

RESEARCH ARTICLE

The Role of Transposable Elements in the Origin and Evolution of MicroRNAs in Human

Sheng Qin¹, Ping Jin¹, Xue Zhou³, Liming Chen², Fei Ma^{1*}

1 Laboratory for Comparative Genomics and Bioinformatics & Jiangsu Key Laboratory for Biodiversity and Biotechnology, College of Life Science, Nanjing Normal University, Nanjing, China, **2** The Key Laboratory of Developmental Genes and Human Disease, Ministry of Education, Institute of Life Science, Southeast University, Nanjing, China, **3** School of Chemistry and Biological Engineering, NanJing Normal University TaiZhou College, TaiZhou, China

* mafei01@tsinghua.org.cn



CrossMark
click for updates

OPEN ACCESS

Citation: Qin S, Jin P, Zhou X, Chen L, Ma F (2015) The Role of Transposable Elements in the Origin and Evolution of MicroRNAs in Human. PLoS ONE 10(6): e0131365. doi:10.1371/journal.pone.0131365

Editor: Alfons Navarro, University of Barcelona, SPAIN

Received: February 13, 2015

Accepted: June 1, 2015

Published: June 26, 2015

Copyright: © 2015 Qin et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was jointly supported by grants from the National Natural Science Foundation of China (No. 30970348), the Major Program of Natural Science Research of Jiangsu Higher Education Institutions (No. 12KJA180005), the Ph.D. Programs Foundation of Ministry of Education of China (No. 20113207110009), and the Jiangsu Province Ordinary University Innovative Research Project (No. CXLX13-382).

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

Abstract

MicroRNAs (miRNAs) are crucial regulators of gene expression at the post-transcriptional level in eukaryotes via targeting gene 3'-untranslated regions. Transposable elements (TEs) are considered as natural origins of some miRNAs. However, what miRNAs are and how these miRNAs originate and evolve from TEs remain unclear. We identified 409 TE-derived miRNAs (386 overlapped with TEs and 23 un-overlapped with TEs) which are derived from TEs in human. This indicates that the TEs play important roles in origin of miRNAs in human. In addition, we found that the proportions of miRNAs derived from TEs (MDTEs) in human are more than other vertebrates especially non-mammal vertebrates. Furthermore, we classified MDTEs into three types and found that TE head or tail sequences along with adjacent genomic sequences contribute to generation of human miRNAs. Our current study will improve the understanding of origin and evolution of human miRNAs.

Introduction

Transposable elements (TEs) as important components of many genomes are able to mobilize and replicate in the host genomes [1]. There are two kinds of these elements: retrotransposons, and DNA transposons [2]. The retrotransposons can be further classified into three categories: long terminal repeat (LTR), long interspersed nuclear element (LINE) and short interspersed nuclear element (SINE). TEs were claimed to be an evolutionary force and were found to be related to epigenetic regulatory mechanisms [3–5]. In addition, TE sequences are able to provide TF binding sites during gene expression, and change regulatory networks of gene expression [6, 7].

The discovery of small RNA led to an emerging of prospect for uncovering the functions of TEs. Various small RNAs have been discovered, such as microRNAs (miRNAs), short interfering RNAs, piwi interacting RNAs and so on [8–10]. MiRNA is the first discovered small RNA [10]. MiRNAs are a class of short non-coding RNAs (approximately 22 nt) that are cleaved from longer (approximately 70 to 90 nt) precursor miRNAs (pre-miRNAs) [11, 12]. In animals, most miRNAs regulate gene expression by targeting mRNA-specific regions (known as

miRNA-target sites) via partially complementary manner [13, 14]. These target sites are mainly located in 3'-untranslated regions (3'-UTRs) and complement with miRNAs via sequences of approximately 7 nt at the 5' ends of miRNAs (known as 'seed' regions) [15, 16]. MiRNAs regulate gene expression by degrading mRNAs or repressing mRNA translation via recognizing their target sites [17]. Although many algorithms were developed to predict the target sites of miRNAs, the mechanism of miRNA recognition of target sites is not fully understood [15, 18–20]. Understanding the origin of miRNAs and their target sites will improve the rationale for developing algorithms to optimizing the prediction of miRNA-target genes.

TEs were claimed to provide a natural mechanism for the origin of new miRNAs and the targets of some miRNAs [21–25]. For instance, mir-28, mir-95 and mir-151 are derived from LINE-2 TEs, and mir-548 family is derived from Made1 TEs [21–23]. Alu elements of TEs could be targeted by almost 30 human miRNAs [25]. The hsa-mir-566 was found to be derived from Alu and 80% of its predicted target sites were claimed to be derived from TEs and related to Alu element [23]. However, it remains largely unknown what and how miRNAs originated from TEs in human.

In the current study, we provide evidences to show that TEs are important sources for the origin of miRNAs in human. Our results uncover the evolution of miRNAs derived from TEs in human and provide an insight into the mechanism of the origin of miRNAs.

Materials and Methods

To identify the miRNAs derived from TEs (MDTEs), pre-miRNAs and their associated data were collected and analyzed by following steps:

Firstly, 6845 pre-miRNAs with chromosomal locations of eight vertebrates (*Danio rerio*, *Xenopus tropicalis*, *Gallus gallus*, *Bos taurus*, *Mus musculus*, *Macaca mulatta*, *Pan troglodytes* and *Homo sapiens*) were obtained from miRBase v20 [26]. The pre-miRNAs and their adjacent upstream and downstream 4,000 bp sequences were downloaded using the BioMart tool from Ensembl genome database (Release 68) [27].

Secondly, the pre-miRNAs and their adjacent sequences were used as the query sequences to identify TEs. TEs which locate on the query sequences were identified using the RepeatMasker program based on the repeatmasker libraries 20140131 [28, 29]. The Wu-blast program was used as the searching engine of RepeatMasker and the parameter—s was set to improve the accuracy of identification. The locations of TEs on query sequences were extracted from the “.out” file which is one of the output files of RepeatMasker.

Finally, the locations of TEs were compared with those of pre-miRNAs on query sequences. If a pre-miRNA overlapped with a TE on the query sequence, this miRNA was defined to be a MDTE. The proportions of overlap between pre-miRNAs and TE sequences were calculated for classification of MDTE types.

To identify a human MDTE which lose its sequence feature of TE and has homologous relations of miRNAs among human and seven other vertebrates were analyzed and human MDTEs without TE sequence feature were identified if a human miRNA showing non-overlapping with a TE but has homologies from several other vertebrates overlapping with same TE sequence.

To address whether different TE families have equal contributions to the origin of MDTEs, the proportions of MDTEs generated from different TE families were calculated and compared with the proportions of TE families in human genome. The proportions of MDTEs generated from different TE families were calculated following the procedure described above. The proportions of TE families in human genome were obtained from the published genome sequencing data [30]. Pearson's Chi-squared test was used to evaluate the significance of differences

and $P < 0.01$ indicates that different TE families have different contributions to the origin of MDTEs significantly.

Results and Discussion

Identification of miRNAs derived from TEs in human and seven other vertebrates

The MDTEs were identified from human and seven other vertebrates (*Danio rerio*, *Xenopus tropicalis*, *Gallus gallus*, *Bos taurus*, *Mus musculus*, *Macaca mulatta* and *Pan troglodytes*). Surprisingly, non MDTEs were found in *Xenopus tropicalis*. Proportions of MDTEs in miRNAs increased with the evolution of vertebrates and the proportion in human was more than those in other analyzed vertebrates (Fig 1). Meanwhile, it was observed the proportions of MDTEs in miRNAs bear little relevance to the proportions of TEs in genomes. For example, although more than one-third of the genomes are made up by TEs in *Danio rerio* and *Xenopus tropicalis* [31, 32], the proportions of MDTEs in miRNAs are less than 5%. In comparison, TE sequences constitute 9% of genome in *Gallus gallus*, but 6.98% of miRNAs were MDTEs [33]. The MDTEs account for 19.84% of miRNAs in *Homo sapiens* and TE sequences make up 44.83% of its genome [30]. This observation might be due to the significant differences between the components of TEs in *Danio rerio* and *Xenopus tropicalis* and those in human and other mammals. This argument was supported, at least in part, by the observation that the major TEs are DNA types in *Danio rerio* and *Xenopus tropicalis* compared to retrotransposable elements in mammals [30–32, 34]. Given the contribution of TEs to miRNAs were negligible in *Drosophila* [35], MDTEs mainly present in genomes of human and other mammals. Information of MDTEs in human and seven other vertebrates was summarized and listed in Table 1.

When MDTEs were undergone homology analysis among *Danio rerio*, *Gallus gallus* and mammals, no homology of MDTEs among them were found. Fourteen MDTEs are conserved among five mammals and forty-seven MDTEs are conserved among primates. This finding implies that MDTEs are species-specific due to the difference of TEs among species.

Analysis of MDTEs in human

To further investigate the pattern of MDTEs in human, 1872 miRNA gene sequences of *Homo sapiens* collected from the miRBase v20 were mapped to the human genome and analyzed. In total, 386 MDTEs which completely or partly overlap with TEs show unique relationships to their related TEs. It can be demonstrated via observing the origin of multi-copy MDTEs or MDTE families. Each copy in multi-copy MDTEs or each member of a MDTE family was found to be originated from the same TE. For example, six copies of hsa-mir-3118 in human genome are all partly derived from LINE/L1PA13 and a large miRNA group, hsa-mir-548, is derived from DNA/MADE1 element. Taken together, our findings suggested that a MDTE and its homologies are derived from the same TE. S1 Table lists detailed information of all MDTEs.

When multi-copy MDTEs were excluded, 338 unique MDTEs (UMDTEs) were identified and can be classified into three types (Fig 2A): Type I UMDTEs derived from inverted TE sequences, Type II UMDTEs with sequences partly overlap with TE sequences that are not inverted, and Type III UMDTEs with sequences wholly overlap with TE sequences.

MiRNAs have been identified in various organisms with rapidly increasing number in databases. In humans, except for multi-copies, approximately 19.84% (338/1704) of miRNAs overlapping with TEs are regarded as UMDTEs. Inverted TEs have been claimed as an important configuration of miRNA origin [21, 22]. Consistently, 11.24% of UMDTEs were found to be derived from inverted TEs, 36.98% of the UMDTEs were found to be derived from whole TEs

The proportions of MDTEs in eight vertebrates

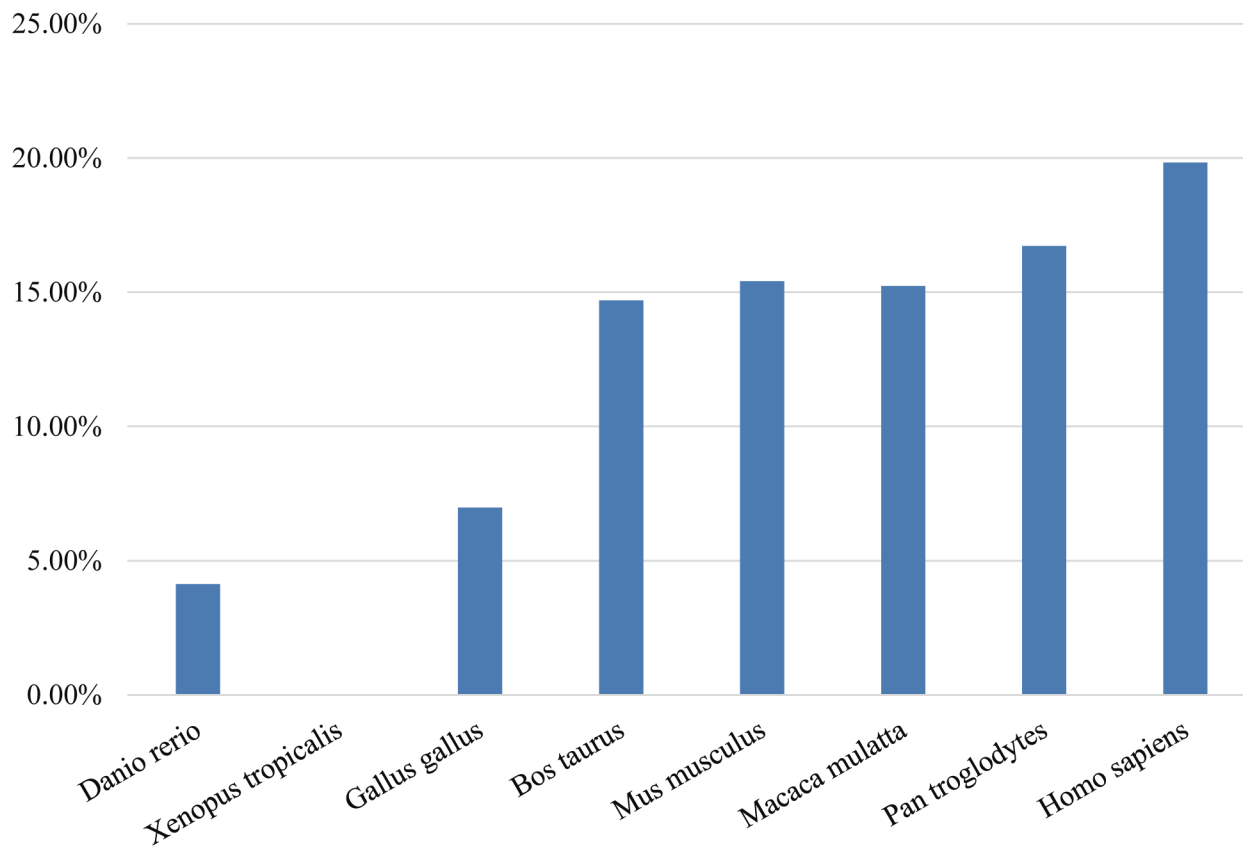


Fig 1. The proportions of MDTEs in human and seven other vertebrates.

doi:10.1371/journal.pone.0131365.g001

Table 1. Summary of MDTEs in eight vertebrates.

	Unique MDTEs	Unique miRNAs	All miRNAs
<i>Danio rerio</i>	9	218	346
<i>Xenopus tropicalis</i>	0	154	189
<i>Gallus gallus</i>	48	688	734
<i>Bos taurus</i>	103	701	798
<i>Mus musculus</i>	169	1097	1186
<i>Macaca mulatta</i>	85	558	615
<i>Pan troglodytes</i>	96	574	656
<i>Homo sapiens</i>	338	1704	1872

Unique MDTEs are the numbers of miRNAs which are derived from TEs excluded multi-copy miRNAs. Unique miRNAs are numbers of all miRNAs excluded multi-copy miRNAs. All miRNAs are total numbers of miRNAs in a species.

doi:10.1371/journal.pone.0131365.t001

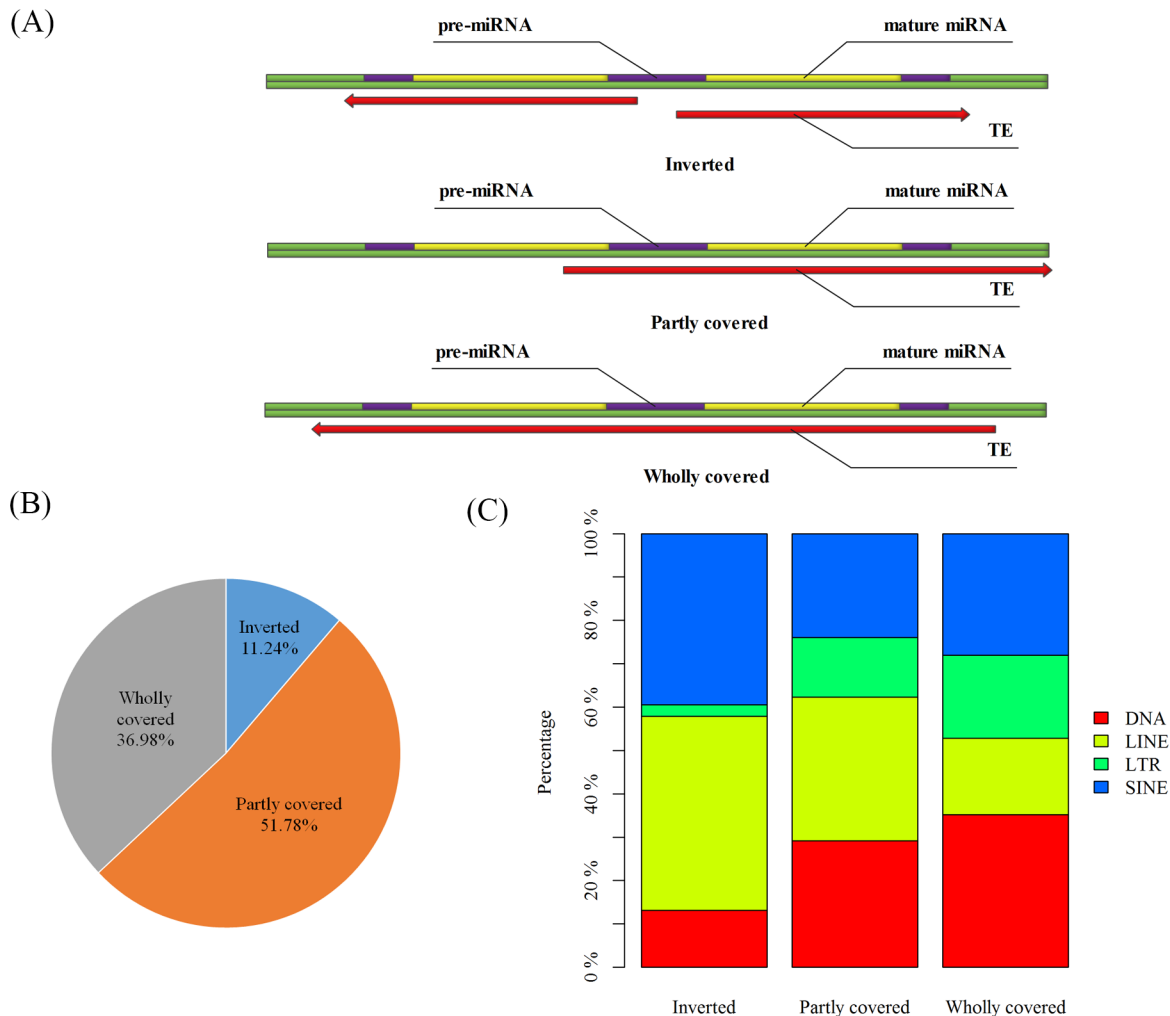


Fig 2. Analysis of three types of MDTEs. (A) Three types of MDTEs. Type I Inverted—The pre-miRNA is derived from two inverted TEs. **Type II Partly overlapped**—One arm is a TE while the other is the complementary sequence of that TE. Part of the pre-miRNA is derived from a TE, but another part is without TE features. **Type III Wholly overlapped**—The pre-miRNA wholly overlaps with a TE. Genome sequence is marked with green bar, pre-miRNA with purple bar and mature miRNA with yellow bar. Red arrow indicates the direction of TE sequence. **(B) Percentage of three types of MDTEs. (C) The proportion of TE families in three types of MDTEs.**

doi:10.1371/journal.pone.0131365.g002

(Fig 2B). It might be due to the abundance of similar fragments and palindromic structure in TEs [4, 36]. These fragments provide the potential to form the hairpin structure of miRNAs. Interestingly, approximately 51.78% of UMDTEs partly overlap with TEs (Fig 2B).

Four TE classes (SINE, LINE, LTR and DNA) have different contributions for three types of MDTEs (Fig 2C). In human, SINEs, LINEs, LTR retroposons and DNA transposon copies comprise 13%, 20%, 8% and 3% of the genome sequences (The proportions of four TEs as a whole: SINE:29.55%; LINE:45.45%; LTR:18.18%; DNA:6.82%;) [30]. SINE and LINE (39.47% & 44.74%) are the major contributors for Type I UMDTEs. The bulk of Type II UMDTEs is composed of LINE, DNA and SINE (33.14%, 29.14% and 24.01%). In Type III UMDTEs, DNA and SINE are main resources (DNA:35.20%, SINE:28.00%), while LTR and LINE take up 19.2% and 17.6% respectively.

Compared to frequencies of TE families in human genome, TE families show different contributions to human MDTEs (Pearson's Chi-squared test: $\chi^2 = 49.65$, $P = 1.69e^{-8}$, Fig 3).

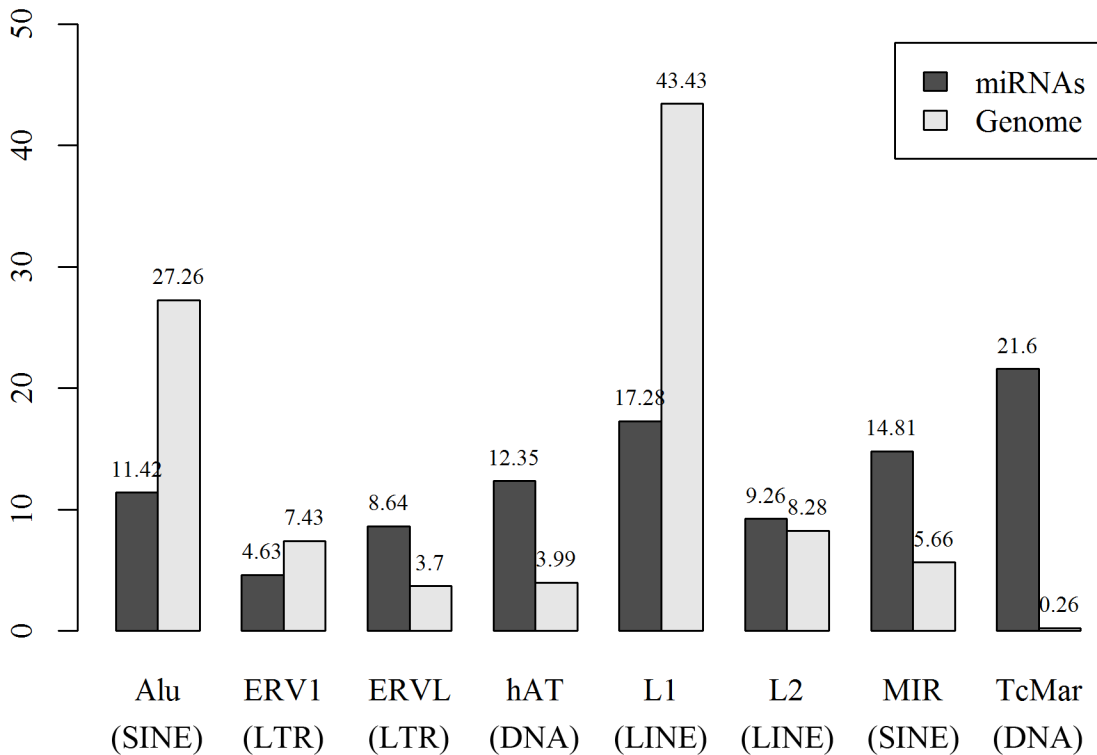


Fig 3. Percentage of MDTEs and different TE families in human genome. The sum of relative percentage is 100% for each group.

doi:10.1371/journal.pone.0131365.g003

Consistent with the result of previous work [23], MIR (SINE) and DNA elements including TcMar and hAT families contribute more to origin of miRNAs compared to their frequencies in human genome. In comparison, Alu (SINE), L1 (LINE) and ERV1 (LTR) families less overlap with miRNAs. Unexpectedly, ERVL (LTR) family shows more overlap than expected regarding its frequency in human genome.

Type II MDTEs are derived from TEs via two patterns in human

In UMDTEs, 51.78% miRNA belongs to Type II MDTEs which partly overlap with TEs. Type II MDTEs was found to be generated by two patterns: Pattern I in which MDTEs loss their TE sequence features from whole TE sequences, and Pattern II in which MDTEs with a part of the pre-miRNA are derived from the head or tail of TEs (Fig 4). In Pattern I, it is evident to observe MDTEs in miRNA homologies or multi-copy miRNAs obviously. The overlap between TEs and MDTEs in pattern I is reduced from 100% to 30% or even less (Table 2).

Compared with Pattern I MDTEs, which account for 77.14% of Type IIMDTEs, TEs form one arm of the pre-miRNAs in Pattern II MDTEs. In this condition, the TEs were inserted in the proximity of appropriate sequences that are similar to the complementary sequences of the TE head or tail to form the hairpin structure of pre-miRNAs, such as hsa-mir-326, hsa-mir-421 and hsa-mir-619 (S2 Table). For Pattern II MDTEs, the mature miRNAs are derived from not only the internal portion of TEs but also non-TE sequences that were complementary with the head or tail of TEs.

Two origin mechanisms of MDTEs can be found across three types of MDTEs. In Type II and Type III MDTEs, some miRNAs were generated via being taken into genomes by TEs and

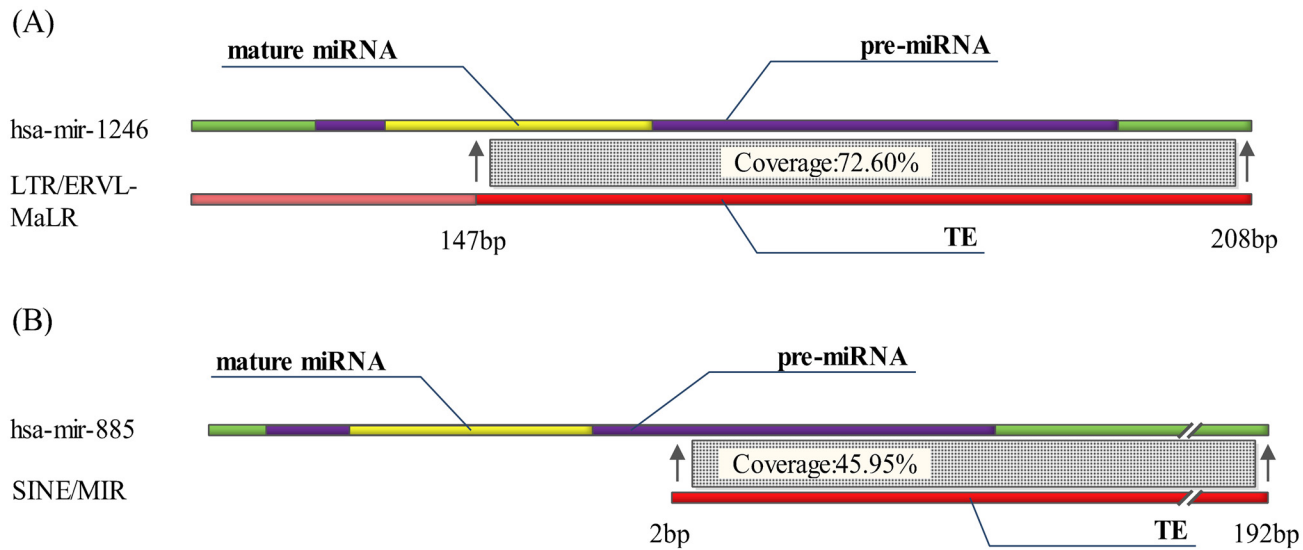


Fig 4. Two patterns of Type II MDTEs. (A) Pattern I MDTEs. The MDTEs lost their TE features of the wholly overlapped type. For example, 72.60% of hsa-mir-1246 with its downstream sequence was derived from 147 bp to 208 bp of MLT1M. The sequence before 147 bp has lost its TE features. **(B) Pattern II MDTEs.** TEs take up half of the miRNAs from their head or tail. For example, the 3' arm of hsa-mir-885 with its downstream sequence is derived from 2 bp to 192 bp of AluSc. The 5' arm is derived from non-TE sequence.

doi:10.1371/journal.pone.0131365.g004

passed on to new species after species differentiation. In Type I and Type II MDTEs, some miRNAs are generated by TEs from current genome.

Identification and characterization of human MDTEs without TE sequence features

About 19.84% miRNAs wholly or partly overlap with TE sequences in human, but the origin of other miRNAs is not very clear. To identify those MDTEs losing their sequence features of

Table 2. Extent of overlap in the same miRNA homologies or multi-copy miRNAs.

miRNAs	TEs	Overlap	miRNAs	TEs	Overlap
hsa-mir-1289-1	MER5A	100.00%	hsa-mir-548ab	MADE1	100.00%
hsa-mir-1289-2	MER5A	78.38%	hsa-mir-548u	MADE1	97.53%
hsa-mir-3118-1	L1PA13	53.33%	hsa-mir-548q	MADE1	96.00%
hsa-mir-3118-2	L1PA13	53.42%	hsa-mir-548f-3	MADE1	95.40%
hsa-mir-3118-3	L1PA13	53.33%	hsa-mir-548e	MADE1	90.91%
hsa-mir-3118-4	L1PA13	41.33%	hsa-mir-548m	MADE1	90.70%
hsa-mir-3118-5	L1PA13	40.79%	hsa-mir-548l	MADE1	89.53%
hsa-mir-3118-6	L1PA13	41.33%	hsa-mir-548v	MADE1	88.75%
hsa-mir-378a	MIRc	65.15%	hsa-mir-548g	MADE1	87.64%
hsa-mir-378b	MIRc	61.40%	hsa-mir-548az	MADE1	84.21%
hsa-mir-378d-1	MIRb	42.59%	hsa-mir-548ap	MADE1	81.25%
hsa-mir-378d-2	MIRc	44.90%	hsa-mir-548x-2	MADE1	80.00%
hsa-mir-378e	MIRc	31.65%	hsa-mir-548a-1	MADE1	77.32%
hsa-mir-378f	MIRc	50.00%	hsa-mir-548f-4	MADE1	71.43%
hsa-mir-378g	MIR3	100.00%	hsa-mir-548k	MADE1	68.10%
hsa-mir-378h	MIRc	100.00%	hsa-mir-548i-1	MADE1	53.02%

doi:10.1371/journal.pone.0131365.t002

TEs, the miRNAs which do not overlap with TEs in human were analyzed and compared with their homologies in other vertebrates. Twenty-three miRNAs were identified as human MDTEs which do not overlap with TEs, while their homologies either wholly or partly overlap with same TE sequence in other species (S3 Table). Although these MDTEs just account for 1.35% of all miRNAs in human, it implies that more miRNAs than expected may be derived from TE sequences in vertebrates.

Conclusion

In summary, we found that TE is an important origin source of human miRNAs. MiRNAs can be brought into genomes during the insertion of TEs or generated by TE sequences via particular mechanisms in current genome. When MDTEs fixed in the genome, sequence features of TE of MDTEs might be lost during the evolution. The observation that some MDTEs partly overlap with TE sequences and some MDTEs do not overlap with TEs implies that there are more MDTEs in genomes of vertebrates than what we previously believed. Our findings provide an insight into the origin and evolution of miRNAs.

Supporting Information

S1 Table. MDTEs overlap with TEs.

(XLSX)

S2 Table. Two patterns of Type II MDTEs.

(XLSX)

S3 Table. MDTEs do not overlap with TEs but their homologies overlap with TEs.

(XLSX)

Author Contributions

Conceived and designed the experiments: SQ FM. Analyzed the data: SQ PJ XZ. Wrote the paper: SQ LMC FM.

References

1. Kazazian HH Jr. Mobile elements: drivers of genome evolution. *Science*. 2004; 303(5664):1626–32. PMID: [15016989](#)
2. Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capi P, Chalhoub B, et al. A unified classification system for eukaryotic transposable elements. *Nature Reviews Genetics*. 2007; 8(12):973–82. PMID: [17984973](#)
3. Biémont C, Vieira C. Genetics: junk DNA as an evolutionary force. *Nature*. 2006; 443(7111):521–4. PMID: [17024082](#)
4. Feschotte C. Transposable elements and the evolution of regulatory networks. *Nature Reviews Genetics*. 2008; 9(5):397–405. doi: [10.1038/nrg2337](#) PMID: [18368054](#)
5. Lynch VJ, Leclerc RD, May G, Wagner GP. Transposon-mediated rewiring of gene regulatory networks contributed to the evolution of pregnancy in mammals. *Nat Genet*. 2011; 43(11):1154–9. doi: [10.1038/ng.917](#) PMID: [21946353](#)
6. Kunarso G, Chia N-Y, Jeyakani J, Hwang C, Lu X, Chan Y-S, et al. Transposable elements have rewired the core regulatory network of human embryonic stem cells. *Nature Genetics*. 2010; 42(7):631–4. doi: [10.1038/ng.600](#) PMID: [20526341](#)
7. Zemojtel T, Vingron M. P53 binding sites in transposons. *Front Genet*. 2012; 3:40. doi: [10.3389/fgene.2012.00040](#) PMID: [22448161](#)
8. Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, Tuschl T. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *nature*. 2001; 411(6836):494–8. PMID: [11373684](#)

9. Girard A, Sachidanandam R, Hannon GJ, Carmell MA. A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature*. 2006; 442(7099):199–202. PMID: [16751776](#)
10. Lee RC, Feinbaum RL, Ambros V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell*. 1993; 75(5):843–54. PMID: [8252621](#)
11. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*. 2004; 116(2):281–97. PMID: [14744438](#)
12. Carrington JC, Ambros V. Role of microRNAs in plant and animal development. *Science*. 2003; 301(5631):336–8. Epub 2003/07/19. PMID: [12869753](#)
13. Lai EC. Predicting and validating microRNA targets. *Genome biology*. 2004; 5.
14. Smalheiser NR, Torvik VI. Complications in mammalian microRNA target prediction. *Methods Mol Biol*. 2006; 342:115–27. PMID: [16957371](#)
15. Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell*. 2005; 120(1):15–20. PMID: [15652477](#)
16. Nilsen TW. Mechanisms of microRNA-mediated gene regulation in animal cells. *Trends Genet*. 2007; 23(5):243–9. PMID: [17368621](#)
17. Valencia-Sanchez MA, Liu J, Hannon GJ, Parker R. Control of translation and mRNA degradation by miRNAs and siRNAs. *Genes & development*. 2006; 20(5):515–24.
18. John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS. Human microRNA targets. *PLoS biology*. 2004; 2(11):e363. PMID: [15502875](#)
19. Krek A, Grun D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, et al. Combinatorial microRNA target predictions. *Nat Genet*. 2005; 37(5):495–500. PMID: [15806104](#)
20. Miranda KC, Huynh T, Tay Y, Ang YS, Tam WL, Thomson AM, et al. A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. *Cell*. 2006; 126(6):1203–17. PMID: [16990141](#)
21. Borchert GM, Holton NW, Williams JD, Hernan WL, Bishop IP, Dembosky JA, et al. Comprehensive analysis of microRNA genomic loci identifies pervasive repetitive-element origins. *Mob Genet Elements*. 2011; 1(1):8–17. PMID: [22016841](#)
22. Piriyaopongsa J, Jordan IK. A family of human microRNA genes from miniature inverted-repeat transposable elements. *PloS one*. 2007; 2(2):e203. PMID: [17301878](#)
23. Piriyaopongsa J, Marino-Ramirez L, Jordan IK. Origin and evolution of human microRNAs from transposable elements. *Genetics*. 2007; 176(2):1323–37. PMID: [17435244](#)
24. Smalheiser NR, Torvik VI. Mammalian microRNAs derived from genomic repeats. *Trends Genet*. 2005; 21(6):322–6. PMID: [15922829](#)
25. Smalheiser NR, Torvik VI. Alu elements within human mRNAs are probable microRNA targets. *Trends Genet*. 2006; 22(10):532–6. PMID: [16914224](#)
26. Kozomara A, Griffiths-Jones S. miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Research*. 2011; 39(suppl 1):D152–D7.
27. Hubbard T, Barker D, Birney E, Cameron G, Chen Y, Clark L, et al. The Ensembl genome database project. *Nucleic Acids Res*. 2002; 30(1):38–41. PMID: [11752248](#)
28. Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J. Repbase Update, a database of eukaryotic repetitive elements. *Cytogenetic and genome research*. 2005; 110(1–4):462–7. PMID: [16093699](#)
29. Smit A, Hubley R, Green P. RepeatMasker Open-3.0. 2004.
30. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001; 409(6822):860–921. PMID: [11237011](#)
31. Hellsten U, Harland RM, Gilchrist MJ, Hendrix D, Jurka J, Kapitonov V, et al. The genome of the Western clawed frog *Xenopus tropicalis*. *Science*. 2010; 328(5978):633–6. doi: [10.1126/science.1183670](#) PMID: [20431018](#)
32. Howe K, Clark MD, Torroja CF, Torrance J, Berthelot C, Muffato M, et al. The zebrafish reference genome sequence and its relationship to the human genome. *Nature*. 2013; 496(7446):498–503. doi: [10.1038/nature12111](#) PMID: [23594743](#)
33. International Chicken Genome Sequencing C. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature*. 2004; 432(7018):695–716. PMID: [15592404](#)
34. Mouse Genome Sequencing C, Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, et al. Initial sequencing and comparative analysis of the mouse genome. *Nature*. 2002; 420(6915):520–62. PMID: [12466850](#)

35. Nozawa M, Miura S, Nei M. Origins and evolution of microRNA genes in *Drosophila* species. *Genome Biol Evol.* 2010; 2:180–9. doi: [10.1093/gbe/evq009](https://doi.org/10.1093/gbe/evq009) PMID: [20624724](https://pubmed.ncbi.nlm.nih.gov/20624724/)
36. Slotkin RK, Martienssen R. Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet.* 2007; 8(4):272–85. PMID: [17363976](https://pubmed.ncbi.nlm.nih.gov/17363976/)