TÜBİTAK

Turkish Journal of Biology

http://journals.tubitak.gov.tr/biology/

Research Article

Turk J Biol
(2016) 40: 1096-1110

doi:10.3906/biy-1602-62

Analysis of global microRNAome profiles of *Caenorhabditis elegans* oocytes and early embryos

Selen GÜÇLÜ DURGUN*, Ahmet Raşit ÖZTÜRK, Arzu ATALAY*****
Biotechnology Institute, Ankara University, Ankara, Turkey

Received: 15.02.2016 • **Accepted/Published Online:** 25.03.2016 • **Final Version:** 08.11.2016

Abstract: One of the most complex cell transformations in biology is the remodeling of a fertilized oocyte into a totipotent zygote. There is no transcription during this event, and therefore mRNAs and proteins accumulated in the oocyte during oogenesis are utilized until zygotic genome activation. In order to differentiate into somatic cells, the embryo must erase the totipotency signature coming from the oocyte cytoplasm in the somatic blastomeres. In this study, the microRNAome profiles of oocytes and early embryos before and after zygotic genome activation were identified. By using next-generation high-throughput pyrosequencing and miRNA microarray techniques, miRNAome profiles of *Caenorhabditis elegans* oocytes (mitotic and meiotic) and early embryos before (2-blastomere stage) and after (8-blastomere stage) zygotic genome activation were generated. Besides observing results similar to those of previous reports, new miRNAs were identified whose expression levels changed during this transformation. After selection and polymerase chain reaction validation of the candidate miRNAs, downstream targets were identified by using bioinformatic tools and they were subjected to pathway analysis. Further analysis of the identified miRNAs in this study may provide a better understanding of the regulation of early embryonic development, germline–soma differentiation, oocyte–embryo transition, and maintenance of totipotency status.

Key words: *Caenorhabditis elegans*, early embryogenesis, microRNA, microRNAomics, maternal RNA, oocyte–embryo transition, totipotent, pyrosequencing, microarray

1. Introduction

Remodeling of a totally differentiated fertilized oocyte to a totipotent egg is a very complex cell transformation since one of its hallmarks is the absence of transcription. Several biological processes such as chromatin remodeling, epigenetic modifications, maternal RNA depletion, and zygotic genome activation take place at the same time during oocyte–embryo transition. During oogenesis RNA binding proteins, transcription factors, and specific RNAs encoding them are synthesized and stored in a translationally repressed state in the oocyte cytoplasm. Through late oogenesis, transcription is shut off and fertilization and early embryonic cleavages are orchestrated by these maternal RNAs and proteins until zygotic gene activation.

Following fertilization, in order to unleash the totipotency of the embryo, zygotic genome activation and degradation of maternal mRNAs and proteins must occur. How maternal proteins are degraded in the early *Caenorhabditis elegans* embryo was identified by De Renzo and Seydoux (2004). Maternal mRNA and protein decay are fundamental for proper embryonic development and differentiation of somatic lineages and germline specification. Presumably, degradation of maternal RNAs allows the zygotic genome to take over and introduce spatial and temporal differences in mRNA distribution. In *Drosophila* maternal mRNA degradation depends on recruitment to 3'UTRs to the CCR4/POP2/NOT1 deadenylase (Semotok et al., 2005) and in zebrafish, a single miRNA, mir-430, expressed at the onset of zygotic transcription, accelerates degradation and clearance of several hundred maternal mRNAs (Giraldez et al., 2006).

In an attempt to identify whether microRNAs or other classes of small RNAs are responsible for degradation of maternal mRNAs in the early embryo, two high-throughput pyrosequencing studies have been performed (Stoeckius et al., 2009, 2014). A lot of small RNA types are present in the early embryo, so for identifying the complexity and composition of small RNA transcriptome changes during the early stages of *C. elegans* embryogenesis, embryos

* Present address: Faculty of Medicine, Gazi University, Ankara, Turkey
** Present address: Genome and Stem Cell Center, Erciyes University, Kayseri, Turkey
*** Correspondence: arzu.atalay@ankara.edu.tr

were sorted with fluorescent activated cell sorting and deep sequencing was performed. Very complex and orchestrated changes within almost all classes of small RNAs, including miRNAs and 26G-RNAs, were described (Stoeckius et al., 2009). The second study combined transcriptome and proteome approaches to uncover the dynamics of oocyte embryo transition and obtained data suggesting a mechanism for maternal RNA degradation via a conserved polyC motif and binding partners. This work also suggested that endogenous siRNAs but not miRNAs promote mRNA clearance during the oocyte–embryo transition (Stoeckius et al., 2014).

Although these two studies provided a considerable amount of data describing changes in small RNA transcriptomes of mature oocytes and several different stages of embryos, we also aimed to identify the totipotency-associated microRNAs. We have combined deep sequencing and microarray techniques in order to analyze the microRNA profiles of oocyte-enriched germlines, 2-blastomere-stage (before zygotic genome activation) embryos, 8-blastomere-stage (after zygotic genome activation) embryos, and a mixture of pregastrula-stage early embryos. The obtained results from this work will be an important contribution to the knowledge of the regulation of germline/soma differentiation, maintaining/unleashing of totipotency, and posttranscriptional gene regulation in early embryos.

2. Materials and methods

2.1. *C. elegans* strains and culture

Wild-type (N2, Bristol, UK) and mutant *fog-2(q71)V C. elegans* strains were maintained using standard methods on OP50-seeded NGM plates at 20 °C unless otherwise noted (Brenner, 1974; Nayak et al., 2005).

2.2. Total RNA isolation from collected germlines and embryos

Egg preparation and synchronization of worms were performed according to Stiernagle (2006). Day 1 adult worms were collected and washed in M9. After addition of 0.2 mM levamisole as an anesthetic, worms were dissected using 25-G needles as scissors. The fully extruded gonads and stage-specific embryos were collected by mouth pipette into TRIzol (Invitrogen) on ice, flash frozen in liquid nitrogen, and stored at –80 °C until RNA extraction. RNA was isolated from all samples according to the manufacturer's instructions. RNA was precipitated by adding glycogen (20 µg glycogen/1 mL TRIzol). The RNAs whose RIN values were found to be higher than 8 (BioAnalyzer, Agilent) were used for further analysis. In order to get enough RNA for microarray and cDNA synthesis, 6000–7000 embryos were collected from each stage. For deep sequencing, much more starting material is needed and 1000 germlines were collected for RNA isolation.

2.3. Next-generation sequencing

Cloning of small RNAs and library preparation for 454 sequencing was performed according to Gregory Hannon's protocol (http://genoseq.ucla.edu/images/a/a9/SmallRNA.pdf). Briefly, small RNAs were separated by PAGE and the corresponding bands for mature miRNAs were excised from the gel. After precipitation of small RNAs, adapters were ligated to both sides (Güçlü, 2013). Following reverse transcription and polymerase chain reaction (PCR) amplification of cDNA, 454 sequencing was performed according to the manufacturer's recommendations using commercially available kits from Roche. Deep sequencing was performed on a Genome Sequencer Titanium FLX (Roche). Each sequence was processed using a set of customized MATLAB codes to clear the 3'- and 5'-end linkers, which were added during the sequencing reaction, and raw miRNA sequences were obtained. The expression profile of each miRNA was determined by another customized MATLAB routine, which finds out how many times each miRNA was counted in the sequence file. The annotations of the known miRNAs were performed using a publicly available database, miRanalyzer (Hackenberg et al., 2009), according to miRBase release 16 (http://bioinfo2.ugr.es/miRanalyzer/miRanalyzer.php).

2.4. MicroRNA microarray

From each 2-blastomere-stage and 8-blastomere-stage embryo, 2.5 µg of total RNA was sent to Exiqon A/S (Vedbaek, Denmark) and microRNA microarray analysis was performed on the miRCURY LNA microRNA Array v11. Data analysis was also performed by Exiqon Services.

2.5. Quantitative real-time polymerase chain reaction (QRT-PCR)

cDNA synthesis was performed using the miScript Reverse Transcription Kit (QIAGEN) according to the manufacturer's recommendations. Briefly, after addition of poly (A) to the 3'-ends of mature miRNAs, reverse transcription was carried out by using oligo-dT primers. QRT-PCR reactions were also carried out using miScript primers and the miScript SYBR Green PCR Kit (QIAGEN) by including biological and technical duplicates. The primer names and catalog numbers (QIAGEN) are as follows: Ce_miR-239a_1 (MS00019446), Ce_miR-48_1 (MS00019852), Ce_lin-4_1 (MS00018900), Ce_miR-71_1 (MS00020006), Ce_miR-241_1 (MS00019467), Ce_miR-237_1 (MS00019432), Ce_miR-51_1 (MS00019873), Ce_miR-64_2 (MS00028714), Ce_miR-52_1 (MS00019880), Ce_miR-1829b_1 (MS00019040), and cel-miR-52-3p (MIMAT0020311). U18 small RNA was used for normalization and the primer sequences are as follows:

CeU18F, 5'- ATAGAAAACCGGCTGAGCCA-3',

CeU18R, 5'- TGGCTCAGCCGGTTTTCTAT -3'.

Heat mapping and cluster analysis were performed with the Cluster 3.0 and TreeView programs.

2.6. In silico analysis of target mRNAs

The potential target mRNAs of each miRNA were found using the microRNA.org website (http://www.microrna.org/microrna/home.do, August 2010 release). After preparing the list of potential targets with mirSVR scores bigger than –0.9, the gene lists were used for pathway analysis using the GeneCodis 2.0 program (http://genecodis.dacya.ucm.es/analysis/) (Carmona-Saez et al., 2007; Nogales-Cadenas et al., 2009; Tabas-Madrid et al., 2012).

3. Results

3.1. RNA isolation from germline and stage-specific embryos

In order to identify totipotency/germline-associated miRNAs, oocyte-enriched germlines and early-stage embryos (before gastrula) were collected from the *fog-2(q71)V C. elegans* strain. For identifying maternal and zygotic miRNAs 2-blastomere-stage (before zygotic genome activation) and 8-blastomere-stage (after zygotic genome activation) embryos were collected from N2 worms. Figure 1 shows the stage specificity of collected embryos and dissected germlines used for RNA isolation.

3.2. Next-generation sequencing

Pyrosequencing reports showed that 74% of the wells (% Short and % Passed filter) on the picotiter plate generated reads outs, which is equal to 1,275,747 successful sequencing reactions (Table 1). Median read length was 56 and longest bead length was 604 (control beads). Each sequence was processed using a set of customized MATLAB codes to clear the 3'- and 5'-end linkers, which were added during the sequencing reaction, and raw small RNA sequences were obtained. Table 2 shows the partial statistics after trimming of germline (oocyte)-associated small RNAs; 72,315 sequences were obtained for annotation and the number of unique sequences was 31079 (Table 2; Supplement 1). Partial statistics after trimming of early-stage mixed embryo-associated small RNAs are given in Table 3; 98,519 sequences were obtained for annotation and the number of unique sequences was 10.627 (Table 3; Supplement 2). The expression profile of each miRNA was determined by another customized MATLAB routine, which finds out how many times each miRNA was counted in the sequence file. The annotations of the known miRNAs were performed by using a publicly

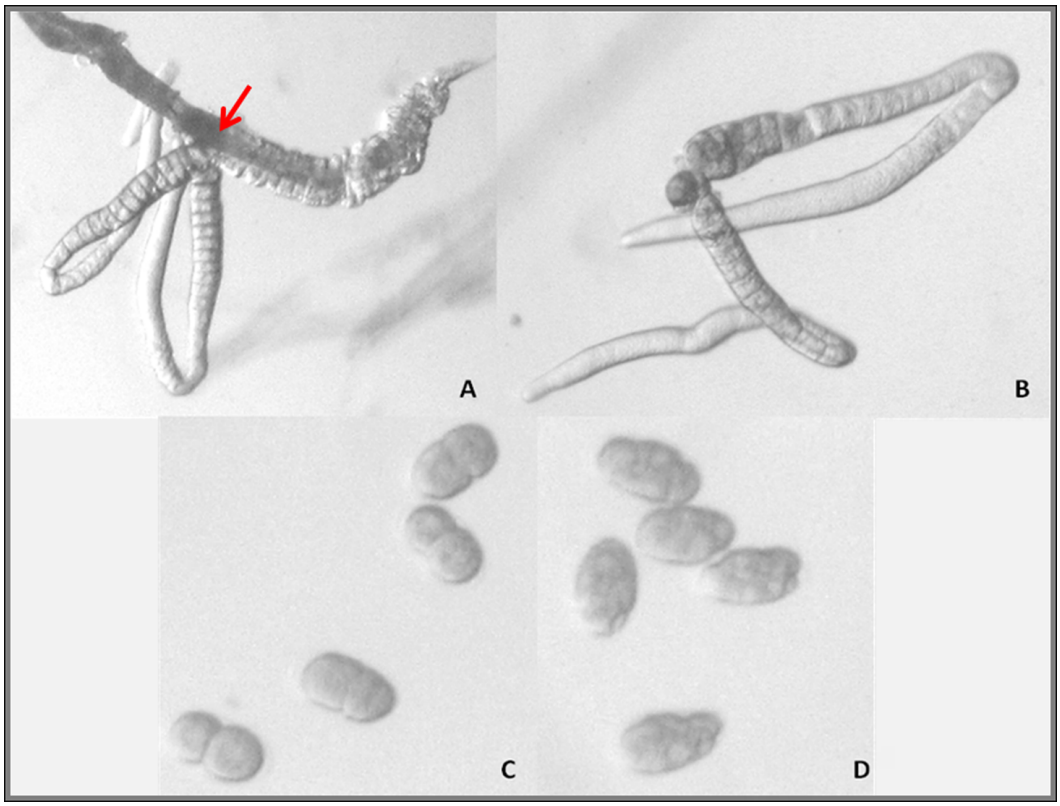

Figure 1. Dissected germlines and stage-specific embryos used for RNA isolation. (A) When dissected with needles, gonad arms were still connected to the intestine. (B) After cutting the intestine, whole gonad arms were picked out for RNA isolation. The *fog-2(q71)V* mutant *C. elegans* strain was used for germline collections since the mature oocyte number is very high compared to wild-type *C. elegans*. (C) Two-blastomere-stage and (D) 8-blastomere-stage embryos were collected from dissected N2 wild-type *C. elegans*.

Table 1. GS FLX reaction report.

TCAG (Library)	Region		
	1*	2*	Total
Raw Wells	728,525	622,015	1,350,540
Key Pass Wells	690,498	585,249	1,275,747
Failed Dot	62,810	50,751	113,561
Mixed	102,087	117,823	219,910
Short Quality	342,299	225,470	567,769
Short Primer	135,953	148,887	284,840
Passed Filter Wells	47,349	42,318	89,667
% Dot + Mixed	23.88	28.80	26.14
% Short	69.26	63.97	66.83
% Passed Filter	6.86	7.23	7.03
Length Average	54.93	57.55	56.36
Length Std. Deviation	12.96	9.61	
Longest Reads Length	604	477	604
Shortest Reads Length	41	41	41
Median Reads Length	55.0	57.0	56.0

*1: Region 1 of picotiter plate containing germline (oocyte) sample.
*2: Region 2 of picotiter plate containing mixed early-stage embryo sample.

available database, miRanalyzer. Eighty-four and 90 microRNAs were annotated for oocytes and early-stage embryos, respectively. According to the read counts, cel-miR-241-5p, cel-miR-246-3p, cel-miR-48-5p, cel-lin-4-5p, cel-miR-78, cel-miR-34-5p, and cel-miR-51-5p had the highest expression levels in the oocyte-enriched germlines. Cel-miR-52-3p, cel-miR-41-3p, cel-miR-42-3p, cel-miR-37-5p, cel-miR-44-3p/cel-miR-45-3p, cel-miR-46-3p, cel-miR-73-3p, cel-miR-37-3p, cel-miR-52-5p, and cel-miR-36-3p had the highest expression levels in the early mixed-stage embryos (before gastrula). While 33 of the miRNAs were not expressed in the oocyte-enriched germlines, 25 of the miRNAs were not expressed in the early embryos (Table 4).

3.3. MicroRNA microarray

Table 5 summarizes the results of the microRNAs analyzed by the miRCURY LNA microRNA Array that are differentially expressed in 2-blastomere-stage and 8-blastomere-stage N2 worms. The normalized log ratio values are shown. Positive logFC indicated the logarithmic foldness of upregulation and negative logFC indicated the logarithmic foldness of downregulation. Although there are no biological replicates of the experiment due to the lack of starting material, this gives some clues about the maternal and zygotic microRNAs in the early embryos. The expression levels of a total of 61 mature microRNAs were analyzed.

3.4. QRT-PCR analysis of differentially expressed miRNAs

After putting together pyrosequencing data, microarray data, and the work published by Stoeckius et al. (2009), a few of the available miScript primers were used for further confirmation of expression levels. QRT-PCR results were expressed as relative quantification based on the $2^{-\Delta\Delta ct}$ method and normalized against the average of the *U18 sno* gene. Heat maps generated from QRT-PCR data reflecting gene expression values in 2-blastomere-stage, 8-blastomere-stage, early-stage mixed embryos, and oocyte-enriched germline samples are shown in Figure 2. Red coloring indicates expression levels higher in the oocyte-enriched germlines and 2- and 8-blastomere-stage embryos compared to mixed-stage early embryos, while green indicates higher expression in the mixed-stage embryos compared to the others.

3.5. In silico analyses of target mRNAs

In order to enrich and classify the target mRNAs regulated by these miRNAs, pathway analyses were performed. First, potential target mRNAs of the differentially expressed miRNAs, for the available ones, were obtained from microRNA.org and pathway analysis was performed

Table 2. Deep sequencing statistics of oocyte-enriched germline sample after trimming linker sequences.

Average length of miRNA sequences		20.97 bases	
# of sequences		77,582	
# of sequences that are filtered		72,315	
Missing sequences		5267 (6.79%)	
Total number of unique sequences		31,079	
		The most popular sequences	Read count
10 bases	45 times	AGCACGGAAACATATGTACGGGTG	4094
11 bases	130 times	TCGCATCTACTGAGCCTACCTCA	3415
12 bases	273 times	TGAGGTAGGCTCAGTAGATGCGA	2617
13 bases	456 times	ATTGCCGTACTGAACGATCTCA	1690
14 bases	1139 times	CACCCGTACATATGTTTCCGTGCT	1637
15 bases	1585 times	TCACCGGGTGTACATCAGCTAA	1334
16 bases	2535 times	TTAGCTGATGTACACCCGGTGA	1058
17 bases	3417 times	ACTGAACTGCCTACATCTTGCCA	1025
18 bases	5002 times	ATGAGTAACGGTTCTAGTCA	805
19 bases	5261 times	TTAGCTGATGTANACCCGGTGA	673
20 bases	8088 times	AGCACGGAACATATGTACGGGTG	659
21 bases	8416 times	TGAGATCGTTCAGTACGGCAAT	612
22 bases	15,517 times	TTACATGTTTCGGGTAGGAG	518
23 bases	14,049 times	CACCGTACATATGTTTCCGTGCT	516
24 bases	8651 times	TGACTAGAACCGTTACTCAT	487
25 bases	2011 times	TGGCAAGATGTAGGCAGTTCAGT	444
26 bases	713 times	AGCACGGAAACATATNTACGGGNG	443
27 bases	181 times	TCACCGGTGTACATCAGCTAA	439
28 bases	51 times	AGCACGGAAACATATGTACGGGGTG	416
29 bases	37 times	TCAGCTATGCCAACATCTTGCC	324
30 bases	13 times	CATGACACTGATTAGGGATGTGA	320
31 bases	6 times	AGCTGAATGTGTCTCTAGTCA	314
32 bases	3 times	TCGCATCTACTGAGCTACCTCA	313
33 bases	1 time	TCAACGTGATATGGTCGTAAGC	302
34 bases	1 time	TCACATCCCTAATCAGTGTCANG	292
37 bases	1 time	TTGCCGTACTGAACGATCTCA	264
		TGGCGACCACGGCA	254
		TGGCGACCACGGCAGGATTCGA	252
		ACTGGCTTTCACGATGATCTCA	229
		AACATGGATAGGAGCTACGGGTA	222

for each of them. Line 1 of Table 6 shows the number of total potential mRNA targets for each miRNA and the number of miRNAs that were found to be associated with any biological processes; approximately one-third of the target mRNAs were found to be associated with biological processes. The remaining lines of the table provide partial information about the number of target mRNAs having roles within the indicated biological processes.

4. Discussion

One of the most important cell fate decisions made in early embryonic development is the choice between becoming

Table 3. Deep sequencing statistics of pregastrula mixed-stage early embryos after trimming linker sequences.

Average length of miRNA sequences	23.50 bases
# of sequences	103,128
# of sequences that are filtered	98,519
Missing sequences	4609 (4.47%)
Total number of unique sequences	10,627

		The most popular sequences	Read count
10 bases	23 times	AGCACGGAAACATATGTACGGGGTG	9416
11 bases	33 times	CACCCGTACATATGTTTTCCGTGCT	6747
12 bases	86 times	AGCACGGAAAACATATGTACGGGGG	6726
13 bases	108 times	ACTGAACTGCCTACATCTTGCCA	5050
14 bases	329 times	AGCACGGAAACATATGTACGGGTG	4801
15 bases	345 times	TGGCAAGATGTAGGCAGTTCAGT	3223
16 bases	374 times	CACCCGTACATATGTTTCCGTGCT	3219
17 bases	574 times	AGCACGGAAAACATATNTACGGGGG	3075
18 bases	1134 times	ACTGAAACTGCCTACATCTTGCCA	2240
19 bases	1982 times	ATTGCCGTACTGAACGATCTCA	1902
20 bases	2406 times	TCACCGGGTGTACATCAGCTAA	1731
21 bases	7323 times	AGCTGAAATGTGTCTCTAGTCA	1521
22 bases	13,727 times	AGCTGAATGTGTCTCTAGTCA	1486
23 bases	18,625 times	TGACTAGAGACACATTCAGCT	1382
24 bases	17,368 times	AGCACGGGAAAACATATNTACGGGGNG	1278
25 bases	21,933 times	ACTGAAACTGCCTACATCTTGCCCA	1028
26 bases	13,598 times	CACCCCGTACATATGTTTTCCGTGCT	984
27 bases	2499 times	TGAGATCGTTCAGTACGGCAAT	795
28 bases	514 times	CACGTTACAATGAAAGGGTAG	763
29 bases	85 times	TGTGGGTGTCCGTTGCGGTGCTA	676
30 bases	37 times	TTAGCTGATGTANACCCCGGTGA	670
31 bases	15 times	AGCACGGAACATATGTACGGGTG	663
32 bases	3 times	TTAGCTGATGTACACCCCGGTGA	582
33 bases	3 times	AGCACGGAAAACATATGTACGGGTG	563
34 bases	1 time	GTAGCAATGGTACTGACGCTGT	529
36 bases	2 times	TTGCCGTACTGAACGATCTCA	520
56 bases	1 time	CTCAGCGGAAAACATTACGGGGA	390
		CTGAATGTGTCTCTAGTCA	370
		TTAGCTGATGTACACCCGGTGA	339
		CTCAGCGGAAAACATTACGGGGTA	339

either a somatic or a germ cell. Only the germ cell lineage passes genetic information onto the next generation and totipotent germ cells are capable of differentiating into all cell lineages while somatic cells become progressively restricted to their fate. Increased understanding of the signaling and modeling events that shape the oocyte–embryo transition is critical to understand what constitutes a healthy egg and ensuring proper embryonic development.

Table 4. The read count number of miRNAs annotated by the miRanalyzer program of the deep-sequenced samples. GL: Oocyte-enriched germlines, EE: Pregastrula mixed-stage early embryos.

miRNA	GL	EE	GL/EE
cel-miR-124-3p	0.00	0.95	0.00
cel-miR-124-5p	0.00	0.95	0.00
cel-miR-2217-3p	0.00	0.95	0.00
cel-miR-228-3p	0.00	6.67	0.00
cel-miR-233-3p	0.00	3.81	0.00
cel-miR-236-3p	0.00	3.81	0.00
cel-miR-2-3p	0.00	3.81	0.00
cel-miR-242	0.00	0.95	0.00
cel-mir-252	0.00	0.95	0.00
cel-miR-2-5p	0.00	3.81	0.00
cel-miR-35-5p	0.00	1.91	0.00
cel-miR-44-5p	0.00	14.30	0.00
cel-miR-46-5p	0.00	0.95	0.00
cel-miR-49-3p	0.00	14.30	0.00
cel-miR-49-5p	0.00	1.91	0.00
cel-miR-51	0.00	0.95	0.00
cel-miR-54	0.00	2.86	0.00
cel-miR-54-5p	0.00	4.77	0.00
cel-miR-56-5p	0.00	33.37	0.00
cel-miR-57-3p	0.00	5.72	0.00
cel-miR-61-5p	0.00	0.95	0.00
cel-miR-62	0.00	21.93	0.00
cel-miR-67-3p	0.00	1.91	0.00
cel-miR-67-5p	0.00	0.95	0.00
cel-miR-73-5p	0.00	6.67	0.00
cel-miR-74-5p	0.00	0.95	0.00
cel-miR-75-5p	0.00	0.95	0.00
cel-miR-792-3p	0.00	1.91	0.00
cel-miR-800-3p	0.00	0.95	0.00
cel-miR-86-5p	0.00	0.95	0.00
cel-miR-90-5p	0.00	5.72	0.00
cel-miR-52-3p	5.24	1011.62	0.01
cel-miR-41-3p	7.34	417.62	0.02
cel-miR-42-3p	3.15	108.69	0.03
cel-miR-37-5p	63.98	726.54	0.09
cel-miR-44-3p/cel-miR-45-3p	225.49	1833.51	0.12
cel-miR-46-3p	12.59	96.30	0.13
cel-miR-73-3p	486.65	3493.49	0.14
cel-miR-37-3p	39.85	273.64	0.15
cel-miR-52-5p	2075.59	10105.75	0.21

Table 4. (Continued).

cel-miR-36-3p	72.37	343.25	0.21
cel-miR-55-3p	15.73	61.02	0.26
cel-miR-50-5p	25.17	95.35	0.26
cel-miR-35-3p	102.78	373.76	0.28
cel-miR-57-5p	6.29	21.93	0.29
cel-miR-228-5p	5.24	14.30	0.37
cel-miR-83-3p	7.34	17.16	0.43
cel-miR-229-5p	2.10	4.77	0.44
cel-miR-39-3p	127.95	268.88	0.48
cel-miR-74-3p	2.10	3.81	0.55
cel-miR-79-3p	1.05	1.91	0.55
cel-miR-65-5p	2.10	2.86	0.73
cel-miR-87-5p	2.10	2.86	0.73
cel-miR-47-3p	18.88	24.79	0.76
cel-miR-58-3p	895.68	1175.62	0.76
cel-miR-40-3p	1514.48	1916.46	0.79
cel-miR-229-3p	22.02	27.65	0.80
cel-miR-36-5p	43.00	51.49	0.84
cel-miR-60-3p	7.34	8.58	0.86
cel-miR-53-5p	125.86	137.30	0.92
cel-miR-87-3p	5.24	5.72	0.92
cel-miR-58-5p	10.49	10.49	1.00
cel-miR-1829b	1.05	0.95	1.10
cel-miR-244-5p	1.05	0.95	1.10
cel-miR-43-3p	1.05	0.95	1.10
cel-miR-51-3p	12.59	11.44	1.10
cel-miR-72-3p	2.10	1.91	1.10
cel-miR-72-5p	409.04	317.50	1.29
cel-miR-82-3p	109.08	71.51	1.53
cel-miR-66-5p	421.62	274.60	1.54
cel-miR-38-3p	7.34	4.77	1.54
cel-miR-65-3p	3.15	1.91	1.65
cel-miR-64-5p	5.24	2.86	1.83
cel-miR-56-3p	87.05	45.77	1.90
cel-miR-61-3p	721.58	370.90	1.95
cel-miR-80-3p	48.25	24.79	1.95
cel-miR-81-3p	87.05	40.05	2.17
cel-miR-1-3p	2.10	0.95	2.20
cel-miR-51-5p	148.93	56.25	2.65
cel-miR-795-5p	3.15	0.95	3.30
cel-miR-90-3p	3.15	0.95	3.30
cel-miR-47-5p	7.34	1.91	3.85
cel-miR-237-5p	4.20	0.95	4.40

Table 4. (Continued).

cel-miR-253-3p	4.20	0.95	4.40
cel-miR-59-3p	13.63	2.86	4.77
cel-miR-34-5p	46.15	8.58	5.38
cel-miR-250-3p	6.29	0.95	6.60
cel-miR-70-3p	9.44	0.95	9.90
cel-miR-78	65.03	4.77	13.64
cel-lin-4-5p	44.05	1.91	23.10
cel-miR-48-5p	2824.44	43.86	64.40
cel-miR-246-3p	808.63	0.95	848.10
cel-let-7-5p	1.05	0.00	INF
cel-miR-1829c	1.05	0.00	INF
cel-miR-230-3p	2.10	0.00	INF
cel-miR-238-3p	6.29	0.00	INF
cel-miR-239a-3p	1.05	0.00	INF
cel-miR-239a-5p	2.10	0.00	INF
cel-miR-241-3p	5.24	0.00	INF
cel-miR-241-5p	44.05	0.00	INF
cel-miR-253	1.05	0.00	INF
cel-miR-41-5p	2.10	0.00	INF
cel-miR-4816-3p	1.05	0.00	INF
cel-miR-4936	2.10	0.00	INF
cel-miR-4937	1.05	0.00	INF
cel-miR-5592-3p	11.54	0.00	INF
cel-miR-5592-5p	7.34	0.00	INF
cel-miR-64	1.05	0.00	INF
cel-miR-71-5p	11.54	0.00	INF
cel-miR-77-3p	4.20	0.00	INF
cel-miR-77-5p	1.05	0.00	INF
cel-miR-786-3p	4.20	0.00	INF
cel-miR-794-5p	1.05	0.00	INF
cel-miR-81-5p	1.05	0.00	INF
cel-miR-84-3p	3.15	0.00	INF
cel-miR-84-5p	2.10	0.00	INF
cel-miR-85-3p	1.05	0.00	INF

Our aim was to identify the miRNAs associated with totipotency and the oocyte–embryo transition. Due to the difficulty of getting enough starting material for deep sequencing on the GS FLX platform, we used the mutant *fog-2(q71)V C. elegans* strain (feminization of germline), which has more mature oocytes compared to the wild-type N2 due to lack of sperm, and we also used pregastrula mixed-stage embryos from this mutant for deep sequencing. On the other hand, although we handpicked thousands of 2- and 8-blastomere-stage embryos, it was not enough for deep sequencing, so we did microarray analysis. According to the data obtained from next-generation sequencing, the majority of totipotency-associated miRNAs are cel-miR-48-5p, lin4-5p, cel-miR-246-3p, cel-miR-71-5p, and cel-miR-241-5p. These miRNAs are maternally expressed in oocytes and their expression levels decrease through embryonic development. On the other hand, major zygotic microRNAs including miR-52-5p, cel-miR-52-3p,

Table 5. Differential expression analysis of the microRNAs in 2- and 8-blastomere-stage embryos on microarray.

miRNA name	2-blastomere embryos	8-blastomere embryos	logFC 2/8
cel-miR-79-3p	–0.16	–0.68	1.43
cel-miR-239a-5p	–0.23	–0.61	1.29
cel-miR-77-3p	–0.14	–0.27	1.09
cel-miR-47-3p	0.21	0.16	1.04
cel-miR-237-5p	–0.23	–0.28	1.04
cel-miR-272	0.32	0.31	1.01
cel-miR-46-3p	–0.04	–0.02	0.99
cel-miR-34-5p	–0.01	0.03	0.98
cel-miR-36-3p	–0.40	–0.36	0.97
cel-miR-1824-5p	0.66	0.72	0.96
cel-miR-60-3p	–0.24	–0.16	0.95
cel-miR-85-3p	–0.19	–0.11	0.94
cel-miR-229-5p	–0.14	–0.05	0.94
cel-miR-44-3p/cel-miR-45-3p	–0.37	–0.25	0.93
cel-miR-786-3p	0.01	0.17	0.90
cel-miR-61-3p	–0.43	–0.27	0.89
cel-miR-354	0.22	0.39	0.89
cel-miR-244-5p	0.33	0.50	0.89
cel-miR-253-3p	–0.26	–0.08	0.89
cel-miR-63-3p	–0.12	0.06	0.88
cel-miR-42-3p	–0.47	–0.28	0.88
cel-miR-66-5p	–0.29	–0.09	0.87
cel-miR-83-3p	–0.36	–0.16	0.87
cel-miR-57-5p	–0.15	0.06	0.87
cel-miR-50-5p	0.39	0.61	0.86
cel-miR-53-5p	–0.12	0.11	0.85
cel-miR-84-5p	–0.20	0.04	0.85
cel-miR-1-3p	–0.17	0.08	0.84
cel-miR-230-3p	–0.26	0.06	0.80
cel-miR-252-5p	–0.19	0.13	0.80
cel-miR-75-3p	–0.46	–0.13	0.79
cel-miR-90-3p	–0.32	0.03	0.78
cel-miR-87-3p	–0.04	0.32	0.78
cel-miR-72-5p	–0.46	–0.07	0.76
cel-miR-795-5p	–0.73	–0.33	0.76
cel-miR-58-3p	–0.58	–0.17	0.75
cel-let-7-5p	–0.24	0.19	0.74
cel-miR-56-3p	–0.13	0.31	0.74
cel-miR-241-5p	–0.42	0.03	0.73
cel-miR-39-3p	–0.72	–0.20	0.70
cel-miR-74-3p	–0.45	0.09	0.69

Table 5. (Continued).

cel-miR-54-3p	–0.51	0.05	0.68
cel-miR-55-3p	–0.29	0.30	0.67
cel-miR-238-3p	–0.51	0.08	0.66
cel-miR-270	–0.68	–0.07	0.66
cel-miR-37-3p	–0.57	0.04	0.66
cel-miR-40-3p	–0.59	0.03	0.65
cel-miR-35-3p	–0.65	0.00	0.64
cel-miR-356	–0.85	–0.19	0.63
cel-miR-48-5p	–0.58	0.19	0.59
cel-miR-81-3p	–0.61	0.19	0.58
cel-miR-73-3p	–0.41	0.41	0.56
cel-miR-38-3p	–0.74	0.09	0.56
cel-miR-52-5p	–0.61	0.26	0.55
cel-miR-51-5p	–0.82	0.06	0.54
cel-miR-82-3p	–0.69	0.19	0.54
cel-miR-71-5p	–0.65	0.24	0.54
cel-miR-80-3p	–0.75	0.17	0.53
cel-miR-65-5p	–0.70	0.27	0.51
cel-miR-64-5p	–0.72	0.30	0.49
cel-miR-789	–2.05	–0.04	0.25

cel-miR-37-3p, cel-miR-46-3p, cel-miR-42-3p, cel-miR-73-3p, cel-miR-55-3p, and cel-miR-35-3p are expressed in embryos after zygotic genome activation.

cel-miR-48-5p (also known as lin-58) expression was increased at the 8-blastomere stage and this is consistent with the literature. Stoeckius et al. (2009) reported miR-48-5p expression in one blastomere that was degraded in 2–4 blastomeres and expression was observed again following zygotic genome activation. On the other hand, our next-generation sequence data showed that miR-48-5p is expressed maternally, which was also consistent with the QRT-PCR assay results (Figure 2). According to our next-generation sequencing data, cel-miR-48-5p, cel-miR-84-5p, and cel-miR-241-5p, which all belong to the let-7 family, are expressed maternally. This result is also consistent with the literature. These miRNAs are degraded at the 2-blastomere stage and resynthesized during the postgastrulation stage in embryos (Stoeckius et al., 2009). The expression level of cel-miR-241-5p was higher in 8-blastomere-stage embryos compared to 2-blastomere-stage embryos in our microarray data; however, a significant difference was not observed in QRT-PCR results. When previously reported studies and our data are considered together, miR-241-5p is not only expressed maternally but it is also resynthesized in the later stages of *C. elegans* embryogenesis.

cel-miR-237-5p shares strong sequence homology with *lin-4* of *C. elegans*. miR-237 belongs to the lin-4 family of microRNAs, which also includes human miR-125b-1 and miR-125b-2. Despite the low number of sequence reads, according to the results of next-generation sequencing, miR-237-5p appears to be maternally expressed and was found as only one read in one-cell-stage embryos in the miRNAome profiling study of Stoeckius et al. as well (2009). Since miR-237-5p is the top fifth differentially expressed miRNA according to our microarray data, we tested the expression level of this miRNA by QRT-PCR. Our results confirmed that miR-237-5p is expressed more in 2-blastomere-stage embryos before zygotic genome activation (Figure 2). Further functional analysis will contribute to the elucidation of the function of cel-miR-237-5p in early embryos.

cel-miR-71-5p is one of the highly expressed microRNAs in 2- and 8-blastomere-stage embryos. Both next-generation sequencing and QRT-PCR results also indicate that cel-miR-71-5p is expressed maternally (Tables 4 and 5; Figure 2). This miRNA is expressed abundantly in larval and early adult forms of *C. elegans*. Overexpression of cel-miR-71-5p increases the lifespan and provides resistance to temperature and oxidative stress (De Lencastre et al., 2010; Boulias and Horvitz, 2012). Further functional analysis will contribute to clarifying the role of cel-miR-71-5p in early embry.

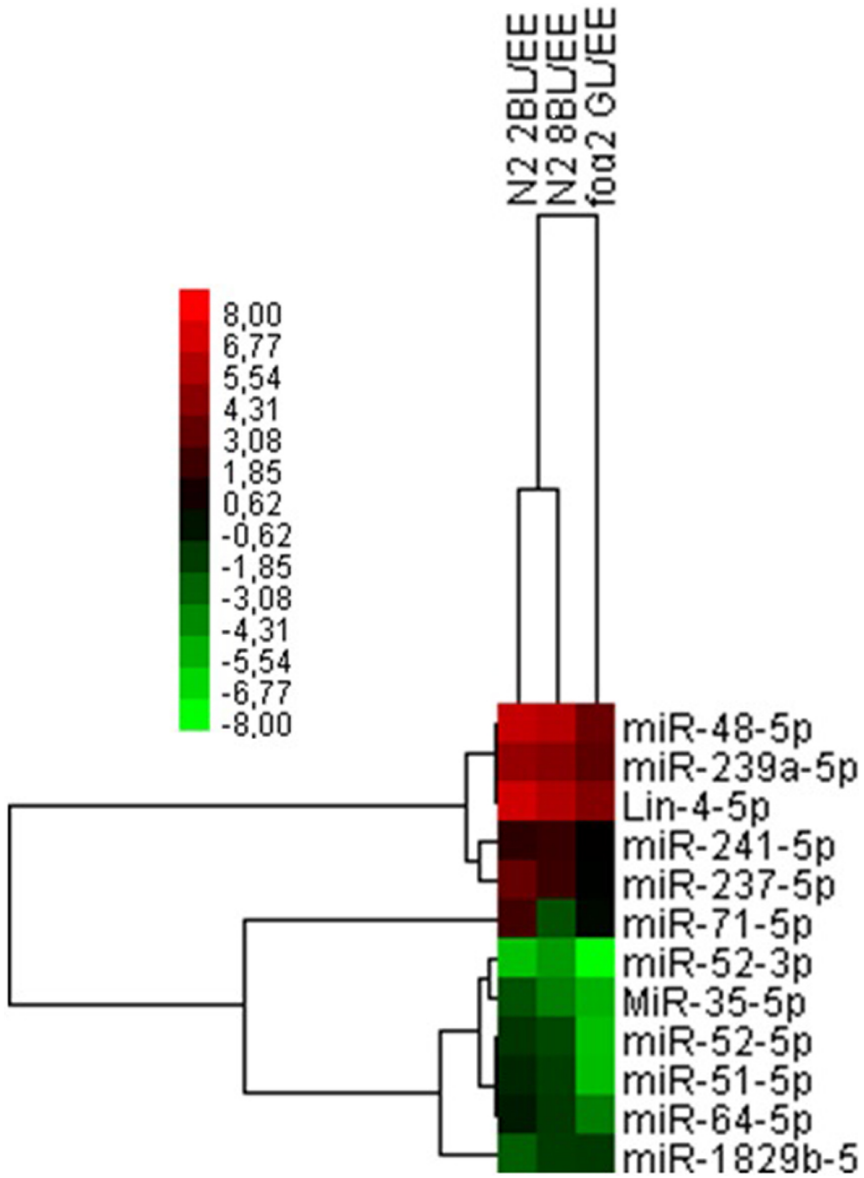

Figure 2. Clustergram generated from QRT-PCR data reflecting gene expression values in oocyte-enriched germlines (GL) and 2 (2B)- and 8 (8B)-blastomere-stage embryos versus pregastrula mixed early-stage embryos (EE). Dendrograms indicate coregulated groups. Magnitude of logarithmic fold change bar is shown on the left.

cel-miR-71 and cel-miR-239 are target components of the IGF signaling pathway and DNA damage checkpoint pathways to regulate aging. cel-miR-71 targets the PI3K AGE-1 and PDK-1 components of the IGF signaling pathway, whereas cel-miR-239 appears to upregulate the gene expression of these proteins through unknown mechanisms (De Lencastre et al., 2010). cel-miR-71 targets CDC-25.1 and CHK-1 in a negative feedback loop (De Lencastre et al., 2010). This suggests that these miRNAs have a function in the cellular response to environmental stress, in addition to promoting longevity. These miRNAs may function during environmental stress.

cel-miR-239-5p is expressed maternally according to next-generation sequencing and microarray data that are consistent with QRT-PCR results (Figure 2). According to the stage-specific microRNA profile that was determined by Stoeckius et al. in 2009, it was shown that cel-miR-239a is observed in the 1-blastomere-stage and it was degraded through embryogenesis, but later on the expression levels were increased again in late embryogenesis. The function

Table 6. Pathway analysis of differentially expressed microRNAs.

	cel-miR-52-5p	cel-miR-239a-5p	cel-miR-48-5p	cel-Lin-4-5p	cel-miR-71-5p	cel-miR-241-5p	cel-miR-237-5p	cel-miR-51-5p	cel-miR-64-5p	cel-miR-56-5p
Associated mRNA number in GeneCodis/Potential target mRNA number (microRNA.org)	143/364	189/486	128/407	36/78	552/1853	147/461	37/84	130/327	256/707	137/357
Biological processes that target mRNAs of each miRNA play a role in										
Embryonic development	30	46	66		240	74	11		97	33
Reproduction	23	36	47	10	197	55	11		66	27
Locomotion	23		35	10	173	40	11	20	60	24
Oviposition	14	15	15		57	16		12		13
Nematode larval development		33	37	11	161	42	12		56	
Growth		32	28	7	115		8			
Hermaphrodite genitalia development		32	21		78	23				

of cel-miR-239-5p is not clear so further investigations will contribute to understanding its role in early embryonic development.

According to the next-generation sequencing and microarray data, cel-miR-52-5p is expressed during the early embryonic stage. This result correlates with the literature. cel-miR-52-5p was observed at the highest level in *C. elegans* L1 larvae and its expression was also observed in different stages of development. Biochemical studies indicate that miR-52 interferes with ribonucleoprotein complexes (Lee and Ambros 2001; Caudy et al., 2003). MicroRNAs are integrated with ribonucleoprotein complexes as the RISC complex (RNA-induced silencing complex) and this complex is created with specific mRNA translational repression or degradation. *C. elegans* miR-52-5p has 62.5% overall similarity to human miR-99 according to sequence relationships (Ibáñez Ventoso et al., 2008).

cel-lin-4-5p is a maternally expressed microRNA based on our next-generation sequencing data (Table 4) and QRT-PCR experiment results (Figure 2). lin-4-5p, which is an important developmental regulator, is also known as miR-125 in *Drosophila* and mammals.

cel-miR-52-3p is expressed during embryonic development based on next-generation sequencing results. According to our data, a 193-fold increase was observed in the mixed stage of early embryos compared to oocytes (Table 4) and also QRT-PCR data were found to be consistent with the sequencing results (Figure 2). This miRNA was not detected in a previous microRNA profiling study on the Illumina platform (Stoeckius et al., 2009) and it will be further investigated during our future studies.

cel-miR-51-5p is expressed maternally in *C. elegans* according to our next-generation sequence read counts, microarray results, and a previous miRNAome profiling study (Stoeckius et al., 2009). QRT-PCR expression levels between 2 and 8 blastomeres did not exhibit significant changes, but zygotic expression profiles were observed when compared to oocytes and mixed stages of early embryos (Figure 2). cel-miR-51-5p was detected in the mid-embryo and mostly anterior in late embryos. cel-miR-51-5p was detected in channel nerve cells in the late embryo until the L1 larval stage. cel-miR-51-5p expression was found in the muscles in the head and intestines, suggesting that this microRNA has many functions.

According to our microarray data, cel-miR-64-5p is the second highly expressed miRNA after zygotic genome activation. It had a very low number of sequence read counts in next-generation sequencing. A previous miRNAome profiling study in early *C. elegans* embryos showed that expression level increases after zygotic genome activation (Stoeckius et al., 2009), which is consistent with our miRNA microarray and QRT-PCR results.

A homology analysis study in 2008 was performed with *C. elegans*, *D. melanogaster*, and *H. sapiens* miRNAs. MicroRNAs that are preserved in all of these species are miR-48, lin-4, miR-51, miR-52, miR-73, miR-79, miR-237, and miR-241 (Ibáñez Ventoso et al., 2008). Some miRNAs that were selected for further analysis are members of the same family and also they include the same "seed regions", which are conserved in many organisms (Ruby et al., 2006). miR-52, miR-51, miR-55, and miR-56 include the ACCCGU motif and show homology with *C. briggsae*, *D. melanogaster*, *D. rerio*, and mammalian miRNAs. Similarly, lin-4 and miR-237 include the CCCUG motif and show homology with all animals mentioned above. miR-48 and miR-241 have the GAGGUA motif and have homology with all species. miR-73 has the GGCAAG motif, which was found in all species mentioned above except zebrafish, and its tumor suppressor feature was highlighted in other studies. The miR-35, miR-37, and miR-42 families, all of which include the GAAAGA motif, are just found in *C. elegans* and *C. briggsae.* miR-239a has the UUGUAC motif and miR-79 has the UAAAGC motif, and these two are conserved in *C. elegans*, *C. briggsae*, and *D. melanogaster* only (Ibáñez Ventoso et al., 2008).

Two previous high-throughput pyrosequencing studies were performed in an attempt to identify the small RNA transcriptome in oocytes and stage-specific embryos (Stoeckius et al., 2009, 2014). Obtained data suggested a mechanism for maternal RNA degradation via a conserved polyC motif, and binding partners and endogenous siRNAs but not miRNAs promote mRNA clearance during the oocyte–embryo transition (Stoeckius et al., 2014). These studies provided a lot of data about the changes in small RNA transcriptome in mature oocytes and several different stage embryos. Some of our results are consistent with these studies, but we have identified new miRNAs expressed in these developmental stages. Mutant analysis and identification of downstream mRNA targets will provide a better understanding of regulation of totipotency and posttranscriptional gene regulation in early embryos. We have performed pathway analysis in order to enrich downstream potential mRNA targets, and among the clustered ones, the majority of them were associated with embryonic development and reproduction. Our pyrosequencing data were analyzed with an earlier version miRBase (release 16); a newer version such as release 21 will annotate many more microRNAs with our trimmed sequences, since the number of mature *C. elegans* miRNAs is around 434.

Acknowledgment

We thank the *Caenorhabditis* Genetics Center for worm strains. This study was funded by the Scientific and Technological Research Council of Turkey (grant numbers 108T918 and 108T254).

References

Boulias K, Horvitz HR (2012). The *C. elegans* microRNA mir-71 acts in neurons to promote germline-mediated longevity through regulation of DAF-16/FOXO. Cell Metab 15: 439-450.

Brenner S (1974). The genetics of *Caenorhabditis elegans*. Genetics 77: 71-94.

Carmona Saez P, Chagoyen M, Tirado F, Carazo JM, Pascual Montano A (2007). GENECODIS: A web-based tool for finding significant concurrent annotations in gene lists. Genome Biol 8: R3.

Caudy AA, Ketting RF, Hammond SM, Denli AM, Bathoorn AM, Tops BB, Silva JM, Myers MM, Hannon GJ, Plasterk RH (2003). A micrococcal nuclease homologue in RNAi effector complexes. Nature 425: 411-414.

De Lencastre A, Pincus Z, Zhou K, Kato M, Lee SS, Slack FJ (2010). MicroRNAs both promote and antagonize longevity in *C. elegans*. Curr Biol 20: 2159-2168.

De Renzo C, Seydoux G (2004). A clean start: degradation of maternal proteins at the oocyte-to-embryo transition. Trends Cell Biol 14: 420-426.

Giraldez AJ, Mishima Y, Rihel J, Grocock RJ, Van Dongen S, Inoue K, Enright AJ, Schier AF (2006). Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. Science 312: 75-79.

Güçlü S (2013). Analysis of global microRNAome profiles in *Caenorhabditis elegans* oocytes and embryos. MSc, Ankara University, Ankara, Turkey (in Turkish with abstract in English).

Hackenberg M, Sturm M, Langenberger D, Falcon Perez JM, Aransay AM (2009). miRanalyzer: a microRNA detection and analysis tool for next-generation sequencing experiments. Nucleic Acids Res 37: W68-W76.

Ibáñez Ventoso C, Vora M, Driscoll M (2008). Sequence relationships among *C. elegans*, *D. melanogaster* and human microRNAs highlight the extensive conservation of microRNAs in biology. PLoS One 3: e2818.

Lee RC, Ambros V (2001). An extensive class of small RNAs in *Caenorhabditis elegans*. Science 294: 862-864.

Nayak S, Goree J, Schedl T (2005). fog-2 and the evolution of self-fertile hermaphroditism in *Caenorhabditis*. PLoS Biol 3: e6.

Nogales Cadenas R, Carmona Saez P, Vazquez M, Vicente C, Yang X, Tirado F, Carazo JM, Pascual Montano A (2009). GeneCodis: interpreting gene lists through enrichment analysis and integration of diverse biological information. Nucleic Acids Res 37: W317-W322.

Ruby JG, Jan C, Player C, Axtell MJ, Lee W, Nusbaum C, Ge H, Bartel DP (2006). Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in *C. elegans*. Cell 127: 1193-1207.

Semotok JL, Cooperstock RL, Pinder BD, Vari HK, Lipshitz HD, Smibert CA (2005). Smaug recruits the CCR4/POP2/NOT deadenylase complex to trigger maternal transcript localization in the early *Drosophila* embryo. Curr Biol 15: 284-294.

Stiernagle T (2006). Maintenance of *C. elegans*. WormBook 2006: 1-11.

Stoeckius M, Grun D, Kirchner M, Ayoub S, Torti F, Piano F, Herzog M, Selbach M, Rajewsky N (2014). Global characterization of the oocyte-to-embryo transition in *Caenorhabditis elegans* uncovers a novel mRNA clearance mechanism. EMBO J 33: 1751-1766.

Stoeckius M, Maaskola J, Colombo T, Rahn HP, Friedländer MR, Li N, Chen W, Piano F, Rajewsky N (2009). Large-scale sorting of *C. elegans* embryos reveals the dynamics of small RNA expression. Nat Methods 6: 745-751.

Tabas Madrid D, Nogales Cadenas R, Pascual Montano A (2012). GeneCodis3: a non-redundant and modular enrichment analysis tool for functional genomics. Nucleic Acids Res 40: W478-W483.