Laboratoire IBISC

Biologie Intégrative et Systèmes Complexes

An automatic method for identifying TE-derived miRNAs

Sébastien Tempel, Fariza Tahi

IBISC University of Evry-Val d'Essonne, Genopole, France

RAPPORT DE RECHERCHE
IBISC-RR-2011-02
mars 2011

An automatic method for identifying TE-derived miRNAs

Sébastien TEMPEL[1] and Fariza TAHI[1]

Laboratoire IBISC, Université d'Evry/Genopole,Tour Evry 2, 523, Place des Terrasses, 91000 Evry, France
sebastien.tempel@ibisc.univ-evry.fr
fariza.tahi@ibisc.univ-evry.fr

Abstract *MicroRNAs and transposable elements (TEs) share numerous characteristics: size, stable secondary structure, maturation into shorter sequence and regulation of binding target genes. TE-derived miRNAs are microRNAs that are generated from transposable elements. TE-derived miRNAs and TEs have a similar distribution of their occurrences that is distinct from 'classical' miRNAs: a TE-derived miRNA have many occurrences spread in many chromosomes, and a 'classical' miRNA has generally one occurrence of few ones appearing in a cluster.*

We developed an automatic method called miRNAcheck to distinguish a TE-derived miRNA from a 'classical' miRNA. Given a miRNA candidate, miRNAcheck calculates in a first step the number of occurrences of the candidate in the genome. The ten occurrences the most similar to the candidate sequence are then extended and a consensus sequence is created. Finally, the consensus sequence is compared to TE sequences in RepBase, a database of TEs.

From the 1048 human miRNAs of miRBase, we selected the miRNAs that have at least 5 similar occurrences in the genome. We get only 67 candidates. Among them, 59 are identified by miRNAcheck as TE-derived miRNAs.

miRNAcheck is available at the Web site: http://EvryRNA.ibisc.univ-evry.fr/

Keywords miRNA, Transposable Element, TE-derived miRNA

1 Introduction

Recent studies show the whole genomes of higher eukaryotes are transcribed [14] while the genes represent only few percentages of these genomes. The non-genic regions are mainly composed by non-coding RNAs (ncRNAs) and transposable elements (TEs) that represent a substantial fraction of many eukaryotic genomes. For example, about 50% of the human genome is derived from transposable element sequences [10].

Transposable elements are present in nearly all genomes that have been studied to date and in some cases represent most of the genome [17]. They move or are copied from one genomic location to another [4]. TEs are characterized and classified on the basis of terminal or sub-terminal remarkable structures or of their protein-coding capacity. TEs are conventionally divided into two classes [30]: Class I and Class II. Class I is represented by the retrotransposons LINEs, SINEs, LTRs, and ERVs, both requiring reverse transcription from an RNA intermediate. Class II includes "cut-and-paste" DNA transposons, which are characterized by terminal inverted repeats (TIRs) and are mobilized by a transposase [4]. Many families of both classes do not show any coding capacity and are called non-autonomous transposable elements. They have cumulated so many mutations, insertions or deletions so they are generally solely defined by their extremities [29,5]. In Class I, Short INterspersed Elements (SINEs) are short sequences (100 to 500 nt) and present a stable secondary structure similar to the fusion of tRNA structure and hairpin structure [15,28]. In Class II, Miniature Inverted-repeat Transposable Elements (MITEs) are non-autonomous transposable elements characterized by a small size (80-500 bp), a stable secondary structure, generally hairpin structure, and an insertion into A + T-rich regions [3]. MITEs could generate small interfering RNAs (22-24 bp) by a pathway similar to that required for TE-derived siRNA biogenesis and by DICER-like proteins [18].

siRNAs are non-coding RNAs generated from a biological response to double-stranded RNAs (dsRNAs) called RNA interferences (RNAi) [8,25]. Long dsRNA molecules (for example TE) initiate RNAi by being

converted to smaller 21-23 nt siRNAs by the Dicer enzyme. Therefore, hairpin RNAs have been commonly used to induce RNAi [8].

MicroRNAs (miRNAs) are non-coding RNAs with only 21-25 nt in sequence length that are present in all sequenced higher eukaryotes [1,9]. They are involved as negative regulators of gene expression at the post-transcriptional level by binding to specific mRNA targets whose translations are inhibited or down-regulated [9,21]. According to the current understanding of miRNA biogenesis, miRNA genes are transcribed and then are cleaved into a 60-80 bp long precursor of miRNA sequences (pre-miRNAs) by the Drosha/Pasha complex. Pre-miRNAs, structured as hairpins, are transported into the cytoplasm by Exportin5 and cleaved by Dicer into mature miRNAs [1]. In the RISC complex, a miRNA binds with a specific mRNA transcript and leads to the cleavage or the degradation of the mRNA.

TEs and miRNAs share numerous characteristics (Fig. 1), especially the similarity of their biogenesis and the regulation of their targets [23,11]. Moreover, some recent bioinformatic studies show that some miRNAs share their sequences or an important part of their sequences with TEs [11,23,27]. These miRNAs, annotated in miRBase [7], are called TE-derived miRNAs and present a high number of occurrences in the genome [11]. Both classes of TEs could be involved in TE-derived miRNAs [24].

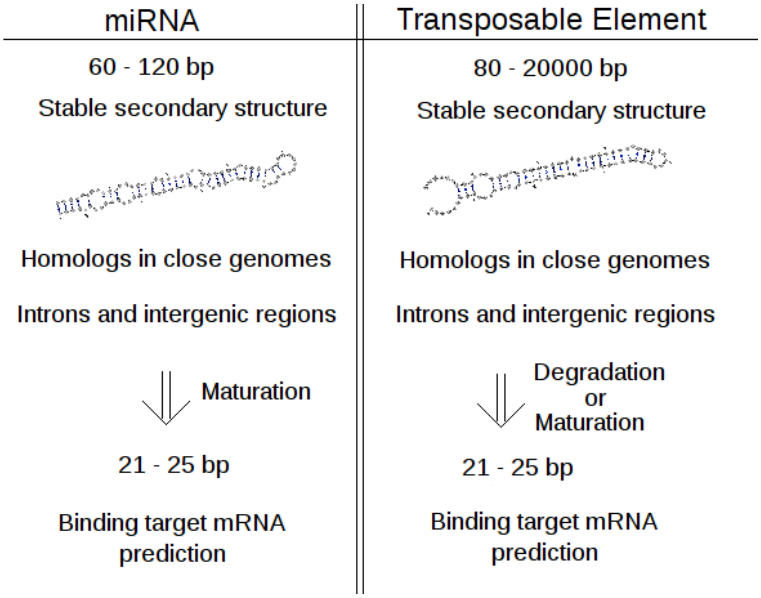

Figure 1. Bioinformatic characteristics of (pre-)miRNAs and transposable elements. The size and the secondary structure of pre-miRNAs and TEs are similar. Mature miRNAs down regulate the target binding genes [1] and the siRNAs generated from TE targets the TE-genes [18].

In this article, we present an automatic method for identifying TE-derived miRNAs, called miRNAcheck.

2 Our approach

2.1 How to identify TE-derived miRNAs

The main criteria that identifies TE-derived miRNA candidates from other miRNAs is the number and the distribution of the candidate occurrences. miRNAs do not have a transposition mechanism like TEs, and are not widespread in all chromosomes, not even widespread in one chromosome [2,19]. A mechanism that can copy miRNAs is an error of chromosome replication that can give a cluster of miRNAs [26]. This difference in the copy mechanism changes the localization of occurrences and allows to distinguish the TEs from the satellites (tandem repeats) [4].

The study of a miRNA occurrences distribution depends on:

1. The number of occurrences in the whole genome. Excepted TE-derived miRNAs, a miRNA has few occurrences in the whole genome [2,19]. Therefore, we can consider that a miRNA candidate that occurs several times in the genome has a strong probability to be a TE-derived miRNA.
2. The number of distinct chromosomes where appear the occurrences. The tandem repeat mechanism does not allow a sequence to jump to another chromosome [4]. Then, very few miRNAs are found in two chromosomes. The presence on a second chromosome could be explained by the chromosomic rearrangement during the evolution. Therefore, we can consider that a miRNA candidate present in several chromosomes have a strong probability to be a TE-derived miRNA.
3. The distance between the occurrences. Some recent studies show that some similar miRNAs are clustered in a small distance [26] and that the tandem repeat mechanism creates copies close to the original sequence [4]. For example, there is a cluster of 49 miRNAs in human chromosome 19 spread on only 150 kb. Sewer *et al* approximated the maximal distance of a miRNA cluster to 20kb [26]. Therefore, we can consider that if two or more similar occurrences are distant of more than 20 kb, there is a strong probability that the candidate is a TE-derived miRNA.

2.2 Description of our method

In order to identify TE-derived miRNAs, we developed an automatic method called miRNAcheck which works as follows.

Given a miRNA candidate, the first step of our method consists to study the distribution of the candidate occurrences, using BLAT [16] of UCSC Genome Browser [6]. We calculate the number of occurrences, named ”hits”, of the candidate in the genome, and more particularly the number of ”similar hits”. Similar hits are hits whose similarity with the candidate is greater than 80% and whose size is between 80% and 120% of the candidate size. This definition is similar to the identification definition of transposable elements [24]. We calculate also the number of chromosomes where occur the different similar hits.

After the study of the occurrences distribution, the second step of our method looks for a possible similarity with transposable elements. However, the size of pre-miRNA candidates (60-80 bp) could be too short for an identification by Censor [13]. To extend the candidate sequence, our method extracts the ten best similar hits (or all similar hits if they are less than ten). Thanks to UCSC genome browser [6], we get the surrounding sequence around each hit: 100 nt left to the hit and 100 nt right to the hit. These sequences are then aligned with ClustalW [20] and a consensus sequence is created. The nucleotide consensus at position i corresponds to the nucleotide present at least 7 times in the alignment at same position; otherwise there is the character N. We assume that ten hits are sufficient to create a consensus sequence since the hits have a similarity with the candidate greater than 80%.

Finally, we compare the consensus sequence to a TE database: RepBase [12]. The candidate is a TE-derived miRNA if the consensus is similar to a TE in RepBase.

2.3 miRNAcheck tool

Our method was implemented in JAVA. The obtained tool, called miRNACheck, is available on the Web site: http://EvryRNA.ibisc.univ-evry.fr.

The interface of miRNAcheck (Fig. 2) works as follows: the user enters the sequence of a miRNA candidate in STADEN format, enters a miRNA name and chooses the genome where the sequence belongs. miRNAcheck sends a request to BLAT of UCSC Genome Browser and gets the hits of the sequence in the genome (2 in Fig. 2). The line above the hits table resume the BLAT result (the number of hits returned by BLAT, the number of chromosomes where appear the hits, and the number of similar hits, i.e. hits that have a size between 80% and 120% of the miRNA size and that have a similarity greater than 80% with the candidate sequence). The user can check the hits obtained from BLAT with a link to the results obtained by BLAT. miRNAcheck selects then the 10 most similar hits (or all if there are less than 10 similar hits) and extends the hits in the genome sequence. The extended hits sequences are then aligned by ClustalW and a consensus sequence is generated

(3 in Fig. 2). Finally, the consensus is sent to the RepBase database [12] in order to identify a TE candidate associated to the consensus sequence. The alignment between the consensus and the most similar TE is then shown (4 in Fig. 2). A pop-up resumes the results and specifies if the candidate is a TE-derived miRNA.

Figure 2. JAVA interface of miRNAcheck. It shows here the results obtained for the human miRNA has-mir-1273e.

3 Results and discussion

For our tests, we considered the 1048 human miRNAs present in MiRBase version 16. The first step was to calculate the number of hits of each of the miRNAs. Only 67 miRNAs have more than 5 similar hits in the genome. These miRNAs are listed in Fig. 3. Thanks to miRNAcheck, we found that among these 67 miRNAs, 59 are TE-derived miRNAs (Fig. 3).

As shown in Fig. 3, 9 miRNAs (that have more than 5 similar hits) do not correspond to a RepBase TE. 3 of them (HSA-MIR-3179-1, HSA-MIR-3179-2 and HSA-MIR-3179-3) occur in one chromosome only. However, the hits of these miRNAs are not close to each other (there is a hit that is at more than 2 million nt from an other). These 9 miRNAs require therefore more study in order to know if they are TE-derived or not.

miRNA name	TE name	Hits	Chrom	miRNA name	TE name	Hits	Chrom
HSA-MIR-548A-1	MADE1	55	25	HSA-MIR-548X	MADE1	89	19
HSA-MIR-548A-2	MADE1	91	28	HSA-MIR-548Y	MADE1	13	29
HSA-MIR-548A-3	MADE1	102	26	HSA-MIR-548Z	MADE1	145	27
HSA-MIR-548AA-1	MADE1	86	23	HSA-MIR-566	AluSg	108	20
HSA-MIR-548AA-2	MADE1	157	25	HSA-MIR-622	???	43	28
HSA-MIR-548B	MADE1	14	21	HSA-MIR-1233-1	???	27	10
HSA-MIR-548C	MADE1	145	27	HSA-MIR-1233-2	???	27	10
HSA-MIR-548D-1	MADE1	86	23	HSA-MIR-1244-1	???	22	21
HSA-MIR-548D-2	MADE1	157	25	HSA-MIR-1244-2	???	22	21
HSA-MIR-548E	MADE1	62	21	HSA-MIR-1244-3	???	22	21
HSA-MIR-548F-1	MADE1	89	25	HSA-MIR-1255B-1	TIGGER1	160	23
HSA-MIR-548F-2	MADE1	103	23	HSA-MIR-1268	AluY	87	13
HSA-MIR-548F-3	MADE1	89	20	HSA-MIR-1273	AluY	148	24
HSA-MIR-548F-5	MADE1	47	19	HSA-MIR-1273D	AluY	84	21
HSA-MIR-548G	MADE1	27	26	HSA-MIR-1273E	AluY	108	24
HSA-MIR-548H-1	MADE1	45	22	HSA-MIR-1285-2	AluJr	36	23
HSA-MIR-548H-2	MADE1	179	24	HSA-MIR-1302-1	MER53	54	20
HSA-MIR-548H-4	MADE1	34	23	HSA-MIR-1302-10	MER53	9	8
HSA-MIR-548I-1	MADE1	17	22	HSA-MIR-1302-11	MER53	9	8
HSA-MIR-548I-2	MADE1	16	21	HSA-MIR-1302-2	MER53	9	8
HSA-MIR-548I-3	MADE1	18	23	HSA-MIR-1302-3	MER53	10	8
HSA-MIR-548I-4	MADE1	128	22	HSA-MIR-1302-4	MER53	49	27
HSA-MIR-548K	MADE1	5	5	HSA-MIR-1302-9	MER53	9	8
HSA-MIR-548L	MADE1	85	22	HSA-MIR-3118-1	L1PA13_5	14	8
HSA-MIR-548M	MADE1	74	21	HSA-MIR-3118-2	L1PA13_5	14	8
HSA-MIR-548N	MADE1	180	22	HSA-MIR-3118-3	L1PA13_5	14	8
HSA-MIR-548O	MADE1	90	23	HSA-MIR-3118-4	L1PA13_5	13	8
HSA-MIR-548P	MADE1	80	25	HSA-MIR-3118-5	L1PA13_5	12	8
HSA-MIR-548Q	MADE1	5	5	HSA-MIR-3118-6	L1PA13_5	13	8
HSA-MIR-548S	MADE1	8	7	HSA-MIR-3135	AluJb	29	10
HSA-MIR-548T	MADE1	104	23	HSA-MIR-3179-1	???	6	1
HSA-MIR-548V	MADE1	58	22	HSA-MIR-3179-2	???	6	1
HSA-MIR-548W	MADE1	132	24	HSA-MIR-3179-3	???	6	1
				HSA-MIR-3929	AluJr	8	2

Figure 3. Human miRNAs that have at least 5 similar hits in the genome. Most of them are identified as TE-derived miRNAs. miRNAs that have "???" in 'TE name' column are not similar to known TEs listed in Repbase. The columns 'Hits' and 'Chrom' correspond respectively to the number of similar hits and to the number of chromosomes where appear these hits. There are 24 genomic chromosomes and 9 haplotype chromosomes in UCSC Genome Browser [16].

Fig. 3 shows also that miRNAs having a same name prefix (e.g. HSA-MIR-548A-1, HSA-MIR-548A-2, HSA-MIR-548B, etc.) correspond to a same TE, which is not surprising since these miRNAs have similar sequences.

One important remark is that very few miRNAs have several similar hits; only 77 among 1048 have more than 5 similar hits. This observation confirms the fact that a miRNA is normally unique or with very few and close similar hits.

King Jordan *et al.* have previously discussed the origin of miRNAs and the possibility that they come from the evolution of MITEs [11,23]. Their hypothesis is supported by the similarity between their secondary structures and by the similarity between their targeting mechanism. Moreover, Smalheiser *et al.* shown that some mammal miRNAs have a small fragment of L2 transposable element in their sequence [27]. If the hypothesis of miRNA TE-derived origin seems possible, some studies stipulate that miRNAs derive from genomic loci distinct from any other recognized elements [1,22] and Yan *et al.* think for instance that mir4441 and mir4446 are misannoted as miRNAs but are siRNAs [31].

Our automatic method confirmed the previously result obtained manually by Jordan *et al.* which shown that 6 human miRNAs 'hsa-mir-548' are TE-derived [11]. Thanks to our tool miRNAcheck, we identified 62 new TE-derived human miRNAs.

4 Conclusion

In this paper, we present an automatic method called miRNAcheck for identifying TE-derived miRNAs. TE-derived miRNAs are miRNAs that are derived from transposable elements (TEs).

Our method is based on the hypothesis that a miRNA that have several occurrences widespread in the genome has a high probability to be TE-derived. The first step of miRNAcheck is to calculate the number of occurrences of the miRNA candidate, the number of chromosomes where appear the different occurrences and the distance between the occurrences. The second step is then to calculate a consensus sequence to the ten occurrence sequences the more similar to the miRNA sequence. Finally, the last step consists to check if the consensus sequence corresponds to a TE in RepBase database.

We tested our method on human miRNAs of miRBase. There are a total of 1048 human miRNAs and only 77 have more than 5 occurrences (with high similarity). Almost all these 77 miRNAs are identified by miRNAcheck as TE-derived, i.e. corresponding to TEs in RepBase.

Thanks to miRNAcheck, one could check very quickly if a miRNA candidate is a TE-derived miRNA. It requires between 30 seconds to 1 minute to treat a miRNA sequence (depending on the number of occurrences in UCSC and on the access to RepBase).

miRNAcheck is available at the Web site: http://EvryRNA.ibisc.univ-evry.fr/

Acknowledgements

This work was funded by the Council of Essonne Region (Pôle System@tic, OpenGPU project).

References

[1] D. Bartel, MicroRNAs: genomics, biogenesis, mechanism and function. *Cell*, 116:281-197, 2004.

[2] E. Berezikov, N. Robine, A. Samsonova, J.O. Westholm, A. Naqvi, J-H. Hung, K. Okamura, Q. Dai, D. Bortolamiol-Becet, R. Martin, Y. Zhao, P.D. Zamore, G.J. Hannon, M.A. Marra, Z. Weng, N. Perrimon and E.C. Lai, Deep annotation of Drosophila melanogaster microRNAs yields insights into their processing, modification, and emergence. *Genome Res.*, 21:203-215, 2011.

[3] Y. Chen and F. Zhou and G. Li and Y. Xu, A recently active miniature inverted-repeat transposable element, Chunjie, inserted into an operon without disturbing the operon structure in Geobacter uraniireducens Rf4. *Genetics*, 179:2291-7, 2008.

[4] N.L. Craig and R. Gragie and M. Gellert and A.M. Lambowitz, Mobile DNA II Second Edition. *ASM Press*, 2002.

[5] C. Feschotte and C. Mouches, Evidence that a family of miniature inverted-repeat transposable elements (MITEs) from the *Arabidopsis thaliana* genome has arisen from a pogo-like DNA transposon. *Mol. Biol. Evol*, 17:4051-730-737, 2000.

[6] P.A. Fujita and B. Rhead and A.S. Zweig and A.S. Hinrichs and D. Karolchik and M.S. Cline and M. Goldman and G.P. Barber and H. Clawson and A. Coelho and M. Diekhans and T.R. Dreszer and B.M. Giardine and R.A. Harte and J. Hillman-Jackson and F. Hsu and V. Kirkup and R.M. Kuhn and K. Learned and C.H. Li and L.R. Meyer and A. Pohl and B.J. Raney and K.R Rosenbloom and K.E. Smith and D. Haussler and W.J. Kent, The UCSC Genome Browser database: update 2011. *Nucleic Acids Res.*, 39:D876-82, 2011.

[7] S. Griffiths-Jones and H. Saini and S. van Dongen and A. Enright, miRBase: tools for microRNA genomics. *Nucleic Acids Res.*, 36:D154-D158, 2008.

[8] G.J. Hannon and J.J. Rossi, Unlocking the potential of the human genome with RNA interference. *Nature*, 431:371-378, 2004.

[9] L. He and G. Hannon, microRNAs: small RNAs with a big role in gene regulation. *Nat. Rev. Genet*, 5:522-531, 2004.

[10] International Human Genome Sequencing Consortium, Initial sequencing and analysis of the human genome. *Nature*, 409:860-921, 2001.

[11] J. Piriyapongsa and I.K. Jordan, A Family of Human MicroRNA Genes from Miniature Inverted-Repeat Transposable Elements. *PLoS ONE*, 2:e203, 2007.

[12] J. Jurka and V.V. Kapitonov and A. Pavlicek and P. Klonowski and O. Kohany and J. Walichiewicz, Repbase Update, a database of eukaryotic repetitive elements. *Cytogentic and Genome Research*, 110:462-467, 2005.

[13] J. Jurka and P. Klonowski and V. Dagman and P. Pelton., CENSOR - a program for identification and elimination of repetitive elements from DNA sequences. *Comput Chem.*, 20:119-21, 1996.

[14] P. Kapranov and A.T. Willingham and T.R. Gingeras, Genome-wide transcription and the implications for genomic organization. *Nat Rev Genet*, 8:413-23, 2007.

[15] H. Kawagoe-Takaki and N. Nameki and M. Kajikawa and N. Okada, Probing the secondary structure of salmon SmaI SINE RNA. *Gene*, 365:67-73, 2006.

[16] W.J. Kent, BLAT – The BLAST-Like Alignment Tool. *Genome Research*, 4:656-664, 2002.

[17] M.G. Kidwell, and D.R. Lisch, Perspective: transposable elements and host genome evolution. *Trends Ecol. Evol*, 15:95-99, 2001.

[18] H. Kuang and C. Padmanabhan and F. Li and A. Kamei and P.B. Bhaskar and S. Ouyang Jiang and C. Robin Buell and B. Baker, Identification of miniature inverted-repeat transposable elements (MITEs) and biogenesis of their siRNAs in the Solanaceae: New functional implications for MITEs. *Genome Res.*, 19:42-56, 2008.

[19] P. Landgraf *et al,*, A Mammalian microRNA Expression Atlas Based on Small RNA Library Sequencing. *Cell*, 129:1401-1414, 2007.

[20] M.A. Larkin and G Blackshields and NP Brown and R. Chenna and PA McGettigan and H McWilliam and F Valentin and IM Wallace and A Wilm and R Lopez and JD Thompson and TJ Gibson and DG Higgins., Clustal W and Clustal X version 2.0. *Bioinformatics*, 23:2947-8, 2007.

[21] Y. Lee and M. Kim and J. Han and K. Yeom,K. and S. Lee and S. Baek and V. Kim, microRNA genes are transcribed by RNA polymerase II. *EMBO J.*, 23:4051-4060, 2004.

[22] BC. Meyers and MJ. Axtell and B. Bartel and DP. Bartel and D. Baulcombe and JL. Bowman and X. Cao and JC. Carrington and X. Chen and PJ. Green and S. Grifitths-Jones and SE. Jacobsen and AC. Mallory and RA. Martienssen and R. Scott Poethig and Y. Qi and H. Vaucheret and O. Voinnet and Y. Watanabe and D. Weigel and JK Zhui, Criteria for Annotation of Plant MicroRNAs. *The Plant Cell*, 20:3186-3190, 2008.

[23] J. Piriyapongsa and I.K. Jordan, Dual coding of siRNAs and miRNAs by plant transposable element. *RNA*, 14:814-821, 2008.

[24] J. Piriyapongsa and L. Marino-Ramirez and I.K. Jordan, Origin and Evolution of Human microRNAs From Transposable Elements. *Genetics*, 176:1323-1337, 2007.

[25] R. Rana, Illuminating the silence: understanding the structure and function of small RNAs. *Molecular Cell Biology*, 8:23-36, 2007.

[26] A. Sewer and N. Paul and P. Landgraf and A. Aravin and S. Pfeffer and M.J. Brownstein and T. Tuschl and E. van Nimwegen and M. Zavolan, Identification of clustered microRNAs using an ab initio prediction method. *BMC Bioinformatics*, 6:267, 2005.

[27] N.R. Smalheiser and V.I. Torvik, Mammalian microRNAs derived from genomic repeats. *Trends Genet.*, 21:322-326, 2005.

[28] J.D. Suntera and S.P. Patela and R.A. Skiltona and N. Githakaa and D.P. Knowlesb ande G.A. Scolesb and V. Nened and E de Villiers and R.P. Bishopa, A novel SINE family occurs frequently in both genomic DNA and transcribed sequences in ixodid ticks of the arthropod sub-phylum Chelicerata. *Genet. Dev.*, 415:13-22, 2008.

[29] S.R. Wessler and T.E. Bureau and S.E. White, LTR-retrotransposons and MITEs: important players in the evolution of plant genomes. *Genet. Dev.*, 5:814-821, 1995.

[30] T. Wicker and F. Sabot and A. Hua-Van and J.L. Bennetzen and P. Capy and B. Chalhoub and A. Flavell and P. Leroy and M. Morgante and O. Panaud and E. Paux and P. SanMiguel and A.H. Schulman, A unified classification system for eukaryotic transposable elements. *Nat Rev Genet*, 8:973-82, 2007.

[31] Y. Yan and Y. Zhang and K. Yang and Z. Sun and Y. Fu and X. Chen and R. Fang, Small RNAs from MITE-derived stem-loop precursors regulate abscisic acid signaling and abiotic stress responses in rice. *The Plant Journal*, 65:820-828, 2011.