

Expansion of the Human Micro RNA Gene Repertoire in the Human Lineage after the Divergence of Old World and New World Monkeys

Reiko F Kikuno*

Kazusa DNA Research Institute, 2-6-7, Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan

Abstract

It has been widely postulated that micro RNAs (miRNAs) have played an important role in achieving morphological complexity during the course of eumetazoan evolution, because the miRNA repertoire has continually increased along with greater complexity in the genomes of bilaterians and vertebrates. In particular, the miRNA repertoire is dramatically expanded in human. To examine how miRNAs changed during the process of primate evolution, we estimated the evolutionary origins of 1,527 present human miRNAs along the human lineage, after the divergence of eutherian mammals. We constructed an evolutionary profile of the miRNA genes, based on the presence and absence of the known and predicted functional orthologues in six non-human primate species: chimpanzee, gorilla, orangutan, gibbon, macaque, and marmoset and three representative eutherian mammals: bovine, dog and mouse. The functional orthologues were predicted by multiple genome sequence alignments and BLAST searches, followed by four filtering steps. The number of miRNA genes and gene families drastically expanded in the lineage of old world monkey, after the divergence from the common ancestor with new world monkey, at approximately seven-fold higher rates than the earlier stage of primate evolution. The rates of expansion of miRNA and gene families have been slightly declined in hominoids. Comparisons of the number of genes with that of the gene families acquired in each evolutionary period suggested that *de novo* gene generation, rather than gene duplication, considerably contributed to the generation of new miRNA genes.

Keywords: Human microRNA; Molecular evolution; Primates; Evolutionary origin; Comparative genomics; Sequence analysis

Introduction

MicroRNAs (miRNAs) are short endogenous non-coding RNAs that function as negative regulators of gene expression at the post-transcriptional level, in a broad range of cellular process in animals and plants [1-3]. In animal cells, the mature ~ 20 nt long miRNAs are derived from a 70-100-nt precursor (pre-miRNA), which forms a distinctive stem-loop structure. The seed region (nucleotides 2-8) of the mature miRNA is critical for its target recognition, through complementary base-pairing to the binding sites on the targets [4,5].

The acquisition of new miRNA genes in the genome is considered to correlate with dramatic increases in morphological complexity. In contrast, an expanded protein-coding repertoire and increased genome duplication events are not sufficiently related to differences in morphological complexity. For instance, although miRNA genes have been continually acquired in each eumetazoan lineage, other regulatory elements, such as transcription factors, do not show this propensity [6,7]. The increase in miRNA families is reportedly not metronomic in eumetazoan lineages, but occurs according to increased complexity [7]. To obtain clues to explain the possible association of miRNA with the many phenotypic changes that have occurred during primate evolution, we tried to identify the lineage-specific miRNAs, and to examine how the miRNA genes and families have increased along the lineage of primate evolution.

Recently, Iwama et al. [8] estimated the evolutionary period in which each of the present human miRNA genes originated, by partitioning the time period from human to the common ancestral branch between human and platypus into 15 portions. They then computed the rate of generation of human miRNA genes along vertebrate evolution, and proposed that two peak periods occurred when the human miRNA genes originated at significantly accelerated rates. They found homologues of human miRNA genes based on synteny information, using multiple and pair-wise genome sequence alignments. However, synteny

conservation is not the sole means to predict the correct orthologue, in general. In this study, we incorporated both synteny information and BLAST search [9] results to find homologues of human miRNA genes through each genome. Since BLAST fundamentally produces a local alignment to the query, we tried to obtain a global alignment by supplementary pair-wise alignments between the query and the local genomic sequence, including the BLAST hit region.

Since we are interested in the diversification process of miRNA genes in the primate lineage, we predicted the functional orthologues of human miRNA genes in the genomes of six non-human primates, including chimpanzee (*Pan troglodytes*), gorilla (*Gorilla gorilla*), orangutan (*Pongo pygmaeus*), gibbon (*Nomascus leucogenys*), macaque (*Macaca mulatta*) and marmoset (*Callithrix jacchus*), as well as three non-primate mammals, bovine (*Bos taurus*), dog (*Canis lupus familiaris*) and mouse (*Mus musculus*) as an outgroup. To distinguish the functional orthologues of human miRNAs from other homologues, we applied four different kinds of filters: identity of sequences of miRNA precursors (pre-miRNAs), conservation of the secondary structure, and conservation of the mature and seed sequences. We considered the evolutionary distance between human and each species to determine the thresholds of sequence and secondary structure conservation for screening. Based on the presence and absence of the known and predicted orthologues of human miRNA, we constructed

*Corresponding author: Reiko F Kikuno, Kazusa DNA Research Institute, 2-6-7, Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan, Tel: +81-438-52-3919; Fax: +81-438-52-3907; E-mail: kikuno@kazusa.or.jp

Received August 14, 2013; Accepted October 05, 2013; Published October 15, 2013

Citation: Kikuno RF (2013) Expansion of the Human Micro RNA Gene Repertoire in the Human Lineage after the Divergence of Old World and New World Monkeys. J Theor Comput Sci 1: 102. doi: [10.4172/2376-130X.1000102](https://doi.org/10.4172/2376-130X.1000102)

Copyright: © 2013 Kikuno RF. 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.

an evolutionary profile of human miRNAs in these species. We divided the stage of primate evolution into seven periods, by the divergences of species with genome sequences determined with high coverage. We then estimated the time when each human miRNA gene originated, and that when miRNA gene family to which the human miRNA belongs generated within the evolutionary period thus specified.

According to TimeTree [10], we presumed the following divergence times. Human and chimpanzee diverged 6.3 million years (Myr) ago; human and gorilla 8.8 Myr ago; human and orangutan 15.7 Myr ago; human and gibbon 20.4 Myr ago; hominoids and old world monkeys (such as macaque) about 29.0 Myr ago; old world monkeys and new world monkeys (such as *Callithrix jacchus*) 42.6 Myr ago; and primates and rodents 92.3 Myr ago. We estimated the acquisition rates of new human miRNA genes and gene families, based on these assumptions.

Approximately, 73% of the present human miRNA genes have originated within the simian lineage to human, and approximately 27% of the genes have originated within the hominoid lineage. At least 8% of the genes were supposed to be present only in the human genome. We also found that the number of miRNAs expanded significantly in the lineage of old world monkey, after the divergence from the common ancestor with new world monkey. The rate was accelerated approximately seven-fold, as compared to that estimated in the earlier stage, and it has been slightly declining during the course of hominoid evolution to date.

Materials and Methods

Detection of human pre-miRNA homologues in the other nine species

The known pre-miRNAs in human and seven other species (chimpanzee, gorilla, orangutan, macaque, dog, bovine and mouse) were fetched from the miRBase database (release 18.0) [11]. Data for gibbon and marmoset were not available there. Among 1,527 human miRNA genes, 875 were classified into 474 miRNA gene families, while the remaining 652 miRNA genes were not categorized. The seed sequence was defined as nucleotides 2-8 from the 5'-end of each mature miRNA sequence.

Multiple sequence alignments in the syntenic regions from six primate genomes (human, chimpanzee, gorilla, orangutan, macaque and marmoset) were selected from a synteny-oriented Ensembl Enredo-Pecan-Orteus (EPO) 6-way multiple alignment [12]. We excised the pieces of the genome alignments containing the human pre-miRNAs by our original Perl script, using Ensembl Compara API [13]. To avoid overlooking homologues in the regions where conserved synteny was not defined, we searched for human pre-miRNA homologues against the genome sequences of the six species, as well as four other species (gibbon, dog, bovine and mouse), using NCBI BLAST [9] version 2.2.24. All of the genome sequences were downloaded from the NCBI ftp site [14]. When the top hit sequence was aligned only locally by BLAST, we fetched the genomic sequence containing this sequence along with 100 nt of both its 5'- and 3'-flanking sequences. We then re-aligned this sequence with the query by ClustalW [15], to obtain the global sequence alignment. If we still did not obtain an alignment without terminal gaps, we did not deem this homologue to be an orthologue of the query.

We calculated the sequence difference corrected for multiple hits of substitutions (K^c) from aligned sequences, obtained by both

the genome alignment and BLAST search, using dnadist with Jukes-Cantor correction method [16] in the Phylip Package version 3.6 [17].

Secondary structure analysis

Secondary structure folding with maximum expected accuracy (MEA) of the pre-miRNA was calculated by RNAfold, in Vienna Package [18]. To measure the secondary structure conservation of the human pre-miRNA and the detected homologue, we used RNAforester [19], which calculates the structure alignment by considering both sequence and secondary structure similarities. We used the base pair substitution matrix RIBOSUM85-60 and obtained a structure alignment score (σ). Since the alignment score cannot be used directly to compare the similarity among the structures with different sequence lengths, we normalized the similarity score for each pre-miRNA homologue found, according to Wang et al. [20]. The normalized similarity score of structures C and h was defined as

$$str_sim(C, h) = \frac{\sigma(C, h)}{\sigma(h, h)} \times 100$$

where C is the candidate sequence in the non-human species, h is the known pre-miRNA sequence in human, $\sigma(C, h)$ represents the raw local alignment score of the two structures C and h, and $\sigma(h, h)$ is the self-alignment score of h. The normalized score, str_sim, ranges from 0 to 100, and a higher value reflects more similarity between the two structures.

Filtering

We adopted the following filtering processes to predict the functional orthologues of human pre-miRNAs, among the homologues found in the previous step.

Smirnov-Grubbs rejection test: To determine the thresholds to distinguish functional orthologues from other homologues, we applied the Smirnov-Grubbs rejection test to the distributions of sequence difference (K^c) and structure score (str_sim) values from pairs of known pre-miRNAs between human and each other species from miRBase. Many of the known pre-miRNAs were considered to be functional, because they have homologues in at least one of the other species. Conservation of sequence and structure depends on the strength of the functional constraint operating on each pre-miRNA and the divergence time. To exclude the effects of the divergence time on the extents of sequence identity and structure similarity, we examined the distributions of K^c and str_sim values, and determined the thresholds to define outliers at significance levels of $p < 0.01$ and $p < 0.05$, for each pair of human and other species. We excluded the homologues with lower sequence or secondary structure conservation values than the thresholds from the candidates of functional orthologues. The threshold obtained from the comparison between human and chimpanzee was used to select functional orthologues in gorilla, and that obtained from the comparison between human and macaque was used to select functional orthologues in gibbon and marmoset, because there were either none or only a limited number of known miRNAs for these species in miRBase. We used the threshold from the human-bovine comparison for that of human-dog. The thresholds obtained from distributions of K^c and str_sim values in human-dog comparison were too much strict (0.142 for K^c and 36.0 str_sim) than those from human-bovine comparison, although the divergence time of the two species pairs were supposed to be the same. We thought these were artifacts caused by the limited number of known miRNAs in dog.

We then applied the following two filters.

Mature sequence filter: No gaps in the mature sequence.

Seed filter: No mutations or gaps in the seed region (positions 2-8 from the 5' end of the mature sequence).

Results

Prediction of functional orthologues of human pre-miRNAs in other species

There were 1,527 human pre-miRNA sequences registered in miRBase (release 18.0) [11]. Using each human pre-miRNA sequence as a query, we searched for the homologues in the genomes of the six non-human primates by multiple genomic alignment and BLAST searches [9]. For three non-primate mammalian genomes, we searched by BLAST only, because synteny was less conserved in these species. Homologues were aligned with the query, based on the similarities of the sequences and the predicted secondary structures, respectively. While one or more homologues of 1,523 out of the 1,527 human pre-miRNAs were detected in at least one of the genomes of the nine species, no homologues were found for the four pre-miRNAs (hsa-mir-629, hsa-mir-3622a, hsa-mir-3622b and hsa-mir-3690), by either the genome sequence alignment or BLAST search. Figure 1 shows the

frequency distribution of the nucleotide difference (K^c) calculated from comparisons between the human pre-miRNAs and the homologues found in the other species. The distribution of K^c between pairs of known pre-miRNAs in human and other species fetched from miRBase were plotted in the same panel, except for gibbon and marmoset, since there are no miRNA sequence data for these species in miRBase. The threshold to determine outliers among the K^c values between pairs of human pre-miRNAs, estimated at a significance level of $p < 0.01$, is shown in each panel.

We also investigated the secondary structure conservation of the homologues. The frequency distributions of the str_sim values (see Methods), obtained from comparisons of human pre-miRNAs and the homologues in each species were plotted (Figure 2). The str_sim values were calculated from the predicted structure-based alignment, followed by normalization according to the alignment length. The threshold to exclude outliers among the str_sim values, estimated at a significance level of $p < 0.01$, is also shown in each panel. When a homologue was classified as an outlier by the tests on K^c or str_sim , it was excluded from the candidates of functional orthologues of human pre-miRNAs, and was not subjected to further filtering. For the remaining candidates of functional orthologues, we finally applied two additional filters, a

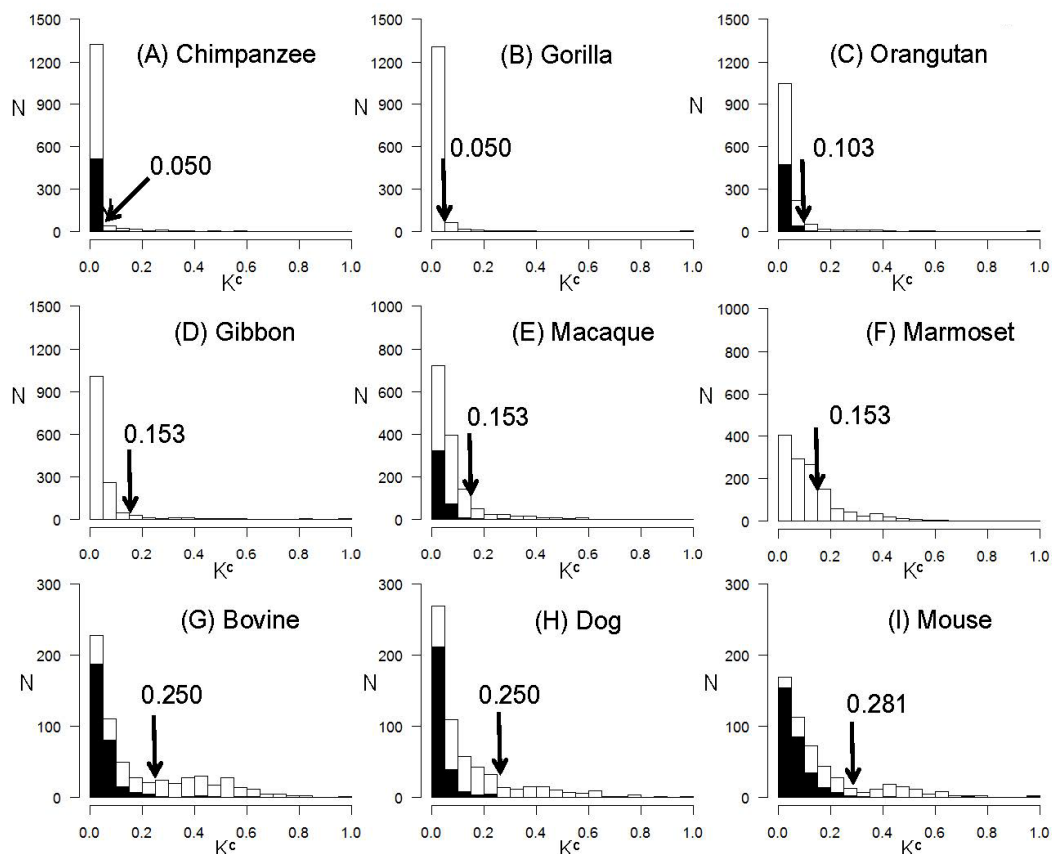


Figure 1: Distribution of the sequence differences, calculated from the comparisons between human pre-miRNAs and homologues in nine other species.

The nucleotide sequences of human pre-miRNAs were compared with the homologues of (A) chimpanzee, (B) gorilla, (C) orangutan, (D) gibbon, (E) macaque, (F) marmoset, (G) bovine, (H) dog and (I) mouse. The sequence difference (K^c) was calculated and corrected for multiple hits of substitutions. N stands for the number of pre-miRNA homologues over K^c values between human and each of other species.

Closed bars indicate the distribution of K^c from the sequence pairs of known miRNAs that were registered in miRBase. Open bars represent the distribution of K^c from those of human known miRNAs and the homologues found in the genomes of other species. For gibbon and marmoset, no known miRNA sequence data were available. The thresholds to define the outliers were calculated by the Smirnov-Grubbs rejection test from the K^c distributions of known pre-miRNA pairs, at a significance level of $p < 0.01$, and they are also shown in the panels.

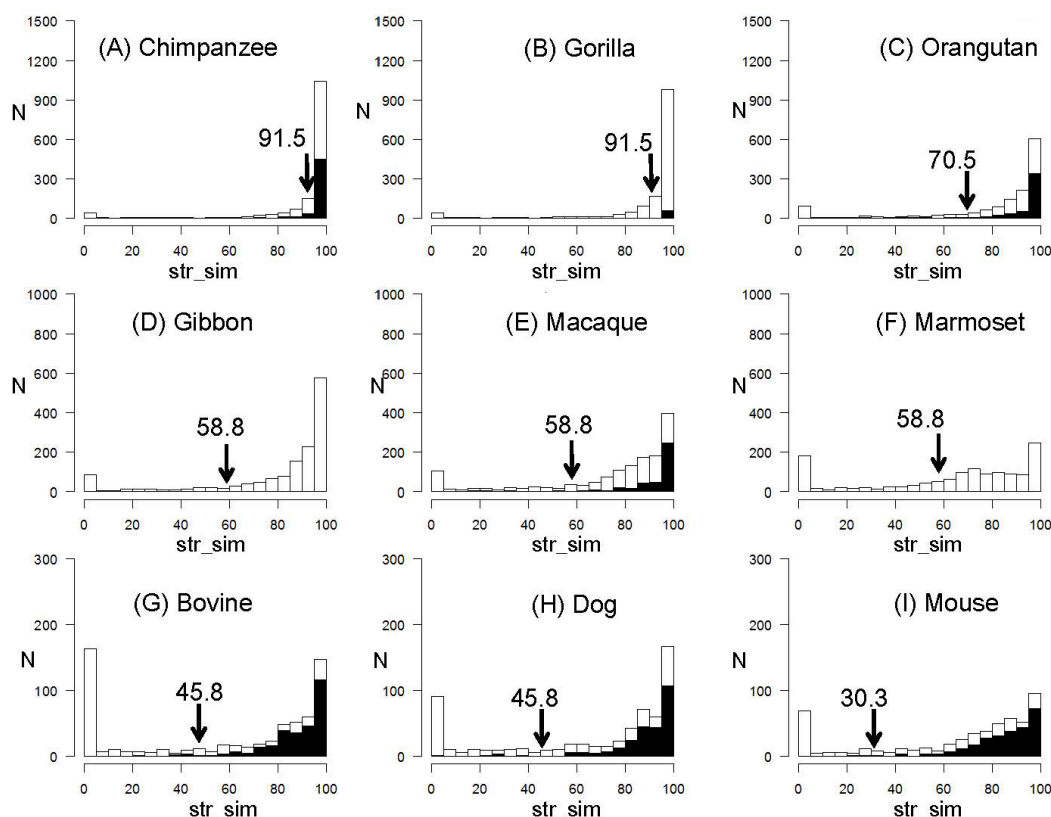


Figure 2: Distribution of the secondary structure conservation scores between human pre-miRNA and homologues in nine other species.

Secondary structure conservation (*str_sim*) was calculated for the pairs of predicted secondary structures from the pre-miRNA sequences of human and nine other species. Panels (A) to (I) indicate the distributions of *str_sim* values from the same species pairs shown in the legend of Figure 1. N stands for the number of pre-miRNA homologues over *str_sim* values between human and each of other species. Open bars show the distributions of *str_sim* values from those of human known miRNAs and homologues found in the genomes of other species. No known miRNA sequence data were available for gibbon and marmoset. The thresholds to define outliers were calculated based on the *str_sim* distributions from known miRNA gene pairs at a significance level of $p < 0.01$, and they are also shown in the panels.

mature sequence filter: no gaps were allowed in the mature sequence, and a seed filter: no mutations were allowed in the seed region.

Table 1 shows the numbers of human miRNA homologues detected by our genome sequence alignment and BLAST searches among the nine other species, and the numbers of functional orthologues predicted at significance level of $p < 0.01$ and $p < 0.05$. Hereafter, in the text, we described the results using the significance level of $p < 0.01$, unless otherwise mentioned. There were few differences in the numbers of human miRNA homologues in primates before the selection by filtering. For example, we found 1,307 homologues from marmoset and 1,441 homologues from chimpanzee. On the contrary, we observed considerably large differences among the numbers of predicted functional orthologues from these species, as the result of filtering. The numbers of predicted functional orthologues varied from 534 in marmoset to 1,088 in chimpanzee. It was noteworthy that the numbers of homologues found in the three non-primate mammals were less than half of those found in primates even before adopting the filters: 620 in bovine, 605 in dog and 522 in mouse.

Estimation of the evolutionary period when each human miRNA gene originated

We constructed an evolutionary profile of the human miRNA genes, based on the presence and absence of the known and functional

orthologues of human miRNA predicted in the genome of each non-human species (Table S1 and S2). When a human miRNA orthologue was considered to be present as a functional one in the genome of a non-human species, we assumed that this miRNA gene was also functional in the common ancestor between these organisms and human. The evolutionary origin of this miRNA gene can be estimated as far back as the divergence time between human and the most distantly related species possessing the orthologue. This is basically the same idea proposed by Iwama et al. [8]. When we predicted the functional orthologues of human miRNA genes in any of the three non-primate mammalian genomes, we assumed that they originated before simian evolution, probably before the eutherian radiation. We did not further classify them in terms of the evolutionary origin.

Rates of generation of new miRNA genes and gene families

We classified all of the present 1,527 human pre-miRNAs into eight groups, according to the estimated time period when the functional miRNA gene was generated (Table 2, S1 and S2). The eight time periods along the evolutionary lineage were defined as follows.

Period 1: Time period from the divergence between human and chimpanzee (Node 1) to date.

Period 2: Time period from divergence between human and gorilla (Node 2) to Node 1.

Period 3: Time period from divergence between human and orangutan (Node 3) to Node 2.

Period 4: Time period from divergence time between human and gibbon (Node 4) to Node 3.

Period 5: Time period from divergence time between human and macaque (Node 5) to Node 4.

Period 6: Time period from divergence between old world monkey and new world monkey (Node 6) to Node 5.

Period 7: The initial phase of the simian lineage; time period from divergence between primate and rodent (Node 7) to Node 6.

Period 8: Any time before Node 7.

In total, 122 human pre-miRNAs, including the four pre-miRNAs for which no homologues were found in any other species, were supposed to originate within Period 1 and to be functional only in human. We found that the miRNA genes gained variations about seven-times faster (23.2/3.4) in the lineage toward human, after the divergence between old world monkey and new world monkey, as compared with the earlier stage of primate evolution. The rate has been slightly decreased in the hominoid lineage.

We also estimated the evolutionary period when each miRNA gene family appeared. According to miFam.dat in the miRBase, 875 out of 1,527 human miRNA genes were classified into 474 miRNA gene families, while the remaining 652 miRNA genes were not categorized in any families. We found three additional homologous gene pairs among these 652 miRNA genes: hsa-mir-1295a and hsa-mir-1295b; hsa-mir-3184 and hsa-mir-423; and hsa-mir-499a and hsa-mir-499b, by a mutual similarity search of the 652 pre-miRNA sequences. They were regarded as three additional gene families. We classified the functional orthologues of each human miRNA gene predicted in this study as the additional members of the miRNA gene family to which the human miRNA belongs. We then estimated the evolutionary period when each of these gene families originated, by determining the most distantly-related species to human. As a result, 419 miRNA genes (27.4%, 419/1,527) and 228 gene families (25.6%, 288/1,123) were supposed to have originated before primates branched from the common ancestor of other eutherian mammals, because at least one orthologue was found in either the dog, bovine or mouse genome. The growth rate of the mRNA gene families showed a similar expansion to that of the miRNA genes; they increased much faster (approximately eight-fold, 19.3/2.4) after the divergence between old world monkey and new world monkey, and decreased by approximately 65.6% after the divergence between human and orangutan. We also found that the emergence of the majority of the new miRNA genes could be explained by the generation of new miRNA families through the course of primate evolution.

Discussion

In this study, we estimated the evolutionary origin of all the present human miRNA genes, based on the evolutionary profiles constructed according to the presence and absence of the known and predicted functional orthologues in the other nine genomes analyzed. The presence of homologues of human pre-miRNA that were filtered out was not considered when estimating the evolutionary age of the miRNA. It was not clear if the miRNA assumed a function in the common

ancestor, unless a functional orthologue that is more distantly related to human, than the homologue was predicted. We followed the leads of previous studies, which proposed that once miRNA gene families were acquired in a functional form, they have been highly conserved and seldom lost in eumetazoan lineages [6,7]. We supposed that a major part of the miRNA homologues that were filtered out were generated in the genome, but did not yet assume a functional role, rather than considering them to be pseudogenes that were once functional, but no longer so. Therefore, the origins of the current human miRNA genes estimated in this study should be regarded as the time periods when the miRNA genes appeared with similar functions to their present ones. We, however, cannot exclude the possibility that a certain amount of miRNA genes may have been disappeared in each lineage. Genome sequence data of multiple species in each clade in primates will enable us to estimate the number of 'dead' genes, and to predict more precisely the origin of human miRNA genes in future.

The 122 miRNA genes that originated in Period 1 were supposed to be present only in human. Some of them may not be functional, though. They may have appeared by chance and just expressed hairpins that can give rise to short RNAs. We predicted many more human-specific miRNA genes than the previous study [8], while we found fewer miRNAs that originated before the simian divergence to other mammals. The differences in the results between Iwama et al. [8] and our study could be attributed to the different processes for detecting homologues and the filtering to select out functional orthologues from other homologues. Our screening filters worked particularly well in selecting functional orthologues in distantly related species (Table 1). It should be noted that the predicted origins of each human miRNA gene and gene family could largely shift, depending on the methods for searching homologues and the filters applied.

We found that the diversification of miRNA genes was often accompanied by the generation of new gene families in each evolutionary period (Table 2). Previous studies proposed that *de novo* gene generation, rather than gene duplication, considerably contributed to the generation of new miRNA genes [21,22]. Our study supported their results in each step of primate evolution. The mechanism of *de novo* miRNA generation is not clear, although several processes have been proposed thus far, such as appearance by chance from intergenic or intronic sequences [21,23], and from genomic repeats or transposable elements [24,25].

We also found that the miRNA genes increased in a slightly different mode from that of the miRNA gene families along hominoid evolution. The miRNA genes and gene families both varied dramatically (approximately seven- to eight-fold higher) in the lineage of old world monkey, after the divergence from the common ancestor to new world monkey. The rates of expansion of miRNA and gene families

Species compared	Number of homologues	No. of orthologues (Ratio)			
		p<0.01		p<0.05	
Chimpanzee	1441	1088	(75.5)	1009	(70.0)
Gorilla	1434	1065	(74.3)	976	(68.1)
Orangutan	1404	929	(66.2)	907	(64.6)
Gibbon	1412	934	(66.1)	909	(64.4)
Macaque	1415	807	(57.0)	789	(55.8)
Marmoset	1307	534	(40.9)	498	(38.1)
Bovine	620	339	(54.7)	335	(54.0)
Dog	605	372	(61.5)	366	(60.5)
Mouse	522	331	(63.4)	325	(62.3)

Table 1: Predicted numbers of human miRNA gene homologues and orthologues in the genomes of other nine species.

Period	From (Myr ago)	To (Myr ago)	Time range (Myr)	Origin of miRNA Genes				Origin of miRNA gene families			
				Number		Rate		Number		Rate	
				p<0.01	p<0.05	p<0.01	p<0.05	p<0.01	p<0.05	p<0.01	p<0.05
1	6.3	0	6.3	122	160	19.4	25.4	72	98	11.4	15.6
2	8.8	6.3	2.5	59	63	23.6	25.2	38	39	12.8	15.6
3	15.7	8.8	6.9	134	122	19.4	17.7	88	81	11.0	11.7
4	20.4	15.7	4.7	89	91	18.9	19.4	92	98	19.5	20.9
5	29.0	20.4	8.6	219	220	25.4	25.6	161	161	18.7	18.7
6	42.6	29.0	13.6	316	317	23.2	23.3	263	261	19.3	19.2
7	92.3	42.6	49.7	169	144	3.4	2.9	121	104	2.4	2.1

The numbers of human miRNA orthologues predicted at significance levels of p<0.01 and p<0.05 were shown for each evolutionary period of birth. The generation rate of miRNA genes per Myr was calculated based on the divergence time estimated by TimeTree [10].

Table 2: Classification of human miRNAs by evolutionary origin.

have tended to be declining in hominoids. Our results did not yield an essential conflict with those reported by Iwama et al. [8], in terms of the growth rate of human miRNA gene generation. We followed TimeTree to obtain the divergence times between human and the other species analyzed, while Iwama et al. [8] estimated the evolutionary distance by referring to the number of neutral substitutions per site, calculated based on the EPO-35 alignment. When we considered the evolutionary distance among the organisms, as Iwama et al. [8] did, we also found a peak of accelerated rate before the divergence of gibbon and human. The rate of miRNA gene acquisition was substantially affected by the method of evolutionary distance estimation. However, it was consistent in noting that the expansion of miRNA genes and gene families occurred in the lineage of old world monkey, after the divergence of new world monkey.

The diversification process of miRNA genes was revealed to be quite different from that of protein-coding genes. Clamp et al. [26] suggested that there are only 168 human-specific genes out of ~ 21,000 human protein-coding genes, and the mammalian protein-coding genes have been largely stable, with relatively little invention of truly novel genes. This observation was confirmed by the comparative analysis based on Compara in Ensembl Release 69 [27]. We found approximately 80% of human protein coding genes have orthologues in each of the nine other species from mouse to chimpanzee. The ratio was roughly maintained in these species. Expansion of genes in the lineage of old world monkey was supposed to be unique to miRNA genes. Further studies on characterization of miRNAs taking account of the evolutionary age, such as classification of target genes, expression level and tissue specificity will give clues to explain the possible association of miRNA genes in getting morphological complexity in primate evolution.

Acknowledgment

This work was supported by Kazusa DNA Research Institute.

References

- Bartel DP (2004) MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* 116: 281-297.
- He L, Hannon GJ (2004) MicroRNAs: Small RNAs with a big role in gene regulation. *Nat Rev Genet* 5: 522-531.
- Kloosterman WP, Plasterk RH (2006) The diverse functions of microRNAs in animal development and disease. *Dev Cell* 11: 441-450.
- Nielsen CB, Shomron N, Sandberg R, Hornstein E, Kitzman J, et al. (2007) Determinants of targeting by endogenous and exogenous microRNAs and siRNAs. *RNA* 13: 1894-1910.
- Bartel DP (2009) MicroRNAs: Target recognition and regulatory functions. *Cell* 136: 215-233.
- Heimberg AM, Sempere LF, Moy VN, Donoghue PC, Peterson KJ (2008) MicroRNAs and the advent of vertebrate morphological complexity. *Proc Natl Acad Sci U S A* 105: 2946-2950.
- Peterson KJ, Dietrich MR, McPeck MA (2009) MicroRNAs and metazoan macroevolution: Insights into canalization, complexity, and the Cambrian explosion. *Bioessays* 31: 736-747.
- Iwama H, Kato K, Imachi H, Murao K, Masaki T (2013) Human microRNAs originated from two periods at accelerated rates in mammalian evolution. *Mol Biol Evol* 30: 613-626.
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* 25: 3389-3402.
- Hedges SB, Dudley J, Kumar S (2006) TimeTree: A public knowledge-base of divergence times among organisms. *Bioinformatics* 22: 2971-2972.
- Kozomara A, Griffiths-Jones S (2011) miRBase: Integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res* 39: D152-D157.
- ftp://ftp.ensembl.org/pub/release-69/emf/ensembl-compara/epo_6_primate/
- <http://www.ensembl.org/info/docs/api/compara/index.html#api>
- <ftp://ftp.ncbi.nih.gov/genomes/>
- 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.
- Jukes TH, Cantor CR (1969) Evolution of protein molecules: Mammalian protein metabolisms, III. Academic Press, New York, USA.
- <http://evolution.genetics.washington.edu/phylip/getme.html>
- Hofacker IL (2003) Vienna RNA secondary structure server. *Nucleic Acids Res* 31: 3429-3431.
- Höchsmann M, Töller T, Giegerich R, Kurtz S (2003) Local similarity in RNA secondary structures. *Proc IEEE Comput Soc Bioinform Conf* 2: 159-168.
- Wang X, Zhang J, Li F, Gu J, He T, et al. (2005) MicroRNA identification based on sequence and structure alignment. *Bioinformatics* 21: 3610-3614.
- Chen K, Rajewsky N (2007) The evolution of gene regulation by transcription factors and microRNAs. *Nat Rev Genet* 8: 93-103.
- Roux J, González-Porta M, Robinson-Rechavi M (2012) Comparative analysis of human and mouse expression data illuminates tissue-specific evolutionary patterns of miRNAs. *Nucleic Acids Res* 40: 5890-5900.
- Campo-Paysaa F, Sémon M, Cameron RA, Peterson KJ, Schubert M (2011) microRNA complements in deuterostomes: Origin and evolution of microRNAs. *Evol Dev* 13: 15-27.
- Smalheiser NR, Torvik VI (2005) Mammalian microRNAs derived from genomic repeats. *Trends Genet* 21: 322-326.
- Piriyapongsa J, Mariño-Ramírez L, Jordan IK (2007) Origin and evolution of human microRNAs from transposable elements. *Genetics* 176: 1323-1337.
- Clamp M, Fry B, Kamal M, Xie X, Cuff J, et al. (2007) Distinguishing protein-coding and noncoding genes in the human genome. *Proc Natl Acad Sci USA* 104: 19428-19433.
- Flicek P, Amodè MR, Barrell D, Beal K, Brent S, et al. (2012) Ensembl 2012. *Nucleic Acids Res* 40: D84-D90.