

RESEARCH ARTICLE

PeTMbase: A Database of Plant Endogenous Target Mimics (eTMs)

Gökhan Karakulah^{1*}, Kuaybe Yücebilgili Kurtoğlu², Turgay Unver^{1*}

1 İzmir International Biomedicine and Genome Institute (iBG-izmir), Dokuz Eylül University, Inciralti, İzmir, Turkey, **2** Department of Molecular Biology and Genetics, İstanbul Medeniyet University, İstanbul, Turkey

* gokhan.karakulah@deu.edu.tr (GK); turgay.unver@deu.edu.tr (TU)

Abstract

MicroRNAs (miRNA) are small endogenous RNA molecules, which regulate target gene expression at post-transcriptional level. Besides, miRNA activity can be controlled by a newly discovered regulatory mechanism called endogenous target mimicry (eTM). In target mimicry, eTMs bind to the corresponding miRNAs to block the binding of specific transcript leading to increase mRNA expression. Thus, miRNA-eTM-target-mRNA regulation modules involving a wide range of biological processes; an increasing need for a comprehensive eTM database arose. Except miRSponge with limited number of Arabidopsis eTM data no available database and/or repository was developed and released for plant eTMs yet. Here, we present an online plant eTM database, called PeTMbase (<http://petmbase.org>), with a highly efficient search tool. To establish the repository a number of identified eTMs was obtained utilizing from high-throughput RNA-sequencing data of 11 plant species. Each transcriptome libraries is first mapped to corresponding plant genome, then long non-coding RNA (lncRNA) transcripts are characterized. Furthermore, additional lncRNAs retrieved from GREENC and PNRD were incorporated into the lncRNA catalog. Then, utilizing the lncRNA and miRNA sources a total of 2,728 eTMs were successfully predicted. Our regularly updated database, PeTMbase, provides high quality information regarding miRNA: eTM modules and will aid functional genomics studies particularly, on miRNA regulatory networks.



OPEN ACCESS

Citation: Karakulah G, Yücebilgili Kurtoğlu K, Unver T (2016) PeTMbase: A Database of Plant Endogenous Target Mimics (eTMs). PLoS ONE 11 (12): e0167698. doi:10.1371/journal.pone.0167698

Editor: Baohong Zhang, East Carolina University, UNITED STATES

Received: October 28, 2016

Accepted: November 20, 2016

Published: December 9, 2016

Copyright: © 2016 Karakulah et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: We have uploaded our data to Mendeley with DOI: <http://dx.doi.org/10.17632/htgxyrcv2.1>.

Funding: This work was supported by Turkish Academy of Science (TUBA) with grant type GEBIP.

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

Introduction

MicroRNAs (miRNAs) are a class of small RNA (~21 nt) molecules, that has emerged as key regulators of gene expression at post-transcriptional level [1]. In the current molecular framework for miRNA biogenesis, miRNA coding genes (*MIR* genes) produce primary miRNAs (pri-miRNAs) utilizing RNA polymerase II. Pri-miRNAs then undergoes two subsequent cleavage steps by DICER-LIKE1 (DCL1) enzymes to end up with miRNA/miRNA* duplex [2]. Then, the double stranded RNA molecules are loaded onto RNA-induced silencing complex (RISC) which can recognize the specific target mRNA by sequence complementarity in which mature miRNA strand is directed to its target strand and induce silencing mechanism through translational repression or site-specific cleavage of the mRNA molecule [3]. In plants, miRNAs

have vital roles in diverse set of biological processes such as growth, development, biotic/abiotic stress responses and signal transduction [4–6].

miRNA activity is also finely regulated by a recently discovered mechanism called endogenous target mimicry (eTM) [7]. Target mimics, also described as miRNA decoys, sponges or competing endogenous RNAs (ceRNAs), are generally belonged to long non-coding RNA (lncRNA) class [8]. lncRNAs are RNA transcripts that are longer than 200 nt in length and lack open reading frame [9]. Despite the small number of lncRNAs examined for their biological functions, findings indicate a regulatory role in gene expression at both transcriptional and post-transcriptional levels [10–12]. Some lncRNAs bind miRNA binding sites and blocks the interaction between miRNA and its specific target mRNA for further regulation. In target mimicry, eTMs bind to miRNAs with a three-nucleotide bulge between the 5' end 10th and 11th positions via sequence complementarity. In this way, the eTM-miRNA pairing rescues the real target transcripts to be cleaved by its complementary miRNA, leading to increased expression levels of target mRNAs. This ceRNA hypothesis that has recently gained attention, presumes the enhancing the expression of corresponding target transcripts through sequestering miRNA activity [13].

The first discovered eTM in plants, *Induced by Phosphate Starvation 1 (IPS1)* transcript was characterized in *Arabidopsis* as an endogenous lncRNA [14]. With a 3-nt bulge in the miRNA cleavage site, IPS1 binds to phosphate starvation-induced miRNA, ath-miR399. Due to the presence of the loop in base pairing, ath-miR399 cannot bind and specifically cleave the target transcript. Therefore, IPS1 serves as a miRNA target mimic (or decoy) and inhibits binding of ath-miR399 to its target transcript, PHO2. Franco-Zorilla and his colleagues observed that the IPS1 overexpressor plants accumulated increased amounts of target transcript PHO2¹⁴. Since then, several eTMs for some of the conserved miRNAs have been computationally identified in *Arabidopsis*, rice and other sequenced plant genomes [15–17]. Artificial eTMs can also induce target mimicry, thus through functional studies, biological roles of miRNAs were revealed by manipulating miRNA activity [18–20].

As this novel gene regulation network plays vital roles in a wide range of biological processes, an increasing need for a comprehensive eTM database arose. Although, a number of lncRNA databases such as TAIR (for *Arabidopsis*) [21], PlantNATsDB (natural antisense transcripts) [22], lncRNAdb (functional lncRNA database) [10], NONCODE (noncoding RNAs in only *Arabidopsis* as plant species) [23], PLncDB (with a limited number of lncRNAs in only *Arabidopsis*) [24], GreenC (lncRNA repository in plants and algae) [25], CANTATdb (lncRNA database covering ten plant species) [26] and PNRD for non-coding RNAs in four plant species [27] are available, except miRSponge covering only a very limited number of eTM data for *Arabidopsis thaliana* [28], there is no any comprehensive available database specifically developed for plant eTMs yet. To provide a high quality tool for plant scientist that can accelerate the research on this unique class of transcripts, we established an online database called PeTMbase, which deposits computationally predicted eTMs in plants.

Materials and Methods for Database Design

Collection and Processing of Plant Transcriptome Libraries

To build a comprehensive catalogue of lncRNA transcripts from available plant transcriptomes, RNA-sequencing libraries generated from leaf samples of 11 selected monocot/dicot plant species, including *Arabidopsis thaliana*, *Brachypodium distachyon*, *Glycine max* (soybean), *Hordeum vulgare* (barley), *Medicago truncatula*, *Oryza sativa* (rice), *Populus trichocarpa* (poplar), *Sorghum bicolor*, *Solanum lycopersicum* (tomato), *Triticum aestivum* (wheat), and *Zea mays* (maize), were downloaded from Sequence Read Archive (SRA) of National Center for

Biotechnology Information (NCBI) [29] by SRA Tool Kit v.2.7.0. Consequently, a total of 21 transcriptome libraries produced by the Illumina sequencing platform were included in the study (S1 Table).

The quality graphs of each library were visually examined through SRA browser, and sequence files were converted to FASTQ files using fastq-dump utility of SRA Tool Kit with—skip-technical,—clip and—split-files (for paired end libraries only) options. Sequencing reads in FASTQ files were mapped to the corresponding reference genomes using STAR aligner v2.5.1 [30] with default parameters, and the mean mapping rate of total 21 libraries was found 88.50% (min: 43.64%, max: 94.04%), indicating that the libraries were suitable for performing *ab initio* transcriptome assembly (S2 Table). To build STAR aligner indexes, the corresponding reference genome assemblies in multi-FASTA format were downloaded from the ftp site of Ensembl Plants database release 31 [31]. The *ab initio* transcriptome assembly for each plant species was performed with Cufflinks v2.2.1 [32] to generate a comprehensive catalog of novel lncRNA transcripts. Subsequently, the transcript.gtf files generated in each Cufflinks run were merged with Cuffmerge tool of the Cufflinks suite. To identify novel intergenic transcripts, each merged GTF file was then compared to the corresponding Ensembl v.31 reference annotation using cuffcompare tool. Coding potential of novel transcripts were examined using Transdecoder v.1 (<http://transdecoder.github.io/>), and sequences longer than 200 nucleotides and containing Open Reading Frame (ORF) < 100 amino acids were classified as novel lncRNAs. We also manually retrieved and included known lncRNA species, annotated by GRNC [25] and PNRD [27] plant non-coding RNA databases, to generate a complete list of lncRNAs for further analyses. A schematic diagram on the workflow of our study is shown in Fig 1.

Identification of eTMs

A total of 4,323 mature miRNA sequences of 11 plant species were retrieved from miRBase release 21 [33]. To determine putative plant eTMs, possible miRNA target mimic sites were scanned within lncRNA sequences, and marked where they were identified through a custom script written in R programming language [34] as per the following rules previously described in [17]: (i) the 2nd to 8th positions at the 5' end of a miRNA sequence must perfectly match to target lncRNA sequence, (ii) three unpaired nucleotides (bulges) are allowed between the 9th to 12th positions at the 5' end of the miRNA sequence, and (iii) at most 3 nucleotide mismatch (excluding bulge region) can be between miRNA and lncRNA sequences. Fig 2A shows an example of pairing *Z. mays* eTM, zma_eTM_miR528b-5p-19, and its potential target miRNA, zma-miR528b-5p. Consequently, a large number of lncRNA sequences meeting the binding rules described above were classified as eTMs. Using this approach, several conserved target mimic sites within eTM sequences were discovered among the selected plant species particularly, *A. thaliana*, *O. sativa*, *S. bicolor*, *T. aestivum*, and *Z. mays* (Fig 2B).

Database Construction and Development of User Interfaces

To provide a unique and searchable resource of plant eTMs, and to bring currently discovered miRNA mimicry molecules, we established an online database called PeTMbase. Computational analysis of the plant lncRNA sequences from a variety of data sources led to the identification of several novel eTMs in multiple species. Therefore, we have assigned a unique PeTMbase ID for each predicted eTM as follows: (i) the source of eTM sequence is represented by the first three letters of PeTMbase ID. For instance, it is given “zma” prefix if eTM is discovered in *Zea mays*, (ii) the following three letters are “eTM”, and (iii) the rest of PeTMbase ID comprise miRBase ID of target miRNA. In case multiple eTM sequences targeting the same miRNA, incremental

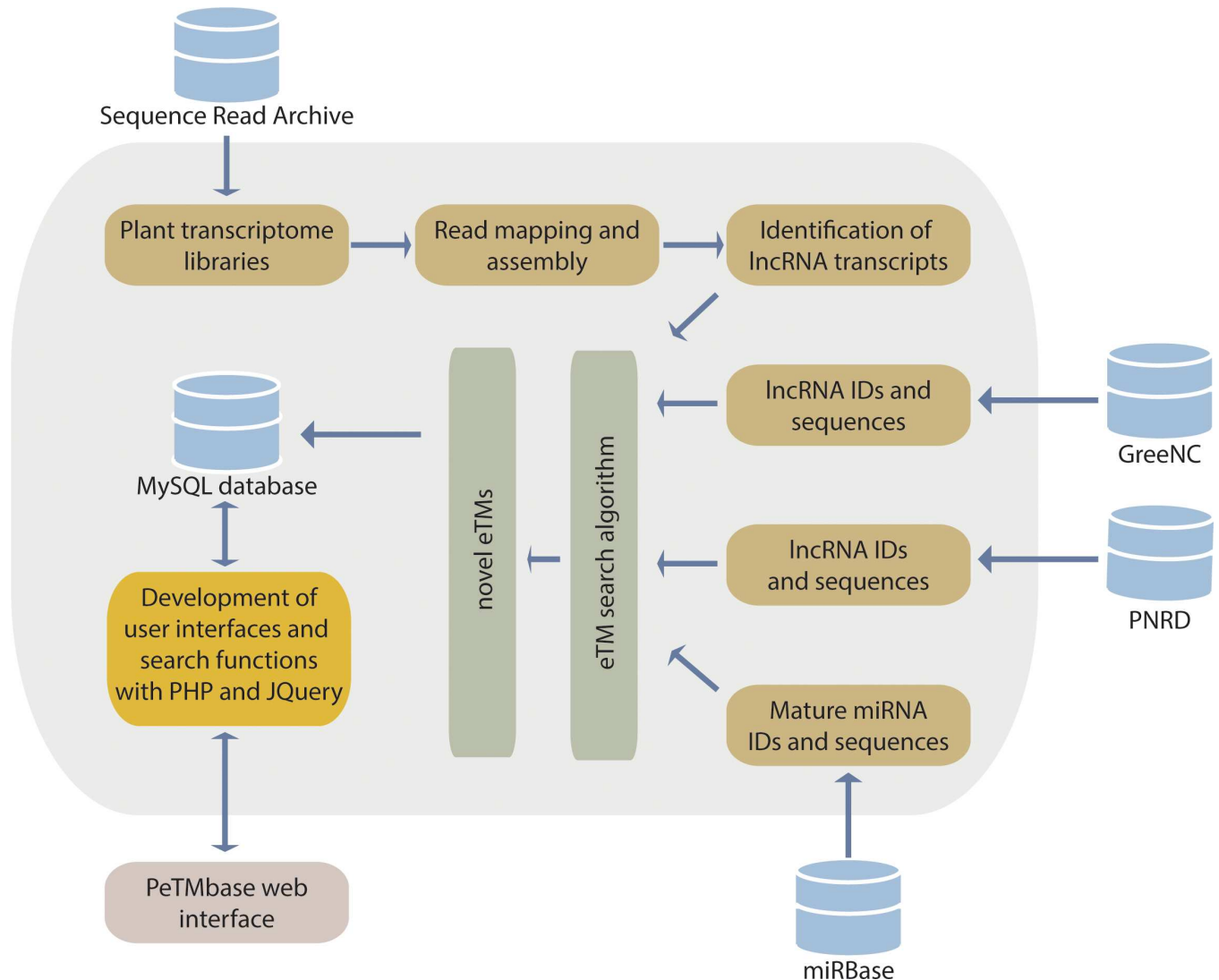


Fig 1. Workflow diagram followed to construct PeTMbase. Raw data sets from diverse public databases were incorporated to extract the biologically relevant information. Sequence libraries were downloaded from SRA database, and known lncRNA sequences retrieved from non-coding RNA databases. MiRNA binding sites within lncRNA sequences were identified with a custom R script, and flexible and user-friendly interfaces were developed diverse scripting languages and web technologies.

doi:10.1371/journal.pone.0167698.g001

integer starting from 1 is suffixed to corresponding PeTMbase ID. The database stores individual information including, PeTMbase ID, sequence, miRNA target(s), transcript properties, and associated external links (if available), for each eTM into the underlying MySQL v5.7 database (<https://www.mysql.com/>). The user interfaces were developed with Hyper Text Markup Language (HTML) 5.0, and database search operations of PeTMbase were coded in PHP v7.1 (<http://www.php.net/>). The communication between the user interface of PeTMbase and MySQL database was implemented with PHP scripts. To offer a simple and an easy-to-use searching experience for the users, it was utilized JQuery (<https://jquery.com/>) AJAX (asynchronous HTTP) methods.

eTM ID	Long non-coding RNA ID	microRNA ID
zma_eTM_miR827-5p-1	GRNC_zma_lcl Zmays_GRMZM2G449065_T01	zma-miR827-5p
<p>Long non-coding RNA sequence - 5' -> 3' GCAACTGATATATCATCATGAAGATGAGCAATGGCCAATGGACACACATGTTTCTATTTTCCAGTATATTATGATCAGAGCCCGCATATGGTACAACCTAGCAACATGAGTGG GATTAATGAACACCACCAACAAC AATAACGGCATCAAACTAAGCAACACGTTACGAATTTTCAAGCCTTTGGAAGACGAACACCAGTGTCTCGGCGACCTCAGTATGAAA TGAATGGATTTGTAACACAATTGTGATACATTATGATTATCTGGGGTAGTTTCCCATGTGGCAGCAACCCTTCGGTAAATCTGGGGTAGTTTCTCCAAGGTGAG CGCCTTATGTCCATCAATAACAATGAGATAAAAAAACATGCTTGGCCTGTTGTTTCCAACCTATTTGTACAGGCTGAGTCCAGGCCCCCAATATACAGAACAGTATACCT CTTGATGCTCTCTATCTATGACCTTCTGATGACATACAGAGGTGGCATGTATCTCAAGGAGTGTCTCCGGAATGTAAGCAAGGC</p> <p>Reverse complementary microRNA sequence - 3' -> 5' GGTTAAATGACCACCAACAAA</p> <p>microRNA Sequence UUUGUUGGUGGUCAUUUAACC</p>		
zma_eTM_miR827-3p-2	zma_TCONS_00047750	zma-miR827-3p
<p>Long non-coding RNA sequence - 5' -> 3' TATATCTAGATTCATATATGTAATCATGGTTTATTAAGAGCTTTTTATAAAGCTACGCGAGAGTGGTGGGAGGATTTAAATGGTAAATATACTTAAACATGACTCCAGC AACTAGGAACATCACTGAAGGCGCTAAATTAACGGAAATTAAGCAGATCCTTTTCTACAATCCATTGATCAGCATTGCTCAATTATAGGTACCTGTTGCTCGCTACAGTTTACT CAAAAACGAATAAAACGGTGTATCAGGCGCTGGCTCTTATCCCCTGACACCTTTTCCCTTCCATCTGTTTGTCTGGCACTCCAAACAGGGGGCTTAACTCGCGTCCAA AATGATCCACATCAGGTCTCAATTGGCAAAGTTCAAAGCCACATGCCAGTGGCCAGGCCATGGGCGTTTCTCAATCTTCCATCCAACCTATTTTATAGACACCGCTCA ATAAAGCATCCAGGTGCATCAGTTAAATATAAAAAACAGCTCACTCCTTTGGCTTTTGTCTTGCAGATCTGCTGATCACTGAATCCTCAAGCACTGCTCAATAGAAAGTG GACAACACTTTAGATGCAATTGGGCATCAATAAAAGTGATAAAAAATAGGAAATAGCAAGTTAGGTAATCTGAGGTTGACCCCTTCCATCAATTAAGTACTGATCAG CAGCATAGGATAGCCTTTGAAATCACAAACATGAAGGGACAGCAGGCAGCAACACATATGGAAATGATGAGTCAATCCACTTAACTAAATGGTCAATCCACTTTAACT AAATGGGACACTTGTGAGATGCAACCGTTTCTCAAGTCAATGTTTCCAAAGGAGTACCCGCAAGTTCCAAAACATGATAAGTTCTACGTTGTTGTTGTTCTTCACT TATT TGCCATGGATGTCATCTAT CGCTTCCATTGTCACTTATTTACTTCCATCAATCATTTTATAGACACCGCTCAATAAAGCATCAAGGTGCATCAGTTAAATATAGA AACAGCTCACTCCTTTAAAGCTGACATTGAATATGTTGGTATGATGTCGCAAAAGAAGCCGCATGACAGGCACACAGGACATGCTCAGAGCACACAGCTATGTTCTCAA ATAAGCATCACCCAAAGGAAAGCCACAACCTTGGCAGGCACAGCGAAACAAGTTGGCATTGTCAAAACACGACAATCTTCAAGCAACGGATTGCATATTAAGCTCGACGG GGCAAGGGATCTGCGACCCACCAACAGACTGAGGTGGCCGAAATGATGGCCGGGCGGAGAAGTAAATGGAGTACCGCATCTGCGGCAAGATCAACACCCAGGCTAGACTTC AGGGACAGCGGCAAGGGACCCAGCGCCCTGAGCATGCCGTGGCAGGCGAGCTGTTGGGCGCCCTCAAAACAGCGGTGCCATTGCATGCGACAGAGACACCGTGAAGGCCG CTCGCTTCTGTACGATAAAGAGGTGGAAGCGCCAATGCGCGTGGTGTGGTAGGCGAGGTCGCGCAAGGTGGACCGGTCAGGGCTAGGGTTGAGGGCAAGCAGACAG AAGAAAT</p> <p>Reverse complementary microRNA sequence - 3' -> 5' TGTTTGTGATGGTCTATCTAA</p> <p>microRNA Sequence UUAGAUGACCAUCAGCAAACA</p>		
zma_eTM_miR827-3p-1	zma_TCONS_00047749	zma-miR827-3p
zma_eTM_miR529-5p-2	GRNC_zma_lcl Zmays_GRMZM2G474370_T01	zma-miR529-5p
zma_eTM_miR529-5p-1	zma_TCONS_00079923	zma-miR529-5p
zma_eTM_miR528b-5p-9	PNRD_zma_TCONS_00002233	zma-miR528b-5p

Fig 3. An eTM:IncRNA sequence representation in PeTMbase. The full cDNA sequence of the predicted eTMs, and their target miRNA sequences can be viewed following the database search. miRNA binding site within the eTM sequences is highlighted and the sequence properties of lncRNAs can be retrieved by clicking over the lncRNA ID.

doi:10.1371/journal.pone.0167698.g003

csv) format, which allows incorporating provided information with other bioinformatics tools such as blast.

Conclusions and Future Works

Here, we constructed a living database of plant eTMs, which is the most comprehensive eTM repository mining previously sequenced plant transcriptomes. PeTMbase has the following features: i) to search eTM sequences by miRNA name, ii) to search eTM sequences in selected plant species, iii) to retrieve a comprehensive information such as gene ID, exon number, genomic coordinates of the lncRNA and detailed information of the corresponding miRNA sequence, iv) to download eTM data per plant species.

This open-access database enables its users to reach eTM sequences and informs plant researchers on miRNA:eTM modules. Therefore, our up-to-date database will help to conduct functional genomics studies on miRNA and target genes. The modular and extensible architecture of PeTMbase is open to grow and to integrate more plant species as we include more plant species and dissect new RNA-sequencing data. Thus the number of eTMs belonged to a

diverse set of plant species is increased. In addition to that, an analysis and a submission module will be incorporated, so that the users will be able to perform eTM prediction by sequence and register their research findings to our public database. PeTMbase will be useful in studies particularly on miRNA regulatory networks and to provide insights into the regulatory roles of eTMs.

Supporting Information

S1 Table. SRA data sets utilized for the identification of novel lncRNA sequences.
(PDF)

S2 Table. Reference genomes and alignment statistics. Reference genome assemblies utilized for short read mapping and alignment statistics associated with the transcriptome libraries.
(PDF)

Acknowledgments

Authors kindly acknowledge the database construction support by Omer Faruk Gerdan (PhD candidate). Database URL: <http://petmbase.org/>.

Author Contributions

Conceptualization: TU.

Data curation: GK.

Formal analysis: GK KYK.

Funding acquisition: TU.

Investigation: TU GK KYK.

Methodology: GK KYK TU.

Project administration: TU GK.

Software: GK KYK.

Supervision: TU GK.

Visualization: GK KYK.

Writing – review & editing: TU GK KYK.

References

1. Jones-Rhoades MW, Bartel DP, Bartel B. MicroRNAs AND THEIR REGULATORY ROLES IN PLANTS. *Annual Review of Plant Biology*. 2006; 57(1):19–53.
2. Kurtoglu KY, Kantar M, Budak H. New wheat microRNA using whole-genome sequence. *Funct Integr Genomics*. 2014; 14(2):363–79. doi: [10.1007/s10142-013-0357-9](https://doi.org/10.1007/s10142-013-0357-9) PMID: [24395439](https://pubmed.ncbi.nlm.nih.gov/24395439/)
3. Unver T, Namuth-Covert DM, Budak H. Review of current methodological approaches for characterizing microRNAs in plants. *Int J Plant Genomics*. 2009; 2009:262463. PubMed Central PMCID: PMC2760397. doi: [10.1155/2009/262463](https://doi.org/10.1155/2009/262463) PMID: [19834623](https://pubmed.ncbi.nlm.nih.gov/19834623/)
4. Rodriguez RE, Mecchia MA, Debernardi JM, Schommer C, Weigel D, Palatnik JF. Control of cell proliferation in *Arabidopsis thaliana* by microRNA miR396. *Development*. 2010; 137(1):103–12. doi: [10.1242/dev.043067](https://doi.org/10.1242/dev.043067) PMID: [20023165](https://pubmed.ncbi.nlm.nih.gov/20023165/)
5. Sunkar R, Chinnusamy V, Zhu J, Zhu J-K. Small RNAs as big players in plant abiotic stress responses and nutrient deprivation. *Trends in plant science*. 2007; 12(7):301–9. doi: [10.1016/j.tplants.2007.05.001](https://doi.org/10.1016/j.tplants.2007.05.001) PMID: [17573231](https://pubmed.ncbi.nlm.nih.gov/17573231/)

6. Wang JW, Czech B, Weigel D. miR156-regulated SPL transcription factors define an endogenous flowering pathway in *Arabidopsis thaliana*. *Cell*. 2009; 138(4):738–49. doi: [10.1016/j.cell.2009.06.014](https://doi.org/10.1016/j.cell.2009.06.014) PMID: [19703399](https://pubmed.ncbi.nlm.nih.gov/19703399/)
7. Gupta PK. MicroRNAs and target mimics for crop improvement. *Curr Sci India*. 2015; 108:1624–33.
8. Zhang H, Chen X, Wang C, Xu Z, Wang Y, Liu X, et al. Long non-coding genes implicated in response to stripe rust pathogen stress in wheat (*Triticum aestivum* L.). *Molecular biology reports*. 2013; 40(11):6245–53. doi: [10.1007/s11033-013-2736-7](https://doi.org/10.1007/s11033-013-2736-7) PMID: [24065539](https://pubmed.ncbi.nlm.nih.gov/24065539/)
9. Liu X, Hao L, Li D, Zhu L, Hu S. Long non-coding RNAs and their biological roles in plants. *Genomics, proteomics & bioinformatics*. 2015; 13(3):137–47.
10. Quek XC, Thomson DW, Maag JL, Bartonicek N, Signal B, Clark MB, et al. IncRNAdb v2. 0: expanding the reference database for functional long noncoding RNAs. *Nucleic acids research*. 2014:gku988.
11. Wapinski O, Chang HY. Long noncoding RNAs and human disease. *Trends in cell biology*. 2011; 21(6):354–61. doi: [10.1016/j.tcb.2011.04.001](https://doi.org/10.1016/j.tcb.2011.04.001) PMID: [21550244](https://pubmed.ncbi.nlm.nih.gov/21550244/)
12. Wang KC, Yang YW, Liu B, Sanyal A, Corces-Zimmerman R, Chen Y, et al. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature*. 2011; 472(7341):120–4. doi: [10.1038/nature09819](https://doi.org/10.1038/nature09819) PMID: [21423168](https://pubmed.ncbi.nlm.nih.gov/21423168/)
13. Thomson DW, Dinger ME. Endogenous microRNA sponges: evidence and controversy. *Nature Reviews Genetics*. 2016; 17(5):272–83. doi: [10.1038/nrg.2016.20](https://doi.org/10.1038/nrg.2016.20) PMID: [27040487](https://pubmed.ncbi.nlm.nih.gov/27040487/)
14. Franco-Zorrilla JM, Valli A, Todesco M, Mateos I, Puga MI, Rubio-Somoza I, et al. Target mimicry provides a new mechanism for regulation of microRNA activity. *Nat Genet*. 2007; 39(8):1033–7. doi: [10.1038/ng2079](https://doi.org/10.1038/ng2079) PMID: [17643101](https://pubmed.ncbi.nlm.nih.gov/17643101/)
15. Ye C, Xu H, Shen E, Liu Y, Wang Y, Shen Y, et al. Genome-wide identification of non-coding RNAs interacted with microRNAs in soybean. *Frontiers in Plant Science*. 2014; 5.
16. Banks IR, Zhang Y, Wiggins BE, Heck GR, Ivashuta S. RNA decoys: an emerging component of plant regulatory networks? *Plant signaling & behavior*. 2012; 7(9):1188–93.
17. Wu HJ, Wang ZM, Wang M, Wang XJ. Widespread long noncoding RNAs as endogenous target mimics for microRNAs in plants. *Plant physiology*. 2013; 161(4):1875–84. Epub 2013/02/23. PubMed Central PMCID: [PMC3613462](https://pubmed.ncbi.nlm.nih.gov/PMC3613462/). doi: [10.1104/pp.113.215962](https://doi.org/10.1104/pp.113.215962) PMID: [23429259](https://pubmed.ncbi.nlm.nih.gov/23429259/)
18. Todesco M, Rubio-Somoza I, Paz-Ares J, Weigel D. A collection of target mimics for comprehensive analysis of microRNA function in *Arabidopsis thaliana*. *PLoS Genet*. 2010; 6(7):e1001031. PubMed Central PMCID: [PMC2908682](https://pubmed.ncbi.nlm.nih.gov/PMC2908682/). doi: [10.1371/journal.pgen.1001031](https://doi.org/10.1371/journal.pgen.1001031) PMID: [20661442](https://pubmed.ncbi.nlm.nih.gov/20661442/)
19. Yan J, Gu Y, Jia X, Kang W, Pan S, Tang X, et al. Effective small RNA destruction by the expression of a short tandem target mimic in *Arabidopsis*. *Plant Cell*. 2012; 24(2):415–27. PubMed Central PMCID: [PMC3315224](https://pubmed.ncbi.nlm.nih.gov/PMC3315224/). doi: [10.1105/tpc.111.094144](https://doi.org/10.1105/tpc.111.094144) PMID: [22345490](https://pubmed.ncbi.nlm.nih.gov/22345490/)
20. Ivashuta S, Banks IR, Wiggins BE, Zhang Y, Ziegler TE, Roberts JK, et al. Regulation of gene expression in plants through miRNA inactivation. *PloS one*. 2011; 6(6):e21330. doi: [10.1371/journal.pone.0021330](https://doi.org/10.1371/journal.pone.0021330) PMID: [21731706](https://pubmed.ncbi.nlm.nih.gov/21731706/)
21. Poole RL. The TAIR database. *Methods Mol Biol*. 2007; 406:179–212. PMID: [18287693](https://pubmed.ncbi.nlm.nih.gov/18287693/)
22. Chen D, Yuan C, Zhang J, Zhang Z, Bai L, Meng Y, et al. PlantNATsDB: a comprehensive database of plant natural antisense transcripts. *Nucleic Acids Res*. 2012; 40(Database issue):D1187–93. PubMed Central PMCID: [PMC3245084](https://pubmed.ncbi.nlm.nih.gov/PMC3245084/). doi: [10.1093/nar/gkr823](https://doi.org/10.1093/nar/gkr823) PMID: [22058132](https://pubmed.ncbi.nlm.nih.gov/22058132/)
23. Zhao Y, Li H, Fang S, Kang Y, Wu W, Hao Y, et al. NONCODE 2016: an informative and valuable data source of long non-coding RNAs. *Nucleic Acids Res*. 2016; 44(D1):D203–8. PubMed Central PMCID: [PMC4702886](https://pubmed.ncbi.nlm.nih.gov/PMC4702886/). doi: [10.1093/nar/gkv1252](https://doi.org/10.1093/nar/gkv1252) PMID: [26586799](https://pubmed.ncbi.nlm.nih.gov/26586799/)
24. Jin J, Liu J, Wang H, Wong L, Chua NH. PLncDB: plant long non-coding RNA database. *Bioinformatics*. 2013; 29(8):1068–71. PubMed Central PMCID: [PMC3624813](https://pubmed.ncbi.nlm.nih.gov/PMC3624813/). doi: [10.1093/bioinformatics/btt107](https://doi.org/10.1093/bioinformatics/btt107) PMID: [23476021](https://pubmed.ncbi.nlm.nih.gov/23476021/)
25. Paytavi Gallart A, Hermoso Pulido A, Anzar Martinez de Lagran I, Sanseverino W, Aiese Cigliano R. GREENC: a Wiki-based database of plant lncRNAs. *Nucleic acids research*. 2016; 44(D1):D1161–6. Epub 2015/11/19. PubMed Central PMCID: [PMC4702861](https://pubmed.ncbi.nlm.nih.gov/PMC4702861/). doi: [10.1093/nar/gkv1215](https://doi.org/10.1093/nar/gkv1215) PMID: [26578586](https://pubmed.ncbi.nlm.nih.gov/26578586/)
26. Szczesniak MW, Rosikiewicz W, Makalowska I. CANTATAdb: A Collection of Plant Long Non-Coding RNAs. *Plant & cell physiology*. 2016; 57(1):e8. PubMed Central PMCID: [PMC4722178](https://pubmed.ncbi.nlm.nih.gov/PMC4722178/).
27. Yi X, Zhang Z, Ling Y, Xu W, Su Z. PNRD: a plant non-coding RNA database. *Nucleic Acids Res*. 2015; 43(Database issue):D982–9. PubMed Central PMCID: [PMC4383960](https://pubmed.ncbi.nlm.nih.gov/PMC4383960/). doi: [10.1093/nar/gku1162](https://doi.org/10.1093/nar/gku1162) PMID: [25398903](https://pubmed.ncbi.nlm.nih.gov/25398903/)
28. Wang P, Zhi H, Zhang Y, Liu Y, Zhang J, Gao Y, et al. miRSponge: a manually curated database for experimentally supported miRNA sponges and ceRNAs. *Database: The Journal of Biological Databases and Curation*. 2015; 2015:bav098.

29. Leinonen R, Sugawara H, Shumway M. The Sequence Read Archive. *Nucleic Acids Research*. 2011; 39(suppl 1):D19–D21.
30. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013; 29(1):15–21. doi: [10.1093/bioinformatics/bts635](https://doi.org/10.1093/bioinformatics/bts635) PMID: [23104886](https://pubmed.ncbi.nlm.nih.gov/23104886/)
31. Kersey PJ, Allen JE, Armean I, Boddu S, Bolt BJ, Carvalho-Silva D, et al. Ensembl Genomes 2016: more genomes, more complexity. *Nucleic Acids Research*. 2016; 44(D1):D574–D80. doi: [10.1093/nar/gkv1209](https://doi.org/10.1093/nar/gkv1209) PMID: [26578574](https://pubmed.ncbi.nlm.nih.gov/26578574/)
32. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotech*. 2010; 28(5):511–5.
33. Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Research*. 2006; 34(suppl 1):D140–D4.
34. Team RC. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2015. URL <http://www.R-project.org>. 2016.