

DATABASE

Open Access

TarDB: an online database for plant miRNA targets and miRNA-triggered phased siRNAs



Jing Liu, Xiaonan Liu, Siju Zhang, Shanshan Liang, Weijiang Luan and Xuan Ma*

Abstract

Background: In plants, microRNAs (miRNAs) are pivotal regulators of plant development and stress responses. Different computational tools and web servers have been developed for plant miRNA target prediction; however, in silico prediction normally contains false positive results. In addition, many plant miRNA target prediction servers lack information for miRNA-triggered phased small interfering RNAs (phasiRNAs). Creating a comprehensive and relatively high-confidence plant miRNA target database is much needed.

Results: Here, we report TarDB, an online database that collects three categories of relatively high-confidence plant miRNA targets: (i) cross-species conserved miRNA targets; (ii) degradome/PARE (Parallel Analysis of RNA Ends) sequencing supported miRNA targets; (iii) miRNA-triggered phasiRNA loci. TarDB provides a user-friendly interface that enables users to easily search, browse and retrieve miRNA targets and miRNA initiated phasiRNAs in a broad variety of plants. TarDB has a comprehensive collection of reliable plant miRNA targets containing previously unreported miRNA targets and miRNA-triggered phasiRNAs even in the well-studied model species. Most of these novel miRNA targets are relevant to lineage-specific or species-specific miRNAs. TarDB data is freely available at <http://www.biosequencing.cn/TarDB>.

Conclusions: In summary, TarDB serves as a useful web resource for exploring relatively high-confidence miRNA targets and miRNA-triggered phasiRNAs in plants.

Keywords: Plant, miRNA target, PhasiRNA, Degradome, Database

Background

In plants, microRNAs (miRNAs) are endogenous ~ 21 nucleotide (nt) non-coding RNAs, which are loaded into ARGONAUTE1 (AGO1) forming RNA-induced silencing complex (RISC) to direct RNA cleavage or translational repression of target transcripts [1–5]. Early studies well established that plant miRNAs pair with their target RNAs in a near-complementary manner [6], and demonstrated that plant miRNAs act through endonucleolytic cleavage of target RNAs [7, 8]. Meanwhile, emerging

evidence suggests that translational repression is an important mode of miRNA actions in plants [9–11].

To fully understand miRNA-target RNA interactions, miRNA target prediction and validation become vital. Plant miRNA targets can be more readily predicted as compared with animal miRNA targets, due to the extensive complementarity between miRNAs and target RNAs [12, 13]. Bioinformatics tools or web servers such as Targetfinder, psRNATarget, psRobot, comTAR, TAPIR and TarHunter have been developed to predict miRNA targets in plants [14–19]. The detailed protocols of implementing these tools were recently reviewed [20].

All above plant miRNA target prediction programs are based on in silico analysis, while the rapid development

* Correspondence: skyxma@tjnu.edu.cn

College of Life Sciences, Tianjin Key Laboratory of Animal and Plant Resistance, Tianjin Normal University, Tianjin 300387, China



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

of high throughput degradome/PARE (Parallel analysis of RNA ends) sequencing techniques have enabled to experimentally characterize miRNA cleavage sites at genome-wide scale. Accordingly, a few computational pipelines such as CleaveLand, PARESnip and sPARTA were developed to analyse degradome/PARE-seq datasets [21–23].

In addition, miRNA initiated *trans*-acting small interfering RNAs (tasiRNAs) or phased small interfering RNAs (phasiRNAs) have been implicated to play crucial roles in regulating plant growth and stress responses [24–26]. In *Arabidopsis*, phasiRNAs are predominantly 21-nt in length and are produced from limited numbers of gene loci including *TAS*, *PPR*, *AFB* and *NBS-LRR* [27, 28]. In other non-model plants, both 21-nt and 24-nt phasiRNAs were found; they are derived from hundreds to thousands of genomic loci, and a subset of them are particularly enriched in the reproductive tissues [29–38]. 22-nt miRNA has been recognized as a trigger for phasiRNA production [28, 39]. A “two-hit” model for miR390 triggered phasiRNAs at *TAS3* locus was well characterized, and miR390-*TAS3* interaction occurs in evolutionarily conserved manner [40]. By analysing polysome-bound small RNAs (sRNAs) in *Arabidopsis*, Li et al. showed that endoplasmic reticulum (ER) is an important site of phasiRNA initiation [41]. Recently, Yang et al. showed that miRNA-induced cleavage occurs on ER-bound polysomes in maize and rice [42].

Given the essential regulatory roles of miRNAs and phasiRNAs, it is highly necessary to systematically integrate miRNA target prediction, degradome/PARE-seq analysis and miRNA-triggered phasiRNA identification to create a high-confidence miRNA target database in plants. Currently, a few plant miRNA databases such as miRBase [43] and PmiREN [44] have been established; PmiREN also contains miRNA target data. PmiREN extensively focuses on miRNAs, while the miRNA target data on PmiREN are incomplete; for example, *Oryza sativa* miR2118 has over 1000 target sites in the genome, whereas PmiREN collects very limited numbers of miR2118 targets. Several plant miRNA target prediction web servers such as psRNATarget [17], psRobot [16] and WPMIAS [45] have been reported, but they lack miRNA-initiated phasiRNA information. The pipelines PHASIS (<https://github.com/atulkakrana/PHASIS>) and PhaseTank [46] were developed to predict phasiRNAs in plants. Recently, Chen et al. developed sRNAanno, a database that has comprehensive collection of phasiRNA loci in plants [47]. sRNAanno does not indicate which phasiRNA sites are triggered by miRNAs.

To this end, we have systematically analysed plant miRNA targets and miRNA-triggered phasiRNAs, and constructed TarDB database, which collects 62,888 cross-species conserved miRNA targets, 4304 degradome/

PARE-seq supported miRNA targets and 3182 miRNA-triggered phasiRNA loci. TarDB collects high-confidence miRNA targets and serves as a useful resource for future studies in plant sRNA field.

Construction and content

Data resource

The degradome/PARE-seq data used to create TarDB were downloaded from NCBI GEO or SRA databases (<http://www.ncbi.nlm.nih.gov>). For some raw sequencing data, the adaptor sequences were detected by FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>), and then were trimmed using Cutadapt (<https://cutadapt.readthedocs.io/en/stable/>). The sRNA-seq data were retrieved from NCBI GEO or Donald Danforth Plant Science Center (<http://smallrna.danforthcenter.org/>) or Dr. Blake Meyers’s lab website (<https://mpss.meyerslab.org/>). Plant genomic and transcript sequences as well as annotations were derived from JGI Phytozome (<https://phytozome.jgi.doe.gov/>). Gene ontology terms for each transcript were downloaded from Phytozome BioMart (version 12). The mature and precursor miRNA sequences were derived from miRBase (<http://www.mirbase.org/>) or PmiREN (<http://www.pmiren.com/>) or Plant sRNA Gene Server at Pennsylvania State University (<https://plantsmallrnagenes.science.psu.edu/>). The secondary structures of precursor miRNAs were generated using Perl module RNA::HairpinFigure (<https://metacpan.org/pod/RNA::HairpinFigure>). The graphs of different plant species were downloaded from www.plantgenera.org. The details of the data resources used for constructing TarDB are included in Supplementary Table S1.

Analysis procedure

Our workflow of creating TarDB is depicted on the “Guide” page (<http://www.biosequencing.cn/TarDB/guide/guide.html>), which includes three parts. In part I, the cross-species conserved miRNA targets were identified using TarHunter [18] with homo mode and score ≤ 5 . The homo mode requires the 50-nt upstream and downstream regions of miRNA target sites are cross-species conserved. Then, the results were parsed by in-house Perl scripts to generate the webpages in HTML format. In part II, The degradome/PARE-seq supported targets were identified by CleaveLand4 [21] with category ≤ 2 , Allen et al. score [12, 14] ≤ 5 and *P*-value ≤ 0.05 . The degradome signature plots in PDF format were converted to PNG format using ImageMagick (<https://imagemagick.org/index.php>) with density of 100. In part III, the phasiRNA loci were identified following previously well-documented approach [27, 28, 30, 31, 48] with minor modifications. Briefly, the processed sRNA reads were first mapped to genome using

ShortStack (<https://github.com/MikeAxtell/ShortStack>) allowing no mismatch, and the assignment of multi-mapping reads was guided by unique mapping reads (option `--mmap u`). The key parameters for executing ShortStack is as follows: `--bowtie_m 100 --ranmax 50 --mmap u --mismatches 0 --nostitch`. Next, the sRNA reads from genomic Watson and Crick strands were unified and the phasing scores were calculated as previously described [48]. Subsequently, the hypergeometric test (P -value < 0.01) was performed to obtain candidate phasiRNA loci [28, 49]. PhasiRNA analysis algorithms and scripts have been reported previously, such as *PHASIS* (<https://github.com/atulkakrana/PHASIS>) and Phase-Tank [46]. We implemented TarHunterL [18] to predict possible miRNA target sites at each phasiRNA locus, and then retrieved the loci with predicted miRNA slicing site locating at phasing positions. We performed the above steps using in-house Perl and R scripts, which enabled to automatically generate the graphs of sRNA reads profiles and phasing score plots at different phasiRNA loci. Finally, we manually inspected the graph of each phasiRNA locus to guarantee the phasing quality.

Database construction

TarDB database was placed on a web server with Linux CentOS6.2 operating system. The webpages at TarDB were created using HTML (Hypertext Markup Language) and CSS (Cascading Style Sheets), and were rendered by Bootstrap version 4.4 (<https://getbootstrap.com/>), Layui (<https://www.layui.com/>) and jQuery (<https://jquery.com/>). Several plugins were downloaded for interactive displaying, such as jsTree (<https://www.jstree.com/>) for showing interactive tree. TarDB database was managed by MySQL (<https://www.mysql.com>). PHP (pre hypertext processor, version 5.6) scripts were implemented at server end for querying MySQL database.

Utility and discussion

Database details

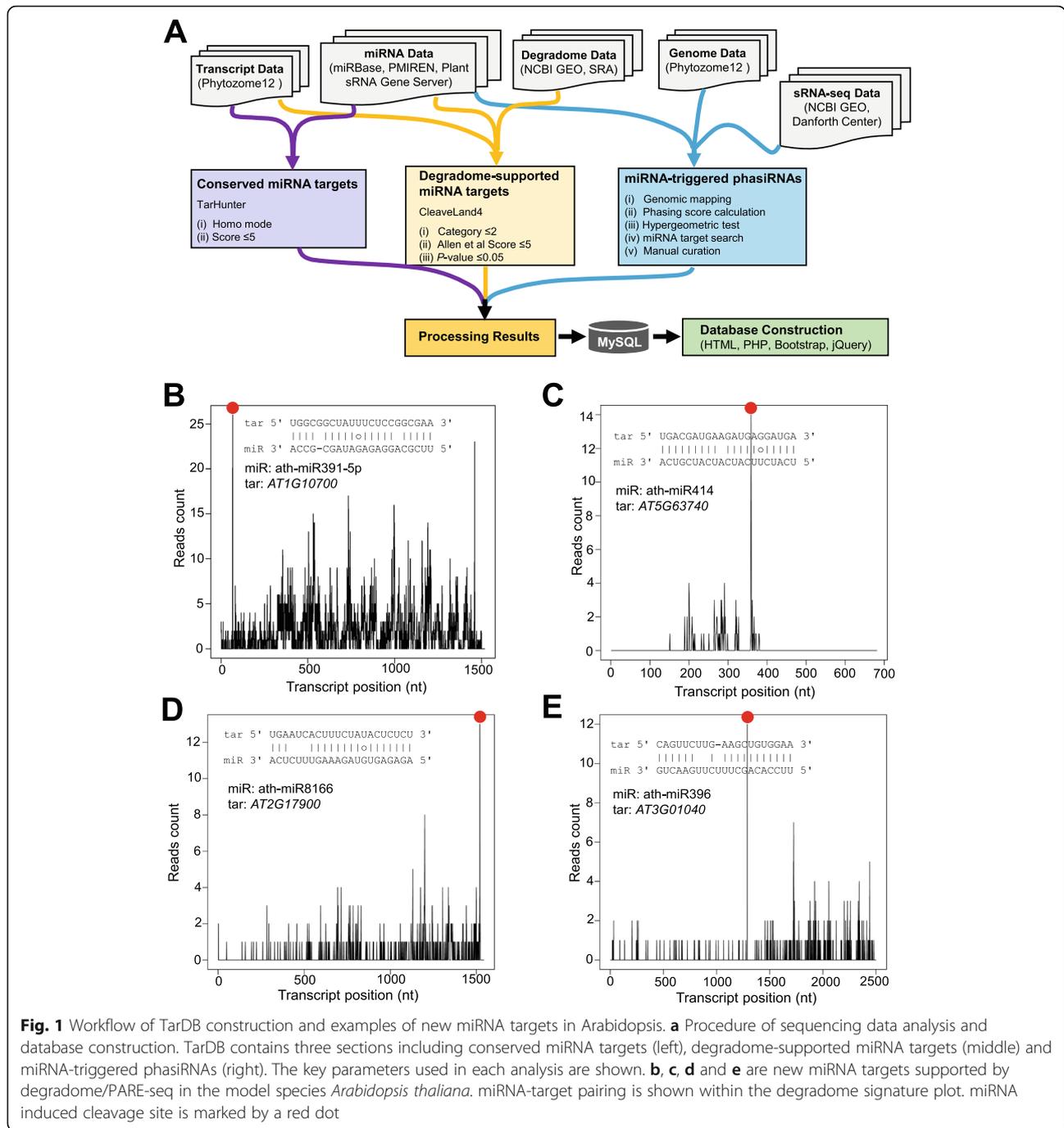
Our workflow of constructing TarDB database is depicted in Fig. 1a. The miRNA target data deposited at TarDB consist of three categories: cross-species conserved miRNA targets, degradome/PARE-seq supported miRNA targets and miRNA-triggered phasiRNAs.

The conserved miRNA targets were identified by TarHunter, our previously reported tool that is based on the rationale that homologous miRNAs target homologous sequences among diverse species [18]. TarDB collects a total of 62,888 conserved miRNA targets with cutoff score of 5, which fall into 4775 conserved groups from 43 plant species. These species range from green algae to higher flowering plants, including 24 dicotyledonous and 12 monocotyledonous plants, 1 basal

angiosperm, 1 gymnosperm, 3 bryophytes and 2 algae species. The phylogenetic relationships of these 43 species are shown in Additional file 1: Supplementary Fig. S1. Without conservation filter, TarHunter identified 539,420 miRNA-target pairs; thus, the conservation filter greatly narrows down the target gene list and increases the prediction confidence. It is worth noting that TarHunter analysis is based on *in silico* prediction of cross-species conserved miRNA target sites, and may produce false positive results. If users aim to obtain highly reliable miRNA-targeted transcripts, they can choose the degradome/PARE-seq option on TarDB.

The degradome/PARE-seq analysis was based on Phytozome annotated transcript database. Degradome/PARE-seq supported miRNA targets were identified by CleaveLand4 [21] with score ≤ 5 and P -value ≤ 0.05 . Only the data belonging to degradome categories 0, 1 and 2 data are displayed on TarDB, since these categories represent relatively reliable cleavage sites. Degradome/PARE-seq has been the most effective and high-throughput approach for capturing miRNA target sites at genome-wide scale in plants. Through analysis of 51 published degradome/PARE-seq datasets (Additional file 2: Supplementary Table S1), we obtained a total of 4304 degradome-supported high-confidence miRNA targets from 18 plants. TarDB collects novel degradome-supported miRNA targets even in the well-studied model species. Take *Arabidopsis thaliana* as an example: we identified 233 miRNA-target pairs (gene isoforms were counted once) in *A. thaliana* using the following criteria: (i) category 0 or 1; (ii) score ≤ 5 ; (iii) P -value ≤ 0.01 . The majority of these miRNA-target interactions have been characterized previously, but there remains a handful of potential new miRNA targets that need further investigations as shown below. In *Arabidopsis*, miR391 targets *PRS3* (*AT1G10700*), a P-independent phosphoribosyl pyrophosphate (PRPP) synthase gene (Fig. 1b); miR414 targets *AT5G63740*, a gene encoding RING/U-box superfamily protein (Fig. 1c); miR8166 targets *ASHR3* (*AT2G17900*) that confers histone H3 lysine-36 methylation (Fig. 1d); miR396 regulates *AT3G01040* encoding a putative galacturonosyltransferase (Fig. 1e). In addition to model species, TarDB also collects many novel degradome/PARE-seq supported miRNA targets in diverse non-model species, a few of which will be mentioned in the “Case study” section.

The miRNA-triggered phasiRNA loci were identified following previously well-documented criteria [28, 30, 31, 48] and by manual curation. Currently, many plant miRNA target prediction tools or servers (e.g., Targetfinder, psRNATarget, psRobot) lack phasiRNA analysis function. Therefore, we incorporated phasiRNA data on TarDB platform allowing users to conveniently query miRNA-triggered phasiRNAs in plants. Through analysis of 176 published sRNA-seq datasets, we obtained 2275 21-nt and 338 24-nt



miRNA-triggered phasiRNA loci from 21 species, and most of the phasiRNA triggering miRNAs are lineage specific (Additional file 1: Supplementary Fig. S1). Note that we identified a large numbers of phasiRNA candidate loci, but miRNA-triggered phasiRNAs only represent a small portion. Additionally, we discarded the phasiRNA loci with the predicted miRNA cleavage site not locating at phasiRNA register positions.

Database interface

TarDB web database has six main interfaces including “Home”, “Browse”, “Search”, “Download”, “Guide” and “Contact”. The “Home” interface presents an overview of TarDB database. It contains an introduction of miRNA target regulations, and briefly describes the prior studies on conserved miRNA targets, degradome/PARE-seq technique and miRNA-triggered

phasiRNAs in plants. It also consists of the basic statistics of TarDB data.

The “Browse” interface allows users to browse various miRNA families, diverse plant species and the three types of miRNA targets on TarDB. The miRNA sequence data are mostly derived from miRBase (release 22) [43]. Some miRNA data are from PmiREN [44] and Plant sRNA Gene server [50]. Users can view the sequences and secondary structures of mature/precursor miRNAs, and click on the corresponding external links to obtain more miRNA information (Fig. 2a). The “Browse Targets” section offers users an easy three-step way to browse any miRNA target data on TarDB (Fig. 2b). First, users need to choose miRNA target type, and then select a species which will automatically generates a miRNA list. Finally, users can click a specific miRNA on the list to get access to relevant miRNA target data.

The “Search” interface is the key section of TarDB, and it comprises three modes. In the “Search target” mode, users can search conserved miRNA targets, degradome/PARE-seq supported miRNA targets and miRNA-triggered phasiRNAs with customizable parameters such as penalty scores, maximum mispairs, degradome category, *P*-value cutoff and phasiRNA types (Fig. 2c). In the “Search locus” mode, users can query different types of miRNA targets at a specific genomic locus in a specified species (Fig. 2d). In the “keyword search” mode, users can search miRNA targets by entering a keyword, e.g., species name, miRNA or transcript IDs (Fig. 2d). The searching results are displayed in tabular format (Fig. 3a). The results can be further narrowed down using a filtering box (red dashed-line box in Fig. 3a). Each resultant record has hyperlinks that navigate to specific species, miRNA, target and evidence webpages (red arrows in Fig. 3a). The “Target” page contains transcript sequence, functional annotation and Gene Ontology (GO) information (Fig. 3b). Users can also get access to JGI Phytozome transcript website or JGI genome browser to visualize gene structure in genomic content (Fig. 3b). The “Evidence” page presents detailed supporting information for certain miRNA-target regulations. For conserved miRNA targets, miRNA-target pairing patterns and sequence alignment of homologous target sites from various species are displayed (Fig. 3c). For miRNA targets with degradome/PARE-seq evidence, the Allen et al. score [12, 14], CleaveLand4 *P*-value and the degradome signature plot highlighting miRNA cleavage position are shown (Fig. 3d). For phasiRNA loci, the sRNA-seq reads profile and phasing score plot are displayed (Fig. 3e). Within the transcript sequence, the miRNA target site is marked in red color.

The “Download” interface shows a phylogenetic tree of various plant species. Clicking on each species node

allows users to download the corresponding miRNA target data as a zip compressed file. The “Guide” interface presents our workflow of sequencing data manipulation and database construction, as well as a step-to-step guidance for exploring the key features of TarDB. The “Frequently Asked Questions (FAQs)” section on the “Guide” page provides explanations for the parameters in searching different types of miRNA targets. The “Guide” page also contains the hyperlinks that navigate to related miRNA target web resources.

Case study

Next, we present four case studies to illustrate the process of mining TarDB for identifying novel conserved miRNA targets, degradome-supported miRNA targets and miRNA-triggered phasiRNAs in plants.

Case 1

In the “Search Conserved miRNA Targets” section on the “Search” page, users can query conserved miRNA targets using a combination of parameters. The default score cutoff is set to 4. Smaller scores indicate more stringent miRNA-target complementarities. Users can set total mispair cutoff value, i.e., total mismatches and Indels (insertions and deletions). Users can also adjust seed mispair cutoff value, i.e., total mispairs at miRNA 5′ positions 2–7. miRNA seed region is crucial for miRNA-target interaction in animals and plants [51, 52]; thus, we included this parameter in miRNA target search. The “predicted cleavage” is based on the previous observation that perfect match at miRNA 5′ positions 9–11 is crucial for miRNA-mediated cleavage [53].

Actually, TarDB provides flexible ways for searching conserved miRNA targets. We take miR391 as an example. We have mentioned above that *Arabidopsis* miR391 targets a PRPP synthase gene *AT1G10700* by means of degradome/PARE-seq analysis (Fig. 1b). To view miR391-*AT1G10700* interaction in phylogenetic way, users can simply enter “miR391” in the keyword box on the “Search” page and select “cross-species conserved” target type, and then all conserved miR391 targets among different species will be displayed (Additional file 1: Supplementary Fig. S2A). Users can choose “*AT1G10700*” record to view its details (red circle in Additional file 1: Supplementary Fig. S2A). Clearly, the regulation between miR391 and PRPP synthase gene is conserved in four *Brassicaceae* species including *Arabidopsis thaliana*, *Arabidopsis lyrata*, *Capsella rubella* and *Brassica rapa* (Additional file 1: Supplementary Fig. S2B). Collectively, we can deduce that miR391 regulates PRPP synthase gene, which, to the best of our knowledge, has not been reported yet.

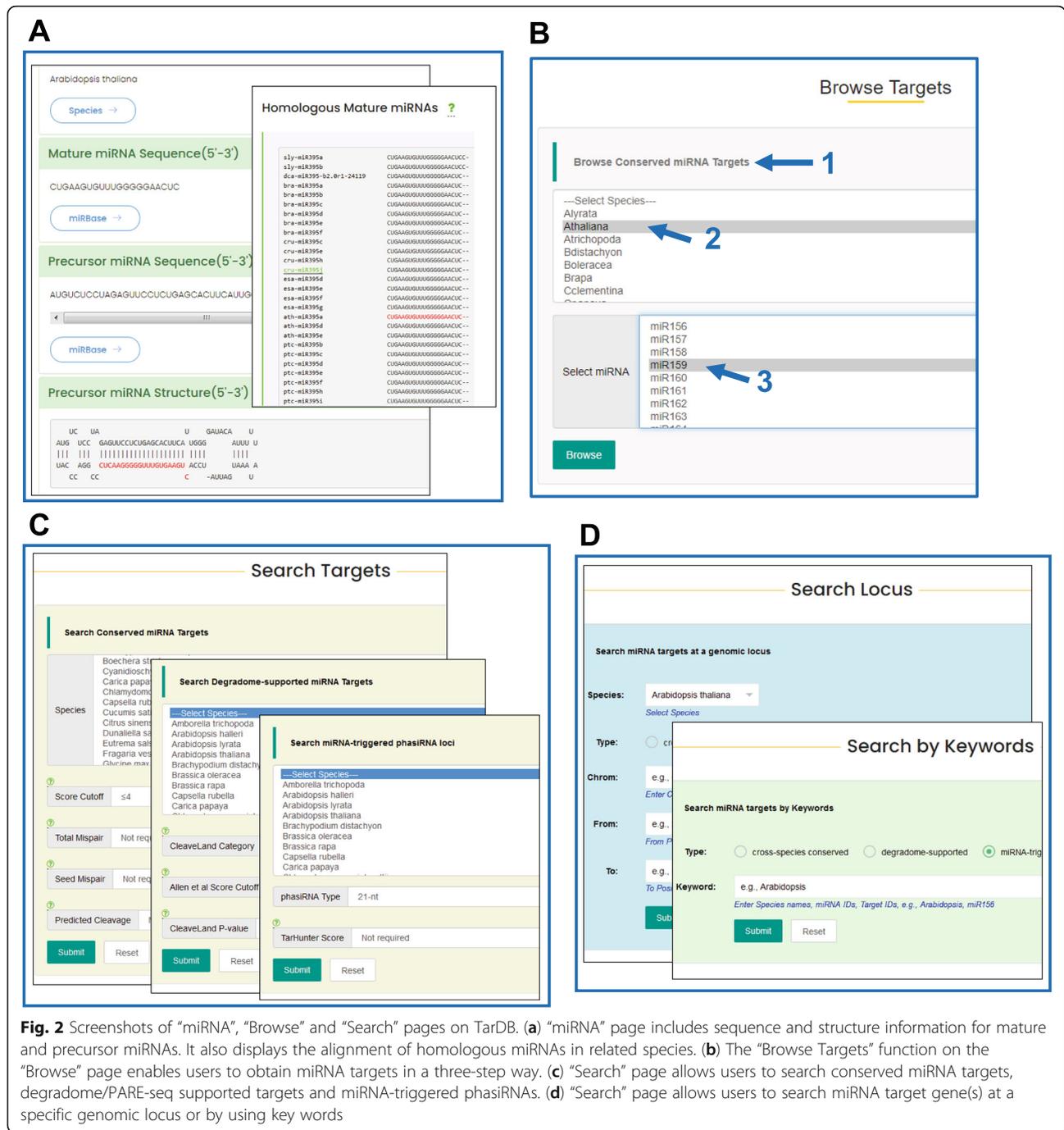


Fig. 2 Screenshots of “miRNA”, “Browse” and “Search” pages on TarDB. (a) “miRNA” page includes sequence and structure information for mature and precursor miRNAs. It also displays the alignment of homologous miRNAs in related species. (b) The “Browse Targets” function on the “Browse” page enables users to obtain miRNA targets in a three-step way. (c) “Search” page allows users to search conserved miRNA targets, degradome/PARE-seq supported targets and miRNA-triggered phasiRNAs. (d) “Search” page allows users to search miRNA target gene(s) at a specific genomic locus or by using key words

Case II

Degradome/PARE-seq provides a robust experimental evidence for miRNA directed cleavage of target RNAs in plants [54, 55]. One of the functionalities of TarDB is to search degradome/PARE-seq supported miRNA targets in various plants especially for non-model species. Take bread wheat, an important global cereal, as an example: in the “Degradome supported miRNA target” section on the “Search” page, users can choose “*Triticum aestivum*”

from the species selection box, and then simply click the “Submit” button. This returns a list of 122 wheat miRNA-target pairs with degradome/PARE-seq evidence. Normally, our default settings are sufficiently strict to identify relatively high-confidence miRNA targets. Users can adjust appropriate parameters such as increasing Allen et al. score which identifies miRNA targets with relaxed pairing. Users can also select “Category 2”, which still identifies statistically significant

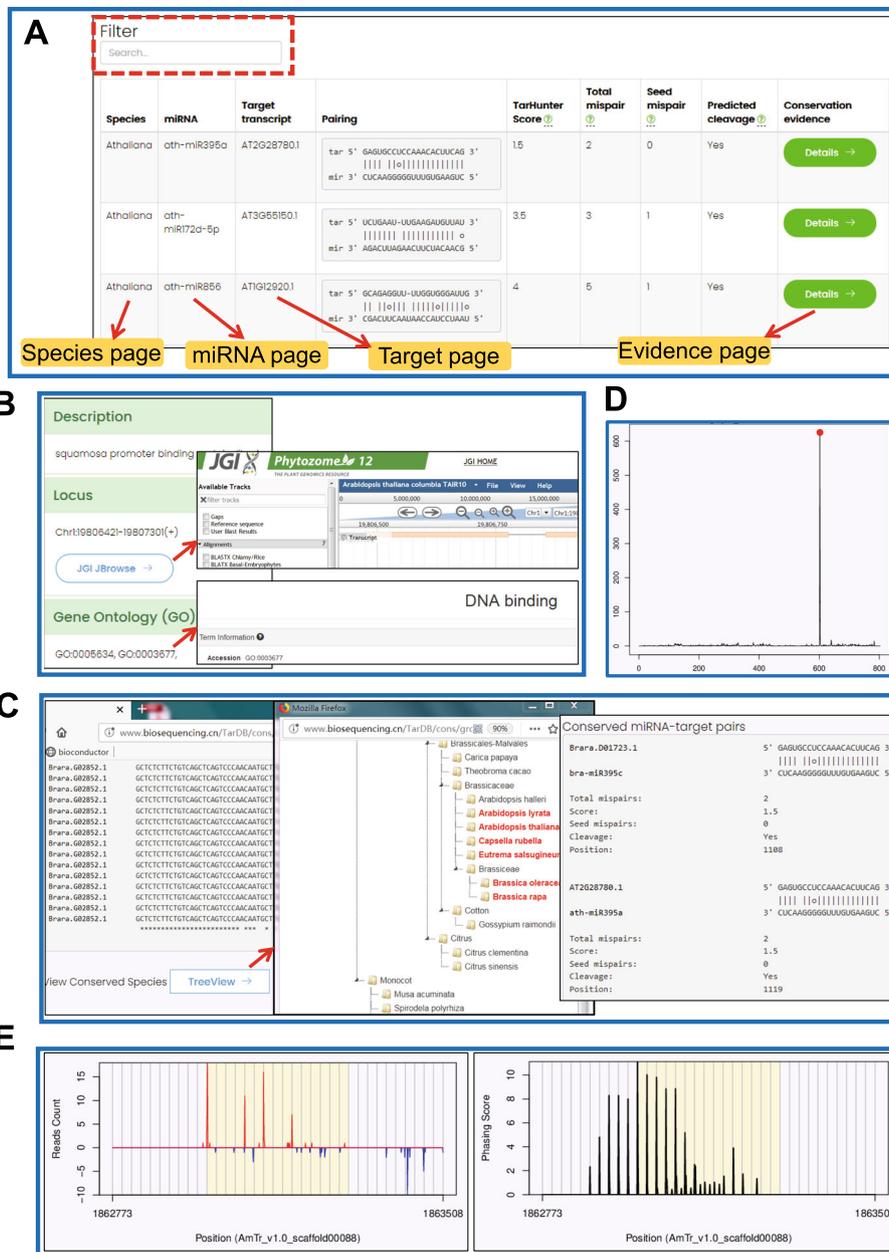
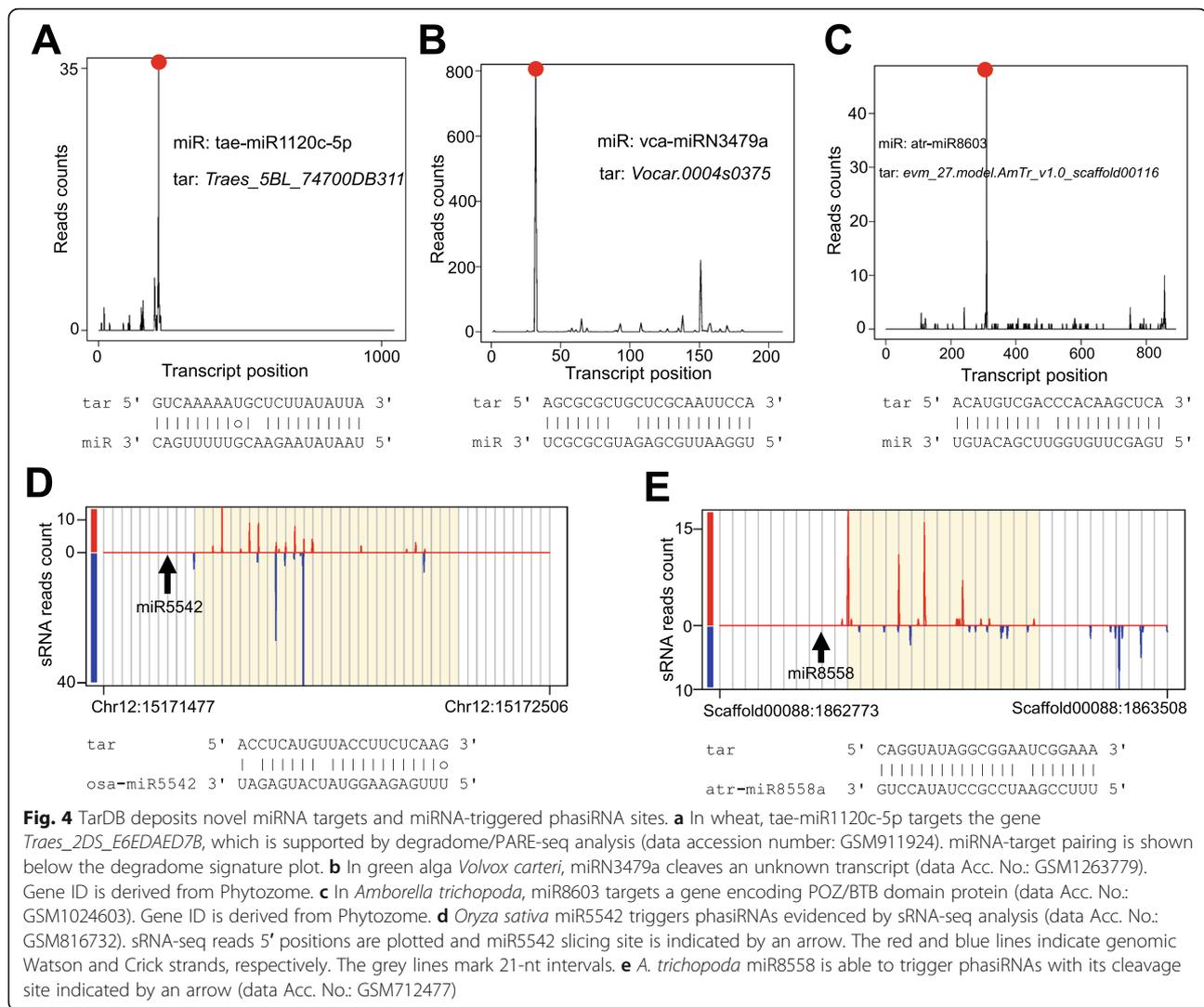


Fig. 3 Screenshots of searching results and hyperlinked pages. **a** Searching results are shown in tabular format. Red dashed-line box indicates filtering function. Querying results can be further linked to species, miRNA, target and evidence pages. **b** “Target” page has hyperlinks to Phytozome Genome Browser and contains the information of GO identifiers. **c** Alignment of conserved miRNA target sites. Clicking the “Treeview” button displays the species having conserved miRNA targets. **d** Screenshot of degradome signature plot. **e** Screenshot of sRNA-seq reads (left) and phasing score (right) profiles. The reads mapping signals at genomic Watson and Crick strands are shown in red and blue colors, respectively. 21/24-nt intervals are marked by grey lines in phasing score plot

degradome peaks but at the risk of getting false positives. Although wheat miRNA targets have been well reported [56–58], TarDB contains novel unreported wheat miRNA targets; for instance, miR1120 regulates a gene (*Traes_2DS_E6EDAED7B*) encoding peroxidase superfamily protein in wheat (Fig. 4a). Through mining TarDB, we could

identify previously undocumented miRNA targets particularly in non-model species; for examples, miRN3479a cleaves an unknown transcript in the multicellular alga *Volvox carteri* (Fig. 4b), and miR8603 targets a gene encoding POZ/BTB domain protein in the ancient angiosperm species *Amborella trichopoda* (Fig. 4c).



Case III

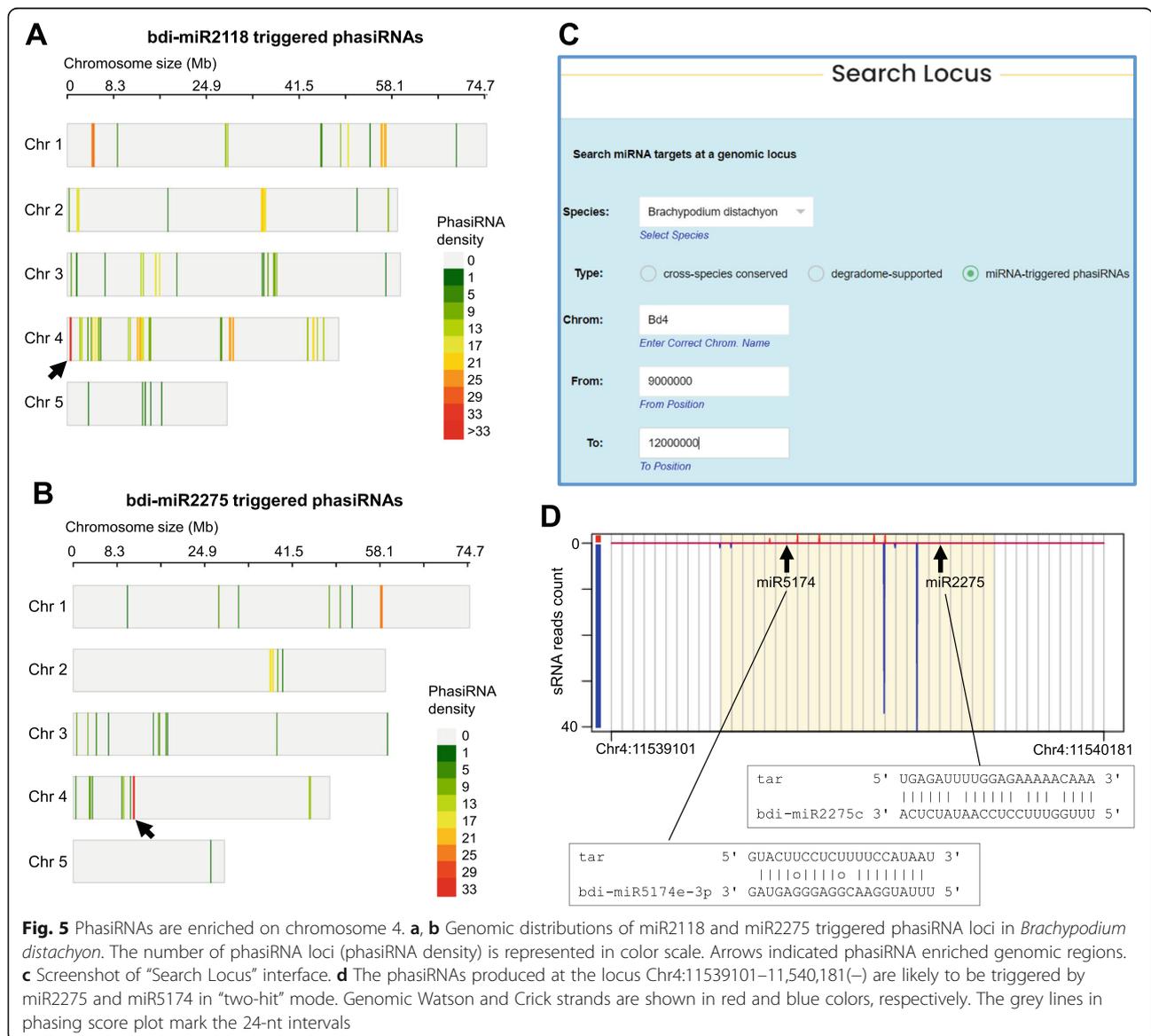
tasiRNAs/phasiRNAs belong to one of the classic types of siRNAs found in plants [25, 59]. Although many phasiRNA triggers are unknown, compelling evidence suggests that 22-nt miRNAs are capable of triggering phasiRNA biogenesis [39]. In the “two-hit” model, 21-nt miRNAs such as miR161 and miR400 are also involved in phasiRNA production [40]. The rapid increase of sRNA-seq data facilitates identification of miRNA-initiated phasiRNAs in diverse plant species. Take miR2118, which triggers 21-nt phasiRNA production in a wide spectrum of angiosperms [25], as an example. A recent study in rice (*Oryza sativa*) reveals that miR2118-dependent phasiRNAs in the anther wall are characteristic of abundant U, and are essential for anther wall development [60]. To browse miR2118-triggered phasiRNAs in rice, users can simply enter “miR2118” in the keyword box on the “Search” page, and then filter the results by entering “Osativa”, which outputs 1112

candidate miR2118-triggered phasiRNA loci in rice genome. Apart from the known phasiRNA triggers, TarDB collects new triggers; for instance, *O. sativa* specific miR5542 is able to trigger phasiRNAs with its cleavage site locating exactly at phasing position (Fig. 4d). Take *Amborella trichopoda*, which represents a basal lineage of flowering plants, as another example. To probe miRNA-triggered phasiRNAs in this species, users can choose the species “*Amborella trichopoda*” and select a phasiRNA type (21-nt or 24-nt), and then click the “Submit” button. This identifies 27 21-nt phasiRNA loci in *A. trichopoda*. Interestingly, 11 of them are triggered by miR8558, a 22-nt miRNA (Fig. 4e). Users can click “miR8558” to get the details of this miRNA; evidently, it is specific in *A. trichopoda* and has no homologs in other species. In plants, miR8558-triggered phasiRNAs has not been reported; thus, TarDB provides a platform for mining novel miRNA-triggered phasiRNAs in plants.

Case IV

It has been well studied that in grass species, miR2118 and miR2275 trigger 21-nt and 24-nt phasiRNAs, respectively, during plant reproductive stage [33, 35]. Recently, Tian et al. reports evolutionary analysis of miR2118 and miR2275 triggered phasiRNAs in five *Oryza* species, and their study strengthens the viewpoint that phasiRNAs are able to cleave phasiRNA precursor in *cis* manner [61]. We examined miR2118 and miR2275 triggered phasiRNAs in *Brachypodium distachyon*, a cereal grain species closely related to the Triticeae crops. TarDB collects 499 miR2118 triggered phasiRNA loci and 194 miR2275 triggered phasiRNA loci in *B. distachyon*. Analysis of the genomic distributions of these phasiRNA loci shows that they are enriched on

Chromosome 4 (Fig. 5a,b). Notably, both miR2118 and miR2275 triggered phasiRNAs are abundantly produced from two distinct regions on Chromosome 4 (indicated by arrows in Fig. 5a,b), which may be similar to the phasiRNA supercluster found in rice. TarDB has a “Search locus” function on the “Search” page, allowing users to search miRNA targets in specific genomic regions. For instance, with this function, users can query the phasiRNAs at positions 9–12 Mb on Chromosome 4 in *B. distachyon* (Fig. 5c). This returns a full list of 21-nt and 24-nt phasiRNA loci; 12 of them are triggered by miR2118, 4 are triggered by miR5163, and 51 are triggered by miR2275. Hence, miR2275-initiated phasiRNA loci are remarkably enriched at this 3 Mb region on Chromosome 4. Among them, the locus Chr4:



11539101–11,540,181(–) may produce a transcript cleaved by miR2275 and miR5174 to induce 24-nt phasiRNAs in “two-hit” manner, as shown in Fig. 5d.

Taken together, TarDB will be useful for plant biologists who seek for high-confidence miRNA target sites. TarDB collects the targets of a comprehensive list of plant miRNAs, including both highly conserved and less conserved miRNAs. According to the data deposited on TarDB, we compiled the phylogenetic distributions of different miRNA families in plants, and labeled the miRNAs that have degradome/PARE-seq supported targets or are able to trigger phasiRNAs (Additional file 1: Supplementary Fig. S1). Many previous efforts were devoted to characterizing highly conserved miRNAs and their targets in plants; in recent years, the lineage-specific miRNAs have been attracting more attentions. On TarDB, the great majority of lineage-specific miRNAs have either degradome or phasiRNA evidence (Additional file 1: Supplementary Fig. S1), indicating lineage-specific miRNAs are evolutionarily functional and may need future investigations. TarDB also collects the targets for a large number of species-specific miRNAs having degradome-seq or sRNA-seq supported targets, which are not shown in Additional file 1: Supplementary Fig. S1. These species-specific miRNAs have proved to play pivotal roles in plant development and stress responses [5, 62, 63], and TarDB offers a platform to search their targets.

During degradome/PARE-seq analysis, we found many datasets of non-model species have limited sequencing depth; comparatively, the model species *A. thaliana* has numerous degradome/PARE-seq data derived from various tissues and treatments. Therefore, with the rapid development of NGS (next-generation sequencing) technologies, sequencing-based efforts toward non-model plants are highly needed. Acquisition of the degradomes/small RNAs from a great diversity of non-model species will provide further insights into miRNA-target co-evolution in plants. With the increasing genomic and sequencing data, we will update TarDB regularly and include more plant species and miRNA target data in the future.

Conclusions

Here, we introduce TarDB, a miRNA target and miRNA-triggered phasiRNA database, which implements cross-species conservation and experimental filters to obtain relatively reliable miRNA targets. TarDB provides rich information and serves as a useful web resource for exploring high-confidence miRNA targets in plants. TarDB can be freely accessed at <http://www.biosequencing.cn/TarDB>.

Abbreviations

sRNA: Small RNA; miRNA: microRNA; phasiRNA: Phased small interfering RNA; tasiRNA: Trans-acting siRNA; AGO: ARGONAUTE; RISC: RNA-Induced silencing

complex; PARE: Parallel analysis of RNA ends; ER: Endoplasmic reticulum; PRPP: Phosphoribosyl pyrophosphate

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-021-07680-5>.

Additional file 1: Supplementary Fig. S1. Phylogenetic overview of plant miRNA targets with degradome or phasiRNA support. Phylogenetic relationship of 43 plant species on TarDB. Different plant groups are shaded in different colors. Plant miRNAs are generally divided into highly conserved and lineage-specific miRNA families; those species-specific miRNAs are not shown. According to TarDB data, the miRNAs that have degradome/PARE-seq supported targets are underlined in red, and the 22-nt miRNAs that trigger phasiRNAs are underlined in blue. **Supplementary Fig. S2.** The regulation between miR391 and PRPP synthase gene is conserved in *Brassicaceae* species. **(A)** Screenshot of miR391 target searching result. Red dashed line circle indicates the link to view the details of miR391-*AT1G10700* regulation. **(B)** The regulation between miR391 and PRPP synthase gene is conserved in four *Brassicaceae* species that are highlighted in red colors. Sequence alignments of different miR391 target sites are shown.

Additional file 2: Supplementary Table S1. Degradome/PARE-seq and small RNA-seq data used for TarDB database construction.

Acknowledgements

The authors would like to thank other members in Dr. Weijiang Luan's lab for helps with this work.

Authors' contributions

X.M. and W.L. conceived the study; J.L., X.L., S.Z. and S.L. downloaded the high throughput sequencing datasets, J.L., X.L., S.Z., S.L. and X.M. performed bioinformatics analysis and constructed the database; X.M. and W.L. wrote the manuscript. All authors have read and approved the manuscript.

Funding

This work was supported by grants from National Science Foundation of China (No. 31871321) and Tianjin Rice Industrial Technology System of China (No. ITTRS2018006).

Availability of data and materials

TarDB can be freely accessed at <http://www.biosequencing.cn/TarDB>.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 31 January 2021 Accepted: 5 May 2021

Published online: 13 May 2021

References

1. Axtell MJ. Classification and comparison of small RNAs from plants. *Annu Rev Plant Biol.* 2013;64(1):137–59. <https://doi.org/10.1146/annurev-arplant-050312-120043>.
2. Borges F, Martienssen RA. The expanding world of small RNAs in plants. *Nat Rev Mol Cell Biol.* 2015;16(12):727–41. <https://doi.org/10.1038/nrm4085>.
3. Voinnet O. Origin, biogenesis, and activity of plant microRNAs. *Cell.* 2009; 136(4):669–87. <https://doi.org/10.1016/j.cell.2009.01.046>.
4. Yu Y, Jia T, Chen X. The 'how' and 'where' of plant microRNAs. *New Phytologist.* 2017;216(4):1002–17. <https://doi.org/10.1111/nph.14834>.

5. Song X, Li Y, Cao X, Qi Y. MicroRNAs and their regulatory roles in plant-environment interactions. *Annu Rev Plant Biol.* 2019;70(1):489–525. <https://doi.org/10.1146/annurev-arplant-050718-100334>.
6. Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel DP. Prediction of plant microRNA targets. *Cell.* 2002;110(4):513–20. [https://doi.org/10.1016/S0092-8674\(02\)00863-2](https://doi.org/10.1016/S0092-8674(02)00863-2).
7. Llave C, Xie Z, Kasschau KD, Carrington JC. Cleavage of *Scarecrow-like* mRNA targets directed by a class of *Arabidopsis* miRNA. *Science.* 2002;297(5589):2053–6. <https://doi.org/10.1126/science.1076311>.
8. Tang G, Reinhart BJ, Bartel DP, Zamore PD. A biochemical framework for RNA silencing in plants. *Genes Dev.* 2003;17(1):49–63. <https://doi.org/10.1101/gad.1048103>.
9. Brodersen P, Sakvarelidze-Achard L, Bruun-Rasmussen M, Dunoyer P, Yamamoto YY, Sieburth L, et al. Widespread translational inhibition by plant miRNAs and siRNAs. *Science.* 2008;320(5880):1185–90. <https://doi.org/10.1126/science.1159151>.
10. Chen X. A microRNA as a translational repressor of APETALA2 in *Arabidopsis* flower development. *Science.* 2004;303(5666):2022–5. <https://doi.org/10.1126/science.1088060>.
11. Li S, Liu L, Zhuang X, Yu Y, Liu X, Cui X, et al. MicroRNAs inhibit the translation of target mRNAs on the endoplasmic reticulum in *Arabidopsis*. *Cell.* 2013;153(3):562–74. <https://doi.org/10.1016/j.cell.2013.04.005>.
12. Allen E, Xie Z, Gustafson AM, Carrington JC. microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell.* 2005;121(2):207–21. <https://doi.org/10.1016/j.cell.2005.04.004>.
13. Jones-Rhoades MW, Bartel DP. Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Mol Cell.* 2004;14(6):787–99. <https://doi.org/10.1016/j.molcel.2004.05.027>.
14. Fahlgren N, Carrington JC. miRNA Target Prediction in Plants. *Method Mol Biol (Clifton, NJ).* 2010;592:51–7.
15. Chorostecki U, Palatnik JF. comTAR: a web tool for the prediction and characterization of conserved microRNA targets in plants. *Bioinformatics (Oxford, England).* 2014;30(14):2066–7.
16. Wu HJ, Ma YK, Chen T, Wang M, Wang XJ. PsRobot: a web-based plant small RNA meta-analysis toolbox. *Nucleic Acids Res.* 2012;40(Web Server issue):W22–8.
17. Dai X, Zhuang Z, Zhao PX. psRNATarget: a plant small RNA target analysis server (2017 release). *Nucleic Acids Res.* 2018;46(W1):W49–54. <https://doi.org/10.1093/nar/gky316>.
18. Ma X, Liu C, Gu L, Mo B, Cao X, Chen X. TarHunter, a tool for predicting conserved microRNA targets and target mimics in plants. *Bioinformatics (Oxford, England).* 2018;34(9):1574–6.
19. Bonnet E, He Y, Billiau K, Van de Peer Y. TAPIR, a web server for the prediction of plant microRNA targets, including target mimics. *Bioinformatics (Oxford, England).* 2010;26(12):1566–8.
20. Sablok G, Yang K, Wen X. Protocols for miRNA target prediction in plants. *Method Mol Biol (Clifton, NJ).* 2019;1970:65–73.
21. Addo-Quaye C, Miller W, Axtell MJ. CleaveLand: a pipeline for using degradome data to find cleaved small RNA targets. *Bioinformatics (Oxford, England).* 2009;25(1):130–1.
22. Kakrana A, Hammond R, Patel P, Nakano M, Meyers BC. sPARTA: a parallelized pipeline for integrated analysis of plant miRNA and cleaved mRNA data sets, including new miRNA target-identification software. *Nucleic Acids Res.* 2014;42(18):e139. <https://doi.org/10.1093/nar/gku693>.
23. Folkes L, Moxon S, Woolfenden HC, Stocks MB, Szittyta G, Dalmay T, et al. PAREsnip: a tool for rapid genome-wide discovery of small RNA/target interactions evidenced through degradome sequencing. *Nucleic Acids Res.* 2012;40(13):e103. <https://doi.org/10.1093/nar/gks277>.
24. Fei Q, Xia R, Meyers BC. Phased, secondary, small interfering RNAs in posttranscriptional regulatory networks. *Plant Cell.* 2013;25(7):2400–15. <https://doi.org/10.1105/tpc.113.114652>.
25. Liu Y, Teng C, Xia R, Meyers BC. PhasiRNAs in plants: their biogenesis, genetic sources, and roles in stress responses, development, and reproduction. *Plant Cell.* 2020;32(10):3059–80. <https://doi.org/10.1105/tpc.20.00335>.
26. Bélanger S, Pokhrel S, Czymbek K, Meyers BC. Pre-meiotic, 24-nucleotide reproductive phasiRNAs are abundant in anthers of wheat and barley but not rice and maize. *Plant Physiol.* 2020;184(3):1407–23. <https://doi.org/10.1104/pp.20.00816>.
27. Howell MD, Fahlgren N, Chapman EJ, Cumbie JS, Sullivan CM, Givan SA, et al. Genome-wide analysis of the RNA-DEPENDENT RNA POLYMERASE6/DICER-LIKE4 pathway in *Arabidopsis* reveals dependency on miRNA- and tasiRNA-directed targeting. *Plant Cell.* 2007;19(3):926–42. <https://doi.org/10.1105/tpc.107.050062>.
28. Chen HM, Chen LT, Patel K, Li YH, Baulcombe DC, Wu SH. 22-nucleotide RNAs trigger secondary siRNA biogenesis in plants. *Proc Natl Acad Sci U S A.* 2010;107(34):15269–74. <https://doi.org/10.1073/pnas.1001738107>.
29. Arikiti S, Xia R, Kakrana A, Huang K, Zhai J, Yan Z, et al. An atlas of soybean small RNAs identifies phased siRNAs from hundreds of coding genes. *Plant Cell.* 2014;26(12):4584–601. <https://doi.org/10.1105/tpc.114.131847>.
30. Xia R, Xu J, Arikiti S, Meyers BC. Extensive families of miRNAs and PHAS loci in Norway spruce demonstrate the origins of complex phasiRNA networks in seed plants. *Mol Biol Evol.* 2015;32(11):2905–18. <https://doi.org/10.1093/molbev/msv164>.
31. Zhai J, Jeong DH, De Paoli E, Park S, Rosen BD, Li Y, et al. MicroRNAs as master regulators of the plant NB-LRR defense gene family via the production of phased, trans-acting siRNAs. *Genes Dev.* 2011;25(23):2540–53. <https://doi.org/10.1101/gad.177527.111>.
32. Xia R, Chen C, Pokhrel S, Ma W, Huang K, Patel P, et al. 24-nt reproductive phasiRNAs are broadly present in angiosperms. *Nat Commun.* 2019;10(1):627. <https://doi.org/10.1038/s41467-019-08543-0>.
33. Song X, Li P, Zhai J, Zhou M, Ma L, Liu B, et al. Roles of DCL4 and DCL3b in rice phased small RNA biogenesis. *Plant J Cell Mol Biol.* 2012;69(3):462–74. <https://doi.org/10.1111/j.1365-3113X.2011.04805.x>.
34. Johnson C, Kasprzewska A, Tennessen K, Fernandes J, Nan GL, Walbot V, et al. Clusters and superclusters of phased small RNAs in the developing inflorescence of rice. *Genome Res.* 2009;19(8):1429–40. <https://doi.org/10.1101/gr.089854.108>.
35. Zhai J, Zhang H, Arikiti S, Huang K, Nan GL, Walbot V, et al. Spatiotemporally dynamic, cell-type-dependent premeiotic and meiotic phasiRNAs in maize anthers. *Proc Natl Acad Sci U S A.* 2015;112(10):3146–51. <https://doi.org/10.1073/pnas.1418918112>.
36. Kakrana A, Mathioni SM, Huang K, Hammond R, Vandivier L, Patel P, et al. Plant 24-nt reproductive phasiRNAs from intramolecular duplex mRNAs in diverse monocots. *Genome Res.* 2018;28(9):1333–44. <https://doi.org/10.1101/gr.228163.117>.
37. Zhang YC, Lei MQ, Zhou YF, Yang YW, Lian JP, Yu Y, et al. Reproductive phasiRNAs regulate reprogramming of gene expression and meiotic progression in rice. *Nat Commun.* 2020;11(1):6031. <https://doi.org/10.1038/s41467-020-19922-3>.
38. Jiang P, Lian B, Liu C, Fu Z, Shen Y, Cheng Z, et al. 21-nt phasiRNAs direct target mRNA cleavage in rice male germ cells. *Nat Commun.* 2020;11(1):5191. <https://doi.org/10.1038/s41467-020-19034-y>.
39. Cuperus JT, Carbonell A, Fahlgren N, Garcia-Ruiz H, Burke RT, Takeda A, et al. Unique functionality of 22-nt miRNAs in triggering RDR6-dependent siRNA biogenesis from target transcripts in *Arabidopsis*. *Nat Struct Mol Biol.* 2010;17(8):997–1003. <https://doi.org/10.1038/nsmb.1866>.
40. Axtell MJ, Jan C, Rajagopalan R, Bartel DP. A two-hit trigger for siRNA biogenesis in plants. *Cell.* 2006;127(3):565–77. <https://doi.org/10.1016/j.cell.2006.09.032>.
41. Li S, Le B, Ma X, Li S, You C, Yu Y, et al. Biogenesis of phased siRNAs on membrane-bound polysomes in *Arabidopsis*. *Elife.* 2016;5. <https://doi.org/10.7554/eLife.22750>.
42. Yang X, You C, Wang X, Gao L, Mo B, Liu L, et al. Widespread occurrence of microRNA-mediated target cleavage on membrane-bound polysomes. *Genome Biol.* 2021;22(1):15. <https://doi.org/10.1186/s13059-020-02242-6>.
43. Kozomara A, Birgaoanu M, Griffiths-Jones S. miRBase: from microRNA sequences to function. *Nucleic Acids Res.* 2019;47(D1):D155–d162. <https://doi.org/10.1093/nar/gky1141>.
44. Guo Z, Kuang Z, Wang Y, Zhao Y, Tao Y, Cheng C, et al. PmiREN: a comprehensive encyclopedia of plant miRNAs. *Nucleic Acids Res.* 2020;48(D1):D1114–21. <https://doi.org/10.1093/nar/gkz2894>.
45. Fei Y, Mao Y, Shen C, Wang R, Zhang H, Huang J. WPMIAS: whole-degradome-based plant microRNA-target interaction analysis server. *Bioinformatics.* 2020;36(6):1937–9.
46. Guo Q, Qu X, Jin W. PhaseTank: genome-wide computational identification of phasiRNAs and their regulatory cascades. *Bioinformatics (Oxford, England).* 2015;31(2):284–6.
47. Chen C, Li J, Feng J, Liu B, Feng L, Yu X, et al. sRNAanno-a database repository of uniformly annotated small RNAs in plants. *Horticulture Res.* 2021;8(1):45. <https://doi.org/10.1038/s41438-021-00480-8>.
48. De Paoli E, Dorantes-Acosta A, Zhai J, Accerbi M, Jeong DH, Park S, et al. Distinct extremely abundant siRNAs associated with cosuppression in petunia. *RNA (New York, NY).* 2009;15(11):1965–70.

49. Chen HM, Li YH, Wu SH. Bioinformatic prediction and experimental validation of a microRNA-directed tandem trans-acting siRNA cascade in Arabidopsis. *Proc Natl Acad Sci U S A*. 2007;104(9):3318–23. <https://doi.org/10.1073/pnas.0611119104>.
50. Lunardon A, Johnson NR, Hagerott E, Phifer T, Polydore S, Coruh C, et al. Integrated annotations and analyses of small RNA-producing loci from 47 diverse plants. *Genome Res*. 2020;30(3):497–513. <https://doi.org/10.1101/gr.256750.119>.
51. Mallory AC, Reinhart BJ, Jones-Rhoades MW, Tang G, Zamore PD, Barton MK, et al. MicroRNA control of PHABULOSA in leaf development: importance of pairing to the microRNA 5' region. *EMBO J*. 2004;23(16):3356–64. <https://doi.org/10.1038/sj.emboj.7600340>.
52. Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB. Prediction of mammalian microRNA targets. *Cell*. 2003;115(7):787–98. [https://doi.org/10.1016/S0092-8674\(03\)01018-3](https://doi.org/10.1016/S0092-8674(03)01018-3).
53. Liu Q, Wang F, Axtell MJ. Analysis of complementarity requirements for plant microRNA targeting using a *Nicotiana benthamiana* quantitative transient assay. *Plant Cell*. 2014;26(2):741–53. <https://doi.org/10.1105/tpc.113.120972>.
54. Addo-Quaye C, Eshoo TW, Bartel DP, Axtell MJ. Endogenous siRNA and miRNA targets identified by sequencing of the Arabidopsis degradome. *Curr Biol*. 2008;18(10):758–62. <https://doi.org/10.1016/j.cub.2008.04.042>.
55. German MA, Pillay M, Jeong DH, Hetawal A, Luo S, Janardhanan P, et al. Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. *Nat Biotechnol*. 2008;26(8):941–6. <https://doi.org/10.1038/nbt1417>.
56. Song G, Zhang R, Zhang S, Li Y, Gao J, Han X, et al. Response of microRNAs to cold treatment in the young spikes of common wheat. *BMC Genomics*. 2017;18(1):212. <https://doi.org/10.1186/s12864-017-3556-2>.
57. Sun F, Guo G, Du J, Guo W, Peng H, Ni Z, et al. Whole-genome discovery of miRNAs and their targets in wheat (*Triticum aestivum* L.). *BMC Plant Biol*. 2014;14(1):142. <https://doi.org/10.1186/1471-2229-14-142>.
58. Tang Z, Zhang L, Xu C, Yuan S, Zhang F, Zheng Y, et al. Uncovering small RNA-mediated responses to cold stress in a wheat thermosensitive genic male-sterile line by deep sequencing. *Plant Physiol*. 2012;159(2):721–38. <https://doi.org/10.1104/pp.112.196048>.
59. Axtell MJ, Meyers BC. Revisiting criteria for plant MicroRNA annotation in the era of big data. *Plant Cell*. 2018;30(2):272–84. <https://doi.org/10.1105/tpc.17.00851>.
60. Araki S, Le NT, Koizumi K, Villar-Briones A, Nonomura KI, Endo M, et al. Komiya R: miR2118-dependent U-rich phasiRNA production in rice anther wall development. *Nat Commun*. 2020;11(1):3115. <https://doi.org/10.1038/s41467-020-16637-3>.
61. Tian P, Zhang X, Xia R, Liu Y, Wang M, Li B, et al. Evolution and diversification of reproductive phased small interfering RNAs in *Oryza* species. *New Phytol*. 2021;229(5):2970–83.
62. Tang J, Chu C. MicroRNAs in crop improvement: fine-tuners for complex traits. *Nat Plant*. 2017;3(7):17077. <https://doi.org/10.1038/nplants.2017.77>.
63. Zhang Y, Xia R, Kuang H, Meyers BC. The diversification of plant NBS-LRR defense genes directs the evolution of microRNAs that target them. *Mol Biol Evol*. 2016;33(10):2692–705. <https://doi.org/10.1093/molbev/msw154>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

