



Database tool

PncStress: a manually curated database of experimentally validated stress-responsive non-coding RNAs in plants

Wenyi Wu¹, Yan Wu¹, Dahui Hu¹, Yincong Zhou¹, Yanshi Hu¹,
Yujie Chen^{1,2} and Ming Chen^{1,2,*}

¹Department of Bioinformatics, State Key Laboratory of Plant Physiology and Biochemistry, College of Life Sciences, Zhejiang University, Hangzhou 310058, China ²College of Life Sciences and Food Engineering, Inner Mongolia University for the Nationalities, Tongliao 028043, China

*Corresponding author: Tel/Fax: +86(0)571-88206612; E-mail: mchen@zju.edu.cn

Citation details: Wu,W., Wu,Y., Hu,D. *et al.* PncStress: a manually curated database of experimentally validated stress-responsive non-coding RNAs in plants. *Database* (2020) Vol. 2020: article ID baaa001; doi:10.1093/database/baaa001

Received 8 October 2019; Revised 10 December 2019; Accepted 2 January 2020

Abstract

Non-coding RNAs (ncRNAs) are recognized as key regulatory molecules in many biological processes. Accumulating evidence indicates that ncRNA-related mechanisms play important roles in plant stress responses. Although abundant plant stress-responsive ncRNAs have been identified, these experimentally validated results have not been gathered into a single public domain archive. Therefore, we established PncStress by curating experimentally validated stress-responsive ncRNAs in plants, including microRNAs, long non-coding RNAs and circular RNAs. The current version of PncStress contains 4227 entries from 114 plants covering 48 biotic and 91 abiotic stresses. For each entry, PncStress has biological information and network visualization. Serving as a manually curated database, PncStress will become a valuable resource in support of plant stress response research.

Database URL: <http://bis.zju.edu.cn/pncstress>

Introduction

Plants growing in nature are constantly exposed to ever-changing environmental conditions including a wide array of abiotic and biotic stress stimuli. Such unpredictable stress conditions would make a severe impact on the growth and development of plants and even affect the yield of crops. For example, abiotic stress like temperature extremes, heat or cold, would result in the accelerated

generation of reactive oxygen species, leading to oxidative stress damage, irreparable metabolic dysfunction and death (1,2). And various pathogens would challenge the host's immune system, which is a major biotic threat other than insect pests (3). To survive and accommodate to external conditions, plants have evolved an elaborate defense system through specificity or cross-talk mechanisms based on transcriptional, post-transcriptional, translational and post-translational processes (4–6).

With the popularity of the next-generation sequencing technologies and bioinformatics approaches, increasing research have shown that plant non-coding RNAs (ncRNAs) are widely involved in stress-responsive regulatory processes (7–9). ncRNAs constitute a class of RNAs that can function but not translated into proteins, including microRNAs (miRNAs), long non-coding RNAs (lncRNAs) and circular RNAs (circRNAs) (10–14). These types have been confirmed to play important roles in plant stress responses (1,15–19). miRNAs have emerged as vital regulatory molecules in plants and animals (20,21). Plant miRNAs and their regulatory targets are involved in plant responses to the environment. There is evidence that the expression of miR395 increases while its target ATP sulfurylase 1 (APS1) declines upon sulfate starvation in *Arabidopsis thaliana* (22). Studies illustrate miR169-mediated regulation is critical for adaptation to drought stress in *A. thaliana* (23), *Oryza sativa* (24) and *Medicago truncatula* (25). Other experiments have confirmed that ta-siRNAs show down-regulation in drought and high-salinity stress in *A. thaliana*, which are derived from TAS ncRNAs and targeted by miRNA (26). Recent studies also show that lncRNAs perform biological roles in various development processes. One of the typical examples is the regulation exerted by cold-induced long antisense intragenic RNA and cold-assisted intronic noncoding RNA transcripts, which are initiated

from the FLOWERING LOCUS C antisense transcripts that are repressed by prolonged cold (27,28). Hidden treasure 1 (HID1) has been proved to act as a positive lncRNA regulator of photomorphogenesis in continuous red light (29). Under dehydration, phosphate deficiency and heat stress, circRNAs have been reported to act as sponges of miRNAs (30).

Due to the expanding research efforts on plant ncRNAs by high-throughput and low-throughput techniques, a large number of ncRNA candidates have been deposited. Examples include PNRD (31), NONCODE (32) and lncRNAdb (33). However, the existing databases have not systematically curated stress-responsive ncRNA in plants that have been experimentally validated by low-throughput experiments. Therefore, we manually curated 2523 miRNAs, 444 lncRNAs and 52 circRNAs in PncStress, which have been validated to respond to 139 stresses in 114 plants. We provide each entry with specific information gathered from the literature, such as stress, species, expression pattern, experimental method, functional description, genomic position and sequence. PncStress is the most comprehensive database after comparing with PASmiR (34), lncRNAdb (33), EVLncRNAs (35) and circFunBase (36). For those published databases and tools, we provide an integrated interface based on different research fields.

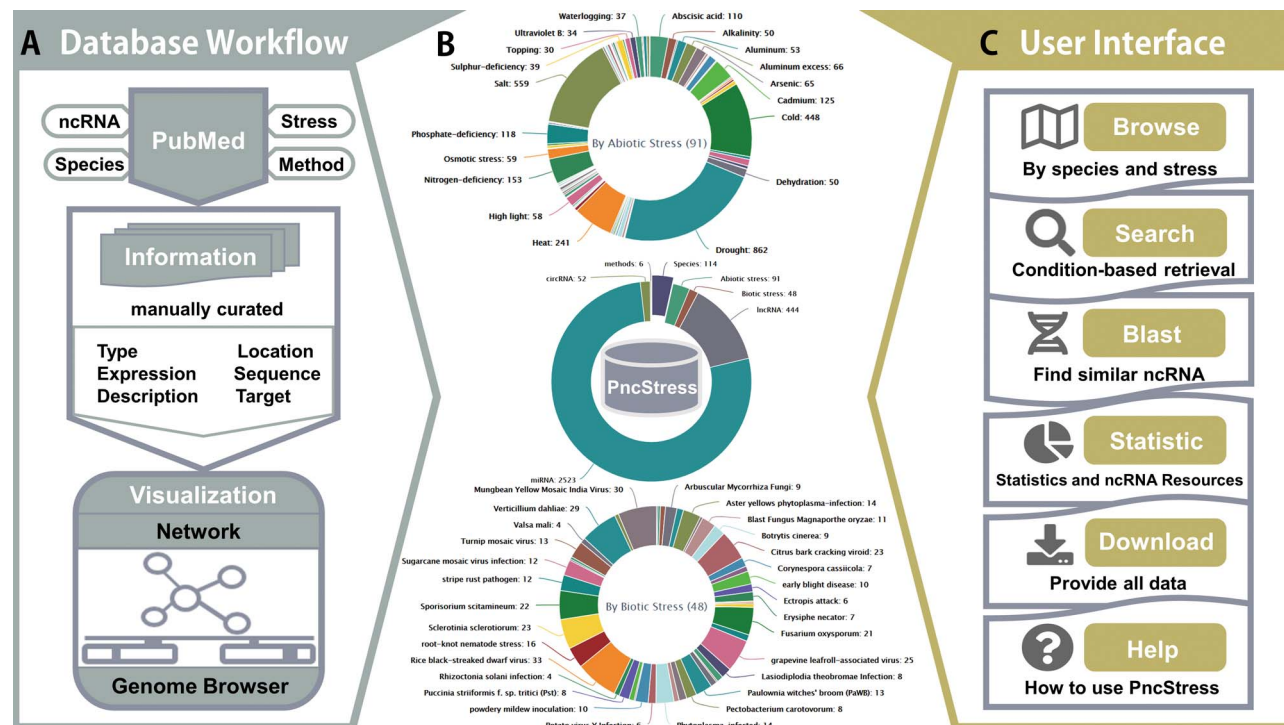


Figure 1. The architecture and statistics of PncStress.

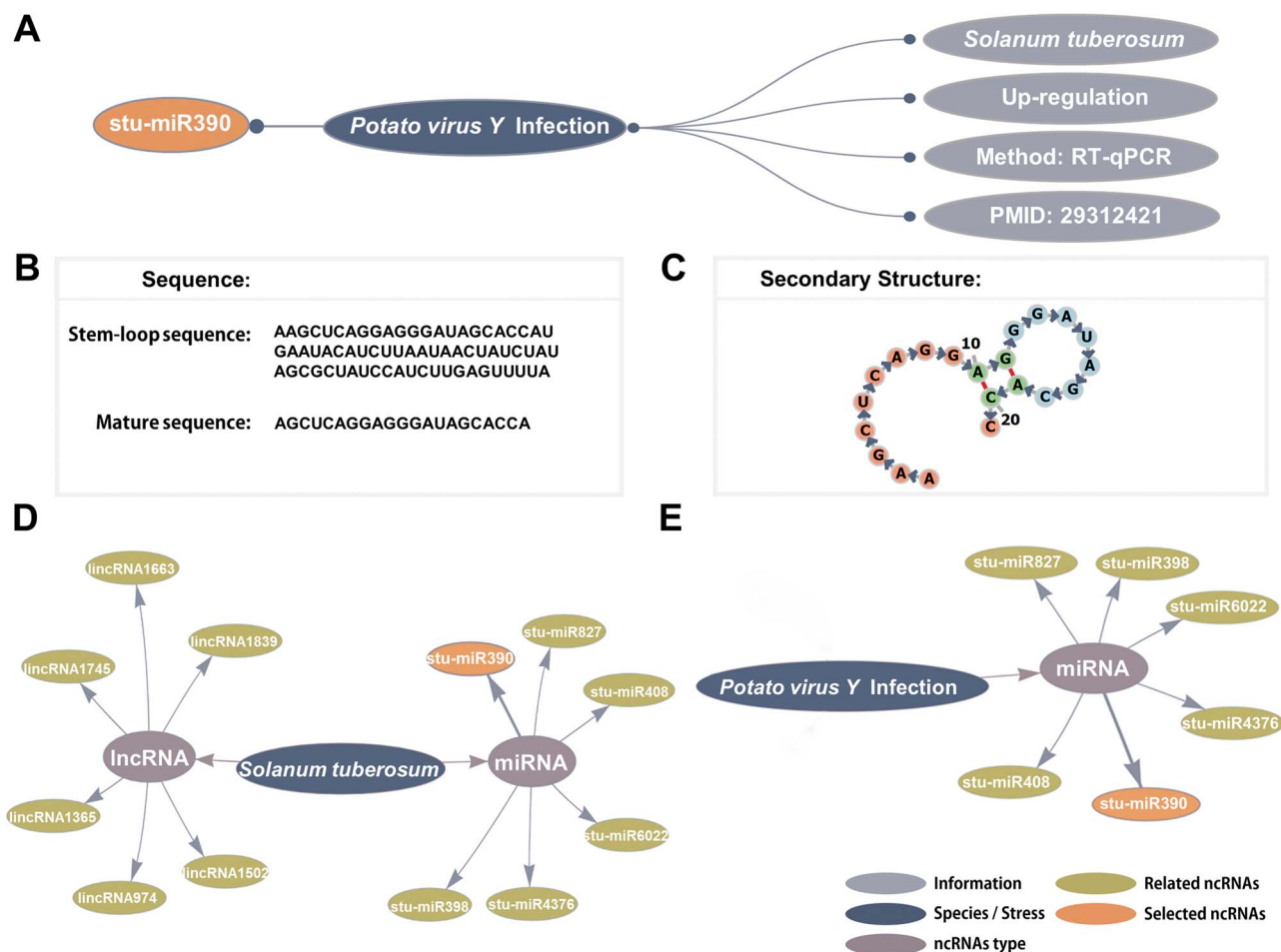


Figure 2. Detailed information and network visualization. (A) The basic information of stu-miR408. (B) The stem-loop sequence and mature sequence of stu-miR390. (C) The secondary structure of stu-miR390. (D) The *S. tuberosum*-ncRNAs network. (E) The *Potato virus Y* infection-ncRNAs network.

Materials and Methods

Data compilation

The collection of all PncStress entries was performed as follows. We firstly filtered out articles related to miRNA, lncRNA and circRNA from PubMed that were connected with stress, response, defense and adaptation. Considering the ncRNA alternate name, we used a list of keywords containing all the synonyms during the query. Take the term ‘lncRNA’ as an example, ‘lncRNA’, ‘long noncoding RNA’, ‘long non-coding RNA’ or ‘long non-protein coding RNA’ were extensively searched. Then both scientific and generic names of plants were added as keywords to filter the results, like ‘*Vitis vinifera*’ and ‘Grape’, ‘*Zea may*’ and ‘maize’. From these literatures, we manually curated ncRNAs and its biological information confirmed by low-throughput methods [e.g. reverse transcription-polymerase chain reaction (RT-PCR), quantitative RT-PCR, stem-loop RT-PCR or northern blot]. For more functional information, we predicted the secondary structure by RNAfold (37),

interaction candidates by psRNATarget (38) and miRNA-triggered phasiRNAs by PhaseTank (39).

Database construction

The PncStress database was implemented using HTML and PHP. MySQL was used for data storage and efficient management. Bootstrap and JavaScript libraries jQuery were used to make the interface application and network visualization. Dalliance (40) was used to view the genome, in which the genomic context of ncRNAs was based on the annotation from Phytozome (41). The BLAST module was implemented using SequenceServer (42).

Results

Database architecture and data statistics

Figure 1 shows the database architecture of PncStress. All entries are processed through the ‘Data Workflow’

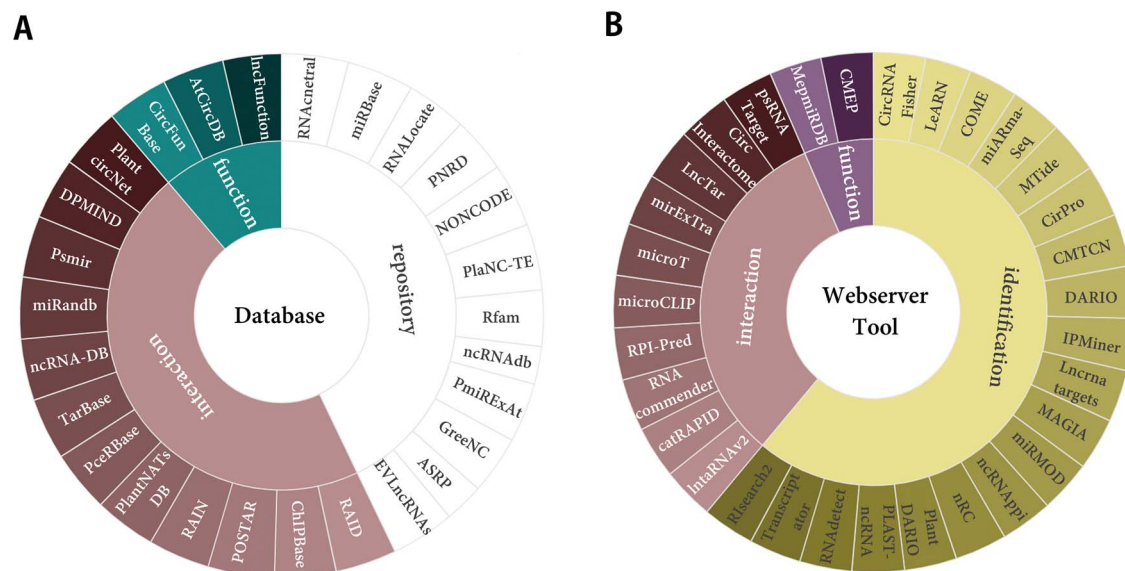


Figure 3. Plant ncRNAs research resources. (A) The 27 popular databases of ncRNAs. (B) The 31 useful webservers and tools of ncRNAs.

(Figure 1A). Each validated ncRNA contains manually curated information from PubMed, including stress, species, expression pattern, experimental method, genomic position, concrete sequence, functional description, target gene and reference literature. The visualization consists of networks and genome browser. ‘User Interface’ provides browsing, search, statistics and downloads as well as online BLAST service (Figure 1C). Each section contains documentation to ensure effective use.

The statistics of PncStress are summarized in three parts (Figure 1B). In the current version, PncStress possesses 4227 entries, including 2523 miRNAs, 444 lncRNAs and 52 circRNAs in response to 48 biotic stresses and 91 abiotic stresses, which have been validated by six experimental methods in 114 species. The statistics also show the number of each abiotic and biotic stress-responsive ncRNAs. Drought, salt, cold and heat are the most abundant stresses with 862, 559, 448 and 241 associated ncRNAs, respectively.

Data retrieval

PncStress provides a three-part retrieval system. Firstly, users can retrieve all stress-responsive ncRNAs by species or stresses in the ‘Browse’ page. Species are sorted in alphabetical order, while stresses are divided into biotic and abiotic. Users can find information of every items from each result table, such as species name, stress name, expression pattern, validated method and PMID, which are also part of network visualization in the ‘Detail’ page (Figure 2A). The ‘Detail’ page contains additional information of each ncRNA, including the location, sequence, validated target, secondary structure, genome browser, func-

tional description and interaction candidates. In addition, stem-loop sequence and mature sequence from miRBase (43), predicted alignment files predicted by psRNATarget and miRNA triggered phasiRNAs predicted by PhaseTank are provided special for miRNAs. As an example, the sequence and secondary structure of stu-miR390 are displayed in Figure 2B and C.

Secondly, we constructed species–ncRNAs and stress–ncRNAs networks that directly retrieve stress-responsive ncRNAs in each species and stresses. Six miRNAs and six lncRNAs respond to *Solanum tuberosum* (Figure 2D), while six miRNAs respond to *Potato virus Y* infection (Figure 2E). Both networks are associated with stu-miR390. The researchers found up-regulation of miR408, miR827, miR398 and miR390 in potato cultivar Désirée, which is tolerant to *Potato virus Y* infection as well as down-regulation of miR4376 and miR6022 (44). Users can click the interested node to its detailed page.

In addition, PncStress provides four advanced search options to retrieve data and enable auto-completion and fuzzy search for the text fields. Users can search by species, stress, ncRNA name and three ncRNA types, or combine two or more options.

Additional ncRNA resources

Given the more practical requirements in plant ncRNAs research field, PncStress provides a collection interface of the most popular and topic-specific ncRNA resources in the ‘Statistics’ page. The 58 web resources were grouped into categories of identification, function, interaction and repository, including RNAcentral (45), NONCODE (32), PccRBase (46) and MTide (47). They were divided into

two parts: 27 databases and 31 webservers/tools (Figure 3). Users can hover over the desired item to get its description and click to access its web page.

Discussion

Plant stress-responsive biomolecules have attracted increasing attention, and the involvement of ncRNAs has expanded to further understanding plant stress response mechanisms. Thus, obtaining experimentally validated stress-responsive ncRNAs directly from a large number of studies and their biological information would become a problem for individual researchers. Here, we have developed PncStress, a repository that manually curated the three most crucial stress-responsive ncRNAs in plants. PncStress provides detailed page with tables and visualization for users to obtain stress-responsive ncRNAs information. Take *stu-miR390* for example, it was up-regulated in response to *Potato virus Y* infection in *S. tuberosum*, which has been validated through RT-qPCR in a study that was comprised in PubMed with a unique identifier (PMID 29312421). This information can be retrieved from the table and is part of network visualization in the ‘Detail’ page. For other ncRNAs that respond to several stress conditions, we divided their stress conditions into biotic and abiotic. *Stu-miR390* associated with other miRNAs and lncRNAs in *S. tuberosum*-ncRNAs network and *Potato virus Y* infection-ncRNAs network. The intertwining of different networks indicates the crucial regulatory role of ncRNAs in plant survival and adaptation to external conditions. We also have constructed ncRNA-related networks for other plants and stresses. Additional manually curated information from PubMed are displayed for each validated ncRNA. Those predicted interaction candidates can serve as functional annotation to gain more insight into ncRNA regulation.

PncStress contains most of the plants and stress conditions. For instance, in *A. thaliana*, 305 miRNAs, 186 lncRNAs and 5 circRNAs were deposited in PncStress, which are experimentally validated to be associated with 26 abiotic stresses and 4 biotic stresses. PncStress is expected to be a fundamental study resource that includes data on the most important crops such as maize, wheat, cotton and potato.

In summary, PncStress is a manually curated knowledge-base of experimentally validated stress-responsive miRNAs, lncRNAs and circRNAs in plants. It will bridge the gap in ncRNAs functional research and further facilitate biologists in revealing the roles of ncRNAs in diverse stress conditions. We will update PncStress regularly based on newly published data. The improvement of PncStress can facilitate future plant study in stress responses.

Funding

National Key Research and Development Program of China (nos. 2018YFC0310602 and 2016YFA0501704); National Natural Science Foundation of China (nos. 31771477 and 31571366); the Fundamental Research Funds for the Central Universities; Jiangsu Collaborative Innovation Center for Modern Crop Production and Collaborative Innovation Center for Modern Crop Production co-sponsored by province and ministry.

Acknowledgement

The authors are grateful to the anonymous experts for insightful reviews and members of Ming Chen’s laboratory, especially Yongjing Liu and Peijing Zhang, for helpful discussions and valuable comments.

Conflict of interest

None declared.

References

1. Hasanuzzaman, M., Nahar, K., Alam, M.M. *et al.* (2013) Physiological, biochemical, and molecular mechanisms of heat stress tolerance in plants. *Int. J. Mol. Sci.*, **14**, 9643–9684.
2. Del Rio, L.A. (2015) ROS and RNS in plant physiology: an overview. *J. Exp. Bot.*, **66**, 2827–2837.
3. Islam, W., Noman, A., Qasim, M. *et al.* (2018) Plant responses to pathogen attack: small RNAs in focus. *Int. J. Mol. Sci.*, **19**, 515.
4. Vaucheret, H. (2006) Post-transcriptional small RNA pathways in plants: mechanisms and regulations. *Genes Dev.*, **20**, 759–771.
5. Guerra, D., Crosatti, C., Khoshro, H.H. *et al.* (2015) Post-transcriptional and post-translational regulations of drought and heat response in plants: a spider’s web of mechanisms. *Front. Plant Sci.*, **6**, 57.
6. Ng, D.W., Abeysinghe, J.K. and Kamali, M. (2018) Regulating the regulators: the control of transcription factors in plant defense signaling. *Int. J. Mol. Sci.*, **19**, 3737.
7. Xie, Z., Kasschau, K.D. and Carrington, J.C. (2003) Negative feedback regulation of dicer-Like1 in Arabidopsis by microRNA-guided mRNA degradation. *Curr. Biol.*, **13**, 784–789.
8. Borsani, O., Zhu, J., Verslues, P.E. *et al.* (2005) Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in Arabidopsis. *Cell*, **123**, 1279–1291.
9. Yan, K., Liu, P., Wu, C.A. *et al.* (2012) Stress-induced alternative splicing provides a mechanism for the regulation of microRNA processing in Arabidopsis thaliana. *Mol. Cell*, **48**, 521–531.
10. Reinhart, B.J., Slack, F.J., Basson, M. *et al.* (2000) The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature*, **403**, 901–906.
11. Hamilton, A.J. and Baulcombe, D.C. (1999) A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science*, **286**, 950–952.

12. Lagos-Quintana, M., Rauhut, R., Lendeckel, W. *et al.* (2001) Identification of novel genes coding for small expressed RNAs. *Science*, **294**, 853–858.
13. Kapusta, A. and Feschotte, C. (2014) Volatile evolution of long noncoding RNA repertoires: mechanisms and biological implications. *Trends Genet.*, **30**, 439–452.
14. Jeck, W.R., Sorrentino, J.A., Wang, K. *et al.* (2013) Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA*, **19**, 141–157.
15. Knight, H. and Knight, M.R. (2001) Abiotic stress signalling pathways: specificity and cross-talk. *Trends Plant Sci.*, **6**, 262–267.
16. Hirayama, T. and Shinozaki, K. (2010) Research on plant abiotic stress responses in the post-genome era: past, present and future. *Plant J.*, **61**, 1041–1052.
17. Zhang, Y.C. and Chen, Y.Q. (2013) Long noncoding RNAs: new regulators in plant development. *Biochem. Biophys. Res. Commun.*, **436**, 111–114.
18. Contreras-Cubas, C., Palomar, M., Arteaga-Vazquez, M. *et al.* (2012) Non-coding RNAs in the plant response to abiotic stress. *Planta*, **236**, 943–958.
19. Zhao, J., He, Q., Chen, G. *et al.* (2016) Regulation of non-coding RNAs in heat stress responses of plants. *Front. Plant Sci.*, **7**, 1213.
20. Sunkar, R., Li, Y.F. and Jagadeeswaran, G. (2012) Functions of microRNAs in plant stress responses. *Trends Plant Sci.*, **17**, 196–203.
21. Covarrubias, A.A. and Reyes, J.L. (2010) Post-transcriptional gene regulation of salinity and drought responses by plant microRNAs. *Plant Cell Environ.*, **33**, 481–489.
22. Jones-Rhoades, M.W. and Bartel, D.P. (2004) Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Mol. Cell*, **14**, 787–799.
23. Li, W.X., Oono, Y., Zhu, J. *et al.* (2008) The Arabidopsis NFYA5 transcription factor is regulated transcriptionally and posttranscriptionally to promote drought resistance. *Plant Cell*, **20**, 2238–2251.
24. Zhao, B., Liang, R., Ge, L. *et al.* (2007) Identification of drought-induced microRNAs in rice. *Biochem. Biophys. Res. Commun.*, **354**, 585–590.
25. Trindade, I., Capitao, C., Dalmay, T. *et al.* (2010) miR398 and miR408 are up-regulated in response to water deficit in *Medicago truncatula*. *Planta*, **231**, 705–716.
26. Matsui, A., Mizunashi, K., Tanaka, M. *et al.* (2014) tasiRNA-ARF pathway moderates floral architecture in Arabidopsis plants subjected to drought stress. *Biomed Res Int.*, **2014**: 303451.
27. Swiezewski, S., Liu, F., Magusin, A. *et al.* (2009) Cold-induced silencing by long antisense transcripts of an Arabidopsis Polycomb target. *Nature*, **462**, 799–802.
28. Heo, J.B. and Sung, S. (2011) Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. *Science*, **331**, 76–79.
29. Wang, Y., Fan, X., Lin, F. *et al.* (2014) Arabidopsis noncoding RNA mediates control of photomorphogenesis by red light. *Proc. Natl. Acad. Sci. U. S. A.*, **111**, 10359–10364.
30. Litholdo, C.G. Jr. and da Fonseca, G.C. (2018) Circular RNAs and plant stress responses. *Adv. Exp. Med. Biol.*, **1087**, 345–353.
31. Yi, X., Zhang, Z., Ling, Y. *et al.* (2015) PNRD: a plant non-coding RNA database. *Nucleic Acids Res.*, **43**, D982–D989.
32. Zhao, Y., Li, H., Fang, S. *et al.* (2016) NONCODE 2016: an informative and valuable data source of long non-coding RNAs. *Nucleic Acids Res.*, **44**, D203–D208.
33. Quek, X.C., Thomson, D.W., Maag, J.L. *et al.* (2015) lncRNADB v2.0: expanding the reference database for functional long noncoding RNAs. *Nucleic Acids Res.*, **43**, D168–D173.
34. Zhang, S., Yue, Y., Sheng, L. *et al.* (2013) PASmiR: a literature-curated database for miRNA molecular regulation in plant response to abiotic stress. *BMC Plant Biol.*, **13**, 33.
35. Zhou, B., Zhao, H., Yu, J. *et al.* (2018) EVLncRNAs: a manually curated database for long non-coding RNAs validated by low-throughput experiments. *Nucleic Acids Res.*, **46**, D100–D105.
36. Meng, X., Hu, D., Zhang, P. *et al.* (2019) CircFunBase: a database for functional circular RNAs. *Database*, **36**, baz003.
37. Lorenz, R., Bernhart, S.H., Honer Zu Siederdisen, C. *et al.* (2011) ViennaRNA package 2.0. *Algorithms Mol. Biol.*, **6**, 26.
38. Dai, X., Zhuang, Z. and Zhao, P.X. (2018) psRNATarget: a plant small RNA target analysis server (2017 release). *Nucleic Acids Res.*, **46**, W49–W54.
39. Guo, Q., Qu, X. and Jin, W. (2015) PhaseTank: genome-wide computational identification of phasiRNAs and their regulatory cascades. *Bioinformatics (Oxford England)*, **31**, 284–286.
40. Down, T.A., Piipari, M. and Hubbard, T.J. (2011) Dalliance: interactive genome viewing on the web. *Bioinformatics*, **27**, 889–890.
41. Goodstein, D.M., Shu, S., Howson, R. *et al.* (2012) Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.*, **40**, D1178–D1186.
42. Priyam, A., Woodcroft, B.J., Rai, V. *et al.* (2019) Sequenceserver: a modern graphical user interface for custom BLAST databases. *Mol. Biol. Evol.*, **36**, 2922–2924.
43. Kozomara, A. and Griffiths-Jones, S. (2014) miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.*, **42**, D68–D73.
44. Kriznik, M., Petek, M., Dobnik, D. *et al.* (2017) Salicylic acid perturbs sRNA-gibberellin regulatory network in immune response of potato to *Potato virus Y* infection. *Front. Plant Sci.*, **8**, 2192.
45. The RNAcentral Consortium (2019) RNAcentral: a hub of information for non-coding RNA sequences. *Nucleic Acids Res.*, **47**, D1250–D1251.
46. Yuan, C., Meng, X., Li, X. *et al.* (2017) PceRBase: a database of plant competing endogenous RNA. *Nucleic Acids Res.*, **45**, D1009–D1014.
47. Zhang, Z., Jiang, L., Wang, J. *et al.* (2015) MTide: an integrated tool for the identification of miRNA-target interaction in plants. *Bioinformatics*, **31**, 290–291.