

*Full Length Research Paper*

# Computational analysis of human miRNAs phylogenetics

Raisa Bano and Hamid Rashid\*

Department of Bioinformatics, Mohammad Ali Jinnah University, Islamabad, Pakistan.

Accepted 18 February, 2011

Over one third of the human genome appears to be regulated by miRNAs which are involved in fundamental cellular processes and are negative regulators of gene expression. MicroRNAs (miRNAs) are evolutionary conserved across broad phylogenetic distances and have gained considerable attention about evolution, genetic and phylogenetic analysis. Comparing sequences of miRNA precursors within a species and between closely related species should thus help to determine patterns of molecular evolution and the timescales at which different aspects of the evolution of miRNAs can be best investigated. In this study, miRNAs were analyzed for their relationship with each other and pattern of variations among different organisms, and it is hoped that it will enhance our understandings on the use of miRNAs in therapeutics for the treatment of various diseases which are the main focus of modern research these days in molecular studies. Four different sequences of human miRNAs (*Homo sapiens* miR-99-precursor-21, *H. sapiens* microRNA let-7a-1, *H. sapiens* microRNA 145 and *H. sapiens* microRNA 18a) were retrieved from the NCBI, and multiple sequence alignment and phylogenetic analysis of the miRNAs through UPGMA was performed which showed their relationship among each other and pattern of variations among different organisms.

**Key words:** MicroRNA (miRNA), phylogenetics, miR-99, microRNA let-7a-1, microRNA 145, microRNA 18a.

## INTRODUCTION

MicroRNAs (miRNAs) are noncoding RNAs of approximately 22 nucleotides that function as post-transcriptional regulators. By base-pairing with the complementary sites in the 3' untranslated region (UTR) of mRNA, miRNAs can control the mRNA stability and the efficiency of translation (Filipowicz, 2008). These tiny molecules are important regulators in animals and plants for gene expression by translational repression, mRNA cleavage and decay (Zhang et al., 2007). More than 400 miRNAs have been identified in humans and are evolutionally conserved from plants to animals, regulating various biological processes including development, cell differentiation, cell proliferation and cell death (Yukihiro et al., 2007). Computational analysis predicts that mammalian miRNAs regulate approximately 30% of all

protein-coding genes (Lewis et al., 2005; Lewis et al. 2003). Their expression profiles can be used for the classification, diagnosis and prognosis of human malignancies (George and Carlo, 2006) acting either as tumor suppressors or oncogenes (Kristen and Glen, 2008).

miRNAs are evolutionary conserved across broad phylogenetic distances (Lagos-Quintana et al., 2001; Lee and Ambros, 2001). They have gained considerable attention in evolution, genetic and phylogenetic analysis (Liu et al., 2008; Guo et al., 2009). Some miRNAs in a single animal species are similar in sequence and produce the same or similar mature miRNA sequences, and these miRNAs always compose miRNA gene family. These family members may be derived from ancestral miRNA gene directly or indirectly through duplication, but the duplication process maybe complex and unclear based on limited miRNA data across animal species. Nonetheless, miRNA gene evolution might provide potential implication for selection of miRNA and fate of miRNA (Guo and Lu, 2010). Comparing sequences of

\*Corresponding author. E-mail: drhamid@jinnah.edu.pk. Tel: +92-111878787. Ext: 135.

miRNA precursors within a species and between closely related species should thus help to determine patterns of molecular evolution and the timescales at which different aspects of the evolution of miRNA precursors are best investigated (Warthmann et al., 2008).

Molecular studies have considerably enhanced our understanding of the use of miRNAs in therapeutics for cancer treatments and the evolution of miRNAs in human genome. Prediction of relationship between evolutionary pattern and distribution pattern of miRNAs in human genome is the main focus of this study.

## MATERIALS AND METHODS

### Sequence retrieval

Four different sequences of human miRNA: *Homo sapiens* miR-99-precursor-21 micro RNA, complete sequence, *H. sapiens* microRNA let-7a-1 (MIRLET7A1), *H. sapiens* microRNA 145 (MIR145), and *H. sapiens* microRNA 18a (MIR18A) were retrieved from the NCBI (<http://www.ncbi.nlm.nih.gov>) in FASTA format.

### Sequence alignment

BLAST (Altschul et al., 1990) was performed for each of the four human miRNA sequences retrieved from NCBI using the online Geneious 4.8.3 software (<http://www.geneious.com>). Sequences were studied for their similarity patterns by feeding the FASTA format of sequence into the software.

### Phylogenetic analysis

Phylogenetic analysis of the miRNA sequences through UPGMA was obtained using Geneious software. A tree was constructed by the software showing the ancestral relationship among the sequences.

The tree gives different clusters, showing their relationship with each other. The sequences which lie in the same cluster are closely related with those lying in distant clusters.

## RESULTS AND DISCUSSION

### Sequence retrieval

*H. sapiens* miR-99-precursor-21 micro RNA, *H. sapiens* microRNA let-7a-1 (MIRLET7A1), *H. sapiens* microRNA 145 (MIR145) and *H. sapiens* microRNA 18a (MIR18A) were retrieved from the NCBI in FASTA format.

### Sequence-1

>gi|20805650|gb|AF480571.1| *Homo sapiens* miR-99-precursor-21 micro RNA, complete sequence:

```
CCCATGGCATAAACCCGTAGATCCGATCTTGTGGT
GAAGTGGACCGCACAAAGCTCGCTTCTATGGGTCTGT
GTCAGTGTG.
```

### Sequence-2

>gi|262205583|ref|NR\_029476.1| *Homo sapiens* microRNA let-7a-1 (MIRLET7A1), microRNA:

```
TGGGATGAGGTAGTAGGTTGTATAGTTTTAGGGTCA
CACCCACCACTGGGAGATAACTATACAATCTACTGTC
TTTCCTA.
```

### Sequence-3

>gi|262205329|ref|NR\_029686.1| *Homo sapiens* microRNA 145 (MIR145), microRNA:

```
CACCTTGTCTCCTCACGGTCCAGTTTTCCAGGAATCC
CTTAGATGCTAAGATGGGGATTCCTGGAAATACTGTT
CTTGAGGTCATGGTT.
```

### Sequence-4

>gi|262205635|ref|NR\_029488.1| *Homo sapiens* microRNA 18a (MIR18A), microRNA:

```
TGTTCTAAGGTGCATCTAGTGCAGATAGTGAAGTAGA
TTAGCATCTACTGCCCTAAGTGCTCCTTCTGGCA,
```

### Sequence alignment

Each of the four human miRNA sequences was studied for their similarity patterns and BLAST was therefore performed by feeding the FASTA format of sequence into the online Geneious 4.8.3 software. After performing BLAST for each of four sequences, the software produced BLAST tables showing their accession numbers, percent similarity, e-value, etc (Tables 1 to 4). The sequences having lowest e-value are more closely related, while the difference in e-value shows the dissimilarity among them.

*H. sapiens* 18a microRNA (MIR18A) shows different similarity patterns with different organisms. It showed 100% pair wise similarities with most of the organisms having e-value 1.94E-28, whereas, few are found to be distant from miRNA 18a which include *Gallus gallus* (mir-18a (MIR18A)), *Mus musculus* (microRNA 18 (Mir18)), Mouse DNA (sequence from clone RP24-298F18 on chromosome 14, complete sequence) with e-value 9.03E-27 and *Tetraodon nigroviridis* (full-length cDNA) with e-value 1.51E-24 showing 98.60 and 98.50% similarities. Also *Equus caballus* (mir-18a (MIR18A)) with e-value 7.03E-23, *H. sapiens* (piRNA piR-37525) with e-value 1.20E-05 and *H. sapiens* (piR-37524) with e-value 4.32E-05 shows 100% pairwise identity (Table 1) (Guo and Lu, 2010; Wang et al., 2005). *H. sapiens* microRNA let-7a-1 (MIRLET7A1) on the basis of wise percent pair

**Table 1.** *H. sapiens* microRNA 18a (MIR18A).

E-value	Name	Description	Organism	Sequence length	Pairwise identity (%)
1.94E-28	NR_032077	<i>Macaca mulatta</i> microRNA mir-18 (MIR18), microRNA	<i>Macaca mulatta</i>	71	100.00
1.94E-28	NR_032071	Pan troglodytes microRNA mir-18a (MIR18A), microRNA	<i>Pan troglodytes</i>	71	100.00
1.94E-28	NR_032230	Monodelphis domestica microRNA mir-18 (Mir18), microRNA	<i>Monodelphis domestica</i>	71	100.00
1.94E-28	NR_030892	<i>Bos taurus</i> microRNA mir-18a (MIR18A), microRNA	<i>Bos taurus</i>	71	100.00
1.94E-28	NR_031536	<i>Sus scrofa</i> microRNA mir-19a (MIR19A), microRNA	<i>Sus scrofa</i>	71	100.00
1.94E-28	NR_029488	<i>Homo sapiens</i> microRNA 18a (MIR18A), microRNA	<i>Homo sapiens</i>	71	100.00
1.94E-28	NR_027350	<i>Homo sapiens</i> MIR17 host gene (non-protein coding) (MIR17HG), non-coding RNA	<i>Homo sapiens</i>	71	100.00
1.94E-28	BC109082	<i>Homo sapiens</i> cDNA clone MGC:126270 IMAGE:40034315, complete cds	<i>Homo sapiens</i>	71	100.00
1.94E-28	BC109081	<i>Homo sapiens</i> microRNA host gene 1 (non-protein coding), mRNA (cDNA clone IMAGE:40034314)	<i>Homo sapiens</i>	71	100.00
1.94E-28	AY875971	<i>Sus scrofa</i> clone bh_64922 mir17 microRNA cluster, partial sequence	<i>Sus scrofa</i>	71	100.00
1.94E-28	AF480553	<i>Homo sapiens</i> miR-18-precursor-13 micro RNA, complete sequence	<i>Homo sapiens</i>	71	100.00
1.94E-28	AL162375	Human DNA sequence from clone RP11-282D2 on chromosome 13 Contains the 5' end of a novel gene, 5 miR-precursor-13-micro RNAs and a CpG island, complete sequence	<i>Human DNA</i>	71	100.00
1.94E-28	AL138714	Human DNA sequence from clone RP11-121J7 on chromosome 13q32.1-32.3 Contains the 3' end of a novel gene, the 5' end of the GPC5 gene for glypican 5, 5 miR-precursor-13 micro RNAs and a CpG island, complete sequence.	<i>Human DNA</i>	71	100.00

identity, showed that the most distant organisms are *G. gallus* (microRNA let-7a-1 (MIRLET7A-1)) with e-value 2.36E-28 showing 96.30% similarities, *Danio rerio* and

Zebrafish DNA (let7a-1 (mirlet7a-1) and (sequence from clone DKEY-189A9 in linkage group 11) with e-value 8.47E-28 showing 97.40 similarities. Zebrafish DN

Table 1. cont.

	AB176708	<i>Homo sapiens</i> C13orf25 v_2 mRNA, complete cds, miR-91-precursor-13 micro RNA, microRNA miR-91, microRNA miR-17, miR-18-precursor-13 micro RNA, microRNA miR-18, miR-19a-precursor-13 micro RNA, microRNA miR-19a, microRNA miR-20, miR-19b-precursor-13 micro RNA, microRNA miR-19b, miR-92-precursor-13 micro RNA, microRNA miR-92	<i>Homo sapiens</i>	71	100.00
1.94E-28	AY866318	<i>Macaca nemestrina</i> microRNA mir-17, microRNA mir-18, microRNA mir-19a, microRNA mir-20, microRNA mir-19b-1 genes, complete sequence; and microRNA mir-92-1 gene, partial sequence	<i>Macaca nemestrina</i>	71	100.00
1.94E-28	AY866317	<i>Lagothrix lagotricha</i> microRNA mir-17, microRNA mir-18, microRNA mir-19a, microRNA mir-20, microRNA mir-19b-1, and microRNA mir-92-1 genes, complete sequence	<i>Lagothrix lagotricha</i>	71	100.00
1.94E-28	AY866316	<i>Saguinus labiatus</i> microRNA mir-17, microRNA mir-18, microRNA mir-19a, microRNA mir-20, microRNA mir-19b-1, and microRNA mir-92-1 genes, complete sequence	<i>Saguinus labiatus</i>	71	100.00
1.94E-28	AY866315	<i>Macaca mulatta</i> microRNA mir-17, microRNA mir-18, microRNA mir-19a, microRNA mir-20, microRNA mir-19b-1, and microRNA mir-92-1 genes, complete sequence	<i>Macaca mulatta</i>	71	100.00
1.94E-28	AY866314	<i>Pan troglodytes</i> microRNA mir-17, microRNA mir-18, microRNA mir-19a, microRNA mir-20, microRNA mir-19b-1, and microRNA mir-92-1 genes, complete sequence	<i>Pan troglodytes</i>	71	100.00
1.94E-28	AY866313	<i>Pongo pygmaeus</i> microRNA mir-17, microRNA mir-18, microRNA mir-19a, microRNA mir-20, microRNA mir-19b-1, and microRNA mir-92-1 genes, complete sequence	<i>Pongo pygmaeus</i>	71	100.00
1.94E-28	AY866312	<i>Pan paniscus</i> microRNA mir-17, microRNA mir-18, microRNA mir-19a, microRNA mir-20, microRNA mir-19b-1, and microRNA mir-92-1 genes, complete sequence	<i>Pan paniscus</i>	71	100.00
1.94E-28	AY866311	<i>Ateles geoffroyi</i> microRNA mir-17, microRNA mir-18, microRNA mir-19a, microRNA mir-20, microRNA mir-19b-1, and microRNA mir-92-1 genes, complete sequence	<i>Ateles geoffroyi</i>	71	100.00

Table 1. cont

1.94E-28	AY866310	Lemur catta microRNA mir-17, microRNA mir-18, microRNA mir-19a, microRNA mir-20, microRNA mir-19b-1, and microRNA mir-92-1 genes, complete sequence	<i>Lemur catta</i>	71	100.00
1.94E-28	AY866309	Gorilla gorilla microRNA mir-17, microRNA mir-18, microRNA mir-19a, microRNA mir-20, microRNA mir-19b-1, and microRNA mir-92-1 genes, complete sequence	<i>Gorilla gorilla</i>	71	100.00
9.03E-27	NR_031407	<i>Gallus gallus</i> microRNA mir-18a (MIR18A), microRNA	<i>Gallus gallus</i>	71	98.60
9.03E-27	NR_029736	<i>Mus musculus</i> microRNA 18 (Mir18), microRNA	<i>Mus musculus</i>	71	98.60
9.03E-27	AC163295	<i>Mus musculus</i> BAC clone RP23-132K20 from chromosome 14, complete sequence	<i>Mus musculus</i>	71	98.60
9.03E-27	CT033786	Mouse DNA sequence from clone RP24-298F18 on chromosome 14, complete sequence	<i>Mouse DNA</i>	71	98.60
1.51E-24	CR693207	Tetraodon nigroviridis full-length cDNA	<i>Tetraodon nigroviridis</i>	67	98.50
7.03E-23	NR_032960	<i>Equus caballus</i> microRNA mir-18a (MIR18A), microRNA	<i>Equus caballus</i>	61	100.00
1.20E-05	DQ599459	<i>Homo sapiens</i> piRNA piR-37525, complete sequence	<i>Homo sapiens</i>	30	100.00
4.32E-05	DQ599458	<i>Homo sapiens</i> piRNA piR-37524, complete sequence	<i>Homo sapiens</i>	29	100.00

(sequence from clone CH211-218P2 in linkage group 25) having e-value 3.09E-17 showed 88.80%, *G. gallus* (let-7a-3 (MIRLET7A-3) with e-value 1.11E-16 showed 88.50% similarities. Human DNA (sequence from clone RP4-695O20 on chromosome 22q13.1-13.33), *H. sapiens* (hypothetical LOC400931 (LOC400931), non-coding RNA), Human DNA (sequence from clone RP4-695O20 on chromosome 22q13.1-13.33 contains the 3' end of a novel gene and a CpG island), *Pan paniscus* (let-7a-3 gene) and Pan troglodytes (let-7a-3 gene) having e-value 5.21E-10 showed 84.00% similarities. *Macaca mulatta* (let-7a-3 (MIRLET7A-3), *Bos Taurus* (let-7a-3 (MIRLET7A-3)) and *H. sapiens* (let-7a-3 (MIRLET7A3) which have e-value 1.87E-09 showed 83.80% similarities. Zebrafish DNA (sequence from clone CH211-233F11 in linkage group 6) with 3.13E-07 showed 82.40%, Zebrafish DNA (sequence from clone DKEYP-116F3) and Zebrafish DNA (sequence from clone CH211-194J14) has e-value 4.06E-06 and showed 80.50%, whereas *Tetraodon nigroviridis* (full-length cDNA) has e-value 5.25E-05 and shows 80.00% pairwise

similarities, whereas organism with e-value 2.34E-33 showed 100% pair wise identity (Table 2) (Johnson et al., 2005; Lin et al., 2007; Chang et al., 2008).

Similarity and alignment analysis was also done for *H. sapiens* microRNA 145 (MIR145). The results obtained showed 100% pair wise similarities among different organisms carrying miRNA 145 gene with e-value 9.49E-38. Whereas, few are found distant including *M. mulatta* (mir-145 (MIR145), *Macaca nemestrina* (mir-145 gene) and *M. mulatta* (mir-145 gene, *Mus musculus* (chromosome 18, clone RP24-286B14) having e-value 4.42E-36 and showing 98.90% identity. *M. musculus* (BAC clone RP24-390G17 from chromosome 18), *M. musculus* (BAC clone RP23-202I21 from 18) having e-value 5.71E-35 shows 98.80% pairwise similarities. *M. musculus* (microRNA 145 (Mir145)) and *Monodelphis domestica* (mir-145 (Mir145)) which has e-value 4.48E-26, shows 98.60% pairwise similarities (Table 3) (Lin et al., 2007; John et al., 2004; Huang and Gu, 2007). *H. sapiens* miR-99-precursor-21 micro RNA was searched for its similarities with different organisms using BLAST.

**Table 2.** *H. sapiens* microRNA let-7a-1 (MIRLET7A1), microRNA.

<b>E-value</b>	<b>Name</b>	<b>Description</b>	<b>Organism</b>	<b>Sequence length</b>	<b>Pairwise identity (%)</b>
2.34E-33	NR_032352	<i>Macaca mulatta</i> microRNA let-7a-1 (MIRLET7A-1), microRNA	<i>Macaca mulatta</i>	80	100.00
2.34E-33	NR_030971	<i>Bos taurus</i> microRNA let-7a-1 (MIRLET7A-1), microRNA	<i>Bos taurus</i>	80	100.00
2.34E-33	NR_029476	<i>Homo sapiens</i> microRNA let-7a-1 (MIRLET7A1), microRNA	<i>Homo sapiens</i>	80	100.00
2.34E-33	NR_029725	<i>Mus musculus</i> microRNA let7a-1 (Mirlet7a-1), microRNA	<i>Mus musculus</i>	80	100.00
2.34E-33	AC147162	<i>Mus musculus</i> BAC clone RP24-270A10 from chromosome 13, complete sequence	<i>Mus musculus</i>	80	100.00
2.34E-33	AC124517	<i>Mus musculus</i> BAC clone RP23-387K15 from chromosome 13, complete sequence	<i>Mus musculus</i>	80	100.00
2.34E-33	AL158152	Human DNA sequence from clone RP11-2B6 on chromosome 9q22.2-31.1 Contains a novel gene and a CpG island, complete sequence	<i>Human DNA</i>	80	100.00
2.34E-33	AY866358	<i>Macaca nemestrina</i> microRNA let-7a-1 and microRNA let-7f-1 genes, complete sequence	<i>Macaca nemestrina</i>	80	100.00
2.34E-33	AY866357	<i>Pan troglodytes</i> microRNA let-7a-1 and microRNA let-7f-1 genes, complete sequence	<i>Pan troglodytes</i>	80	100.00
2.34E-33	AY866356	<i>Pongo pygmaeus</i> microRNA let-7a-1 and microRNA let-7f-1 genes, complete sequence	<i>Pongo pygmaeus</i>	80	100.00
2.34E-33	AY866355	<i>Macaca mulatta</i> microRNA let-7a-1 and microRNA let-7f-1 genes, complete sequence	<i>Macaca mulatta</i>	80	100.00
2.34E-33	AY866354	Lemur catta microRNA let-7a-1 and microRNA let-7f-1 genes, complete sequence	<i>Lemur catta</i>	80	100.00
2.34E-33	AY866353	<i>Pan paniscus</i> microRNA let-7a-1 and microRNA let-7f-1 genes, complete sequence	<i>Pan paniscus</i>	80	100.00
2.34E-33	AY866352	<i>Ateles geoffroyi</i> microRNA let-7a-1 and microRNA let-7f-1 genes, complete sequence	<i>Ateles geoffroyi</i>	80	100.00
2.34E-33	AY866351	<i>Saguinus labiatus</i> microRNA let-7a-1 and microRNA let-7f-1 genes, complete sequence	<i>Saguinus labiatus</i>	80	100.00
2.34E-33	AY866350	Gorilla gorilla microRNA let-7a-1 and microRNA let-7f-1 genes, complete sequence.	<i>Gorilla gorilla</i>	80	100.00
2.34E-33	AY866349	<i>Lagothrix lagotricha</i> microRNA let-7a-1 and microRNA let-7f-1 genes, complete sequence	<i>Lagothrix lagotricha</i>	80	100.00

Table 2. cont.

1.09E-31	NR_032235	Monodelphis domestica microRNA let-7a-1 (Mirlet7a-1), microRNA	<i>Monodelphis domestica</i>	77	100.00
2.36E-28	NR_031458	<i>Gallus gallus</i> microRNA let-7a-1 (MIRLET7A-1), microRNA	<i>Gallus gallus</i>	80	96.30
8.47E-28	NR_029976	Danio rerio microRNA let7a-1 (mirlet7a-1), microRNA	<i>Danio rerio</i>	77	97.40
8.47E-28	CR548622	Zebrafish DNA sequence from clone DKEY-189A9 in linkage group 11, complete sequence	<i>Zebrafish DNA</i>	77	97.40
3.09E-17	CR354430	Zebrafish DNA sequence from clone CH211-218P2 in linkage group 25, complete sequence	<i>Zebrafish DNA</i>	80	88.80
1.11E-16	NR_031396	<i>Gallus gallus</i> microRNA let-7a-3 (MIRLET7A-3), microRNA	<i>Gallus gallus</i>	78	88.50
5.21E-10	FP325332	Human DNA sequence from clone RP4-695O20 on chromosome 22q13.1-13.33, complete sequence	<i>Human DNA</i>	81	84.00
5.21E-10	NR_027033	<i>Homo sapiens</i> hypothetical LOC400931 (LOC400931), non-coding RNA	<i>Homo sapiens</i>	81	84.00%
5.21E-10	AL049853	Human DNA sequence from clone RP4-695O20 on chromosome 22q13.1-13.33 Contains the 3' end of a novel gene and a CpG island, complete sequence	<i>Human DNA</i>	81	84.00%
5.21E-10	AY866052	Pan paniscus microRNA let-7a-3 gene, complete sequence	<i>Pan paniscus</i>	81	84.00%
5.21E-10	AY866050	Pan troglodytes microRNA let-7a-3 gene, complete sequence	<i>Pan troglodytes</i>	81	84.00%
1.87E-09	NR_032354	<i>Macaca mulatta</i> microRNA let-7a-3 (MIRLET7A-3), microRNA	<i>Macaca mulatta</i>	80	83.80%
1.87E-09	NR_031384	<i>Bos taurus</i> microRNA let-7a-3 (MIRLET7A-3), microRNA	<i>Bos taurus</i>	80	83.80%
1.87E-09	NR_029478	<i>Homo sapiens</i> microRNA let-7a-3 (MIRLET7A3), microRNA	<i>Homo sapiens</i>	80	83.80%
3.13E-07	CR847944	Zebrafish DNA sequence from clone CH211-233F11 in linkage group 6, complete sequence	<i>Zebrafish DNA</i>	74	82.40%
4.06E-06	CR936540	Zebrafish DNA sequence from clone DKEYP-116F3, complete sequence	<i>Zebrafish DNA</i>	82	80.50%
4.06E-06	BX897734	Zebrafish DNA sequence from clone CH211-194J14, complete sequence	<i>Zebrafish DNA</i>	82	80.50%

**Table 2.** cont

1.46E-05	DQ581032	<i>Homo sapiens</i> piRNA piR-49144, complete sequence	<i>Homo sapiens</i>	30	100.00%
5.25E-05	CR716857	<i>Tetraodon nigroviridis</i> full-length cDNA	<i>Tetraodon nigroviridis</i>	80	80.00%

**Table 3.** *H. sapiens* microRNA 145 (MIR145), microRNA.

E-value	Name	Description	Organism	Sequence length	Pairwise identity (%)
9.49E-38	NR_031976	<i>Pan troglodytes</i> microRNA mir-145 (MIR145), microRNA	<i>Pan troglodytes</i>	88	100.00
9.49E-38	NR_030906	<i>Bos taurus</i> microRNA mir-145 (MIR145), microRNA	<i>Bos taurus</i>	88	100.00
9.49E-38	NR_029686	<i>Homo sapiens</i> microRNA 145 (MIR145), microRNA	<i>Homo sapiens</i>	88	100.00
9.49E-38	NR_027180	<i>Homo sapiens</i> hypothetical LOC728264 (LOC728264), non-coding RNA	<i>Homo sapiens</i>	88	100.00
9.49E-38	AK093957	<i>Homo sapiens</i> cDNA FLJ36638 fis, clone TRACH2018950	<i>Homo sapiens</i>	88	100.00
9.49E-38	AC131025	<i>Homo sapiens</i> chromosome 5 clone RP11-394O4, complete sequence	<i>Homo sapiens</i>	88	100.00
9.49E-38	AC008681	<i>Homo sapiens</i> chromosome 5 clone CTB-53M17, complete sequence	<i>Homo sapiens</i>	88	100.00
9.49E-38	AY865899	<i>Pongo pygmaeus</i> microRNA mir-145 gene, complete sequence	<i>Pongo pygmaeus</i>	88	100.00
9.49E-38	AY865898	Gorilla gorilla microRNA mir-145 gene, complete sequence	<i>Gorilla gorilla</i>	88	100.00
9.49E-38	AY865897	<i>Pan troglodytes</i> microRNA mir-145 gene, complete sequence	<i>Pan troglodytes</i>	88	100.00
4.42E-36	NR_031975	<i>Macaca mulatta</i> microRNA mir-145 (MIR145), microRNA	<i>Macaca mulatta</i>	88	98.90
4.42E-36	AY865900	<i>Macaca nemestrina</i> microRNA mir-145 gene, complete sequence	<i>Macaca nemestrina</i>	88	98.90
4.42E-36	AY865896	<i>Macaca mulatta</i> microRNA mir-145 gene, complete sequence	<i>Macaca mulatta</i>	88	98.90
5.71E-35	AC115691	<i>Mus musculus</i> chromosome 18, clone RP24-286B14, complete sequence	<i>Mus musculus</i>	86	98.80



Table 3. cont.

5.71E-35	AC125157	<i>Mus musculus</i> BAC clone RP24-390G17 from chromosome 18, complete sequence	<i>Mus musculus</i>	86	98.80
5.71E-35	AC148011	<i>Mus musculus</i> BAC clone RP23-202I21 from 18, complete sequence	<i>Mus musculus</i>	86	98.80
9.63E-28	NR_032936	<i>Equus caballus</i> microRNA mir-145 (MIR145), microRNA	<i>Equus caballus</i>	70	100.00
4.48E-26	NR_032181	<i>Monodelphis domestica</i> microRNA mir-145 (Mir145), microRNA	<i>Monodelphis domestica</i>	70	98.60
4.48E-26	NR_029557	<i>Mus musculus</i> microRNA 145 (Mir145), microRNA	<i>Mus musculus</i>	70	98.60
9.49E-38	NR_031976		<i>Pan troglodytes</i>	88	100.00

The organisms showed 100% pairwise identity with each other and having e-value 6.63E-34 for miR-99-precursor-21 miRNA as compared to *Equus caballus* (mir-99a-2 (MIR99A-2)) and *Pongo pygmaeus* (mir-99a gene) which have e-value 3.09E-32 and shows 98.80% pairwise identity. *M. musculus* (BAC clone RP24-252G15 from 16) and *M. musculus* (BAC clone RP23-156A3 from chromosome 16) having e-value 3.11E-27, shows 95.10% identity. *G. gallus* (mir-99a (MIR99A) shows 94.70% which has e-value 6.73E-24 and *M. musculus* (99a (Mir99a) with e-value 1.13E-21 shows 96.90% pairwise identity (Table 4) (Lagos-Quintana et al., 2001; Lau et al., 2001; Sempere et al., 2006). Das (2009) also performed BlastN to find hsa-miR-650 homologs in the draft genome sequences of 10 mammalian species including human, chimpanzee, orangutan, and macaque, and find that miRNAs were conserved across the animal kingdom.

### Phylogenetic analysis

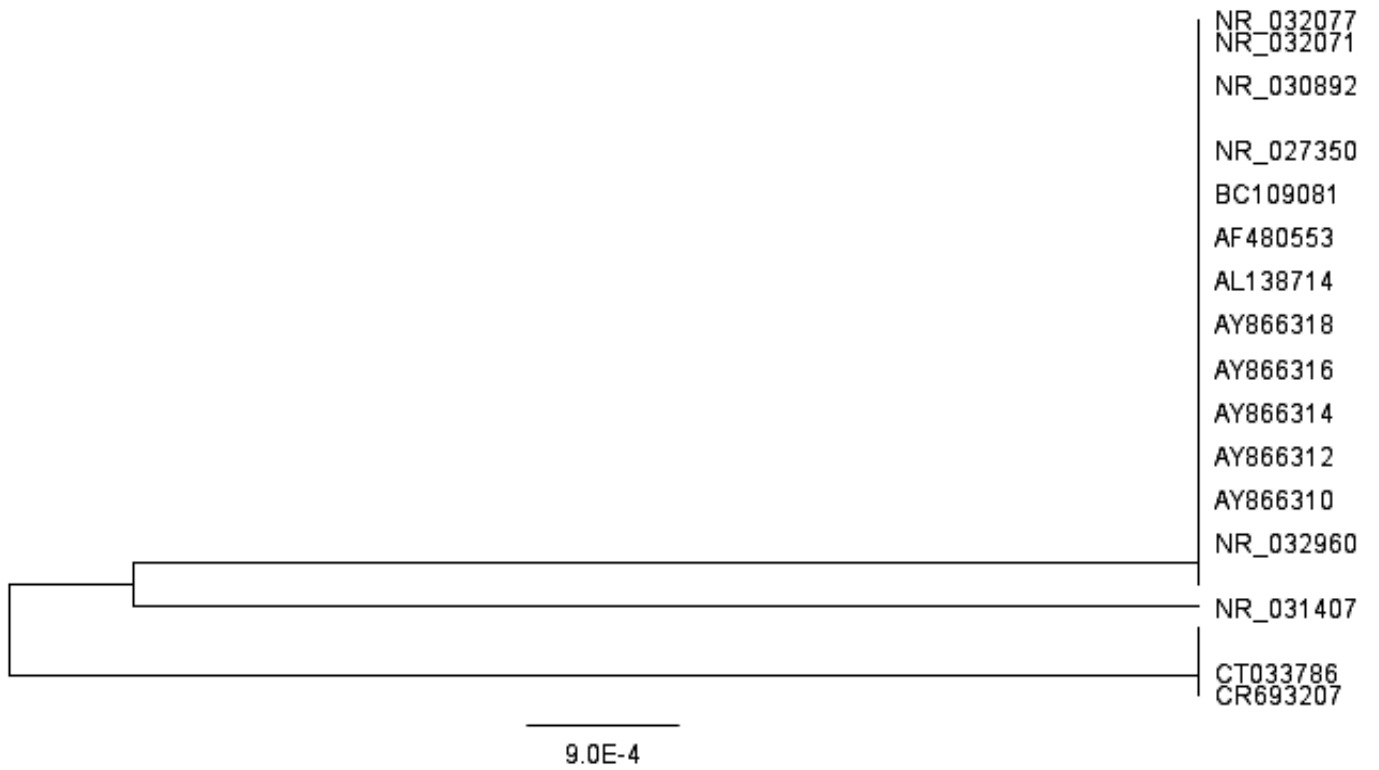
Phylogenetic analysis of the four *H. sapiens* miRNA (18a (MIR18A), let-7a-1 (MIRLET7A1), 145 (MIR145) and miR-99-precursor-21) has been done through Geneious software using Tamura-Nei Algorithm (Tamura and Nei, 1993). The UPGMA rooted tree diagram of miRNA 18a (MIR18A) shows different clusters formation. Organism that originated from same ancestors having same e-value and 100% pairwise identity, are placed in same clusters, whereas, those which are distant from each other are placed in separate clusters. Organisms with accession numbers NR\_032077, DQ\_599458, AY\_366317 and AL\_162375 lie in the same clusters. Those organisms which have accession number CT\_033786 lie in separate cluster on the basis of their e-value and percentage pairwise identity. Whereas, organisms with accession

clusters are distant from the rest of the clusters, it is number NR\_031407 lie in different cluster. As these two obvious that the rest of the organisms evolved from these two clusters (Figure 1).

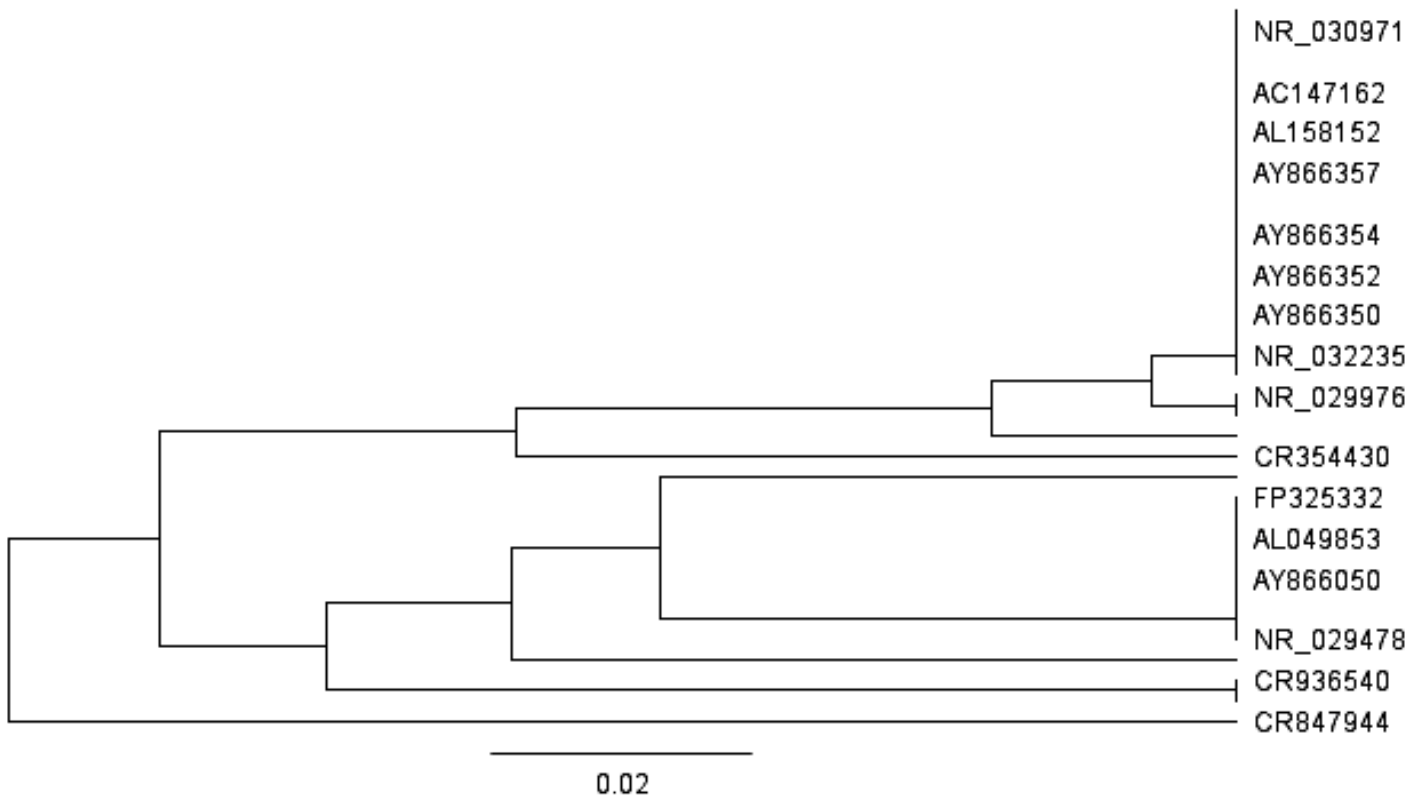
The miRNA molecules showed different divergence patterns (Figures 1 to 4) (Guo and Lu, 2010). *H. sapiens* microRNA let-7a-1 (MIRLET7A1) was analyzed for its evolutionary relationship with other organisms through Geneious software by rooted tree diagram using Tamura-Nei algorithm. Different clusters were obtained showing the ancestral relationship of different organisms. Organisms under accession number CR\_847944 are the distinct organisms found in one cluster. The rest of the organisms forming one cluster are thought to be evolved from this cluster (Figure 2). *H. sapiens* microRNA 145 (MIR145) has been analyzed for phylogenetic studies using Tamura-Nei algorithm. The cluster formation predicts the ancestral relationship of different organisms. Organisms with accession number NR\_032181 are the most distant cluster followed by NR\_029557 and NR\_031975. Accession numbers AY\_865897, NR\_032936, NR\_027180, NR\_031976, NR\_030906, NR\_029686, AC\_131025, AY\_865900, AC\_148011, AC\_125157 and AC\_115691 related organism lie in the same clusters based on their e-value and percentage pairwise identity are evolved from the distant clusters (Figure 3). The data obtained for *H. sapiens* miR-99-precursor-21 micro RNA shows that organisms with accession number NR\_031398 are distinct in the UPGMA rooted tree diagram. Similarly, organisms with accession number NR\_029535 are also distinct from the rest of the clusters. Whereas, those organisms having accession numbers NR\_029514, BS\_000028 and NR\_032012 lie close in one cluster, showing their evolutionary relevancy and also, these organisms are evolved from those distinct organisms (Figure 4) (Berezikov et al., 2005; Bentwich et al., 2005; Hertel et al., 2006).

**Table 4.** *H. sapiens* miR-99-precursor-21 micro RNA complete sequence.

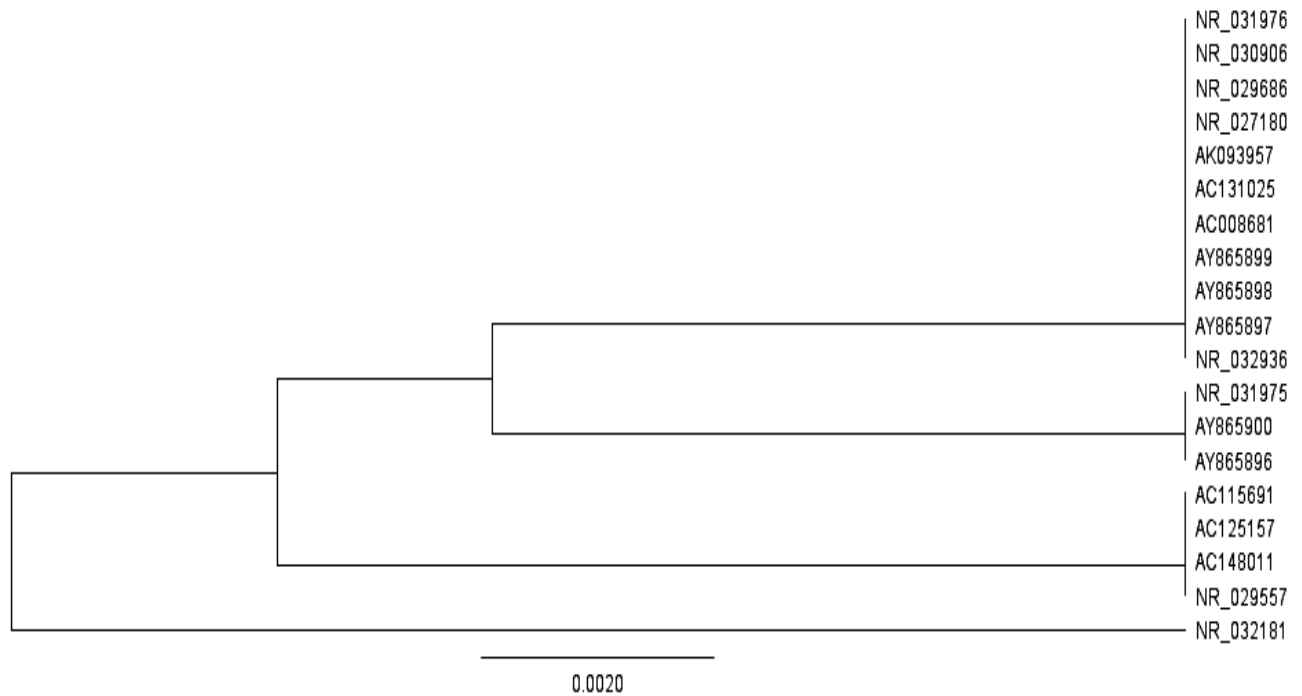
E-value	Name	Description	Organism	Sequence length	Pairwise identity (%)
6.63E-34	NR_032012	Pan troglodytes microRNA mir-99a (MIR99A), microRNA	<i>Pan troglodytes</i>	81	100.00
6.63E-34	NR_032011	<i>Macaca mulatta</i> microRNA mir-99a (MIR99A), microRNA	<i>Macaca mulatta</i>	81	100.00
6.63E-34	NR_030927	<i>Bos taurus</i> microRNA mir-99a (MIR99A), microRNA	<i>Bos taurus</i>	81	100.00
6.63E-34	NR_029514	<i>Homo sapiens</i> microRNA 99a (MIR99A), microRNA	<i>Homo sapiens</i>	81	100.00
6.63E-34	AF480571	<i>Homo sapiens</i> miR-99-precursor-21 micro RNA, complete sequence	<i>Homo sapiens</i>	81	100.00
6.63E-34	BS000028	Pan troglodytes chromosome 22 clone:PTB-087F19, map 22, complete sequences	<i>Pan troglodytes</i>	81	100.00
6.63E-34	AY866059	Pan paniscus microRNA mir-99a gene, complete sequence	<i>Pan paniscus</i>	81	100.00
6.63E-34	AY866058	<i>Macaca nemestrina</i> microRNA mir-99a gene, complete sequence	<i>Macaca nemestrina</i>	81	100.00
6.63E-34	AY866057	<i>Lagothrix lagotricha</i> microRNA mir-99a gene, complete sequence	<i>Lagothrix lagotricha</i>	81	100.00
6.63E-34	AY866055	<i>Gorilla gorilla</i> microRNA mir-99a gene, complete sequence	<i>Gorilla gorilla</i>	81	100.00
6.63E-34	AY866054	Pan troglodytes microRNA mir-99a gene, complete sequence	<i>Pan troglodytes</i>	81	100.00
6.63E-34	AY866053	<i>Macaca mulatta</i> microRNA mir-99a gene, complete sequence	<i>Macaca mulatta</i>	81	100.00
6.63E-34	AP001667	<i>Homo sapiens</i> genomic DNA, chromosome 21q, section 11/105	<i>Homo sapiens</i>	81	100.00
6.63E-34	AP000962	<i>Homo sapiens</i> genomic DNA, chromosome 21q21.1-q21.2, clone:B821P16, LL56-APP region, complete sequence	<i>Homo sapiens</i>	81	100.00
3.09E-32	NR_033057	<i>Equus caballus</i> microRNA mir-99a-2 (MIR99A-2), microRNA	<i>Equus caballus</i>	81	98.80
3.09E-32	AY866056	<i>Pongo pygmaeus</i> microRNA mir-99a gene, complete sequence	<i>Pongo pygmaeus</i>	81	98.80
3.11E-27	AC122451	<i>Mus musculus</i> BAC clone RP24-252G15 from 16, complete sequence	<i>Mus musculus</i>	81	95.10
3.11E-27	AC161457	<i>Mus musculus</i> BAC clone RP23-156A3 from chromosome 16, complete sequence	<i>Mus musculus</i>	81	95.10
6.73E-24	NR_031398	<i>Gallus gallus</i> microRNA mir-99a (MIR99A), microRNA	<i>Gallus gallus</i>	75	94.70
1.13E-21	NR_029535	<i>Mus musculus</i> microRNA 99a (Mir99a), microRNA	<i>Mus musculus</i>	65	96.90



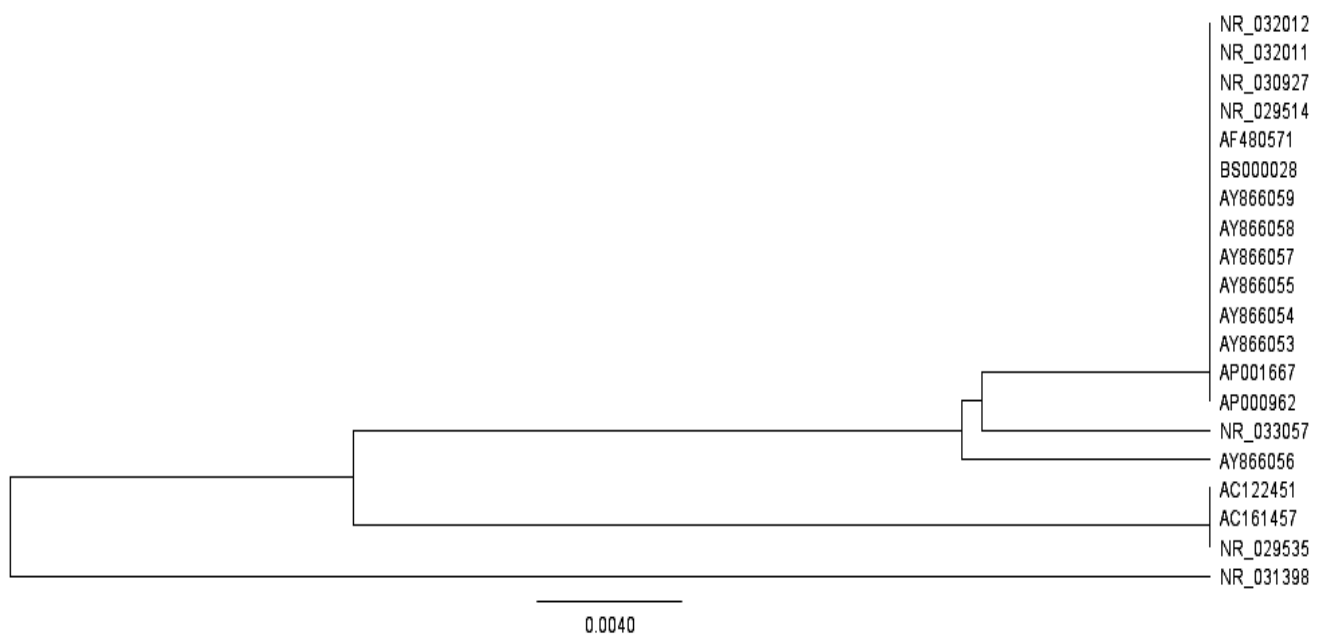
**Figure 1.** UPGMA tree diagram of *H. sapiens* microRNA 18a (MIR18A), microRNA.



**Figure 2.** UPGMA tree diagram of *H. sapiens* microRNA let-7a-1 (MIRLET7A1), microRNA.



**Figure 3.** UPGMA tree diagram of *H. sapiens* microRNA 145 (MIR145), microRNA.



**Figure 4.** UPGMA tree diagram of *H. sapiens* miR-99-precursor-21 micro RNA, complete sequence.

**REFERENCES**

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990). Basic local alignment search tool. *J. Mol. Biol.* 215: 403-410.

Bentwich I, Avniel AA, Karov Y, Aharonov R, Gilad S, Barad O, Barzilai A, Einat P, Einav U, Meiri E, Sharon E, Spector Y, Bentwich Z (2005). Identification of hundreds of conserved and nonconserved human microRNAs. *Nat. Genet.* 37: 766-770.

Berezikov E, Guryev V, van de Belt J, Wienholds E, Ronald Plasterk HA (2005). Phylogenetic Shadowing and Computational Identification of Human microRNA Genes. *Cell*, 120: 21-24.

Chang TC, Duonan Y, Yun SL, Erik AW, Dan EA, Kristin MW, Chi VD,

- Andrei TT, Joshua TM (2008). Widespread microRNA repression by Myc contributes to tumorigenesis. *Nat. Genet.* 40(1): 43-50.
- Das S (2009). Evolutionary Origin and Genomic Organization of MicroRNA Genes in Immunoglobulin Lambda Variable Region Gene Family. *Mol. Biol. Evol.* 26(5): 1179-1189.
- Filipowicz W, Bhattacharyya SN, Sonenberg N (2008). Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat. Rev. Genet.* 9: 102-114.
- George AC, Carlo MC (2006). MicroRNA-Cancer Connection: The Beginning of a New Tale. *Canc. Res.* 66: 7390-7394.
- Guo L, Lu Z (2010) The Fate of miRNA Strand through Evolutionary Analysis: Implication for Degradation as Merely Carrier Strand or Potential Regulatory Molecule? *PLoS ONE*, 5(6): p. 11387.
- Guo L, Sun BL, Sang F, Wang W, Lu ZH (2009). Haplotype Distribution and Evolutionary Pattern of miR-17 and miR-124 Families Based on Population Analysis. *PLoS ONE*, 4: p. 7944.
- Hertel J, Manuela L, Kristin M, Claudia F, Andrea T, Christoph F, Ivo LH, Peter FS (2006). The expansion of the metazoan microRNA repertoire. *BMC. Genomics*, 7: p. 25.
- Huang Y, Gu X (2007). A bootstrap based analysis pipeline for efficient classification of phylogenetically related animal miRNAs. *BMC. Genomics*, 8: p. 66.
- John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS (2004). Human MicroRNA Targets. *PLoS Biol.* 2(11): p. 363.
- Johnson SM, Grosshans H, Shingara J, Byrom M, Jarvis R, Cheng A, Labourier E, Reinert KL, Brown D, Slack FJ (2005). RAS is regulated by the let-7 microRNA family. *Cell*, 120: 635-647.
- Kristen MN, Glen JW (2008). MicroRNAs and cancer: past, present, and potential future. *Mol. Cancer Ther.* 7: 3655-3660.
- Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T (2001). Identification of novel genes coding for small expressed RNAs. *Science*, 294: 853-858.
- Lau NC, Lim LP, Weinstein EG, Bartel DP (2001). An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science*, 294: 858-862.
- Lee RC, Ambros V (2001). An extensive class of small RNAs in *Caenorhabditis elegans*. *Science*, 294: 862-864.
- Lewis BP, Burge CB, Bartel DP (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell*, 120: 15-20.
- Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB (2003). Prediction of mammalian microRNA targets. *Cell*, 115: 787-798.
- Lin YC, Li CH, Ming WK, John Y, Huan HK, Wan LL, Ruey JL, Alice LY, Wen HL (2007). Human *TRIM71* and Its Nematode Homologue Are Targets of *let-7* MicroRNA and Its Zebrafish Orthologue Is Essential for Development. *Mol. Biol. Evol.* 24(11): 2525-2534.
- Liu N, Okamura K, Tyler DM, Phillips MD, Chung WJ, et al. (2008). The evolution and functional diversification of animal microRNA genes. *Cell Res.* 18: 985-996.
- Sempere LF, Cole CN, Mcpeek MA, Peterson KJ (2006). The phylogenetic distribution of metazoan microRNAs: insights into evolutionary complexity and constraint. *J. Exp. Zool.* 306B: 1-14.
- Tamura K, Nei M (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* 10: 512-526.
- Wang X, Jing Z, Fei L, Jin G, Tao H, Xuegong Z, Yanda L (2005). MicroRNA identification based on sequence and structure alignment. *Bioinformatics*, 21(18): 3610-3614.
- Warthmann N, Sandip D, Christa L, Detlef W (2008). Comparative analysis of the mir319a microRNA locus in *Arabidopsis* and related Brassicaceae. *Mol. Biol. Evol.* 25(5): 892-902.
- Yukihiro A, Yoshihito N, Tomoki N (2007). MicroRNA-143 and -145 in Colon Cancer. *DNA. Cell Bio.* 26(5): 311-320.
- Zhang B, Pan X, Cobb GP, Anderson TA (2007). microRNAs as oncogenes and tumor suppressors. *Dev. Biol.* 302(1): 1-12.