COMPUTATIONAL IDENTIFICATION OF PUTATIVE miRNA HOMOLOGS IN ZEBRAFINCH REVEALS PLATYPUS SIGNATURES

Ganesh Sathyamurthy* N. Ramachandra Swamy

*Department of Biochemistry, Bangalore University, Bangalore-560 001, India

ABSTRACT: MicroRNAs (miRNAs) constitute an extensive class of non coding, gene regulating RNAs. Till date, there are no detailed reports of miRNAs in Zebra finch (*Taeniopygia guttata*). Here we adopt the method of homology search for the identification of miRNAs in *T. guttata* by using the pre-miRNAs of all known vertebrates. We report 221 miRNAs, of which 46 miRNAs are having 100% identity to pre miRNA and 208 miRNA to mature sequences. 104 sequences have mismatches at pre-miRNA level but have conserved mature miRNA. Thus the functions of mature sequences have been conserved during evolution of *T. guttata* genome. We also find that 157 miRNAs are closer to *Gallus gallus*, which is expected. Of which 117 sequences do not have 100% identity at pre-miRNA level. More importantly we find that 23 miRNAs come from platypus giving a remarkable clue to their relation with avian members. We also report conservation of 10 miRNA clusters. Five of these clusters are conserved with *G. gallus*, while the other 5 are altered with respect to gene orientation in the genome of *T. guttata,* and one among them is repositioned to chromosome 4a in *T. guttata.* These 5 altered gene orientations are unique to *T. guttata* genome and could have arisen due to divergence during evolution

Key words: Evolution, Divergence, *Taeniopygia guttata,* miRNA cluster, *Gallus gallus,* Platypus, Phylogenetic shadowing

INTRODUCTION

MicroRNAs (miRNAs) are a class of endogenous, small non-coding, single stranded RNA. They comprise approximately 19-25 nucleotides that are embedded within the stem regions of hairpin transcripts, called pre-miRNAs (1). The first miRNAs lin-5 (2) and let 7 (2), were identified in *C.elegans* about a decade ago. Many endogenously encoded miRNAs have been detected in mammals, plants, insects, worms and viruses through computational methods (3 – 13) and confirmed their expression by cloning, northern blotting, microarray assay and sequencing of short RNA molecules (1, 2, 14- 19).

By March, 2009, 9539 hairpin sequence entries have been registered in miRBase (20) (21) (22). Sequence analysis have shown that some mature miRNAs are phylogenetically conserved, particularly in the first 7 – 8 residues at the 5` end in species of the same kingdom (13) . Quite a few mature miRNA sequences are conserved between animals and plants. For example, mir-854, has been identified in *C. elegans*, mouse, human and plants (23).

There are two major approaches for miRNA discovery, which are computational identification and cDNA cloning. Since some miRNAs are expressed at a low level and the expression of many miRNAs has spatio-temporal specificity, it is difficult to find them through cDNA cloning. However, computational approaches can predict the miRNAs specifically expressed or with low abundance. The hairpin sequences of precursor miRNAs are phylogenetically diverse. In addition, the genomic locations of miRNA precursor genes and the folding structures of miRNAs have been used to identify previously unknown miRNAs. The three characteristics that allow miRNA genes to be identified using computational approaches are: i. miRNAs are generally derived from 70-100 nucleotide precursor transcripts having an extended stem-loop structure; ii. miRNAs are usually conserved between genomes of related species; and iii. miRNAs display a characteristic pattern of evolutionary divergence (3-7, 15-18). Additionally genomic mapping of known miRNAs have enabled identification of orthologous miRNAs in other species where genomic annotations are lacking (24-25).

Experimental evidence reveals that miRNAs play important roles in a variety of disease, such as cancer, diabetes, viral infection, cardiac disease, as well as in stem cell biology (3 - 8, 18, 19). Some miRNAs are present in the genome as clusters where multiple miRNAs are aligned in the same orientation and transcribed as polycistronic structure, which may function synchronously and cooperatively. Studies have provided evidence for miRNA regulation of many essential oncogenes including *BCL2, RAS, MYC*, *p53* (26 - 30).

In this study, we focused on a computational search for novel miRNA homologs in *T. guttata*. We have searched and analyzed the *T. guttata* homologs from all known vertebrate pre-miRNAs and mature miRNA sequences. Interestingly we also find platypus miRNAs share more homologs with *T. guttata* miRNAs next to *G. gallus.* Further we present the statistical data of *T. guttata* miRNAs in comparison with other vertebrates. Additionally, we report, few clusters have changed the gene orientation when compared to *G. gallus*. Some of these differences may play a role in *T. guttata* divergence during evolution.

MATERIALS AND METHODS

G. gallus pre-miRNA sequences and *T. guttata* genome

G. gallus pre-miRNA sequences and mature miRNAs were obtained from the Welcome Trust Sanger Institute's miRBase, release 13, 2009 (http://www.microrna.sanger.ac.uk). 474 sequences were stored in fastA format and used for further analysis. The *T. guttata* Genome Build 1.1, Assembly 3.2.4, released on March 5, 2009 was used directly from NCBI's Zebrafinch's Genome. The Genomic sequence of Build 1, released in June 2008 was also used from UCSC site.

Identification of putative miRNA precursor regions from *T. guttata* genome

The all known vertebrate pre-miRNAs were used as the query sequences to identify the putative pre-miRNA sequences in the *T. guttata* genome. The BLAST tool (31) was used on *T. guttata* genome. The results obtained were stored in fastA format and used for further analysis. Maximum Bit Score and Identity percent values were used to iterate BLAST hits to reduce all false positive hits, thereby increasing most putative pre-miRNA homologs. The complete list of pre-miRNA of *T. guttata* hits are shown in Supplementary data.

Prediction of secondary structures

Due to the differences in the different structure analysis tools, 2D de novo analysis was performed for both *G. gallus* and *T. guttata* sequences. The tool used was RNAshapes (32), which was executed locally to use. The query sequences of both *G. gallus* and *T. guttata* were fed into the tool as Multi FastA format and the results were obtained in .ps extension file, which contained the secondary structure and the free energy value. The Post script file was used to visualize the secondary structure and the free energy, (Figure 1).

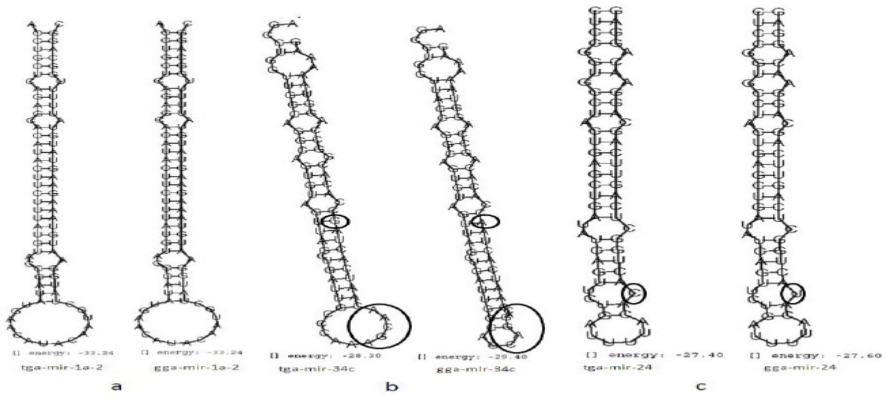

Figure 1. Secondary structure of putative miRNAs sequences of T. guttata and G. gallus compared with Free energy. RNAshapes was used for the predicition. A. mir-1a-2 have 100% identity in structure and sequence, b. mir-34c have 100% identity mature sequence and 89% identity in pre-miRNA sequence, c. mir-24 have 100% identity in mature and 98% identity in pre-miRNA sequence.

Mature miRNA analysis

To analyze the conservation of mature sequences in the pre-miRNA a local alignment was performed, using LALIGN (33), by comparing mature query miRNA sequence with putative pre-miRNA sequences from *T. guttata* genomic hits. The 100% and between 95% - 89% Identity of the miRNA genes are shown in supplementary Table 1.

Table 1. Comparison of pre-miRNA clusters between G.gallus and *T. guttata*, threshold distance is ~500 nt.

pre-miRNA	chromosome	Strand orientation		Finch		chicken	
		finch	chicken	Start	End	Start	End
mir-92	1	-	-	299703	299626	152248070	152248147
mir-19b	1	-	-	299826	299740	152248183	152248269
mir-20a	1	-	-	299963	299866	152248306	152248403
mir-19a	1	-	-	300130	300033	152248492	152248572
mir-18a	1	-	-	300270	300178	152248626	152248718
mir-17	1	-	-	300417	300333	152248781	152248865
mir-23b	z	-	-	3993631	3992713	41157406	41157491
mir-27b	z	-	-	3993317	3993225	41157642	41157738
mir-24	z	-	-	3992780	3992713	41158175	41158242
mir-218-1	4	-	-	7961587	7961479	77774698	77774806
mir-218-3	4	-	-	7961587	7961479	77774698	77774806
let-7j	26	-	-	532880	532798	1442697	1442779
let-7k	26	-	-	533068	532986	1442897	1442979
mir-367	4	-	+	2681318	2681246	58652350	58652422
mir-302d	4	-	+	2681471	2681404	58652214	58652282
mir-302c	4	-	+	2681803	2681747	58651576	58651640
mir-20b	4a	+	-	2095136	2095220	3970047	3970131
mir-18b	4a	+	-	2094938	2095018	3970228	3970311
mir-181a-1	8	-	+	5108744	5108641	2001561	2001664
mir-181b-1	8	-	+	5108556	5018468	2001750	2001838
let-7a-1	12	+	-	738816	738905	6302911	6303000
let-7f	12	+	-	739227	739313	6302497	6302583
mir-15a	1	+	-	13007147	13007229	173700493	173700575
mir-16-1	1	+	-	13007288	13007371	173700351	173700434

miRNA cluster analysis

The pre-miRNA sequences of *T. guttata* and *G. gallus* pre-miRNA sequences were used as queries for BLAT (http://genome.ucsc.edu/cgi-bin/hgBlat) analysis against their respective genomes in the UCSC Genome Server to identify genomic location. The genomic regions were identified and reported as in Table 1. Similarly to check the homology with other species, human and Zebra fish had the similar clusters conserved as shown in Table 4. 10 clusters were identified which are highly conserved between *G. gallus* and *T. guttata*. A graph was generated for mir-92/17 cluster using customized genome browser from UCSC server for comparing conservation between the two genomes which shows the comparison in human, *D. rerio*, *G. gallus* and *T. guttata* (Supplementary Data).

RESULTS

miRNAs in *T. guttata* Genome

The potential pre-miRNAs search in *T. guttata* was performed using BLAST tool. Pre miRNA from *T. guttata* were considered if the minimum cut off Identity was 85% with query pre-miRNA sequences apart from maximum bit score value.

Using these criteria, 221 putative *T. guttata* pre-miRNA sequences were identified (Figure 2). Of this 221, 46 miRNAs had 100% identity percent with query sequences followed by 103 miRNAs between 95-99.9 identity percent, 63 between 90-94.9% and 11 miRNAs up to 89%. At the mature sequence level, 208 miRNAs have 100% identity, followed by 11 between 95-99.9 identity percent, 2 each between 90-94.9% and upto 89%. The maximum gap at pre-miRNA level was 7, yet in this sequence the mature sequence had 100% identity. Some degree of mismatch was also observed in the mature sequence (Figure 3). Thus showing degree of divergence to *T. guttata* and making them species specific.

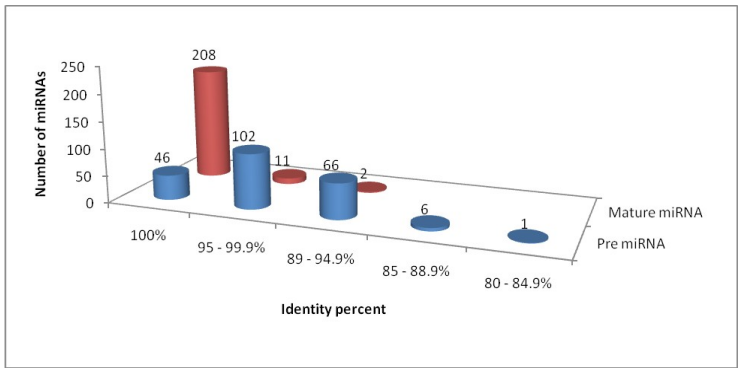

Figure. 2. Histogram showing number of pre- and mature miRNAs identified in *T. guttata.*

```
gga-mir-302b   cccuuuuACUUUAACAUGGAGGUGCUUUCUgugauuuuacaaaaguaagugcuuccauguuuuaguagagg
               :::::::::::::: :::::::::::::::::::::  :   :::::::::::::::::::::::::::::::
tgu-mir-302b   cccuuuuACUUUACCAUGGAGGUGCUUUCugugacauagaaaaaguaagugcuuccauguuuuaguagagg
```

a

```
tgu-mir-193 gugguuccagaguCGGGGUUUUGGGGGCGaGaUGagcuuauguuuuauccaacuggcccacaaagucccgcuuuugguggucacuuug
            ::: :::::::::::::::::::: :::::::::::: :: ::::::::::::::::::::::::::::::::::::::::::::::
oan-mir-193 gugauuccagaguCGGGGUUUUGAGGGCGaGaUGaguuuguguuuuauccaacuggcccacaaagucccgcuuuugguggucacuuug
```

b

```
oan-mir-1399  gugaaauuacUGaaUaGaUGUGaUUUGGggaaauuacaguuccuuguuuaguauguauugugcaugcccuacagccuucccuaaccaguagcucccaauu
              :::::::::::::: :::::::::::::::::::::::::::::::: :::::::::: :::::::::::::::::::::::::::::::::::::
tgu-mir-1399  gugaaauuacUGaaaaGaUGUGaUUUGGggaaauuacaguuccuugcuuaguauguacugugcaugcccuacagccuucccuaaccaguagcucccaauu
```

c

Figure 3. pre-miRNA sequence of T. guttata which is not 100% identical. pre-miRNA sequence of *T. guttata* which is not 100% identical. The mature sequences also have altered during evolution at one or two position. **a.** tgu-mir-302a and gga-mir-302a; **b**. tgu-mir-193 and oan-mir-193; **c**. tgu-mir-1399 and oan-mir-1399.

Phylogenetic Shadowing

This method was previously used to identify the human miRNAs conservation with mouse and rat (34). In our study, we tried to look at the query sequences from which we have got the final miRNAs after iterations. We find that of 221 sequences, 157 have come from G. gallus the closest vertebrate. Further, 40 sequences have 100% identity with pre-miRNAs, followed by 76 between 95-99.9 identity percent, 39 between 90-94.9% and 2 upto 89%. Similar observation was seen at mature sequence level, More interestingly, we also find that the next maximum number of 23 miRNAs have arrived from platypus making a remarkable signature that Platypus is more closely associated with *T. guttata* next to chicken. Our observations further support the previous reports (35) (36). Further to strength our observations 21 miRNAs of 23 have 100% identity at mature sequence level which shows the degree of conservation from platypus to *T. guttata*. (Figure 4, 5)

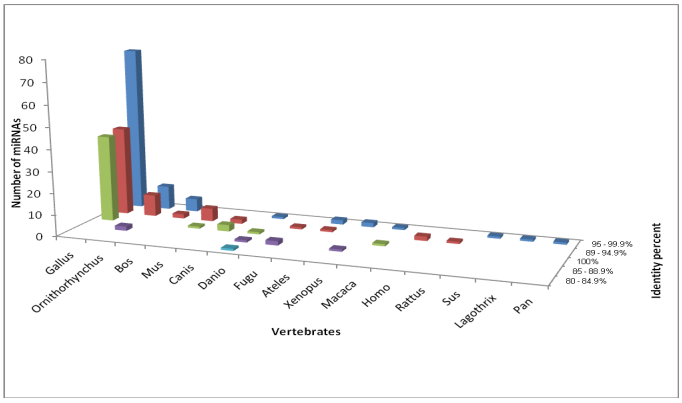

Figure 4. Histogram showing the number of identical pre-miRNAs distributed across the vertebrates species. The histogram is based on the identity percent between the query sequence of vertebrates and hit sequences of *T. guttata.*

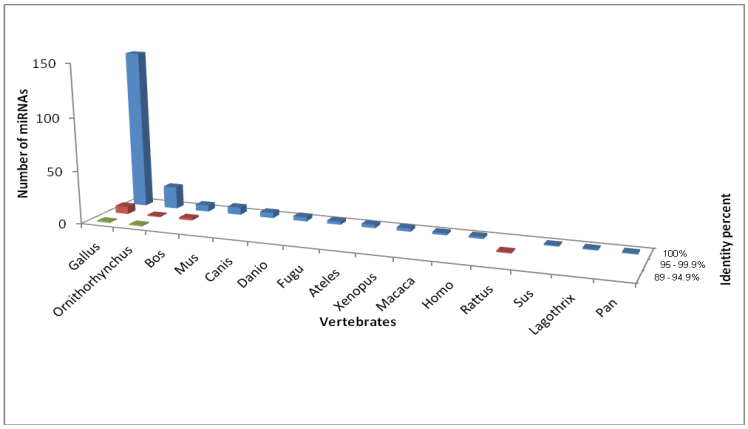

Figure 5. Histogram showing the number of identical mature miRNAs distributed across the vertebrates species. The histogram is based on the identity percent between the query sequence of vertebrates and hit sequences of *T. guttata.*

miRNA Gene Family

Typically a miRNA family contains several mature miRNAs which are identical or almost identical. In most cases, a minimum of two or three members are experimentally proved, while others are considered based on *in silico* prediction as mentioned previously (1, 37). It is not known that if miRNA with some differences in sequence at 5` and 3` end affect their function. In *G. gallus*, 475 miRNAs are present in ~395 families. In case of *T. guttata* we find 185 mature miRNAs are present in 119 gene families (Table 2). Out of these, 34 families have been experimentally validated in chicken, 31 in human and few in zebrafish (38 - 39). Also, we find that 11 miRNA gene families share common between platypus and *T. guttata*.

Table 2. Gene familes in T. guttata. Seed region is conserved among the family members.

Seed Sequence	Members
GAGGUA	let-7a-1; let-7a-2; let-7a-3; let-7b; let-7c; let-7d; let-7e; let-7e-2; let-7f; let-7g; let-7i; let-7j; let-7k
GGAAUG	miR-1; miR-1-1; miR-1-2
ACCCGU	miR-100; miR-100-2
CAGUAC	miR-101; miR-101-2; miR-101a-1; miR-101b
GCAGCA	miR-103-1; miR-103-2
ACCCUG	miR-10b; miR-10b-1
GGAGUG	miR-122; miR-122-1; miR-122-2
AAGGCA	miR-104-2; miR-104-3; miR-124; miR-124-1; miR-124-3; miR-124-5; miR-124a; miR-124a-2
CCCUGA	miR-125; miR-125b; miR-125b-1
CACAGU	miR-128; miR-128-2; miR-128a
CAGUGC	miR-130a; miR-130c
UGGUCC	miR-133a; miR-133a-1; miR-133a-2; miR-133b
AUGGCU	miR-135a-1; miR-135a-2; miR-135a-3
GCUGGU	miR-138-1; miR-138-2; miR-138a;
GUGUGC	miR-147; miR-147-1; miR-147-2; miR-147b
UGCAUA	miR-153; miR-153-1
GCAGCA	miR-15a; miR-15c; miR-16-1; miR-16-2
AACAUU	miR-181a-1; mir-181a-2; miR-181b-1; miR-181b-2;
UAAGGU	miR-18a; miR-18b
AACUGG	miR-193a; miR-193b
AGGUAG	miR-196-2; miR-196-3
CCCAGU	miR-199; miR-199-1; miR-199-2
UGGAAU	miR-1a-2; miR-1a-1
UAACAC	miR-200a; miR-200b
UCCCUU	miR-204-1; miR-204-2
AAAGUG	miR-20a; miR-20b
AAUCUC	miR-216; miR-216b
UUGUGC	miR-218-1; miR-218-2miR-218-3
CCACCA	miR-220b; miR-220d
UAGCAC	miR-29a; miR-29b-1; miR-29b-2; miR-29c
UAAACA	miR-30a; miR-30b; miR-30c; miR-30c-1; miR-30d; miR-30e; miR-30f
GUGCAU	miR-33-1; miR-33-2
AAUGCC	miR-365; miR-365-1; miR-365-2
AACCGU	miR-451; miR-454
CCUGCA	miR-460; miR-460b
GGAAGA	miR-7; miR-7-1; miR-7-2; miR-7-3; miR-7a-1
UCUUUG	miR-9; miR-9-2; miR-9-3; miR-9-4; miR-9-5
UAUUGC	miR-92; miR-92-1; miR-92a-1
AACCCG	miR-99; miR-99a

miRNA Clusters

miRNA clusters are defined as miRs present in the same direction and are transcribed in polycistronic transcriptional unit. It is well known that miRNAs quite often form clusters in the genome(40). We observed at least 9 different clusters at a threshold between 100 and 500 nt, which are found in Chromosome 1, 2, 4, 4a, 26 and Z, in *T. guttata.* Similar clusters were reported in case of *G. gallus* (Table 1). However, interestingly we find that between *G. gallus* and *T. guttata* there is no conservation of the orientation in mir-367/320, mir-18/20, mir-181, let7 and mir 15/16 clusters, but chromosome conservation have been retained, except for cluster mir-18/20. We also looked upon the conservation of these clusters in human and zebrafish and the results showed translocation of cluster across the genomes (Table 3).

Table 3. Comparison of miR clusters at 500nt threshold across species for conservation. The orientation and chromosome for Human, D. rerio, chicken were taken from miRGen database and finch's orientation and chromosome was taken for the BLAST hit region.

miRNA	Human		D. rerio		Chicken		Finch	
	Chr*	direction	chr	direction	chr	direction	chr	direction
mir-92	13	+	8	-	1	-	1	-
mir-19b	13	+	8	-	1	-	1	-
mir-20a	13	+	8	-	1	-	1	-
mir-19a	13	+	8	-	1	-	1	-
mir-18a	13	+	8	-	1	-	1	-
mir-17	13	+	8	-	1	-	1	-
mir-23b	9	+	10	+	z	-	z	-
mir-27b	9	+	10	+	z	-	z	-
mir-24	9	+	nil	nil	z	-	z	-
mir-757-1	nil	nil	nil	nil	2	-	2	-
mir-757-2	nil	nil	nil	nil	2	-	2	-
mir-367	4	-	nil	nil	4	+	4	-
mir-302d	4	-	nil	nil	4	+	4	-
mir-302c	4	-	nil	nil	4	+	4	-
mir-181a1	1	-	8	-	8	+	8	-
mir-181b1	1	-	8	-	8	+	8	-
let-7a1	9	+	11	+	12	-	12	+
let-7f	9	+	11	+	12	-	12	+
let-7j	nil	nil	6	+	26	-	26	-
let-7k	nil	nil	nil	nil	26	-	26	-
mir-15a	13	-	1	-	1	-	1	+
mir-16-1	13	-	1	-	1	-	1	+
mir-20b	x	-	14	-	4	-	4a	+
mir-18b	x	-	14	-	4	-	4a	+
mir-218-1	5	-	nil	nil	4	-	4	-
mir-218-3	nil	nil	nil	nil	4	-	4	-

*Chr= chromosome.

DISCUSSION

221 miRNAs were identified on the basis of homology search using all known vertebrate pre-miRNA sequences. They are all homologous and arranged in orthologous positions with *G. gallus* in most of the miRNA genes. All of the miRNAs genes, and clusters are highly conserved between both the genomes. These factors with G. gallus are more likely to be excepted for the simple reason that they belong to the same class. But the most interesting part relies on data of phylogenetic shadowing when compared with the query vertebrates. 23 miRNAs originate from the Platypus miRNAs. Our data is an add on data to strength the previous reports of platypus evolutionary observations (35, 36). Nearly 11 gene families are shared between platypus and *T. guttata*. Further, 21 miRNAs have 100% homologous with mature sequence indicating the fact of evolutionarily function conservation between phylogenetically separated species.

Within the avain family, 20.5% of the pre-miRNAs and 88.6% of the mature miRNAs are conserved in *G. gallus* miRNAs with 100% Identity. 124 (75.6%) miRNA sequences have altered nucleotides in the precursor sequences leaving behind the mature sequence with 100% identity. It important to point out that only 40 pre-miRNAs out of 185 have complete hairpin conservation, while 164 of 185 sequences have been modified leaving mature miRNA sequence untouched (Fig. 6). In pre-miRs, the folding and free energy have altered to near identical, accounting for the thermodynamic stability for the change in base sequences (Fig. 1, 6).

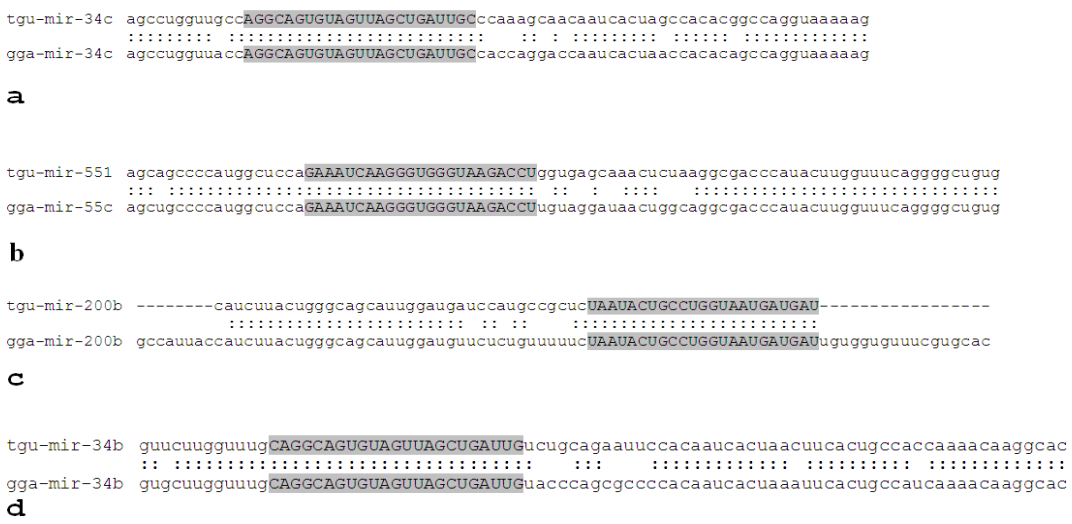

Figure 6. pre-miRNA sequence of T. guttata and G. gallus are not 100% identical, but the mature sequence remains unchanged during evolution. a. Tg- mir-34c and G. gallus; b. Tg-mir-551 and G. gallus; c.Tg- mir20b and gga-mir-20b; d:Tg-mir-34b and gga-mir-34b.

The observed 1 to 2 bases modifications in the 21 mature miRNAs seems to be unique for *T. guttata* genome (Figure. 3), which include both Indel and substitution in the mature sequence. Interestingly, the miRNAs at the seed region are untouched conserving its functionality. It could be noted that, 10 miRNAs which have altered sequences in pre-miRNAs and mature miRNAs could represent either new family or they might have undergone evolutionary changes. Further experimental validation is required to such a hypothesis, which is beyond the scope of this paper. These changes could be due to divergent evolution from *G. gallus*. Since most of the miRNAs have not been affected due to such a divergence, it indicates the importance of these miRNAs in the survival of the genome.

Conservation of gene cluster and orientation

The miR clusters (mir-17/92, mir-23/27, mir-757, mir-218, mir-let7j/k) identified on Chromosome 1, 2, 4, 26 and Z (Table 3 and 4) are conserved in Gene orientation and chromosome number in both *G. gallus* and *T. guttata*. In case of the clusters mir 302/367, mir-181, mir-let 7a/f, mir-15/16, although location of chromosome is same (1, 8, 12 and 26), the orientation in which miRNA genes oriented are different. Interestingly, we find that the cluster mir-18/20 is repositioned on chromosome 4a of *T. guttata* suggesting interchromosomal rearrangements. This observation is also supported by other studies previously reported (41). Later 5 clusters suggest the uniqueness of *T. guttata* from *G. gallus.* Therefore, this miRNA analysis could be also used for understanding changes occur during evolution (Supplementary Data).

More interestingly, among the 10 clusters described above, most have also been indentified in human and zebrafish. However, they are all located in different chromosomes. A few clusters such as mir-757, let-7j/k and mir-302/367 have been unique for avian alone (Table 4). Cluster mir-17/92 is reported in humans. miR- 17/92 cluster composed of 6 miRNAs (miR-17, 18a, 19a, 20a, 19b, 92) was found to be related to tumorigenesis and promoted tumor angiogenesis through targeting the anti-angiogenic thrombospondin-1 (Tsp 1) by miR-19 or connective tissue growth factor (CTGF) by miR-18 to down regulate their functions (13, 42). This cluster, termed as “Oncomirs”, is conserved in *G. gallus* and *T. guttata*, but the gene orientation is not well conserved with respect to human and zebrafish again making it unique to avian.

Further observed conservation of gene clusters with human and zebrafish indicate the functional conservation of these clusters. We also found that many of the miR genes were not present in the *T. guttata*, due to the fact that the sequencing and assembly are yet to be completed and may have to be identified for looking into the divergence during evolution. More importantly, *T. guttata* miRNAs give an addon clue to platypus relation to aves.

REFERENCES

1. Ambros, V., Bartel, B., Bartel, D.P., Burge, C.B., Carrington, J.C., Chen, X., Dreyfuss, G., Eddy, S.R., Griffiths-Jones, S., Marshall, M. *et al.* (2003) A uniform system for microRNA annotation. *Rna*, **9**, 277-279.
2. Reinhart, B.J., Slack, F.J., Basson, M., Pasquinelli, A.E., Bettinger, J.C., Rougvie, A.E., Horvitz, H.R. and Ruvkun, G. (2000) The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature*, **403**, 901-906.
3. 3.Grad, Y., Aach, J., Hayes, G.D., Reinhart, B.J., Church, G.M., Ruvkun, G. and Kim, J. (2003) Computational and experimental identification of C. elegans microRNAs. *Mol Cell*, **11**, 1253-1263.
4. Lai, E.C., Tomancak, P., Williams, R.W. and Rubin, G.M. (2003) Computational identification of Drosophila microRNA genes. *Genome Biol*, **4**, R42.
5. Wang, J.F., Zhou, H., Chen, Y.Q., Luo, Q.J. and Qu, L.H. (2004) Identification of 20 microRNAs from Oryza sativa. *Nucleic Acids Res*, **32**, 1688-1695.
6. Jones-Rhoades, M.W. and Bartel, D.P. (2004) Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Mol Cell*, **14**, 787-799.
7. Bonnet, E., Wuyts, J., Rouze, P. and Van de Peer, Y. (2004) Detection of 91 potential conserved plant microRNAs in Arabidopsis thaliana and Oryza sativa identifies important target genes. *Proc Natl Acad Sci U S A*, **101**, 11511-11516.
8. Yoon, S. and De Micheli, G. (2005) Prediction of regulatory modules comprising microRNAs and target genes. *Bioinformatics*, **21 Suppl 2**, ii93-100.
9. Wang, X., Zhang, J., Li, F., Gu, J., He, T., Zhang, X. and Li, Y. (2005) MicroRNA identification based on sequence and structure alignment. *Bioinformatics*, **21**, 3610-3614.
10. Nam, J.W., Shin, K.R., Han, J., Lee, Y., Kim, V.N. and Zhang, B.T. (2005) Human microRNA prediction through a probabilistic co-learning model of sequence and structure. *Nucleic Acids Res*, **33**, 3570-3581.
11. Dezulian, T., Remmert, M., Palatnik, J.F., Weigel, D. and Huson, D.H. (2006) Identification of plant microRNA homologs. *Bioinformatics*, **22**, 359-360.
12. Cao, J., Tong, C., Wu, X., Lv, J., Yang, Z. and Jin, Y. (2008) Identification of conserved microRNAs in Bombyx mori (silkworm) and regulation of fibroin L chain production by microRNAs in heterologous system. *Insect Biochem Mol Biol*, **38**, 1066-1071.
13. Helvik, S.A., Snove, O., Jr. and Saetrom, P. (2007) Reliable prediction of Drosha processing sites improves microRNA gene prediction. *Bioinformatics*, **23**, 142-149.
14. Lim, L.P., Lau, N.C., Weinstein, E.G., Abdelhakim, A., Yekta, S., Rhoades, M.W., Burge, C.B. and Bartel, D.P. (2003) The microRNAs of Caenorhabditis elegans. *Genes Dev*, **17**, 991-1008.
15. Carrington, J.C. and Ambros, V. (2003) Role of microRNAs in plant and animal development. *Science*, **301**, 336-338.
16. Lee, R.C., Feinbaum, R.L. and Ambros, V. (1993) The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell*, **75**, 843-854.
17. Lagos-Quintana, M., Rauhut, R., Lendeckel, W. and Tuschl, T. (2001) Identification of novel genes coding for small expressed RNAs. *Science*, **294**, 853-858.
18. Lau, N.C., Lim, L.P., Weinstein, E.G. and Bartel, D.P. (2001) An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. *Science*, **294**, 858-862.
19. Lee, R.C. and Ambros, V. (2001) An extensive class of small RNAs in Caenorhabditis elegans. *Science*, **294**, 862-864.

20. Griffiths-Jones, S., Grocock, R.J., van Dongen, S., Bateman, A. and Enright, A.J. (2006) miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res*, **34**, D140-144.
21. Griffiths-Jones, S. (2004) The microRNA Registry. *Nucleic Acids Res*, **32**, D109-111.
22. Ambros, V. (2004) The functions of animal microRNAs. *Nature*, **431**, 350-355.
23. Arteaga-Vazquez, M., Caballero-Perez, J. and Vielle-Calzada, J.P. (2006) A family of microRNAs present in plants and animals. *Plant Cell*, **18**, 3355-3369.
24. Xia, Q., Zhou, Z., Lu, C., Cheng, D., Dai, F., Li, B., Zhao, P., Zha, X., Cheng, T., Chai, C. *et al.* (2004) A draft sequence for the genome of the domesticated silkworm (Bombyx mori). *Science*, **306**, 1937-1940.
25. He, P.A., Nie, Z., Chen, J., Chen, J., Lv, Z., Sheng, Q., Zhou, S., Gao, X., Kong, L., Wu, X. *et al.* (2008) Identification and characteristics of microRNAs from Bombyx mori. *BMC Genomics*, **9**, 248.
26. Cimmino, A., Calin, G.A., Fabbri, M., Iorio, M.V., Ferracin, M., Shimizu, M., Wojcik, S.E., Aqeilan, R.I., Zupo, S., Dono, M. *et al.* (2005) miR-15 and miR-16 induce apoptosis by targeting BCL2. *Proc Natl Acad Sci U S A*, **102**, 13944-13949.
27. Talotta, F., Cimmino, A., Matarazzo, M.R., Casalino, L., De Vita, G., D'Esposito, M., Di Lauro, R. and Verde, P. (2009) An autoregulatory loop mediated by miR-21 and PDCD4 controls the AP-1 activity in RAS transformation. *Oncogene*, **28**, 73-84.
28. Solomon, A., Mian, Y., Ortega-Cava, C., Liu, V.W., Gurumurthy, C.B., Naramura, M., Band, V. and Band, H. (2008) Upregulation of the let-7 microRNA with precocious development in lin-12/Notch hypermorphic Caenorhabditis elegans mutants. *Dev Biol*, **316**, 191-199.
29. Hosako, H., Martin, G.S., Barrier, M., Chen, Y.A., Ivanov, I.V. and Mirkes, P.E. (2009) Gene and MicroRNA Expression in p53-Deficient Day 8.5 Mouse Embryos. *Birth Defects Research Part a-Clinical and Molecular Teratology*, **85**, 546-555
30. Georges, S.A., Chau, B.N., Braun, C.J., Zhang, X. and Dobbelstein, M. (2009) Cell cycle arrest or apoptosis by p53: are microRNAs-192/215 and -34 making the decision? *Cell Cycle*, **8**, 680-681.
31. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol*, **215**, 403-410.
32. Steffen, P., Voss, B., Rehmsmeier, M., Reeder, J. and Giegerich, R. (2006) RNAshapes: an integrated RNA analysis package based on abstract shapes. *Bioinformatics*, **22**, 500-503.
33. Huang, X.Q. and Miller, W. (1991) A Time-Efficient, Linear-Space Local Similarity Algorithm. *Advances in Applied Mathematics*, **12**, 337-357.
34. Berezikov, E., Guryev, V., van de Belt, J., Wienholds, E., Plasterk, R.H. and Cuppen, E. (2005) Phylogenetic shadowing and computational identification of human microRNA genes. *Cell*, **120**, 21-24.
35. Warren, W.C., Hillier, L.W., Marshall Graves, J.A., Birney, E., Ponting, C.P., Grutzner, F., Belov, K., Miller, W., Clarke, L., Chinwalla, A.T. *et al.* (2008) Genome analysis of the platypus reveals unique signatures of evolution. *Nature*, **453**, 175-183.
36. Grutzner, F., Rens, W., Tsend-Ayush, E., El-Mogharbel, N., O'Brien, P.C., Jones, R.C., Ferguson-Smith, M.A. and Marshall Graves, J.A. (2004) In the platypus a meiotic chain of ten sex chromosomes shares genes with the bird Z and mammal X chromosomes. *Nature*, **432**, 913-917.
37. Zhang, L., Ding, L., Cheung, T.H., Dong, M.Q., Chen, J., Sewell, A.K., Liu, X., Yates, J.R., 3rd and Han, M. (2007) Systematic identification of C. elegans miRISC proteins, miRNAs, and mRNA targets by their interactions with GW182 proteins AIN-1 and AIN-2. *Mol Cell*, **28**, 598-613.
38. Weber, M.J. (2005) New human and mouse microRNA genes found by homology search. *Febs J*, **272**, 59-73.

39. Glazov, E.A., Cottee, P.A., Barris, W.C., Moore, R.J., Dalrymple, B.P. and Tizard, M.L. (2008) A microRNA catalog of the developing chicken embryo identified by a deep sequencing approach. *Genome Research*, **18**, 957-964.
40. Fu, H., Tie, Y., Xu, C., Zhang, Z., Zhu, J., Shi, Y., Jiang, H., Sun, Z. and Zheng, X. (2005) Identification of human fetal liver miRNAs by a novel method. *FEBS Lett*, **579**, 3849-3854.
41. Stapley, J., Birkhead, T.R., Burke, T. and Slate, J. (2008) A linkage map of the zebra finch Taeniopygia guttata provides new insights into avian genome evolution. *Genetics*, **179**, 651-667.
42. Tong, C.Z., Jin, Y.F. and Zhang, Y.Z. (2006) Computational prediction of microRNA genes in silkworm genome. *J Zhejiang Univ Sci B*, **7**, 806-816.
43. Kadener, S., Rodriguez, J., Abruzzi, K.C., Khodor, Y.L., Sugino, K., Marr, M.T., 2nd, Nelson, S. and Rosbash, M. (2009) Genome-wide identification of targets of the drosha-pasha/DGCR8 complex. *Rna*, **15**, 537-545.
44. Bruneau, B.G. and Black, B.L. (2007) The heart's Da Vinci code: a Renaissance at Keystone. *Development*, **134**, 1631-1633.

****