RESEARCH

Daphnia magna egg piRNA cluster expression profiles change as mothers age

Jack Hearn[1*] and Tom J. Little[2]

Abstract

Background: PiRNAs prevent transposable elements wreaking havoc on the germline genome. Changes in piRNA expression over the lifetime of an individual may impact on ageing through continued suppression, or release, of transposable element expression. We identified piRNA producing clusters in the genome of *Daphnia magna* by a combination of bioinformatic methods, and then contrasted their expression between parthenogenetically produced eggs representing maternally-deposited germline piRNAs of young (having their 1st clutch) and old (having their 5th clutch) mothers. Results from eggs were compared to cluster expression in three generations of adults.

Results: As for other arthropods, *D. magna* encodes long uni-directionally transcribed non-coding RNAs consisting of fragmented transposable elements which account for most piRNAs expressed. Egg tissues showed extensive differences between clutches from young mothers and those from old mothers, with 578 and 686 piRNA clusters upregulated, respectively. Most log fold-change differences for significant clusters were modest, however. When considering only highly expressed clusters, there was a bias towards 1st clutch eggs at 41 upregulated versus eight clusters in the eggs from older mothers. F_0 generation differences between young and old mothers were fewer than eggs, as 179 clusters were up-regulated in young versus 170 old mothers. This dropped to 31 versus 22 piRNA clusters when comparing adults in the F_1 generation, and no differences were detected in the F_3 generation. Inter-generational losses of differential piRNA cluster were similar to that observed for *D. magna* micro-RNA expression.

Conclusions: Little overlap in differentially expressed clusters was found between adults containing mixed somatic and germline (ovary) tissues and germ-line representing eggs. A cluster encompassing a Tudor domain containing gene important in the piRNA pathway was upregulated in the eggs from old mothers. We hypothesise that regulation of this gene could form part of a feedback loop that reduces piRNA pathway activity explaining the reduced number of highly-expressed clusters in eggs from old mothers.

Keywords: piRNA, Epigenomics, Small RNA, Life-history, Ageing, Daphnia, Arthropod small RNAs, Differential expression, Bioinformatics

Background

Roles of piRNAs and their production

Piwi-interacting RNAs (piRNAs) are 21-35 nucleotide long small RNAs that are maternally deposited into oocytes to provide immunity to complementary transposable elements (TEs) by suppressing new insertions into the germline. This protects the developing embryo from transposon-mediated illegitimate recombination, double-stranded breaks, and disruptive insertions into coding sequences promoters which can cause aberrant gene expression [1, 2]. Although efficient against recognised TEs of maternal origin, piRNAs are less effective against TEs of paternal origin [2]. In arthropods, somatically expressed piRNAs are an ancestral mechanism of protection against TEs [3], although research has

*Correspondence: Jack.Hearn@lstmed.ac.uk

[1] Department of Vector Biology, Liverpool School of Tropical Medicine, Liverpool, UK
Full list of author information is available at the end of the article

focussed on germline piRNAs. This transposon suppression role may even represent a deeper ancestral trait of bilaterians [4, 5]. Additional functions of piRNAs include regulating gene expression, protecting against viruses and telomere maintenance [2, 6], and in silkworms piRNAs act as master-regulators of sex-determination [7].

PiRNAs originate from long precursors RNAs in arthropods and mammals [6], and are first transcribed and then further processed in the cytoplasm into mature piRNAs [6]. These piRNA-producing clusters contain contiguous remnants and nested fragments of transposons consigned to genomic 'graveyards' [2, 6]. They are transcribed in a single direction in most arthropods surveyed to date, with *Drosophila* species also having evolved a unique dual-strand stranded system expressed by germline cells [8, 9]. After transcription, piRNA cluster transcripts enter either the ping-pong cycle or phased piRNA pathways which both occur in the cytoplasm [2, 6]. Phased piRNA production can occur in germline and somatic cells and occurs when Piwi (P-element induced wimpy testis) is loaded at the 5' end of piRNA precursors and cleaved by the Zucchini protein [10]. The process is then repeated in a step-by-step process through Piwi loading of newly created 5' ends to produce distinct piRNAs along the precursor RNA [11]. The ping-pong cycle occurs in germline cells when a sense orientated piRNA guides the Argonaute 3 (Ago3) protein to a complementary cluster RNA and cleaves it [11]. Once Ago3 and a piRNA are bound, the Aubergine (Aub) protein binds to the 5' end of the Ago3-piRNA targeted mRNA cleavage products and slices it into a mature anti-sense piRNA which recognises and cuts complementary RNAs such as TE mRNA. Ago3 recognises the 5' ends of these cleaved RNAs thus propagating the ping-pong loop which ultimately results in the post-transcriptional silencing of transposons [11]. Maternal deposition of the Aub protein and associated piRNAs into oocytes initiates the ping-pong cycle intergenerationally in *Drosophila* [1, 12].

Together, the ping-pong loop and phased piRNA pathways create a diverse population of piRNAs optimised for their RNA silencing role [2]. In addition to post-transcriptional suppression of targets, both pathways also can direct transcriptional silencing in the nucleus [2]. PiRNAs guide Ago3 to complementary transposable element insertion sites in the nucleus and promote H3K9me3 modification of histone tails which leads to heterochromatin formation and ultimately silencing of the locus [2, 6, 11]. The characteristic size of piRNAs is determined by loading of intermediate piRNAs into Piwi and Aub proteins. Due to intrinsic preferences of Piwi proteins in each pathway, phased pathway piRNAs have a bias for Uridine at the 1st base (1U) and ping-pong cycle piRNAs for Adenine at the 10th position (10A) with no associated 1st base bias [6]. Together, piRNA length distributions and base position biases are useful characteristics for classifying piRNAs versus other sRNA species.

The association between Piwi and piRNAs is weaker than that for miRNAs with Ago1 [13]. PiRNAs therefore require a longer region of matching with their target RNA than the seven base pairs in miRNAs to form a stable association. As a result, piRNA target recognition is more selective than for miRNAs, but beyond an essential seed-matching region, mismatches between piRNA and their targets is tolerated. This can lead to evolutionarily conserved piRNA-target interactions despite the accumulation of mutations in transposable element sequences over time [13].

Ageing and the piRNA pathway in arthropods

Ageing is associated with increased expression of transposable elements in a variety of animals due to progressive genomic dysregulation [14, 15], with age-related phenotypes resulting from negative effects of transposition on cell and genome integrity [14]. Among arthropods, lifespan has been shown to increase in *D. melanogaster* through the application of reverse transcriptase inhibitors which reduced TE activity [14, 16]. The piRNA pathway was observed to differ intergenerationally with age in *D. melanogaster*, as egg chambers of older mothers were upregulated for piRNA pathway genes (27 of 31 tested) [17]. PiRNA pathway genes have also been implicated in somatic ageing of the termite *Macrotermes bellicosus* [18]. The heads of older, worker caste termites had lower expression of four such genes and higher expression of several hundred TEs versus younger workers suggestive of reduced suppressive activity [14, 18]. By contrast much longer-lived termite kings and queens maintain stable TE and gene expression levels throughout lifetimes [18]. In line with termite workers, increased TE activity with age occurs in *Drosophila* somatic tissues [19–21], albeit with the caveat that TE insertion rates are prone to overestimation in sequencing data [22]. It has been hypothesised that age-related misexpression of TEs is restricted to somatic tissues due to efficient policing of the germline by the piRNA response through life [17, 18, 23].

The piRNA pathway in Branchiopod crustaceans

The focal species of this study, *Daphnia magna*, is a member of the Crustacean class Branchiopoda. The Crustacea is a paraphyletic sub-phylum and the Branchiopoda is more closely related to insects (subphylum Hexapoda) than crabs for example (class Malacostraca) [24, 25]. *Daphnia* species have been the focus of limited piRNA studies in this group to date. Notably, the piRNA pathway is likely to underly the evolution of obligately

parthenogenetic reproduction from cyclical parthenogenesis in the cladoceran crustacean *Daphnia pulex* [26]. A transposon insertion upstream of the meiotic cohesion factor *Rec8* in *D. pulex* was found to correlate perfectly with obligate parthenogenesis in effected strains [26]. The piRNA pathway is hypothesised to silence the TE-inserted copy of *Rec8* and several other wildtype paralogs of *Rec8* through sequence homology to the generated piRNAs [26]. Little further is known of the presence and action of piRNAs in *Daphnia* or generally across the crustacea, nor what impact a cyclically-parthenogenetic lifecycle has on piRNA dynamics. *Daphnia magna* and *Daphnia pulex* each encode seven essential piRNA pathway Piwi-like genes in their genomes [27, 28], alongside three Argonaute genes in *D. magna* to two in *D. pulex* [27]. Other crustaceans are likely to encode piRNA pathway genes, with evidence for piRNA expression and piRNA pathway genes in *Triops cancriformis* (tadpole shrimps, class Branchiopoda, order Notostraca) [29]. Seven copies of Piwi in *Daphnia* species represents a gene expansion versus other arthropods, which is perhaps associated with sub-functionalisation across Piwi copies to the soma and splitting of germline roles between meiosis and parthenogenesis of cyclically parthenogenetic species [30]. A similar association between expanded Piwi-like genes and reproductive plasticity was observed in the genome of the pea aphid *Acyrthosiphon pisum* [31] (class Insecta, order Hemiptera). In addition to gene expansion, genes of the piRNA pathway exhibit strong signals of positive selection and expansion in *Drosophila* species, probably in response to repeated transposable element invasions and in antiviral defence [32, 33].

Daphnia magna piRNA responses to ageing

Prior research has established that micro-RNA (miRNA) expression and DNA methylation status respond to ageing and caloric restriction in *D. magna* [34–36], with caloric restriction resulting in pervasive effects on gene expression [37]. The miRNA and DNA methylation profile of mothers changed with age in parthenogenetically reproducing individuals [34, 36]. The eggs of old mothers also showed a different miRNA profile from the eggs of young mothers, but this difference was greatly reduced once eggs hatched and grew to adulthood, and was not evident in great granddaughters [34]. Hence, the miRNA profiles of adult *Daphnia* reflect the age of the individual and not their parental generation, and due to maternal provisioning of eggs, will also reflect their mother's age. Here, we take advantage of that pre-existing resource to rigorously identify piRNAs in eggs and adult tissues of *D. magna* for the first time, which we used to predict putatively piRNA producing loci along the *D. magna* genome. By quantifying clusters in eggs from primiparous and multiparous (specifically, having their 5th clutch) mothers, we tested for age-related changes in piRNA expression in eggs and their maternal and descendent generation adult *Daphnia*. There were extensive differences between eggs from young and old mothers with little overlap in differentially expressed clusters in their mothers, akin to the pattern observed in miRNAs. PiRNA expression is therefore regulated to some extent by ageing-related processes in *D. magna*.

Results

piRNA cluster identification and annotation

Quality and length filtering of reads had the largest effect on egg small RNA libraries with 38-72% of data remaining after this step, removal of miRNA, other non-coding RNAs and piRNA classification had a lesser effect (Table S1). After all filtering steps 35-64% of original reads were retained, and absolute numbers of reads per egg library ranged from 5.2 to 10.5 million. Results for adult piRNA libraries were similar with one outlier (replicate Y1A_F0, Table S1). There were 15,719 ShortStack clusters after aligning egg and adult libraries (Table 1), of these 4,747 had a significant (adjusted *p*-value < 0.05) PingPongPro predicted ping-pong cycle signature. ProTRAC predicted

Table 1 Numbers of piRNA clusters predicted after each filtering step

Predictions	Number
ShortStack predicted clusters.	15719
ShortStack clusters containing PingPongPro predicted transposons.	4747
Adult proTRAC cluster predictions.	19
Egg proTRAC cluster predictions.	19
Merged proTRAC cluster predictions.	23
PingPongPro and proTRAC merged predictions.	4732
Average read length filtered clusters.	4653
Final count after clusters overlapping other RNA species were removed.	4606

The final row gives the number of clusters input to differential expression analysis

19 piRNA producing loci each for egg and adult libraries analysed separately (File S1). On combining egg and adult proTRAC results 22 unique piRNA producing loci remained, all of which were transcribed uni-directionally. After merging ShortStack and proTRAC predicted loci, removal of clusters with an average read length of 24 or less, and clusters overlapping other species of RNA in the genome annotation 4,606 putative piRNA loci remained for expression-based analyses, including 20 proTRAC clusters. These 4,606 clusters had an average aligned read length of 26.27 bp (standard deviation, ±0.53). The proTRAC predicted clusters spanned 2.1% of all filtered putatively piRNA producing loci (466,865/22,296,837 bp), whilst accounting for on an average of 47% and 56% total expression (as proportion of normalised counts) in eggs and adults respectively (Table S2). Cluster lengths varied from 31 to 91,046 bp and were present on all 10 linkage groups of the *D. magna* genome.

EggNOG annotated 2,832 of 4,606 piRNA clusters, of which 2,214 received a description (Table S3, piRNA clusters with entries in the "Description" column). Among these clusters 976 received eggNOG annotations corresponding to transposable elements (Table S4). A further 74 clusters received Ribonuclease H domain annotations (Table S4) which are likely to form part of transposable elements. RepeatMasker annotated more clusters with TE elements at 86% (3,948/4,606) containing at least one. This dropped to 67% (3,058/4,606) of clusters when restricting RepeatMasker predictions to known TE families.
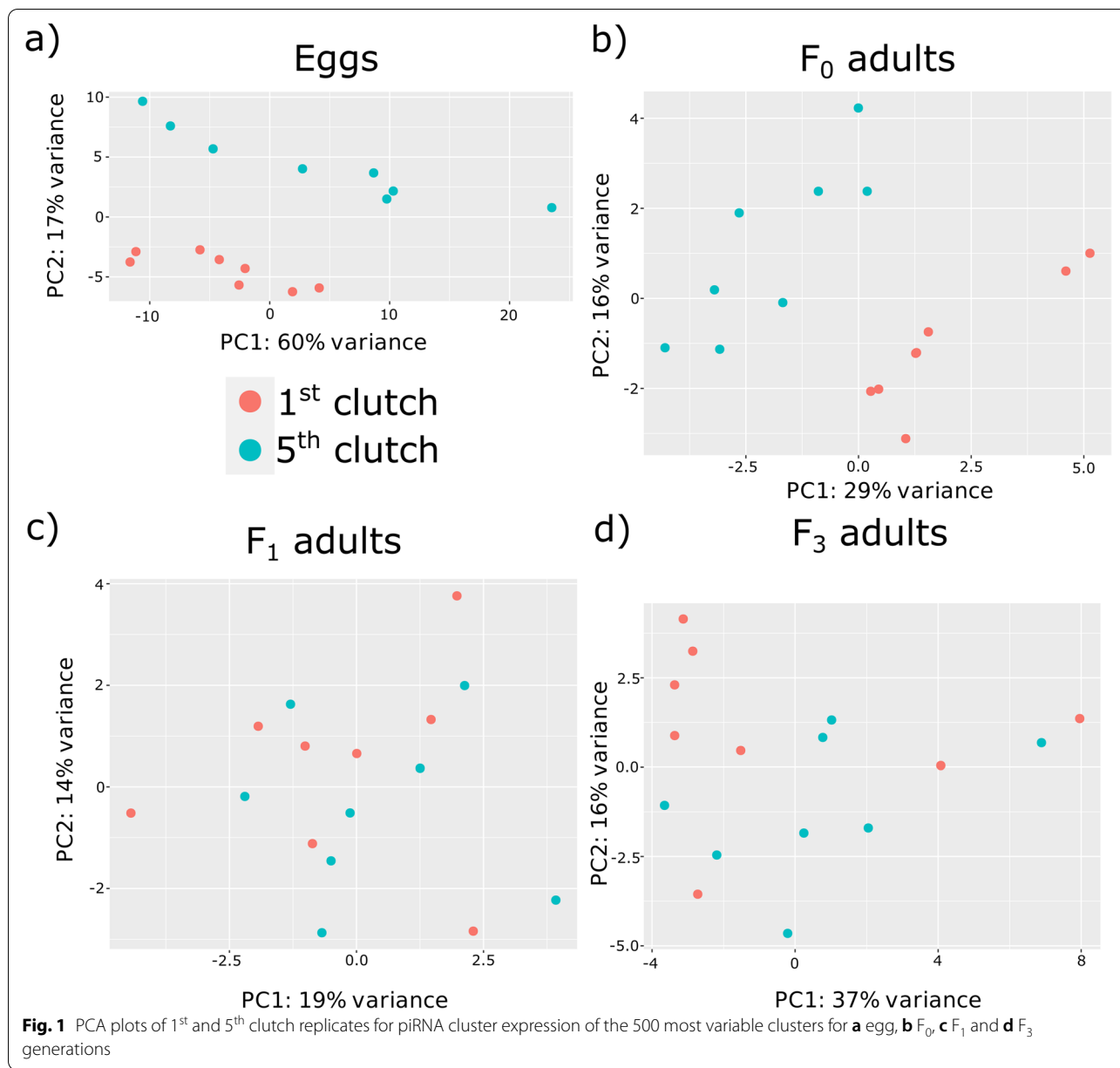
Egg versus adult piRNA cluster expression
Many clusters had large expression differences between egg replicates combined and adult replicates combined ($F_0 + F_1 + F_3$ generations). For egg transcripts per million (TPM) per cluster averaged across replicates, 33 had a TPM of 1000 or more greater than adult averages. Conversely, 37 clusters had a TPM greater than 1000 or more in the average of adult replicates relative to eggs (Table S5). Three of the clusters upregulated in eggs had much greater differences in TPM (13,354-20,060 TPM) versus the remainder (1,059-6,130 TPM). Two of these most upregulated clusters were proTRAC predicted piRNA producing loci occurring on the negative strand. NC_046175.1:10253003-10266628 at 13.6 kb long encompasses a long non-coding RNA (lncRNA) which act as precursors to piRNAs flanked by two protein coding genes. Most piRNAs aligned in the region within the lncRNA and were biased toward primary piRNAs, with 96% of 1st read positions being uracil and 27% adenine at the 10th position. The second cluster had a similarly biased 1U to 10A ratio (88% to 22%). For this cluster, no overlapping lncRNA has been annotated and most reads align across and downstream of an uncharacterised gene 'LOC116926960'. The third cluster not within a proTRAC prediction was very short at 35 bp long and did not overlap a feature, the closest gene-body was ~3 kb downstream of the cluster and encodes a lncRNA locus. Seven adult clusters had a TPM difference greater than 10,000 in favour of adults versus eggs. Of these, six are proTRAC clusters and the remaining locus was annotated with retro-element domains by eggNOG. Further to these high-expression clusters, two consecutive clusters separated by only 2.8 kb had very low average TPMs (< 10) in eggs versus adults (> 1000). Both clusters overlap a single gene (LOC116928002) which encodes a vitellogenin-2-like protein.

Differential expression of piRNA clusters in eggs
Replicates separated by clutch in Eggs and the F_0 maternal generation (PCA plots, Fig. 1a,b), with principal component 1 accounting for 60% of variation in eggs versus 29% in F_0 adults. This was lost in F_1 adults and F_3 generations which were inter-mixed by clutch (Fig. 1c,d). There were 578 differentially expressed clusters with expression greater in 1st clutch eggs and 686 such clusters in 5th clutch eggs (Fig. 2, Table 2 and Files S2 and S3). Although a large number of differentially expressed loci in each clutch, the magnitude of difference was modest for the majority of clusters (Fig. 2a, contrast with 2b), with an average $\log_2$-fold change of 0.56 for significant egg clusters (File S2, $\log_2$-fold change column averaged). Bias is increased towards 5th clutch eggs when considering clusters of longer than 10 kb, at 86 with expression greater in 1st clutch eggs and 213 in 5th clutch eggs. Only 130 of all 1264 differentially expressed egg piRNA loci exhibited a doubling in expression in one clutch versus the other ($\log_2$-fold change greater than one or less than minus one), and the majority of these had low absolute expression levels.
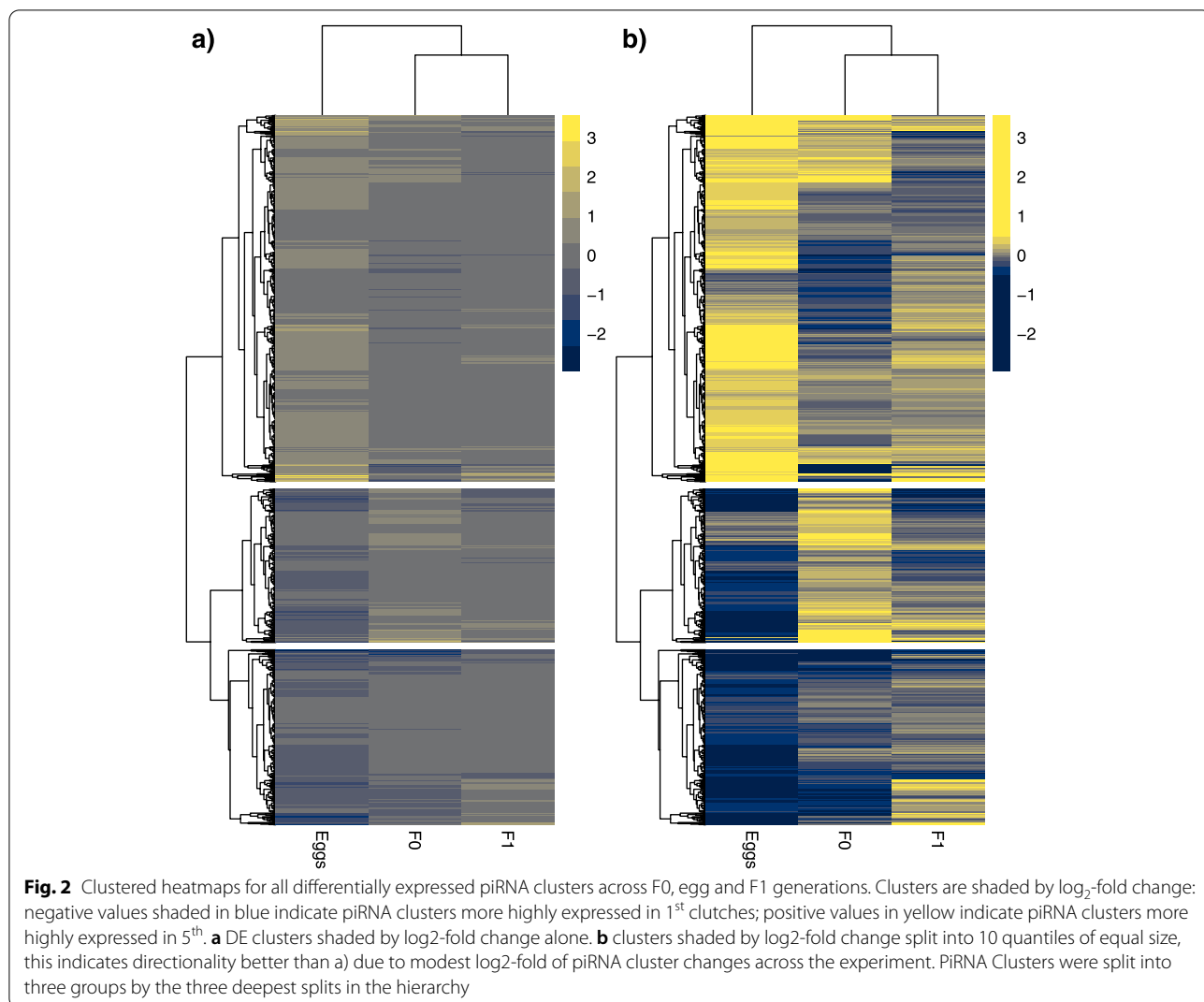
When restricting differentially expressed loci to those with a count of greater than 1000 transcripts per million (TPM) 41 and eight 1st and 5th clutch clusters were significant respectively (heatmap, Figure S1, Table S6). Of the eight 5th clutch loci (Table S6), seven overlap a gene-body in the genome annotation four of which are lncRNA loci and four also contain TE elements by RepeatMasker annotation. One cluster was annotated (NC_046177.1:4348409-4351253) as containing a Tudor domain (TDRD6) which has roles in germ cell development and the piRNA ping-pong cycle [38, 39]. Related to the ping-pong cycle, three further clusters overlap ATP-binding helicase genes. However, none of these were differentially expressed in eggs or

Fig. 1 PCA plots of 1st and 5th clutch replicates for piRNA cluster expression of the 500 most variable clusters for **a** egg, **b** F_0, **c** F_1 and **d** F_3 generations

adults nor were they of high TPM expression. For the 41 clusters more highly expressed in the 1st clutch, 22 overlapped predicted protein coding genes, 10 lncR-NAs and two pseudogenes. Annotated protein-coding genes included Spermatogenesis-associated protein 20 (NC_046183.1:4720450-4720481), but no genes directly implicated in piRNA production or regulation. Additionally, 33 of the 41 1st clutch clusters overlap a known transposable element versus zero for the 5th clutch clusters, which was also reflected in 7 egg-NOG annotations to repetitive element domains for 1st clutch clusters.

Differential expression of piRNA clusters in adults

In the maternal (F_0) generation, 179 clusters were up-regulated in mothers on their 1st clutch versus 170 in those on their 5th clutches (Table 2, File S3). This dropped to 31 and 22 clusters in the F_1 generation adults (whose mothers produced them when on either their 1st or 5th clutch) and zero differences between clutches in the F_3 generation (whose great grandmothers were on either their 1st or 5th clutch). There was low concordance in cluster expression between adult and egg generations (Fig. 3, Table S7), with less than 10% of egg clusters being shared with F_0 adults and a smaller

Fig. 2 Clustered heatmaps for all differentially expressed piRNA clusters across F0, egg and F1 generations. Clusters are shaded by log₂-fold change: negative values shaded in blue indicate piRNA clusters more highly expressed in 1st clutches; positive values in yellow indicate piRNA clusters more highly expressed in 5th. **a** DE clusters shaded by log2-fold change alone. **b** clusters shaded by log2-fold change split into 10 quantiles of equal size, this indicates directionality better than a) due to modest log2-fold of piRNA cluster changes across the experiment. PiRNA Clusters were split into three groups by the three deepest splits in the hierarchy

Table 2 Differentially expressed piRNA clusters upregulated in upregulated 1st or 5th clutch in each generation

Comparison	Upregulated in 1st clutch	Upregulated in 5th clutch
Eggs	578	686
Eggs, TPM > 1000	41	8
F_0	179	170
F_1	31	22
F_3	0	0

TPM > 1000 = an equal to or greater than 1000 average transcripts per million threshold in the upregulated condition for egg comparisons

amount with F_1 adults. This distinction between adult and egg libraries was supported by 14 of the highly-expressed proTRAC clusters being significantly differentially expressed in F_0 adults. Whereas only one proTRAC cluster (NW_022654559.1:2-5240) was differentially expressed in eggs, in that case more highly in 1st clutch eggs. Of the 14 significant F_0 clusters, 13 were more highly expressed in mothers on their 5th clutch than those on their 1st, however the single proTRAC cluster more highly expressed in young mothers was the same as that for eggs and had the lowest adjusted *p*-value of the 14 (4.84 x 10⁻¹³). This cluster is annotated as an RNA polymerase II regulatory region and possibly encodes a transposase due to the presence of an Activator family domain (NW_022654559.1:2-5240, Table S3).

Transposable element encoded in differentially expressed piRNA clusters

Most differentially expressed piRNA clusters in all comparisons overlapped with RepeatMasker predicted locations of complex repeats (Table 3, File S4), percentages which remained high when including only known

Fig. 3 Venn diagrams of overlaps between eggs and F_0 and F_1 adult differentially expressed clusters. **a** shows clusters upregulated in the 1st clutch versus 5th clutch and **b** clusters upregulated in the 5th clutch versus 1st clutch
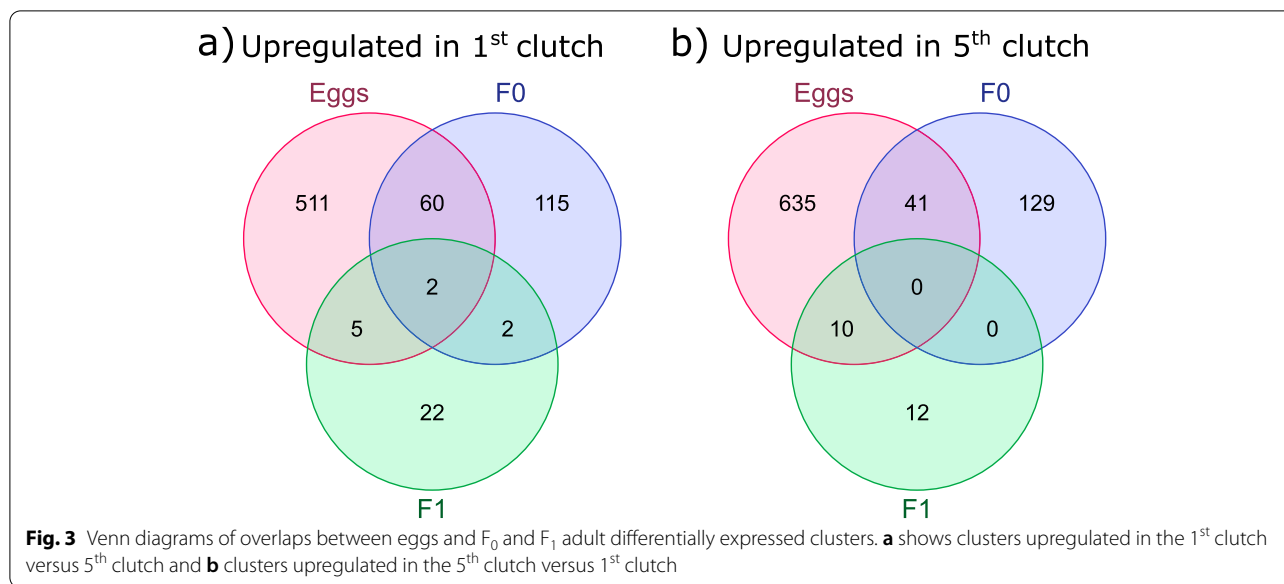
Table 3 PiRNA clusters annotated with TEs by RepeatMasker and eggNOG

Comparison	Upregulated in 1st clutch			Upregulated in 5th clutch		
Program	RepeatMasker	RepeatMasker known	EggNOG	RepeatMasker	RepeatMasker known	EggNOG
Eggs	539 (93%)	422 (73%)	140 (24%)	623 (91%)	587 (86%)	294 (42%)
Eggs, TPM > 1000	32 (78%)	22 (54%)	7 (17%)	4 (50%)	4 (50%)	0 (0%)
F_0	167 (93%)	154 (86%)	79 (44%)	161 (95%)	144 (85%)	53 (31%)
F_1	30 (97%)	24 (77%)	10 (31%)	21 (95%)	17 (77%)	7 (31%)

repeat families. Differentially expressed Egg clusters with TPM > 1000 had the fewest TE annotations. For eggnog annotations, in piRNA clusters upregulated in 1st clutch *D. magna* 346/578 (60%) were annotated by eggNOG mapping, of these 24% (140/578) were annotated with TE-like sequences (Table 3). This rose to 90% (614/686) of clusters upregulated in 5th clutches with 42% (294/686) encoding TE sequences. Most clusters were also annotated for piRNA clusters upregulated in F_0s at 85% (152/179) and 84% (142/170) for 1st and 5th clutches respectively. This was also true for the much-reduced number of significant clusters in the F1, as 81% (25/31) and 91% (20/22) for 1st and 5th clutches respectively.

Percentages are percentage of the total differentially expressed clusters in that comparison given in Table 2. "RepeatMasker" annotations are repetitive elements remaining after filtering for low-complexity and simple repeats; "RepeatMasker known" are annotations to known TE elements by filtering RepeatMasker results for "unknown" repeat families; "EggNOG" piRNA cluster TE annotations, including Ribonuclease H, domains.

Discussion

By integrating two methods of predicting small RNA loci a high-confidence, well-replicated piRNA dataset was created for 1st and 5th clutch eggs, their parents (the F_0), and descendent generations of parthenogenetically reproducing *D. magna* adults (the F_1 and F_3 generations). Among 4,606 piRNA clusters we identified almost 4,000 clusters containing transposable element (TE) sequences of which over 3,000 belong to known families. PiRNA clusters were then quantified and contrasted between clutches representing different ages of *D. magna*. A large number of clusters were differentially regulated in eggs, and much less so in whole adult tissues of each generation tested. Eggs and their parental F_0s showed good separation of expression patterns (Fig. 1), a result in line with miRNAs derived from the same dataset [34]. Hence, the piRNA profile of eggs is likely dictated by maternal provisioning of the eggs as in *Drosophila* [1, 40, 41]. This commonality was lost by adulthood, as in the F_1 generation, adults were inter-mixed in expression profiles and differential expression was much reduced between clutches compared to F_0s and eggs. Therefore, F1 adults are more

similar to one another in their piRNA profiles than their maternal generation, as was the case for miRNAs [35]. This indicates that differences in piRNA profiles in the egg resulting from maternal-provisioning are mostly lost or 'reset' as part of *D. magna* development. Future research will determine if the remaining differences have a functional effect or represent 'expression noise'.

D. magna piRNA cluster profiles

Despite covering only 2.1% of the piRNA producing loci, proTRAC-predicted clusters were responsible for approximately ~50% of piRNAs. In this, *D. magna* is similar to *Drosophila* where such long clusters of fragmented transposons, which proTRAC was designed to detect, are also responsible for producing the majority of piRNAs [8, 42]. This similarity to *Drosophila* may reflect monophyly as *Daphnia* (class Branchiopoda) are more closely related to the Hexapoda than other crustacean classes except the Remipedia [24, 25]. Only mono-directionally expressed piRNA clusters were predicted by proTRAC in *D. magna* egg and adult tissues. Bi-directionally expressed piRNA clusters found in *Drosophila* may represent a fly-specific adaptation [9, 43], however a minority of piRNA clusters (15%) in the mud crab crustacean, *Scylla paramamosain*, were also expressed in this manner [44]. As such, we cannot rule out the presence of bi-directionally expressed clusters in *D. magna*.

Parthenogenetically-reproducing *D. magna* piRNA clusters

All egg data surveyed was from embryos destined to become asexual females. The advantage of which is the genetic identity of replicates, which was expected to reduce variance in response to age. Conversely, this means males were not sampled in this study and testis are known to be a location of piRNA expression in arthropods [44–46]. Hence, this study is not an exhaustive survey of piRNAs in *D. magna,* as distinct piRNA clusters may well be expressed in male germlines and/or somatic tissues. In *S. paramamosain,* piRNA expression and cluster activity was much greater in ovaries than testes [44], and predicted piRNA clusters sizes were longer and more numerous in ovaries than testis. Indeed, although not an exact comparison, the number of expressed ovary piRNA clusters in *S. paramamosain* was much greater than that for *D. magna* eggs (19). Sensitivity to proTRAC sliding window parameters may contribute to these different estimates [47]. In [44], 300 bp windows with increments of 100 bp were applied versus 5000bp windows across 1000bp increments here. The phylogenetic divergence between *S. paramamosain* and *D. magna* may also play a role, as they belong to the Malacostraca and Branchiopoda, respectively.

Tudor domain genes regulated by piRNAs

Tudor domain genes are thought to form a molecular scaffold that connects elements of the ping-pong cycle [48], a piRNA cluster containing such a domain (TDRD6) was upregulated in 5th clutch eggs. From this, we hypothesise that piRNAs may regulate the ping-pong cycle through down-regulation of pathway genes. If piRNAs are down-regulating a component of their own pathway in one treatment we would expect to see lower general expression of clusters relative to comparisons. This is indeed what we observed here with more clusters with a TPM > 1000 over-expressed in 1st clutch versus TDRD6 containing cluster-upregulated 5th clutch eggs at 41 and eight clusters respectively. This hypothesis is speculative and would require implication of the identified TDRD6 domain encoding gene in piRNA production in *Daphnia* and supporting gene expression data for 1st and 5th clutch eggs. Furthermore, 32 of 41 clusters with TPM > 1000 upregulated in 1st clutches were annotated with a TE element or domain versus four of eight in 5th clutches, suggestive of more efficient TE control in 1st clutch parthenogenetically-reproducing *D. magna*. This is in line with *Drosophila* somatic tissues [19–21], but not with the idea of continued efficient policing of reproductive tissues with age [17, 18, 23]. No other differentially expressed piRNA clusters overlapped piRNA pathway genes. Three piRNA clusters of low TPM expression in eggs and adults did overlap ATP-binding helicase genes which act during oogenesis alongside piRNAs to methylate and repress transposable elements [49].

PiRNA cluster versus piRNA targeting

We restricted this analysis to regions of the genome from which piRNAs originate. This is because most small RNA prediction algorithms have been designed to identify miRNA targets according to the well understood seed matching rules of miRNA-target interactions. Such methods are prone to false positives [50, 51]. Because of this, Fridrich et al [50] recommend biological interpretation of miRNA interactions in combination with experimental result. Currently, it is hard to justify using such imprecise algorithms for piRNA interactions as they were not designed for this (but see [52–57] among others). Despite this, recent advances in understanding the seed-matching rules of piRNAs indicate longer and therefore more specific piRNA-target interactions [13, 58, 59], rules which appear to be shared with *Aedes* mosquitoes [60]. If a common property across animals, more limited potential targets per piRNA than miRNA will perhaps make computational inference of piRNA targets a more fruitful exercise than it has been for miRNAs.

Conclusions

By strict and multi-step filtering it was possible to enrich a small RNA dataset for piRNAs. Clusters of transposable element fragments covered a small fraction of the total predicted piRNA-producing loci but were responsible for most piRNA expression. These clusters were all transcribed in a single direction. Differential expression results were interpreted by piRNA clusters versus targets due to false-positive issues with miRNA-based targeting approaches. Recent insights into piRNA seed-matching to targets, however, may make piRNAs more amenable than miRNAs to computational inference of targets in future. The dynamics of piRNA cluster expression in *Daphnia magna* changes with age in adults and their eggs, but not in the resulting adults or subsequent generations, as was observed for miRNAs in this *D. magna*. Our results suggest more efficient control of TEs in younger 1st clutch *Daphnia* than in older *Daphnia* on their 5th clutch, perhaps through a piRNA-mediated negative feedback on a Tudor domain containing components of the piRNA pathway, a hypothesis requiring further investigation.

Methods

Maternal ageing experiment and sequencing

This experiment was first described in Hearn et al 2018 [34] and is summarized here. A clone of *Daphnia magna* (C32) originating in Kaimes pond near Leitholm in the Scottish Borders, United Kingdom [61] shows a maternal effect pattern where large offspring are produced when mothers are older [62]. To create experimental lines, 24 groups of five female *Daphnia* were placed into jars of 200 ml of artificial culture medium and fed 2.5 x 10^6 cells of the single-celled green algae *Chlorella vulgaris* daily for three generations. After three generations, five second clutch new-born females were designated F_0 of the experiment and fed *ad libitum* (5 x 10^6 cells *C. vulgaris* daily). Eggs of the 1st and 5th clutch of these acclimatized F_0 mothers were collected (Fig. 4a). Eggs were collected by flushing brood chambers with medium using a hypodermic syringe, pipetted onto tissue paper to dry, and then ground by motorized pestle in 350 μl Qiazol. Eggs from six jars each containing five *Daphnia* were combined to form a biological replicate from 30 *Daphnia* mothers in total. Eight biological replicates from each 1st and 5th clutch mother were created resulting in 16 egg libraries in total.

Adult reference libraries were created separately from eggs (Fig. 4b). First, replicate young (i.e., on their first clutch) and replicate old mothers (on their fifth clutch) were harvested to produce the F_0 set of RNA samples. Prior to RNA harvesting, new-born from these young and old F_0 mothers were isolated and grown to adulthood. This next generation of adults constituted the F_1, and RNA from F_1 adults was harvested just after they had their first clutch, the idea being to test for a maternal effect of having been born to a young or old (F_0) mother [32]. The first clutch offspring of F_1 adults were used to seed an F_2 generation, from which RNA was not harvested. The first clutch offspring of F_2 adults were used to seed an F_3 generation, from which we harvested RNA when they had their first clutch, in order to test for great-grandmaternal effects, where great grandmothers (i.e., the F_0) were either young or old when they produced eggs. Each replicate consisted of five adults per jar and there were eight replicates per treatment per generation. All generations were fed *ad libitum* throughout the experiment.

In total, 64 single-end small RNA libraries of 50 bp length were prepared with the CleanTag Small RNA Library kit (16 each F_0 adults, F_0 eggs , F_1 adults, and F_3 adults) and sequenced at Edinburgh Genomics (Edinburgh, United Kingdom) to 50 bp length [34]. Raw sequence data were deposited under bioproject PRJEB22591 in the European Nucleotide Archive.

Identification of *D. magna* piRNAs and differential expression analysis

Raw reads were trimmed with fastp (v0.20.1) under defaults and reads 21 bp longer and 35 bp or under retained and converted to fasta using seqret (EMBOSS v6.5.7.0). Reads were aligned to all *Daphnia* sequences in the RFAM database (release 14.5) and to pre-miRNA sequences identified in [34, 63] using SortMeRNA (v4.3.3). The filtered reads were classified as piRNA or not in piRNN using the *Drosophila melanogaster* trained model [64]. This classifier is based on a convolutional neural network framework and was recommended for use with none model organisms after comparison with other classifiers [47].
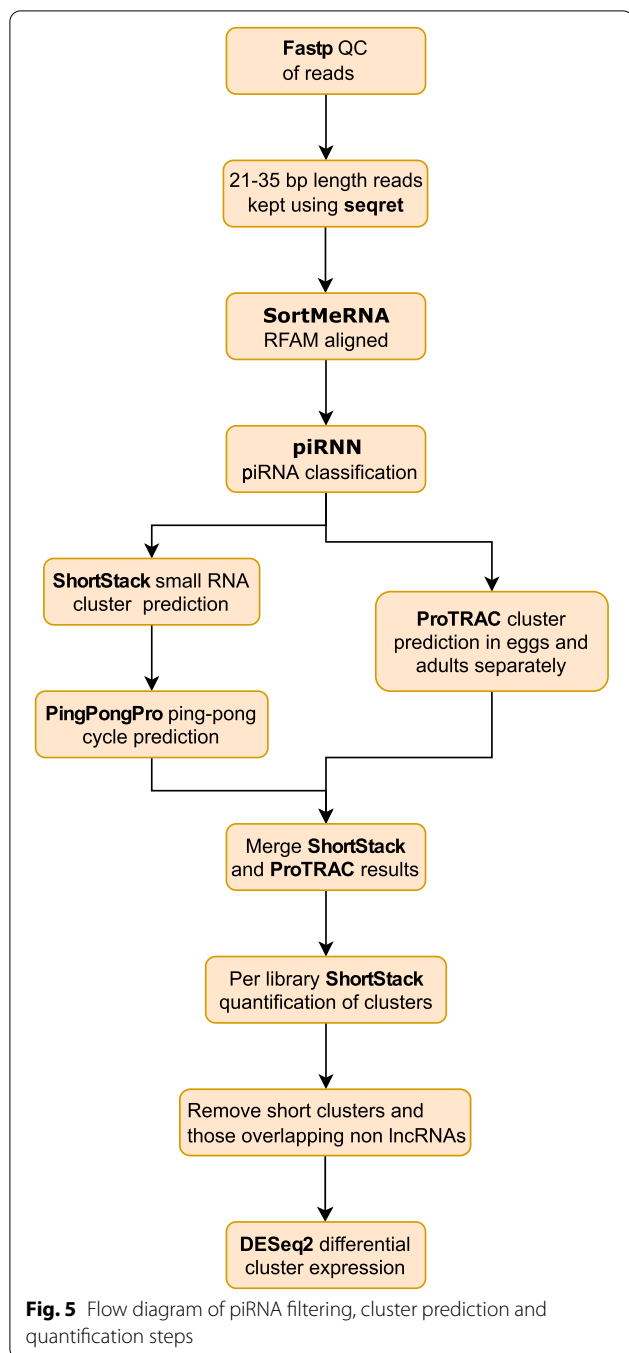
Two approaches were combined to identify piRNA producing loci across the *Daphnia magna* chromosomal genome assembly [27] using reads classified as piRNAs (Fig. 5, piRNA filtering flow diagram). Firstly, piRNA producing clusters were identified for adult (F_0 + F_1 + F_3) and egg (F_0 egg) libraries separately in proTRAC (v2.4.2) [65]. The proTRAC pipeline removes collapses redundant reads ("TBr2_collapse.pl"), removes low-complexity sequences ("TBr2_duster.pl"), aligns reads to the genome (using "sRNAmapper.pl") and reallocates multi-mapping reads according to local transcription levels ("reallocate.pl") before the proTRAC algorithm itself is run. Each step was run with the proTRAC v2.4.2 documentation example settings. The second approach was to align the piRNA-classified reads to the genome using the small RNA aligner ShortStack
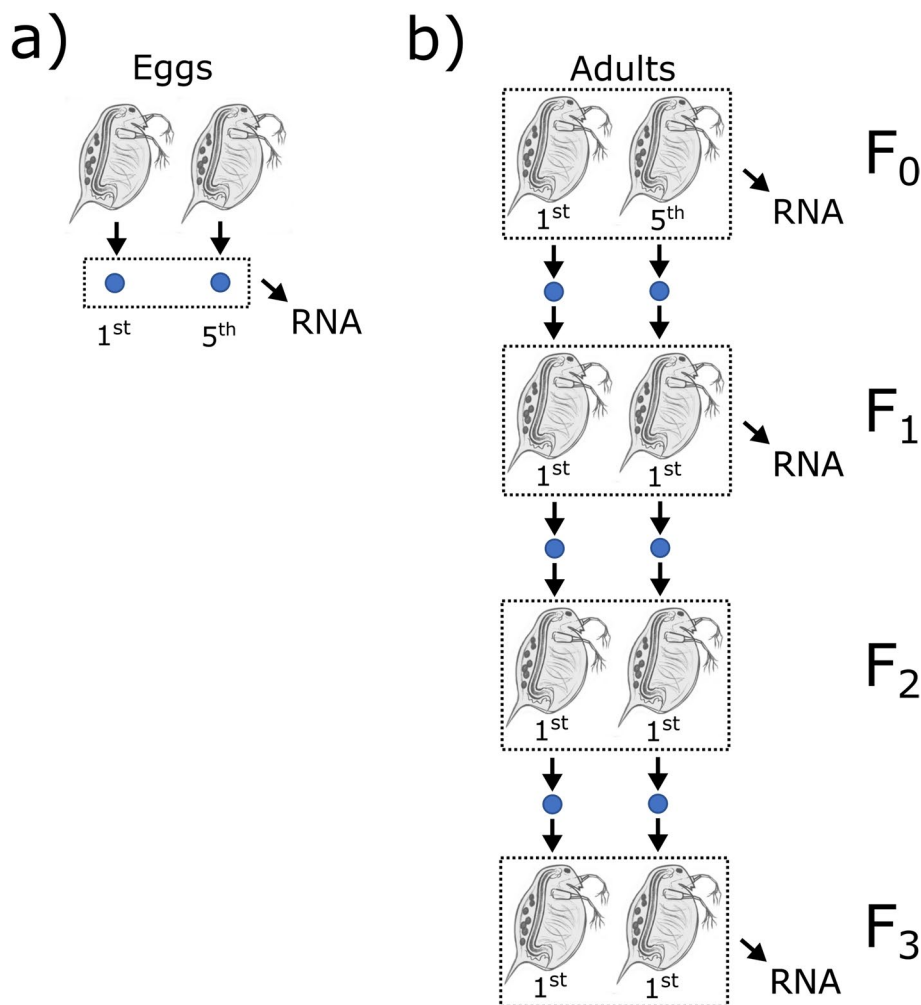
Fig. 4 Schematic showing design for experiments to generate sRNA pools. **a** Eggs and **b** adults of different ages (F_0) or whose parents or great grandparents were of different ages (F_1 and F_3). Numbers below Daphnia mothers/eggs indicate the clutch sampled from for that generation. Adapted from Hearn et al [34], Fig. 2

(v3.8.5) [66] allowing the maximum two mismatches per mapping to account for differences between the genome assembly and strain C32. Each ShortStack-defined cluster was checked for the characteristic 10 bp 1U and 10A overlap (the ping-pong signature) in read-alignment stacks with in PingPongPro (v1.0), and any such sequences combined with others if within a range of 1000 bp of one another (option "T 1000"). ShortStack clusters with significant ping-pong signals (adjusted *p*-value <- 0.05) were combined with the Adult and Egg proTRAC predictions using bedtools merge (v2.23.0). Reads were then re-aligned to the genome in Short-Stack using the combined piRNA cluster co-ordinates to quantify each cluster per library for differential expression analysis. Finally, clusters were removed if they overlapped another species of RNA in the *Daphnia magna* genome annotation (defined as "guide_RNA", "rRNA",

"snoRNA", and "snRNA" in the assembly annotation file "GCF_003990815.1_ASM399081v1_genomic.gff") and, to avoid undetected miRNAs, if the average read length aligned to a cluster was less than 24 bp. Differential gene expression between 1st and 5th clutches was performed on counts per cluster per library separately for eggs and adult generations in the R Bioconductor package DESeq2 (v1.32.0) [67]. *P*-values were adjusted using Independent Hypothesis Weighting [68] in the R Bioconductor package IHW (1.20.0) as part of the DESeq2 workflow, and a significance threshold for adjusted *p*-values of < 0.05 applied. For comparison of expression between piRNA clusters, counts were converted into transcripts per million (TPM) values using "counts_to_TPM.R" (https://gist.github.com/slowkow/c6ab0348747f86e2748b). Mean fragment length per cluster used to calculate TPM was the average read length of reads aligned to that cluster.

Fig. 5 Flow diagram of piRNA filtering, cluster prediction and quantification steps

not appropriate to perform a DESeq2 analysis due to the distinct and time-separated library preparation between eggs and adults undertaken in [34]. A flow chart of bioinformatic steps from raw-reads to DESeq2 differential expression testing was created using https://app.diagrams.net/. $\log_2$-fold change clustered heatmaps of piRNA loci were created using the R package pheatmap (https://CRAN.R-project.org/package=pheatmap) for all differentially expressed piRNA loci across each comparison and for piRNA clusters with average TPM greater than 1000 in egg replicates. Two heatmaps were created to aid interpretation. The first heatmap shaded piRNAs by log2-fold change alone. The second combined clusters by ten log-fold change quantiles containing equal numbers of clusters in order to show direction of change for each cluster. Venn diagrams intersecting genes upregulated in 1st and 5th clutches between eggs, F_0 and F_1 adults were created using https://www.molbiotools.com/listcompare.php. R code for differential expression analyses and TPM count generation is given in File S5.

Annotating piRNA clusters

In order to identify piRNA regulated transposable elements, the *D. magna* genome was annotated with RepeatModeler (v2.0.2a) and RepeatMasker (4.1.2) [69, 70] using predicted repeats and Dfam database transposable elements (version 02/09/2020). The Repeatmasker GFF (General Feature Format) file was then filtered to remove tRNAs, simple and low-complexity repeats prior to intersection with piRNA cluster locations. A second round of filtering removed all RepeatModeler families that were "unknown" to produce a more stringent set of TE predictions. All piRNA clusters were also annotated independently using eggNOG-mapper.

Abbreviations

Ago3: Argonaute 3; Aub: Aubergine; C32: *Daphnia magna* Clone 32; GFF: General Feature Format; lncRNA: Long non-coding RNA; miRNA: microRNA; PCA: Principal component analysis; Piwi: P-element induced wimpy testis; piRNA: Piwi-interacting RNA; RNA: Ribonucleic acid; sRNA: Small RNA; TDRD6: Tudor Domain Containing 6; TE: Transposable element; TPM: Transcripts per million.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-022-08660-z.

Additional file 1: Figure S1. Clustered heatmaps for differentially expressed piRNA clusters with a TPM > 1000 in egg replicates across F0, egg and F1 generations.

Additional file 2: Table S1. Raw and filtered read counts per library.

Additional file 3: Table S2. Total and proTRAC-predicted piRNA cluster normalised read counts per library.

Additional file 4: Table S3. EggNOG annotations for clusters included in the experiment.

Unlike most single-end sequencing experiments, the length of the molecule of cDNA sequenced (fragment length) corresponds to the RNA sequenced and was shorter than the read-length sequenced (50 bp), in this case putative piRNAs of approximately 26 bp. Differences in TPM values averaged across replicates were used to compare egg and combined (F0, F1 and F3) adult differences to identify tissue-biased piRNA clusters. It was

Availability of data and materials
The datasets generated and/or analysed during the current study are available in the European Nucleotide Archive repository under bioproject PRJEB22591, https://www.ebi.ac.uk/ena/browser/view/PRJEB22591.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Department of Vector Biology, Liverpool School of Tropical Medicine, Liverpool, UK. [2]Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, Edinburgh, UK.

References
1. Brennecke J, Malone CD, Aravin AA, Sachidanandam R, Stark A, Hannon GJ. An epigenetic role for maternally inherited piRNAs in transposon silencing. Science. 2008;322:1387–92.
2. Ozata DM, Gainetdinov I, Zoch A, O'Carroll D, Zamore PD. PIWI-interacting RNAs: small RNAs with big functions. Nat Rev Genet. 2019;20:89–108.
3. Lewis SH, Quarles KA, Yang Y, Tanguy M, Frézal L, Smith SA, et al. Pan-arthropod analysis reveals somatic piRNAs as an ancestral defence against transposable elements. Nat Ecol Evol. 2018;2:174–81. https://doi.org/10.1038/s41559-017-0403-4.
4. Jehn J, Gebert D, Pipilescu F, Stern S, Kiefer JST, Hewel C, et al. PIWI genes and piRNAs are ubiquitously expressed in mollusks and show patterns of lineagespecific adaptation. Commun Biol. 2018;1:137. https://doi.org/10.1038/s42003-018-0141-4.
5. Waldron FM, Stone GN, Obbard DJ. Metagenomic sequencing suggests a diversity of RNA interference-like responses to viruses across multicellular eukaryotes. PLoS Genet. 2018;14:e1007533.
6. Czech B, Hannon GJ. One Loop to Rule Them All: The Ping-Pong Cycle and piRNA-Guided Silencing. Trends Biochem Sci. 2016;41:324–37. https://doi.org/10.1016/j.tibs.2015.12.008.
7. Kiuchi T, Koga H, Kawamoto M, Shoji K, Sakai H, Arai Y, et al. A single female-specific piRNA is the primary determiner of sex in the silkworm. Nature. 2014;509:633–6.
8. Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, et al. Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. Cell. 2007;128:1089–103.
9. Mohn F, Sienski G, Handler D, Brennecke J. The rhino-deadlock-cutoff complex licenses noncanonical transcription of dual-strand piRNA clusters in *Drosophila*. Cell. 2014;157:1364–79.
10. Gainetdinov I, Colpan C, Arif A, Cecchini K, Zamore PD. A single mechanism of biogenesis, initiated and directed by PIWI proteins, explains piRNA production in most animals. Mol Cell. 2018;71:775–90.
11. Mérel V, Boulesteix M, Fablet M, Vieira C. Transposable elements in *Drosophila*. Mob. DNA. 2020;11:23. https://doi.org/10.1186/s13100-020-00213-z.
12 Grentzinger T, Armenise C, Brun C, Mugat B, Serrano V, Pelisson A, et al. piRNA-mediated transgenerational inheritance of an acquired trait. Genome Res. 2012;22(10):1877–88.
13. Anzelon TA, Chowdhury S, Hughes SM, Xiao Y, Lander GC, MacRae IJ. Structural basis for piRNA targeting. Nature. 2021;597:285–9. https://doi.org/10.1038/s41586-021-03856-x.
14. Gilbert C, Peccoud J, Cordaux R. Transposable Elements and the Evolution of Insects. Annu Rev Entomol. 2021;66:355–72. https://doi.org/10.1146/annurev-ento-070720-074650.
15. Sturm Á, Perczel A, Ivics Z, Vellai T. The Piwi-piRNA pathway: road to immortality. Aging Cell. 2017;16:906–11.
16. Wood JG, Jones BC, Jiang N, Chang C, Hosier S, Wickremesinghe P, et al. Chromatin-modifying genetic interventions suppress age-associated transposable element activation and extend life span in *Drosophila*. Proc Natl Acad Sci. 2016;113:11277–82. https://doi.org/10.1073/pnas.1604621113.
17. Erwin AA, Blumenstiel JP. Aging in the *Drosophila* ovary: contrasting changes in the expression of the piRNA machinery and mitochondria but no global release of transposable elements. BMC Genomics. 2019;20:1–15.
18 Elsner D, Meusemann K, Korb J. Longevity and transposon defense, the case of termite reproductives. Proc Natl Acad Sci. 2018;115:5504–9. https://doi.org/10.1073/pnas.1804046115.
19. Chen H, Zheng X, Xiao D, Zheng Y. Age-associated de-repression of retrotransposons in the *Drosophila* fat body, its potential cause and consequence. Aging Cell. 2016;15:542–52.
20. Li W, Prazak L, Chatterjee N, Grüninger S, Krug L, Theodorou D, et al. Activation of transposable elements during aging and neuronal decline in *Drosophila*. Nat Neurosci. 2013;16:529–31.
21. Jones BC, Wood JG, Chang C, Tam AD, Franklin MJ, Siegel ER, et al. A somatic piRNA pathway in the *Drosophila* fat body ensures metabolic homeostasis and normal lifespan. Nat Commun. 2016;7:13856. https://doi.org/10.1038/ncomms13856.
22. Treiber CD, Waddell S. Resolving the prevalence of somatic transposition in *Drosophila*. Elife. 2017;6:e28297.

23   Fabian DK, Dönertaş HM, Fuentealba M, Partridge L, Thornton JM. Transposable Element Landscape in *Drosophila* Populations Selected for Longevity. Genome Biol Evol. 2021;13:evab031. https://doi.org/10.1093/gbe/evab031.

24   Giribet G, Edgecombe GD. The Phylogeny and Evolutionary History of Arthropods. Curr Biol. 2019;29:R592-602. https://doi.org/10.1016/j.cub.2019.04.057.

25   Lozano-Fernandez J, Giacomelli M, Fleming JF, Chen A, Vinther J, Thomsen PF, et al. Pancrustacean Evolution Illuminated by Taxon-Rich Genomic-Scale Data Sets with an Expanded Remipede Sampling. Genome Biol Evol. 2019;11:2055–70. https://doi.org/10.1093/gbe/evz097.

26   Eads BD, Tsuchiya D, Andrews J, Lynch M, Zolan ME. The spread of a transposon insertion in *Rec8* is associated with obligate asexuality in *Daphnia*. Proc Natl Acad Sci. 2012;109:858–63. https://doi.org/10.1073/pnas.1119667109.

27   Lee B-Y, Choi B-S, Kim M-S, Park JC, Jeong C-B, Han J, et al. The genome of the freshwater water flea *Daphnia magna*: A potential use for freshwater molecular ecotoxicology. Aquat Toxicol. 2019;210:69–84.

28   Ye Z, Xu S, Spitze K, Asselman J, Jiang X, Ackerman MS, et al. A new reference genome assembly for the microcrustacean Daphnia pulex. G3. 2017;7:1405–16.

29   Ikeda KT, Hirose Y, Hiraoka K, Noro E, Fujishima K, Tomita M, et al. Identification, expression, and molecular evolution of microRNAs in the "living fossil" *Triops cancriformis* (tadpole shrimp). RNA. 2015;21:230–42.

30   Schurko AM, Logsdon JM, Eads BD. Meiosis genes in *Daphnia pulex* and the role of parthenogenesis in genome evolution. BMC Evol Biol. 2009;9:78. https://doi.org/10.1186/1471-2148-9-78.

31   Lu H, Tanguy S, Rispe C, Gauthier J-P, Walsh T, Gordon K, et al. Expansion of genes encoding piRNA-associated argonaute proteins in the pea aphid: diversification of expression profiles in different plastic morphs. PLoS One. 2011;6:e28051.

32   Lewis SH, Webster CL, Salmela H, Obbard DJ. Repeated Duplication of Argonaute2 Is Associated with Strong Selection and Testis Specialization in *Drosophila*. Genetics. 2016;204:757–69. https://doi.org/10.1534/genetics.116.192336.

33   Crysnanto D, Obbard DJ. Widespread gene duplication and adaptive evolution in the RNA interference pathways of the *Drosophila obscura* group. BMC Evol Biol. 2019;19:99.

34   Hearn J, Chow FW-N, Barton H, Tung M, Wilson P, Blaxter M, et al. *Daphnia magna* microRNAs respond to nutritional stress and ageing but are not transgenerational. Mol Ecol. 2018;27:1402–12.

35   Hearn J, Pearson M, Blaxter M, Wilson PJ, Little TJ. Genome-wide methylation is modified by caloric restriction in *Daphnia magna*. BMC Genomics. 2019;20:197. https://doi.org/10.1186/s12864-019-5578-4.

36   Hearn J, Plenderleith F, Little TJ. DNA methylation differs extensively between strains of the same geographical origin and changes with age in *Daphnia magna*. Epigenetics Chromatin. 2021;14:4. https://doi.org/10.1186/s13072-020-00379-z.

37   Hearn J, Clark J, Wilson PJ, Little TJ. *Daphnia magna* modifies its gene expression extensively in response to caloric restriction revealing a novel effect on haemoglobin isoform preference. Mol Ecol. 2020;29:3261–76. https://doi.org/10.1111/mec.15557.

38   Pandey RR, Tokuzawa Y, Yang Z, Hayashi E, Ichisaka T, Kajita S, et al. Tudor domain containing 12 (TDRD12) is essential for secondary PIWI interacting RNA biogenesis in mice. Proc Natl Acad Sci. 2013;110:16492–7.

39   Sato K, Iwasaki YW, Siomi H, Siomi MC. Tudor-domain containing proteins act to make the piRNA pathways more robust in *Drosophila*. Fly (Austin). 2015;9:86–90.

40   Megosh HB, Cox DN, Campbell C, Lin H. The role of PIWI and the miRNA machinery in *Drosophila* germline determination. Curr Biol. 2006;16:1884–94.

41   Guzzardo PM, Muerdter F, Hannon GJ. The piRNA pathway in flies: highlights and future directions. Curr Opin Genet Dev. 2013;23:44–52.

42   Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V, Zamore PD. A distinct small RNA pathway silences selfish genetic elements in the germline. Science. 2006;313:320–4.

43   Zhang Z, Wang J, Schultz N, Zhang F, Parhad SS, Tu S, et al. The HP1 homolog rhino anchors a nuclear complex that suppresses piRNA precursor splicing. Cell. 2014;157:1353–63.

44   Waiho K, Fazhan H, Zhang Y, Li S, Zhang Y, Zheng H, et al. Comparative profiling of ovarian and testicular piRNAs in the mud crab *Scylla paramamosain*. Genomics. 2020;112:323–31.

45   Quénerch'du E, Anand A, Kai T. The piRNA pathway is developmentally regulated during spermatogenesis in *Drosophila*. Rna. 2016;22:1044–54.

46   Li L, Wang S, Huang K, Zhang Y, Li Y, Zhang M, et al. Identification and Characterization of MicroRNAs in Gonads of *Helicoverpa armigera* (Lepidoptera: Noctuidae). Insects. 2021;12:749.

47   Liu Y, Li A, Xie G, Liu G, Hei X. Computational methods and online resources for identification of pirna-related molecules. Interdiscip Sci Comput Life Sci. 2021;13:176–91. https://doi.org/10.1007/s12539-021-00428-5.

48   Czech B, Munafò M, Ciabrelli F, Eastwood EL, Fabry MH, Kneuss E, et al. piRNA-guided genome defense: from biogenesis to silencing. Annu Rev Genet. 2018;52:131–57.

49   Ge DT, Wang W, Tipping C, Gainetdinov I, Weng Z, Zamore PD. The RNA-binding ATPase, Armitage, couples piRNA amplification in nuage to phased piRNA production on mitochondria. Mol Cell. 2019;74:982–95.

50   Fridrich A, Hazan Y, Moran Y. Too Many False Targets for MicroRNAs: Challenges and Pitfalls in Prediction of miRNA Targets and Their Gene Ontology in Model and Non-model Organisms. Bioessays. 2019;41:1800169.

51   Pinzón N, Li B, Martinez L, Sergeeva A, Presumey J, Apparailly F, et al. microRNA target prediction programs predict many false positives. Genome Res. 2017;27:234–45.

52   Zuo Y, Liang Y, Zhang J, Hao Y, Li M, Wen Z, et al. Transcriptome analysis identifies piwi-interacting RNAs as prognostic markers for recurrence of prostate cancer. Front Genet. 2019;10:1018.

53   Weng W, Liu N, Toiyama Y, Kusunoki M, Nagasaka T, Fujiwara T, et al. Novel evidence for a PIWI-interacting RNA (piRNA) as an oncogenic mediator of disease progression, and a potential prognostic biomarker in colorectal cancer. Mol Cancer. 2018;17:1–12.

54   Gou L-T, Dai P, Yang J-H, Xue Y, Hu Y-P, Zhou Y, et al. Pachytene piRNAs instruct massive mRNA elimination during late spermiogenesis. Cell Res. 2014;24:680–700.

55   Ghosheh Y, Seridi L, Ryu T, Takahashi H, Orlando V, Carninci P, et al. Characterization of piRNAs across postnatal development in mouse brain. Sci Rep. 2016;6:1–9.

56   Roy J, Sarkar A, Parida S, Ghosh Z, Mallick B. Small RNA sequencing revealed dysregulated piRNAs in Alzheimer's disease and their probable role in pathogenesis. Mol Biosyst. 2017;13:565–76.

57   Liu Y, Zhang J, Li A, Liu Z, He Z, Yuan X, et al. Prediction of cancer-associated piRNA-mRNA and piRNA-lncRNA interactions by integrated analysis of expression and sequence data. Tsinghua Sci Technol. 2018;23:115–25.

58   Zhang D, Tu S, Stubna M, Wu W-S, Huang W-C, Weng Z, et al. The piRNA targeting rules and the resistance to piRNA silencing in endogenous genes. Science. 2018;359:587–92.

59   Shen E-Z, Chen H, Ozturk AR, Tu S, Shirayama M, Tang W, et al. Identification of piRNA binding sites reveals the argonaute regulatory landscape of the C. elegans germline. Cell. 2018;172:937–51.

60   Halbach R, Miesen P, Joosten J, Tacsköprü E, Rondeel I, Pennings B, et al. A satellite repeat-derived piRNA controls embryonic development of *Aedes*. Nature. 2020;580:274–7.

61   Auld SKJR, Hall SR, Housley Ochs J, Sebastian M, Duffy MA. Predators and patterns of within-host growth can mediate both among-host competition and evolution of transmission potential of parasites. Am Nat. 2014;184:S77-90.

62   Garbutt JS, Little TJ. Bigger is better: changes in body size explain a maternal effect of food on offspring disease resistance. Ecol Evol. 2017;7:1403–9.

63   Coucheron DH, Wojewodzic MW, Bøhn T. MicroRNAs in *Daphnia magna* identified and characterized by deep sequencing, genome mapping and manual curation. Sci Rep. 2019;9:15945. https://doi.org/10.1038/s41598-019-52387-z.

64   Wang K, Hoeksema J, Liang C. piRNN: deep learning algorithm for piRNA prediction. PeerJ. 2018;6:e5429–e5429. https://doi.org/10.7717/peerj.5429.

65   Rosenkranz D, Zischler H. proTRAC–a software for probabilistic piRNA cluster detection, visualization and analysis. BMC Bioinformatics. 2012;13:5. https://doi.org/10.1186/1471-2105-13-5.

66   Johnson NR, Yeoh JM, Coruh C, Axtell MJ. Improved Placement of Multimapping Small RNAs. G3. 2016;6:2103–11.

67. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. Genome Biol. 2014;15:550.
68. Ignatiadis N, Klaus B, Zaugg JB, Huber W. Data-driven hypothesis weighting increases detection power in genome-scale multiple testing. Nat Methods. 2016;13:577–80.
69. Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, et al. RepeatModeler2 for automated genomic discovery of transposable element families. Proc Natl Acad Sci. 2020;117:9451–7. https://doi.org/10.1073/pnas.1921046117.
70. Smit A, Hubley R, Green P. RepeatMasker Open-4.0. 2013–2015. http://www.repeatmasker.org.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.