Insights into mechanisms of fertility through study of *Caenorhabditis elegans* piRNAs

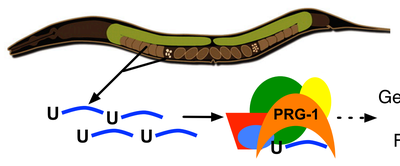
Mallory Freeberg[1,2], Allison C. Billi[2], and John K. Kim[2]
[1]Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI
[2]Life Sciences Institute and Department of Human Genetics, University of Michigan, Ann Arbor, MI

Abstract

piRNAs are a class of short interfering RNAs (siRNAs) characterized by 21 nt length, 5' uridine, and no sequence similarity or conservation.[1] piRNAs require Piwi protein PRG-1 for accumulation and are involved in germline maintenance and fertility.[2-4] piRNAs suppress harmful genetic elements, including transposable elements, by priming biogenesis of secondary siRNAs that target these elements for degradation.[5-8] piRNAs map primarily to two broad regions of chromosome IV, but their mechanism of biogenesis remains largely unknown. A recently characterized class of germline endogenous siRNAs, 26G RNAs, show distinct expression in either spermatogenic cells or oocytes and embryos.[9,10] We wondered whether piRNAs also show patterns of male or female germline specificity. Although some piRNAs have been detected in both male and female germlines,[2,3] what might distinguish germline-specific patterns of enrichment among all piRNAs is poorly understood. Here, we computationally analyze published sequencing datasets to assess germline specificity of piRNAs. More than 70% of piRNAs are >5-fold enriched in male or female germline. Initial piRNA studies identified a short 8 nt motif (CTGTTTCA) located 46 nt upstream of piRNA loci.[1,2] Strikingly, 80% of male piRNA upstream regions contain the 5 nt core motif GTTTC, compared to less than half for female piRNAs. Additionally, position 1 of the short motif is enriched for C upstream of male piRNAs, which is associated with more robust male piRNA expression; no such pattern is observed for female piRNAs. By generating transgenic *C. elegans* animals expressing a synthetic piRNA sequence, we show that placing a C-containing short motif upstream of a female piRNA alters its expression pattern to look more like a male piRNA. Furthermore, placing a non-C-containing motif upstream of a male piRNA alters its expression pattern to look more like a female piRNA. These data suggest that the upstream motif orchestrates germline expression patterns of piRNAs in *C. elegans*.

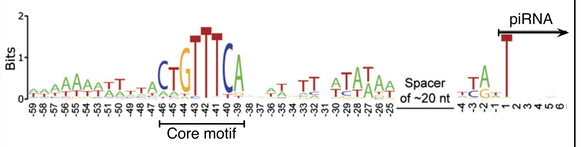
Background

piRNAs expressed in germline and require PRG-1.[1-4]



Tc3 repression
Germ cell proliferation/ maintenance
Fertility at elevated temperatures

piRNA-coding loci have distinct features.[1-4]



- Conserved 5' U (encoded by T in genome)
- Conserved 8 nt upstream core motif
- 21 nt long

- How can sequencing data from multiple sources be integrated to discover biologically relevant patterns?

References

1. Ruby JG et al. *Cell* **127**:1193 (2006) 2. Batista PJ et al. *Mol Cell* **31**:67 (2008). 3. Das PP et al. *Mol Cell* **31**:79 (2008). 4. Wang G et al. *Curr Biol* **18**:861 (2008). 5. Bagijn MP et al. *Science* **337**:574 (2012). 6. Lee H et al. *Cell* **150**:78 (2012). 7. Shirayama M et al. *Cell* **150**:65 (2012). 8. Ashe A et al. *Cell* **150**:88 (2012). 9. Conine CC et al. *PNAS* **107**:3588 (2010). 10. Han T et al. *PNAS* **106**:18674 (2009).

Funding

MAF was supported by the National Science Foundation Open Data IGERT Grant 0903629. ACB was supported by the NIH Genetics Training Grant graduate fellowship T32GM007544-34.

Methods

1. Download public small RNA sequencing datasets from Gene Expression Omnibus

- Sequences mapped to known piRNAs using Bowtie
- piRNA counts normalized to reads per million mapped reads (RPM)

2. Calculate Enrichment score for each piRNA

Small RNA sequencing
14 male libraries
3 female libraries
56.4M total reads -- 1.3M piRNA reads

piRNA filters
15,093 annotated piRNAs
13,711 piRNAs more highly expressed in N2 compared to *prg-1(tm872)*

Enrichment score calculation
Enrichment score for each piRNA initialized to 0

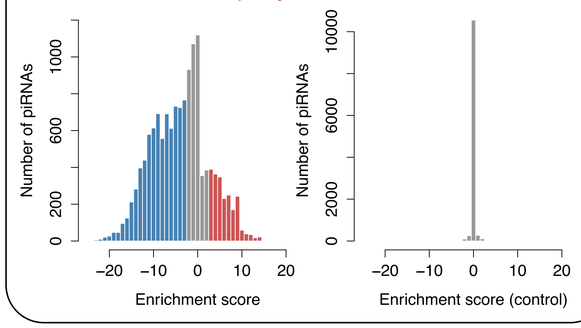
5'-monoP$_i$ dependent
5 male vs. 1 female (5 comparisons)

5'-monoP$_i$ independent
9 male vs. 2 female (18 comparisons)

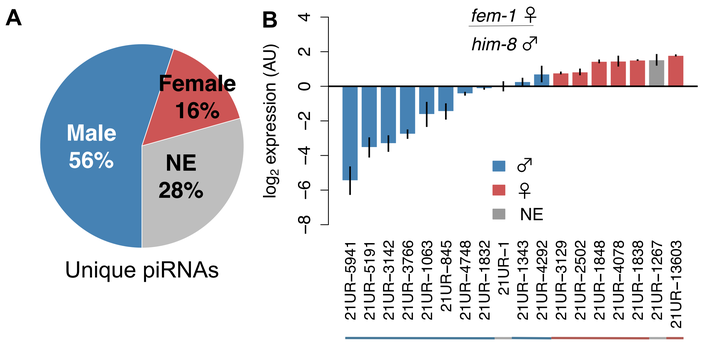
In each comparison for each piRNA:
>5-fold abundant in male: score - 1
>5-fold abundant in female: score +1

3. Classify piRNAs as germline-enriched

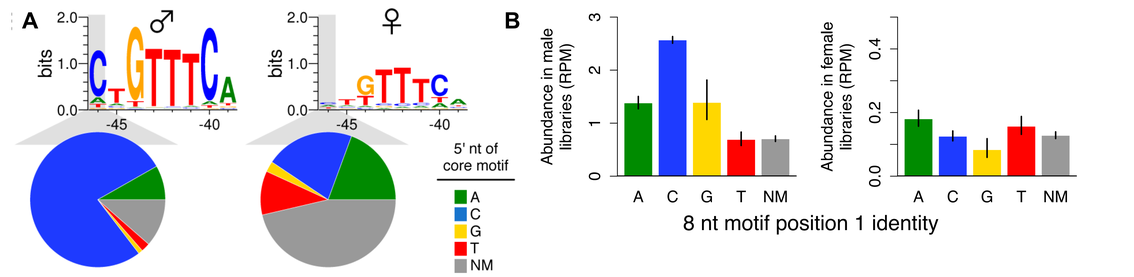
Male (♂) piRNAs: score ≤ -3
Non-enriched (NE) piRNAs: -3 > score > 3
Female (♀) piRNAs: score ≥ 3



Results

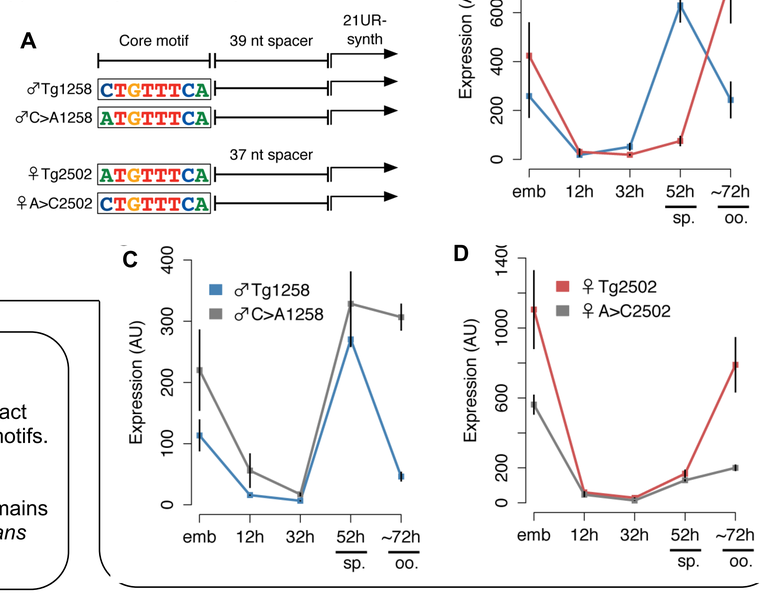


More than 70% of piRNAs are male or female germline-enriched.
(**A**) piRNAs are classified as male germline-enriched (blue), female germline-enriched (red), or not enriched (grey) by a computational meta-analysis of published sequencing datasets.
(**B**) Germline classification of piRNAs was validated by Taqman RT-qPCR. Relative expression of representative piRNAs were assayed in *fem-1(hc17)* female and *him-8 (e1489)* male animals. Error bars: +/- 1 SD of two biological replicates. NE: not enriched. AU: arbitrary units.



Male and female piRNAs have distinct regulatory motifs that correlate with piRNA abundance.
(**A**) Top: Male, but not female, piRNA loci are enriched for 8 nt motifs with C at position 1. Bottom: Significantly fewer female piRNA loci have a motif than male piRNA loci (compare colored to grey areas in pie charts). NM: no motif. (**B**) Male piRNAs downstream of CNGTTTCN motifs are significantly more abundant than all other male piRNAs (Student's two-tailed *t*-test p-values <0.01). Female piRNA abundance is not significantly different with any motif (p-values >0.05).

Position 1 of motif orchestrates piRNA germline expression patterns.
(**A**) Upstream motif of male piRNA Tg1258 has C at position 1 which is mutated to A (C>A1258). Upstream motif of female piRNA Tg2502 has A at position 1 which is mutated to C (A>C2502). (**B**) Male piRNA 1258 peaks at 52 hours during male germline development (spermatogenesis, sp). Female piRNA 2502 peaks at 72h during female germline development (oogenesis, oo.). emb: embryo; AU: arbitrary units. Error bars: +/- 1 SD of two biological replicates. (**C**) Mutating male piRNA upstream motif from C>A results in sustained expression during oogenesis (compare at 72h). (**D**) Mutating female piRNA upstream motif from A>C results in lack of increase in expression during oogenesis (compare at 72h).



Future Directions

What factors recognize the upstream motif to orchestrate male and female piRNA expression?
Some transcription factors (TFs) have been proposed to interact with the upstream motif; however, *in vivo* evidence of this interaction is lacking. Our lab is currently investigating a family of TFs that could potentially interact with the motif. Different members of this family may interact differently with C-containing and non-C-containing motifs.

How does this mechanism of piRNA biogenesis reflect mechanisms in higher organisms?
Higher organisms, such as flies and mammals, also exhibit germline-specific expression of piRNAs. What remains elusive is how germline-specific expression is achieved. A complete understanding of the mechanism in *C. elegans* piRNA expression could provide insight into mechanisms in higher organisms.