OPEN ACCESS Freely available online

PLoS one

The Fate of miRNA* Strand through Evolutionary Analysis: Implication for Degradation As Merely Carrier Strand or Potential Regulatory Molecule?

Li Guo[1], Zhuhong Lu[1,2]*

1 State Key Laboratory of Bioelectronics, School of Biological Science and Medical Engineering, Southeast University, Nanjing, China, 2 Key Laboratory of Child Development and Learning Science of Ministry of Education, Southeast University, Nanjing, China

Abstract

Background: During typical microRNA (miRNA) biogenesis, one strand of a ~22 nt RNA duplex is preferentially selected for entry into a silencing complex, whereas the other strand, known as the passenger strand or miRNA* strand, is degraded. Recently, some miRNA* sequences were reported as guide miRNAs with abundant expression. Here, we intended to discover evolutionary implication of the fate of miRNA* strand by analyzing miRNA/miRNA* sequences across vertebrates.

Principal Findings: Mature miRNAs based on gene families were well conserved especially for their seed sequences across vertebrates, while their passenger strands always showed various divergence patterns. The divergence mainly resulted from divergence of different animal species, homologous miRNA genes and multicopy miRNA hairpin precursors. Some miRNA* sequences were phylogenetically conserved in seed and anchor sequences similar to mature miRNAs, while others revealed high levels of nucleotide divergence despite some of their partners were highly conserved. Most of those miRNA precursors that could generate abundant miRNAs from both strands always were well conserved in sequences of miR-#-5p and miR-#-3p, especially for their seed sequences.

Conclusions: The final fate of miRNA* strand, either degraded as merely carrier strand or expressed abundantly as potential functional guide miRNA, may be destined across evolution. Well-conserved miRNA* strands, particularly conservation in seed sequences, maybe afford potential opportunities for contributing to regulation network. The study will broaden our understanding of potential functional miRNA* species.

Citation: 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): e11387. doi:10.1371/journal.pone.0011387

Editor: Jason E. Stajich, University of California Riverside, United States of America

Received March 15, 2010; **Accepted** June 8, 2010; **Published** June 30, 2010

Copyright: © 2010 Guo et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The work is supported by the project 30871393 from National Natural Science Foundation of China and funded by Tsinghua National Laboratory for Information Science and Technology (TNList) Cross-discipline Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: zhlu@seu.edu.cn

Introduction

MicroRNAs (miRNAs) are an abundant class of small non-protein-coding RNAs that have emerged as key post-transcriptional regulators of gene expression in animals and plants [1,2]. Metazoan miRNA genes are transcribed by either RNA polymerase II or RNA polymerase III into primary miRNA transcripts (pri-miRNAs) as single genes or in clusters [1,3,4,5]. The pri-miRNAs contain stem-loop structures (hairpins) that harbor the miRNAs in the 5' or 3' half of the stem. These primary miRNA gene transcripts are typically, but not always, recognized and cut by the endonuclease Drosha in the cell nucleus to produce miRNA hairpin precursors that are then exported to the cytosol, where the hairpin structures are cut by the endonuclease Dicer at relatively fixed positions and released as short double-stranded RNA duplexes [6,7,8,9,10,11,12]. Although both strands of duplexes are necessarily produced in equal amounts by transcription, their accumulation is asymmetric at steady state [13]. Based on the thermodynamic stability of each end of this duplex, one of the strands is thought to be a biologically active miRNA, and the other is considered as an inactive strand and a carrier strand called miRNA* (miRNA star) or passenger strand [14]. Generally, the miRNA* strand is typically degraded, whereas the mature miRNA strand is taken up into the microribonucleoprotein complex (miRNP) [6] (Figure 1A and Figure 1B). The mature miRNA strand is used as a guide to direct negatively post-transcriptional regulation by the binding of 5'-seed (nucleotides 2–8) and anchor (nucleotides 13–16) with target sequences in the 3' untranslated region (UTR) of cognate mRNAs [1,15]. Once bound to Ago proteins, miRNAs are more stable than average mRNAs and the half-life of most miRNAs is greater than 14 hours [16]. They may be produced by 5' (left arms) or 3' arms (right arms) of the miRNA precursors, and the nonrandom nature of miRNA strand selection might reflect an active process that minimizes the population of silencing complexes with illegitimate miRNA* species [13] (Figure 1). The mechanism of strand selection maybe correlates with the relative free energies of the duplex ends [11,13,17].

However, recently, some miRNA* sequences were reported as mature functional miRNAs with abundant expression, and miRNA/miRNA* ratios may vary dramatically among develop-

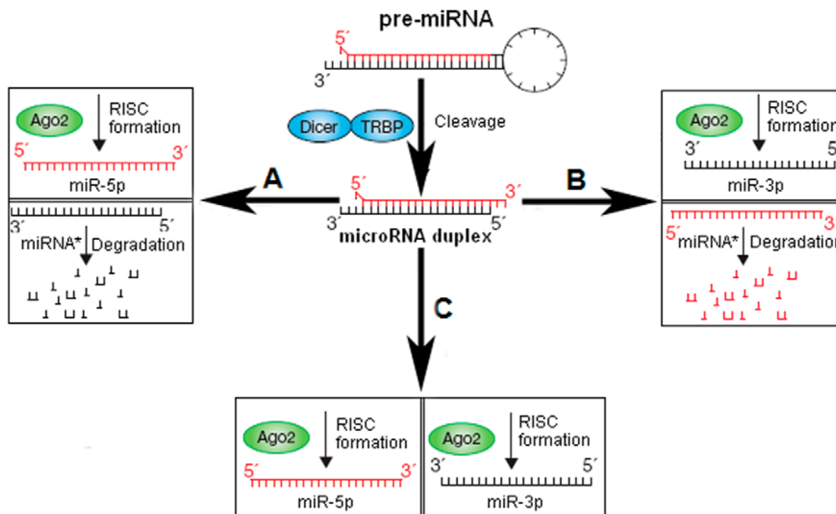


Figure 1. The selection of mature miRNA strand and the fate of its passenger strand. (A) The mature miRNA sequence is miR-#-5p, and its passenger sequence is miRNA* (star) which is degraded. (B) The mature miRNA sequence is miR-#-3p, and its passenger sequence is miRNA*. (C) Both the miR-#-5p and miR-#-3p are mature miRNA sequences that guide RISC to silence target mRNAs through mRNA cleavage, translational repression or deadenylation.
doi:10.1371/journal.pone.0011387.g001

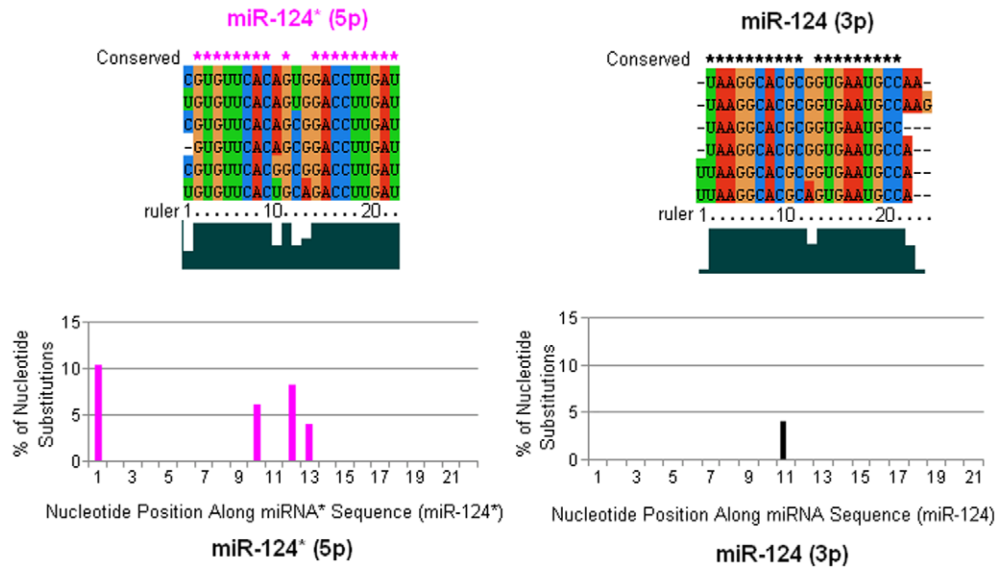
mental stages [13,18,19]. Most *Drosophila* miRNAs are bound to Ago1, and miRNA* strands accumulate bound to Ago2 [20]. The rarer partner of the mature miRNA, has been recognized both in terms of increasing the complexity of regulatory networks and in governing miRNA and messenger RNA evolution [13,19,21,22,23]. Some hairpins produce miRNAs from both strands at comparable frequencies because strand selection is often not a stringent process [24]. These abundant miRNA* species are often present at physiologically relevant levels and can associate with Argonaute proteins [13] (Figure 1). During Drosophilid evolution, more than 40% miRNA* sequences resist nucleotide divergence, and at least half of these well-conserved miRNA* species select for conserved 3' untranslated region seed matches well above background noise [13]. The miRNA* species diverge much more slowly than miRNA terminal loops, and conserved miRNA* sequences are almost perfectly conserved in sequences similar to mature miRNA sequences [13,22]. According to miRBase database (version 14.0, http://www.mirbase.org/), about 80 kinds of human miRNA precursors can yield two kinds of abundant mature miRNAs (left-arm, miR-#-5p; right-arm, miR-#-3p) with different seed sequences and target mRNAs, while most miRNA precursors only yield abundant mature miRNAs from left-arms or right-arms and rare miRNA* sequences. Most miRNA* species still keep fewer sequence counts despite of their mature miRNA sequences are detected higher expression levels based on high-throughput method [25]. Therefore, those specific miRNA precursors that yield two kinds of abundant functional miRNAs from different arms maybe reflect evolutionary implication across miRNA gene evolution. Although evolutionary patterns of miRNA* are consistent with their regulatory potential across Drosophilid evolution [13], limited knowledge about evolutionary information of miRNA/miRNA* has been discussed especially across different animal species.

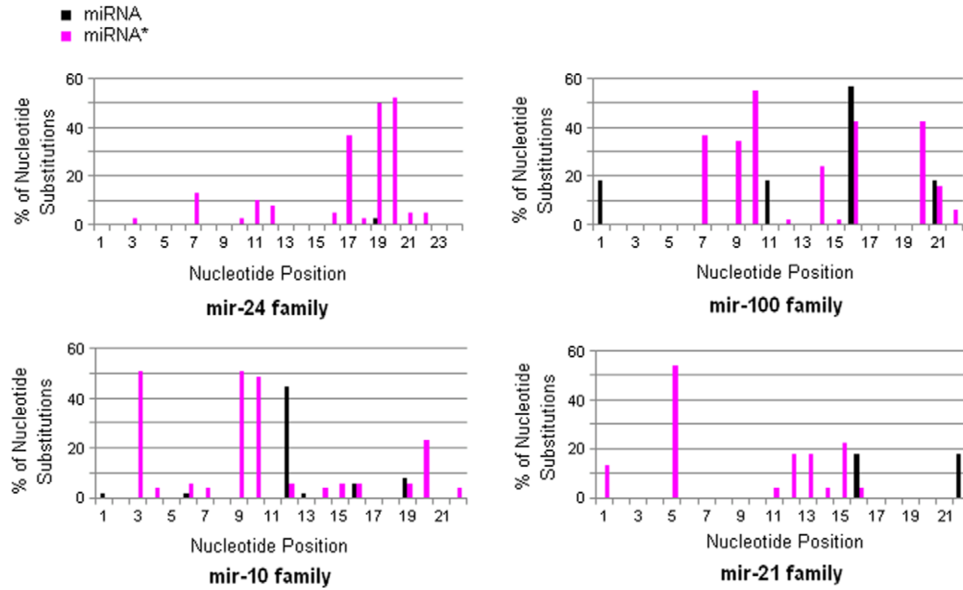
miRNAs are evolutionary conserved across broad phylogenetic distances [26,27,28], and they have gained considerable attention about evolution, genetic and phylogenetic analysis [15,23,29,30,31,32,33]. The non-coding small RNAs are strongly conserved in primary sequence and rarely secondarily lost once integrated into a gene regulatory network [22,31,34]. Recent study suggested an explosive increase in the miRNA repertoire in vertebrates [35]. 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*. The miRNA* strand with lower expression level because of degradation, or as functional mature miRNA with abundant clones, maybe get evolutionary implication by analyzing their evolutionary patterns. In the study, we intended to discover potential relationship between evolutionary pattern and selection of mature miRNA by analyzing miRNA/miRNA* based on miRNA gene families and single miRNA gene across vertebrates. Simultaneously, we also analyzed a complex miRNA gene family from a single animal species to study divergence trends of miRNA/miRNA* and discover potential evolutionary implication across evolution. Finally, because different miRNAs showed different distribution spectrums and evolutionary patterns across vertebrates, evolutionary analysis of miRNA/miRNA* based on single miRNA gene was performed across the same kinds of typical animals.

Results

Divergence patterns of miRNA/miRNA* based on miRNA gene families

Mature miRNAs always were highly conserved across vertebrates, especially in seed sequences (nucleotides 2–8) and anchor sequences (nucleotides 13–16), while their passenger strands showed higher nucleotide divergence (Figure 2). Some miRNA* species were well conserved across vertebrates although they showed a higher level of nucleotide divergence than their partners (Figure 2A and Figure 2B). The divergence mainly resulted from divergence of different animal species, homologous miRNA genes and multicopy hairpin precursors. For example, in complex miRNA gene families, such as let-7 family, miRNA* sequences


A
miR-124* (5p)
miR-124 (3p)
Conserved
ruler
% of Nucleotide Substitutions
Nucleotide Position Along miRNA* Sequence (miR-124*)
Nucleotide Position Along miRNA Sequence (miR-124)
miR-124* (5p)
miR-124 (3p)


B
miRNA
miRNA*
% of Nucleotide Substitutions
Nucleotide Position
mir-24 family
mir-100 family
mir-10 family
mir-21 family

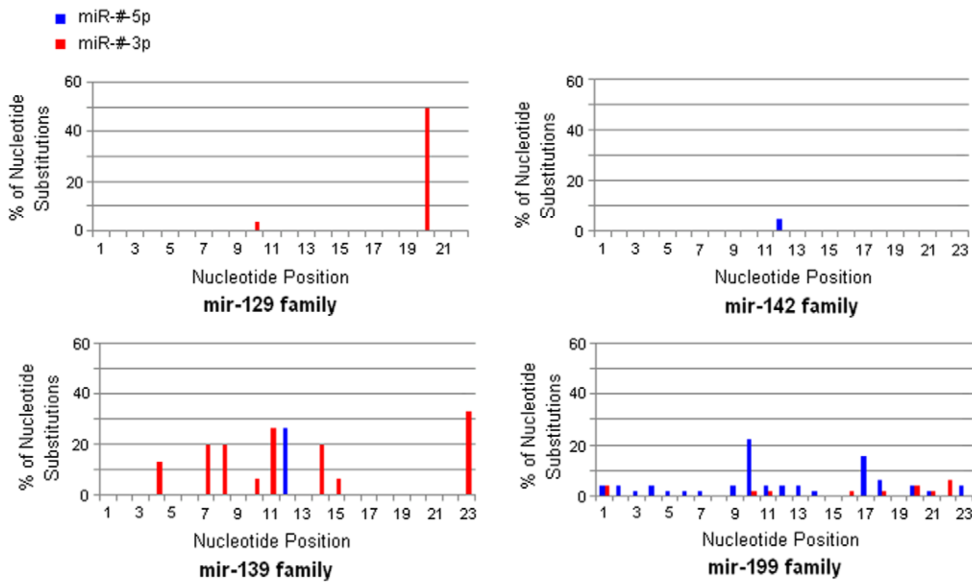

C
miR-#-5p
miR-#-3p
% of Nucleotide Substitutions
Nucleotide Position
mir-129 family
mir-142 family
mir-139 family
mir-199 family

Figure 2. Clustal multiple sequence alignment and mutational profile of miRNA and miRNA* sequences across vertebrates. (A) Multiple sequence alignment of mir-124 family and mutational profile of miRNA and miRNA* sequences. Repeated sequences were discarded. Conserved: conserved positions in sequence alignment. The bottom figures were the percentage of nucleotide substitutions from positions 1 to ~22 in miRNA and miRNA* sequences. Both miR-124 and miR-124* sequences were well conserved, especially miR-124. (B) Mutational profile of miRNA/miRNA* in miRNA gene families. Both miRNA and miRNA* sequences showed different evolutionary patterns. miRNA* sequences always were involved more nucleotide substitutions than miRNAs, even in positions 2–8. (C) Mutational profile of miR-#-5p/miR-#-3p in miRNA gene families. These miRNAs (miR-#-5p and miR-#-3p) were reported as abundant miRNAs according to miRBase database (version 14.0). Although miR-#-5p had different divergence pattern with miR-#-3p, both of them always were well conserved especially for the seed sequences. In mir-199 family, nucleotide divergences of miR-199-5p mainly resulted from gga-mir-199b and bta-mir-199c.
doi:10.1371/journal.pone.0011387.g002

were less conserved because of wide distribution spectrum in vertebrates, multiple homologous genes and multicopy precursors. Different miRNA* sequences showed different levels of nucleotide divergence. miR-124* sequences were well conserved, while miR-100* and miR-10* sequences showed greater nucleotide divergences than their miRNAs (Figure 2A and Figure 2B). Even in positions 2–8, some miRNA* sequences were involved nucleotide substitutions. According to miRBase database, some miRNA precursors were reported that they could generate two kinds of abundant miRNAs (miR-#-5p and miR-#-3p). Intriguingly, despite involved homologous genes and multicopy precursors (such as mir-142 family and mir-129 family), many of these miR-#-5p/miR-#-3p sequences were well conserved (Figure 2C). However, miR-#-5p and miR-#-3p showed different levels of nucleotide divergence despite both of them always had conserved seed sequences.

Although miRNA gene families may be involved complex evolutionary history across the animal kingdom and in a single animal species, miRNA/miRNA* based on a single animal species might show different levels of nucleotide divergence and imply different fates. Here, we took an example of let-7 family in *Homo sapiens*, which included several homologous members (Figure 3A). Some of these members could be found to have multicopy precursors, for example, hsa-let-7a could be produced by hsa-let-7a-1, hsa-let-7a-2 and hsa-let-7a-3. Mature hsa-let-7 sequences were produced by 5p (left arms) and well conserved especially for the seed sequences (nucleotides 2–8) and anchor sequences (nucleotides 13–16), while hsa-let-7* showed a higher level of nucleotide divergence even positions 2–8 (Figure 3A). These multicopy precursors could yield the same mature miRNA sequences, but their loop sequences and miRNA* strands might show greater divergence than miRNAs (Figure 3A). Interestingly, the similar trend of nucleotide divergence of miRNA and miRNA* could be detected across vertebrates (Figure 3B). Phylogenetic network of hsa-let-7 family was split into several clades based on different miRNA genes (Figure 4). Multicopy precursors for a single miRNA, such as hsa-let-7a-1, hsa-let-7a-2 and hsa-let-7a-3, might be reconstructed in different clusters (Figure 4).

Divergence patterns of miRNA/miRNA* based on single miRNA gene

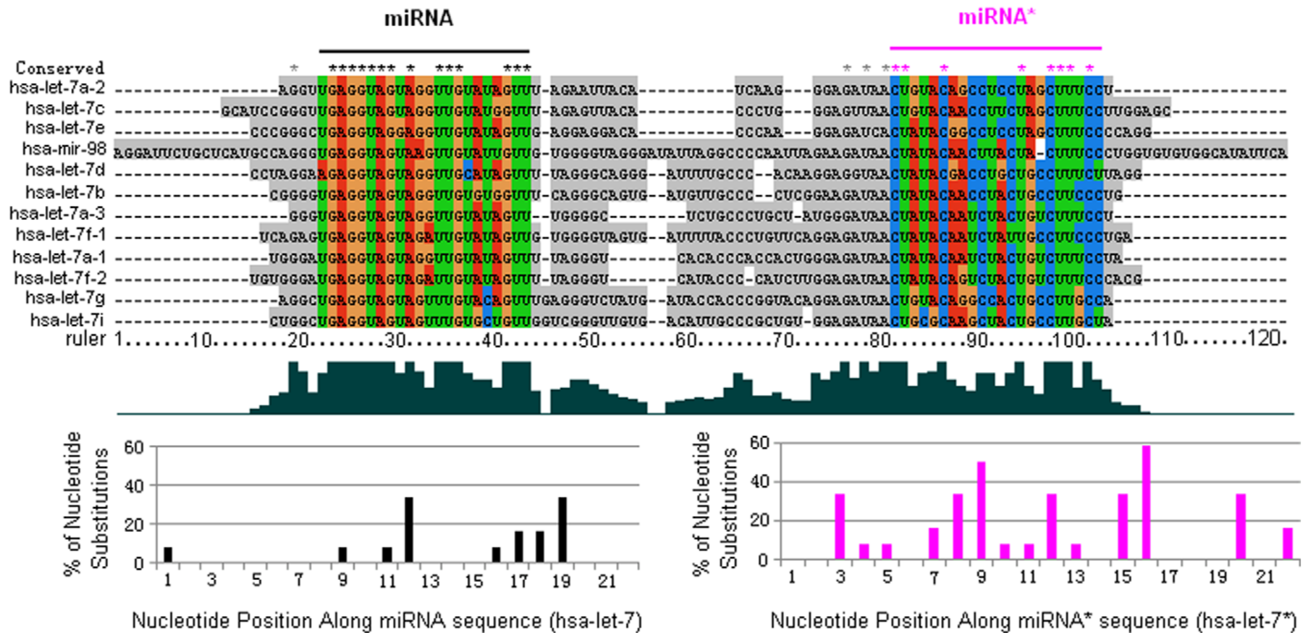
We observed different amounts of nucleotide divergence between miRNA and miRNA* sequences based on single miRNA gene, such as miR-125a-5p/miR-125a-3p and miR-210/miR-210* (Figure S1). Generally, more sites of miRNA* were involved divergence despite miRNAs were highly conserved (Figure S1). Different levels of divergence pattern also were detected in mammalian-specific miRNAs. Similarly, loop sequences showed different levels of divergence between various miRNA genes (Figure S1). According to human miRNAs in miRBase database, 80 kinds of miRNA precursors were reported two kinds of abundant miRNAs (miR-#-5p and miR-#-3p). Sequence analysis based on miRNA precursor sequences revealed that >80% of these miR-#-5p and miR-#-3p sequences ensured conserved seed sequences throughout evolution.

Because miRNAs always showed different distribution spectrums across the animal kingdom, we also analyzed miRNA and miRNA* across several kinds of typical vertebrate animals: *Danio rerio* (Pisces), *Homo sapiens* (Mammalia), *Gallus gallus* (Aves) and *Xenopus tropicalis* (Amphibia). Mature miRNAs were highly conserved across these animal species, while their passenger strands showed different evolutionary patterns (Figure 5). Some miRNA* sequences were less conserved even in positions 2–8, such as miR-31*, miR-100* and miR-125b*, while their terminus regions (5' and 3') were more conserved than their central regions. Other miRNA* sequences were well conserved similar to mature miRNAs, such as miR-18a*, miR-18b*, miR-17-3p and miR-455-3p (Figure 5). Although mature miR-100 and miR-125b were highly conserved, their star sequences showed greater divergence across species. Some well-conserved miRNA* species were reported as functional guide miRNAs with abundant expression, which were well conserved particularly in seed and anchor sequences (Figure 5C). Homologous miRNA genes maybe showed different divergence patterns in the same kinds of animals, such as miR-18a* and miR-18b* (Figure 5B). The loop sequences also showed different divergence trends although they maybe showed greater divergence than miRNA and miRNA*(Figure 5).

Discussion

Mature miRNAs (miR-#-5p or miR-#-3p) were evolutionarily conserved across the animal kingdom [26,27,28], while their passenger strands, either as typically degraded miRNA* or abundant mature miRNAs, always showed conservation across vertebrates with higher nucleotide divergences than their partners (Figure 2). Different miRNA* sequences showed various divergence patterns. Data analysis revealed that some mature miRNAs and their passenger strands were well-conserved, especially in their seed sequences (Figure 2). For example, miR-124 (Figure 2A), a phylogenetic conserved miRNA from *Caenorhabditis* to *Homo*, is one of the most abundantly expressed miRNAs in the nervous system and contributes to the development of nervous system [36,37,38]. However, some miRNA* sequences showed higher level of nucleotide divergence, even though in their positions 2–8, such as miR-10* and miR-100* (Figure 2B). Even if multicopy hairpin precursors could yield the same mature miRNA sequences, another product, termed as miRNA* species, always diverged, such as hsa-let-7a-1* and hsa-let-7a-2*, hsa-let-7f-1* and hsa-let-7f-2* (Figure 3A). Despite of greater divergence than miRNAs, we also found miRNA* diverged much more slowly than terminal loops, which maybe strongly aid the identification of functional animal miRNA hairpins as ''saddle'' structures [39,40]. Interestingly, similar divergence trends of human let-7/let-7* could be detected by sequence analysis across vertebrates (Figure 3), which might reveal historical miRNA gene divergence and similar evolutionary trend across different animals. High divergence levels could be detected among these homologous miRNA genes

A — hsa-let-7 family



B — let-7 family

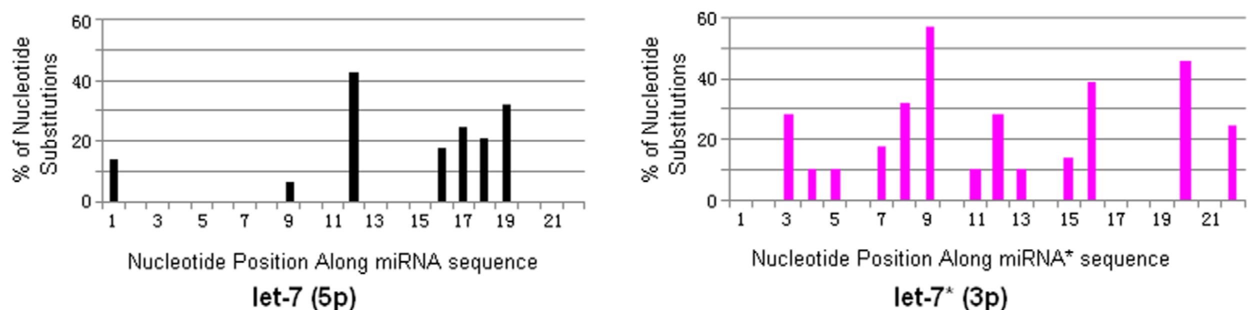


Figure 3. Clustal multiple sequence alignment of let-7 family and mutational profile of miRNA and miRNA* sequences. (A) Hsa-let-7 showed high conservation in positions 2-8 (the seed sequence) and relatively conservation in positions 13–16, while miRNA* sequence showed frequent substitutions and insertions/deletions. (B) Let-7 sequences across vertebrates showed high conservation in positions 2–8 (the seed sequence), while let-7* sequences showed frequent substitutions.
doi:10.1371/journal.pone.0011387.g003

(Figure 3 and Figure 4). Although the divergence mainly resulted from the loop regions, the divergence of miRNA* strands also contributed partly to the high divergence level (Figure 3A and Figure 4). On the other hand, we selected several typical vertebrate animals to analyze miRNA/miRNA* sequences because different miRNAs had different distribution spectrums across the animal kingdom. Similarly, some miRNA* strands were highly conserved, but others were less conserved despite their mature miRNAs were well conserved (Figure 5). Those miRNAs that reported both miR-#-5p and miR-#-3p could be mature functional miRNAs always were well conserved especially for seed and anchor sequences, such as miR-17, miR-140 and miR-455 (Figure 5). Nevertheless, some miRNA* strands were diverged even in their seed sequences, such as miR-31*, miR-100* and miR-125b* (Figure 5). Therefore, across miRNA gene evolution, functional mature miRNAs still were well conserved especially for their seed sequences, while miRNA* sequences showed various evolutionary patterns. Some miRNA* maybe showed high divergence levels between different precursors even between different multicopy precursors, but others ensured well-conserved seed sequences, especially for those miRNA genes that generated abundant miRNAs from two arms of hairpins (Figure 2, Figure 5 and Figure S1). Evolutionary conservation of passenger strand might result from two plausible reasons. Firstly, evolutionary process would be influenced because it maybe contributed to stable stem-loop structure of miRNA hairpin precursor. Secondly, the well conservation of passenger strand might afford an opportunity to be mature miRNA to bind target mRNA similar to its partner. Therefore, the evolutionary patterns of miRNA* might be a pivotal implication (discussed below).

According to miRNA biogenesis, as miRNA partners, the miRNA passenger strands should be more tightly constrained at their 3' ends which pair with the miRNA seed sequences (nucleotides 2–8). However, similar to Okamura et al. [13], systematic analysis showed that some miRNA* sequences were notably analogous to miRNA strands: well conserved in seed

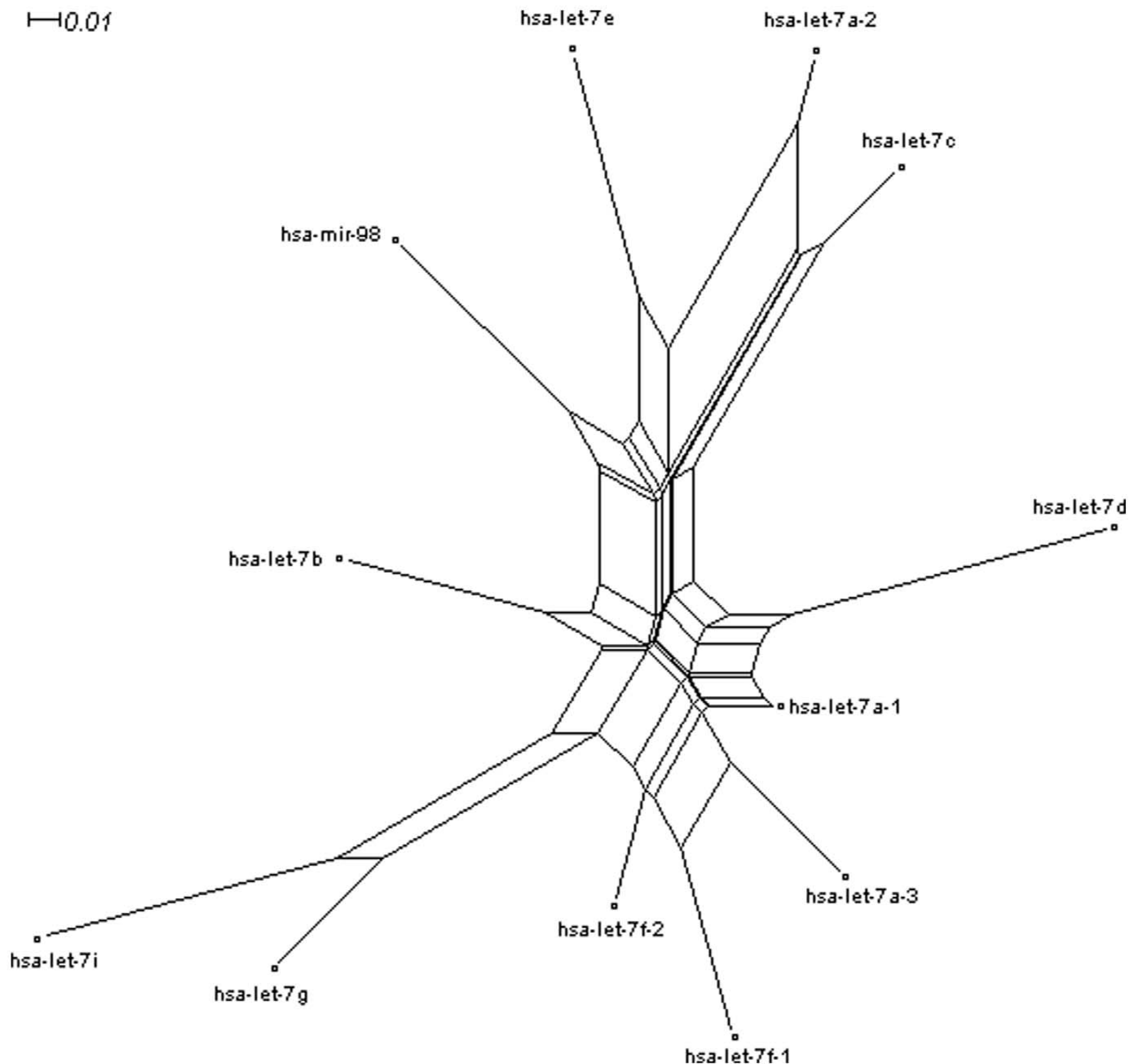


Figure 4. Phylogenetic network of hsa-let-7 family using neighbor-net method. Based on miRNA precursor sequences of let-7 family in *Homo sapiens*, 12 miRNA genes were split into several clusters. Multicopy miRNA precursors for a single miRNA, for example, hsa-let-7a-1, hsa-let-7a-2 and hsa-let-7a-3, were reconstructed in different clusters.
doi:10.1371/journal.pone.0011387.g004

(nucleotides 2–8) and anchor sequences (nucleotides 13–16) (Figure 2, Figure 5 and Figure S1). They also showed patterns of nucleotide divergence that were consistent with their selection for regulatory activity [13]. Therefore, the evolutionary pattern of miRNA* afforded an opportunity to be abundant functional guide miRNAs based on well conserved seed sequences that reflected their sequence-based, trans-regulatory activity [13,41]. Indeed, earlier computational efforts for miRNA genes finding hinted the possibility of trans-acting activity for miRNA* species [40]. Some miRNA* strands were abundant because they were degraded more slowly than others, and the miRNA:miRNA* ratio of many loci became increasingly skewed as development proceeded [13]. According to miRBase database, we found two kinds of mature products (miR-#-5p from left-arm and miR-#-3p from right-arm) were reported from some miRNA precursors, such as mir-199a and mir-17. Analysis of miRNA based on high-throughput sequencing data also showed abundant miRNA* although less abundant than their partners [25]. Recent study revealed that miRNA:miRNA* ratios were flexible in different development stages and both of them resisted nucleotide divergence across Drosophilid evolution [13]. The expression level of miRNA passenger strand mainly relied on degradation degree and degradation rate, because both strands of miRNA duplex were necessarily produced in equal amounts by transcription. We found different miRNA* showed various divergence patterns despite their mature miRNAs were highly conserved (Figure 2, Figure 5 and Figure S1). Generally, those less conserved miRNA* strands were not reported as mature functional miRNAs. The divergence

A

```
hsa-mir-31      GGAGAGGAGGCAAGAUGCUGGCAUAGCUGUUGAA--CUGGGAACCUGCUAUGCCAACAUAUUGCCAUCUUUC
gga-mir-31      GGAGGGGAGGCAAGAUGUUGGCAUAGCUGUUAAC--CUAAAAACCUGCUAUGCCAACAUAUUGUCAUCUUUC
xtr-mir-31      AGAGAGGAGGCAAGAUGUUGGCAUAGCUGUUGCA--UCUGAAACCAGUUGUGCCAACCUAUUGCCAUCUUUC
dre-mir-31      GAAGAGAUGGCAAGAUGUUGGCAUAGCUGUUAAUGUUUAUGGGCCUGCUAUGCCUCCAUAUUGCCAUUUCUG

hsa-mir-100     GUUGCCACAAACCCGUAGAUCCGAACUUGUGGUAU-UAGUCCGCACAAGCUUGUAUCUAUAGGUAUGUGUCUGUU
gga-mir-100     GUUGCCACAAACCCGUAGAUCCGAACUUGUGGUCA-UAUUCCACACAAGCUUGUAUCUAUAGGUAUGUGUCUGUC
dre-mir-100-1   GCUGACACAAACCCGUAGAUCCGAACUUGUGGUGA-CUGUCCACACAAGCUUGUAUCUAUAGGUAUCUGUCUGUG
dre-mir-100-2   GCUGCCACAAACCCGUAGAUCCGAACUUGUGGUGU-CUCUGUGCACAAGCUCGUGUCUAUAGGUAUGUGUCUGCG
xtr-mir-100     GUUGCCAUAAACCCGUAGAUCCGAACUUGUGCUGUGCCCCUCUCACAAGCUCGAGUGUGCGGGUCUGUGUCGGCU

hsa-mir-125b-1  CGCUCCUCUCAGUCCCUGAGACCCUAACUUGUGAUGUUUACCGUUUAAA-UCCACGGGUUAGGCUCUUGGGAGCUGCGAGUCGUG
xtr-mir-125b-1  CACCCCUCUCAAUCCCUGAGACCCUAACUUGUGAUGUUUAGCUUUAAAAAUCCACGGGUUAGGCUCUUGGGAGCUGUGAGUUGUG
dre-mir-125b-1  UGCCUCUCACAAUCCCUGAGACCCUAACUUGUGACGUUUUCCUGUUAUG-UGCACGGGUUAGGUUCUUGGGAGCUGAGAGGGGUG
dre-mir-125b-2  UGCCCCUCUCCUUCCCUGAGACCCUAACUUGUGACGUUCUGCUUCGAUG-UCCACGGGUUGGGUUCUCGGGAGCUGUGAGAGGCA
gga-mir-125b    GACUUUUCCUAGUCCCUGAGACCCUAACUUGUGAGG-UUUUGUAGCAACAAUCACAAGUCAGGCUCUUGGGACCUAGGCGGAGGG
hsa-mir-125b-2  GACUUUUCCUAGUCCCUGAGACCCUAACUUGUGAGG-UAUUUUAGUAACA-UCACAAGUCAGGCUCUUGGGACCUAGGCGGAGGG
xtr-mir-125b-2  GACUUUUCCUAGUCCCUGAGACCCUAACUUGUGAGGAUUUUUUAGCAACAAUCACAAGUUAGGCUCUUGGGACCUAGGCGGAGGG
dre-mir-125b-3  CACUCCUCCUGGUCCCUGAGACCCUAACUUGUGAGCUUUGUGUGCUAAAAAUCACAGGUUAAGCUCUUGGGACCUGGGCAGAGGG
```

B

```
gga-mir-122-1   GAGCUGUGGAGUGUGACAAUGGUGUUUGUGUC-CAAUCUAUCAAACGCCAUUAUCACACUAAAUAGCUAC
gga-mir-122-2   GAGCUAUGGAGUGUGACAAUGGUGUUUGUGUC-CAAUCUAUCAAACGCCAUUAUCACACUAAAUAGCUAC
hsa-mir-122     GAGCUGUGGAGUGUGACAAUGGUGUUUGUGUC-UAAACUAUCAAACGCCAUUAUCACACUAAAUAGCUAC
xtr-mir-122     GAGCUAUGGAGUGUGACAAUGGUGUUUGUGUC-AGAGCUAUCAAACGCCAUUAUCACACUAAUGAGCUAC
dre-mir-122     GAGCUGUGGAGUGUGACAAUGGUGUUUGUAUCAUCUGUCGUCAAACGCCAUUAUCACACUAAAUAGCCAC

xtr-mir-18a     UGUCCUAAGGUGCAUCUAGUGCAGAUAGUGAAGUAGAUUAGCAUCUACUGCCCUAAGUGCUCCUUCUGGCA
gga-mir-18a     UGUACUAAGGUGCAUCUAGUGCAGAUAGUGAAGUAGAUUAGCAUCUACUGCCCUAAGUGCUCCUUCUGGCA
hsa-mir-18a     UGUUCUAAGGUGCAUCUAGUGCAGAUAGUGAAGUAGAUUAGCAUCUACUGCCCUAAGUGCUCCUUCUGGCA
dre-mir-18a     UGUGCUAAGGUGCAUCUAGUGCAGAUAGUGAAGUAGACUAGCACCUACUGCCCUAAGUGCUCCUUCUGGCA

xtr-mir-18b     UGUGUUAAGGUGCAUCUAGUGCAGUUAGUGACAUAGUGUAGCAUCUACUGCCCUAAAUGCUCCUUUUGGCA
gga-mir-18b     UGUGUUAAGGUGCAUCUAGUGCAGUUAGUGACGUAGCGUAGAAUCUACUGCCCUAAAUGCUCCUUCUGGCA
hsa-mir-18b     UGUGUUAAGGUGCAUCUAGUGCAGUUAGUGAAGCAGCUUAGAAUCUACUGCCCUAAAUGCCCCUUCUGGCA
dre-mir-18b     UAUGCUAAGGUGCAUUUAGUGCAGAUAGUGAAAUAGACUAGCACCUACUGCCCUAAGUGCCCCUUCUGGCA
```

C

```
hsa-mir-17      GUCA-GAAUAAUGUCAAAGUGCUUACAGUGCAGGUAGUGAUAUG-U-GCAUCUACUGCAGUGAAGGCACUUGUAGCAUUAU-GGUGAC
gga-mir-17      GUCA-GAGUAAUGUCAAAGUGCUUACAGUGCAGGUAGUGAUAUA-UAGAACCUACUGCAGUGAAGGCACUUGUAGCAUUAU-GUUGAC
xtr-mir-17      GUCAAGAGUAAUGUCAAAGUGCUUACAGUGCAGGUAGUGAUUUAACAGAACCUACUGCAGUGAAGGCACUUGUAGCAUUAU-AUUGAC
dre-mir-17a-1   GUCA-AUGUAUUGUCAAAGUGCUUACAGUGCAGGUAGUAUUAUG-GAAUAUCUACUGCAGUGGAGGCACUUCUAGCAAUACACUUGAC
dre-mir-17a-2   GUCA-CUGUAGUGUCAAAGUGCUUACAGUGCAGGUAGUUCAAUA-UAA--UCUACUGCAGUGGAGGCACUUCAAGCUUUAC-CGUGAC

gga-mir-455     UCCCUGGUGUGAGGGUAUGUGCCCUUGGACUACAUCGUGGAAGCCAGCACCAUGCAGUCCAUGGGCAUAUACACUUGCCUCAAGGU
dre-mir-455     UCCCUGGUGUGAGGGUAUGUGCCCUUGGACUACAUCGUGGAAGCCAGCACCAUGCAGUCCAUGGGCAUAUACACUUGCCUCAAGGC
xtr-mir-455     UCCCUGGCGUGAGGGUAUGUGCCCUUGGACUACAUCGUGGAAGCCAGCACCAUGCAGUCCAUGGGCAUAUACACUUGCCUCAAGGU
hsa-mir-455     UCCCUGGCGUGAGGGUAUGUGCCUUUGGACUACAUCGUGGAAGCCAGCACCAUGCAGUCCAUGGGCAUAUACACUUGCCUCAAGGC

dre-mir-140     UUGUCUCCUGUGUCCCGUCAGUGGUUUUACCCUAUGGUAGGUUACGUCAUGCUGUUCUACCACAGGGUAGAACCACGGACGGGAUGUCUGGAG
gga-mir-140     GUCUCUCC-GUGUCCUGCCAGUGGUUUUACCCUAUGGUAGGUUACGUCAUGCUGUUCUACCACAGGGUAGAACCACGGACGGGAUGCCGGGGC
xtr-mir-140     CCCUCUCU-GUGUCCUCCCAGUGGUUUUACCCUAUGGUAGGUUACGUCAGGCUGUUCUACCACAGGGUAGAACCACGGACAGGAUAC-GGGGA
hsa-mir-140     CUCUCUCU-GUGUCCUGCCAGUGGUUUUACCCUAUGGUAGGUUACGUCAUGCUGUUCUACCACAGGGUAGAACCACGGACAGGAUACCGGGGC
```

Figure 5. Nucleotide divergences of miRNA and miRNA* across several typical vertebrates. In order to avoid influence of various distribution spectrums of miRNAs, we selected several typical vertebrates to analyze miRNA evolutionary trends: dre (*Danio rerio*, Pisces), hsa (*Homo sapiens*, Mammalia), gga (*Gallus gallus*, Aves) and xtr (*Xenopus tropicalis*, Amphibia). Red conserved sites indicated mature miRNAs (including less abundant but functional miRNA* strands), pink conserved sites indicated miRNA* sequences that were not reported as functional guide miRNAs. Some flank sequences of miRNA sequences were deleted. All the mature miRNAs were highly conserved in these animals, while their miRNA* strands showed different divergence trends. (A) miRNA* sequences were less conserved due to divergence of different animals and multicopy miRNA precursors. (B) miRNA* sequences were well conserved similar to mature miRNAs though they were not reported as abundant functional miRNAs. (C) miR-#-5p and miR-#-3p were well conserved across different animals, and both of them were reported as abundant functional guide miRNAs.
doi:10.1371/journal.pone.0011387.g005

of less-conserved miRNA* always resulted from individual animal and/or multicopy precursors (Figure 5 and Figure S1). Evolutionary trends of the miRNA* strands might be potential implication for their final fates: degradation as by-products or functional regulatory molecules as mature miRNAs. It is plausible that non-functional miRNA* strands maybe involved higher rates of nucleotide substitution during evolution, while functional miRNA* sequences would be strictly regulated that were critical during binding target mRNAs. The correlation between the evolutionary constraint of miRNA* and their expression levels might reflect their potential function as endogenous regulatory RNAs. Some miRNA* strands might become functional guide strands and they were phylogenetically conserved similar to their mature miRNAs. Those well conserved miRNA* strands might also play important roles in regulating network in different development stages, but limited miRNA data cannot afford enough experimental evidences. Therefore, evolutionary patterns of many miRNA* strands were consistent with their regulatory potential [13,23], and the final fate, degradation as merely carrier strand or becoming potential functional guide miRNAs, might be got some implication throughout miRNA gene evolution. Some passenger strands were well-conserved in positions 2–8 similar to their mature miRNAs, and the phylogenetic conservation of miRNA* may be evolutionary implication to become abundant guide miRNAs and play important roles in particular developmental contexts at specific times. The systematic evolutionary analysis maybe broaden our understanding of miRNA* strands, especially for those potential regulatory miRNA* species.

Materials and Methods

All the miRNA and miRNA* sequences, and their miRNA precursor sequences from different animal species were obtained in miRBase database (version 14.0, http://www.mirbase.org/). We denoted the miRNA precursors by mir-#, the mature miRNAs by miR-#, and miRNA* (miRNA star) by miR-#* in accordance with the convention in miRBase database. If the miRNA* strands were reported as abundant mature miRNA, miR-#-5p or miR-#-3p was denoted. In the study, miR-#-5p and miR-#-3p were identified according to human miRNAs in miRBase database. These sequences were aligned with Clustal X 2.0 [42] by using the multiple sequence alignment. Phylogenetic network of miRNA genes was reconstructed using the neighbor-net method [43] based on Jukes-Cantor model as implemented in SplitsTree 4.10 [44]. For human let-7 family, we attempted to reconstruct the evolutionary history from the gene tree and discover potential evolutionary implications of let-7 and let-7*. All the gaps/missing data were deleted in phylogenetic network.

Because miRNA* sequences always degraded, there were limited miRNA* sequences in miRBase database. In order to discover detailed evolutionary information, we analyzed predicted consensus sequences as miRNA* sequences according to known miRNA* based on their precursor sequences. Because of imprecise and alternative cleavage of Dicer and Drosha, multiple isomiRs, the population of variants of known miRNAs, have been identified from the sequencing data by applying high-throughput DNA sequencing technologies [25,45,46,47,48]. Therefore, in the study, we only analyzed nucleotide substitutions of internal sequences of miRNA and miRNA* without considering gaps/missing sites in the terminus regions. Percentage of nucleotide substitution at positions (from 1 to ~22) was estimated for miRNAs and miRNA* sequences by analyzing all the miRNA precursors from miRBase database. In order to estimate substitution trend more precisely, we selected the most abundant nucleotide at each position as reference nucleotide.

Supporting Information

Figure S1 Patterns of nucleotide divergence of miRNA and miRNA* across vertebrates. (A) and (B) showed well conserved miR-#-5p and miR-#-3p based on miRNA gene family. (C) and (D) showed divergence patterns of miR-#-5p/miR-#-3p and miRNA/miRNA* based on single miRNA gene.
Found at: doi:10.1371/journal.pone.0011387.s001 (3.99 MB TIF)

Author Contributions

Conceived and designed the experiments: LG ZL. Analyzed the data: LG. Contributed reagents/materials/analysis tools: LG. Wrote the paper: LG.

References

1. Bartel DP (2004) MicroRNAs: Genomics, biogenesis, mechanism, and function. Cell 116: 281–297.
2. Plasterk RHA (2006) Micro RNAs in animal development. Cell 124: 877–881.
3. Borchert GM, Lanier W, Davidson BL (2006) RNA polymerase III transcribes human microRNAs. Nature Structural & Molecular Biology 13: 1097–1101.
4. Cai XZ, Hagedorn CH, Cullen BR (2004) Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. Rna-a Publication of the Rna Society 10: 1957–1966.
5. Lee Y, Kim M, Han JJ, Yeom KH, Lee S, et al. (2004) MicroRNA genes are transcribed by RNA polymerase II. Embo Journal 23: 4051–4060.
6. Filipowicz W, Bhattacharyya SN, Sonenberg N (2008) Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? Nature Reviews Genetics 9: 102–114.
7. Friedlander MR, Chen W, Adamidi C, Maaskola J, Einspanier R, et al. (2008) Discovering microRNAs from deep sequencing data using miRDeep. Nature Biotechnology 26: 407–415.
8. Grishok A, Pasquinelli AE, Conte D, Li N, Parrish S, et al. (2001) Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C-elegans developmental timing. Cell 106: 23–34.
9. Hutvagner G, McLachlan J, Pasquinelli AE, Balint E, Tuschl T, et al. (2001) A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Science 293: 834–838.
10. Lund E, Guttinger S, Calado A, Dahlberg JE, Kutay U (2004) Nuclear export of microRNA precursors. Science 303: 95–98.
11. Schwarz DS, Hutvagner G, Du T, Xu ZS, Aronin N, et al. (2003) Asymmetry in the assembly of the RNAi enzyme complex. Cell 115: 199–208.
12. Schwarz DS, Hutvagner G, Du TT, Xu ZS, Aronin N, et al. (2007) Asymmetry in the assembly of the RNAi enzyme complex (Reprinted from Cell, vol 115, pg 199-208, 2003). Cell 131: 30–40.
13. Okamura K, Phillips MD, Tyler DM, Duan H, Chou YT, et al. (2008) The regulatory activity of microRNA star species has substantial influence on microRNA and 3 ' UTR evolution. Nature Structural & Molecular Biology 15: 354–363.

14. O'Toole AS, Miller S, Haines N, Zink MC, Serra MJ (2006) Comprehensive thermodynamic analysis of 3 ' double-nucleotide overhangs neighboring Watson-Crick terminal base pairs. Nucleic Acids Research 34: 3338–3344.
15. Grimson A, Srivastava M, Fahey B, Woodcroft BJ, Chiang HR, et al. (2008) Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. Nature 455: 1193–U1115.
16. Hwang HW, Wentzel EA, Mendell JT (2007) A hexanucleotide element directs microRNA nuclear import. Science 315: 97–100.
17. Khvorova A, Reynolds A, Jayasena SD (2003) Functional siRNAs and rniRNAs exhibit strand bias. Cell 115: 209–216.
18. Ro S, Park C, Young D, Sanders KM, Yan W (2007) Tissue-dependent paired expression of miRNAs. Nucleic Acids Research 35: 5944–5953.
19. Jagadeeswaran G, Zheng Y, Sumathipala N, Jiang HB, Arrese EL, et al. (2010) Deep sequencing of small RNA libraries reveals dynamic regulation of conserved and novel microRNAs and microRNA-stars during silkworm development. Bmc Genomics 11: 52.
20. Ghildiyal M, Xu J, Seitz H, Weng ZP, Zamore PD (2010) Sorting of Drosophila small silencing RNAs partitions microRNA* strands into the RNA interference pathway. Rna-a Publication of the Rna Society 16: 43–56.
21. Jazdzewski K, Liyanarachchi S, Swierniak M, Pachucki J, Ringel MD, et al. (2009) Polymorphic mature microRNAs from passenger strand of pre-miR-146a contribute to thyroid cancer. Proceedings of the National Academy of Sciences of the United States of America 106: 1502–1505.
22. Wheeler BM, Heimberg AM, Moy VN, Sperling EA, Holstein TW, et al. (2009) The deep evolution of metazoan microRNAs. Evolution & Development 11: 50–68.
23. Liu N, Okamura K, Tyler DM, Phillips MD, Chung WJ, et al. (2008) The evolution and functional diversification of animal microRNA genes. Cell Research 18: 985–996.
24. Kim VN, Han J, Siomi MC (2009) Biogenesis of small RNAs in animals. Nature Reviews Molecular Cell Biology 10: 126–139.
25. Guo L, Lu Z (2010) Analysis of isomiR from high-throughput DNA sequencing data. in press. .
26. Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T (2001) Identification of novel genes coding for small expressed RNAs. Science 294: 853–858.
27. 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.
28. Lee RC, Ambros V (2001) An extensive class of small RNAs in Caenorhabditis elegans. Science 294: 862–864.
29. Chen K, Rajewsky N (2007) The evolution of gene regulation by transcription factors and microRNAs. Nature Reviews Genetics 8: 93–103.
30. 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: e7944.
31. Hertel J, Lindemeyer M, Missal K, Fried C, Tanzer A, et al. (2006) The expansion of the metazoan microRNA repertoire. Bmc Genomics 7: 25.
32. Niwa R, Slack FJ (2007) The evolution of animal microRNA function. Current Opinion in Genetics & Development 17: 145–150.
33. Sempere LF, Cole CN, McPeek MA, Peterson KJ (2006) The phylogenetic distribution of metazoan microRNAs: Insights into evolutionary complexity and constraint. Journal of Experimental Zoology Part B-Molecular and Developmental Evolution 306B: 575–588.
34. Heimberg AM, Sempere LF, Moy VN, Donoghue PCJ, Peterson KJ (2008) MicroRNAs and the advent of vertebrate morphological complexity. Proceedings of the National Academy of Sciences of the United States of America 105: 2946–2950.
35. Bompfunewerer AF, Flamm C, Fried C, Fritzsch G, Hofacker IL, et al. (2005) Evolutionary patterns of non-coding RNAs. Theory in Biosciences 123: 301–369.
36. Cheng LC, Pastrana E, Tavazoie M, Doetsch F (2009) miR-124 regulates adult neurogenesis in the subventricular zone stem cell niche. Nature Neuroscience 12: 399–408.
37. Nelson PT, Baldwin DA, Kloosterman WP, Kauppinen S, Plasterk RHA, et al. (2006) RAKE and LNA-ISH reveal microRNA expression and localization in archival human brain. Rna-a Publication of the Rna Society 12: 187–191.
38. Smirnova L, Grafe A, Seiler A, Schumacher S, Nitsch R, et al. (2005) Regulation of miRNA expression during neural cell specification. European Journal of Neuroscience 21: 1469–1477.
39. Berezikov E, Guryev V, van de Belt J, Wienholds E, Plasterk RHA, et al. (2005) Phylogenetic shadowing and computational identification of human microRNA genes. Cell 120: 21–24.
40. Lai EC, Tomancak P, Williams RW, Rubin GM (2003) Computational identification of Drosophila microRNA genes. Genome Biology 4: R42.
41. Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB (2003) Prediction of mammalian microRNA targets. Cell 115: 787–798.
42. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, et al. (2007) Clustal W and clustal X version 2.0. Bioinformatics 23: 2947–2948.
43. Bryant D, Moulton V (2004) Neighbor-Net: An agglomerative method for the construction of phylogenetic networks. Molecular Biology and Evolution 21: 255–265.
44. Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. Molecular biology and evolution 23: 254.
45. Kuchenbauer F, Morin RD, Argiropoulos B, Petriv OI, Griffith M, et al. (2008) In-depth characterization of the microRNA transcriptome in a leukemia progression model. Genome Research 18: 1787–1797.
46. Lagos-Quintana M, Rauhut R, Yalcin A, Meyer J, Lendeckel W, et al. (2002) Identification of tissue-specific microRNAs from mouse. Current Biology 12: 735–739.
47. Morin RD, O'Connor MD, Griffith M, Kuchenbauer F, Delaney A, et al. (2008) Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. Genome Research 18: 610–621.
48. Ruby JG, Jan C, Player C, Axtell MJ, Lee W, et al. (2006) Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C-elegans. Cell 127: 1193–1207.