Analysis of conserved miRNAs in cynomolgus macaque genome using small RNA sequencing and homology searching

Xia Huang[1,*], Shijia Li[1,*], Xiaoming Liu[2], Shuting Huang[1], Shuang Li[1] and Min Zhuo[1]

[1] School of Biology and Biological Engineering, South China University of Technology, Guangzhou, Guangdong, China
[2] Guangzhou Tulip Information Technologies Ltd., Guangzhou, Guangdong, China
[*] These authors contributed equally to this work.

ABSTRACT

MicroRNAs (miRNAs) are important regulators that fine-tune diverse cellular activities. Cynomolgus macaques (*Macaca fascicularis*) are used extensively in biomedical and pharmaceutical research; however, substantially fewer miRNAs have been identified in this species than in humans. Consequently, we investigated conserved miRNA profiles in cynomolgus macaques by homology searching and small RNA sequencing. In total, 1,455 high-confidence miRNA gene loci were identified, 408 of which were also confirmed by RNA sequencing, including 73 new miRNA loci reported in cynomolgus macaques for the first time. Comparing miRNA expression with age, we found a positive correlation between sequence conservation and expression levels during miRNA evolution. Additionally, we found that the miRNA gene locations in cynomolgus macaque genome were very flexible. Most were embedded in intergenic spaces or introns and clustered together. Several miRNAs were found in certain gene locations, including 64 exon-resident miRNAs, six splice-site-overlapping miRNAs (SO-miRNAs), and two pairs of distinct mirror miRNAs. We also identified 78 miRNA clusters, 68 of which were conserved in the human genome, including 10 large miRNA clusters predicted to regulate diverse developmental and cellular processes in cynomolgus macaque. Thus, this study not only expands the number of identified miRNAs in cynomolgus macaques but also provides clues for future research on the differences in miRNA repertoire between macaques and humans.

Subjects Biochemistry, Genetics, Molecular Biology, Zoology
Keywords microRNA, RNA sequencing, Cynomolgus macaque

INTRODUCTION

MiRNAs are small (18–25 nucleotides long), non-coding RNAs that regulate the expression of target genes at the post-transcriptional level. Generally, miRNA genes are transcribed by RNA polymerase II and form primary miRNA hairpins (pri-miRNAs). A region of a pri-miRNA forming hairpin structure is recognized by a microprocessor complex consisting of Drosha and DGCR8 proteins, which produce a stem-loop precursor miRNA

(pre-miRNA) (*Lee et al., 2003*). Pre-miRNAs are then exported to the cytoplasm and are further processed by the enzyme Dicer, which cleaves off the terminal loop, leaving a miRNA duplex about 22 nucleotides in length. Many hairpins produce functional mature miRNAs from both duplex arms (*Kobayashi & Tomari, 2016*; *Rorbach, Unold & Konopka, 2018*). Mature miRNAs assemble with Argonaute proteins into miRNA-induced silencing complexes (miRISCs) to mediate mRNA targeting, primarily through pairing between the miRNA seed region (nucleotides 2–8) and complementary site within the 3′ UTR of mRNA (*Berezikov, 2011*). In addition to the seed region, certain sites in the 3′ part of miRNAs can also contribute to target recognition (i.e., nucleotides 13–16 known as the 'supplemental region') (*Gebert & MacRae, 2019*). In animals, miRISCs induce gene silencing through translation repression and mRNA decay (*Jonas & Izaurralde, 2015*; *Treiber, Treiber & Meister, 2019*). A single miRNA usually regulates a large number of genes simultaneously, and one gene may be regulated by multiple miRNAs. In particular, transcription factors or RNA-binding proteins can be regulated by individual miRNAs and entire cellular pathways can be regulated by miRNA clusters. MiRNA binding of neighboring target sites on a target mRNA can result in cooperative repression (*Gebert & MacRae, 2019*; *Uhlmann et al., 2012*). It has been estimated that about 60% of mammalian protein-coding genes are conserved targets of miRNAs (*Friedman et al., 2009*). The comprehensive interactions between miRNAs and mRNAs often play significant roles in composing complex genetic networks and fine-tuning diverse biological functions, such as organism development, body patterning, tissue homeostasis, differentiation, and cell cycle regulation (*Berezikov, 2011*; *Pinhal et al., 2018*). Evolutionarily, many miRNAs are highly conserved due to their functional importance. However, excessive lineage-specific miRNAs have also emerged in various taxa (*Heimberg et al., 2008*; *Bartel, 2018*). It is worth noting that miRNA expansion appears to be associated with body-plan innovations and other phenotypic changes in bilaterians and vertebrates (*Heimberg et al., 2008*; *Meunier et al., 2013*; *Christodoulou et al., 2010*). Therefore, they contribute considerably to phenotypic evolution in animals.

Cynomolgus macaques (*Macaca fascicularis*), also known as crab-eating or long-tailed monkeys, are distributed widely in Southeast Asia, including Indonesia, Vietnam, Philippines, Cambodia, Myanmar, Malaysia, Brunei, and Thailand (*Shiina et al., 2015*). The species belongs to the Cercopithecinae subfamily of Old-world monkeys and has emerged as an alternative animal model after the banning of rhesus macaques (*Macaca mulatta*) from India. Cynomolgus macaques have been utilized extensively in biomedical and pharmaceutical research, including studies on aging, neurological diseases, diabetes, pregnancy, and drug evaluation (*Theriault et al., 1999*; *Carter, 2011*; *Grigsby, 2016*; *Yang et al., 2014*; *Willard & Shively, 2012*), thus prompting the sequencing of its genome (*Ebeling et al., 2011*; *Yan et al., 2011*). Based on comparative genomic analysis, cynomolgus macaques share 99.2% and 92.8% genomic identity with rhesus macaques and humans, respectively (*Veeranagouda et al., 2015*). Furthermore, recent animal study indicates a strong correlation between increasing metazoan morphological disparity and miRNA repertoires, although no correlation with protein domain diversity (*Deline et al., 2018*). Species-specific miRNAs are considered to be important contributors to cancers, autoimmune diseases, and dissimilar responses of species to toxicant exposure (*Koufaris & Gooderham, 2013*; *Koufaris,*

2016; *Jimenez & Piera-Velazquez, 2013*). Therefore, elucidating the differences in miRNA repertoires between monkeys and humans could have valuable implications for the use of nonhuman primate models in biomedical studies. So far, larger-scale annotations of miRNAs have only been conducted for a few species (*Dannemann et al., 2012a*; *McCreight et al., 2017*) and only 20 of the ~300 known primate species have entries in miRBase (*Kozomara, Birgaoanu & Griffiths-Jones, 2019*; *Perelman et al., 2011*). Furthermore, the number of characterized human pre-miRNAs ($n = 1,917$) is three times larger than that of rhesus macaques ($n = 617$). In addition, cynomolgus macaque miRNAs remain limited to several hundred (*Yang et al., 2014*; *Veeranagouda et al., 2015*; *Yang et al., 2012*). Thus, further studies are needed to extend the miRNA profile of this species. Recently, technological advances such as bioinformatics and high-throughput RNA sequencing allowed the identification of a great number of putative miRNAs in different organisms. The computational methods are better to discover conserved miRNAs comprehensively (*Demirci, Baumbach & Allmer, 2017*), while the RNA-seq method is more advantageous in identifying novel miRNAs with non-canonical structures. In this study, we used the two complementary approaches mentioned above to identify miRNAs in cynomolgus macaques. This study should not only expand the number of recognized miRNAs in cynomolgus macaques but also provide clues to elucidate the differences in miRNA repertoires between macaques and humans, and thus improve our understanding of the evolution and function of miRNAs.

MATERIAL AND METHODS

Sample collection

Peripheral whole blood samples were taken from four unrelated Vietnamese-origin cynomolgus macaques (2–4 years old, one male, three females). The monkeys were healthy and housed at Guangdong Landau Biotechnology Co., Ltd. (Guangzhou China). All experiments were reviewed and approved by the Institutional Animal Care and Use Committee (IACUC) of Guangdong Landau Biotechnology Co., Ltd. (project number: IACUC-003). Peripheral whole blood samples (2 ml) were used for small RNA extraction. Firstly, red blood cells were selectively lysed using an EZNA Blood RNA Kit (Omega, USA) and white blood cells were collected by centrifugation (450 g, 10 min, 4 °C). To avoid interference of 18S and 28S rRNA, small RNAs (18–200 nt) were isolated from white blood cells using RNAiso (TaKaRa, China). The samples were dissolved in 30 µl of diethylpyrocarbonate-treated water and stored at −80 °C prior to small RNA sequencing at the Beijing Genomics Institute (BGI, Shenzhen, China).

Small RNA library construction and sequencing

Four small RNA libraries were prepared and sequenced using standard methods at BGI. Before library construction, the integrity and concentration of samples were checked using an Agilent 2100 Bioanalyzer (Agilent RNA 6000 Nano Kit, USA). The RNAs were size-fractionated on a 15% PAGE gel, and small fragments (range of 18–30 nt) were collected and then amplified by real-time polymerase chain reaction (RT-PCR) using adaptor primers for 17 cycles. The fragments (140–160 bp in length) were then fractionated with a 15% PAGE

gel. The final library was quantified by determining average sequence length using the Agilent 2100 Bioanalyzer (Agilent DNA 1000 Reagents, USA) and quantifying the library using quantitative RT-PCR (StepOnePlus Real-Time PCR System, Applied Biosystems, Foster, USA). The fragment sizes of the four libraries were all between 140 bp and 160 bp and the concentration of each library was more than 8.5 nM, thus meeting the minimum requirements for the construction of a sequencing library. Finally, the libraries were used for sequencing analysis on an HiSeq 4000 System (Illumina, USA). The primary data were then subjected to low-quality filtering. Reads that satisfied any of the following parameters were removed: (1) reads aligned to adaptors or primers with less than three mismatches, (2) reads with >10% unknown bases (N bases), (3) reads with >50% low-quality bases in one read, and (4) reads <18 nt. The remaining clean reads were used in the following analyses. The clean reads of the four samples were submitted to the European Nucleotide Archive (ENA) under study accession number PRJEB38080.

Analysis of sequencing data

The cynomolgus macaque reference genome (v. macFas5) was downloaded from the UCSC Genome Browser (http://genome.ucsc.edu). Clean reads were aligned to the reference genome using BGI software SOAP2 (*Li et al., 2009b*). Sequences perfectly mapped to the genome along their entire length were further annotated by alignment against macaque genome annotations and known annotation documents from the NCBI/UCSC/Ensembl databases (*Benson et al., 2013*; *Kalvari et al., 2018*). Sequences matching rRNA, tRNA, snRNA, snoRNA, and repeat associated sequences or overlapping with exons were discarded, and the remaining sRNA tags were aligned to a reference set of all animal miRNAs from miRBase (release 22) using BLAST by allowing no more than two mismatches outside the seed region. To distinguish authentic miRNAs from false annotations, candidates corresponding to known miRNAs from miRBase were considered conserved miRNA orthologs in the cynomolgus macaques when they were supported by at least 10 reads of mature sequences from the four samples and detected in at least three individuals according to criteria described previously (*Veeranagouda et al., 2015*; *Kozomara & Griffiths-Jones, 2014*). The pre-miRNA sequences were predicted based on the presence of hairpin structures analyzed by MIREAP under default settings (https://sourceforge.net/projects/mireap/) (*Tan et al., 2013*).

Detection of miRNA genes by homology searching

BLAT and BLAST softwares were used for homology searching (*Kent, 2002*). All mammalian pre-miRNA sequences from miRBase (release 22) were aligned to the cynomolgus macaque genome. The candidates were assumed to be conserved pre-miRNA sequences in cynomolgus macaques only when they fulfilled three criteria: (1) similarity of full length pre-miRNA >93%; (2) gene location corresponding to only one miRNA (although can have multiple gene mapping information); and (3) in mature sequences, no more than two mismatches outside the seed region and no more than one mismatch in the seed region compared with existing miRNAs (*Veeranagouda et al., 2015*). Furthermore, 10 nt up and downstream of the conserved pre-miRNA sequences were analyzed by MIREAP (*Tan et al.,*

2013) and RNAfold (*Lorenz et al., 2011*). Stem-loop hairpins were considered acceptable only when they fulfilled three criteria: (1) the mature miRNA present in one arm of the stem; (2) absent of large internal bulges ($n < 6$); and, (3) a minimal folding free energy (MFE) of predicted pre-miRNA secondary structures of no more than -15 kcal/mol. After mapping all pre-miRNA candidates with acceptable secondary structures to the genome, they were further annotated by alignment against known annotation documents from the NCBI/UCSC/Ensembl databases (*Benson et al., 2013*; *Kalvari et al., 2018*). Small RNAs were sorted by their chromosomal locations on the genome. Based on criteria described previously (*Marco et al., 2013*), if two neighboring miRNA loci were located within 10 kb and on the same strand, they were considered clustered miRNA genes. The miRNA orthologs were grouped into different miRNA families according to sequence similarity to known miRNA family members. ClustalW was used for the alignment of sequences. Data patch processing was implemented in python.

Target prediction

The target genes of miRNAs were predicted by TargetScan 7.2 (http://www.targetscan.org), miRanda (http://www.microrna.org), and PITA (https://tools4mirs.org/). Gene ontology (GO) was applied for functional annotation analysis of the predicted targets. Enriched pathways involved with the miRNAs were determined according to the Kyoto Encyclopedia of Genes and Genomes (KEGG). Both GO and KEGG analyses were performed using the clusterProfilers package in R (*Yu et al., 2012*), and an adjusted *P*-value (Benjamini–Hochberg false discovery rate correction) of less than 0.05 was considered significant.

Correlation between age and miRNA expression

Considering the variable RNA concentrations and biases in the RNA-seq procedures for each sample, the expression levels of mature miRNAs were normalized by transcripts per million, i.e., dividing the number of reads for each miRNA by the total number of clean reads for each sample and multiplying it by 10^6 (*Xu et al., 2013*). The ages of the miRNAs were assessed based on the set of homologous sequences identified in other vertebrate species in miRBase (*Meunier et al., 2013*). If the miRNA homologous sequences were shared among vertebrates or mammals, these miRNAs were considered old and grouped into vertebrate or mammalian categories, respectively. If the miRNAs were specific to primates or Cercopithecidae, they were considered young and grouped into primate or Cercopithecidae categories, respectively. The average expression levels of miRNAs in different age groups of the four cynomolgus macaques were statistically analyzed by GraphPad Prism 8. Spearman correlation coefficients (R values) between the age of miRNAs and average expression levels of mature miRNAs were computed by matrix analysis in GraphPad and *P*-values were derived at a confidence interval of 95%.

RESULTS

Identification of miRNA genes in cynomolgus macaque genome by homology searching

To identify previously unknown miRNA sequences in macaques (*Yang et al., 2014*; *Veeranagouda et al., 2015*), we used all mammalian pre-miRNA sequences deposited in miRBase as a reference for homology searching. According to the criteria described above, a total of 1,697 homologous sequences were perfectly mapped to the cynomolgus macaque genome with nonoverlapping gene locations and were then used to analyze their secondary structures. 1,526 of the 1,697 homologous sequences were predicted to form acceptable stem-loop hairpins and were considered conserved miRNA candidates. Additionally, we annotated the 1,526 candidates in the cynomolgus macaque genome using the NCBI/UCSC/Ensembl databases and miRBase for screening of homologies in other species. A candidate identified by homology searching was deemed to be of high confidence if it met any of the following conditions: (1) annotated as miRNA in other species or homolog has experimental support; and, (2) reported as a *Macaca fascicularis* miRNA previously. In total, 1,455 of the 1,526 pre-miRNA candidates showed high confidence. The remaining 71 pre-miRNA candidates lacked support in the other species and were excluded from the following analysis. 672 of the 1,455 candidates were reported previously and the remaining 783 gene loci were firstly annotated in cynomolgus macaque genome (*Yang et al., 2014*; *Veeranagouda et al., 2015*). Compared the 1,455 pre-miRNAs with their homologs in other mammals, the stem-loop sequence similarities of 1,186 candidates (81.5% of 1,455) exceeded 95%. Furthermore, 1,074 candidates (73.8% of 1,455) contained 0 mismatches in the -5 p or -3 p mature sequences, with the remaining 299 and 82 having one and more than one mismatch, respectively. In comparison to human miRNAs, 1,083 of the 1,917 human pre-miRNAs deposited in miRBase had orthologs in the cynomolgus macaque genome. Thus, most pre-miRNA and mature miRNA sequences were highly conserved. Most of the 1,455 miRNA sequences in the cynomolgus macaque were single copies, although others had paralogs at several gene locations. For instance, mir-548, a poorly conserved primate-specific miRNA gene family, was represented by 53 paralogs distributed across 18 chromosomes, similar to that observed in humans (*Liang, Guo & Liu, 2012*). The names, stem-loop sequences, gene locations, homology similarities, mismatches in mature sequences, and MFE values of the predicted hairpin structures for the 1,455 pre-miRNA candidates are listed in Table S1.

Identification of miRNAs in cynomolgus macaque using small RNA sequencing

A major obstacle in research on primate miRNAs is the lack of experimentally verified miRNAs in nonhuman primates. To better characterize the identified miRNAs in cynomolgus macaques, small RNA-seq was performed using peripheral blood leucocytes from the four animals. On average, 14 million clean reads were obtained and about 9 million small RNAs (~5,000 unique reads) were mapped to miRBase for each sample (Table 1). As described in the methods, sequences that matched against orthologs in miRBase were considered as known miRNAs, otherwise, they were considered novel. A

Table 1   Clean reads of RNA sequences in cynomolgus macaques mapped to mature miRNAs in miRBase.

	CE01	CE02	CE03	CE04	Average
Clean reads	11640333	10922961	11573957	20255682	13598233
Unique sRNAs	365682	469182	426731	219056	370163
Clean reads for miRNAs	7686565	6746516	6772031	15351472	9139146
Unique sRNAs for miRNAs	5845	4705	5319	4116	4996

Notes.

CE01, CE02, CE03, and CE04 represent four unrelated cynomolgus macaques. Clean reads are total sRNA reads after filtering noise. Unique sRNAs represent numbers of sRNA types. Clean reads for miRNAs mean clean reads mapped to known miRNAs in miRBase.

total of 546 known mature miRNA sequences and 136 potentially novel mature miRNAs (Table S2) were identified. As the 136 potentially novel miRNAs need further confirmation, only the 546 known miRNAs were used in subsequent analyses. To understand whether these miRNAs resulted from one or both of arms of a given precursor, we assigned the 546 mature miRNAs based on pre-miRNA information. As shown in Table S3, 337 mature miRNAs were expressed from both the -5 p and -3 p arms of 170 pre-miRNAs, and the remaining 209 mature miRNAs were expressed from either the -5 p or -3 p arms of 213 pre-miRNAs. As seven mature miRNAs were derived from more than one pre-miRNA, the 546 mature miRNAs were processed from 383 stem-loop miRNA sequences. We identified 379 of the 383 pre-miRNAs by homology searching (as above), including 357 pre-miRNAs with mapping information for single gene and 22 with mapping information for 2–4 genes. We re-screened the remaining four pre-miRNAs, mir-383, mir-378c, mir-4536, and mir-5698, which were deposited in miRBase against the cynomolgus macaque genome. All showed sequence similarities below 60%, indicating that these four pre-miRNA sequences were distinct to cynomolgus macaques. Analysis of sequencing data indicated that 408 of the 1,455 miRNA loci were expressed in peripheral blood leucocytes. These 408 miRNA loci demonstrated high confidence as they were identified by both RNA sequencing and homology searching and showed homolog support in other species. In total, 335 have been reported in cynomolgus macaques previously (*Yang et al., 2014*; *Veeranagouda et al., 2015*), with the remaining 73 considered to be new pre-miRNA loci in cynomolgus macaques (Table 2).

Thus, in the current study, we confirmed 1,455 high-confidence pre-miRNA sequences by homology screening and RNA sequencing. According to their sequence similarities (*Dostie et al., 2003*; *Yue, Sheng & Orwig, 2008*; *Bartel, 2009*), 915 were grouped into 491 families (Table S1). Another 540 were not clustered into currently known families, suggesting new targeting properties. To assess the age of the miRNAs, their precursor sequences were searched in miRBase. The age of miRNAs was based on the group of species in which miRNA homologous sequences were identified. We classified the 1,455 miRNAs into four groups: i.e., 267 conserved in vertebrates, 332 conserved in mammals, 806 primate-specific, and 50 Cercopithecidae-specific. Among the 408 high-confidence pre-miRNA sequences, 179 were conserved in vertebrates, 140 were conserved in mammals, 77 were conserved in primates, and 12 were conserved in Cercopithecidae. We then

Peerℱ

Table 2 Seventy three newly identified pre-miRNA genes loci in macaque genome.

pre-miRNAs	Accession numbers	Sequences	Average expression level	
			-5p	-3p
mir-10399	LR745789	<u>AATTACAGATTGTCTCAGAGA</u>AAACAGATGAGTTACTCTCTCAGACAAGCTGTAGGTC	14.04	
mir-1246	LR594712	TGCATCCTTG <u>AATGGATTTTTGGAGCAGG</u>AGTGGACACCTGACCCAAAGGAAATCAATCCATAGGCTAGCAAT	0.96	
mir-1248	LR594713	TTT <u>ACCTTCTTGTATAAGCACTGTGCT</u>AAAATTGCAGACACTAGGACTATGTCTTGGTTTTTGCAATAAT GCTAGCAGAGTACACACAAGAAGAAAAGTAACAGCA	2.55	
mir-124-3	LR745790	TGAGGGCCCCTCTGCGTGTTCACAGCGGACCTTGATTTAATGTCTATACAA***TTAAGGCACGCGGTGAATGCCA*** AGAGAGGCGCCTCC		0.22
mir-1294	LR594714	CACTTAATATGTGCCAAGATCTGTTCATTTATGATCTCACTGAGTCC <u>TGTGAGGTTGGCATTGTTGTCT</u> GGCGTTGTCTGATATACAACAGTGCCAACTTCACAGGACTCAGTGAAGTGAAGCTGAGGATTAGGAAGGTGTG	1.04	
mir-1301	LR594715	CTGCCAAGCGACCCCTAGAATGGGGATTGTGGGGGGGTCGCTCTAGGCACCGCAGCACTGTGCTGGGGATG ***TTGCAGCTGCCTGGGAGTGACTTC*** ACACAGTCCTCTCTGCCT		30.95
mir-1307	LR594716	CATCAAGACCCAGCTGAGTCACTGTCACTGCCTACCAATC <u>TCGACCGGACCTCGACCGGCT</u> CGTCTGTGTTGC-CAATCG ***ACTCGGCGTGGCGTCGGTCGTG*** GTAGATAGGCGGTCATGCATACGAATTTTCAGCTCTTGTTCTGGTGAC	1.02	543.87
mir-1327	LR745791	GGGGACTGCTCTT <u>TAAAAGGCTGTTAAATGGTGAT</u> ACTATATTTTTTAACATCAGCATCATTTAGAAACCTTTTAGA-GAACAGTTCTC	1.08	
mir-138-1	LR745792	<u>AGCTGGTGTTGTGAATCAGGCCG</u> TTGCCAATCAGAGAACGGCTACTTCACAACACCAGGGCC	2.20	
mir-1468	LR745793	GGCGGGCGGTTT <u>CTCCGTTTGCCTGTTTTGCTGA</u> TGTACATTCAACTCATTCTCAGCAAAATAAGCAAATG-GAAAATTTGTCCATC	3.87	
mir-151b	LR594718	ACCTCTGATGTATCAATCTCTCTTCGGGGCTCCCGAGACACAGAAACAGACACCTGCCC ***TCGAGGAGCTCACAGTCT*** AGACAAACAAACCCAGGGT		3.50
mir-1839	LR745794	<u>GAAAAGGTAGATAGAACAGGTCTT</u>GTTTGCAAAATAAACTC***AAGGCCTACTTATCTACCAAC***AG	249.11	1.85
mir-1843	LR745796	<u>TATGGAGGTCTCTGTCTGGCT</u>TAGGACAGCTGGCTAAG ***TCTGATCGTTCCCCTCCATAC***A	33.89	51.60

Table 2 (continued)

pre-miRNAs	Accession numbers	Sequences	-5p	-3p
mir-1976	LR745798	TGGCCTCTGGGCACGGGGGTTGGGTGTGCAAAGGGTGGCAGCAAGGAAGGCAGGGTTCCTGAGGTGTGTC CTCCTGCCCTCCTTGCTGTAGACTTTGGCCTGAGCAAGGAG	1.60	
mir-199b	LR594719	CCAGAGGACACCTCCACTCCGTCTA CCCAGTGTTTAGACTATCTGTTC AGGACTCCCAAATTGTACAGTAGTCTGCA-CATTGGTTAGGCTGGGCTGGGTTAGACCCTCGG	230.72	
mir-200b	LR745799	CCAGCTCGGGCAGCCGTGGC CATCTTACTGGGCAGCATTGGA TGGAGTCAGGTCTC *TAATACTGCCTGGTAAT-GATGA* CGGCGGGGCCCTGCACG	0.79	13.71
mir-219b	LR594722	GGAGCTCAGCCAC AGATGTCCAGCCACAATTCTCG GTTGGCCGCAGACTCGTACAAGAATTGCGTTTGGA-CAATCAGTGGCGAAGCCC	1.33	
mir-2355	LR594723	ATCCCCAGATACAGTGGACAATATGCTATTATAATTGTATGGC *ATTGTCCTTGCTGTTTGGAGAT* AA		10.11
mir-3074	LR594725	GGGCTCGACTCCT GTTCCTGCTGAACTGAGCCAG TGTGTAAAATGAGAACTGATATCAGCTCAGTAGGCACCG-GAGGGCGGGTCC	11.20	
mir-3074-2	LR745800	CAGGCTCCAAGGGGGCTTGACTCCT GTTCCTGCTGAACTGAGCCAG TGTGCACAAACCAACTGT-GTTTCAGCTCAGTAGGCACGGGAGGCAGAGCCCAGGGAGGCCA	11.20	
mir-30e	LR745801	TTCTGGGCAGTCTTTGCTAC TGTAAACATCCTTGACTGGAAG CTGTAAGGTGTTCAGAGGAGC *TTTCAGTCGGATGTTTACAGC* GGCAGGCTGCCACGGTCGTCCCCAGCTAC	3875.85	1023.80
mir-3120	LR594726	GGCTGGGTTGTCATGTGACTG CCTGTCTGTGCCTGCTGTACA GGTGAGCGGATGTTCTG *CACAGCAAGTGTA-GACAGGCA* GACACATGACAACTCTGTCCAGCC	4.30	9.52
mir-3155a	LR745803	CACTTTTGAGACGCCTGTTCCGGGCATCA GCTCCCACTGCAGAGGCTGG GGAGCCGGACAGCTCCCTTCC *CAAGCTCTGCAGTGGGAACTGA* TGCCTGGAACAGTTCCTGCA	0.44	7.83
mir-3158	LR594729	ATTCAGGCTGGTCCTGCAGAGAGGAAGCCCTTCTGCTTCCAGGTATTGG *AAGGGCTTCCTCTCTGCAGGAC* CAGCCTGAAT		14.74
mir-3158-2	LR745804	ATTCAGGCTGGTCCTGCAGAGAGGAAGCCCTTCCAATACCTGGAAGCAG *AAGGGCTTCCTCTCTGCAGGAC* CAGCCTGAAT		14.74
mir-3160-1	LR594730	GGACCTACCCTGGGCTTTCTAGTCTCAGCTCTCCTCCAGCTCAACTGGTCAGG *AGAGCTGAGACTAGAAAGCCCA* GGGCAGGTTC		0.72
mir-3160-2	LR745805	ACCTGCCCTGGGCTTTCTAGTCTCAGCTCTCCTGACCAGTTGAGCTGGAGG *AGAGCTGAGACTAGAAAGCCCA* GGGTAGGT		0.72

(*continued on next page*)

Table 2 (*continued*)

pre-miRNAs	Accession numbers	Sequences	Average expression level	
			-5p	-3p
mir-3164	LR745806	CTTGGAAAC <u>TGTGACTTTAAGGGAAATGACG</u> CACAGCAGGCCCTGGAATCACG ***CCGTTTTGCTTGAAGTTGCAGT*** TTCCCAGG	0.97	0.56
mir-3174	LR745808	CCAGCATCAGCATTACCTGG <u>TAGTGAGTTAGAAATGCAGAGC</u> CCCGGGCTTCTCAGCAAACCTACTGGATCTG CATTTTAATTCACATGCATGGTAATGTCTGTAAAGCACT	3.28	
mir-3191	LR745809	GGGGTCACCTGTCTGGCCGTCTACCTTCCACACTGACAAGGGCCG ***TGGGGACGTAGCTGGCCAGACAG*** GTGACCCC		1.24
mir-320b-1	LR745810	AATTAATCCCTCTCTTTCTAGTTCTTCCTAGAGTGAGG ***AAAAGCTGGGTTGAGAGGGCAA*** ACAAATTAACTAATTAATT		119.04
mir-320c-1	LR745811	AAAAATGAGGCCTTATCTTCCCAGTTCTTCCCAGAGTCAGG ***AAAAGCTGGGTTGAGAGGGT*** AGAAAAAAAAT		41.88
mir-320d-1	LR745812	TTCTCTTCCCAGTTCTTCCCAAAGTTGAG ***AAAAGCTGGGTTGAGAGGA***		25.86
mir-320d-3	LR745813	TCTCTTCCTGGTTCTTCCCGAAGTCAGG ***AAAAGCTGGGTTGAGAGGA***		25.86
mir-320e	LR745814	CTCCATGGGGCTTTCTCTTCCCAGTTCTTCCTGGAGTCGGGG ***AAAAGCTGGGTTGAGAGGGTGA*** ACAGAAAAA		2.08
mir-326	LR594733	CTCATCTGTCTGTTGGGCTGGAGGCAGGGCCTTTGTGAAGGCG GGTGGTGCTCAGATCG ***CCTCTGGGCCCTTCCTCCAG*** CCCCGAGGCGGATTCA		7.38
mir-3609	LR594734	GTAACATTAACTTTTATTCTCGTTTTCCTTTTCTCTACCTTGTAGAGAAG ***CAAAGTGATGAGTAATACTGGCTG*** GAGCCC		1.38
mir-3613	LR745815	TGGTTGGGTTTGGAT <u>TGTTGTACTTTTTTTTTTGTTC</u> GTTGCATTTTTAGGAACAAAAAAAAAAAGCCCAACCCTTCA CACCACTTCA	115.70	
mir-3620	LR745816	GTGAGGTGGGGGCCAGCAGGGA <u>GTGGGCTGGGCTGGGCTGGGCC</u> AAGGTACAAGGCCTCACCCTGCATCCCGCAC-CCAG	0.81	
mir-3691	LR594736	TTGAGGCACTGGGT <u>AGTGGATGATGGAGACTCGGTAC</u> CCACTGCAGAGGGTGGGGACCAAGTCTGCATCAT CCACTCCTCAGTGCCTCAG	0.89	
mir-371b	LR745817	GGTAAC <u>ACTCAAAACATGGCGGCACTT</u> TTTTCACCAGAGAGCAGAAAGTGCCCCCACAGTTTGAGTGCC	1.38	
mir-378g	LR745818	C <u>ACTGGGCTTGGAGTCAGAAG</u> ACCTGACTCCAGCCCAGGTC	0.40	
mir-3913-1	LR745819	TTGTTTATAATAAACTGAAATA <u>TTTGGGACTGATCTTGACACTCT</u> TACATAAAATGTTTTGGCAGACATCAA GATCAGTCCCAAATATTTCAGTTTATTATAGACAG	8.91	
mir-4485	LR594738	AGAGGCACCGCCTGCCCAGTGACACATGTT ***TAACGGCCGCGGTACCCTGA*** CTGTGCA		0.38
mir-4659a	LR745820	GAAACTGATGAAGCTGCCATGTCTAAGAAGAAAACTTTGGAGAAAAAT ***TTTCTTCTTAGACATGGCAGCG*** TCAACAGTTTC		0.93

Table 2 (*continued*)

pre-miRNAs	Accession numbers	Sequences	Average expression level	
			-5p	-3p
mir-4661	LR745821	TTTACTCTG ***AACTAGCTCTGCGGATCCTGAT*** AGACAGCCTGATAGACAGTATCCACAGAGCTAGTCCAGAGTAAA	0.75	
mir-4676	LR594739	TGAACAAAA <u>GAGCCAGTGGTGAGACAGTGA</u> GTTGATTACTTCTCACTGTTTCACCACTGGCTCTTTGGTTCA	16.80	
mir-4684	LR594740	GCACCAGGGGTAC <u>CTCTCTACTGACTTGCAACATA</u> CATTTGTATTGGTG ***TGTTGCAAGTCAGTGGAGAGGT*** ACCCTTGGTGT	0.71	0.56
mir-4791	LR594741	TAAGAAC <u>TGGATATGAAGACTGAAA</u> TAAGCTCCATATCAATGAGAATTTCAATGGGATTATGTATAGTCAAT GTCCAGTAATTA	0.63	
mir-486-1	LR745823	TCTCCATCCTCCCTGGGGCA <u>TCCTGTACTGAGCTGCCCCGAG</u> GCCCTTCATGCTGCCCAGCT ***CGGGGCAGCTCAGTACAGGAT*** ACCTCGGGGTGGGAGTCAGCAGGAGGTGA	7013.65	332.08
mir-4999	LR745825	ATAGAAAATAAAACACATACT <u>GCCGTATTGTCAGGTAGTGATA</u> GGATTTA ***TCACTACCTGACAATACAATAT*** GTGTTTGTTTTATTTTATGT	1.41	0.26
mir-548ad	LR594743	CTATTAGGTTGGTGC <u>AAAAGTAATTGTGGTTTTTG</u> AAAGTAACTTGGCGAAAACCACAATGACTTT GCACCAACCTAATAC	2.20	
mir-548at	LR745827	TAGGTTGGTG <u>CAAAAGTTGTTGCGGTTTTGGC</u> CGCCAAAAGAAATGGCCAAAACCGCAATAACTTTT GTACCAACCTAA	1.34	
mir-548av	LR594747	AAAAGTACTTGTGGATTTGCCATTACCTTTACCTTTAATGGC ***AAAACTGCAGTTACTTTGC***		0.36
mir-548ay	LR594748	AGAAGATGCTTACTACTAGGTTGGTGC <u>AAAAGTAATTGTGGTTTTTGC</u> ATTTAAAGTAATGGCCAAAACCGCG ATTACTCTTGCACGAACCTAACGGTAACACTTCT	3.98	
mir-548l	LR745828	TATTAGGTTGGTGCAAAAGTATTTGCGGGTTTTGTCATTGAAAGTAATGG ***CAAAAACTGCAATTACTTCTGC*** ACCAACCTAATGC		1.08
mir-5690	LR594755	CTTTTAATT <u>TCAGCTACTACCTCTATTAGG</u> ATTTGGGAATTATACTAATAGAGGTAGTAGTTGAAATTAAGAG	4.21	
mir-610	LR594756	TCTATTTGTCTTAGG <u>TGAGCTAAATGTGTGCTAGGA</u> CACATTTGAGCCAAATGTCCCAGCACACATTTAGC TCACATAAGAAAAATGAACTCTAGT	8.26	
mir-641	LR594757	CAGGCTGGGTGAAAGGAAGG <u>AAAGACATAGGATAGAGTCACCTC</u> TGTCCTCTGTCCTCCACCTATAGAGGTGA CTATCCTATGTCTTTCCTTCCTCTCACCCCTGAGTCTCA	1.06	

PeerJ

Table 2 (*continued*)

pre-miRNAs	Accession numbers	Sequences	Average expression level	
			-5p	-3p
mir-6501	LR594758	GG <u>AGTTGCCAGGGCTGCCTTTGGT</u> GACAGCAGCAGTAGAGTTGCCAGAGCAGCCTGCGGTAACAGTA	1.74	
mir-6503	LR745829	AATGGTCCCCCAGGG <u>AGGTCTGCGTTCTAATCCCCA</u> GAAGCTAAGGATTAGG ***GGGACAAGGATGCAGACCTCC*** CTGGGGGACCGTT	7.58	18.52
mir-6503-2	LR745831	AATGGTCCCCGAG <u>GGAGGTCTGCATTCTAATC</u> CCCAGAAGCGAAGGATTAGG ***GGGACAAGGATGCAGACCTCC*** CTGGGGGACCATT	0.17	12.61
mir-6516	LR745833	TGGGTTTTGAA <u>TTTGCAGTAACAGGTGTGAGCA</u> TTCTAGCAGCAGTTTGGTGATCATGTATGATACTG-CAAACAGGACCTA	0.32	
mir-655	LR594759	TTCGTTTCAGAACTATTCAAGGATATTTGAGGAGAGGTTATCCGT GTTATGTTCGCTTCATTCATCATGA ***ATAATACATGGTTAACCTCTTT*** TTGAATATCAGACTCT		2.35
mir-6731	LR594760	ACAGG <u>TGGGAGAGCAGGGTATTGTGGA</u> AGCTCCAGGTGCCAACTGCCTGCCTCTATCCCCCACTCTCCCCAG	1.20	
mir-6735	LR745834	GCAGCC <u>AGGGCAGAGAGCACAGGAATCTGA</u> GGTGACTGGCACAGAAGACTC ***AGGCCTGTGGCTCCTCCCCCAG***	1.14	0.09
mir-6813	LR594761	ACAGG <u>CAGGGGCTGGGGTTTCAGGTTCT</u> CAGTCAGAACCTTGGCCCCTCTCCCCAG	0.69	
mir-6816	LR745836	CCGAGTGGGGCGGGGTGGGTCCCTGCAGGGACTGTGACACT ***GAAGGACCTGCACCTCCGCCCACA*** G		0.82
mir-6866	LR745837	CCATT <u>TTAGAGGCTGGAATAGGGATT</u> ATTGAGTCTGGAAGAGTAAGGATCCCTTTATCTGTCCTCTAG	0.69	
mir-7155	LR594762	TCTGGGGTCTTGGGCCATCTGGTTGTGACAGCCCCGA ***TGGCCCAAGACCTCAGACC***		0.68
mir-744	LR594763	TTGGGCAAGG <u>TGCGGGGCTAGGGCTAACAGCA</u> GTCTTACTGAAGGTTTCCTGGAAACCACGCACATG ***CTGTTGC-CACTAACCTCAACCT*** TACTCGGTC	1140.67	1.71
mir-7848	LR594765	GCTGGAGCTGGGTGGGTGTGGCAGGCCCACCGTGGGTATGC AAAGCTCTGACAATGTTTTACTTG ***CTACCCTCGGTCTGCTTACCACA*** CTCCCAGTTCCAC		18.23
mir-935	LR745838	GGCGGGGGCGCGGGCGGCAGTGGCGGGAGCGGCCCCTC GGCCATCCTCCGTCTGC ***CCAGTTACCGCTTCCGCTACCGC*** CGCCGCTCCCGCT		1.44

Notes.

Average expression levels for each mature miRNA were normalized by transcripts per million. Single or double underlines indicate mature miRNAs expressed from -5 p or -3 p arms of pre-miRNAs, respectively.
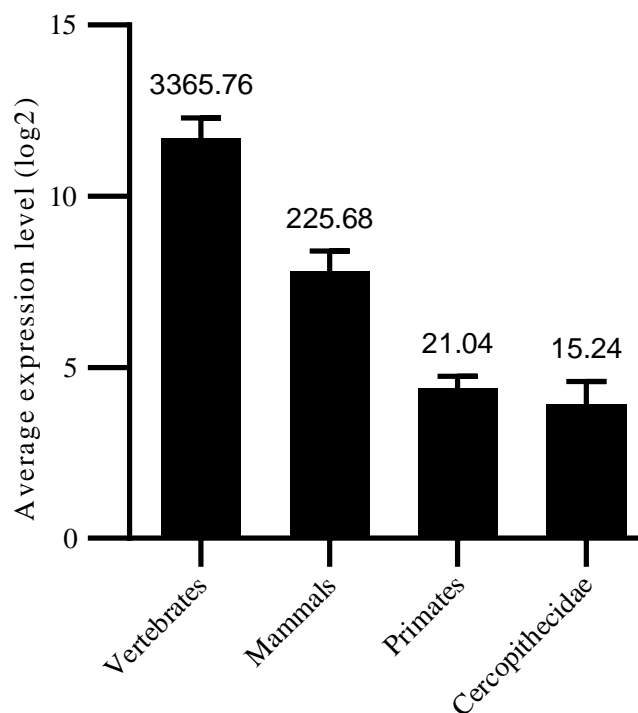
Figure 1 **Age of mature miRNAs relative to expression levels.** Primate-specific miRNAs showed approximately 10-fold and 160-fold lower average expression levels than ancient miRNAs from mammals and vertebrates, respectively.

Full-size ⊡ DOI: 10.7717/peerj.9347/fig-1

examined the relationship between average expression level of miRNAs and their age and found a positive correlation (Spearman R: 0.27 to 0.42, $P < 10^{-4}$; Fig. 1). Primate-specific miRNAs demonstrated lower average expression levels (10-fold and 160-fold, respectively) than ancient miRNAs in mammals and vertebrates. In addition, Cercopithecidae-specific miRNAs showed the lowest expression level among the four groups.

Characteristics of miRNA gene locations in cynomolgus macaques

According to the annotations of the 1,455 high-confidence pre-miRNA sequences (Table S1), the miRNA encoding genes in the cynomolgus macaque genome were also distributed in diverse locations, similar to that found in other mammals (*Kim & Kim, 2007*; *Morin et al., 2008*). Of the 1,455 pre-miRNAs, approximately 46% were located in intergenic regions, 49% were located in introns and 5% were located in RNA coding regions. It is worth noting that 64 pre-miRNA sequences were processed from exons of protein-coding transcripts or non-coding transcripts, including 46 derived from untranslated regions. Six pre-miRNA sequences were found overlapping intron-exon junctions, i.e., SO-miRNAs. Most miRNA gene is encoded by one strand of one locus as above, but in rare cases, two distinct pre-miRNAs are found transcribed and processed on both strands of one locus (*Tyler et al., 2008*; *Picao-Osorio et al., 2015*; *Mohammed et al., 2018*). Two pairs of these so-called antisense or mirror miRNAs were identified in our study, i.e., miR-3120/miR-214 and miR-24-1/miR-3074 (Table S1). Here, miR-3120 was located within an intron of

the dynamin-3 gene and miR-214 was embedded on the antisense strand of the same locus. For the miR-24-1/miR-3074 pair, we found that miR-24-1 was embedded within an intron of the *AOPEP* gene and miR-3074 was processed from the antisense strand. The miR-24-1/miR-3074 mirror miRNAs were first reported in this study in cynomolgus macaques (*Cordero et al., 2015*).

MiRNA gene clusters in cynomolgus macaques

Previous studies have revealed that miRNAs are significantly enriched in clusters in introns or intergenic regions (*Marco et al., 2013*; *Wang et al., 2016*; *Altuvia et al., 2005*). To profile the clustering patterns of miRNAs in cynomolgus macaques, we grouped pre-miRNAs into distinct clusters following previous study (*Marco et al., 2013*). If two neighboring miRNA loci were located within 10 kb and on the same strand, they were considered to be clustered miRNA genes. Among the 1,455 pre-miRNAs, 285 miRNA genes were grouped into 78 distinct clusters. When comparing the 78 miRNA clusters to their human counterparts, if a human miRNA was not represented in the macaque clusters, it was screened cross the macaque genome in order to validate the existence of an ortholog in macaques (*Wang et al., 2016*). As such, 62 miRNA clusters were found to be similar to that in the human genome and six miRNA clusters showed minor changes (Table 3), indicating that most of the clustered miRNAs were evolutionarily conserved. Ten of the 78 clusters were newly emerged in the cynomolgus macaques. According to the similarity of seed regions of the miRNA sequences in each cluster, we grouped the 78 clusters into two sub-classes: 21 homo-seed clusters (miRNAs having identical sequences in the seed region) and 57 hetero-seed clusters (miRNAs having distinct sequences in the seed region). These results are consistent with previous studies suggesting that the composition of miRNA clusters is heterogeneous with a diversiform evolutionary mechanism (*Wang et al., 2016*). Homo-seed clusters are usually products of local duplications, and hetero-seed clusters are mainly shaped by a combination of *de novo* formations and local duplications (*Wang et al., 2016*).

For most miRNA clusters revealed in this study, the number of pre-miRNAs ranged from two to three. Nonetheless, 10 large clusters with no less than five miRNA precursors were also observed (Table 3), which were composed of members from different miRNA families but with the same expression state. Among the 10 large clusters, most miRNAs in the six of them were detected by RNA-seq in this study, but none was detected by RNA-seq in the rest of the four clusters (Table 3). These results indicate that each of the 10 clusters was transcribed as a polycistronic element (*Ryazansky, Gvozdev & Berezikov, 2011*; *Saini, Griffiths-Jones & Enright, 2007*). Comparing the compositions of the 10 miRNA-clusters in cynomolgus macaques to their orthologs in humans, eight showed identical composition, whereas the mir-379/mir-656 and mir-512/mir-517 clusters showed minor changes. The mir-379/mir-656 cluster in macaques was positioned on chromosome 7 and contained 42 miRNA genes. Compared to humans, 40 out of the 42 miRNA genes were the same as humans. Simultaneously, a mir-376a copy was added and a mir-300 ortholog was lost. In the macaques, the mir-512/mir-517 cluster was located on chromosome 19 and contained 40 miRNA genes. Most shared common seed regions of the miR-515 family

Table 3 MiRNA gene clusters in cynomolgus macaques (*Mafa*).

Cluster in *Mafa*	miRNA number in *Mafa*	miRNAs in *Mafa*	Type	Detected by RNA-seq	Cluster in humans	miRNA number in humans	miRNAs in humans
chr1:128814989-128815918+	2	137/2682	hetero	no	chr1:98045270-98046113-	2	2682/137
chr1:186962061-186965108-	2	30c-1/30e	homo	yes	chr1:40754370-40757360+	2	30e/30c-1
chr1:227255265-227256419-	2	429/200a[a]	homo	yes	chr1:1167123-1169076+	3	200b/200a/429
chr1:30111736-30117576-	2	214/199a-2	hetero	yes	chr1:172138815-172144614-	2	214/199a
chr1:57185008-57185289-	2	181b-1/181a-1	homo	yes	chr1:198858886-198859130-	2	181b/181a
chr1:66367833-66368516-	2	29c/29b-1	homo	yes	chr1:207801864-207802513-	2	29c/29b-2
chr1:78854733-78855125-	2	215/194-1	hetero	yes	chr1:220117876-220118227-	2	215/194
chr10:30075091-30075493-	2	130b/301b	homo	yes	chr22:21652989-21653375+	2	301b/130b
chr10:31357983-31358388-	2	1306/3618	hetero	yes	chr22:20085796-20086129+	2	3618/1306
chr10:90709107-90709791-	2	296/298	hetero	yes	chr20:58817625-58818303-	2	296/298
chr11:112931512-112932150-	2	7705/10527[b]	hetero	no	not detected		
chr11:56821914-56822517-	2	26a-2/9947[b]	hetero	yes	not detected		
chr11:7274754-7275246+	2	200c/141	homo	yes	chr12:6963702-6964176+	2	200c/141
chr12:65115697-65116202+	2	7164/10b[b]	hetero	yes	not detected		
chr13:54048315-54054337+	2	216a/217	hetero	no	chr2:55983019-55989041-	2	217/216a
chr14:107474425-107475005+	2	34b/34c	homo	no	chr11:111512949-111513505+	2	34b/34c

Huang et al. (2020), *PeerJ*, DOI 10.7717/peerj.9347

Peer J

Table 3 (*continued*)

Cluster in *Mafa*	miRNA number in *Mafa*	miRNAs in *Mafa*	Type	Detected by RNA-seq	Cluster in humans	miRNA number in humans	miRNAs in humans
chr14:6660857-6661975-	2	4691/7113	hetero	no	chr11:68032864-68033971+	2	7113/4691
chr15:108224172-108226742-	2	3596d/3596b[b]	homo	no	not detected		
chr15:10960721-10961048+	2	3154/199b	hetero	yes	chr9:128244744-128244977-	2	199b/3154
chr15:11398154-11399489-	2	181b-2/181a-2	homo	yes	chr1:198858886-198859130-	2	181b/181a
chr16:18135796-18136249-	2	6777/33b	hetero	no	chr17:17813479-17813916-	2	6777/33b
chr16:1954993-1955463-	2	132/212	homo	yes	chr17:2049928-2050350-	2	132/212
chr16:33365362-33370038+	2	10226/148c[b]	hetero	no	not detected		
chr16:43850674-43855488-	2	142/4736	hetero	yes	chr17:58331244-58336022-	2	142/4736
chr16:61839928-61840468-	2	3064/5047	hetero	yes	chr17:64500774-64501250-	2	3064/5047
chr16:7046633-7047042-	2	195/497	hetero	yes	chr17:7017627-7017999-	2	195/497
chr17:29263944-29264174-	2	16-1/15a	homo	yes	chr3:160404606-160404818+	2	15b/16-2
chr18:60473826-60477188+	2	1-2/133a-1	hetero	no	chr18:21825711-21829036-	2	133a/1
chr19:14206540-14206853+	2	181c/181d	homo	yes	chr1:198858886-198859130-	2	181b/181a
chr19:14522034-14522404+	2	8994/9307[b]	hetero	no	not detected		
chr19:2102695-2104529-	2	1227/6789	hetero	no	chr19:2234061-2235922-	2	1227/6789
chr2:104189414-104189980-	2	425/191	hetero	yes	chr3:49020158-49020694-	2	425/191
chr2:62802629-62802867+	2	15b/16-2	homo	yes	chr3:160404606-160404818+	2	15b/16-2
chr20:14704691-14710146+	2	193b/365a	hetero	yes	chr16:14303979-14309361+	2	193b/365a
chr20:44876971-44882123+	2	138/7181[b]	hetero	yes	not detected		
chr3:118484198-118485386+	2	489/653	hetero	no	chr7:93482783-93484006-	2	653/489
chr3:13210687-13210931-	2	6501/9922[b]	hetero	yes	not detected		
chr3:163632511-163633319-	2	29a/29b-2	homo	yes	chr1:207801864-207802513-	2	29c/29b-2
chr3:30725019-30725854-	2	let-7c/99	hetero	yes	chr21:16539100-16539904+	2	99a/let-7c
chr4:118507476-118512141-	2	133b/206	hetero	yes	chr6:52144400-52149009+	2	206/133b

Huang et al. (2020), *PeerJ*, DOI 10.7717/peerj.9347

17/34

Peerl

Table 3 (*continued*)

Cluster in *Mafa*	miRNA number in *Mafa*	miRNAs in *Mafa*	Type	Detected by RNA-seq	Cluster in humans	miRNA number in humans	miRNAs in humans
chr6:148755825-148757619+	2	143/145	hetero	yes	chr5:149428943-149430720+	2	143/145
chr6:54606090-54606301-	2	449a/449b[a]	homo	yes	chr5:55170585-55172337-	3	449a/449b/ 449c
chr7:68810880-68814721+	2	1179/7−2	hetero	no	chr15:88608120-88611917+	2	1179/7−2
chr8:100682570-100682796-	2	599/875	hetero	yes	chr8:99536650-99536851-	2	599/875
chr8:136192174-136197663-	2	30b/30d	homo	yes	chr8:134800531-134804940-	2	30b/30d
chrun_ke145894:2386-3511-	2	10396a/663-10[b]	hetero	no	not detected		
chrx:111594304-111595504+	2	1912/1264	hetero	no	chrx:114651588-114652718+	2	1912/1264
chrx:144101311-144108532-	2	513b-1/513b-2	homo	no	chrx:147189716-147199114-	2	513c/513b
chrx:148922686-148923813-	2	224/452	hetero	no	chrx:151958583-151959699-	2	224/452
chrx:44130132-44131103-	2	221/222	homo	yes	chrx:45746179-45747094-	2	221/222
chrx:52628444-52629475-	2	98/let-7f-2	homo	yes	chrx:53556240-53557267-	2	98/let-7f-2
chrx:71471342-71471593-	2	421/374b	hetero	yes	chrx:74218391-74218608-	2	421/374b
chrx:71540122-71540378-	2	545/374a	hetero	yes	chrx:74287125-74287346-	2	545/374a
chrx:71841808-71842408-	2	7705/10527[b]	hetero	no	not detected		
chr10:4843233-4844250-	3	let-7b/**4763**/let-7a	homo	yes	chr22:46112751-46113766+	3	let-7a/4763/let-7b
chr14:118172242-118178007-	3	let-7a-2/ 10526 /100[a]	hetero	yes	chr11:122146522-122152296-	2	let-7a-2/100
chr15:108223782-108226749+	3	let-7a-2/let-7f/let-7d	homo	yes	chr9:94175961-94178916+	3	let-7a/let-7f-1/let-7d
chr15:109228189-109229083+	3	23b/27b/24-2	hetero	yes	chr9:95085226-95086085+	3	23b/27b/24-1
chr16:25123023-25123392-	3	451/144/4732	hetero	yes	chr17:28861376-28861722-	3	451a/144/4732
chr16:78700471-78708445-	3	657/338/1250	hetero	yes	chr17:81125290-81133285-	3	657/338/1250

Huang et al. (2020), *PeerJ*, DOI 10.7717/peerj.9347

18/34

Table 3 (*continued*)

Cluster in *Mafa*	miRNA number in *Mafa*	miRNAs in *Mafa*	Type	Detected by RNA-seq	Cluster in humans	miRNA number in humans	miRNAs in humans
chr19:14166279-14166650-	3	24-2/27a/23a	hetero	yes	chr9:95085226-95086085+	3	23b/27b/24-1
chr19:52420893-52421627+	3	99b/let-7e/125a	hetero	yes	chr19:51692617-51693327+	3	99b/let-7e/125a
chr19:54627979-54628968+	3	371/372/373	hetero	yes	chr19:53787677-53788770+	3	371a/372/373
chr3:162480895-162485756-	3	182/96/183	hetero	yes	chr7:129770405-129774988-	3	182/96/183
chr3:44449530-44450041-	3	25/93/106	hetero	yes	chr7:100093570-100094063-	3	25/93/106b
chr8:146049004-146055940-	3	939/1234/6849	hetero	yes	chr8:144394160-144400340-	3	939/1234/6849
chrx:149364083-149366356-	3	105-1/767/105-2	hetero	no	chrx:152392227-152394480-	3	105−1/105−2/767
chr14:9660229-9660537+	4	**6750/6749**/194-2/192	hetero	yes	chr11:64891158-64902450-	4	192/194−2/6750/6749
chr5:111988697-111989381-	5	367/302d/302a/302c/302b	hetero	no	chr4:112647876-112648547-	5	367/302d/302a/302c/302b
chr17:74024859-74025648+	6	17/18a/19a/20a/19b-1/92a-1	hetero	yes	chr13:91350617-91351382+	6	17/18a/19a/20a/19b-1/92a-1
chrx:131080800-131081707-	6	363-2/92a-1/19b-2/20b/18b/106a	hetero	yes	chrx:134169381-134170266-	6	363/92a-2/19b-2/20b/18b/106a
chrx:131455304-131461841-	6	450b/450a-1/450a-2/542/503/424	hetero	yes	chrx:134540193-134546701-	6	450b/450a-1/450a-2/542/503/424
chrx:142887921-142895601-	6	892c/890/888/ 892a/892b/891b	hetero	no	chrx:145992759-146001121-	6	892c/890/888/892a/892b/891b
chrx:48324406-48334989+	7	532/188/500a/362/501/660/502	hetero	yes	chrx:50003166-50014670+	8	532/188/500a/362/501/500b/660/502
chr7:166087997-166103699+	8	493/337/665/431/433/127/432/136	hetero	yes	chr14:100869074-100884771+	8	493/337/665/431/433/127/432/136

Table 3 (*continued*)

Cluster in *Mafa*	miRNA number in *Mafa*	miRNAs in *Mafa*	Type	Detected by RNA-seq	Cluster in humans	miRNA number in humans	miRNAs in humans
chr7:166241788-166285152+	42	379/411/299/ 380/1197/ 323a/758/ 329-1/329-2/ 494/1193/543 /495/ 376a-3 / 376c/376a-2/ 654/376b /376a/1185-2/ 1185-1/381/487b/ 539/889/544a/ 655/487a/382/ 134/668/485 /323b/154/ 496/377/541/ 409/412/ 369/410/656[a]	hetero	yes	chr14:101022070-101066786+	42	379/411/ 299/380/ 1197/323a/ 758/329-1/ 329-2/494/ 1193/543/495/ 376c/376a-2/ 654/376b/ 376a-1/ 300/ 1185-1/1185-2/ 381/487b/539/889/ 544a/655/487a/ 382/134/668/485/323b/ 154/496/377/541/409/ 412/369/410/656
chrx:144120330-144177313-	11	513a-1/513a-2/ 513a-3/506/ 507/508/ 514b/509-1/ 509-2/510/514a[a]	hetero	no	chrX:147225859-147284712-	11	513a-2/506/507/508/514b/ 509-2/509-3/509-1/510/514a-1/ 514a-2

Table 3 (*continued*)

Cluster in *Mafa*	miRNA number in *Mafa*	miRNAs in *Mafa*	Type	Detected by RNA-seq	Cluster in humans	miRNA number in humans	miRNAs in humans
chr19:54493931-54598758+	40	512/512-2 /1323/498 /515-2/519e/ 519e/519e /519c/1283/ 520ak/526b/519a-2 /525/523a/518f/ 519a/518b /526a-1/518c/524/517a/ 519d/518a-1/520g/ 518d-1/ 518g /518d-2/ 523b /516b-1/518a-3/ 517c/520g/519b/ 521/518e/ 518a-4/516a-2/ 516a-1/517b[a]	hetero	no	chr19:53666691-53762418+	46	512-1/512-2/1323/498/ 520e /515-1/519e/ 520f/515-2/519c/1283-1/520a/ 526b/519b/525/ 523/518f/ 520b /518b/526a-1/ 520c/518c/524/517a/519d/521-2/ 520d/517b/520g/516b-2/ 526a-2/518e/518a-1/518d/ 516b-1/518a-2/517c/ 520h /521-1/ 522/519a-1/ 527/516a-1/1283-2/ 516a-2/519a-2

Notes.

[a]MiRNA gene clusters with minor changes in cynomolgus macaques compared to their counterparts in humans.

[b]MiRNA gene clusters that newly identified in cynomolgus macaques. MiRNAs that have homologous sequences in macaque genome are marked by bold. Comparing the compositions of clusters in cynomolgus macaque to their orthologs in humans, differences in monkey and human are marked by underscore and double underscore, respectively.

hetero, hetero-seed, miRNAs with distinct sequences in seed region; homo, homo-seed, miRNAs with identical sequences in seed region.
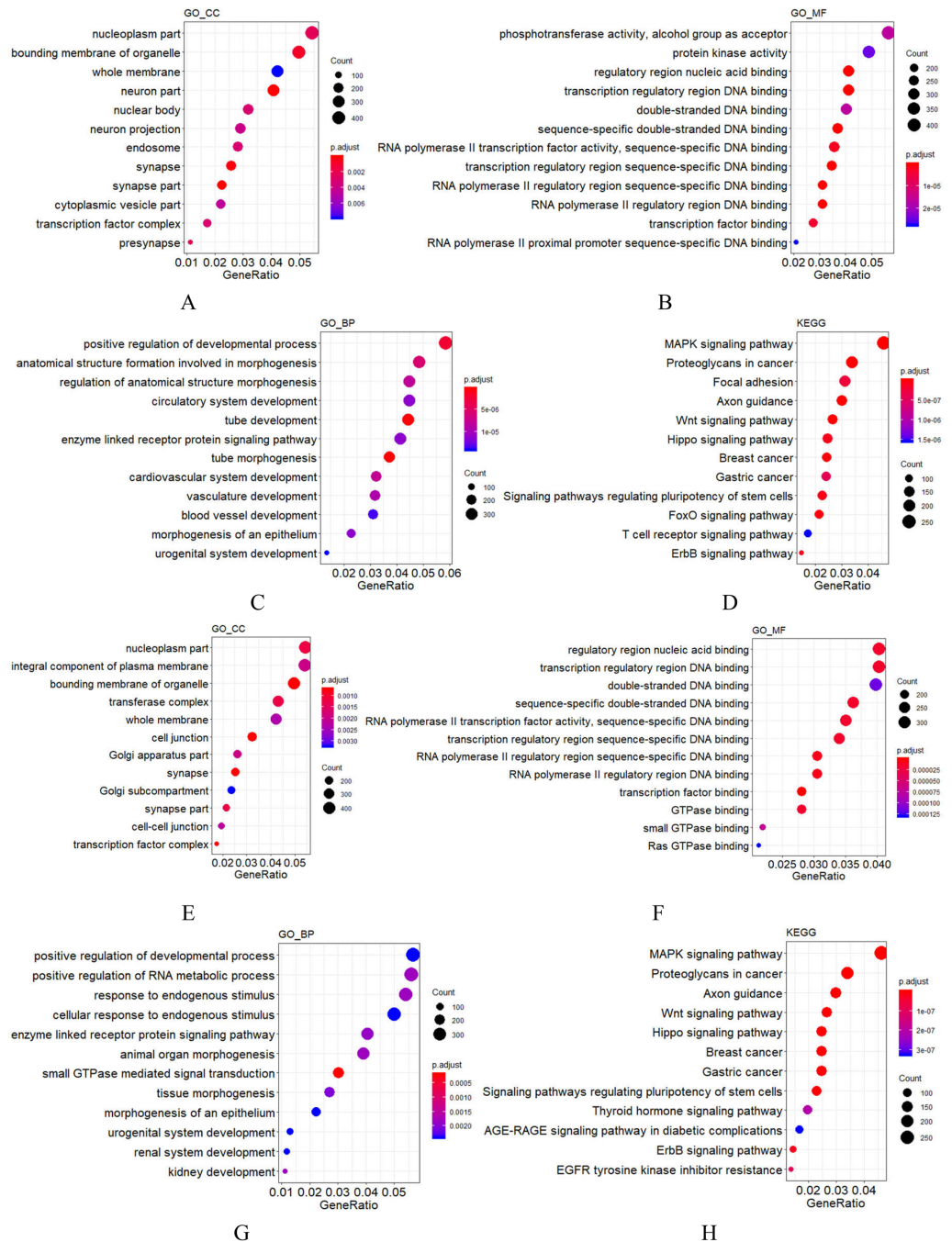
Figure 2 GO and KEGG analysis of mir-379/mir-656 (A–D) and mir-512/mir-517 (E–H) gene clusters.
Vertical axis: GO terms and KEGG pathway names. CC, MF and BP represent cellular component, molecular function and biological process, respectively. The larger the point, the higher the degree of enrichment. The greater the number of candidate target genes in this GO term and KEGG pathway, the color of the points corresponds to different *p*-adjust ranges.

Full-size 🖼 DOI: 10.7717/peerj.9347/fig-2

and likely originated from a common ancestor (*Rodriguez et al., 2004*; *Zhang, Wang & Su, 2008*). Comparing the miRNA composition of the mir-512/mir-517 cluster to the C19MC in humans, 38 of the 40 miRNA genes had orthologs in humans, two (mir-518g and mir-523b) were unique to macaques, and eight, including mir-522, mir-527, and six mir-520 variants, were not found in the macaque genome. These data show that although most C19MC members are conserved throughout the primate kingdom, frequent miRNA gains and losses, such as the gain of mir-518 and mir-523 copies and loss of mir-520 copies in macaques, have occurred (*Berezikov et al., 2006*; *Iwama et al., 2013*; *Awan et al., 2017*). These findings are consistent with previous results suggesting that the C19MC region is under rapid evolution (*Iwama et al., 2013*; *Bentwich et al., 2005*). Specifically, this region is enriched by dispersed Alu elements, which are believed to have facilitated the expansion of C19MC (*Zhang, Wang & Su, 2008*; *Lehnert et al., 2009*). To understand the functions of miRNA clusters, the potential targets of the two very large miRNA clusters were predicted using TargetScan, followed by GO functional annotation. Among the target genes of the mir-379/mir-656 cluster, those relevant to cellular components (CC) were enriched in synapse, membrane, and cell projection; those relevant to molecular function (MF) were mainly clustered in binding and catalytic activity; those relevant to biological process (BP) were mainly enriched in cellular process, biological regulation, metabolic process, development process, and response to stimulus (Fig. 2). For the mir-512/mir-517 cluster, the functional annotations of their target genes were similar to those of the mir-379/mir-656 cluster. KEGG analysis of the two very large miRNA clusters suggested that both regulate normal growth, development, and pathological processes through many pathways, such as MAPK signaling and Axon guidance (Fig. 2), consistent with previous studies on humans (*Haller et al., 2010*; *Lavon et al., 2010*; *Cortez & Calin, 2009*; *Grün et al., 2005*; *Donker et al., 2012*; *Noguer-Dance et al., 2010*; *Tsai et al., 2009*; *Li et al., 2009a*; *Kleinman et al., 2014*; *Ward et al., 2014*). The other eight large miRNA clusters in macaques are with conserved compositions to that in humans, indicating that their functions are also evolutionarily conserved.

DISCUSSION

The majority of reported cynomolgus macaque miRNAs are predicted based on homology to the rhesus genome (*Veeranagouda et al., 2015*). In this study, we firstly identified conserved miRNAs in cynomolgus macaques based on homology to all mammalian pre-miRNA sequences deposited in miRBase. To prevent mis-annotation of homologous sequences as miRNAs, we proposed a more stringent analysis process than previously reported (*Veeranagouda et al., 2015*). First, it requires that the full-length sequence similarity of homologous sequences exceed 93%. Secondly, it requires the homologous sequences to form a stable hairpin-like secondary structure with the mature sequences present in the stem. Next it also requires another supporting information: experimental supports for homolog in other species, or annotated as miRNA in other species, or reported in macaques previously. Using these criteria, 1,455 of the 1,697 homologous sequences were considered pre-miRNA gene loci of high confidence, more than half of which were

firstly annotated in cynomolgus macaque genome (*Yang et al., 2014*; *Veeranagouda et al., 2015*). The data indicated that homology searching is a useful tool for high-throughput identification of conserved miRNAs (*McCreight et al., 2017*). Using these criteria, the false-positive annotations were reduced, but also some real pre-miRNA sequences were missed, such as poorly conserved miRNAs, or miRNAs that were not expressed at sufficient levels in the sequenced samples, or miRNAs that did not form acceptable hairpins. Thus, we just predicted a subset of conserved miRNAs in cynomolgus macaque. Using small RNA sequencing, 408 of the 1,455 genes were detected in peripheral blood leucocytes, which aided the confirmation of -5 p and -3 p mature miRNA sequences. The remaining 1,047 not supported with reads may be expressed in other tissues or under other circumstances, because large number of miRNAs exhibit noticeable spatiotemporal and tissue-specific expression. In addition, environmental stress or physiological disorders will significantly alter miRNA expression profiles (*Pinhal et al., 2018*). Therefore, we just experimentally verified a subset of miRNAs and further experimental confirmations of many miRNAs are needed in cynomolgus macaques.

When contrasting expression levels of miRNAs with their age, our results were in good agreement with previous reports that a positive correlation was found between sequence conservation and expression levels during miRNA evolution (*Dannemann et al., 2012b*). Newly emerged miRNAs are often expressed at lower levels compared with broadly conserved miRNAs (*Wang et al., 2016*), which suggest that miRNAs with low levels of conservation and expression may still be under evolutionary selection. Furthermore, the number and expression of miRNAs increased simultaneously during evolution, consistent with that observed in previous studies (*Meunier et al., 2013*; *Berezikov et al., 2011*; *Lee, Risom & Strauss, 2007*). The miRNA expansion may correlate with organismal complexity and body-plan innovation (*Meunier et al., 2013*; *Christodoulou et al., 2010*; *Mohammed et al., 2018*; *Grimson et al., 2008*).

We also systematically investigated the miRNA gene locations in the cynomolgus macaque genome. Unlike protein-coding genes, the locations of miRNA genes were highly polytropic. Most miRNAs resided in intergenic spaces or within introns of the hosts. Very few of miRNAs were found embedded in exons, overlapping-splice-sites, or antisense regions. Such unusual genomic arrangements of the miRNAs intensively raise our attention to explore their possible regulatory functions. For example, the *MIR1306* gene overlaps the DGCR8 coding region and the *MIR675* gene is embedded in the first exon of H19 lncRNA, both of which are conserved in primates (*Morin et al., 2008*; *Friedländer et al., 2008*; *Keniry et al., 2012*; *Wang et al., 2018*). It has been reported that the mir-1306 hairpin is cleaved by the Drosha-DGCR8 complex, which plays an important negative-feedback role in controlling DGCR8 expression (*Morin et al., 2008*; *Friedländer et al., 2008*; *Keniry et al., 2012*; *Wang et al., 2018*). *H19* is an important imprinted gene involved in placental development prior to birth, which limits the growth of the placenta by regulating miR-675 processes (*Keniry et al., 2012*; *Wang et al., 2018*; *Han et al., 2009*). Thus, it can be inferred that there is an expressive and functional interrelationship between miRNAs and host genes. For example, host genes can act as miRNA reservoirs and miRNA can negatively regulate its host. Among the exon-resident miRNAs identified in this study, it would be interesting

to investigate whether such exonic miRNA hairpins serve dual roles as miRNA precursors as well as RNA instability elements as miRNA hairpin processing could destabilize host transcripts, as mentioned above. We also found some pre-miRNAs overlapping intron-exon junctions. These SO-miRNAs may be involved in regulating gene expression in cynomolgus macaques, whereby microprocessor complex-dependent cleavage of SO-miRNA exons could result in premature transcriptional termination of coding genes, as described in humans (*Pianigiani et al., 2018*). Two pairs of distinct antisense miRNAs were found in this study. The sequences and locations of miR-3120/miR-214 pair were conserved in mammals. Using miRNA target prediction software, we found that miR-3120-3p was predicted to interact with the 3′ UTR of heat shock protein family A member 5 (HSPA5) and miR-214-3p was predicted to target the 3′ UTR of phosphatase and tensin homolog (PTEN), consistent with previous research in humans (*Scott et al., 2012*). HSPA5 is a chaperone for protein folding in the endoplasmic reticulum and PTEN interacts with the mir-3120 host gene dynamin-3 to regulate synaptic proteins involved in receptor cycling and synaptic plasticity (*Yang et al., 2008*). The data suggest that the miR-3120/miR-214 pair represent a genetic unit in cynomolgus macaques that evolved to regulate the complex neuronal pathways associated with synaptic vesicle function and neuronal plasticity (*Scott et al., 2012*). For the miR-24-1/miR-3074 antisense miRNAs, although they have different seed sequences, KEGG analysis indicated that the targets of miR-24-3p and miR-3074-5p contribute to similar biological pathways, including Axon guidance and EGFR tyrosine kinase inhibitor resistance (Table S4). The miR-24-1/miR-3074 antisense miRNAs may regulate essential aspects of cellular function in coordination (*Cordero et al., 2015*). These overlapping regulatory functions indicate that direct or indirect coordinated regulation exist between these antisense miRNA pairs. Future research may identify other miRNAs resident in unusual genomic regions, such as exon-resident miRNAs, SO-miRNAs, and antisense miRNAs, which work in coordination with their host genes to regulate diverse and complex cellular functions.

Another characterization of miRNA genomic locations is that many are clustered together. In total, 78 miRNA clusters were found in this study, 68 of which were conserved in the human genome, including 10 large clusters with no less than five miRNA precursors. It was reported that these 10 large miRNA clusters have significant functions in humans. For example, miRNAs in the mir-17/92 cluster are potential human oncogenes (*He et al., 2005*) and critical regulators of normal development and disease, such as body height and cardiomyocyte proliferation (*Chen et al., 2013*; *Bai et al., 2019*). The X-linked mir-506/514 cluster has regulatory roles in testis development and spermatogenesis (*Zhang et al., 2007*). Furthermore, the very large mir-379/mir-656 and mir-512/mir-517 clusters both participate in development. The former, known as the chromosome 14 miRNA cluster (C14MC), is comprised of 42 miRNA genes in humans and is located at the DLK-DIO3 imprinted domain (*Seitz et al., 2004*; *Morales-Prieto et al., 2013*; *Schmidt, Block & Golos, 2018*; *Sadovsky et al., 2015*; *Cavaillé et al., 2002*; *Glazov et al., 2008*). The latter, known as the chromosome 19 miRNA cluster (C19MC), is comprised of 46 miRNA genes in humans and is imprinted in the placenta (*Noguer-Dance et al., 2010*). C14MC is reported to not only play important roles in the control of neurogenesis, embryonic development, and

pregnancy, but is also associated with progression of hematopoietic and solid tumors (*Haller et al., 2010*; *Lavon et al., 2010*; *Cortez & Calin, 2009*; *Grün et al., 2005*; *Donker et al., 2012*; *Morales-Prieto et al., 2013*). In addition, variations in the expression of miRNAs in C19MC are associated with embryogenesis and tumorigenesis, suggesting a critical role of these miRNAs in development and cancer (*Noguer-Dance et al., 2010*; *Tsai et al., 2009*; *Li et al., 2009a*; *Kleinman et al., 2014*; *Ward et al., 2014*). Eight of the 10 miRNA-clusters in macaques are with conserved compositions to that in humans, indicating that their functions are evolutionarily conserved. The other two miRNA-clusters, the mir-379/mir-656 and mir-512/mir-517, showed minor changes to that in humans and then were used for functional prediction. The results indicated that both of the large miRNA-clusters regulate normal growth, development, and pathological processes through many pathways, consistent with previous studies on humans. Therefore, the miRNA clusters are important as they are involved in vital biological functions, including physiological and pathological processes, by targeting overlapping sets of genes cooperatively (*Marco et al., 2013*; *Wang et al., 2016*; *Kim & Nam, 2006*).

CONCLUSIONS

In summary, 1,455 high-confidence miRNA gene loci were identified in the genome of cynomolgus macaques, 408 of which were also confirmed by RNA-seq. Due to the stringent annotation criteria and limited sequenced samples, we just obtained a subset of miRNAs in macaque. The miRNA genes in cynomolgus macaque were distributed in diverse locations. Very few of miRNAs, such as exon-resident miRNAs, SO-miRNAs, and antisense miRNAs, were predicated to regulate diverse functions in coordination with their host genes. Simultaneously, miRNAs processed from the large clusters were predicted to have essential cellular and developmental functions. This study not only expands the number of identified miRNAs in cynomolgus macaques but also provides clues for future research to improve our understanding of the evolution and function of miRNAs in primate. In the future, a great number of macaque miRNAs need to be verified by experiments. Further investigations on the functional role of primate-specific miRNAs are also needed because evolutionarily young miRNAs may correlate with organismal complexity and body-plan innovation and have the potential for development as novel biomarkers or therapeutic targets in complex diseases, especially in neuropsychiatric disorders.

ACKNOWLEDGEMENTS

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

Competing Interests

Xiaoming Liu is an employee of the Guangzhou Tulip information Technologies Ltd. All authors declare that they have no competing interests.

Author Contributions

- Xia Huang and Shijia Li performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Xiaoming Liu and Shuting Huang performed the experiments, analyzed the data, prepared figures and/or tables, and approved the final draft.
- Shuang Li analyzed the data, prepared figures and/or tables, and approved the final draft.
- Min Zhuo conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.

Ethics

The following information was supplied relating to ethical approvals (i.e., approving body and any reference numbers):

The experiments were reviewed and approved by the Institutional Animal Care and Use Committee (IACUC) of Guangdong Landau Biotechnology Co. Ltd. (project number: IACUC-003).

DNA Deposition

The following information was supplied regarding the deposition of DNA sequences:

The 73 new miRNA gene loci in cynomolgus macaques detected in this study were processed into 88 mature miRNAs. Mature miRNAs are available at the European Nucleotide Archive: LR594712-LR594765, LR745789-LR745838. This data is also available as a Supplemental File.

Data Availability

The following information was supplied regarding data availability:

The raw data are available as a Supplemental File.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.9347#supplemental-information.

REFERENCES

Altuvia Y, Landgraf P, Lithwick G, Elefant N, Pfeffer S, Aravin A, Brownstein M.J, Tuschl T, Margalit H. 2005. Clustering and conservation patterns of human microRNAs. *Nucleic Acids Research* **33**(**8**):2697–2706 DOI 10.1093/nar/gki567.

Awan HM, Shah A, Rashid F, Shan G. 2017. Primate-specific Long Non-coding RNAs and MicroRNAs. *Genomics Proteomics Bioinformatics* **15**(**3**):187–195 DOI 10.1016/j.gpb.2017.04.002.

Bai X, Hua S, Zhang J, Xu S. 2019. The MicroRNA family both in normal development and in different diseases: The miR-17-92 cluster. *BioMed Research International* **2019**:9450240 DOI 10.1155/2019/9450240.

Bartel DP. 2009. MicroRNAs: target recognition and regulatory functions. *Cell* **136**(**2**):215–233 DOI 10.1016/j.cell.2009.01.002.

Bartel DP. 2018. Metazoan microRNAs. *Cell* **173**(**1**):20–51 DOI 10.1016/j.cell.2018.03.006.

Benson DA, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2013. GenBank. *Nucleic Acids Research* **41**(**Database issue**):D36–D42 DOI 10.1093/nar/gks1195.

Bentwich I, Avniel A, Karov Y, Aharonov R, Gilad S, Barad O, Barzilai A, Einat P, Einav U, Meiri E, Sharon E, Spector Y, Bentwich Z. 2005. Identification of hundreds of conserved and nonconserved human microRNAs. *Nature Genetics* **37**(**7**):766–770 DOI 10.1038/ng1590.

Berezikov E. 2011. Evolution of microRNA diversity and regulation in animals. *Nature Reviews Genetics* **12**(**12**):846–860 DOI 10.1038/nrg3079.

Berezikov E, Robine N, Samsonova A, Westholm JO, Naqvi A, Hung JH, Okamura K, Dai Q, Bortolamiol-Becet D, Martin R, Zhao Y, Zamore PD, Hannon GJ, Marra MA, Weng Z, Perrimon N, Lai EC. 2011. Deep annotation of Drosophila melanogaster microRNAs yields insights into their processing, modification, and emergence. *Genome Research* **21**(**2**):203–215 DOI 10.1101/gr.116657.110.

Berezikov E, Thuemmler F, Van Laake LW, Kondova I, Bontrop R, Cuppen E, Plasterk RH. 2006. Diversity of microRNAs in human and chimpanzee brain. *Nature Genetics* **38**(**12**):1375–1377 DOI 10.1038/ng1914.

Carter A.M. 2011. Comparative studies of placentation and immunology in non-human primates suggest a scenario for the evolution of deep trophoblast invasion and an explanation for human pregnancy disorders. *Reproduction* **141**(**4**):391–396 DOI 10.1530/REP-10-0530.

Cavaillé J, Seitz H, Paulsen M, Ferguson-Smith AC, Bachellerie JP. 2002. Identification of tandemly-repeated C/D snoRNA genes at the imprinted human 14q32 domain reminiscent of those at the Prader-Willi/Angelman syndrome region. *Human Molecular Genetics* **11**(**13**):1527–1538 DOI 10.1093/hmg/11.13.1527.

Chen J, Huang ZP, Seok HY, Ding J, Kataoka M, Zhang Z, Hu X, Wang G, Lin Z, Wang S, Pu WT, Liao R, Wang DZ. 2013. mir-17-92 cluster is required for and sufficient to induce cardiomyocyte proliferation in postnatal and adult hearts. *Circulation Research* **112**(**12**):1557–1566 DOI 10.1161/CIRCRESAHA.112.300658.

Christodoulou F, Raible F, Tomer R, Simakov O, Trachana K, Klaus S, Snyman H, Hannon G.J, Bork P, Arendt D. 2010. Ancient animal microRNAs and the evolution of tissue identity. *Nature* **463(7284)**:1084–1088 DOI 10.1038/nature08744.

Cordero F, Ferrero G, Polidoro S, Fiorito G, Campanella G, Sacerdote C, Mattiello A, Masala G, Agnoli C, Frasca G, Panico S, Palli D, Krogh V, Tumino R, Vineis P, Naccarati A. 2015. Differentially methylated microRNAs in prediagnostic samples of subjects who developed breast cancer in the European Prospective Investigation into Nutrition and Cancer (EPIC-Italy) cohort. *Carcinogenesis* **36(10)**:1144–1153 DOI 10.1093/carcin/bgv102.

Cortez MA, Calin GA. 2009. MicroRNA identification in plasma and serum a new tool to diagnose and monitor diseases. *Expert Opinion on Biological Therapy* **9(6)**:703–711 DOI 10.1517/14712590902932889.

Dannemann M, Nickel B, Lizano E, Burbano HA, Kelso J. 2012a. Annotation of primate miRNAs by high throughput sequencing of small RNA libraries. *BMC Genomics* **13**:116 DOI 10.1186/1471-2164-13-116.

Dannemann M, Prüfer K, Lizano E, Nickel B, Burbano HA, Kelso J. 2012b. Transcription factors are targeted by differentially expressed miRNAs in primates. *Genome Biology and Evolution* **4(4)**:552–564 DOI 10.1093/gbe/evs033.

Deline B, Greenwood JM, Clark JW, Puttick MN, Peterson KJ, Donoghue PCJ. 2018. Evolution of metazoan morphological disparity. *Proceedings of the National Academy of Sciences of the United States of America* **115(38)**:E8909–E8918 DOI 10.1073/pnas.1810575115.

Saçar Demirci MD, Baumbach J, Allmer J. 2017. On the performance of pre-microRNA detection algorithms. *Nature Communications* **8(1)**:330 DOI 10.1038/s41467-017-00403-z.

Donker RB, Mouillet JF, Chu T, Hubel CA, Stolz DB, Morelli AE, Sadovsky Y. 2012. The expression profile of C19MC microRNAs in primary human trophoblast cells and exosomes. *Molecular Human Reproduction* **18(8)**:417–424 DOI 10.1093/molehr/gas013.

Dostie J, Mourelatos Z, Yang M, Sharma A, Dreyfuss G. 2003. Numerous microRNPs in neuronal cells containing novel microRNAs. *RNA* **9(2)**:180–186 DOI 10.1261/rna.2141503.

Ebeling M, Küng E, See A, Broger C, Steiner G, Berrera M, Heckel T, Iniguez L, Albert T, Schmucki R, Biller H, Singer T, Certa U. 2011. Genome-based analysis of the nonhuman primate Macaca fascicularis as a model for drug safety assessment. *Genome Research* **21(10)**:1746–1756 DOI 10.1101/gr.123117.111.

Friedländer M.R, Chen W, Adamidi C, Maaskola J, Einspanier R, Knespel S, Rajewsky N. 2008. Discovering microRNAs from deep sequencing data using miRDeep. *Nature Biotechnology* **26(4)**:407–415 DOI 10.1038/nbt1394.

Friedman RC, Farh KK, Burge CB, Bartel DP. 2009. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Research* **19(1)**:92–105 DOI 10.1101/gr.082701.108.

Gebert LFR, MacRae IJ. 2019. Regulation of microRNA function in animals. *Nature Reviews Molecular Cell Biology* **20(1)**:21–37 DOI 10.1038/s41580-018-0045-7.

Glazov EA, McWilliam S, Barris WC, Dalrymple BP. 2008. Origin, evolution, and biological role of miRNA cluster in DLK-DIO3 genomic region in placental mammals. *Molecular Biology and Evolution* **25(5)**:939–948 DOI 10.1093/molbev/msn045.

Grigsby PL. 2016. Animal models to study placental development and function throughout normal and dysfunctional human pregnancy. *Seminars in Reproductive Medicine* **34(1)**:11–16 DOI 10.1055/s-0035-1570031.

Grimson A, Srivastava M, Fahey B, Woodcroft BJ, Chiang HR, King N, Degnan BM, Rokhsar D.S, Bartel DP. 2008. The early origins of microRNAs and Piwi-interacting RNAs in animals. *Nature* **455(7217)**:1193–1197 DOI 10.1038/nature07415.

Grün D, Wang YL, Langenberger D, Gunsalus KC, Rajewsky N. 2005. microRNA target predictions across seven Drosophila species and comparison to mammalian targets. *PLOS Computational Biology* **1(1)**:e13 DOI 10.1371/journal.pcbi.0010013.

Haller F, Von Heydebreck A, Zhang JD, Gunawan B, Langer C, Ramadori G, Wiemann S, Sahin O. 2010. Localization- and mutation-dependent microRNA (miRNA) expression signatures in gastrointestinal stromal tumours (GISTs), with a cluster of co-expressed miRNAs located at 14q32. *Journal of Pathology* **220(1)**:71–86 DOI 10.1002/path.2610.

Han J, Pedersen JS, Kwon SC, Belair CD, Kim YK, Yeom KH, Yang WY, Haussler D, Blelloch R, Kim VN. 2009. Posttranscriptional crossregulation between Drosha and DGCR8. *Cell* **136(1)**:75–84 DOI 10.1016/j.cell.2008.10.053.

He L, Thomson JM, Hemann MT, Hernando-Monge E, Mu D, Goodson S, Powers S, Cordon-Cardo C, Lowe SW, Hannon GJ, Hammond SM. 2005. A microRNA polycistron as a potential human oncogene. *Nature* **435(7043)**:828–833 DOI 10.1038/nature03552.

Heimberg AM, Sempere LF, Moy VN, Donoghue PC, 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(8)**:2946–2950 DOI 10.1073/pnas.0712259105.

Iwama H, Kato K, Imachi H, Murao K, Masaki T. 2013. Human microRNAs originated from two periods at accelerated rates in mammalian evolution. *Molecular Biology and Evolution* **30(3)**:613–626 DOI 10.1093/molbev/mss262.

Jimenez SA, Piera-Velazquez S. 2013. Potential role of human-specific genes, human-specific microRNAs and human-specific non-coding regulatory RNAs in the pathogenesis of systemic sclerosis and Sjogren's syndrome. *Autoimmunity Reviews* **12(11)**:1046–1051 DOI 10.1016/j.autrev.2013.04.004.

Jonas S, Izaurralde E. 2015. Towards a molecular understanding of microRNA-mediated gene silencing. *Nature Reviews Genetics* **16(7)**:421–433 DOI 10.1038/nrg3965.

Kalvari I, Argasinska J, Quinones-Olvera N, Nawrocki EP, Rivas E, Eddy SR, Bateman A, Finn RD, Petrov AI. 2018. Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families. *Nucleic Acids Research* **46(D1)**:D335–D342 DOI 10.1093/nar/gkx1038.

Keniry A, Oxley D, Monnier P, Kyba M, Dandolo L, Smits G, Reik W. 2012. The H19 lincRNA is a developmental reservoir of miR-675 that suppresses growth and Igf1r. *Nature Cell Biology* **14(7)**:659–665 DOI 10.1038/ncb2521.

Kent WJ. 2002. BLAT–the BLAST-like alignment tool. *Genome Research* **12(4)**:656–664 DOI 10.1101/gr.229202.

Kim VN, Nam JW. 2006. Genomics of microRNA. *Trends in Genetics* **22(3)**:165–173 DOI 10.1016/j.tig.2006.01.003.

Kim YK, Kim VN. 2007. Processing of intronic microRNAs. *EMBO Journal* **26(3)**:775–783 DOI 10.1038/sj.emboj.7601512.

Kleinman CL, Gerges N, Papillon-Cavanagh S, Sin-Chan P, Pramatarova A, Quang DA, Adoue V, Busche S, Caron M, Djambazian H, Bemmo A, Fontebasso AM, Spence T, Schwartzentruber J, Albrecht S, Hauser P, Garami M, Klekner A, Bognar L, Montes JL, Staffa A, Montpetit A, Berube P, Zakrzewska M, Zakrzewski K, Liberski PP, Dong Z, Siegel PM, Duchaine T, Perotti C, Fleming A, Faury D, Remke M, Gallo M, Dirks P, Taylor M.D, Sladek R, Pastinen T, Chan JA, Huang A, Majewski J, Jabado N. 2014. Fusion of TTYH1 with the C19MC microRNA cluster drives expression of a brain-specific DNMT3B isoform in the embryonal brain tumor ETMR. *Nature Genetics* **46(1)**:39–44 DOI 10.1038/ng.2849.

Kobayashi H, Tomari Y. 2016. RISC assembly: coordination between small RNAs and Argonaute proteins. *Biochimica et Biophysica Acta/General Subjects* **1859(1)**:71–81 DOI 10.1016/j.bbagrm.2015.08.007.

Koufaris C. 2016. Human and primate-specific microRNAs in cancer: evolution, and significance in comparison with more distantly-related research models: The great potential of evolutionary young microRNA in cancer research. *Bioessays* **38(3)**:286–294 DOI 10.1002/bies.201500135.

Koufaris C, Gooderham NJ. 2013. Are differences in microRNA regulation implicated in species-dependent response to toxicological exposures? *Toxicological Sciences* **131(2)**:337–342 DOI 10.1093/toxsci/kfs302.

Kozomara A, Birgaoanu M, Griffiths-Jones S. 2019. miRBase: from microRNA sequences to function. *Nucleic Acids Research* **47(D1)**:D155–D162 DOI 10.1093/nar/gky1141.

Kozomara A, Griffiths-Jones S. 2014. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Research* **42(Database issue)**:D68–D73 DOI 10.1093/nar/gkt1181.

Lavon I, Zrihan D, Granit A, Einstein O, Fainstein N, Cohen MA, Cohen MA, Zelikovitch B, Shoshan Y, Spektor S, Reubinoff BE, Felig Y, Gerlitz O, Ben-Hur T, Smith Y, Siegal T. 2010. Gliomas display a microRNA expression profile reminiscent of neural precursor cells. *Neuro-Oncology* **12(5)**:422–433 DOI 10.1093/neuonc/nop061.

Lee CT, Risom T, Strauss WM. 2007. Evolutionary conservation of microRNA regulatory circuits: an examination of microRNA gene complexity and conserved microRNA-target interactions through metazoan phylogeny. *DNA and Cell Biology* **26(4)**:209–218 DOI 10.1089/dna.2006.0545.

Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, Lee J, Provost P, Rådmark O, Kim S, Kim VN. 2003. The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425(6956)**:415–419 DOI 10.1038/nature01957.

Lehnert S, Van Loo P, Thilakarathne PJ, Marynen P, Verbeke G, Schuit FC. 2009. Evidence for co-evolution between human microRNAs and Alu-repeats. *PLOS ONE* **4(2)**:e4456 DOI 10.1371/journal.pone.0004456.

Li M, Lee KF, Lu Y, Clarke I, Shih D, Eberhart C, Collins VP, Van Meter T, Picard D, Zhou L, Boutros PC, Modena P, Liang ML, Scherer SW, Bouffet E, Rutka JT, Pomeroy SL, Lau CC, Taylor MD, Gajjar A, Dirks PB, Hawkins CE, Huang A. 2009a. Frequent amplification of a chr19q13.41 microRNA polycistron in aggressive primitive neuroectodermal brain tumors. *Cancer Cell* **16(6)**:533–546 DOI 10.1016/j.ccr.2009.10.025.

Li R, Yu C, Li Y, Lam TW, Yiu SM, Kristiansen K, Wang J. 2009b. SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics* **25(15)**:1966–1967 DOI 10.1093/bioinformatics/btp336.

Liang T, Guo L, Liu C. 2012. Genome-wide analysis of mir-548 gene family reveals evolutionary and functional implications. *Journal of Biomedicine and Biotechnology* **2012**:679563 DOI 10.1155/2012/679563.

Lorenz R, Bernhart SH, HönerZu Siederdissen C, Tafer H, Flamm C, Stadler PF, Hofacker IL. 2011. ViennaRNA package 2.0. *Algorithms for Molecular Biology* **6**:26 DOI 10.1186/1748-7188-6-26.

Marco A, Ninova M, Ronshaugen M, Griffiths-Jones S. 2013. Clusters of microR-NAs emerge by new hairpins in existing transcripts. *Nucleic Acids Research* **41(16)**:7745–7752 DOI 10.1093/nar/gkt534.

McCreight JC, Schneider SE, Wilburn DB, Swanson WJ. 2017. Evolution of microRNA in primates. *PLOS ONE* **12(6)**:e0176596 DOI 10.1371/journal.pone.0176596.

Meunier J, Lemoine F, Soumillon M, Liechti A, Weier M, Guschanski K, Hu H, Khaitovich P, Kaessmann H. 2013. Birth and expression evolution of mammalian microRNA genes. *Genome Research* **23(1)**:34–45 DOI 10.1101/gr.140269.112.

Mohammed J, Flynt AS, Panzarino AM, Mondal MMH, DeCruz M, Siepel A, Lai EC. 2018. Deep experimental profiling of microRNA diversity, deployment, and evolution across the Drosophila genus. *Genome Research* **28(1)**:52–65 DOI 10.1101/gr.226068.117.

Morales-Prieto DM, Ospina-Prieto S, Chaiwangyen W, Schoenleben M, Markert UR. 2013. Pregnancy-associated miRNA-clusters. *Journal of Reproductive Immunology* **97(1)**:51–61 DOI 10.1016/j.jri.2012.11.001.

Morin RD, O'Connor MD, Griffith M, Kuchenbauer F, Delaney A, Prabhu AL, Zhao Y, McDonald H, Zeng T, Hirst M, Eaves CJ, Marra MA. 2008. Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Research* **18(4)**:610–621 DOI 10.1101/gr.7179508.

Noguer-Dance M, Abu-Amero S, Al-Khtib M, Lefèvre A, Coullin P, Moore GE, Cavaillé J. 2010. The primate-specific microRNA gene cluster (C19MC) is imprinted in the placenta. *Human Molecular Genetics* **19(18)**:3566–3582 DOI 10.1093/hmg/ddq272.

Perelman P, Johnson WE, Roos C, Seuánez HN, Horvath JE, Moreira MA, Kessing B, Pontius J, Roelke M, Rumpler Y, Schneider MP, Silva A, O'Brien SJ, Pecon-Slattery J. 2011. A molecular phylogeny of living primates. *PLOS Genetics* **7**(**3**):e1001342 DOI 10.1371/journal.pgen.1001342.

Pianigiani G, Licastro D, Fortugno P, Castiglia D, Petrovic I, Pagani F. 2018. Microprocessor-dependent processing of splice site overlapping microRNA exons does not result in changes in alternative splicing. *RNA* **24**(**9**):1158–1171 DOI 10.1261/rna.063438.117.

Picao-Osorio J, Johnston J, Landgraf M, Berni J, Alonso CR. 2015. MicroRNA-encoded behavior in Drosophila. *Science* **350**(**6262**):815–820 DOI 10.1126/science.aad0217.

Pinhal D, Bovolenta LA, Moxon S, Oliveira AC, Nachtigall PG, Acencio ML, Patton JG, Hilsdorf AWS, Lemke N, Martins C. 2018. Genome-wide microRNA screening in Nile tilapia reveals pervasive isomiRs' transcription, sex-biased arm switching and increasing complexity of expression throughout development. *Scientific Reports* **8**(**1**):8248 DOI 10.1038/s41598-018-26607.

Rodriguez A, Griffiths-Jones S, Ashurst JL, Bradley A. 2004. Identification of mammalian microRNA host genes and transcription units. *Genome Research* **14**(**10A**):1902–1910 DOI 10.1101/gr.2722704.

Rorbach G, Unold O, Konopka BM. 2018. Distinguishing mirtrons from canonical miRNAs with data exploration and machine learning methods. *Scientific Reports* **8**(**1**):7560 DOI 10.1038/s41598-018-25578-3.

Ryazansky SS, Gvozdev VA, Berezikov E. 2011. Evidence for post-transcriptional regulation of clustered microRNAs in Drosophila. *BMC Genomics* **12**:371 DOI 10.1186/1471-2164-12-371.

Sadovsky Y, Mouillet JF, Ouyang Y, Bayer A, Coyne CB. 2015. The function of TrophomiRs and other MicroRNAs in the human placenta. *Cold Spring Harbor Perspectives in Medicine* **5**(**8**):a023036 DOI 10.1101/cshperspect.a023036.

Saini HK, Griffiths-Jones S, Enright AJ. 2007. Genomic analysis of human microRNA transcripts. *Proceedings of the National Academy of Sciences of the United States of America* **104**(**45**):17719–17724 DOI 10.1073/pnas.0703890104.

Schmidt JK, Block LN, Golos TG. 2018. Defining the rhesus macaque placental miRNAome: conservation of expression of placental miRNA clusters between the macaque and human. *Placenta* **65**:55–64 DOI 10.1016/j.placenta.2018.04.003.

Scott H, Howarth J, Lee YB, Wong LF, Bantounas I, Phylactou L, Verkade P, Uney JB. 2012. MiR-3120 is a mirror MicroRNA that targets heat shock cognate protein 70 and auxilin messenger RNAs and regulates clathrin vesicle uncoating. *Journal of Biological Chemistry* **287**(**18**):14726–14733 DOI 10.1074/jbc.M111.326041.

Seitz H, Royo H, Bortolin ML, Lin SP, Ferguson-Smith AC, Cavaillé J. 2004. A large imprinted microRNA gene cluster at the mouse Dlk1-Gtl2 domain. *Genome Research* **14**(**9**):1741–1748 DOI 10.1101/gr.2743304.

Shiina T, Yamada Y, Aarnink A, Suzuki S, Masuya A, Ito S, Ido D, Yamanaka H, Iwatani C, Tsuchiya H, Ishigaki H, Itoh Y, Ogasawara K, Kulski JK, Blancher A. 2015. Discovery of novel MHC-class I alleles and haplotypes in Filipino

cynomolgus macaques (Macaca fascicularis) by pyrosequencing and Sanger sequencing: Mafa-class I polymorphism. *Immunogenetics* **67**(**10**):563–578 DOI 10.1007/s00251-015-0867-9.

Tan TT, Chen M, Harikrishna JA, Khairuddin N, Mohd Shamsudin MI, Zhang G, Bhassu S. 2013. Deep parallel sequencing reveals conserved and novel miRNAs in gill and hepatopancreas of giant freshwater prawn. *Fish and Shellfish Immunology* **35**(**4**):1061–1069 DOI 10.1016/j.fsi.2013.06.017.

Theriault BR, Thistlethwaite Jr JR, Levisetti MG, Wardrip CL, Szot G, Bruce DS, Rilo H, Li X, Gray GS, Bluestone JA, Padrid PA. 1999. Induction, maintenance, and reversal of streptozotocin-induced insulin-dependent diabetes mellitus in the juvenile cynomolgus monkey (Macaca fascilularis). *Transplantation* **68**(**3**):331–337 DOI 10.1097/00007890-199908150-00003.

Treiber T, Treiber N, Meister G. 2019. Regulation of microRNA biogenesis and its crosstalk with other cellular pathways. *Nature Reviews Molecular Cell Biology* **20**(**1**):5–20 DOI 10.1038/s41580-018-0059-1.

Tsai KM, Kao HW, Chen HC, Chen SJ, Lin WC. 2009. Epigenetic control of the expression of a primate-specific microRNA cluster in human cancer cells. *Epigenetics* **4**(**8**):587–592 DOI 10.4161/epi.4.8.10230.

Tyler DM, Okamura K, Chung WJ, Hagen JW, Berezikov E, Hannon GJ, Lai EC. 2008. Functionally distinct regulatory RNAs generated by bidirectional tran-scription and processing of microRNA loci. *Genes and Development* **22**(**1**):26–36 DOI 10.1101/gad.1615208.

Uhlmann S, Mannsperger H, Zhang JD, Horvat EÁ, Schmidt C, Küblbeck M, Henjes F, Ward A, Tschulena U, Zweig K, Korf U, Wiemann S, Sahin O. 2012. Global microRNA level regulation of EGFR-driven cell-cycle protein network in breast cancer. *Molecular Systems Biology* **8**:570 DOI 10.1038/msb.2011.100.

Veeranagouda Y, Rival P, Prades C, Mariet C, Léonard JF, Gautier JC, Zhou X, Wang J, Li B, Ozoux ML, Boitier E. 2015. Identification of microRNAs in Macaca fascicularis (Cynomolgus Monkey) by Homology Search and Experimental Validation by Small RNA-Seq and RT-qPCR Using Kidney Cortex Tissues. *PLOS ONE* **10**(**11**):e0142708 DOI 10.1371/journal.pone.0142708.

Wang F, Liang R, Tandon N, Matthews ER, Shrestha S, Yang J, Soibam B, Yang J, Liu Y. 2018. H19X-encoded miR-424(322)/-503 cluster: emerging roles in cell differentiation, proliferation, plasticity and metabolism. *Cellular and Molecular Life Sciences* **76**(**5**):903–920 DOI 10.1007/s00018-018-2971-0.

Wang Y, Luo J, Zhang H, Lu J. 2016. microRNAs in the same clusters evolve to co-ordinately regulate functionally related genes. *Molecular Biology and Evolution* **33**(**9**):2232–2247 DOI 10.1093/molbev/msw089.

Ward A, Shukla K, Balwierz A, Soons Z, König R, Sahin O, Wiemann S. 2014. MicroRNA-519a is a novel oncomir conferring tamoxifen resistance by targeting a network of tumour-suppressor genes in ER+ breast cancer. *Journal of Pathology* **233**(**4**):368–379 DOI 10.1002/path.4363.

Willard SL, Shively CA. 2012. Modeling depression in adult female cynomolgus monkeys (Macaca fascicularis). *American Journal of Primatology* **74(6)**:528–542 DOI 10.1002/ajp.21013.

Xu Y, Li W, Liu X, Chen H, Tan K, Chen Y, Tu Z, Dai Y. 2013. Identification of dys-regulated microRNAs in lymphocytes from children with Down syndrome. *Gene* **530(2)**:278–286 DOI 10.1016/j.gene.2013.07.055.

Yan G, Zhang G, Fang X, Zhang Y, Li C, Ling F, Cooper DN, Li Q, Li Y, van Gool AJ, Du H, Chen J, Chen R, Zhang P, Huang Z, Thompson JR, Meng Y, Bai Y, Wang J, Zhuo M, Wang T, Huang Y, Wei L, Li J, Wang Z, Hu H, Yang P, Le L, Stenson PD, Li B, Liu X, Ball EV, An N, Huang Q, Zhang Y, Fan W, Zhang X, Li Y, Wang W, Katze MG, Su B, Nielsen R, Yang H, Wang J, Wang X, Wang J. 2011. Genome sequencing and comparison of two nonhuman primate animal models, the cynomolgus and Chinese rhesus macaques. *Nature Biotechnology* **29(11)**:1019–1023 DOI 10.1038/nbt.1992.

Yang H, Kong W, He L, Zhao JJ, O'Donnell JD, Wang J, Wenham RM, Coppola D, Kruk PA, Nicosia SV, Cheng JQ. 2008. MicroRNA expression profiling in human ovarian cancer: miR-214 induces cell survival and cisplatin resistance by targeting PTEN. *Cancer Research* **68(2)**:425–433 DOI 10.1158/0008-5472.CAN-07-2488.

Yang H, Zhang H, Zhu L, Zhang C, Li D. 2012. Identification and Characterization of MicroRNAs in Macaca fascicularis by EST Analysis. *Comparative and Functional Genomics* **2012**:957607 DOI 10.1155/2012/957607.

Yang H, Zhang R, Jing Y, Zhu L, Zhang W, Liu C, Wang J, Yang J, Zhang J, Zen K, Zhang C, Li D. 2014. Identification and characterization of microRNAs in the crab-eating macaque (Macaca fascicularis) using transcriptome analysis. *Gene* **536(2)**:308–315 DOI 10.1016/j.gene.2013.12.006.

Yu G, Wang LG, Han Y, He QY. 2012. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16(5)**:284–287 DOI 10.1089/omi.2011.0118.

Yue J, Sheng Y, Orwig KE. 2008. Identification of novel homologous microRNA genes in the rhesus macaque genome. *BMC Genomics* **9**:8 DOI 10.1186/1471-2164-9-8.

Zhang R, Peng Y, Wang W, Su B. 2007. Rapid evolution of an X-linked microRNA cluster in primates. *Genome Research* **17(5)**:612–617 DOI 10.1101/gr.6146507.

Zhang R, Wang YQ, Su B. 2008. Molecular evolution of a primate-specific microRNA family. *Molecular Biology and Evolution* **25(7)**:1493–1502 DOI 10.1093/molbev/msn094.