

PlanTE-MIR DB: a database for transposable element-related microRNAs in plant genomes

Alan P. R. Lorenzetti¹ · Gabriel Y. A. de Antonio² · Alexandre R. Paschoal² · Douglas S. Domingues³

Received: 9 November 2015 / Revised: 14 January 2016 / Accepted: 19 January 2016 / Published online: 18 February 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract Transposable elements (TEs) comprise a major fraction of many plant genomes and are known to drive their organization and evolution. Several studies show that these repetitive elements have a prominent role in shaping non-coding regions of the genome such as microRNA (miRNA) loci, which are components of post-transcriptional regulation mechanisms. Although some studies have reported initial formation of miRNA loci from TE sequences, especially in model plants, the approaches that were used did not employ systems that would allow results to be delivered by a user-friendly database. In this study, we identified 152 precursor miRNAs overlapping TEs in 10 plant species. PlanTE-MIR DB was designed to assemble this data and deliver it to the scientific community interested in miRNA origin, evolution, and regulation pathways. Users can browse the database through a web interface and search for entries using various parameters. This resource is cross-referenced with repetitive element (Repbase Update) and

miRNA (miRBase) repositories, where sequences can be checked for further analysis. All data in PlanTE-MIR DB are publicly available for download in several file formats to facilitate their understanding and use. The database is hosted at <http://bioinfo-tool.cp.utfpr.edu.br/plantemirdb/>.

Keywords Transposable elements · MicroRNAs · Plant genomes · Database

Introduction

Transposable elements (TEs) are present in almost all living organisms and comprise a significant fraction of most plant genomes (Baidouri and Panaud 2013; Ragupathy et al. 2013). They are known to drive important modifications in host genomes, including the inactivation, creation, and mobilization of genes, chromosomal rearrangement, gene expression modulation, and epigenetic silencing (Lisch 2013; Bennetzen and Wang 2014). TEs are also known to have an important role in shaping long noncoding RNAs (lncRNAs) and small ncRNAs (e.g., piwi-interacting RNAs [piRNAs], small interfering RNAs [siRNAs], and microRNAs [miRNAs]) (Hadjiargyrou and Delilhas 2013; Piriyaongsa and Jordan 2008; Li et al. 2011; Gim et al. 2014).

Mature plant miRNAs are usually 21-nucleotide-long hairpin RNA-derived (hpRNA) sequences that can bind to target mRNAs through Watson-Crick base pairing on the 3' UTR. This pairing results in mRNA destabilization or translational repression, which are effective mechanisms for gene regulation. In plants, RNA polymerase II enzymatic complexes are generally committed to transcribe miRNA loci. Inside the cell nucleus, cleavage of foldback transcribed structures through DICER-LIKE 1 (DCL1) is executed.

Electronic supplementary material The online version of this article (doi:10.1007/s10142-016-0480-5) contains supplementary material, which is available to authorized users.

✉ Douglas S. Domingues
doug@rc.unesp.br

Alan P. R. Lorenzetti
alan.lorenzetti@uel.br

- ¹ Graduation Program in Genetics and Molecular Biology, Universidade Estadual de Londrina, UEL, Londrina, Brazil
- ² Bioinformatics Laboratory, Universidade Tecnológica Federal do Paraná, UTFPR, Cornélio Procopio, Brazil
- ³ Department of Botany, Instituto de Biociências, Universidade Estadual Paulista, UNESP, Rio Claro, Brazil

This step promotes conversion of primary miRNAs (pri-miRNAs) into precursor miRNAs (pre-miRNAs), which is followed by further processing to transform them into miRNA/miRNA* duplex. Finally, one of the dissociated miRNAs is loaded into ARGONAUTE (AGO) proteins to assemble the functional RNA-induced silencing complex (RISC) (Axtell 2013; Erson-Bensan 2014).

Hairpin structures are supposed to arise either by inverted duplication of the target gene locus (Allen et al. 2004) or juxtaposed TEs (Roberts et al. 2013). Piriyaopongsa and Jordan (2008) also describe a model in which folded expressed miniature inverted-repeat transposable elements (MITEs) may be processed by the miRNA biogenesis pathway. In another report, Ou-Yang et al. (2013) used AGO1 immunoprecipitation and DCL mutants to find three MITE-associated bona fide miRNAs depending on those proteins, pointing to their functional activity.

Most miRNA loci originate from intergenic genomic sequences, but there is considerable evidence that they were initially formed from TE sequences (Hadjiargyrou and Delihias 2013; Roberts et al. 2014; Budak and Akpinar 2015). The “domestication” of TEs to form miRNA genes has been demonstrated by high throughput sequencing in rice (Li et al. 2011; Barrera-Figueroa et al. 2012), and other plant species were checked for TE-MIRs at the genomic

level (Zhang et al. 2011; Sun et al. 2012; Kurtoglu et al. 2014). Similar findings have been reported for the human genome and other metazoan genomes (Levy et al. 2008; Tempel et al. 2012), for which there are publicly available resources showing matched overlaps between TE and miRNA loci. However, the comparable plant data has not yet been compiled into a user-friendly database enabling search and retrieval of this information.

In this study, we present PlanTE-MIR DB, available at <http://bioinfo-tool.cp.utfpr.edu.br/plantemirdb/>, which provides a user-friendly database for investigation of overlaps between TE and pre-miRNA loci in ten plant genomes.

Material and methods

Pre-miRNA annotation and curation

Our analysis relied on the annotated pre-miRNAs from miRBase (version 21) (Kozomara and Griffiths-Jones 2014) within 15 genome assemblies. Genome assemblies were retrieved from several repositories based on reference versions indicated by miRbase (Table S1). However, due to divergences between miRBase annotation file accession names and assembly headers, a checking step was executed.

Fig. 1 Search section overview. Entries can be searched either by TE or pre-miRNA. Here, we present a search example for TE by Class I (retrotransposons), LTR order, and Gypsy superfamily

| Species Name | TE Name | Repbase Name | Overlapping pre-miRNA | Details | Fetch |
|------------------------|----------------|--------------|-----------------------|-------------------------|-------------------------------------|
| <i>Arabidopsis tha</i> | RLG_ATHILA4C_L | ATHILA4C_LTR | ath-MIR8175 | Details | <input checked="" type="checkbox"/> |
| <i>Arabidopsis tha</i> | RLG_ATHILA4B_L | ATHILA4B_LTR | ath-MIR855 | Details | <input checked="" type="checkbox"/> |
| <i>Arabidopsis tha</i> | RLG_ATHILA4B_L | ATHILA4B_LTR | ath-MIR401 | Details | <input checked="" type="checkbox"/> |
| <i>Arabidopsis tha</i> | RLG_ATHILA6A_I | ATHILA6A_I | ath-MIR854b | Details | <input checked="" type="checkbox"/> |
| <i>Arabidopsis tha</i> | RLG_ATHILA6A_I | ATHILA6A_I | ath-MIR854d | Details | <input checked="" type="checkbox"/> |
| <i>Arabidopsis tha</i> | RLG_ATHILA6A_I | ATHILA6A_I | ath-MIR854c | Details | <input checked="" type="checkbox"/> |
| <i>Arabidopsis tha</i> | RLG_ATHILA6A_I | ATHILA6A_I | ath-MIR854a | Details | <input checked="" type="checkbox"/> |
| <i>Arabidopsis tha</i> | RLG_ATHILA6A_I | ATHILA6A_I | ath-MIR854e | Details | <input checked="" type="checkbox"/> |

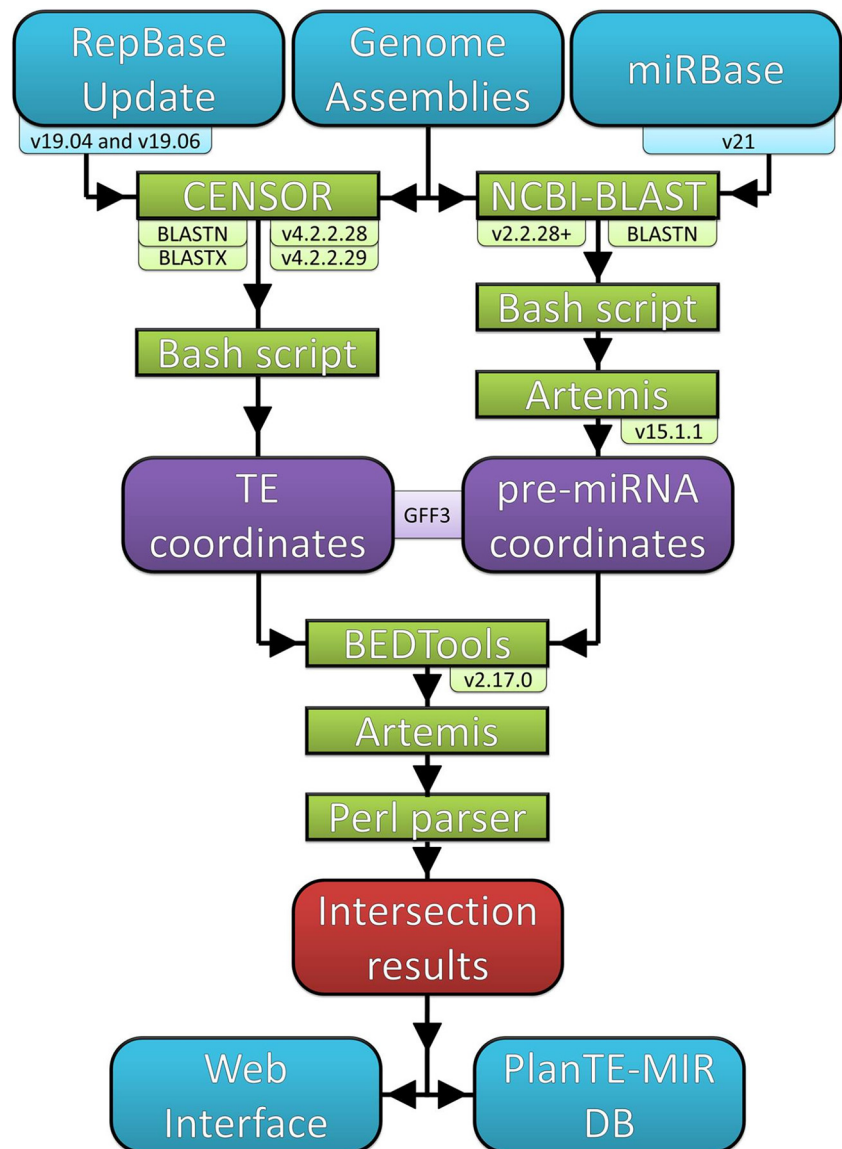
For that, miRBase pre-miRNA sequences in FASTA format were obtained for each of the studied plant species and BLAST (BLASTN, version 2.2.28+) (Camacho et al. 2009) searched against their respective genomes. An in-house bash script was made to run the program and perform the tasks. Only hits with 100 % query coverage and identity were maintained. In the case that some of the annotated pre-miRNAs from the same family showed indistinguishable sequences (e.g., mtr-MIR2669a, mtr-MIR2669b), they were aligned with more than one position. In these cases, we split single position hits from repeated ones and transformed them into GFF3 annotation files. Manual inspection was done using Artemis (version 15.1.1) (Rutherford et al. 2000) through the comparison of previously cited GFF3 files with an accession name corrected miRBase annotation file, all

loaded on source genomes. New manually inspected GFF3 annotation files were created to match the accession names to source assembly headers.

Reference TEs

Plant TE libraries were obtained from Repbase Update (version 19.04 and 19.06 REPET edition) (Jurka et al. 2005). We used CENSOR (version 4.2.28 and 4.2.29) (Jurka et al. 2005) implemented with WU-BLAST (version 2.0 04-May-2006) as a search engine, using BLASTN and BLASTX algorithms according to well-established criteria (Wicker et al. 2007) to stringently remap reference TEs to genome assemblies. Initially, we used early versions of software and libraries with BLASTN, and then later versions with

Fig. 2 Workflow diagram for the identification of TE-MIRs. CENSOR and BLAST programs were used to map TEs and pre-miRNAs. Bash script was used to filter and parse results to GFF3 file format. Using Artemis, pre-miRNA files were checked to confirm names and positions according to miRBase. Positional intersection analysis between TE-miRNA was run using BEDtools and manually checked with Artemis. These results were modelled to build PlanTE-MIR DB



BLASTX. A bash script was written to filter the results, according to the 80-80-80 rule proposed by Wicker and colleagues, and to parse TE coordinates to GFF3 annotation files.

TE-MIR relationship

We found positional overlaps between TEs and pre-miRNAs using the BEDTools (version 2.17.0) (Quinlan and Hall 2010) intersection function. Only pre-miRNAs having at least 36 % of their extension covered by a TE were maintained. Intersections were manually checked using Artemis. Whole sequences were captured from source assemblies through an in-house bash script running EMBOSS tools (version 6.6.0.0) (Rice et al. 2000).

Evolutionary conservation between TE-MIRs across taxa

We used a sequence-based similarity search method (Reciprocal Best BLAST Hit - RBH), following the rationale of Sun et al. (2012) to track evolutionary conservation. Thus, we BLAST (BLASTN, version 2.2.28+) searched our pre-miRNAs against all miRBase (version 21) hairpin sequences. Only hits with E values $\leq 1e-06$ with at least 90 % query alignment and minimum of 80 % identity were maintained. We employed the same criteria used by Zhou et al. (2011) and Sun et al. (2012) to classify TE-MIRs at three levels: highly conserved (when TE-MIR homologs are present in both monocots and eudicots), low conserved (when TE-MIR homologs are present only in monocots

or eudicots), and nonconserved (when a TE-MIR has no homologs outside a single species).

Transcriptional evidence for miRNAs

We checked for transcriptional evidence through browsing only high confidence miRNAs for each one of the plant species in the miRBase (version 21). Our pre-miRNAs were also used as queries and BLAST (BLASTN, version 2.2.28+) searched against the miRNEST (version 2.0) deep sequencing prediction file (Szczęśniak and Makalowska 2014). Only hits presenting full identity and coverage of queries were maintained.

Database and web interface implementation

Annotation files were parsed to table using a Perl in-house script. Additional information was manually introduced using the Kingsoft Office Spreadsheet software. The data were then exported to a comma-separated values table and automatically inserted in MySQL Database Server (version 5.6) relational tables using a PHP script (version 5.3.10).

PlanTE-MIR DB was built on a 64-bit Windows (version 8.1) workstation. XAMPP (version 3.2.1) was executed to integrate Apache HTTP Server (version 2.2) with PHP and MySQL. The back-end was encoded in PHP and HTML5, using JavaScript jQuery library (version 1.11.2), with the plugins jQuery Vegas (version 1.3.5) and Ajax. For website design and structure customizing, we used Cascading Style Sheets (CSS3) as front-end. Except for the Windows operating system, only open source and cross-platform software

Table 1 Overall numbers of TEs, miRNAs, and TE-MIRs for the plant genomes analyzed in this study. Ten species presented at least one TE-MIR in their genome

| Species | TEs ^a | TEs ^b | pre-miRNAs ^c | TE-MIRs ^d |
|--------------------------------|------------------|------------------|-------------------------|----------------------|
| <i>Arabidopsis thaliana</i> | 6837 | 314 | 325 | 22 |
| <i>Brachypodium distachyon</i> | 3991 | 1402 | 317 | 2 |
| <i>Glycine max</i> | 18035 | 18290 | 573 | 4 |
| <i>Medicago truncatula</i> | 7481 | 1128 | 672 | 20 |
| <i>Oryza sativa</i> | 66794 | 2930 | 592 | 56 |
| <i>Physcomitrella patens</i> | 23744 | 3855 | 229 | 1 |
| <i>Populus trichocarpa</i> | 14587 | 2143 | 352 | 10 |
| <i>Sorghum bicolor</i> | 178106 | 29885 | 205 | 35 |
| <i>Solanum tuberosum</i> | 5530 | 14814 | 224 | 1 |
| <i>Vitis vinifera</i> | 16240 | 4787 | 163 | 1 |

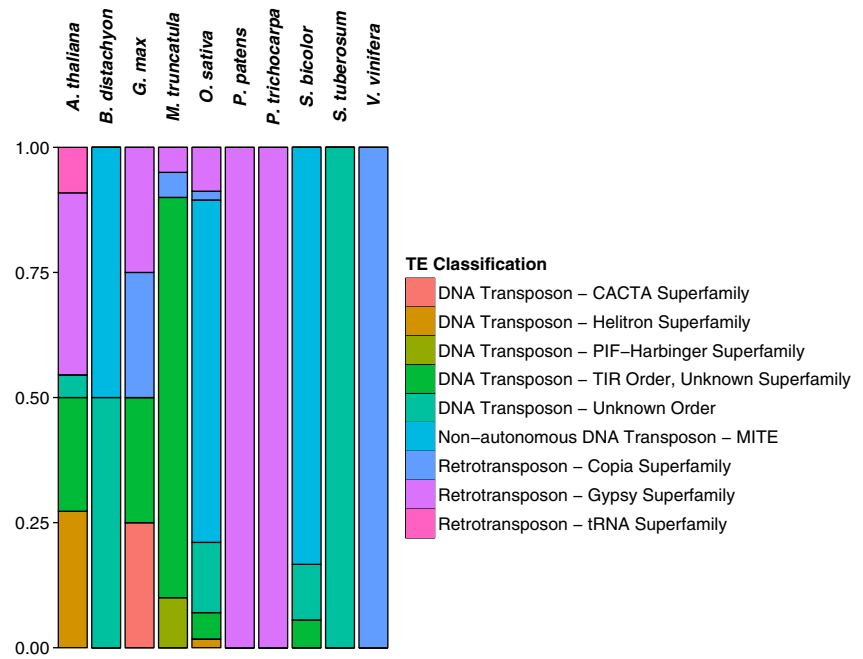
^aAnnotated TEs using CENSOR with BLASTN

^bAnnotated TEs using CENSOR with BLASTX

^cPre-miRNAs retrieved from miRBase (version 21)

^dMiRNAs intersecting at least one TE in each species

Fig. 3 Database composition by plant species and TE classification. MITEs are the most frequent type of repetition associated to miRNAs. DNA transposons were also highly related to this phenomenon. Classification data were collected from the Repbase Update



were used for database and web interface implementation. The complete system is hosted in the Information Office of the Federal University of Technology, Paraná, Brazil (UTFPR) and available at <http://bioinfo-tool.cp.utfpr.edu.br/plantemirdb/>.

Results and discussion

PlanTE-MIR DB: system and database overview

Plant Transposable Element-related miRNA Database (PlanTE-MIR DB, Fig. 1) was built as a resource for researchers interested in the evolution of TE and miRNA and their relationship. In this section, we detail the website and its functionalities.

The web interface was designed to be user-friendly, prioritizing easy ways of finding desired data through the use of filters, and providing alternative file formats when downloading entries. Accordingly, the website is divided in five sections: Home, About, Search, Download, and Team.

The Home and About sections concisely describe the purpose of the repository and its methods, and briefly instruct the user on how to interact with the search and download tools. They also contain information about assembly versions and reference libraries employed in the analysis.

In the Search section, users have a web interface for searching TE-MIR entries. The page was designed as an intuitive step-by-step form where the user can select options by name of the organism and TE or pre-miRNA attributes (Fig. 1). TEs can be found (1) by selecting the reference

name (as supplied by the Repbase Update), (2) by TE name according to the nomenclature of Wicker and colleagues, (3) by TE position in the genome assembly, and (4) by TE class (Wicker et al. 2007). The last option is a hierarchical filter that allows the user to choose among TE classes, orders, and superfamilies. Similarly, pre-miRNAs may be found by miRBase ID, miRBase name, or position in the

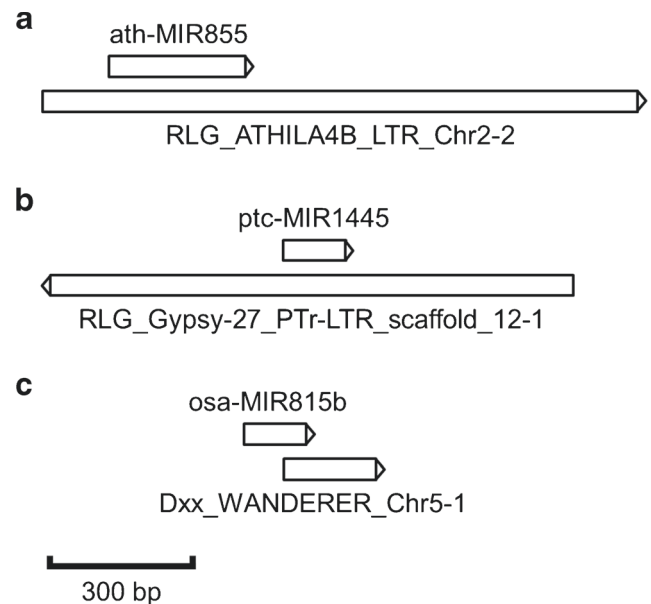


Fig. 4 Representative example of intersection patterns found by our analysis. Pre-miRNAs may intersect long terminal repeat (LTR) regions (a, b) or terminal regions of DNA transposons (c)

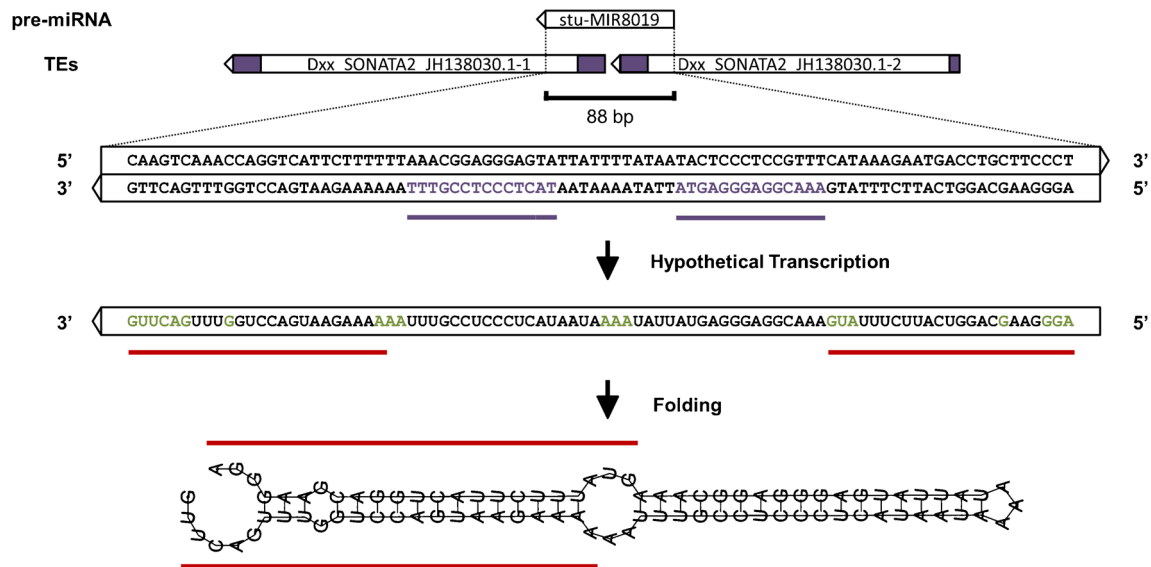


Fig. 5 Example of two juxtaposed non-autonomous DNA transposons overlapping a pre-miRNA in *Solanum tuberosum*. Terminal inverted repeats (TIRs) are highlighted in purple. Dxx_SONATA2_JH138030.1-1 is an intact element and Dxx_SONATA2_JH138030.1-2 lacks part

of the 5' TIR. A hairpin structure may emerge due to transcript complementarity. Light green letters show loop regions, and red lines emphasize mature miRNA regions. The secondary structure was plotted using RNAfold Webserver with the Minimum Free Energy (MFE) prediction method (Lorenz et al. 2011)

assembly. Next, a list of hits is shown to the user, allowing him or her to download search results by selecting Table file format, GFF3 file format, or FASTA file format. Furthermore, the user may access a detailed page containing information about the organism as well as annotations and cross-references obtained for each result. The description table shows species name, common name, assembly version, TE name, TE classification, Repbase Name, TE annotation details (such as Repbase version, CENSOR coverage, CENSOR similarity, start position, end position, and strand), overlapping pre-miRNA, pre-miRNA ID, and pre-miRNA annotation details. Further information relative to these items can be found in electronic supplementary material.

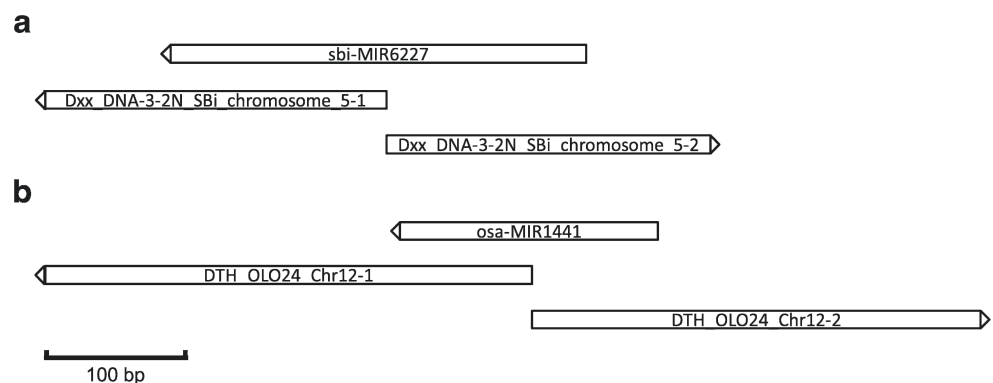
Bulk data for each species are provided through the Download section. All data are available in three formats: table, GFF3, and FASTA. GFF3 annotation files for TEs and

pre-miRNAs can be directly loaded into publicly available assemblies using a genome browser tool (e.g., Artemis).

Identification of TE-MIRs

Since we found some inconsistencies between genome versions in miRBase and those in TE curated databases, we started our analyses by remapping reference TEs to standardized versions of plant genomes (Table S1). Figure 2 summarizes our approach to finding TE-MIR associations; in brief, we searched for positional intersections between pre-miRNAs and TEs. Among 10 species, we obtained a total of 152 pre-miRNAs that overlapped at least one TE (Table 1). Most pre-miRNAs were associated to MITES (Fig. 3) and DNA transposons (Fig. 4). Our analysis provided 60 new cases of high confidence repetition-related miRNAs (Fig. 4, Fig. S1).

Fig. 6 Juxtaposed TEs possibly structuring pre-miRNAs in grasses. Both cases show an inverted insertion of the same TE. In *Sorghum bicolor*, two DNA-3-2N.SBi non-autonomous DNA transposons span sbi-MIR6227 locus (a). In *Oryza sativa*, two OLO24 non-autonomous DNA transposons intersect osa-MIR1441 (b)



Nine pre-miRNAs (ath-MIR401, ath-MIR854a, ath-MIR854b, ath-MIR854c, ath-MIR854d, ath-MIR855, osa-MIR812b, osa-MIR814a, and osa-MIR814b) indicated by Piriyaopongsa and Jordan (2008) as the products of siRNA-miRNA dual coding TEs were confirmed by our analyses. Four members in a rice miRNA family (osa-MIR812f, osa-MIR812h, osa-MIR812i, and osa-MIR812j) and osa-MIR1850 were formerly classified as typical TE-MIRs (Li et al. 2011), since they are in conformity to a standardized protocol of miRNA annotation rules (Meyers et al. 2008). Ten miRNAs are indicated by miRBase to have transcriptional evidence, and ten were found in miRNEST deep-seq predictions. Only two of them are present in both repositories.

To our knowledge, few small RNA precursors have been reported to be formed by juxtaposed TE insertions in plant species (Kuang et al. 2009; Li et al. 2011; Zhang et al. 2011). One of these cases is osa-MIR1879, which was classified as a bona fide miRNA spanning two short non-autonomous retrotransposons. Other two pre-miRNAs (osa-MIR815b and osa-MIR815c) have similar structures, but were suggested as potential pre-evolved miRNAs (Li et al. 2011). Our analysis detected these pre-miRNAs, but they intersected only one TE. However, we found three cases of stem-loop structures formed by TE juxtaposition in potato, sorghum, and rice. In *Solanum tuberosum*, two SONATA2 non-autonomous DNA transposons on the same strand compose the stu-MIR8019 foldback structure (Fig. 5). In *Sorghum bicolor*, two DNA-3-2N.Sbi non-autonomous DNA transposons give rise to sbi-MIR6227 (Fig. 6a). Within *Oryza sativa*, the insertion of two OLO24 non-autonomous DNA transposons on opposite strands probably gave rise to osa-MIR1441 (Fig. 6b).

Using an adapted Reciprocal Best BLAST Hit method, we found that 92.11 % of the matches were species-specific. This result emphasizes that the repetitive element-related miRNAs tend to be species-specific (Sun et al. 2012).

Conclusions

To our knowledge, PlanTE-MIR DB is the first resource storing the putative relationship between TEs and miRNAs in plants. The database delivers, through a user-friendly web interface, several file formats to facilitate understanding and use of the available data. Future versions will update the database to support data provided by other studies. The discovery of new TE-MIRs strongly relies on comprehensive TE annotation, which is still a drawback for several species. Thus, de novo TE annotation for organisms for which there is available data in miRBase would be a valuable resource that would promote future discoveries. Also, new releases of Repbase Update, miRBase, and

plant species genome assemblies should be considered in the next versions of PlanTE-MIR DB. In conclusion, we present a new resource, PlanTE-MIR DB, which allowed us to find new TE-MIR overlaps. We believe that PlanTE-MIR DB can supply insight into the evolution of TEs and miRNAs that will be of great value to the scientific community interested in this subject.

Acknowledgments We thank Romain Guyot (*Institute de recherche pour le développement* - IRD, Montpellier, France) for insightful comments on the TE annotation methods and for his web interface suggestions. APRL received a CAPES fellowship, and GYAdA received a *Fundação Araucária* fellowship. DSD studies on transposable elements are funded by a CAPES/CNPq “Science without borders” grant (process 084/13). This manuscript was reviewed by a professional science editor and by a native English-speaking copy editor to improve readability.

References

- Allen E, Xie Z, Gustafson AM, Sung GH, Spatafora JW, Carrington JC (2004) Evolution of microRNA genes by inverted duplication of target gene sequences in *Arabidopsis thaliana*. *Nat Genet* 36(12). doi:10.1038/ng1478
- Axtell MJ (2013) Classification and comparison of small RNAs from plants. *Annual review of plant biology* 64(January), doi:10.1146/annurev-arplant-050312-120043
- Baidouri ME, Panaud O (2013) Comparative genomic paleontology across plant kingdom reveals the dynamics of TE-driven genome evolution. *Genome Biol Evol* 5(5). doi:10.1093/gbe/evt025
- Barrera-Figueroa BE, Gao L, Wu Z, Zhou X, Zhu J, Jin H, Liu R, Zhu JK (2012) High throughput sequencing reveals novel and abiotic stress-regulated microRNAs in the inflorescences of rice. *BMC Plant Biol* 12(1). doi:10.1186/1471-2229-12-132
- Bennetzen JL, Wang H (2014) The contributions of transposable elements to the structure, function, and evolution of plant genomes. *Annual review of plant biology* 65(February), doi:10.1146/annurev-arplant-050213-035811
- Budak H, Akpinar BA (2015) Plant miRNAs: biogenesis, organization and origins. *Funct Integr Genom* 15(5). doi:10.1007/s10142-015-0451-2
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. *BMC Bioinform* 10. doi:10.1186/1471-2105-10-421
- Erson-Bensan AE (2014) Introduction to microRNAs in biological systems. 26. In: *miRNomics: MicroRNA Biology and Computational Analysis*. Springer
- Gim JA, Ha HS, Ahn K, Kim DS, Kim HS (2014) Genome-wide identification and classification of microRNAs derived from repetitive elements. *Genom Inf* 12(4)
- Hadjjargyrou M, Delilhas N (2013) The intertwining of transposable elements and non-coding RNAs. *Int J Mol Sci* 14(7). doi:10.3390/ijms140713307
- Jurka J, Kapitonov VV, Pavlicek a KlonowskiP, Kohany O, Walichewicz J (2005) Repbase Update, a database of eukaryotic repetitive elements. *Cy35 Togenet Genome Re* 110(1-4). doi:10.1159/000084979
- Kozomara A, Griffiths-Jones S (2014) MiRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res* 42(D1). doi:10.1093/nar/gkt1181

- Kuang H, Padmanabhan C, Li F, Kamei A, Bhaskar PB, Ouyang S, Jiang J, Robin Buell C, Baker B (2009) Identification of miniature inverted-repeat transposable elements (MITEs) and biogenesis of their siRNAs in the Solanaceae: New functional implications for MITEs. *Genome Res* 19(1). doi:[10.1101/gr.078196.108](https://doi.org/10.1101/gr.078196.108)
- Kurtoglu KY, Kantar M, Budak H (2014) New wheat microRNA using whole-genome sequence. *Funct Integr Genom* 14(2). doi:[10.1007/s10142-013-0357-9](https://doi.org/10.1007/s10142-013-0357-9)
- Levy A, Sela N, Ast G (2008) TranspoGene and microTranspoGene: Transposed elements influence on the transcriptome of seven vertebrates and invertebrates. *Nucleic Acids Res* 36(SUPPL. 1). doi:[10.1093/nar/gkm949](https://doi.org/10.1093/nar/gkm949)
- Li Y, Li C, Xia J, Jin Y (2011) Domestication of transposable elements into microRNA genes in plants. *PLoS ONE* 6(5). doi:[10.1371/journal.pone.0019212](https://doi.org/10.1371/journal.pone.0019212)
- Lisch D (2013) How important are transposons for plant evolution? *Nat Rev Genet* 14(1). doi:[10.1038/nrg3374](https://doi.org/10.1038/nrg3374)
- Lorenz R, Bernhart SH, Höner zu Siederdisen C, Tafer H, Flamm C, Stadler PF, Hofacker IL (2011) ViennaRNA Package 2.0. *Algo Mol Biol* 6(1). doi:[10.1186/1748-7188-6-26](https://doi.org/10.1186/1748-7188-6-26)
- Meyers BC, Axtell MJ, Bartel B, Bartel DP, Baulcombe D, Bowman JL, Cao X, Carrington JC, Chen X, Green PJ, Griffiths-Jones S, Jacobsen SE, Mallory AC, Ra M, Poethig RS, Qi Y, Vaucheret H, Voinnet O, Watanabe Y, Weigel D, Zhu JK (2008) Criteria for annotation of plant MicroRNAs. *Plant cell* 20(12). doi:[10.1105/tpc.108.064311](https://doi.org/10.1105/tpc.108.064311)
- Ou-Yang F, Luo QJ, Zhang Y, Richardson CR, Jiang Y, Rock CD (2013) Transposable element-associated microRNA hairpins produce 21-nt sRNAs integrated into typical microRNA pathways in rice. *Funct Integr Genom* 13(2). doi:[10.1007/s10142-013-0313-8](https://doi.org/10.1007/s10142-013-0313-8)
- Piriyapongsa J, Jordan IK (2008) Dual coding of siRNAs and miRNAs by plant transposable elements. *RNA (New York, NY)* 14(5). doi:[10.1261/rna.916708](https://doi.org/10.1261/rna.916708)
- Quinlan AR, Hall IM (2010) BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* 26(6). doi:[10.1093/bioinformatics/btq033](https://doi.org/10.1093/bioinformatics/btq033)
- Ragupathy R, You FM, Cloutier S (2013) Arguments for standardizing transposable element annotation in plant genomes. *Trends Plant Sci* 18(7). doi:[10.1016/j.tplants.2013.03.005](https://doi.org/10.1016/j.tplants.2013.03.005)
- Rice P, Longden I, Bleasby A (2000) EMBOSS: the European molecular biology open software suite. *Trends Genet* 16(6). doi:[10.1016/j.cocis.2008.07.002](https://doi.org/10.1016/j.cocis.2008.07.002)
- Roberts JT, Ea Cooper, Favreau CJ, Howell JS, Lane LG, Mills JE, Newman DC, Perry TJ, Russell ME, Wallace BM, Borchert GM (2013) Continuing analysis of microRNA origins: Formation from transposable element insertions and noncoding RNA mutations. *Mob Genet Elem* 3(6). doi:[10.4161/mge.27755](https://doi.org/10.4161/mge.27755)
- Roberts JT, Cardin SE, Borchert GM (2014) Burgeoning evidence indicates that microRNAs were initially formed from transposable element sequences. *Mobile Genet Elem* 4. doi:[10.4161/mge.29255](https://doi.org/10.4161/mge.29255)
- Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Ma Rajandream, Barrell B (2000) Artemis: sequence visualization and annotation. *Bioinformatics (Oxford England)* 16(10). doi:[10.1093/bioinformatics/16.10.944](https://doi.org/10.1093/bioinformatics/16.10.944)
- Sun J, Zhou M, Mao Z, Li C (2012) Characterization and evolution of microRNA genes derived from repetitive elements and duplication events in plants. *PLoS ONE* 7(4). doi:[10.1371/journal.pone.0034092](https://doi.org/10.1371/journal.pone.0034092)
- Szceśniak MW, Makałowska I (2014) MiRNEST 2.0: A database of plant and animal microRNAs. *Nucleic Acids Research* 42(D1). doi:[10.1093/nar/gkt1156](https://doi.org/10.1093/nar/gkt1156)
- Tempel S, Pollet N, Tahi F (2012) ncRNAclassifier: a tool for detection and classification of transposable element sequences in RNA hairpins. *BMC Bioinform* 13(1). doi:[10.1186/1471-2105-13-246](https://doi.org/10.1186/1471-2105-13-246)
- Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O, Paux E, San-Miguel P, Schulman AH (2007) A unified classification system for eukaryotic transposable elements. *Nat Rev Genet* 8(12). doi:[10.1038/nrg2165-c4](https://doi.org/10.1038/nrg2165-c4)
- Zhang Y, Jiang WK, Gao LZ (2011) Evolution of microRNA genes in *Oryza sativa* and *Arabidopsis thaliana*: An update of the inverted duplication model. *PLoS ONE* 6(12). doi:[10.1371/journal.pone.0028073](https://doi.org/10.1371/journal.pone.0028073)
- Zhou M, Sun J, Wang QH, Song LQ, Zhao G, Wang HZ, Yang HX, Li X (2011) Genome-wide analysis of clustering patterns and flanking characteristics for plant microRNA genes. *FEBS J* 278(6). doi:[10.1111/j.1742-4658.2011.08008.x](https://doi.org/10.1111/j.1742-4658.2011.08008.x)