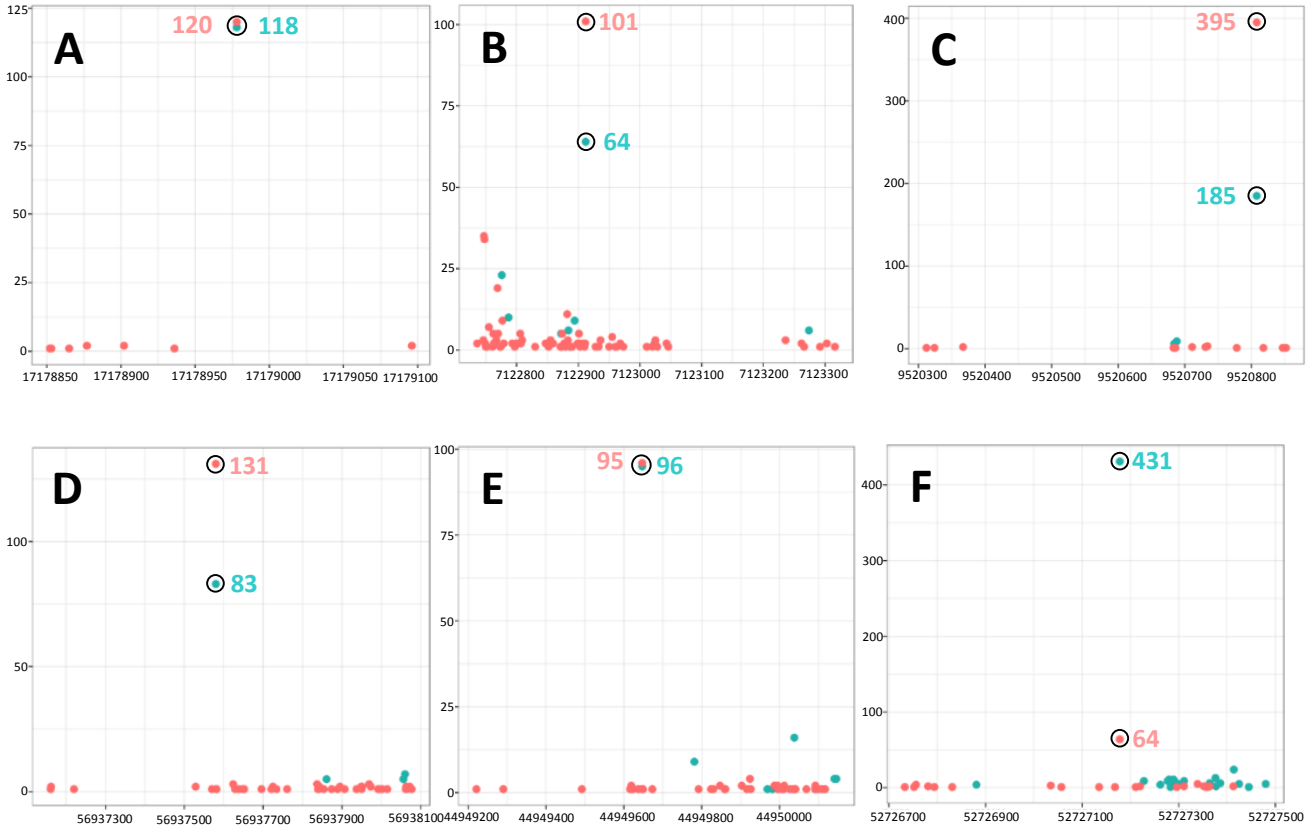
Supplemental Text

sPARTA: A parallelized pipeline for integrated analysis of plant miRNA and cleaved mRNA data sets, including new miRNA target-identification software

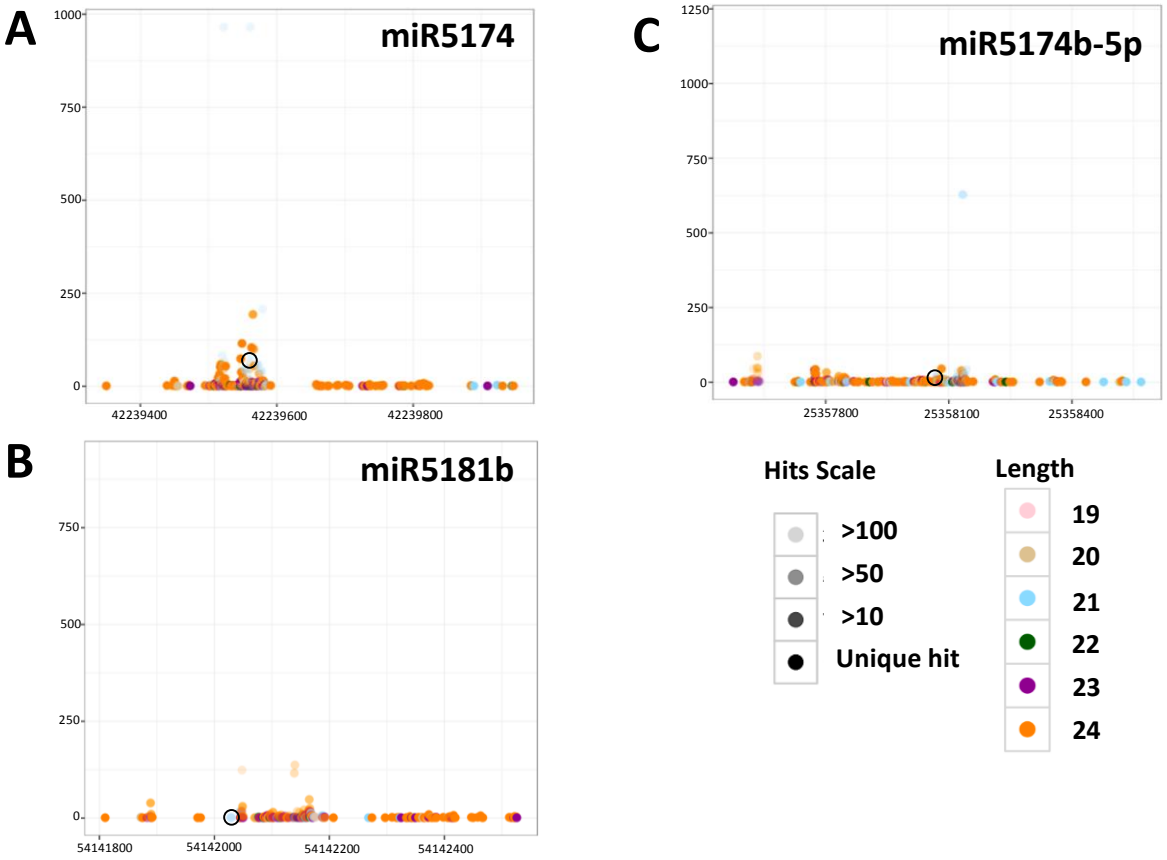
The abundance of miRNAs, crucial for making biological inferences about their potential role at a given time and space, can be easily accessed by clicking on the miRNA sequence. This is available from main results page of our website. Clicking on this opens a new window displaying the miRNA abundance across the different libraries included in the selected database. If there exists more than one locus from which the miRNA could originate, these sites are also displayed; this is useful to check the origin(s) of a small RNA to assess whether it's from a repetitive region, associated with heterochromatic siRNAs, etc. From those data, one can decide whether the miRNA of interest is likely to be real or not. If there exist multiple sRNA databases for a single species, then the user can switch between these databases, either from the main query page (Fig. 10) or by clicking on a "tab" to another website (Supplementary Figure S4). On other hand, miRNA targets can also be scrutinized at much finer level by individually clicking on them, opening a new genome browser window displaying the target site, cleavage site, and PARE reads in the vicinity (Supplementary Figure S4). If the cleavage site is situated in an intergenic region, then the user can navigate to flanking genes to ensure that the small RNA is not derived from unannotated portion of one of these adjacent genes. From this target-specific window, clicking on an individual PARE reads displays its sequence, abundance in different libraries and loci to which this PARE read maps. A more controlled target-specific genome browser instance could be accessed by clicking on the 'PARE read abundance' tab (Supplementary Figure S4). In this page, the user can adjust the 'control panel' to select libraries of interest, abundance view (summed, summed hits-normalized or individual) and library view (consolidated or library-wise). An interesting feature of our MPSS genome browser that greatly complements comPARE is a viewer showing the 'phasing score', which can be enabled from the 'control panel'. This allows the user to identify phasiRNAs in the context of a predicted miRNA cleavage site. An overview (albeit outdated) of the MPSS genome browser features and functionalities was previously published (Nakano *et al.*, 2006). Lastly, each miRNA's nomenclature, sequence information and annotation can be accessed by simply clicking on the miRBase accession identifier from the main results window, which directs the user to miRBase (Griffiths-Jones *et al.*, 2006).

Supplementary Figure 1. miR396 coordinates cell proliferation in leaf meristem by regulating transcription factors belonging to the family of GROWTH-REGULATING FACTOR (GRF).
Plots of PARE data (D-Plots) mapped to genomic regions with cleavage sites highlighted for genic targets of bdi-miR396 in *B. distachyon*. Green dots indicate PARE reads from leaf libraries, and red dots are from panicle libraries. The numbers indicate abundance of reads (in TP15M).
A) Bradi4g16450 (GRF-8 like).
B) Bradi1g09900 (GRF-6 like).
C) Bradi1g12650 (GRF-9 like) is shared between panicle and leaf. These targets encode proteins in the family of Growth Regulating Factor (GRFs).
D, E and F) Examples of novel intergenic targets of miR396 from *B. distachyon* shared between leaf and panicle.



Supplementary Figure 2. Examples of spurious annotations in miRBase.
MicroRNAs from families miR5174 and miR5181 originate from repetitive regions, rich with heterochromatic (24-nt) small RNAs. A resources which allows visualization of miRNAs and their targets in genomic context is sought to allow manual review of miRNAs in online repositories. At the bottom is a legend indicting that the intensity of the fill color indicates the hits (genome matches), while the different colors indicate the small RNA sizes.

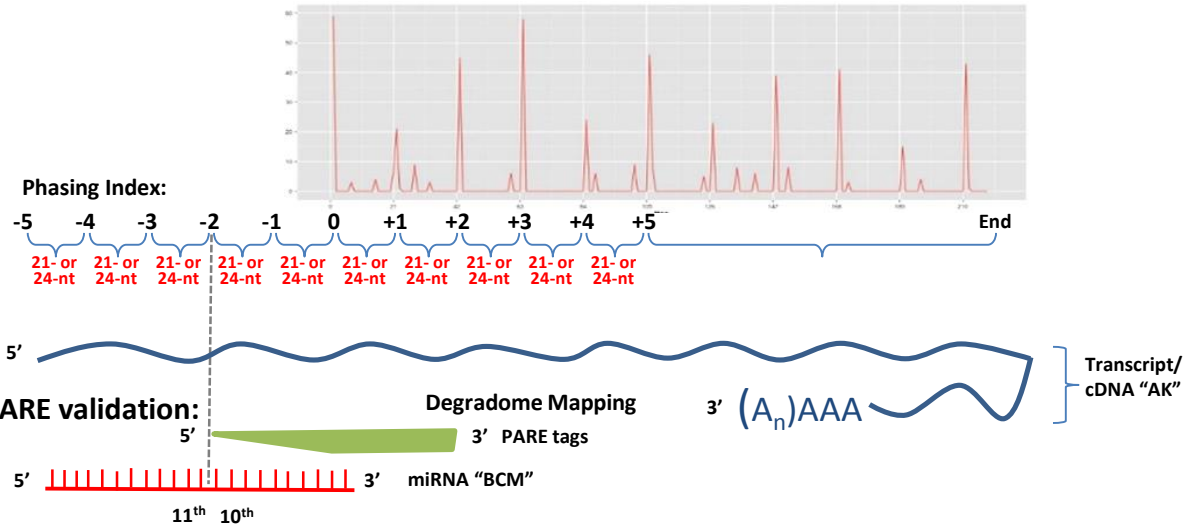
A) bdi-miR5174.
B) bdi-miR5181.
C) bdi-miR5174b.



Supplementary Figure 3. Loci generating phased small RNAs were used in the comparison of predictive power.

The phase-index consisting of 11 coordinates (+/- 5 cycles), corresponding to a phase (21 or 24-nt) periodicity from the initiation site of the phased locus, or site at which the miRNA cleaves to trigger phasiRNA biogenesis. Triggers were identified by searching for miRNA-target interactions with cleavage sites matching a specific phase-index.

1. Validation from identification of phased loci:



3. Literature-based validation:

99% of the PARE-validated triggers of loci generating 21- and 24-nt phased small RNAs were in either the miR2118 or miR2275 families (respectively).

Supplementary Figure 4. Integration of comPARE with the MPSS viewer further enhances its visualization and exploratory functionalities.

Screenshot of the target specific genome browser. i) Different databases could be switched by directly clicking on the database tabs, displaying the same genomic region for a different database and different libraries. ii) Mapped PARE reads are shown at the site of interest, including the cleavage site and in its vicinity. iii) Library-specific PARE abundances could be accessed with a separate web page, the "Library Abundance" viewer. iv) Advanced controls for the display properties are available, to change the viewing properties, altering the selection of libraries, etc.