Published in final edited form as:

Genesis. 2012 March ; 50(3): 286–299. doi:10.1002/dvg.22010.

Identification of novel microRNAs in *Xenopus laevis* metaphase II arrested eggs

Sakthikumar Ambady¹, Zheyang Wu², and Tanja Dominko³

¹Department of Biomedical Engineering, Worcester Polytechnic Institute, Worcester, MA 01609

²Department of Mathematical Sciences, Worcester Polytechnic Institute, Worcester, MA 01609

³Department of Biology and Biotechnology, Worcester Polytechnic Institute, Worcester, MA 01609

Abstract

Using a combination of deep sequencing and bioinformatics approach, we for the first time identify miRNAs and their relative abundance in mature, metaphase II arrested eggs in *X. laevis*. We characterize 115 miRNAs that have been described either in *X. tropicalis* (85), *X. laevis* (9) or other vertebrate species (21) that also map to known *Xenopus* pre-miRNAs and to the *X. tropicalis* genome. Additionally, 72 new *X. laevis* putative candidate miRNAs are identified based on mapping to *X. tropicalis* genome within regions that have the propensity to form hairpin loops. These data expand on the availability of genetic information in *X. laevis* and identifies target miRNAs for future functional studies.

Keywords

Deep sequencing; bioinformatics; miRBase; genetics; amphibian

INTRODUCTION

In recent years, microRNAs (miRNAs) have emerged as an important class of gene regulators. They mediate complex post-transcriptional regulatory activities by interacting with coding and non-coding regions of messenger RNAs (Abdelmohsen *et al.*, 2008; Tay *et al.*, 2008). While they recognize their targets with incomplete complementarity, the hallmark of these short sequences is perfect base pairing of nucleotides 2–7 at the 5' end of a mature miRNA, referred to as “seed” sequence, to its target mRNA. This seed-mediated binding is required for modulation of mRNA expression (Doench and Sharp, 2004; Grimson *et al.*, 2007). Their initial role in post-transcriptional repression was more recently expanded to include post-transcriptional activation/derepression (Cordes *et al.*, 2009; Mortensen *et al.*, 2011) and mRNA stability (Valencia-Sanchez *et al.*, 2006).

Biogenesis of miRNAs is quite elaborate. Primary miR transcripts are produced mostly by RNA Pol II as kilobase long transcripts from either intronic sequences of protein coding genes or from their own transcription units that are not associated with protein coding genes. The primary transcripts (pri-miRNAs) are processed in the nucleus by RNase-III enzyme Drosha and cleaved into 60–100 nt long pre-miRNAs that form a secondary hairpin

Correspondence: Sakthikumar Ambady, Worcester Polytechnic Institute, 100 Institute Road, Worcester, MA 01609 Tel. 508-831-5714, sambady@wpi.edu.

The authors have no conflict of interest to declare.

structure. The pre-miRNAs are exported into the cytoplasm and processed to double stranded stretches of miRNAs by another RNase-III enzyme, Dicer. The passenger strands are removed and degraded and the remaining guide strands are loaded onto RNA induced silencing complex (RISC) where they become single stranded mature miRNAs (reviewed by (Kim *et al.*, 2009). The location of their interaction within the target mRNA as well as the degree of complementarity determines their activity as translational repressors, translational activators or mRNA stabilizers.

MicroRNAs participate in various cellular functions and are increasingly being identified as diagnostic markers in several diseases and cancers (Brase *et al.*, 2010; Chen *et al.*, 2011; Cortez and Calin, 2009; Kroh *et al.*, 2010; Lawrie *et al.*, 2008; Mahn *et al.*, 2011; Shah *et al.*, 2009; Weigel and Dowsett, 2010; Wittmann and Jäck, 2010; Zen and Zhang, 2010). The role of miRNAs in modification of chromatin remodeling complexes and their ability to promote transdifferentiation of cell phenotype have been described more recently (Cardinali *et al.*, 2009; Chen *et al.*, 2006; Fineberg *et al.*, 2009; Yoo *et al.*, 2009; Yoo *et al.*, 2011). MicroRNAs exhibit a high level of conservation across species (Altuvia *et al.*, 2005; Liu *et al.*, 2008). Computational prediction of miRNAs suggests thousands of miRNA molecules for human and other species (Mazière and Enright, 2007; Olena and Patton, 2010). The most current version of miRBase contains 16,772 entries representing hairpin precursor miRNAs, expressing 19,724 mature miRNA products in 153 species (miRBase Registry, release 17; (Kozomara and Griffiths-Jones, 2011).

The African clawed frog *X. laevis* is an important model organism that has been used in developmental biology research for decades. *X. laevis* egg extracts have been invaluable in studying biological processes such as chromatin remodeling and acquisition of transcriptional competence (Blow and Laskey, 1986; Dimitrov and Wolffe, 1996; Kikyo *et al.*, 2000; Lohka and Masui, 1983a, 1983b), cell cycle (Lohka, 1989), and DNA replication (Blow, 1993). One of the most compelling abilities of *X. laevis* metaphase II arrested egg extracts is its ability to reprogram differentiated somatic cells into stem cell gene expressing cells (Alberio *et al.*, 2005; Byrne *et al.*, 2003; Gurdon *et al.*, 2005; Gurdon *et al.*, 1958; Hansis *et al.*, 2004; Miyamoto *et al.*, 2007). While an expanding database of *X. laevis* genomic and genetic information is emerging, sequencing of the *X. laevis* genome has not yet been completed. Similarly, the *X. laevis* transcriptome, small RNAome and proteome remain incomplete compared to other species. For example, 1,902 mature miRNAs have been published for human, 207 for *X. tropicalis* and only 22 for *X. laevis* (miRBase, version 17). All the miRNA sequences in *X. laevis* are derived from a single published study (Watanabe *et al.*, 2005) and from unpublished data from Biasci *et al.* 2008 (miRBase).

Traditionally, miRNAs have been discovered by cloning of small RNAs (Watanabe *et al.*, 2005) following the commonly used Sanger sequencing methods. However, this methodology has limited application in discovering low abundant miRNAs. Another avenue of miRNA discovery is computational analysis based on RNA secondary structure predictions and sequence conservation across species. Even though computational methods are useful in predicting miRNAs, expression of these transcripts still needs to be verified by experimentation. Recent advances in high throughput sequencing technologies are paving the way for discovering new and rare species of expressed small RNAs. Illumina sequencing (formerly Solexa) is an ideal platform for miRNA discovery due to the 35 bp read lengths and the level of sequencing depth afforded by this platform (Szittyá *et al.*, 2008). Therefore, a combination of deep sequencing and computational analysis is an attractive proposition for identifying new and rarely expressed, non-conserved and species-specific miRNAs.

Here we undertook a deep sequencing approach to identify miRNAs in metaphase II arrested *X. laevis* eggs. Combined with bioinformatics and interrogation of genomic sequences

available for *X. tropicalis* we characterize populations of miRNAs in metaphase II egg extracts, describe their likely precursor sequences (pre-miRNAs), identify putative new miRNAs, map their locations to the genomic scaffolds of *X. tropicalis*, and identify their mRNA targets.

RESULTS

Small RNA reads in *X. laevis* metaphase II arrested eggs

A total of 12,526,420 raw reads were obtained from sequencing short RNAs from *X. laevis* metaphase II arrested eggs. Reads were filtered to 11,302,087 mappable reads using the criteria described in Table 1 and assigned to groups described in detail in Figure 1. Only reads between 15–24 nucleotides, corresponding to conventionally accepted miRNA length, and mapping perfectly to the available *X. tropicalis* genome scaffolds were included in the dataset. Distribution of small reads is presented in Figure 3. All identified sequences were able to fold into the hairpin-loop structure characteristic of a folded pre-miRNA. As *X. laevis* genome sequence data becomes available, additional sequences identified (but not presented) in this study may be revisited in the future. The comprehensive dataset is included in Table S1 and available at <http://users.wpi.edu/~dominko/XenopusProject/>.

A total of 115 unique reads in the 15–24 nucleotide size range mapped to mature miRNAs and pre-miRNAs of *X. laevis*, *X. tropicalis* and other vertebrate species in the current version of miRBase and pre-miRNAs further mapped to the *X. tropicalis* genome (Figure 2 Groups 1a,b,c and Group 2). Additional 72 putative candidate – PC miRNAs were identified that did not map to any known pre-miRNAs, but mapped to *X. tropicalis* genome within sequences with propensity to form hairpin structures (Figure 2 Group 4).

Reads mapping to known miRNAs in *X. laevis* or *X. tropicalis* (Group1a)

A total of 94 reads mapped to previously described *X. laevis* (Table 2) or *X. tropicalis* (Table 3) miRNAs/pre-miRNAs in miRBase (miRBase, Release 17) and these pre-miRNAs further mapped to *X. tropicalis* genome (JGI v4.1, www.jgi.doe.gov).

Thirty-nine of the 94 *Xenopus* miRNA identified here (4 out of 9 *laevis* and 35 out of 85 *tropicalis*) showed a perfect match, while the remaining reads contained changes when compared to *Xenopus* miRNAs in miRBase. These changes most frequently included single nucleotide additions or deletions to either the 5' or 3' end of mature reads, base substitutions or a combination of these changes. The majority of the reads containing changes had a single or double base addition or deletion at their 3' end (5 in *laevis* and 39 in *tropicalis*). Three reads contained a single nucleotide substitution. None of these changes affected the seed sequence, located at the 5' end of the miRNA (5' nucleotides 2 through 8). Four reads contained base additions or deletions at their 5' ends and four reads contained changes at both ends; these changes consequently altered the seed sequence. Eight reads showed a 5' or 3' location shift on known *Xenopus* pre-miRNAs (Table S2).

Reads mapping to other known vertebrate (other than *Xenopus*) miRNAs and pre-miRNAs (Group 1b)

Three of twenty-one reads that mapped to *X. tropicalis* genome matched sequences described in zebrafish (dre-mir-24) and human (hsa-mir-98 and hsa-miR-129). Eleven reads matched miRNAs in other species and mapped to new regions within previously described pre-miRNAs; and seven reads matched miRNAs in other species but mapped to novel regions in *X. tropicalis* genome that have not been previously identified as pre-miRNAs (Table 4). Hence, these last seven represent potentially new miRNAs and new pre-miRNAs

for *X. tropicalis* and *X. laevis*. The genomic sequences flanking the reads could form stable hairpin structures (Table S2).

Reads mapping to *X. tropicalis* genomic regions – “putative candidate” miRNAs (Group 4)

A third category of 72 unique reads did not map to any known vertebrate miRNAs or pre-miRNAs in miRBase (Table 5). However they all mapped to *X. tropicalis* genome in regions that showed propensity to form thermodynamically stable hairpin structures. These are considered Putative Candidate – PC - miRNAs for *X. laevis* and *X. tropicalis*. Importantly, none of these mapped to *Xenopus* genomic repeats (rephase). Fifty-eight reads mapped to a single locus and 15 mapped to multiple loci on *X. tropicalis* scaffolds (Table S3).

Genomic clustering and expression of *X. laevis* miRNAs

The location of the reads was analyzed within *X. tropicalis* genomic scaffolds. Clustering was examined for multiple copies of the same read or single copies of multiple reads. Any reads that localized within approximately 15 kilobases within the same genome ID were assigned to a cluster (Table 6 and Table S3). About half of all the known miRNAs localized to one of the clusters containing described previously (Tang and Maxwell, 2008). The size of clusters ranged from 142 bp (cluster #2) to 14,946 bp (cluster #1). The largest cluster (#5) contained 8 known miRNAs, 3 of which have not been previously described for *X. laevis* (Table S3). As observed previously, some miRNA localized to multiple genomic loci, and the rest of the miRNAs appeared to be single copy genes.

Expression of clustered miRNAs varied significantly between different clusters. Some clustered miRNAs exhibited similar expression pattern (e.g. xtr-miR-130b/130c in cluster #11 and miR-23a/27a in cluster #8), consistent with co-transcription of a common pri-miRNA (Baskerville and Bartel, 2005; Tang and Maxwell, 2008). Examination of other *X. laevis* clustered miRNAs, however, indicated their post-transcriptional regulation. One of the most striking examples was expression of seven miRNAs from the 724 bp cluster #5 (miR-363/363-3p/92a/19b/20/20b/18/106). The number of copies for individual miRNAs in this cluster ranged from 1 to 461 (Table S3). If these miRNAs are expressed from a single pri-miRNA then differential processing of the pri-miRNA may be responsible for the observed variation. Alternatively, different miRNAs in the cluster could be expressed under control of different promoters. We observed other expression imbalances including miR-98-p5 and miR-let-7f in cluster #14 (2 and 2162 copies, respectively). Interestingly, both miRNAs are co-expressed from the same pri-miRNA and demonstrate significant quantitative difference in all *X. laevis* tissues examined but could not be detected in *X. laevis* oocytes (Tang and Maxwell, 2008). The expression of miRNA from a specific cluster could reflect developmental stage-specific transcription or differential pre-miRNA post-transcriptional processing.

Only five PC miRNAs localized to 4 clusters, each containing multiple copies of the same PC miRNA (clusters #26, 27 and 29) or two copies of two PC miRNAs (cluster #28). Expression of these PC miRNAs was low and differential expression from different loci could not be determined by sequencing alone. Cluster #25 contained one known miRNAs and one PC miRNA (Table 6).

Some miRNAs are expressed from the same pre-miRNA sequence

Mapping of reads to the genome identified sequences of 17 pre-miRNAs that expressed 2 miRNAs each. Mature and star miRNAs (*) and those with the suffix 5p, 3p and p5, p3 represent miRNAs derived from different regions of the same pre-miRNA (Table S3). We observed differential expression of miRNA pairs derived from the same pre-miRNA. Among the known miRNAs, the expression ratio between the pairs varied widely. For

example, xtr-let-7a and hsalet-7a* were expressed at a ratio of 1055 to 1 while xtr-miR-132 and newly identified hsa-miR-132* exhibited similar copy numbers (Table S3). Expression of the majority of PC miRNA pairs expressed from the same pre-miRNA showed much less variation, possibly due to their relatively low levels of expression.

Target predictions for putative candidate miRNAs and seed-shifted miRNAs

Target prediction was performed using custom TargetScan (Release 5.2, June 2011) based on complementarity between the seed sequences of mature miRNAs (positions 2–8) and frog genes for 72 PC miRNAs identified in the study. Four reads contained seed sequences perfectly matched to seed sequences of known miRNAs (miR-106/93a/302; miR-199/199-5p; miR-139-5p; miR-1/206). Seed sequences of 68 PC miRNAs did not match previously identified miRNA seed sequences but all recognized frog transcript 3' UTR sequences. The number of target genes for individual reads ranged between 1 and 142 (Table S4).

Target prediction was also performed for 12 known miRNAs that showed “seed shift” occurring as a result of base changes (addition or deletion) at the 5' end of the reads (Table S5). Shift in the seed sequence for any miRNA resulted in loss of some target genes, acquisition of some new target genes and in maintenance of some of the same target genes. For example, comparison of target genes for xtr-miR-124_L-1R+1 (242 targets) and its representative xtr-miR-124 (174 targets) reveals that 117 targets remain the same even though the seed sequence has shifted. As a diametrically opposite example, seed sequence of hsa-miR-132* doesn't have any reported targets in *Xenopus* genome, but its seed-shifted variant hsa-miR-132*_L-1, detected in *Xenopus* eggs, recognizes three gene targets.

DISCUSSION

Xenopus laevis represents an important model system in cell and developmental biology. In particular, the availability of its large size oocytes and eggs has established *X. laevis* as a unique source of cytoplasmic extracts. While much has been learned about its smaller cousin *X. tropicalis*, genomic and genetic information for *X. laevis* remains largely unavailable. Here we report the first complete analysis of miRNA populations in *X. laevis* metaphase II arrested eggs using deep sequencing and bioinformatics analysis. As deep sequencing approach is independent of known genome sequence, our data will remain suitable for further analysis when the genome information for *X. laevis* becomes available.

Available bioinformatics tools and approaches for analysis of deep sequencing data in any species rely on comparison of identified sequences to an existing genome sequence database. As sequencing of the genome of *X. laevis* is still in progress, our analysis included only sequences matching to the *X. tropicalis*. The genome assembly for *X. tropicalis* is available (JGI v4.1, www.jgi.doe.gov) with an estimated coverage of 90% on 19,759 scaffolds. While still not assembled into chromosomes, this resource represents the only tool for sequence comparison and analysis for *X. laevis*. Even though they share the same genus, significant differences between the two genomes exist (Cannatella and de S, 1993). *X. laevis* has a genome size of 3.1 gb distributed over 18 pairs of chromosomes, while *X. tropicalis* is diploid with a genome size of 1.7 gb distributed over 10 pairs of chromosomes.

Due to the differences in ploidy and chromosome number, it is reasonable to expect that comparison between *X. laevis* and *X. tropicalis* will identify sequence variations. The majority of *Xenopus* miRNA sequences in miRBase have been deposited based on similarity to vertebrate miRNAs. Recent reports on deep sequencing data for other species indicate that variations have been observed compared to miRBase sequences (Fernandez-Valverde *et al.*, 2010; Lee *et al.*, 2010). Such variations are attributed to RNA degradation, sequencing

errors, and to differences in miRNA processing in a cell- and tissue-dependent manner. In an allopolyploid species such as *X. laevis*, in which genomic sequence information is scant, it is difficult to tease out the exact causes of observed sequence variations. The variations observed in this study possibly represent *X. laevis* specific variations. It is also possible that factors such as handling of small RNAs, depth of sequencing, and default factors set for specific algorithms used for analysis can impact sequence analysis results. Alternatively, dynamic expression of miRNAs can change during development and the identity of small RNA species captured at a particular developmental stage will vary. Twenty-five miRNAs we found in common with Armisen et al. (2009) contained the same sequence variations reported here. Further probing into *Xenopus* miRNA sequences in miRBase revealed that a majority of the sequences were deposited purely based on sequence similarity.

We confirmed expression of many miRNAs described previously in adult *X. laevis* tissues (Michalak and Malone, 2008; Tang and Maxwell, 2008), but describe expression of the majority of these miRNAs now also for the metaphase II arrested *X. laevis* eggs. Some previous studies failed to detect a number of miRNAs in *X. laevis* oocytes and eggs by Northern blotting (Tang and Maxwell, 2008). This is not surprising as the method may not be sensitive enough for low expressing miRNAs. Among the 22 known *X. laevis* miRNAs in miRBase, only 7 have been verified by cloning (Watanabe *et al.*, 2005) and the remaining 15 were presented based on similarity (Biasci *et al.*, 2008, unpublished). Of the 191 *X. tropicalis* miR/pre-miRNAs in miRBase, only 2 sequences have been verified by sequencing (Armisen *et al.*, 2009), forty one miRNAs were verified by Northern blotting (Tang and Maxwell, 2008), and 148 were deposited based on sequence similarity. Therefore miRBase sequences may not be a true representation of expressed *Xenopus* miRNAs. Identical sequence observed in the present study and the study of Armisen et al. (2009), particularly of those reads with sequence changes points to the importance of verifying miRNAs by sequencing before miRBase submission. Of the 94 reads only 9 correspond to *X. laevis* miRs in miRBase. The remaining 85 are therefore new xla-miRs.

Due to high conservation of miRNAs across different species it is not surprising that additional expressed reads in our study matched miRNA sequences described in other vertebrates. Metaphase II arrested eggs represent a very unique developmental stage during oogenesis, and these miRNAs may be specific to this developmental stage. Conversely, we identified expression of miRNAs in metaphase II eggs that were believed to be testis specific, e.g. miR-100 (Michalak and Malone, 2008), or associated with central nervous system miR-124 and miR-23b/miR-24a/miR27b cluster (Walker and Harland, 2008).

The 72 putative candidate miRNA reads identified in this study are genuine expressed sequences with strong secondary structure prediction which makes them putative candidate miRNAs for both *X. laevis* and *X. tropicalis*. Because they are unique to *Xenopus* they are most likely genus-specific. It is unlikely these reads belong to piRNAs, the predominant species of small RNAs in oocytes (Armisen *et al.*, 2009). None of the reads map to genomic repeats (Okamura and Lai, 2008), they are shorter than 25 nucleotides (Grimson *et al.*, 2007), and their precursor genomic sequences have the propensity to form hairpin structures.

Our data indicate the presence of a repertoire of miRNAs that show high level of conservation with *X. tropicalis* and a subset that show sequence differences at the RNA level. Some of the identified miRNAs show differences with *X. tropicalis* genomic sequences suggesting that these sequences may be derived from *X. laevis* specific loci. Therefore it is possible that miRNAs are expressed from one or more loci from multiple paralogs present in *X. laevis* genome. Redundant genes, present in allopolyploid *X. laevis*, could exhibit preferential silencing or differential expression in a cell, tissue or stage-specific manner for the duplicate paralogs, as well as exhibit heterochrony (Adams *et al.*,

2003; Evans, 2007; Hellsten *et al.*, 2007; Morin *et al.*, 2006; Morin *et al.*, 2008; Sémon and Wolfe, 2008; Yanai *et al.*, 2011). Contrary to the gene silencing theory, there is evidence to suggest that duplicated gene functions are preserved rather than silenced (Force *et al.*, 1999; Lynch and Force, 2000; Otto and Yong, 2002).

We also observed many miRNA polymorphisms indicating that they are expressed from more than two genomic loci. The numbers of miR variants, sequence variations and copy numbers are indicative of their origin from different loci. It is unusual to find such a high level of variation in a small stretch of genomic DNA and cannot be explained away as sequencing errors. We hypothesize that the variants with multiple sequence substitutions are derived from *X. laevis* specific loci. However, lack of availability of corresponding *X. laevis* genomic sequences makes it difficult to confirm our hypothesis. Since we are mapping the sequences only to *X. tropicalis*, it is also likely that we are missing out on a number of *X. laevis*-specific miRNAs simply due to the fact that reference sequences are not available.

Deep sequencing allows not only the identification of miRNAs, but also reflects the absolute expression levels based on copy number. The copy numbers of miRs that are derived from the same pre-miR differ considerably, which is indicative of the dynamics and fine-tuned modulation of miRNA expression despite the same genomic region.

One feature of miRs that has not been discussed extensively in published literature is the significance of miRs mapping to multiple loci. Many miRBase sequences show this trend. We observed that 18 known *Xenopus* miRNAs identified in this study (and reported in miRBase) map to multiple genomic coordinates on *X. tropicalis* genome. Sequences flanking the miRNA at all these loci could form thermodynamically stable hairpin structures. Bioinformatics predictions and sequence information of mature sequences alone do not provide sufficient information regarding the locus from which these miRNAs are expressed. The same miRNA could be expressed from different loci depending on its co-localization with gene clusters expressed in specific cell types or at different stages during development, as the majority of *Xenopus* miRNAs are intronic in nature (Tang and Maxwell, 2008). This may reflect a mechanism that allows coordinated expression of genes and corresponding miRNAs in a cell, tissue or stage-specific manner.

Some reads differed from their representative miRNAs by a 5' single nucleotide addition or deletion, therefore shifting the seed sequence. As a consequence, the loss of gene targets, acquisition of new targets and maintenance of some of the same targets can be expected for the shifted miRNA. Depending on the presence of a 3' adenosine preceding the mRNA target sequence (7mer 1A) or true complementarity of the 8th nucleotide in the miRNA seed sequence with the target mRNA (m8), the seed-shifted miRNA can still pair with the same mRNAs (Grimson *et al.*, 2007). These single nucleotide changes in the 5' miRNA sequences may be critically important in the context of developmental regulation.

The current study adds a significant number of novel miRNA sequences to the currently available database for *X. laevis*. However, much experimental work remains to be performed in order to confirm their identity, their expression and their functional significance during oocyte maturation, fertilization and early embryonic development.

METHODS

Preparation of *Xenopus* egg extract and RNA isolation

Five adult *X. laevis* females were superovulated and metaphase II arrested eggs collected, processed and maintained in MII arrest following published protocols (Blow, 1993). Total RNA was isolated using TRIZOL reagent (Invitrogen). Total RNA was resuspended in

RNase/DNase free water and treated with RNase-free DNase (Ambion) to eliminate any traces of contaminating DNA. The mixture was re-extracted with TRIZOL and resuspended in RNase/DNase free water. RNA quality was assessed by gel analysis and Bioanalyzer (Agilent Technologies).

Small RNA sequencing

Deep sequencing of small RNAs extracted from metaphase II arrested *X. laevis* eggs was performed at LC Sciences (Houston, TX) using Illumina's Genome analyzer II platform. Small RNAs were size-fractionated to isolate RNA in the 15–50 nt range and ligated to SRA5' adapter and size fractionated to isolate 41–76 nt fraction. This was followed by SRA3' adapter ligation and size fractionation to isolate 64–99 nt size RNA. This RNA was reverse transcribed to produce single stranded cDNA and PCR amplified for 20 cycles using Illumina's primer set. This pool was once again size-fractionated to isolate cDNAs in 80 to 115 bp range that represented miRNA. This fraction was eluted, precipitated and quantified on Nanodrop (Thermo Scientific). The concentration of cDNA sample was adjusted to 10 nM and 10 µl was used for sequencing reaction. Purified cDNA was used for cluster generation on Illumina's Cluster Station and sequenced using Illumina GAIIx platform. Raw sequencing reads were obtained using Illumina's Pipeline v1.5 software following sequencing image analysis by Pipeline Firecrest Module and base-calling by Pipeline Bustard Module. Filtering of sequences was done as presented in Figure 1.

Computational analysis

Raw sequences were analyzed using the ACGT101-miR v4.2 software package (LC Sciences, Houston, TX). Several digital filters were applied to remove sequencing errors and eliminate unmappable reads. Unique reads in the 15 to 24 nucleotide size range were used for mapping to vertebrate microRNAs/pre-miRNAs (miRBase, Release 17, <http://www.mirbase.org/index.shtml>), to the *X.tropicalis* genome (JGI, version 4.1, <http://genome.jgi-psf.org>) and to *Xenopus* ESTs (Xenbase 2.7, <http://www.xenbase.org>). A computational approach was utilized to predict new miRNAs based on mapping of unique reads to *X. tropicalis* genomic sequences. Flanking sequences of mapped reads were subjected to secondary structure analysis to predict pre-miRNA sequences. Target prediction for putative candidate miRNAs and those miRNAs that showed a seed shift due to additions and deletions of bases on their 5' end was performed using custom target scan (http://www.targetscan.org/vert_50/seedmatch.html).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Support:

This work was funded by NIH award R01GM085456 to TD.

REFERENCES

- Abdelmohsen K, Srikantan S, Kuwano Y, Gorospe M. miR-519 reduces cell proliferation by lowering RNA-binding protein HuR levels. *Proc Natl Acad Sci U S A*. 2008; 105:20297–20302. [PubMed: 19088191]
- Adams KL, Cronn R, Percifield R, Wendel JF. Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. *Proc Natl Acad Sci U S A*. 2003; 100:4649–4654. [PubMed: 12665616]

- Alberio R, Johnson AD, Stick R, Campbell KH. Differential nuclear remodeling of mammalian somatic cells by *Xenopus laevis* oocyte and egg cytoplasm. *Exp Cell Res*. 2005; 307:131–141. [PubMed: 15922733]
- Altuvia Y, Landgraf P, Lithwick G, Elefant N, Pfeffer S, Aravin A, Brownstein MJ, Tuschl T, Margalit H. Clustering and conservation patterns of human microRNAs. *Nucleic Acids Res*. 2005; 33:2697–2706. [PubMed: 15891114]
- Armisen J, Gilchrist MJ, Wilczynska A, Standart N, Miska EA. Abundant and dynamically expressed miRNAs, piRNAs, and other small RNAs in the vertebrate *Xenopus tropicalis*. *Genome Res*. 2009; 19:1766–1775. [PubMed: 19628731]
- Baskerville S, Bartel DP. Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *Rna*. 2005; 11:241–247. [PubMed: 15701730]
- Blow JJ. Preventing re-replication of DNA in a single cell cycle: evidence for a replication licensing factor. *J Cell Biol*. 1993; 122:993–1002. [PubMed: 8354699]
- Blow JJ, Laskey RA. Initiation of DNA replication in nuclei and purified DNA by a cell-free extract of *Xenopus* eggs. *Cell*. 1986; 47:577–587. [PubMed: 3779837]
- Brase JC, Wuttig D, Kuner R, Sülthmann H. Serum microRNAs as non-invasive biomarkers for cancer. *Mol Cancer*. 2010; 9:306. [PubMed: 21110877]
- Byrne JA, Simonsson S, Western PS, Gurdon JB. Nuclei of adult mammalian somatic cells are directly reprogrammed to oct-4 stem cell gene expression by amphibian oocytes. *Curr Biol*. 2003; 13:1206–1213. [PubMed: 12867031]
- Cannatella DC, de S RO. *Xenopus Laevis* as a Model Organism. *Systematic Biology*. 1993; 42:476–507.
- Cardinali B, Castellani L, Fasanaro P, Basso A, Alemà S, Martelli F, Falcone G. MicroRNA-221 and microRNA-222 modulate differentiation and maturation of skeletal muscle cells. *PLoS One*. 2009; 4:e7607. [PubMed: 19859555]
- Chen JF, Mandel EM, Thomson JM, Wu Q, Callis TE, Hammond SM, Conlon FL, Wang DZ. The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. *Nat Genet*. 2006; 38:228–233. [PubMed: 16380711]
- Chen X, Hu Z, Wang W, Ba Y, Ma L, Zhang C, Wang C, Ren Z, Zhao Y, Wu S, Zhuang R, Zhang Y, Hu H, Liu C, Xu L, Wang J, Shen H, Zhang J, Zen K, Zhang CY. Identification of ten serum microRNAs from a genome-wide serum microRNA expression profile as novel non-invasive biomarkers for non-small cell lung cancer diagnosis. *Int J Cancer*. 2011
- Cordes KR, Sheehy NT, White MP, Berry EC, Morton SU, Muth AN, Lee TH, Miano JM, Ivey KN, Srivastava D. miR-145 and miR-143 regulate smooth muscle cell fate and plasticity. *Nature*. 2009; 460:705–710. [PubMed: 19578358]
- Cortez MA, Calin GA. MicroRNA identification in plasma and serum: a new tool to diagnose and monitor diseases. *Expert Opin Biol Ther*. 2009; 9:703–711. [PubMed: 19426115]
- Dimitrov S, Wolffe AP. Remodeling somatic nuclei in *Xenopus laevis* egg extracts: molecular mechanisms for the selective release of histones H1 and H1(0) from chromatin and the acquisition of transcriptional competence. *EMBO J*. 1996; 15:5897–5906. [PubMed: 8918467]
- Doench JG, Sharp PA. Specificity of microRNA target selection in translational repression. *Genes Dev*. 2004; 18:504–511. [PubMed: 15014042]
- Evans BJ. Ancestry influences the fate of duplicated genes millions of years after polyploidization of clawed frogs (*Xenopus*). *Genetics*. 2007; 176:1119–1130. [PubMed: 17435227]
- Fernandez-Valverde SL, Taft RJ, Mattick JS. Dynamic isomiR regulation in *Drosophila* development. *RNA*. 2010; 16:1881–1888. [PubMed: 20805289]
- Fineberg SK, Kosik KS, Davidson BL. MicroRNAs potentiate neural development. *Neuron*. 2009; 64:303–309. [PubMed: 19914179]
- Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J. Preservation of duplicate genes by complementary, degenerative mutations. *Genetics*. 1999; 151:1531–1545. [PubMed: 10101175]
- Grimson A, Farh KK, Johnston WK, Garrett-Engle P, Lim LP, Bartel DP. MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol Cell*. 2007; 27:91–105. [PubMed: 17612493]

- Gurdon JB, Byrne JA, Simonsson S. Nuclear reprogramming by xenopus oocytes. Novartis Found Symp. 2005; 265:129–136. discussion 136–141, 204–111. [PubMed: 16050254]
- Gurdon JB, Elsdale TR, Fischberg M. Sexually mature individuals of *Xenopus laevis* from the transplantation of single somatic nuclei. Nature. 1958; 182:64–65. [PubMed: 13566187]
- Hansis C, Barreto G, Maltry N, Niehrs C. Nuclear reprogramming of human somatic cells by xenopus egg extract requires BRG1. Curr Biol. 2004; 14:1475–1480. [PubMed: 15324664]
- Hellsten U, Khokha MK, Grammer TC, Harland RM, Richardson P, Rokhsar DS. Accelerated gene evolution and subfunctionalization in the pseudotetraploid frog *Xenopus laevis*. BMC Biol. 2007; 5:31. [PubMed: 17651506]
- Kikyo N, Wade PA, Guschin D, Ge H, Wolffe AP. Active remodeling of somatic nuclei in egg cytoplasm by the nucleosomal ATPase ISWI. Science. 2000; 289:2360–2362. [PubMed: 11009424]
- Kim VN, Han J, Siomi MC. Biogenesis of small RNAs in animals. Nat Rev Mol Cell Biol. 2009; 10:126–139. [PubMed: 19165215]
- Kozomara A, Griffiths-Jones S. miRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res. 2011; 39:D152–D157. [PubMed: 21037258]
- Kroh EM, Parkin RK, Mitchell PS, Tewari M. Analysis of circulating microRNA biomarkers in plasma and serum using quantitative reverse transcription-PCR (qRT-PCR). Methods. 2010; 50:298–301. [PubMed: 20146939]
- Lawrie CH, Gal S, Dunlop HM, Pushkaran B, Liggins AP, Pulford K, Banham AH, Pezzella F, Boultonwood J, Wainscoat JS, Hatton CS, Harris AL. Detection of elevated levels of tumour-associated microRNAs in serum of patients with diffuse large B-cell lymphoma. Br J Haematol. 2008; 141:672–675. [PubMed: 18318758]
- Lee LW, Zhang S, Etheridge A, Ma L, Martin D, Galas D, Wang K. Complexity of the microRNA repertoire revealed by next-generation sequencing. RNA. 2010; 16:2170–2180. [PubMed: 20876832]
- Liu N, Okamura K, Tyler DM, Phillips MD, Chung WJ, Lai EC. The evolution and functional diversification of animal microRNA genes. Cell Res. 2008; 18:985–996. [PubMed: 18711447]
- Lohka MJ. Mitotic control by metaphase-promoting factor and cdc proteins. J Cell Sci. 1989; 92(Pt 2): 131–135. [PubMed: 2674165]
- Lohka MJ, Masui Y. Formation in vitro of sperm pronuclei and mitotic chromosomes induced by amphibian ooplasmic components. Science. 1983a; 220:719–721. [PubMed: 6601299]
- Lohka MJ, Masui Y. The germinal vesicle material required for sperm pronuclear formation is located in the soluble fraction of egg cytoplasm. Exp Cell Res. 1983b; 148:481–491. [PubMed: 6605256]
- Lynch M, Force A. The probability of duplicate gene preservation by subfunctionalization. Genetics. 2000; 154:459–473. [PubMed: 10629003]
- Mahn R, Heukamp LC, Rogenhofer S, von Ruecker A, Müller SC, Ellinger J. Circulating microRNAs (miRNA) in serum of patients with prostate cancer. Urology. 2011; 77:1265.e1269–1265.e1216. [PubMed: 21539977]
- Mazière P, Enright AJ. Prediction of microRNA targets. Drug Discov Today. 2007; 12:452–458. [PubMed: 17532529]
- Michalak P, Malone JH. Testis-derived microRNA profiles of African clawed frogs (*Xenopus*) and their sterile hybrids. Genomics. 2008; 91:158–164. [PubMed: 18079091]
- Miyamoto K, Furusawa T, Ohnuki M, Goel S, Tokunaga T, Minami N, Yamada M, Ohsumi K, Imai H. Reprogramming events of mammalian somatic cells induced by *Xenopus laevis* egg extracts. Mol Reprod Dev. 2007; 74:1268–1277. [PubMed: 17474094]
- Morin RD, Chang E, Petrescu A, Liao N, Griffith M, Chow W, Kirkpatrick R, Butterfield YS, Young AC, Stott J, Barber S, Babakaiff R, Dickson MC, Matsuo C, Wong D, Yang GS, Smailus DE, Wetherby KD, Kwong PN, Grimwood J, Brinkley CP, Brown-John M, Reddix-Dugue ND, Mayo M, Schmutz J, Beland J, Park M, Gibson S, Olson T, Bouffard GG, Tsai M, Featherstone R, Chand S, Siddiqui AS, Jang W, Lee E, Klein SL, Blakesley RW, Zeeberg BR, Narasimhan S, Weinstein JN, Pennacchio CP, Myers RM, Green ED, Wagner L, Gerhard DS, Marra MA, Jones SJ, Holt RA. Sequencing and analysis of 10,967 full-length cDNA clones from *Xenopus laevis*

- and *Xenopus tropicalis* reveals post-tetraploidization transcriptome remodeling. *Genome Res.* 2006; 16:796–803. [PubMed: 16672307]
- Morin RD, O'Connor MD, Griffith M, Kuchenbauer F, Delaney A, Prabhu AL, Zhao Y, McDonald H, Zeng T, Hirst M, Eaves CJ, Marra MA. Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Res.* 2008; 18:610–621. [PubMed: 18285502]
- Mortensen RD, Serra M, Steitz JA, Vasudevan S. Posttranscriptional activation of gene expression in *Xenopus laevis* oocytes by microRNA-protein complexes (microRNPs). *Proc Natl Acad Sci U S A.* 108:8281–8286. [PubMed: 21536868]
- Mortensen RD, Serra M, Steitz JA, Vasudevan S. Posttranscriptional activation of gene expression in *Xenopus laevis* oocytes by microRNA-protein complexes (microRNPs). *Proc Natl Acad Sci U S A.* 2011; 108:8281–8286. [PubMed: 21536868]
- Okamura K, Lai EC. Endogenous small interfering RNAs in animals. *Nat Rev Mol Cell Biol.* 2008; 9:673–678. [PubMed: 18719707]
- Olena AF, Patton JG. Genomic organization of microRNAs. *J Cell Physiol.* 2010; 222:540–545. [PubMed: 20020507]
- Otto SP, Yong P. The evolution of gene duplicates. *Adv Genet.* 2002; 46:451–483. [PubMed: 11931235]
- Sémon M, Wolfe KH. Preferential subfunctionalization of slow-evolving genes after allopolyploidization in *Xenopus laevis*. *Proc Natl Acad Sci U S A.* 2008; 105:8333–8338. [PubMed: 18541921]
- Shah PP, Hutchinson LE, Kakar SS. Emerging role of microRNAs in diagnosis and treatment of various diseases including ovarian cancer. *J Ovarian Res.* 2009; 2:11. [PubMed: 19712461]
- Szittyá G, Moxon S, Santos DM, Jing R, Fevereiro MP, Moulton V, Dalmay T. High-throughput sequencing of *Medicago truncatula* short RNAs identifies eight new miRNA families. *BMC Genomics.* 2008; 9:593. [PubMed: 19068109]
- Tang GQ, Maxwell ES. *Xenopus* microRNA genes are predominantly located within introns and are differentially expressed in adult frog tissues via post-transcriptional regulation. *Genome Res.* 2008; 18:104–112. [PubMed: 18032731]
- Tay YM, Tam WL, Ang YS, Gaughwin PM, Yang H, Wang W, Liu R, George J, Ng HH, Perera RJ, Lufkin T, Rigoutsos I, Thomson AM, Lim B. MicroRNA-134 modulates the differentiation of mouse embryonic stem cells, where it causes post-transcriptional attenuation of Nanog and LRH1. *Stem Cells.* 2008; 26:17–29. [PubMed: 17916804]
- Valencia-Sanchez MA, Liu J, Hannon GJ, Parker R. Control of translation and mRNA degradation by miRNAs and siRNAs. *Genes Dev.* 2006; 20:515–524. [PubMed: 16510870]
- Walker JC, Harland RM. Expression of microRNAs during embryonic development of *Xenopus tropicalis*. *Gene Expr Patterns.* 2008; 8:452–456. [PubMed: 18440877]
- Watanabe T, Takeda A, Mise K, Okuno T, Suzuki T, Minami N, Imai H. Stage-specific expression of microRNAs during *Xenopus* development. *FEBS Lett.* 2005; 579:318–324. [PubMed: 15642338]
- Weigel MT, Dowsett M. Current and emerging biomarkers in breast cancer: prognosis and prediction. *Endocr Relat Cancer.* 2010; 17:R245–R262. [PubMed: 20647302]
- Wittmann J, Jäck HM. Serum microRNAs as powerful cancer biomarkers. *Biochim Biophys Acta.* 2010; 1806:200–207. [PubMed: 20637263]
- Yanai I, Peshkin L, Jorgensen P, Kirschner MW. Mapping gene expression in two *Xenopus* species: evolutionary constraints and developmental flexibility. *Dev Cell.* 2011; 20:483–496. [PubMed: 21497761]
- Yoo AS, Staahl BT, Chen L, Crabtree GR. MicroRNA-mediated switching of chromatin-remodelling complexes in neural development. *Nature.* 2009; 460:642–646. [PubMed: 19561591]
- Yoo AS, Sun AX, Li L, Shcheglovitov A, Portmann T, Li Y, Lee-Messer C, Dolmetsch RE, Tsien RW, Crabtree GR. MicroRNA-mediated conversion of human fibroblasts to neurons. *Nature.* 2011; 476:228–231. [PubMed: 21753754]
- Zen K, Zhang CY. Circulating MicroRNAs: a novel class of biomarkers to diagnose and monitor human cancers. *Med Res Rev.* 2010

¹Xenopus – (Specific species)

²Vertebrata – (Selected species)

miR – mature miRNA

mir – pre-miRNA

Figure 1.
Data analysis flowchart

Groups	Criteria	Maps to	Reads	Ref
Group 1	gp1a: Reads mapped to <i>X. tropicalis</i> miRNAs/pre-miRNAs in miRBase and the pre-miRNAs further mapped to <i>Xenopus</i> genome & EST.		94	Tables 2 and 3
	gp1b: Reads mapped to vertebrate miRNAs/pre-miRNAs (non- <i>Xenopus</i>) in miRBase and the pre-miRNAs further mapped to <i>Xenopus</i> genome & EST.		19	Table 4
	gp1c: Reads in gp1a or gp1b that mapped to multiple locations on <i>Xenopus</i> genome (included in the above reads).
Group 2	Reads mapped to vertebrate miRNAs/pre-miRNAs in miRBase. Pre-miRNAs did not further map to genome, but the reads mapped to <i>Xenopus</i> genome. The extended genome sequences from the genome loci form stable hairpins.		2	Table 4
Group 3	Reads mapped to vertebrate miRNAs/pre-miRNAs in miRBase. Pre-miRNAs did not further map to <i>Xenopus</i> genome.		92	Data not presented
Group 4	The unique reads do not map to miRBase of any species. The unique reads map to <i>X. tropicalis</i> genome. Sequences flanking the miRNA at mapped genomic positions have a propensity to form thermodynamically stable hairpins.		72	Table 5

Figure 2.Criteria used to group miRNAs identified in *X. laevis* metaphase II arrested eggs

Figure 3.
Length distribution of sequencing reads between 15 and 25 nucleotides

Table 1

Criteria used for miRNA annotation and hairpin structure determination

miR annotation presented in Supplementary Tables

miRNA_name is the name of detected miRNA sequence.

The miR_name is composed of the 1st known miR name in a cluster, an underscore, and a matching annotation, such as:

L-n means the miRNA_seq (detected) is n bases less than known rep_miRSeq in the left side

R-n means the miRNA_seq (detected) is n bases less than known rep_miRSeq in the right side

L+n means the miRNA_seq (detected) is n bases more than known rep_miRSeq in the left side

R+n means the miRNA_seq (detected) is n bases more than known rep_miRSeq in the right side

2ss5TC13TA means 2 “sequence substitutions” (ss), which are T>C at position 5 and T>A at position 13 of the representative miRNA.

Hairpin determination in Supplementary Tables

Definition of MFEI: $MFEI = -dG * 100 / \text{mirLen} / \text{CG\%}$. Reference: Cell. Mol. Life Sci. 63 (2006) 246–254.

Definition of #base_in Loop: This is the maximum number of bases appearing in hairpin loop region. This number is only for gp1c and gp2.

Criteria:

- 1 number of allowed errors in one bulge in stem: ≤ 12
- 2 number of basepairs (bp) in stem region: ≥ 16
- 3 free energy (dG in kCal/mol): ≤ -15
- 4 length of hairpin (up and down stem + terminal loop): ≥ 50
- 5 length of terminal loop: ≤ 20
- 6 number of allowed errors in one bulge in mature region: ≤ 8
- 7 number of allowed biased errors in one bulge in mature region: ≤ 4
- 8 number of allowed biased bulges in mature region: ≤ 2
- 9 number of basepairs (bp) in mature or mature* region: ≥ 12
- 10 percentage of small RNA in stem region (pm): $\geq 80\%$
- 11 number of allowed errors in mature region: ≤ 7

Qualified hairpin is of flag = 'y' !

The hairpins of flag='n' should belong to gp1a, gp1b & gp3 (these hairpins are known)!

If the flag='n' for hairpins in gp1c, gp2 & gp4 (predicted hairpins), the predicted hairpins may be removed.

Table 2

Nine known *X. laevis* identified in mature miRNAs identified in *X. laevis* metaphase II arrested eggs (Group 1a).

#	Reference miR/ pre-miR_name	miR_seq	Accession #
1	xla-miR-15c	TAGCAGCACATCATGGTTTGTA	JN795018
2	xla-miR-18	TAAGGTGCATCTAGTGCAGTTA	JN795022
3	xla-miR-19b	TGTGCAAATCCATGCAAACTGA	JN795025
4	xla-miR-20	CAAAGTGCTCATAGTGCAGGTAG	JN795026
5	xla-miR-23a	ATCACATTGCCAGGGATTCCA	JN795031
6	xla-miR-92a	TATTGCACTTGTCCTGGCCTGT	JN795051
7	xla-miR-205	TCCTTCATTCCACCGAGTCTGT	JN795102
8	xla-miR-363	CGGGTGGATCACGATGCAATTT	JN795115
9	xla-miR-427	AAAGTGCTTTCTGTTTTGGGCGT	JN795119

Table 3

Eighty-five miRNAs identified in *X. laevis* metaphase II arrested eggs that correspond to known *X. tropicalis* mature miRNAs/pre-miRNAs (Group 1a).

#	Reference miR/ pre-miR_name	miR_seq	New name	Accession #
1	xtr-let-7a	TGAGGTAGTAGTTGTATAGTT	xla-let-7a	JN795004
2	xtr-let-7b	TGAGGTAGTAGTTTGTGTAGTT	xla-let-7b	JN795006
3	xtr-let-7c	TGAGGTAGTAGTTGTATGGTT	xla-let-7c	JN795007
4	xtr-let-7e	TGAGGTAGTAGTTGTTTAGTT	xla-let-7e	JN795008
5	xtr-let-7f	TGAGGTAGTAGATTGTATAGTT	xla-let-7f	JN795009
6	xtr-let-7g	TGAGGTAGTAGTTTGTACAGTT	xla-let-7g	JN795010
7	xtr-let-7i	TGAGGTAGTAGTTTGTGCTGTT	xla-let-7i	JN795011
8	xtr-miR-1a	TGGAATGTAAAGAAGTATGTAT	xla-miR-1a	JN795012
9	xtr-miR-7	TGGAAGACTAGTGATTTTGTGTT	xla-miR-7	JN795013
10	xtr-miR-10a	TACCCTGTAGATCCGAATTTGT	xla-miR-10a	JN795014
11	xtr-miR-10b	TACCCTGTAGAACCGAATTTGT	xla-miR-10b	JN795015
12	xtr-miR-15a	TAGCAGCACATAATGGTTTGTGA	xla-miR-15a	JN795016
13	xtr-miR-15b	TAGCAGCACATCATGATTGCA	xla-miR-15b	JN795017
14	xtr-miR-16b	TAGCAGCACGTAAATATTGGGT	xla-miR-16b	JN795019
15	xtr-miR-16c	TAGCAGCACGTAAATACTGGAG	xla-miR-16c	JN795020
16	xtr-miR-17-5p	CAAAGTGCTTACAGTGCAGGTAG	xla-miR-17-5p	JN795021
17	xtr-miR-18a	TAAGGTGCATCTAGTGCAGATA	xla-miR-18a	JN795023
18	xtr-miR-20a	TAAAGTGCTTATAGTGCAGGTAG	xla-miR-20a	JN795027
19	xtr-miR-22	AAGCTGCCAGTTGAAGAACTGT	xla-miR-22	JN795029
20	xtr-miR-22*	AGTTCTTCAGTGGCAAGCTTT	xla-miR-22*	JN795030
21	xtr-miR-23b	ATCACATTGCCAGGGATT	xla-miR-23b	JN795032
22	xtr-miR-24a	TGGCTCAGTTCAGCAGGAACAG	xla-miR-24a	JN795034
23	xtr-miR-25	CATTGCACTTGTCTCGGTCTGA	xla-miR-25	JN795035
24	xtr-miR-26	TTCAAGTAATCCAGGATAGGCT	xla-miR-26	JN795036
25	xtr-miR-27a	TTACAGTGGCTAAGTTCCG	xla-miR-27a	JN795038
26	xtr-miR-27b	TTACAGTGGCTAAGTTCTGC	xla-miR-27b	JN795039
27	xtr-miR-27c	TTACAGTGGCTAAGTTC	xla-miR-27c	JN795040
28	xtr-miR-29c	TAGCACCATTTGAAATCGGTTA	xla-miR-29c	JN795041
29	xtr-miR-30a-3p	CTTTCAGTCAGATGTTTGCAGC	xla-miR-30a-3p	JN795042
30	xtr-miR-30a-5p	TGTAAACATCCTCGACTGGAAGC	xla-miR-30a-5p	JN795043
31	xtr-miR-30b	TGTAAACATCCTACACTCAGCT	xla-miR-30b	JN795044
32	xtr-miR-30c	TGTAAACATCCTACACTCTCAGCT	xla-miR-30c	JN795045
33	xtr-miR-30c-1	GTGAACATAAGGTGGCTGGGAGAA	xla-miR-30c-1-3p	JN795046
34	xtr-miR-30d	TGTAAACATCCCCGACTGGAAGCT	xla-miR-30d	JN795047
35	xtr-miR-30e	TGTAAACATCCTTGACTGGAAGCT	xla-miR-30e	JN795048

#	Reference miR/ pre-miR_name	miR_seq	New name	Accession #
36	xtr-miR-34a	TGGCAGTGTCTTAGCTGGTTGT	xla-miR-34a	JN795049
37	xtr-miR-34b	AGGCAGTGTAGTTAGCTGATTG	xla-miR-34b	JN795050
38	xtr-miR-92b	TATTGCACTCGTCCCGCCTCC	xla-miR-92b	JN795053
39	xtr-miR-93a	CAAAGTGCTGTTCGTGCAGGTAG	xla-miR-93a	JN795054
40	xtr-miR-99	AACCCGTAGATCCGATCTTGTG	xla-miR-99	JN795057
41	xtr-miR-100	AACCCGTAGATCCGAACCTTGTG	xla-miR-100	JN795058
42	xtr-miR-101a	TACAGTACTGTGATAACTGAA	xla-miR-101a	JN795059
43	xtr-miR-106	AAAAGTGCTTATAGTGCAGGTAG	xla-miR-106	JN795060
44	xtr-miR-107	AGCAGCATTGTACAGGGCTAT	xla-miR-107	JN795061
45	xtr-miR-122	TGGAGTGTGACAAATGGTGT	xla-miR-122	JN795062
46	xtr-miR-124	TAAGGCACGCGTGAATGCCAA	xla-miR-124	JN795064
47	xtr-miR-125a	TCCCTGAGACCCCTAACCTGTGA	xla-miR-125a	JN795065
48	xtr-miR-125b	TCCCTGAGACCCCTAACCTGTGA	xla-miR-125b	JN795066
49	xtr-miR-128	TCACAGTGAACCGGTCTCTTT	xla-miR-128	JN795067
50	xtr-miR-130a	CAGTGCAATGTTAAAAGGGCAT	xla-miR-130a	JN795069
51	xtr-miR-130b	CAGTGCAATGATGAAAGGGCAT	xla-miR-130b	JN795070
52	xtr-miR-130c	CAGTGCAATATTTAAAAGGGCAT	xla-miR-130c	JN795071
53	xtr-miR-132	TAACAGTCTACAGCCATGGTC	xla-miR-132	JN795072
54	xtr-miR-135	TATGGCTTTTATTCCTATGT	xla-miR-135	JN795074
55	xtr-miR-140	CAGTGGTTTTACCCTATGGTA	xla-miR-140	JN795075
56	xtr-miR-143	TGAGATGAAGCACTGTAGCTC	xla-miR-143	JN795077
57	xtr-miR-145	GTCCAGTTTTCCCAGGAATCCCT	xla-miR-145	JN795079
58	xtr-miR-146	TGAGAACTGAATTCC	xla-miR-146	JN795080
59	xtr-miR-146b	TGAGAACTGAATTCCATGGACT	xla-miR-146b	JN795081
60	xtr-miR-148a	TCAGTGCACTACAGAACCTTGT	xla-miR-148a	JN795082
61	xtr-miR-148b	TCAGTGCACTACAGAACCTTGT	xla-miR-148b	JN795084
62	xtr-miR-153	TTGCATAGTCACAAAAGTGATT	xla-miR-153	JN795086
63	xtr-miR-184	TGGACGGAGAAGCTGATAAGGG	xla-miR-184	JN795089
64	xtr-miR-184	TGGACGGAGAAGCTGATAAGG	xla-miR-184-2	JN795090
65	xtr-miR-191	CAACGGAATCCCAAAAGCAGCTGT	xla-miR-191	JN795091
66	xtr-miR-192	ATGACCTATGAATTGACAGCCA	xla-miR-192	JN795092
67	xtr-miR-194-2	CCGGTGGAGATGCTGTTATCTT	xla-miR-194-2-3p	JN795094
68	xtr-miR-199a*	ACAGTAGTCTGCACATTGGTT	xla-miR-199a*	JN795095
69	xtr-miR-200a	TAACACTGTCTGGTAACGATGTT	xla-miR-200a	JN795096
70	xtr-miR-200b	AATACTGCCTGGTAATGATGATT	xla-miR-200b	JN795097
71	xtr-miR-202*	TTCTATGCATATACCTCTTT	xla-miR-202*	JN795098
72	xtr-miR-203	GTGAAATGTTTAGGACCACTTG	xla-miR-203	JN795099
73	xtr-miR-204	TTCCCTTTGTATCCTATGCCT	xla-miR-204	JN795100
74	xtr-miR-206	TGGAATGTAAGGAAGTGTGTGG	xla-miR-206	JN795103

#	Reference miR/ pre-miR_name	miR_seq	New name	Accession #
75	xtr-miR-210	CTGTGCGTGTGACAGCGGCTAA	xla-miR-210	JN795104
76	xtr-miR-215	ATGACCTATGAAATGACAGCCA	xla-miR-215	JN795107
77	xtr-miR-216	TAATCTCAGCTGGCAACTGTGA	xla-miR-216	JN795108
78	xtr-miR-217	ATACTGCATCAGGAACTGATTG	xla-miR-217	JN795109
79	xtr-miR-221	AGCTACATTGTCTGCTGGGTTT	xla-miR-221	JN795110
80	xtr-miR-222	AGCTACATCTGGCTACTGGGTCTC	xla-miR-222	JN795111
81	xtr-miR-338	TCCAGCATCAGTGATTTTGTG	xla-miR-338	JN795113
82	xtr-miR-363-3p	AATTGCACGGTATCCATCTGTAA	xla-miR-363-3p	JN795116
83	xtr-miR-375	TTTGTTTCGTTCCGGCTCGCGTTA	xla-miR-375	JN795118
84	xtr-miR-455	TATGTGCCCTTGGACTACATCG	xla-miR-455	JN795120
85	xtr-miR-499	TTAAGACTTGCACTGATGTTTA	xla-miR-499	JN795124

Table 4

Twenty-one miRNAs identified in *X. laevis* metaphase II arrested eggs that correspond to known vertebrate (non-*Xenopus*) mature miRNAs/pre-miRNAs and map to *X. tropicalis* genome (Groups 1b and Group 2).

#	Reference miR/ pre-miR_name	miR_seq	New name	Accession #
	Group 1b			
1	cfa-miR-212	ACCTTGGCTCTAGACTGCTTACT	xla-miR-212-5p	JN795106
2	dre-miR-210*	AAGCCACTGACTAACGCACATT	xla-miR-210*	JN795105
3	hsa-let-7a*	ATACAATCTACTGTCTTTCTT	xla-let-7a*	JN795005
4	hsa-miR-20b*	ACTGTAATATGGGCACTTACA	xla-miR-20b*	JN795028
5	hsa-miR-26a-2*	CCTATTCTTGATTACTTGTTC	xla-miR-26a-2*	JN795037
6	hsa-miR-92a-1*	AGGTTGGGATTGTTGCAATGCT	xla-miR-92a-1*	JN795052
7	hsa-mir-98-p5	TGAGGTTAGTAAGTTGTATTGTT	xla-miR-98-5p	JN795056
8	hsa-miR-129-3p	AAGCCCTTACCCCAAAAAGCA	xla-miR-129-3p	JN795068
9	hsa-miR-132*	CCGTGGCTTTAGATTGTTACT	xla-miR-132*	JN795073
10	hsa-miR-140	TACCACAGGGTAGAACCACGGA	xla-miR-140-3p	JN795076
11	hsa-miR-143*	GGTGCACTGCTGCATCTCTGG	xla-miR-143*	JN795078
12	hsa-miR-148b*	GAAGTTCTGTTATACACTCCGGCT	xla-miR-148b*	JN795085
13	gga-miR-456	CAGGCTGGTTAGATGGTTGTCA	xla-miR-456	JN795121
14	gga-miR-458	ATAGCTCTTTGAATGGTACTGC	xla-miR-458	JN795122
15	gga-miR-460b-5p	TCCTCATTGTACATGCTGTGTG	xla-miR-460b-5p	JN795123
16	gga-miR-1805-5p	GAGTTGTAGTCTTTCAAACAG	xla-miR-1805-5p	JN795126
17	mmu-miR-204*	GCTGGGAAGGCAAAGGGACGT	xla-miR-204*	JN795101
18	mmu-mir-1983	GCTCCAGTGGCGCAA	xla-mir-1983-5p	JN795127
19	ssc-mir-1285	GTAGTGGGATCGCGC	xla-mir-1285-5p	JN795125
	Group 2			
20	dre-miR-24	TGGCTCAGTTCAGCAGGAACAGGT	xla-miR-24	JN795033
21	mmu-mir-5102	GGGAGTTTGACTGGGGCG	xla-mir-5102-5p	JN795128

Table 5

Seventy-two putative candidate miRNAs identified in *X. laevis* metaphase II arrested eggs that map to *X. tropicalis* genome (Group 4).

#	Read name	miR_seq	New name	Accession #
1	PC-3p-82353	TGGAATGTTAAGAAGTATGTAA	xla-miR-5137	
2	PC-3p-473580	TGCTCTGAGAGACCTTCCTAACA	xla-miR-5139	
3	PC-3p-559940	TGAGAGAAGAGTTGAGTAG	xla-miR-5140	
4	PC-3p-572193	AAAGTGCTTCTCGTTCGGCTGA	xla-miR-5141	
5	PC-3p-590570	AGGGATGGAATGGAAAGGAATGC	xla-miR-5142	
6	PC-3p-722125	AACTAGTACATGCAGCAC	xla-miR-5143	
7	PC-3p-777332	AGCATGATGGGAGTTGTA	xla-miR-5144	
8	PC-3p-912538	TGTATAGCAAAATAG	xla-miR-5145	
9	PC-3p-997070	ATGGGGAATGGTTCTGATTGGC	xla-miR-5146	
10	PC-3p-1026732	TGCTACAGGAGTCCTGTGTCT	xla-miR-5147	
11	PC-3p-1217232	TGCAGTCTGAGGGATGGAGTGGA	xla-miR-5148	
12	PC-3p-1416830	TCTACAGTCCGAGAATC	xla-miR-5149	
13	PC-3p-1429392	TTTCAGGAAAATGACACAAA	xla-miR-5150	
14	PC-3p-1500718	ATTTTGTGGTTTTTCG	xla-miR-5151	
15	PC-3p-1558848	CTGTGATTTTGACTGCAGA	xla-miR-5152	
16	PC-3p-1570955	TAGAAGATCATTGGTTA	xla-miR-5153	
17	PC-3p-1825226	TGAAGAACTGTGTTA	xla-miR-5154	
18	PC-3p-2128722	AGTAGTTTAGAGATGAAAGGAAGC	xla-miR-5155	
19	PC-3p-2265908	TGTATGGCATTGTGGAAA	xla-miR-5156	
20	PC-3p-2419584	TATTTGTCTATAATGTCCT	xla-miR-5157	
21	PC-3p-2500255	TGCTGCTGTGCTAAGTCTGGT	xla-miR-5158	
22	PC-3p-2526314	TGAGGGATGGAATGGAAAGGAAT	xla-miR-5159	
23	PC-3p-2608329	TGACAAGGGATGCTGGGAAA	xla-miR-5160	
24	PC-3p-2838765	TTTTCTTCTAATCGGAATGGGTG	xla-miR-5161	
25	PC-3p-2858122	TGCATTCTTCTCACTG	xla-miR-5162	
26	PC-3p-2866177	CATTAGAGGACATTATAGACA	xla-miR-5163	
27	PC-3p-3225349	TGGATCAATGTGAGA	xla-miR-5164	
28	PC-3p-3396719	TGTAGACACAGAGGGATGATTC	xla-miR-5165	
29	PC-3p-3404189	CACGTAGTCTGAACACTGGGGG	xla-miR-5166	
30	PC-3p-3462114	TCTGTAGGAAAAAGGTAT	xla-miR-5167	
31	PC-3p-3927070	GCTGTACTGCGAACT	xla-miR-5168	
32	PC-5p-183320	CTGGTCAATAGCAGCTGAGCCATG	xla-miR-5169	
33	PC-5p-184903	TAAGTAATGAAGAAGAATATG	xla-miR-5170	
34	PC-5p-377209	TGCACACAGGTCTGGGGAA	xla-miR-5172	
35	PC-5p-429687	CCAGTTGGGTGATAAATGA	xla-miR-5173	
36	PC-5p-508025	AGAGTAGAGATTGTGCAGGGCA	xla-miR-5174	

#	Read name	miR_seq	New name	Accession #
37	PC-5p-524204	TTGAGATGTAGGTTATAAGATT	xla-miR-5175	
38	PC-5p-605340	TGATCGGATCTTGTACAATTGA	xla-miR-5176	
39	PC-5p-661517	GAAAGAAGACCCTGT	xla-miR-5177	
40	PC-5p-661540	AATGAAGAGATAGCATTGCTGGTT	xla-miR-5178	
41	PC-5p-705780	GACAGCAGGACGGTGG	xla-miR-5179	
42	PC-5p-895083	AATTCAAGGAAAAATGTA	xla-miR-5180	
43	PC-5p-904518	ACCATTGGATTGTGGG	xla-miR-5181	
44	PC-5p-1057474	TATGACAGGCTCTTTCAT	xla-miR-5182	
45	PC-5p-1096123	ACCAGTGTGTAGACTACTGTTC	xla-miR-5183	
46	PC-5p-1107030	AGGGATAGGAACTAGTCGTTCA	xla-miR-5184	
47	PC-5p-1212794	TTAACTGTAACAGAAGTC	xla-miR-5185	
48	PC-5p-1258591	TGCACACAGGTCTGGGGA	xla-miR-5186	
49	PC-5p-1564208	CGTAGTGGAAGGAGGCGTGCC	xla-miR-5187	
50	PC-5p-1623845	GTGGAAAACAGACTGTG	xla-miR-5188	
51	PC-5p-1703422	AGTTAAATGCGTCATGACTCA	xla-miR-5189	
52	PC-5p-1750291	TTCATGTAGAATGTTGGTA	xla-miR-5190	
53	PC-5p-1824198	AGGAAGTGGCTAATTC	xla-miR-5191	
54	PC-5p-2103624	TAGTTCTACCGTCTTTCGACC	xla-miR-5193	
55	PC-5p-2169326	AAGCTTCTGAATGTTCCAGTGA	xla-miR-5194	
56	PC-5p-2291026	TCTGTAATACCCACTGAT	xla-miR-5195	
57	PC-5p-2416157	TGTTTTGTGTGTGGATG	xla-miR-5196	
58	PC-5p-2567058	TTATCTATGTTTTTG	xla-miR-5197	
59	PC-5p-2589352	TGTAAGATCTTGAGGGTAT	xla-miR-5198	
60	PC-5p-2703994	GGGAAGTAGAAGTGGGTGG	xla-miR-5199	
61	PC-5p-2773462	TGAGTTAAATGCGTCATGACT	xla-miR-5200	
62	PC-5p-3009997	AGCAGGATCAGATGAACAGACAG	xla-miR-5201	
63	PC-5p-3118828	TAGAGAGGAAGGTGGGGA	xla-miR-5202	
64	PC-5p-3179154	TTAGTTTGTGGATGATA	xla-miR-5203	
65	PC-5p-3372430	TTTAGGCTGAGGGCACAC	xla-miR-5204	
66	PC-5p-3455914	ATCGCGTTAACCGGAAGTCTCA	xla-miR-5205	
67	PC-5p-3532296	TTGGGGTCGTTGTCT	xla-miR-5206	
68	PC-5p-3615992	TGCTGCCAGATCTGATACCGT	xla-miR-5207	
69	PC-5p-3630407	TTTACACTAAACTTGAA	xla-miR-5208	
70	PC-5p-3648555	TATTTAGACGGAACCTA	xla-miR-5209	
71	PC-5p-3770263	CCTGTGGGAATACTGCCAGC	xla-miR-5210	
72	PC-5p-3828717	GCGGCTAGACGAGGCG	xla-miR-5211	

Table 6Expression of miRNAs in *X. laevis* metaphase II arrested eggs localizing to gene clusters

#	miRNA gene cluster	Length (nt)
1	xtr-let-7a/xtr-mir100/miR-125a	14946
2	xtr-let-7a/ xtr-let-7a*/xtr-let-7e	142
3	xtr-mir-15b/xtr-mir-16b	245
4	xtr-mir-15c/xtr-mir-16c	2088
5	xtr-mir-106/xla-mir-18/xla-mir-19b/xla-mir-20/ hsa-mir-20b*/xla-mir-92a/xla-mir-363/xtr-mir-363-3p	724
6	xtr-mir-17-5p/xtr-mir-18a/xtr-mir-19b/xtr-mir-20a/ xtr-mir-92a	729
7	xtr-mir-23b/dre-mir-24/ xtr-mir-24a/xtr-mir-27b	911
8	xtr-mir-23a/xtr-mir-27a	684
9	xtr-mir-25/xtr-mir-93a	206
10	xtr-mir-30a-3p/xtr-mir-30a-5p/xtr-mir-30c	7491
11	xtr-mir-30b/xtr-mir-30d	4325
12	xtr-mir-30c/xtr-mir-30c-1-p3/xtr-mir-30e	1093
13	xtr-mir-34b/xtr-mir-34b	7678
14	hsa-mir-98-p5/xtr-let-7f	286
15	xtr-mir-99/xtr-let-7c	779
16	xtr-mir-130b/xtr-mir-130c	7810
17	xtr-mir-132/hsa-mir-132*/cfa-mir-212	5818
18	xtr-mir-143/hsa-mir-143*/xtr-mir-145	1048
19	xtr-mir-192/xtr-mir-194-2-p3	1664
20	xtr-mir-200a/xtr-mir-200b	1765
21	xtr-mir-202*/xtr-mir-202*	7814
22	xtr-mir-216/xtr-mir-217	460
23	xtr-mir-221/xtr-mir-222	602
24	xla-mir-427/xla-mir-427	493
#	Putative candidate miRNA clusters	Length (nt)
25	xtr-let-7e/PC-5p-183320	7722
26	PC-3p-572193/ PC-3p-572193	380
27	PC-5p-1107030/ PC-5p-1107030	2188
28	PC-3p-590570/PC-3p-2526314	2331
29	PC-3p-2589352/ PC-3p-2589352	11972