architecture, supporting their closer relationship and the Rouphozoa phylogenetic clade [20,21]. On the other hand, our analyses on PIWI protein repertoires in *L. squamata* revealed all necessary protein machinery for piRNA biogenesis [24]. In addition, the presence of several PIWI-like proteins, some of them - at some extend - with similarity to *D. melanogaster Aub* and *AGO3*, suggest that piRNAs might be controlled by a system similar what has been observed in Arthropoda. Although further experimental evidence is needed to support this hypothesis, we demonstrated that *L. squamata* PIWI repertoire is not similar to nematodes studied so far. Taken together, all data inferences covering reciprocal Blast, functional domain annotation, and phylogenetic analysis, strongly support the close phylogenetic relationship between Gastrotricha and Platyhelminthes, which is congruent with miRNA and piRNA results.

In conclusion, we show that the freshwater gastrotrich *Lepidodermella squamata* (Chaetonotida) expresses a comprehensive set of *bona fide* miRNAs and piRNAs and the corresponding biogenesis and RNAi machinery. Furthermore, we demonstrate that the occurrence, structure and sequence of miRNA genes and families is useful to assess the phylogenetic position of gastrotrichs and strongly supports close relatedness to flatworms and not to nematodes. Congruently, by comparing the length and structure of mature and precursor piRNAs, we show that the presence of *bona fide* piRNAs and *responder piRNA* in particular, clearly supports a closer relationship to flatworms, as well. This is consistently followed by the presence / absence distribution and domain architecture of RNAi proteins supporting this Rouphozoa clade, too.

With miRNA and piRNA complements for only one species we certainly cannot make statements about the monophyly of gastrotrichs [20]. Nevertheless, these first small non-coding RNA complements of *L. squamata* add important information for an understanding of the evolution of non-coding RNAs across the metazoan tree of life. We showed that not only the highly conserved miRNAs, but also the piRNAs and the corresponding protein machinery can add valuable phylogenetic information in support of a close relationship of *L. squamata* to flatworms.

Our knowledge on the evolution of ncRNAs is broad but not very detailed because the annotation of *bona fide* ncRNA complements, especially for microscopic organisms, remains a challenge as it requires material from often very sparse samples, expert knowledge, and manual curation. Surprisingly, even major endeavours such as the International Helminth Genomes Consortium have neglected ncRNA analyses despite the obvious availability of high quality genome assemblies [68]. In the absence of such assemblies, our pipeline miRCandRef can be used for miRNA and piRNA annotations that can be complemented by information of key RNAi and biogenesis proteins through readily available transcriptome assemblies for all major metazoan groups toward a better understanding on the evolution of coding and non-coding elements.

bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

## 4. Materials and Methods

### *4.1 Lepidodermella squamata culture*

500 adult specimens of *L. squamata* (Dujardin 1841) [69] were isolated from a purchased culture from Connecticut Valley Biological Supply Co., Inc., Southampton, Mass. Animals were maintained on wheat grains that contained a variety of bacteria, protozoa, and fungi, but no other metazoans.

### *4.2 small-RNA sequencing*

A total RNA sample enriched for small RNA was extracted from 500 pooled individuals using the ZR RNA MicroPrep kit (Zymo Research). miRNA library construction service from total RNA using Illumina TrueSeq Small RNA kit and sequencing on a NextSeq 500 Mid instrument was outsourced to StarSEQ GmbH, Mainz, Germany.

### *4.3 small-RNA bioinformatics*

miRTrace [70] was used to process the raw small RNA-Seq reads. The algorithm miRCandRef (https://hyperbrowser.uio.no/mircandref/) [39] was used for assembling high-quality genomic reads into relatively short genomic contigs (crystal-contigs) that constitute candidate small-RNA loci. Briefly, the pipeline delivers a multifasta file of assembled and clustered crystal-contigs that can be used as a reference for piRNA and miRNA prediction with software like proTRAC [71] and MirMiner [30] or miRDeep2 [72] for miRNAs. To distinguish between missing and not expressed miRNAs, the contigs were additionally screened for absent miRNA families by BLAT [73] as well as for spiralian miRNA families from MirGeneDB 2.0 [35,74].

Putative piRNAs were identified by mapping small RNA-Seq reads to the *L. squamata* genome assembly (Aguilera et al; in prep) with sRNAmapper [75]. By prioritizing specificity over sensitivity to avoid confounding between piRNAs and other ncRNAs or their fragments [76], we removed reads mapping to either miRNA precursors or predicted tRNA and rRNA reference (with a mismatch allowance of up to three nucleotides) obtained with tRNAScan-SE [77] and RNAmmer [78], respectively. Sequences mapped to miRNA genes were also removed. proTRAC [71] was then used to identify non-annotated sequences showing 1T and/or 10A bias (>75% of the sequences), bias towards a piRNA-like size-range (> 75% of the sequences being 24-32 nt) and forming clusters of at least 500 bp with a minimum of 20 hits per cluster and an even distribution in sequence abundance (so that the top 10% of reads account for less than 75% of the reads in the cluster).

bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

*4.4 Identification and domain architecture of RNAi proteins*

We obtained available genome-derived proteomes from a wide range of metazoans focusing on spiralian species (supplementary table S4). For *L. squamata* and other gastrotrich species that predicted proteomes are not available, raw reads were downloaded from NCBI SRA database (supplementary table S5), quality trimmed using Trimmomatic v0.36 [79] (default settings) and assembled using Trinity v2.1.1 [80] (with the -normalized_reads option set for *L. squamata* and *D. aspetos*, otherwise default settings), Trans-ABySS v1.5.3 [81] (default settings), and Bridger v2014-12-01 [82] (default settings). Transcriptome assemblies were combined and redundant transcripts were clustered using CD-HIT v4.6 and TransDecoder v3.0.1 following the strategy described by [83], and the longest ORFs were selected for further analyses.

To search for the presence of RNAi proteins across metazoan species, we downloaded the *C. elegans* proteome from WormBase (www.wormbase.org; release WS267) and a total of 65 proteins involved in small non-coding RNA pathways [44] was used as queries for BLASTp searching with an E-value threshold of $10^{-5}$ [84]. Positive hits were then verified by reciprocal Blast searches against the *C. elegans* proteome (using the same E-value threshold), and only proteins ranked as top hit by both approaches were retained. Those proteins were further verified by Blast search against GenBank nr database.

Proteins retrieved as Dicer, Drosha, Pasha, and Exportin from *L. squamata* and some representatives from Annelida, Platyhelminthes, Rotifera, and Nematoda were thereafter classified and annotated by using Pfam database [85]. Sequence comparisons were performed by visual inspection and all domain architecture figures were created using Adobe Illustrator CC 2015.

*4.5 Identification and phylogenetic analysis of PIWI proteins*

To reconstruct the phylogenetic relationship of gastrotrich PIWI-like proteins, we searched for PIWI genes in different protostome genomes and gastrotrich transcriptomes (supplementary table S6). To this end, we looked for annotated PIWI genes of the nematode *C. elegans* (D2030.6 and C01G5.2a, release WS244), Pacific oyster *Crassostrea gigas* (EKC35279 and EKC29295), fly *D. melanogaster* (AGL81535, AGA18946, and NP_001163498), and human *Homo sapiens* (NP_004755, NP_060538, NP_001008496, and NP_689644) by using Blastp with a cut-off of $10^{-5}$. Those proteins with positive hits were selected and checked for presence of PAZ and Piwi domains using Pfam database [85].

The predicted and annotated PIWI protein sequences from Gastrotricha, Platyhelminthes, Phoronida, Annelida, Nemertea, Orthonectida, together with PIWI paralogs of nematode, fly and human, were aligned using MAFFT v7.407 [86] with the E-INS-i algorithm and the BLOSUM45 scoring matrix. Non-conserved regions were automatically removed with TrimAl v1.4 [87], and maximum likelihood

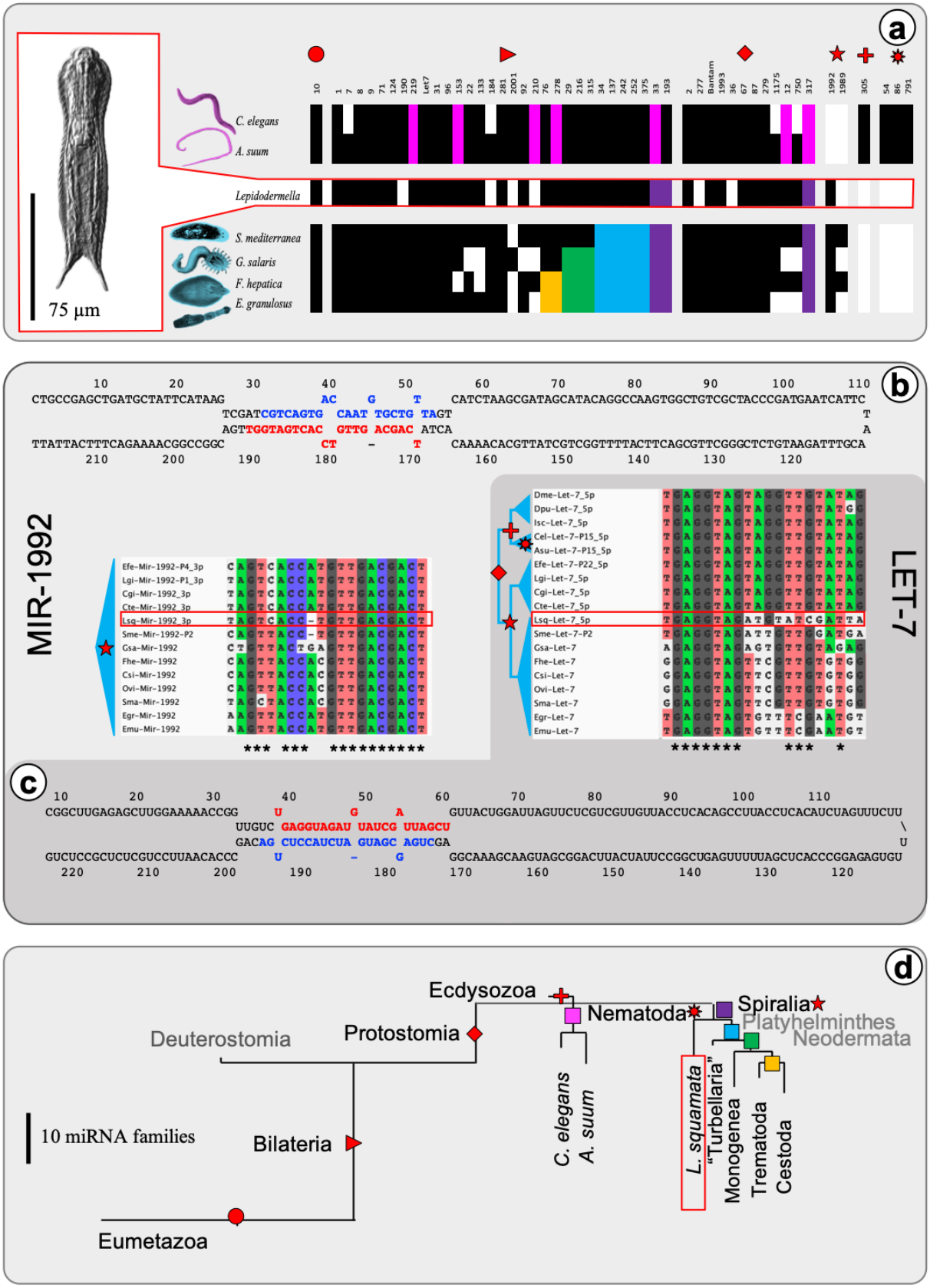
bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(ML) analysis was run using RAxML v8.2.4 [88] with PROTGAMMAAUTO and autoMRE options. Support values were generated by bootstrap with 300 replicates and phylogenetic tree was visualized and edited using the Interactive Tree Of Life web-based tool [89].

bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

## Figures

bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 1. The miRNA complement of *L. squamata* shows a strong affinity to flatworms. (**a**) miRNA distribution in nematodes (top), *L. squamata* (middle and red; depicted in red box to the right) and flatworms (bottom) show pattern of taxa-specific gains and losses of conserved miRNA families for each group. Presence/absence pattern of taxa-specific groups of conserved miRNAs are depicted in red icons (Eumetazoa, Bilateria, Protostomia, and Spiralia; compare to (**d**), black squares represent the presence of a miRNA family. Losses are depicted as white fields or highlighted in colors when interpreted as shared losses (pink: Nematoda, purple: Spiralia, blue: Platyhelminthes, green: Neodermata, yellow: Trematoda+Cestoda). (**b**) (top) Structure of extended precursor of Lsq-Mir-1992, a spiralian-specific miRNA family, including mature (red) and star (blue) annotation and (bottom left) sequence alignment with all known members in Spiralia (3′ - 5′ direction). (**c**) (bottom) Structure of extended precursor of Lsq-Let-7, a bilaterian wide distributed miRNA family, including mature (red) and star (blue) annotation and (top right) sequence alignment with selected invertebrate members. (**d**) gain and loss of miRNA families in Eumetazoa, Bilateria, Protostomia, Nematoda, Platyhelminthes and *L. squamata* branch lengths corresponds to number of gains/losses.

bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

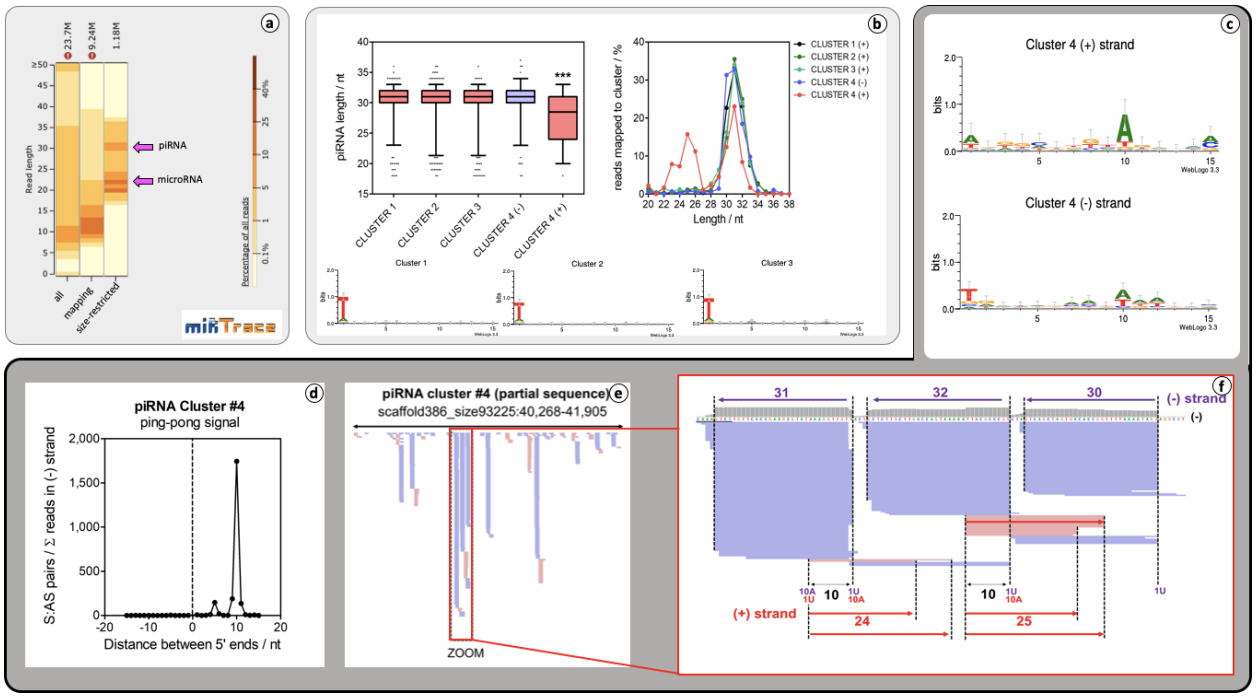



Figure 2: piRNA discovery in *L. squamata.* (**a**) miRTrace read-length distribution of all, genome mapped and genome mapped plus size-selected reads (20-37 nt) show that distinct peaks for *bona fide* miRNA and genuine piRNAs are hidden by a body of non small-RNA fragments. (**b**) top left: size distribution of piRNAs assigned to each cluster, discriminating between piRNAs derived from the plus (red) and minus (purple) strand of piRNA cluster #4. Whiskers represent the 5-95 percentile. ***: $p < 0.001$; Kruskal-Wallis test. (**b**) top right: The percentage of reads of a specified length mapping to the corresponding piRNA clusters. (**b**) bottom: Sequence logo representation based on entropy for reads mapping to representative piRNA clusters: #1, #2 and #3 shows strong 1T bias. Dark grey: Cluster #4. (**c**) Sequence Logo representation of antisense (top) and sense (bottom) mapping reads on cluster #4: (**d**) ping-pong signature (overlap of exactly 10 nucleotides counting from the 5′ side for piRNAs mapping opposite strands from a same cluster). (**e**) Genome browser representation of sequencing reads mapping piRNA cluster #4 (region). Reads corresponding to the minus and plus strands are represented in purple and red, respectively. (**f**) detailed view of a small region consistent with phased piRNA production from a minus-strand derived transcript, and responder piRNAs produced from antisense transcripts by a presumable ping-pong amplification cycle.

bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

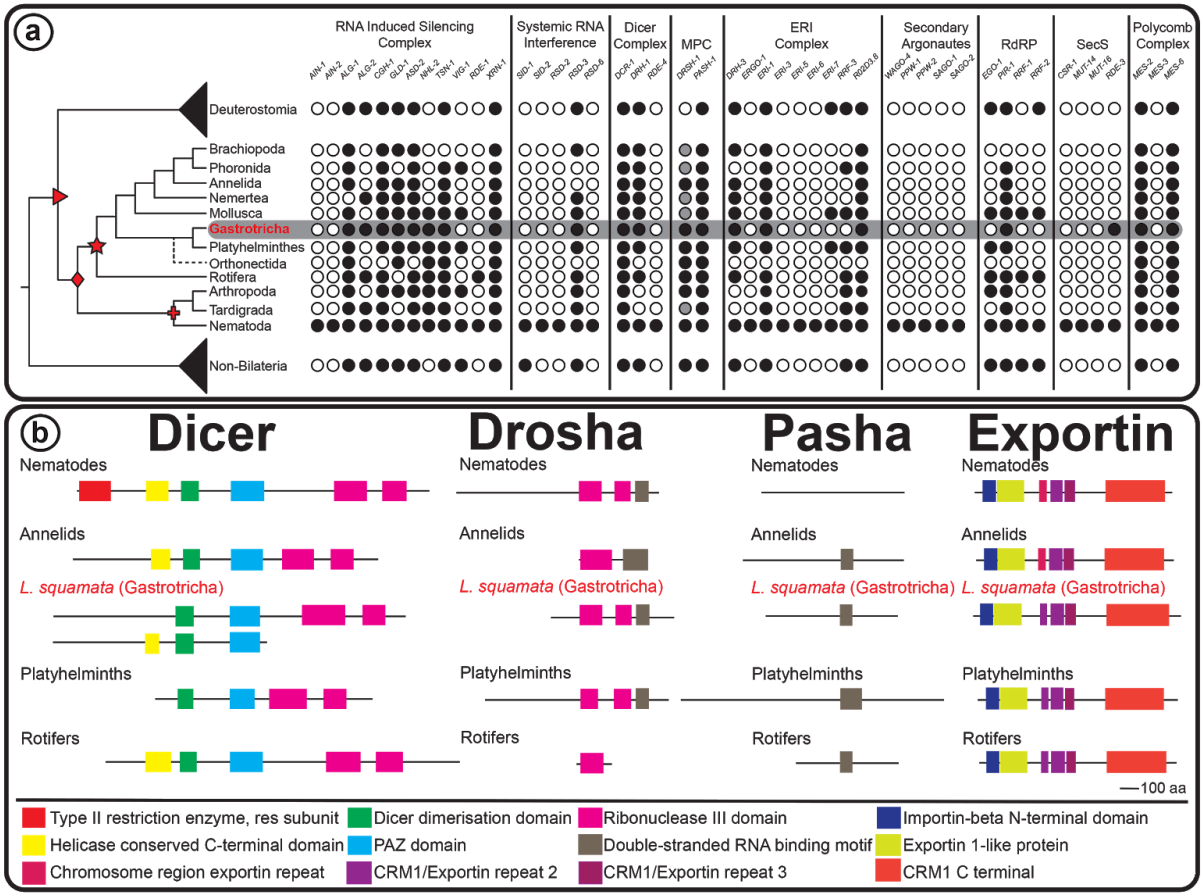



Figure 3: miRNA & piRNA biogenesis proteins and RNAi machinery. (**a**) distribution of RNAi proteins across animals. The cladogram to the left indicates the currently accepted phylogenetic relationships among animal taxa [20,21,90–92], and drawings on particular nodes are according to Figure 1d. Black circles denote the presence of proteins from different small RNA complexes, while white circles illustrates protein absence, and grey circles represent an deduced, but not actually detected presence of Drosha based on the presence of *bona fide* miRNAs. A protein is categorized as present whether we could find it in at least one species from each phylogenetic group. See supplementary table S2 for detailed description of presence/absence of RNAi proteins). Gene names at the top are based on *C. elegans* gene nomenclature. Abbreviations: MPC = Microprocessor Complex; RdRP = RNA dependent RNA polymerases; and SecS = Secondary RNAi Silencing. (**b**) Domain composition of miRNA proteins among selected spiralan and nematode species. The domain composition of Dicer, Drosha, Pasha, and Exportin discovered by reciprocal Blast searches were annotated using Pfam database [85]. All proteins are from genome-derived predicted proteins, except for *L. squamata*. Hence, some proteins might be not full-length due to incomplete assembly.

bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

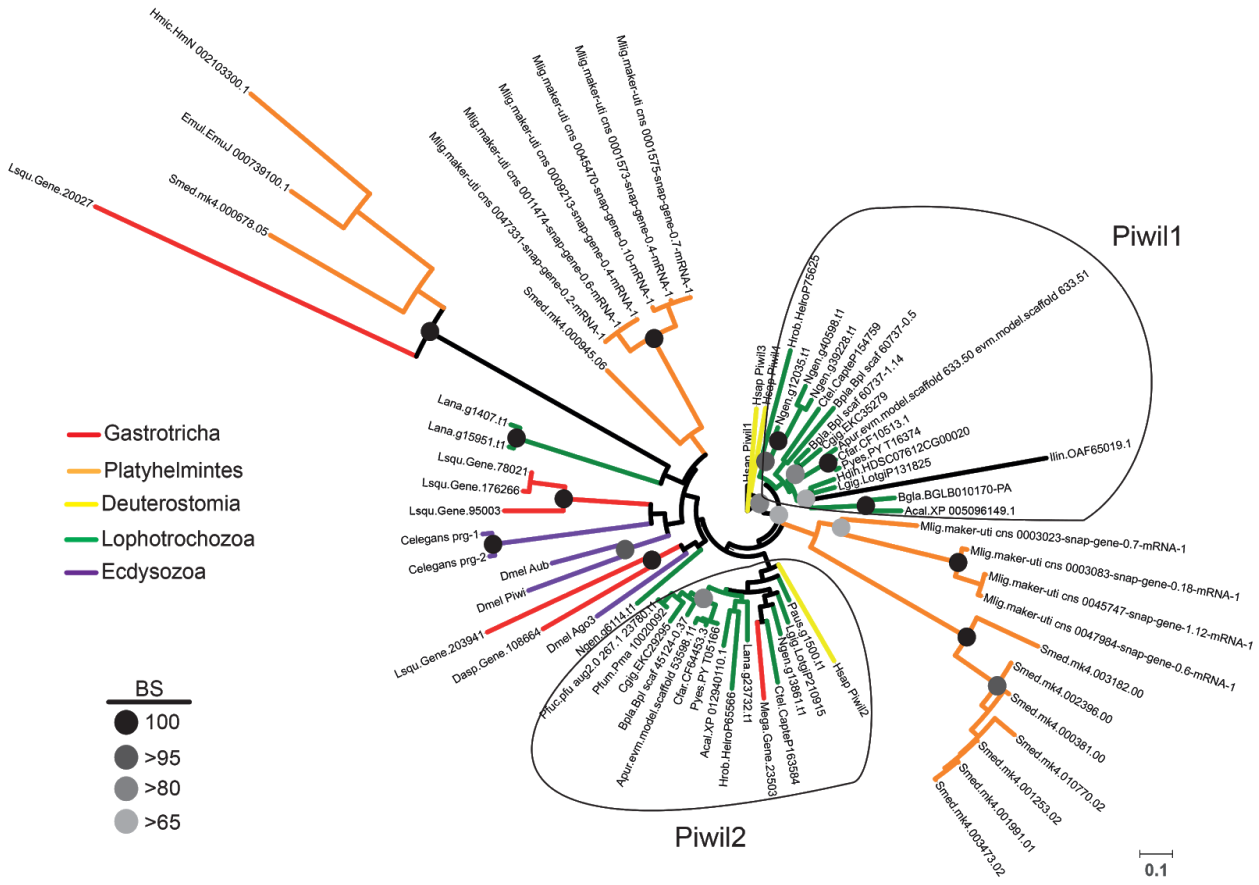



Figure 4. Phylogenetic analysis of PIWI-like proteins. Maximum likelihood analysis of protostome PIWI-like proteins with gastrotrich counterparts highlighted in red. Bootstrap values of particular clades are represented as black-to-grey circles as shown on the lower left. The scale bar on the lower right indicates amino acid substitution rate per site. Color coding on tree branches is shown as follow: Red = gastrotrichs, orange = platyhelminths, yellow = deuterostomes, specifically human PIWI proteins, green = lophotrochozoans (comprising annelids, mollusks, nemerteans, brachiopods, and phoronids), and violet = ecdysozoans (*C. elegans* and *D. melanogaster*).

bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

**Supplementary Material**

**Supplementary File S1**. The miRNA complement of *Lepidodermella squamata* (fasta).

**Supplementary Table S1.** Curated set of 65 RNA proteins from *C. elegans* used for reciprocal Blast searches.

**Supplementary Table S2.** RNAi proteins identified in metazoan predicted proteins and gastrotrich transcriptomes.

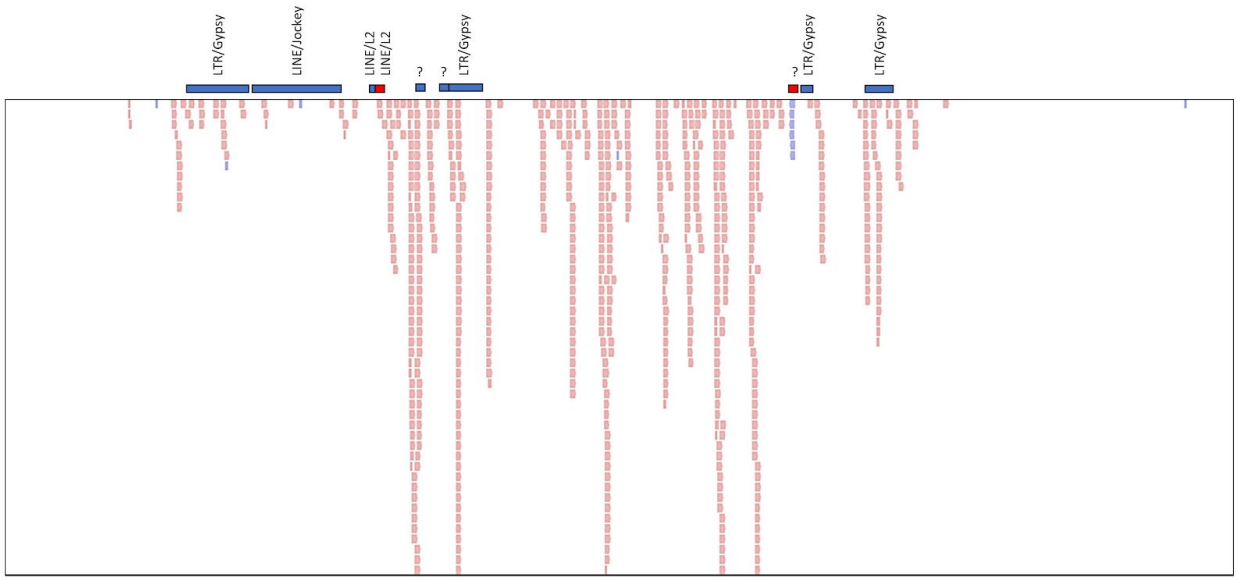
**Supplementary Table S3.** PIWI-like proteins identified in metazoan predicted proteins and gastrotrich transcriptomes.

**Supplementary Table S4.** List of metazoan genomes used for RNAi protein survey.

**Supplementary Table S5.** List of gastrotrich transcriptomes used for RNAi protein survey.

**Supplementary Table S6.** List of protostome genomes and gastrotrich transcriptomes used for PIWI-like phylogenetic analysis.

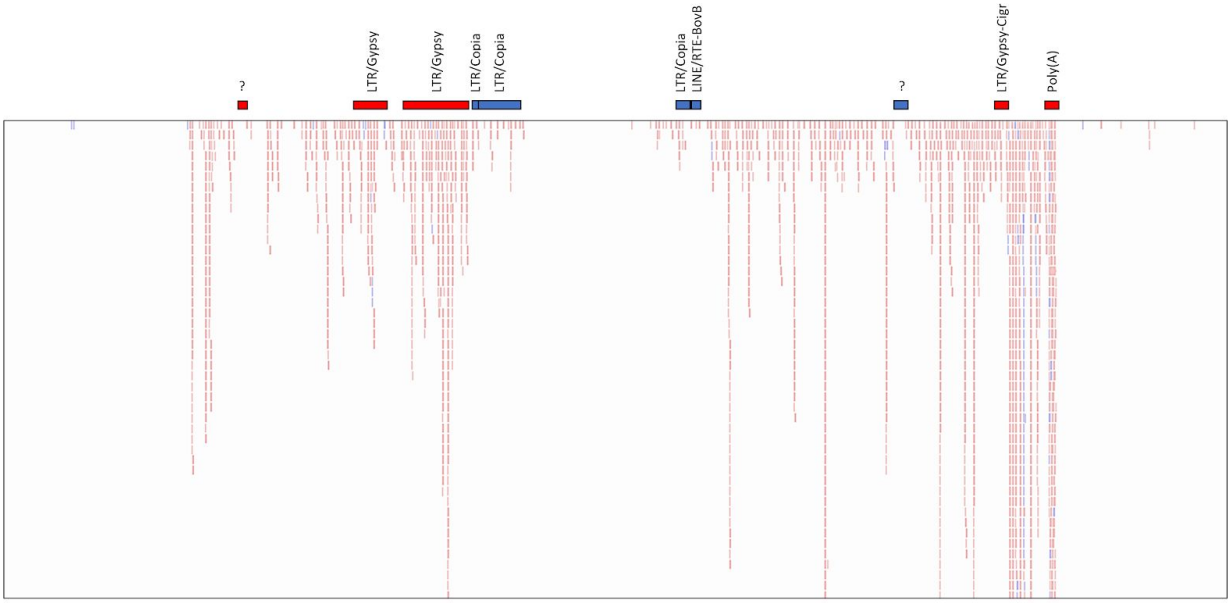
bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.




piRNA cluster #1: scaffold95_size236010:16,013-20,557

**Supplementary Figure S1:** Overlap between sequencing reads corresponding to piRNA cluster #1 and identified repeats in the same genomic regions. Pink (reads) or red (repeats) correspond to the plus strand. Purple (reads) or blue (repeats) correspond to the minus strand. ?: repeated elements of unknown origin.

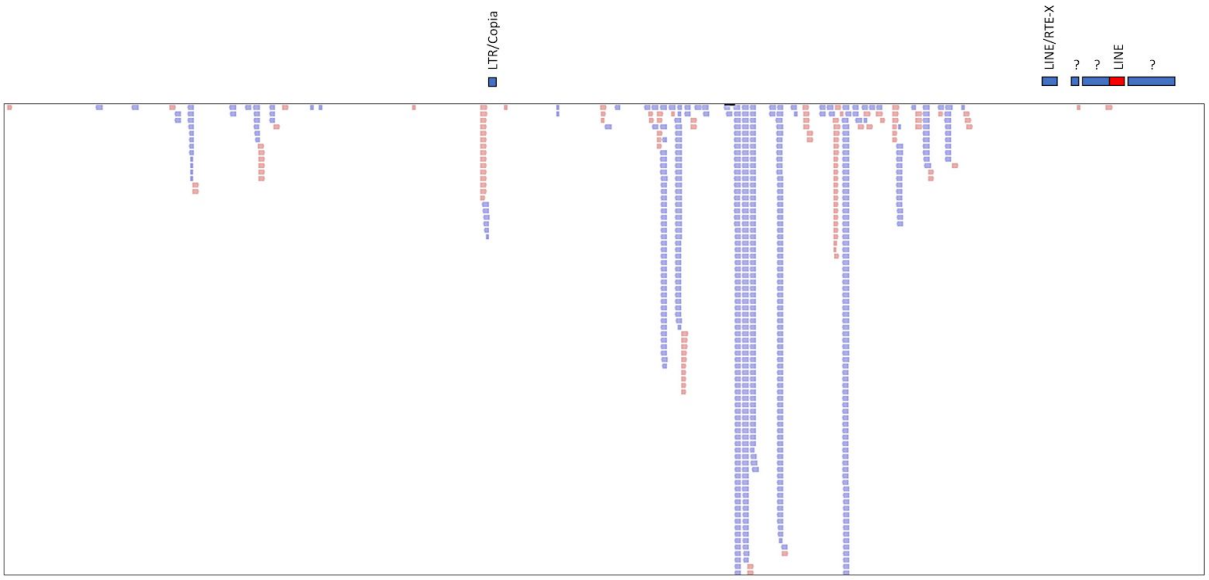
bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.




piRNA cluster #2: scaffold171_size172051:144,160-158,000

**Supplementary Figure S2:** Overlap between sequencing reads corresponding to piRNA cluster #2 and identified repeats in the same genomic regions. Pink (reads) or red (repeats) correspond to the plus strand. Purple (reads) or blue (repeats) correspond to the minus strand. ?: repeated elements of unknown origin.

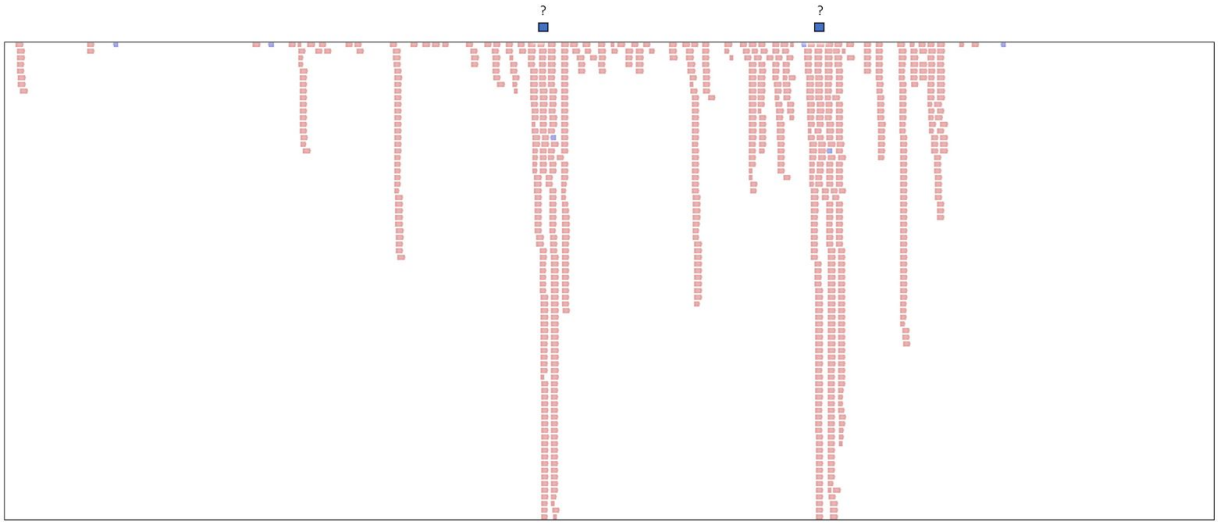
bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.




piRNA cluster #4: scaffold386_size93225:37,926-42,144

**Supplementary Figure S3:** Overlap between sequencing reads corresponding to piRNA cluster #3 and identified repeats in the same genomic regions. Pink (reads) or red (repeats) correspond to the plus strand. Purple (reads) or blue (repeats) correspond to the minus strand. ?: repeated elements of unknown origin.

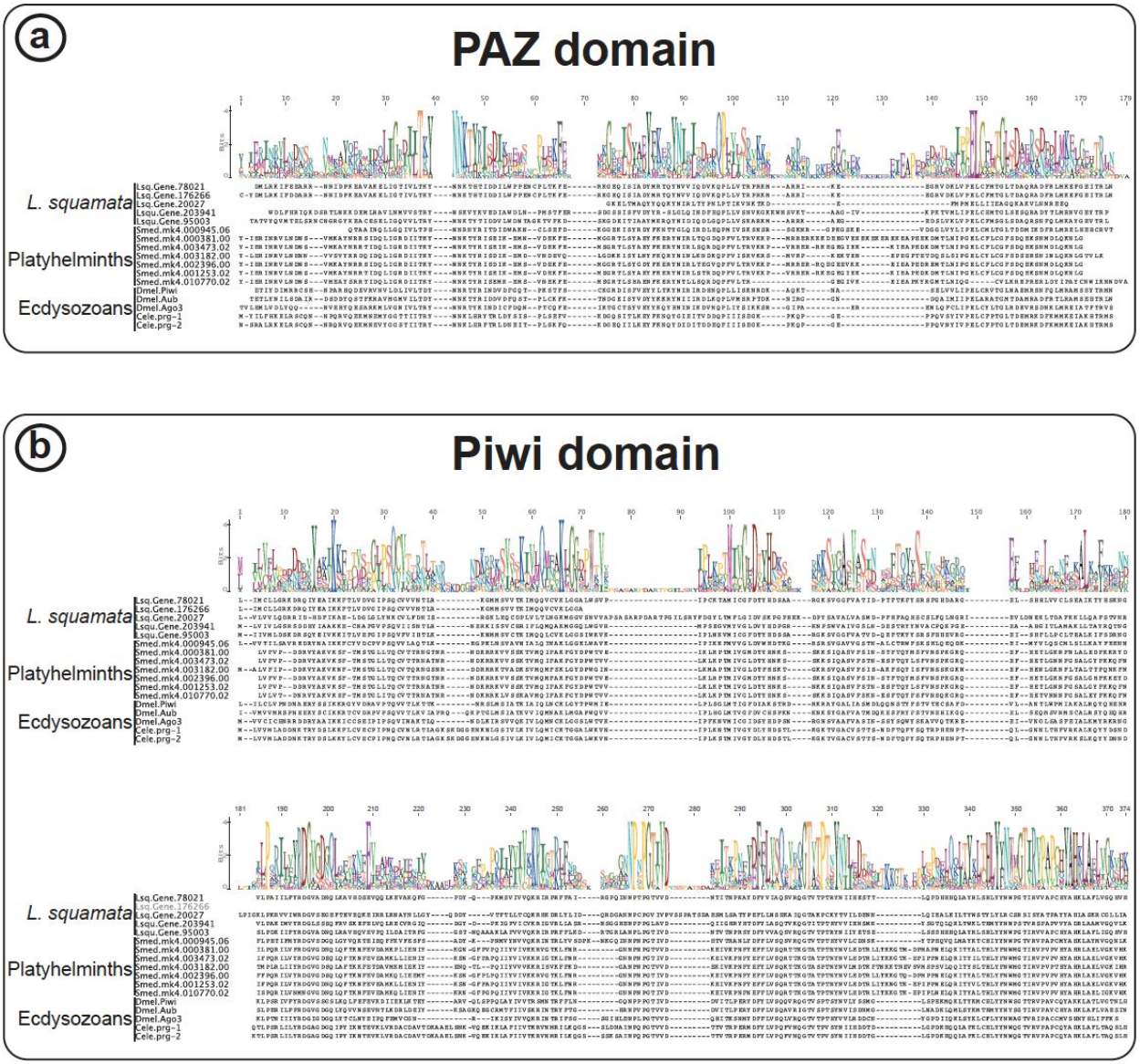
bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.




piRNA cluster #3: scaffold367_size99769: 90,134-93,976

**Supplementary Figure S4:** Overlap between sequencing reads corresponding to piRNA cluster #4 and identified repeats in the same genomic regions. Pink (reads) or red (repeats) correspond to the plus strand. Purple (reads) or blue (repeats) correspond to the minus strand. ?: repeated elements of unknown origin.

bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.




**Supplementary Figure S5.** Alignment of representative PIWI-like protein sequences. (**a**) Comparative alignment of PAZ domain among *L. squamata* and representative species from Platyhelminthes and Ecdysozoa. (**b**) Comparative alignment of Piwi domain among *L. squamata* and representative species from Platyhelminthes and Edysozoa.

bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

**Author Contributions:** Conceptualization, B.F. ; Methodology, B.F., J.P.T, F.A.; Software, B.F., J.P.T, F.A.; Data Curation, B.F., J.P.T., F.A.; Writing-Original Draft Preparation, B.F., J.P.T., F.A.; Writing-Review & Editing, all authors. Visualization, B.F., J.P.T., F.A.; Supervision, A.H..; Project Administration, B.F.; Funding Acquisition, M.R.F., L.B. and A.H.

**Funding:** B.F. and M.R.F acknowledge funding from the Strategic Research Area (SFO) program of the Swedish Research Council (VR) through Stockholm University. F.A. is supported by a CONICYT-PAI Inserción en la Academia grant (79170033) from the Chilean government. A.H. is supported by the European Research Council Community's Framework Program Horizon 2020 (2014–2020) ERC grant agreement 64886.

**Acknowledgments:** The authors thank Julian P.S. Smith III and Kevin J. Peterson for useful comments.

**Conflicts of Interest:** "The authors declare no conflict of interest."

bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

## References


1. Todaro, M.A.; Hummon, W.D.; Others An overview and a dichotomous key to genera of the phylum Gastrotricha. *Meiofauna Marina* **2008**, *16*, 3–20.
2. Balsamo, M.; d'Hondt, J.L.; Kisielewski, J.; Todaro, M.A.; Tongiorgi, P.; Guidi, L.; Grilli, P.; de Jong, Y. Fauna Europea: Gastrotricha. Biodiversity Data Journal 3: e5800 2015.
3. Paps, J.; Riutort, M. Molecular phylogeny of the phylum Gastrotricha: new data brings together molecules and morphology. *Mol. Phylogenet. Evol.* **2012**, *63*, 208–212.
4. Todaro, M.A.; Telford, M.J.; Lockyer, A.E.; Littlewood, D.T.J. Interrelationships of the Gastrotricha and their place among the Metazoa inferred from 18S rRNA genes. *Zool. Scr.* **2006**, *35*, 251–259.
5. Hejnol, A. Gastrotricha. In *Evolutionary Developmental Biology of Invertebrates 2: Lophotrochozoa (Spiralia)*; Wanninger, A., Ed.; Springer Vienna: Vienna, 2015; pp. 13–19 ISBN 9783709118719.
6. Bekkouche, N.; Worsaae, K. Neuromuscular study of early branching Diuronotus aspetos (Paucitubulatina) yields insights into the evolution of organs systems in Gastrotricha. *Zoological Lett* **2016**, *2*, 21.
7. Todaro, M.A.; Dal Zotto, M.; Leasi, F. An Integrated Morphological and Molecular Approach to the Description and Systematisation of a Novel Genus and Species of Macrodasyida (Gastrotricha). *PLoS One* **2015**, *10*, e0130278.
8. Rothe, B.H.; Kieneke, A.; Schmidt-Rhaesa, A. The nervous system of Xenotrichula intermedia and X. velox (Gastrotricha: Paucitubulatina) by means of immunohistochemistry (IHC) and TEM. *Meiofauna Marina* **2011**, *19*, 71–88.
9. Rothe, B.H.; Schmidt-Rhaesa, A.; Kieneke, A. The nervous system of Neodasys chaetonotoideus (Gastrotricha: Neodasys) revealed by combining confocal laserscanning and transmission electron microscopy: evolutionary comparison of neuroanatomy within the Gastrotricha and basal Protostomia. *Zoomorphology* **2011**, *130*, 51–84.
10. Strayer, D.L.; Hummon, W.D.; Hochberg, R. Chapter 7 - Gastrotricha. In *Ecology and Classification of North American Freshwater Invertebrates (Third Edition)*; Thorp, J.H., Covich, A.P., Eds.; Academic Press: San Diego, 2010; pp. 163–172 ISBN 9780123748553.
11. Zrzavy, J. Gastrotricha and metazoan phylogeny. *Zool. Scr.* **2003**, *32*, 61–81.
12. Hyman, L.H. *The Invertebrates: Acanthocephala, Aschelminthes And Entoprocta: The Pseudocoelomate Bilateria (Volume-3)*; Mcgraw-Hill Book Company; New York, Toronto, London, 1951;.
13. Sørensen, M.V. An SEM study of the jaws of Haplognathia rosea and Rastrognathia macrostoma (Gnathostomulida), with a preliminary comparison with the rotiferan trophi. *Acta Zool.* **2000**, *81*, 9–16.
14. Nielsen, C. *Animal Evolution : Interrelationships of the Living Phyla: Interrelationships of the Living Phyla*; OUP Oxford, 2001; ISBN 9780191588525.
15. Rieger, R.M. Monociliated epidermal cells in Gastrotricha: significance for concepts of early metazoan evolution. *J. Zoolog. Syst. Evol. Res.* **1976**, *14*, 198–226.
16. Zrzavy, J.; Mihulka, S.; Kepka, P.; Bezdek, A.; Tietz, D. Phylogeny of the Metazoa Based on Morphological and 18S Ribosomal DNA Evidence. *Cladistics* **1998**, *14*, 249–285.
17. Cavalier-Smith, T. A revised six-kingdom system of life. *Biol. Rev. Camb. Philos. Soc.* **1998**, *73*, 203–266.
18. Giribet, G. Assembling the lophotrochozoan (=spiralian) tree of life. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2008**, *363*, 1513–1522.
19. Edgecombe, G.; Giribet, G.; Dunn, C.; Hejnol, A.; Kristensen, R.; Neves, R.; Rouse, G.; Worsaae, K.; Sørensen, M. Higher-level metazoan relationships: recent progress and remaining questions. *Org. Divers. Evol.* **2011**, *11*, 151–172.
20. Struck, T.H.; Wey-Fabrizius, A.R.; Golombek, A.; Hering, L.; Weigert, A.; Bleidorn, C.; Klebow, S.; Iakovenko, N.; Hausdorf, B.; Petersen, M.; et al. Platyzoan paraphyly based on phylogenomic data supports a noncoelomate ancestry of spiralia. *Mol. Biol. Evol.* **2014**, *31*, 1833–1849.
21. Laumer, C.E.; Bekkouche, N.; Kerbl, A.; Goetz, F.; Neves, R.C.; Sørensen, M.V.; Kristensen, R.M.; Hejnol, A.; Dunn, C.W.; Giribet, G.; et al. Spiralian Phylogeny Informs the Evolution of Microscopic Lineages. *Curr. Biol.* **2015**, *25*, 2000–2006.

bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.


22. Doe, D.A. Comparative ultrastructure of the pharynx simplex in turbellaria. *Zoomorphology* **1981**, *97*, 133–193.
23. Kieneke, A.; Riemann, O.; Ahlrichs, W.H. Novel implications for the basal internal relationships of Gastrotricha revealed by an analysis of morphological characters. *Zool. Scr.* **2008**, *37*, 429–460.
24. Moran, Y.; Agron, M.; Praher, D.; Technau, U. The evolutionary origin of plant and animal microRNAs. *Nat Ecol Evol* **2017**, *1*, 27.
25. Iwasaki, Y.W.; Siomi, M.C.; Siomi, H. PIWI-Interacting RNA: Its Biogenesis and Functions. *Annu. Rev. Biochem.* **2015**, *84*, 405–433.
26. Bartel, D.P. Metazoan MicroRNAs. *Cell* **2018**, *173*, 20–51.
27. Deline, B.; Greenwood, J.M.; Clark, J.W.; Puttick, M.N.; Peterson, K.J.; Donoghue, P.C.J. Evolution of metazoan morphological disparity. *Proc. Natl. Acad. Sci. U. S. A.* **2018**.
28. Halushka, M.K.; Fromm, B.; Peterson, K.J.; McCall, M.N. Big Strides in Cellular MicroRNA Expression. *Trends Genet.* **2018**, *34*, 165–167.
29. Christodoulou, F.; Raible, F.; Tomer, R.; Simakov, O.; Trachana, K.; Klaus, S.; Snyman, H.; Hannon, G.J.; Bork, P.; Arendt, D. Ancient animal microRNAs and the evolution of tissue identity. *Nature* **2010**, *463*, 1084–1088.
30. Wheeler, B.M.; Heimberg, A.M.; Moy, V.N.; Sperling, E.A.; Holstein, T.W.; Heber, S.; Peterson, K.J. The deep evolution of metazoan microRNAs. *Evol. Dev.* **2009**, *11*, 50–68.
31. Tarver, J.E.; Sperling, E.A.; Nailor, A.; Heimberg, A.M.; Robinson, J.M.; King, B.L.; Pisani, D.; Donoghue, P.C.J.; Peterson, K.J. miRNAs: small genes with big potential in metazoan phylogenetics. *Mol. Biol. Evol.* **2013**, *30*, 2369–2382.
32. Fromm, B.; Burow, S.; Hahn, C.; Bachmann, L. MicroRNA loci support conspecificity of Gyrodactylus salaris and Gyrodactylus thymalli (Platyhelminthes: Monogenea). *Int. J. Parasitol.* **2014**, *44*, 787–793.
33. Kenny, N.J.; Sin, Y.W.; Hayward, A.; Paps, J.; Chu, K.H.; Hui, J.H.L. The phylogenetic utility and functional constraint of microRNA flanking sequences. *Proc. Biol. Sci.* **2015**, *282*, 20142983.
34. Helm, C.; Bernhart, S.H.; Höner zu Siederdissen, C.; Nickel, B.; Bleidorn, C. Deep sequencing of small RNAs confirms an annelid affinity of Myzostomida. *Mol. Phylogenet. Evol.* **2012**, *64*, 198–203.
35. Fromm, B.; Billipp, T.; Peck, L.E.; Johansen, M.; Tarver, J.E.; King, B.L.; Newcomb, J.M.; Sempere, L.F.; Flatmark, K.; Hovig, E.; et al. A Uniform System for the Annotation of Vertebrate microRNA Genes and the Evolution of the Human microRNAome. *Annu. Rev. Genet.* **2015**, *49*, 213–242.
36. Tarver, J.E.; Taylor, R.S.; Puttick, M.N.; Lloyd, G.T.; Pett, W.; Fromm, B.; Schirrmeister, B.E.; Pisani, D.; Peterson, K.J.; Donoghue, P.C.J. Well-Annotated microRNAomes Do Not Evidence Pervasive miRNA Loss. *Genome Biol. Evol.* **2018**, *10*, 1457–1470.
37. Thomson, R.C.; Plachetzki, D.C.; Mahler, D.L.; Moore, B.R. A critical appraisal of the use of microRNA data in phylogenetics. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, E3659–68.
38. Bai, Y.; Zhang, Z.; Jin, L.; Kang, H.; Zhu, Y.; Zhang, L.; Li, X.; Ma, F.; Zhao, L.; Shi, B.; et al. Genome-wide sequencing of small RNAs reveals a tissue-specific loss of conserved microRNA families in Echinococcus granulosus. *BMC Genomics* **2014**, *15*, 736.
39. Fromm, B.; Worren, M.M.; Hahn, C.; Hovig, E.; Bachmann, L. Substantial loss of conserved and gain of novel MicroRNA families in flatworms. *Mol. Biol. Evol.* **2013**, *30*, 2619–2628.
40. Philippe, H.; Brinkmann, H.; Copley, R.R.; Moroz, L.L.; Nakano, H.; Poustka, A.J.; Wallberg, A.; Peterson, K.J.; Telford, M.J. Acoelomorph flatworms are deuterostomes related to Xenoturbella. *Nature* **2011**, *470*, 255–258.
41. Friedlaender, M.R.; Adamidi, C.; Ting, H.; Lebedeva, S.; Isenbarger, T.A.; Hirst, M.; Marra, M.; Nusbaum, C.; Lee, W.L.; Jenkin, J.C.; et al. High-resolution profiling and discovery of planarian small RNAs. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 11546–11551.
42. Tsai, I.J.; Zarowiecki, M.; Holroyd, N.; Garciarrubio, A.; Sanchez-Flores, A.; Brooks, K.L.; Tracey, A.; Bobes, R.J.; Fragoso, G.; Sciutto, E.; et al. The genomes of four tapeworm species reveal adaptations to parasitism. *Nature* **2013**, *496*, 57–63.
43. Hahn, C.; Fromm, B.; Bachmann, L. Comparative genomics of flatworms (platyhelminthes) reveals shared genomic features of ecto- and endoparastic neodermata. *Genome Biol. Evol.* **2014**, *6*, 1105–1117.
44. Fontenla, S.; Rinaldi, G.; Smircich, P.; Tort, J.F. Conservation and diversification of small RNA pathways within flatworms. *BMC Evol. Biol.* **2017**, *17*, 215.

bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

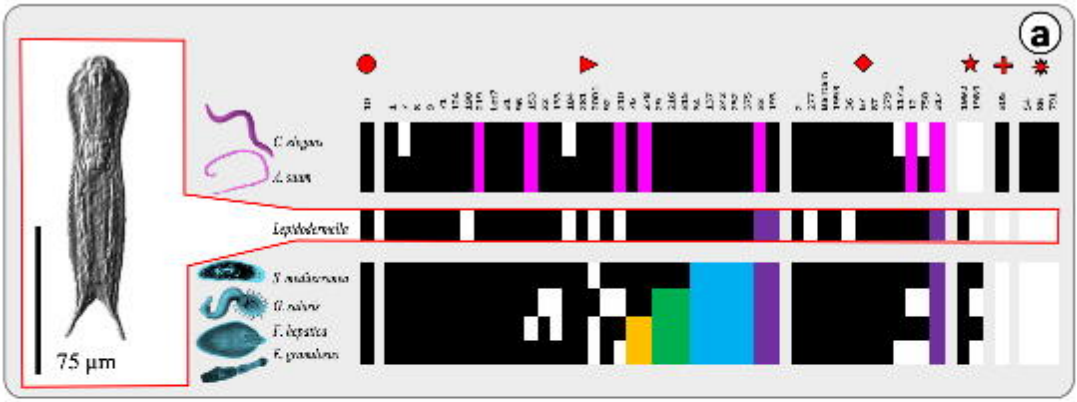

45. Stroehlein, A.J.; Young, N.D.; Korhonen, P.K.; Hall, R.S.; Jex, A.R.; Webster, B.L.; Rollinson, D.; Brindley, P.J.; Gasser, R.B. The small RNA complement of adult Schistosoma haematobium. *PLoS Negl. Trop. Dis.* **2018**, *12*, e0006535.
46. Skinner, D.E.; Rinaldi, G.; Koziol, U.; Brehm, K.; Brindley, P.J. How might flukes and tapeworms maintain genome integrity without a canonical piRNA pathway? *Trends Parasitol.* **2014**, *30*, 123–129.
47. Sarkies, P.; Selkirk, M.E.; Jones, J.T.; Blok, V.; Boothby, T.; Goldstein, B.; Hanelt, B.; Ardila-Garcia, A.; Fast, N.M.; Schiffer, P.M.; et al. Ancient and novel small RNA pathways compensate for the loss of piRNAs in multiple independent nematode lineages. *PLoS Biol.* **2015**, *13*, e1002061.
48. Ovchinnikov, V.Y.; Mordvinov, V.A.; Fromm, B. Extreme conservation of miRNA complements in opisthorchiids. *Parasitol. Int.* **2017**, *66*, 773–776.
49. Fromm, B.; Ovchinnikov, V.; Høye, E.; Bernal, D.; Hackenberg, M.; Marcilla, A. On the presence and immunoregulatory functions of extracellular microRNAs in the trematode Fasciola hepatica. *Parasite Immunol.* **2017**, *39*.
50. Macchiaroli, N.; Cucher, M.; Zarowiecki, M.; Maldonado, L.; Kamenetzky, L.; Rosenzvit, M.C. microRNA profiling in the zoonotic parasite Echinococcus canadensis using a high-throughput approach. *Parasit. Vectors* **2015**, *8*, 83.
51. Ozata, D.M.; Gainetdinov, I.; Zoch, A.; O'Carroll, D.; Zamore, P.D. PIWI-interacting RNAs: small RNAs with big functions. *Nat. Rev. Genet.* **2018**.
52. Faehnle, C.R.; Joshua-Tor, L. Argonautes confront new small RNAs. *Curr. Opin. Chem. Biol.* **2007**, *11*, 569–577.
53. Geldhof, P.; Visser, A.; Clark, D.; Saunders, G.; Britton, C.; Gilleard, J.; Berriman, M.; Knox, D. RNA interference in parasitic helminths: current situation, potential pitfalls and future prospects. *Parasitology* **2007**, *134*, 609–619.
54. Dalzell, J.J.; McVeigh, P.; Warnock, N.D.; Mitreva, M.; Bird, D.M.; Abad, P.; Fleming, C.C.; Day, T.A.; Mousley, A.; Marks, N.J.; et al. RNAi effector diversity in nematodes. *PLoS Negl. Trop. Dis.* **2011**, *5*, e1176.
55. Büssing, I.; Yang, J.-S.; Lai, E.C.; Grosshans, H. The nuclear export receptor XPO-1 supports primary miRNA processing in C. elegans and Drosophila. *EMBO J.* **2010**, *29*, 1830–1839.
56. Fontenla, S.; Dell'Oca, N.; Smircich, P.; Tort, J.F.; Siles-Lucas, M. The miRnome of Fasciola hepatica juveniles endorses the existence of a reduced set of highly divergent micro RNAs in parasitic flatworms. *Int. J. Parasitol.* **2015**, *45*, 901–913.
57. Cucher, M.; Macchiaroli, N.; Kamenetzky, L.; Maldonado, L.; Brehm, K.; Rosenzvit, M.C. High-throughput characterization of Echinococcus spp. metacestode miRNomes. *Int. J. Parasitol.* **2015**, *45*, 253–267.
58. Pérez, M.G.; Macchiaroli, N.; Lichtenstein, G.; Conti, G.; Asurmendi, S.; Milone, D.H.; Stegmayer, G.; Kamenetzky, L.; Cucher, M.; Rosenzvit, M.C. microRNA analysis of Taenia crassiceps cysticerci under praziquantel treatment and genome-wide identification of Taenia solium miRNAs. *Int. J. Parasitol.* **2017**, *47*, 643–653.
59. Girard, A.; Sachidanandam, R.; Hannon, G.J.; Carmell, M.A. A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature* **2006**, *442*, 199–202.
60. Grivna, S.T.; Beyret, E.; Wang, Z.; Lin, H. A novel class of small RNAs in mouse spermatogenic cells. *Genes Dev.* **2006**, *20*, 1709–1714.
61. Aravin, A.; Gaidatzis, D.; Pfeffer, S.; Lagos-Quintana, M.; Landgraf, P.; Iovino, N.; Morris, P.; Brownstein, M.J.; Kuramochi-Miyagawa, S.; Nakano, T.; et al. A novel class of small RNAs bind to MILI protein in mouse testes. *Nature* **2006**, *442*, 203–207.
62. Lee, H.-C.; Gu, W.; Shirayama, M.; Youngman, E.; Conte, D., Jr; Mello, C.C. C. elegans piRNAs mediate the genome-wide surveillance of germline transcripts. *Cell* **2012**, *150*, 78–87.
63. Wang, W.; Yoshikawa, M.; Han, B.W.; Izumi, N.; Tomari, Y.; Weng, Z.; Zamore, P.D. The initial uridine of primary piRNAs does not create the tenth adenine that Is the hallmark of secondary piRNAs. *Mol. Cell* **2014**, *56*, 708–716.
64. Czech, B.; Hannon, G.J. One Loop to Rule Them All: The Ping-Pong Cycle and piRNA-Guided Silencing. *Trends Biochem. Sci.* **2016**, *41*, 324–337.
65. Jehn, J.; Gebert, D.; Pipilescu, F.; Stern, S.; Kiefer, J.S.T.; Hewel, C.; Rosenkranz, D. PIWI genes and piRNAs are ubiquitously expressed in mollusks and show patterns of lineage-specific adaptation.

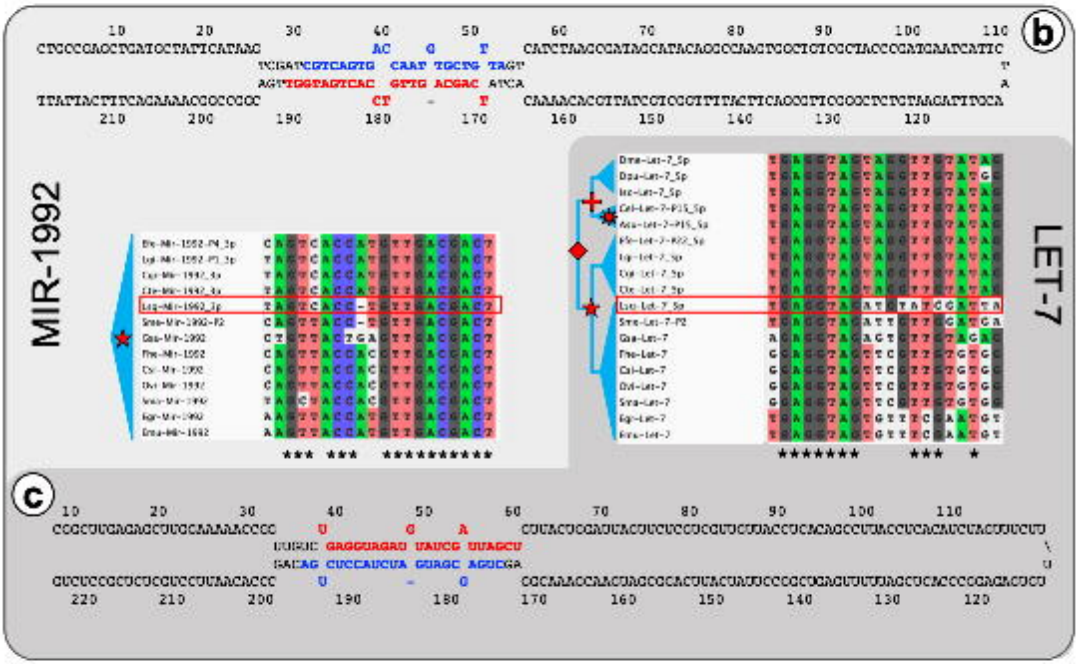
bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

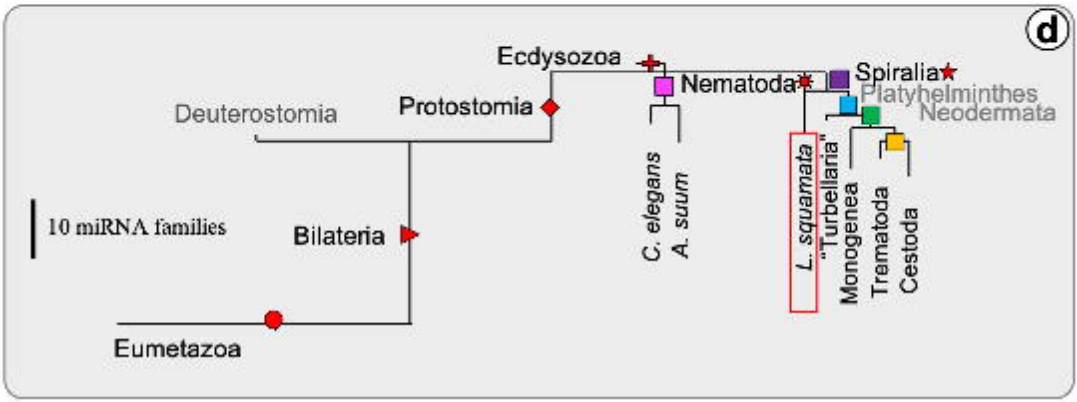
*Commun Biol* **2018**, *1*, 137.
66. Manakov, S.A.; Pezic, D.; Marinov, G.K.; Pastor, W.A.; Sachidanandam, R.; Aravin, A.A. MIWI2 and MILI Have Differential Effects on piRNA Biogenesis and DNA Methylation. *Cell Rep.* **2015**, *12*, 1234–1243.
67. Lewis, S.H.; Quarles, K.A.; Yang, Y.; Tanguy, M.; Frézal, L.; Smith, S.A.; Sharma, P.P.; Cordaux, R.; Gilbert, C.; Giraud, I.; et al. Pan-arthropod analysis reveals somatic piRNAs as an ancestral defence against transposable elements. *Nat Ecol Evol* **2018**, *2*, 174–181.
68. International Helminth Genomes Consortium Comparative genomics of the major parasitic worms. *Nat. Genet.* **2018**.
69. Weiss, M.J. Stabilization in the Name of the Freshwater Gastrotrich Lepidodermella squamata, with Nomenclatural Corrections for Congeneric Species. *Trans. Am. Microsc. Soc.* **1988**, *107*, 369–379.
70. Kang, W.; Eldfjell, Y.; Fromm, B.; Estivill, X.; Biryukova, I.; Friedländer, M.R. miRTrace reveals the organismal origins of microRNA sequencing data. *Genome Biol.* **2018**, *19*, 213.
71. Rosenkranz, D.; Zischler, H. proTRAC--a software for probabilistic piRNA cluster detection, visualization and analysis. *BMC Bioinformatics* **2012**, *13*, 5.
72. Friedlander, M.R.; Mackowiak, S.D.; Li, N.; Chen, W.; Rajewsky, N. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res.* **2012**, *40*, 37–52.
73. Kent, W.J. BLAT--the BLAST-like alignment tool. *Genome Res.* **2002**, *12*, 656–664.
74. Fromm, B.; Domanska, D.; Hackenberg, M.; Mathelier, A.; Hoye, E.; Johansen, M.; Hovig, E.; Flatmark, K.; Peterson, K.J. MirGeneDB2.0: the curated microRNA Gene Database. *bioRxiv* 2018, 258749.
75. Roovers, E.F.; Rosenkranz, D.; Mahdipour, M.; Han, C.-T.; He, N.; Chuva de Sousa Lopes, S.M.; van der Westerlaken, L.A.J.; Zischler, H.; Butter, F.; Roelen, B.A.J.; et al. Piwi proteins and piRNAs in mammalian oocytes and early embryos. *Cell Rep.* **2015**, *10*, 2069–2082.
76. Tosar, J.P.; Rovira, C.; Cayota, A. Non-coding RNA fragments account for the majority of annotated piRNAs expressed in somatic non-gonadal tissues. *Communications Biology* **2018**, *1*, 2.
77. Lowe, T.M.; Eddy, S.R. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **1997**, *25*, 955–964.
78. Lagesen, K.; Hallin, P.; Rødland, E.A.; Staerfeldt, H.-H.; Rognes, T.; Ussery, D.W. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* **2007**, *35*, 3100–3108.
79. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120.
80. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; Thompson, D.A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q.; et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **2011**, *29*, 644–652.
81. Robertson, G.; Schein, J.; Chiu, R.; Corbett, R.; Field, M.; Jackman, S.D.; Mungall, K.; Lee, S.; Okada, H.M.; Qian, J.Q.; et al. De novo assembly and analysis of RNA-seq data. *Nat. Methods* **2010**, *7*, 909–912.
82. Chang, Z.; Li, G.; Liu, J.; Zhang, Y.; Ashby, C.; Liu, D.; Cramer, C.L.; Huang, X. Bridger: a new framework for de novo transcriptome assembly using RNA-seq data. *Genome Biol.* **2015**, *16*, 30.
83. Cerveau, N.; Jackson, D.J. Combining independent de novo assemblies optimizes the coding transcriptome for nonconventional model eukaryotic organisms. *BMC Bioinformatics* **2016**, *17*, 525.
84. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T.L. BLAST+: architecture and applications. *BMC Bioinformatics* **2009**, *10*, 421.
85. Finn, R.D.; Coggill, P.; Eberhardt, R.Y.; Eddy, S.R.; Mistry, J.; Mitchell, A.L.; Potter, S.C.; Punta, M.; Qureshi, M.; Sangrador-Vegas, A.; et al. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* **2016**, *44*, D279–85.
86. Katoh, K.; Misawa, K.; Kuma, K.-I.; Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **2002**, *30*, 3059–3066.
87. Capella-Gutiérrez, S.; Silla-Martínez, J.M.; Gabaldón, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **2009**, *25*, 1972–1973.
88. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313.

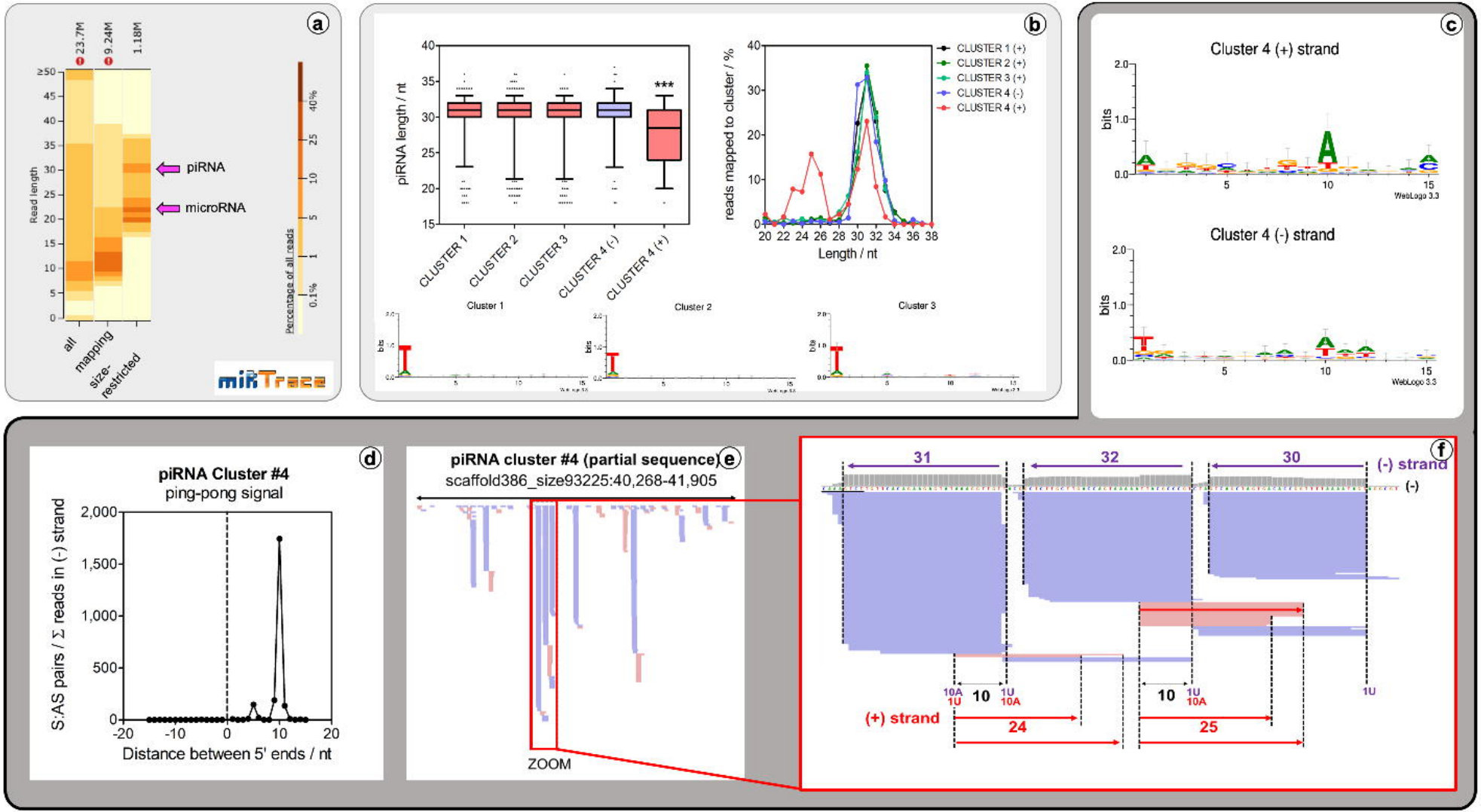
bioRxiv preprint doi: https://doi.org/10.1101/504100; this version posted January 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

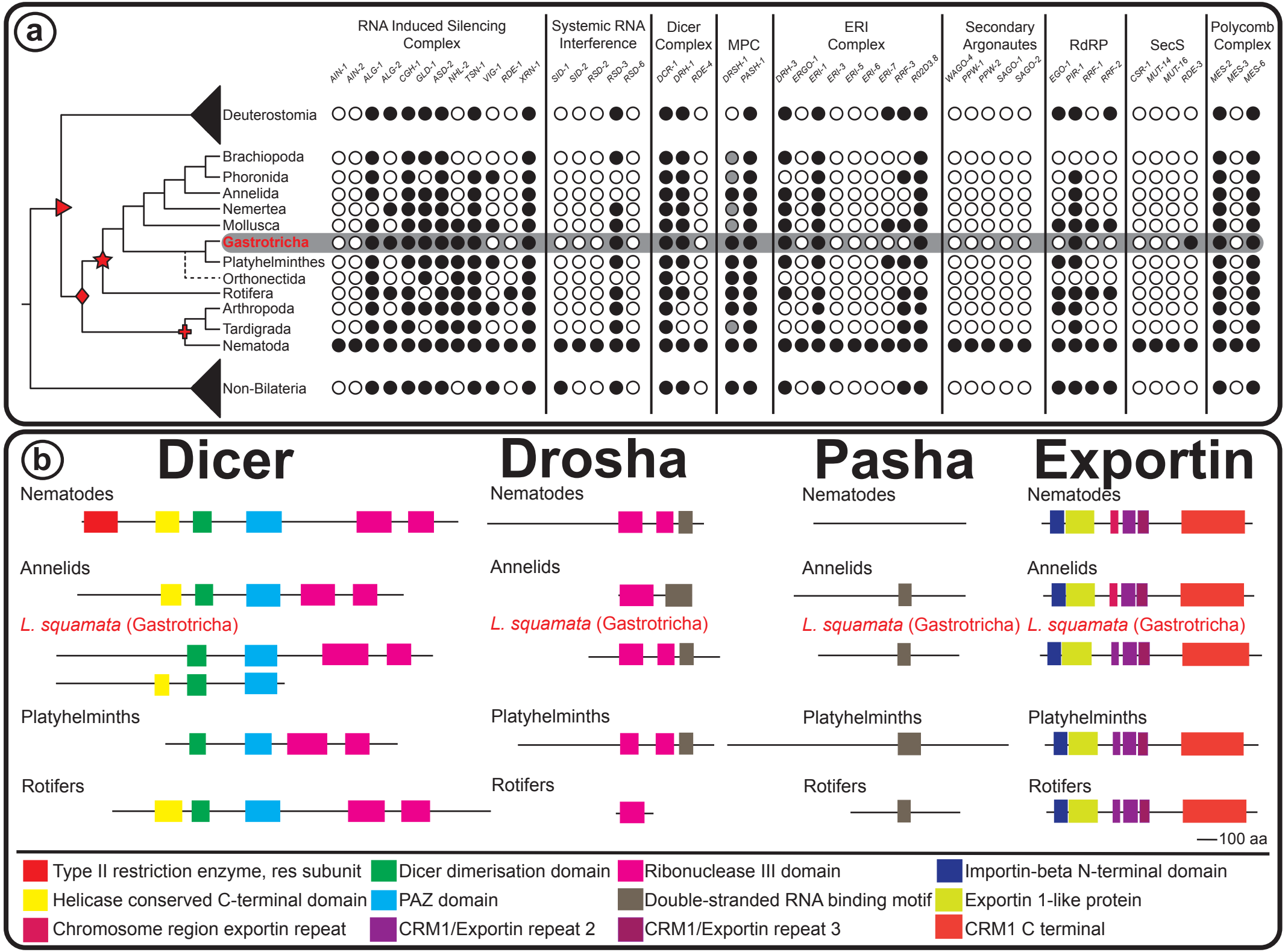
89. Letunic, I.; Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **2016**, *44*, W242–5.
90. Hejnol, A.; Lowe, C.J. Embracing the comparative approach: how robust phylogenies and broader developmental sampling impacts the understanding of nervous system evolution. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2015**, *370*.
91. Dunn, C.W.; Giribet, G.; Edgecombe, G.D.; Hejnol, A. Animal Phylogeny and Its Evolutionary Implications. *Annu. Rev. Ecol. Evol. Syst.* **2014**, *45*, 371–395.
92. Lu, T.-M.; Kanda, M.; Satoh, N.; Furuya, H. The phylogenetic position of dicyemid mesozoans offers insights into spiralian evolution. *Zoological Lett* **2017**, *3*, 6.

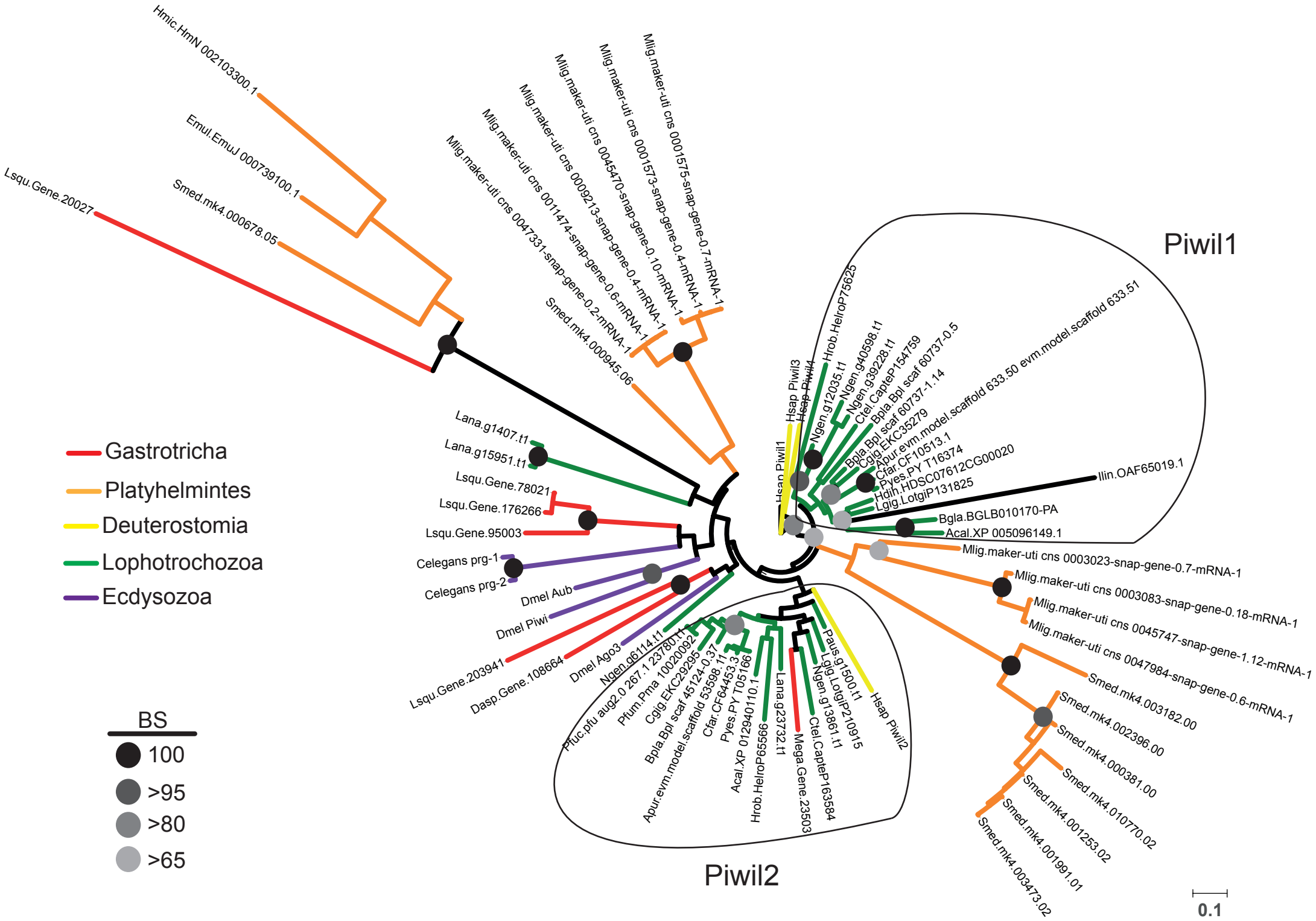

a
75 µm
C. elegans
A. suum
Lepidodermella
N. americanus
H. contortus
F. hepatica
E. granulosus


b
MIR-1992
LET-7
c


d
Ecdysozoa
Nematoda
Spiralia
Platyhelminthes
Neodermata
Deuterostomia
Protostomia
C. elegans
A. suum
L. squamata
"Turbellaria"
Monogenea
Trematoda
Cestoda
10 miRNA families
Bilateria
Eumetazoa


a
23.7M
9.24M
1.18M
Read length
piRNA
microRNA
all
mapping
size-restricted
Percentage of all reads
miRTrace
b
piRNA length / nt
CLUSTER 1
CLUSTER 2
CLUSTER 3
CLUSTER 4 (-)
CLUSTER 4 (+)
***
reads mapped to cluster / %
Length / nt
Cluster 1
Cluster 2
Cluster 3
c
Cluster 4 (+) strand
Cluster 4 (-) strand
bits
WebLogo 3.3
d
piRNA Cluster #4
ping-pong signal
S:AS pairs / Σ reads in (-) strand
Distance between 5' ends / nt
e
piRNA cluster #4 (partial sequence)
scaffold386_size93225:40,268-41,905
ZOOM
f
31
32
30
(-) strand
(-)
(+) strand
10A
1U
10
1U
10A
24
25
1U


a
RNA Induced Silencing Complex
AIN-1 AIN-2 ALG-1 ALG-2 CGH-1 GLD-1 ASD-2 NHL-2 TSN-1 VIG-1 RDE-1 XRN-1
Systemic RNA Interference
SID-1 SID-2 RSD-2 RSD-3 RSD-6
Dicer Complex
DCR-1 DRH-1 RDE-4
MPC
DRSH-1 PASH-1
ERI Complex
DRH-3 ERGO-1 ERI-1 ERI-3 ERI-5 ERI-6 ERI-7 RRF-3 R02D3.8
Secondary Argonautes
WAGO-4 PPW-1 PPW-2 SAGO-1 SAGO-2
RdRP
EGO-1 PIR-1 RRF-1 RRF-2
SecS
CSR-1 MUT-14 MUT-16 RDE-3
Polycomb Complex
MES-2 MES-3 MES-6
Deuterostomia
Brachiopoda
Phoronida
Annelida
Nemertea
Mollusca
Gastrotricha
Platyhelminthes
Orthonectida
Rotifera
Arthropoda
Tardigrada
Nematoda
Non-Bilateria
b
Dicer
Drosha
Pasha
Exportin
Nematodes
Annelids
L. squamata (Gastrotricha)
Platyhelminths
Rotifers
100 aa
Type II restriction enzyme, res subunit
Helicase conserved C-terminal domain
Chromosome region exportin repeat
Dicer dimerisation domain
PAZ domain
CRM1/Exportin repeat 2
Ribonuclease III domain
Double-stranded RNA binding motif
CRM1/Exportin repeat 3
Importin-beta N-terminal domain
Exportin 1-like protein
CRM1 C terminal


Piwil1
Piwil2
Gastrotricha
Platyhelmintes
Deuterostomia
Lophotrochozoa
Ecdysozoa
BS
100
>95
>80
>65
0.1
Hmic.HmN 002103300.1
Emul.EmuJ 000739100.1
Smed.mk4.000678.05
Lsqu.Gene.20027
Mlig.maker-uti cns 0001575-snap-gene-0.7-mRNA-1
Mlig.maker-uti cns 0001573-snap-gene-0.4-mRNA-1
Mlig.maker-uti cns 0045470-snap-gene-0.10-mRNA-1
Mlig.maker-uti cns 0009213-snap-gene-0.4-mRNA-1
Mlig.maker-uti cns 0011474-snap-gene-0.6-mRNA-1
Mlig.maker-uti cns 0047331-snap-gene-0.2-mRNA-1
Smed.mk4.000945.06
Lana.g1407.t1
Lana.g15951.t1
Lsqu.Gene.78021
Lsqu.Gene.176266
Lsqu.Gene.95003
Celegans prg-1
Celegans prg-2
Dmel Aub
Dmel Piwi
Dmel Ago3
Lsqu.Gene.203941
Dasp.Gene.108664
Hsap Piwil1
Hsap Piwil3
Hsap Piwil4
Hrob.HelroP75625
Ngen.g12035.t1
Ngen.g40598.t1
Ngen.g39228.t1
Ctel.CapteP154759
Bpla.Bpl scaf 60737-0.5
Bpla.Bpl scaf 60737-1.14
Cgig.EKC35279
Apur.evm.model.scaffold 633.50 evm.model.scaffold 633.51
Cfar.CF10513.1
Pyes.PY T16374
Hdih.HDSC07612CG00020
Lgig.LotgiP131825
Ilin.OAF65019.1
Bgla.BGLB010170-PA
Acal.XP 005096149.1
Mlig.maker-uti cns 0003023-snap-gene-0.7-mRNA-1
Mlig.maker-uti cns 0003083-snap-gene-0.18-mRNA-1
Mlig.maker-uti cns 0045747-snap-gene-1.12-mRNA-1
Mlig.maker-uti cns 0047984-snap-gene-0.6-mRNA-1
Smed.mk4.003182.00
Smed.mk4.002396.00
Smed.mk4.000381.00
Smed.mk4.010770.02
Smed.mk4.001253.02
Smed.mk4.001991.01
Smed.mk4.003473.02
Ngen.g6114.t1
Pfuc.pfu aug2.0 267.1 23780.1
Pfum.Pma 10020092
Cgig.EKC29295
Bpla.Bpl scaf 45124-0.37
Apur.evm.model.scaffold 53598.11
Cfar.CF64453.3
Pyes.PY T05166
Acal.XP 012940110.1
Hrob.HelroP65566
Lana.g23732.t1
Mega.Gene.23503
Ctel.CapteP163584
Ngen.g13861.t1
Lgig.LotgiP210915
Paus.g1500.t1
Hsap Piwil2