Supporting Online Material for

Developmentally Regulated piRNA Clusters Implicate MILI in Transposon Control

Alexei A. Aravin, Ravi Sachidanandam, Angelique Girard,
Katalin Fejes-Toth, Gregory J. Hannon*

*To whom correspondence should be addressed. E-mail: hannon@cshl.edu

Published 19 April 2007 on *Science* Express

DOI: 10.1126/science.1142612

This PDF file includes:

Materials and Methods

Mouse strains

CD-1 wild-type mice of specific ages were purchased from Charles River Lab and used for purification of Mili RNP complexes and isolation of testis RNA for Northern hybridization. The Mili knock-out strain was obtained from Haifan Lin (Yale University) and is described in (*1*).

Antibody, immunoprecipitation of Mili RNP complexes and small RNA cloning

MILI peptides, DPVRPLFRGPTPVHPSQ (MILI-N, amino acids 2-18) and VRKDREEPRSSLPDPS (MILI-N2, amino acids 107-122), were conjugated to KLH and used for polyclonal antibody production (Covance). Antibodies were affinity purified on a peptide-conjugated resin (Sulfolink, Pierce Biochemicals). Immunoprecipitation of MILI RNP complexes and RNA isolation and labeling was performed as described in (*2*). For small RNA cloning, immunoprecipitations with MILI-N and MILI-N2 antibodies were performed in parallel. Two cDNA libraries were created and sequenced separately. After preliminary analysis, sequences were pooled. For labeling of MILI-bound RNA, the anti-MILI-N2 antibody was used for immunoprecipitation. Small RNA cloning and 454 sequencing of small cDNA libraries were performed as described in (*3*).

Northern hybridization

Northern hybridization was performed as described in (*4*) using either 10 μg of total testis RNA isolated using Trizol (Invitrogen) or RNA isolated from immunoprecipitated MILI complexes. For MILI-bound RNAs, each northern lane corresponded to an IP from 1-2 testes. The oligonucleotide probes used for hybridization were as follows: pre-pachytene piRNA (corresponds to sense SINE B1): 5'-TGGCTGTCCTGGAACTCACTYTGT, pachytene piRNA (corresponds to the chr.17 piRNA cluster): 5'-TCCTTGTTAGTTCTCACTCGTCTTTTA, let-7 miRNA: 5'-ACTATACAACCTACTACCTCA. For LNA probes, hybridization was carried out at 65 degrees.

Semi-quantitative RT-PCR

For RT-PCR 1 μg of total testis RNA was treated with DNaseI (Invitrogen) and reverse- transcribed using an oligo-dT primer and SuperScriptIII (Invitrogen). PCR on L1

and IAP transcripts was done according to standard procedures using SybrGreen PCR master mix (Applied Boisystems) on a Dyad Disciple PCR unit (BioRad) using 60°C for the annealing and polymerization steps. Amplification with primers corresponding to actin mRNA was performed in parallel and used for normalization. Mili-mutant and wild-type samples from the same litters were used for all comparisons. The following primers were used: actin forward: 5' CGGTTCCGATGCCCTGAGGCTCTT, actin reverse: 5' CGTCACACTTCATGATGGAATTGA; L1 forward: 5' GAGAACATCGGCACAACAATC L1 reverse: 5' TTTATTGGCGAGTTGAGACCA, IAP forward: 5' CAGACTGGGAGGAAGAAGCA, IAP reverse: 5' ATTGTTCCCTCACTGGCAAA.

Analysis of methylation status of genomic L1 copies was performed as described in (*5*) using DNA isolated from testes and tails of 14 day-old mice.

Bioinformatic analysis of piRNAs

Analysis of piRNA sequences was done as described in (*3*). Briefly, piRNAs were cloned and sequenced using 454. Sequencing of the pre-pachytene library produced 180,769 reads, 164,739 of which (91%) could be mapped to the mouse genome (release mm8, Feb 2006) with 0-2 mismatches. This population (pre-pachytene piRNAs) was used for further analysis. For comparison, we used 80,952 sequences from adult piRNAs (pachytene piRNAs), 71,230 of which mapped to the genome with 0-2 mismatches. Previous studies and our unpublished analyses have shown that this is reflective of the content of both Mili and Miwi complexes. The size of pre-pachytene piRNAs was similar to that of MILI-associated pachytene piRNAs with an average size of 26-28 nucleotides. Annotation categories were assigned based on the annotation of corresponding genomic sequences extracted from the UCSC genome browser. For subsequent analysis, including cluster extraction, only sequences that had a perfect match to the mouse genome were considered.

To extract piRNA clusters the genome was scanned using a 10 kB window. Windows that had more than 10 uniquely mapping piRNAs were extracted and neighboring windows were merged. The exact boundaries of the clusters were then identified by starting 10 kB upstream and downstream of each candidate cluster and walking inwards until a piRNA was found that was less than 2 kB from its nearest neighbor. These were considered as the cluster boundaries. Clusters for piRNA mapping

to plus and minus genomic strands were extracted separately and then analyzed for relationships to each other.

To identify sequences that match the consensuses for transposable elements, piRNA were aligned to consensus sequences from release 11.08 of Repbase (www.girinst.org/). The following consensuses were used: B1_Mus1 for B1 SINE, L1_MM (Genbank M29324) for LINE L1 and IAPLTR1a_I_MM for the IAP retrotransposon. Matches of piRNA to consensus sequences with up to three mismatches were recovered and included in the analysis. 10nt offset partners and nucleotide biases were calculated for piRNAs matching B1, L1 and IAP consensuses as described in (*3*).

Supplementary figures and Legends

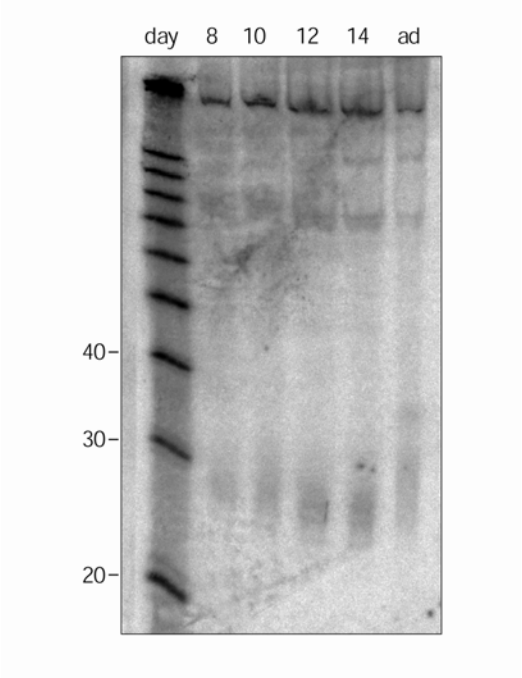

Figure S1. Developmental time course of an IAP piRNA. Total RNA was prepared from testes of mice of the indicated age. Northern blotting was done with an LNA probe (GTCCATCTGACGGCAGAACTGCTGAAA) complementary to a piRNA that matched IAP.

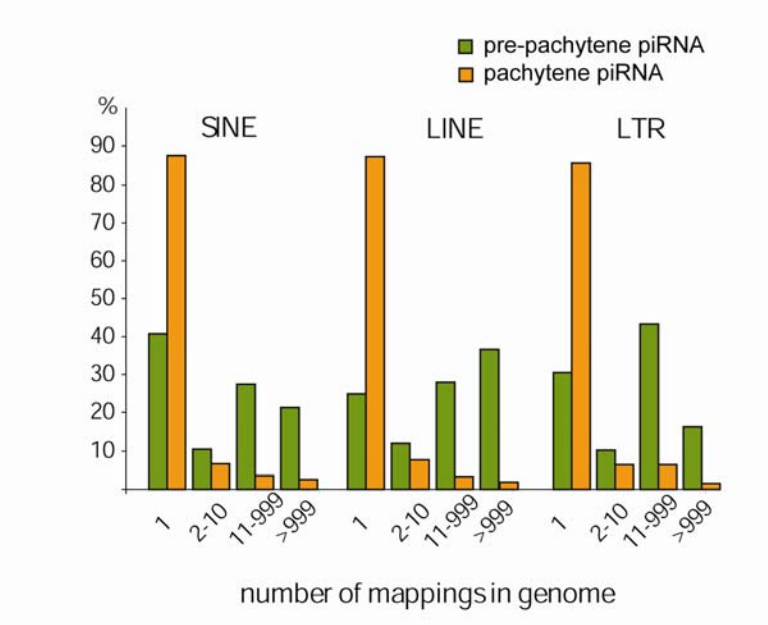

Figure S2. Distribution of pachytene and pre-pachytene piRNAs corresponding to major repeat classes according to their number of genomic mappings.

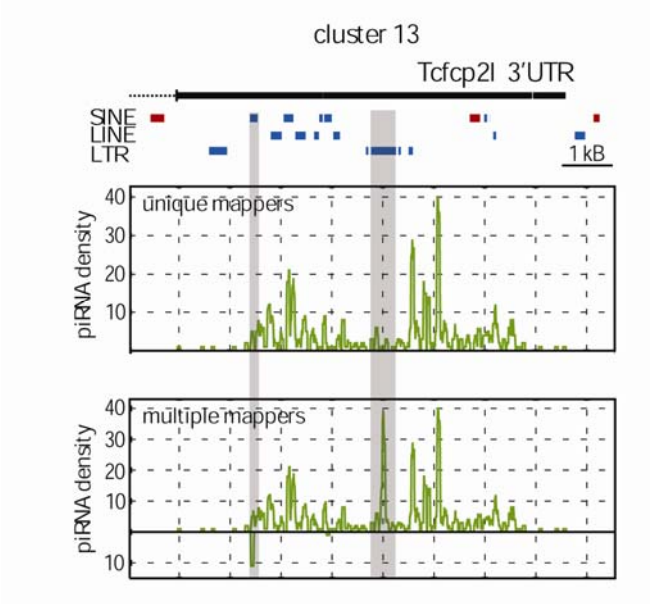

Figure S3. The density of uniquely-mapped or total piRNAs was plotted for a genic pre-pachytene piRNA cluster using a 60 bp sliding window. This cluster corresponds to the 3'UTR of the Tcfcp2l gene and harbors fragments of different transposable elements. Transposable element fragments, separated by type, are indicated as red boxes on the plus genomic strand and as blue boxes on the minus

strand. Some of the transposon fragments have matches to piRNA that have multiple mappings in the genome (shadowed).

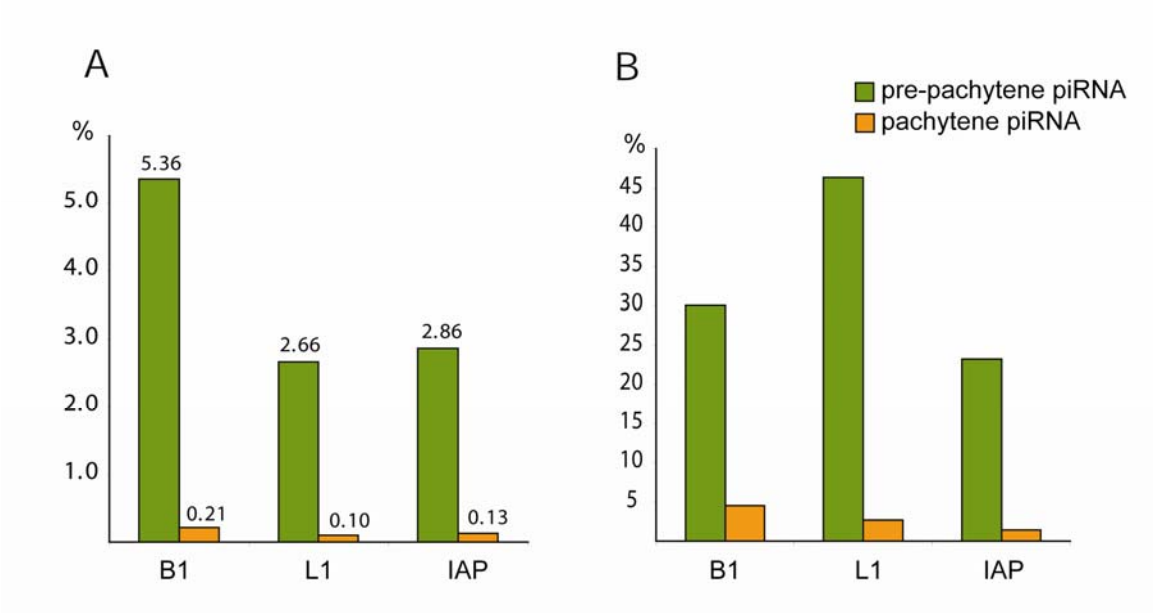

Figure S4. A fraction of pachytene and pre-pachytene piRNAs corresponds to consensus transposon sequences. (A) Fractions of the total piRNA population corresponding to SINE B1, L1 and IAP consensus sequences are shown. A consensus match is defined by a piRNA with up to three mismatches. **(B)** piRNAs were assigned to SINE, LINE and LTR retrotransposon categories based upon the UCSC annotation. The fraction of each category that matched to consensus B1, L1 and IAP sequences with up to three mismatches was plotted, as indicated.

Supplementary Tables and Legends

Table S1

#	Chr	Position, genomic strand	Size, bp	Number of pre-pach. piRNA	Number of pach. piRNA	Annotation
1	chr7	75,642,492 - 75,655,838, -	13,346	1358	3	repeat
2	chr1	133,851,220 - 133,860,030, +	8,810	1108	2	none
3	chr7	6,177,404 - 6,256,525, -	79,121	877	2	repeat
4	chr10	82,790,090 - 82,887,066, +	96,976	711	1	repeat
5	chr1	158,477,753 - 158,490,597, +	12,844	648	3	none
6	chr8	120,091,265 - 120,101,735, +	10,470	641	28	none
7	chr3	95,267,256 - 95,275,985, -	8,729	533	2	repeat
8	chr4	56,961,173 - 56,974,256, -	13,083	493	1	repeat
9	chr11	79,500,448 - 79,514,042, +	13,594	490	4	none
10	chr5	67,671,395 - 67,679,407, -	8,012	479	1	none
11	chr15	103,020,013 - 103,027,778, -	7,765	432	12	intron
12	chr4	107,673,192 - 107,679,793, -	6,601	400	0	intron
13	chr1	120,505,026 - 120,512,597, +	7,571	376	0	exon
14	chr17	27,020,297 - 27,090,241, bi	69,944	363	2431	none
15	chr6	83,328,633 - 83,334,931, -	6,298	312	0	none
16	chr4	134,684,535 - 134,690,030, -	5,495	308	0	none
17	chr9	114,356,025 - 114,362,438, -	6,413	299	2	none
18	chr16	17,170,949 - 17,176,785, +	5,836	297	0	exon
19	chr15	102,107,929 - 102,127,014, +	19,085	294	0	repeat
20	chr17	83,482,678 - 83,510,279, -	27,601	282	0	none
21	chr1	34,757,936 - 34,764,450, -	6,514	278	3	exon
22	chr8	107,237,589 - 107,247,002, -	9,413	272	4	none
23	chr10	60,823,757 - 60,830,390, -	6,633	268	36	none
24	chr1	90,090,090 - 90,097,810, -	7,720	267	1	none
25	chr9	43,894,139 - 43,903,895, -	9,756	267	66	none
26	chr15	83,181,482 - 83,188,123, -	6,641	258	104	none
27	chr6	128,135,608 - 128,166,126, -	30,518	252	160	none
28	chr14	63,766,185 - 63,776,201, -	10,016	250	1	none
29	chr1	162,840,654 - 162,856,978, -	16,324	249	4	none
30	chr9	119,282,055 - 119,289,607, +	7,552	246	1	none
31	chr19	10,049,743 - 10,052,044, +	2,301	244	3	exon
32	chr10	126,394,483 - 126,402,840, +	8,357	240	0	exon
33	chr12	82,633,879 - 82,643,186, -	9,307	233	1	none
34	chr18	60,810,093 - 60,817,376, -	7,283	233	0	exon
35	chr1	172,030,325 - 172,049,163, -	18,838	232	1	none
36	chr15	81,650,732 - 81,654,021, +	3,289	229	0	exon
37	chr17	24,690,047 - 24,692,442, -	2,395	212	1	none
38	chr4	107,727,292 - 107,733,799, -	6,507	211	0	none
39	chr5	3,525,373 - 3,536,813, +	11,440	204	0	none
40	chr8	112,615,680 - 112,622,629, -	6,949	203	2	none
41	chr11	120,441,466 - 120,449,626, -	8,160	200	1	none
42	chr11	97,205,214 - 97,214,156, +	8,942	191	0	exon
43	chr15	89,375,833 - 89,387,924, +	12,091	190	0	intron
44	chr7	63,765,631 - 63,770,015, -	4,384	183	0	none
45	chr16	16,831,275 - 16,843,148, +	11,873	182	1	exon
46	chr16	17,397,339 - 17,400,573, +	3,234	180	0	exon
47	chr18	65,750,674 - 65,762,764, +	12,090	180	0	intron
48	chr12	88,055,528 - 88,063,452, -	7,924	176	1	exon

Table S1

49	chr1	95,642,274 - 95,652,282, +	10,008	175	0	exon
50	chr19	44,621,734 - 44,630,336, +	8,602	174	1	exon
51	chr2	154,222,676 - 154,230,576, +	7,900	172	3	exon
52	chr9	44,042,546 - 44,050,005, -	7,459	169	0	repeat
53	chr11	109,212,267 - 109,216,836, +	4,569	168	2	exon
54	chr7	29,842,606 - 29,848,842, +	6,236	168	44	none
55	chr5	143,132,179 - 143,137,613, -	5,434	167	1	none
56	chr2	28,644,624 - 28,656,948, -	12,324	164	0	none
57	chr1	167,702,059 - 167,711,925, -	9,866	160	2	none
58	chr13	55,321,807 - 55,330,420, +	8,613	156	0	none
59	chr16	24,898,564 - 24,907,454, +	8,890	156	0	none
60	chr14	25,926,547 - 25,933,899, +	7,352	151	0	none
61	chr9	95,989,096 - 95,993,218, +	4,122	151	0	none
62	chr6	120,852,314 - 120,860,725, +/-	8,411	149	3	exon
63	chr4	135,812,545 - 135,821,294, +	8,749	148	0	exon
64	chr6	85,435,439 - 85,450,015, -	14,576	142	0	none
65	chr12	111,710,643 - 111,714,637, +	3,994	141	2	exon
66	chr11	97,490,862 - 97,501,383, +	10,521	139	0	none
67	chr13	55,141,942 - 55,144,453, +	2,511	138	1	exon
68	chr11	98,712,231 - 98,720,202, +	7,971	137	1	exon
69	chr3	51,312,642 - 51,338,292, +	25,650	134	1	intron
70	chr9	27,095,487 - 27,105,823, +	10,336	134	0	none
71	chr10	41,870,382 - 41,874,675, -	4,293	132	0	none
72	chr5	115,404,192 - 115,417,419, -	13,227	129	1	none
73	chr11	60,593,663 - 60,603,926, +	10,263	127	1	none
74	chr4	132,464,195 - 132,471,688, +	7,493	124	2	exon
75	chr9	114,253,503 - 114,261,022, +	7,519	123	0	intron
76	chr4	45,417,318 - 45,421,237, -	3,919	120	0	exon
77	chr4	56,900,267 - 56,903,599, +	3,332	120	6	repeat
78	chr16	13,325,893 - 13,331,028, +	5,135	119	0	none
79	chr19	44,354,546 - 44,359,929, +	5,383	119	0	exon
80	chr11	75,384,589 - 75,399,131, +	14,542	117	4	none
81	chr11	105,870,804 - 105,875,085, +	4,281	117	2	exon
82	chr4	59,993,482 - 59,998,568, +	5,086	114	0	exon
83	chr6	40,331,234 - 40,338,983, -	7,749	114	0	none
84	chr15	81,280,259 - 81,284,591, +	4,332	113	0	repeat
85	chr8	113,890,294 - 113,894,779, -	4,485	113	31	exon
86	chr2	120,384,322 - 120,398,281, +	13,959	111	0	none
87	chr2	33,272,487 - 33,277,113, -	4,626	110	1	exon
88	chr18	33,324,316 - 33,328,485, -	4,169	107	0	exon
89	chr1	95,195,653 - 95,203,598, +	7,945	106	0	repeat
90	chr11	95,580,504 - 95,587,832, +	7,328	105	10	none
91	chr11	117,530,316 - 117,548,410, +	18,094	105	1	none
92	chr15	102,261,049 - 102,263,769, +	2,720	105	0	none
93	chr6	83,772,082 - 83,775,488, -	3,406	105	1	exon
94	chr7	27,893,829 - 27,905,917, +	12,088	105	0	none
95	chr8	114,040,708 - 114,044,123, -	3,415	103	0	none
96	chr10	77,591,109 - 77,596,362, -	5,253	102	1	none
97	chr17	83,520,533 - 83,540,167, -	19,634	102	0	repeat
98	chr2	26,711,197 - 26,730,086, -	18,889	102	0	exon
99	chr8	109,900,441 - 109,908,244, +	7,803	102	2	none
100	chr13	64,202,207 - 64,205,825, -	3,618	101	0	none

Literature Cited

1. S. Kuramochi-Miyagawa *et al.*, *Development* 131, 839 (Feb, 2004).
2. A. Aravin *et al.*, *Nature* 442, 203 (Jul 13, 2006).
3. J. Brennecke *et al.*, *Cell* In Press (2007).
4. A. Girard, R. Sachidanandam, G. J. Hannon, M. A. Carmell, *Nature* 442, 199 (Jul 13, 2006).
5. M. A. Carmell *et al.*, *Dev Cell* in press (2007).