RiceLncPedia: a comprehensive database of rice long non-coding RNAs

Zhengfeng Zhang[1#*], Yao Xu[2#], Fei Yang[3], Benze Xiao[3] and Guoliang Li[2*]

[1]School of Life Sciences, Hubei Key Laboratory of Genetic Regulation and Integrative Biology,CentralChina Normal University, Wuhan 430079, China

[2] National Key Laboratory of Crop Genetic Improvement, Hubei Key Laboratory of Agricultural Bioinformatics, Hubei Engineering Technology Research Center of Agricultural Big Data, College of Informatics, Huazhong Agricultural University, Wuhan 430070, China

[3] College of Plant Science and Technology, HuazhongAgricultural University, Wuhan 430070, China

[*] To whom correspondence should be addressed. Tel: +86 13627259418; Email: zhengfeng@mail.ccnu.edu.cn and Tel: +8627-87285078; Email:guoliang.li@mail.hzau.edu.cn

[#] These authors could be regarded as Joint First Authors.

ABSTRACT

Long non-coding RNAs (lncRNAs) play significant functions in various biological processes including differentiation, development and adaptation to different environments. Although multi research focused on lncRNAs in rice, the systematic identification and annotation of lncRNAs expressed in different tissues, developmental stages under diverse conditions are still scarce. This impacts the elucidation of their functional significance and the further research on them. Here, RiceLncPedia (http://218.199.68.191:10092/) is constructed including rice lncRNAs explored from 2313 publically available rice RNA-seq libraries and characterize them with multi-omics data sets. In the current version, RiceLncPedia shows 6978 lncRNAs with abundant features: (i) expression profile across 2313 rice RNA-seq libraries; (ii) an online genome browser for rice lncRNAs; (iii) genome SNPs in lncRNA transcripts; (iv) lncRNA associations with phenotype; (v) overlap of lncRNAs with transposons; and (vi) LncRNA-miRNA interactions and lncRNAs as the precursors of miRNAs. In total, RiceLncPedia imported numerous of rice lncRNAs during development under various environments as well as their features extracted from multi-omics data and thus serve as a fruitful resource for rice-related research communities. RiceLncPedia will be further updated with experimental validation, functions association and epigenetic characteristics to greatly facilitate future investigation on rice lncRNAs.

Key words: Rice; LncRNAs; Expression; Multi-omics

INTRODUCTION

Long noncoding RNAs (lncRNAs) are referred as RNA molecules with length of at least 200 nucleotides (nt) and usually have low protein coding potential (1). In human, function of lncRNAs is relevant to various important biological processes such as cell differentiation, immune response, diverse cancers and so on (2-7). In plants, emerging evidences indicate that lncRNAs function as key modulators in a wide range of biological processes including development and stress response at the epigenetic, transcriptional and post-transcriptional levels (8-16).

In recent years, the explosive high-throughput sequencing promotes the global discovery of lncRNAs in various processes in animal system as well as plants. Accordingly, multiple databases have been constructed, focusing on different aspect of lncRNAs in human and various animal species. NONCODE 5.0 provides the gene expression pattern from RNA-seq data and includes lncRNAs in 17 species such as human, mouse and so on (17). LncRNAdb v2.0 (18) and LNCipedia (19) are annotation databases based on literature evidence. LncRNAs with the support of coexpression, differential expression, binding proteins and phylogenetic conservation was constructed in database of lncRNAtor (20). lncRNASNP integrated the lncRNA, SNP, GWAS results and miRNA expression profiles in human and mouse (21). Recently, a comprehensive database of human long non-coding RNAs, LncBook was constructed which incorporated multi-omics data including expression profile, sequence variation, association with miRNA, epigenetic features and diseases association to annotate human lncRNAs (22). Several lncRNAs databases were also developed in plants but mainly for model species *Arabidopsis* (23,24). Compared with human and animal lncRNA databases, the number of plant lncRNA databases and lncRNA features referred in the databases were relatively rare. Those plant databases continuously accessible and updated including rice lncRNAs were even fewer (Table 1 Key differences between the existing databases including rice lncRNAs and RiceLncPedia). A long non-coding RNA database of plants (PLncDB) provides rich features for lncRNAs including genomic information, expression profiles in multi developmental stages, mutants and stress conditions, epigenetic modification and small RNA associations (25). However, this database only focused on Arabidopsis without other plant species in the present version. GREENC, a Wiki-based database of plant lncRNAs, identified 120 000 lncRNAs from 37 plant species and six algae, including 5237 lncRNAs from rice (26). The GreeNC database provides information

about sequence, genomic coordinates, coding potential and folding energy for all the identified lncRNAs. However, the lncRNAs in this database were identified from the transcripts in Phytozome v10.3 and therefore the expression profile across tissues, growth conditions were unavailable. Additionally, variations and other genomic features are lack either (26,27). CANTATAdb collected lncRNAs in 36 plant species and 3 algae, among of which, 2788 lncRNAs are collected from only 8 rice RNA-Seq libraries (28). PNRD, a plant non-coding RNA database, collected a total of 25739 entries of 11 different types of ncRNAs from 150 plant species, whereas only harbors 790 lncRNAs in rice (29). PLNlncRbase, A resource for experimentally identified lncRNAs in plants, has manually collected data from nearly 200 published literature, covering 1187 plant lncRNAs in 43 plant species, 1060 of which are stress-related lncRNAs under 17 different abiotic or biotic stress conditions in various plant species (30). Another database including low-throughput experiment validated lncRNAs is EVLncRNAs, which contains 1543 lncRNAs from 77 species, whereas only 428 plant lncRNAs from 44 plant species (31). PLncPRO has discovered lncRNAs responsive to abiotic stresses in rice and chickpea, where nine rice RNA-seq samples were explored, associating with desiccation and salinity stresses from only three rice cultivars (32). However, these databases are small in scale and less comprehensive for rice, a most widely planted staple food crop and model crop. A comprehensive rice database covering lncRNAs from more widely developmental tissues, stages and diverse stress condition as well as integrating multi-omic features is still lacking (Table 1).

In this study, we developed a rice lncRNAs database, RiceLncPedia with the following comprehensive features, attempting to facilitate the understanding and usage of rice lncRNAs: (i) a collection of lncRNAs identified from the majority of the publically available rice RNA-seq dataset (2313 RNA-seq libraries); (ii) lncRNA expression profiles in various tissues, developmental stages, stress treatments and different cultivars; (iii) lncRNA associations with genome variations; (iv) the linkage of lncRNAs with phenotype (inferred from public available QTLs and GWAS results); (v) the overlap information of lncRNAs and transposon elements, since the transposon elements (TEs) play key roles in the generation and function of lncRNAs; (vi) the lncRNAs predicted as miRNA targets or miRNA precursors collected as well.

MATERIALS AND METHODS

Data collection

We identified rice lncRNAs from RNA assemblies based on 2313 public available RNA-seq libraries. To obtain high-confidence lncRNAs, the following steps were adopted from the raw RNA-seq data(Figure 1). Low-quality reads with >5% ambiguous bases were filtered and adapter sequences were trimmed . The clean RNA-seq reads were mapped to the rice reference genome Os-Nipponbare-Reference-IRGSP-1.0 using Hisat 2.1.0 program with the default parameters set (33). Transcriptomes were reconstructed using StringTie v1.3.3 package with the default parameters (33). Transcriptome assemblies generated from the above steps were subsequently merged with StringTie --merge to acquire comprehensive non-redundant transcripts for subsequent analysis: (i) the transcripts assembled using StringTie were then annotated with Cuffcompare program to filter out known protein-coding transcripts, rRNA and tRNA with the comparison code '=' ; (ii) questionable transcripts tagged with codes 'e', 'p' and 's' by Cuffcompare were filtered out; (iii) transcripts with lengths more than 200nt were selected as lncRNA candidates;(iv) transcripts with FPKM scores smaller than 0.5 in all samples were discarded; (v) further screened through the protein-coding score test using Coding Potential Calculator (CPC2) (34), Plant Long Non-Coding RNA Prediction by Random fOrests (PlncPRO) (32) and Protein family database (Pfam) (35). As a consequence, a total of 6978 non-redundant rice lncRNA transcripts were obtained, belonging to 5845 gene loci.

Data integration and annotation

1) Molecular features of rice lncRNAs

Based on the location of lncRNAs relative to proteincoding RNAs, the different types of lncRNAs include intergenic lncRNA (lincRNA), intronic lncRNA, sense lncRNA, antisense lncRNA, which were tagged using Cuffcompare(36)as class code of u, i, o and x, respectively. As code 'j' represents potentially novel isoform: at least one splice junction is shared with a reference transcript (Cuffcompare manuscript), this class of transcripts can be long non-coding isoforms of known genes (37). Thereby, the transcripts with code of u, i, j, o or x were retained for further analysis. The number, length and GC content (%) of lncRNAs were counted by in-house Python scripts.

2) LncRNAs expression

StringTie 1.3.3 was used to calculate FPKMs of lncRNAs. Housekeeping (HK) lncRNAs, tissue-specific (TS) and stress responsive (SR) lncRNAs were determined as following (22) with minor modification. Briefly, based on the expression value of lncRNAs in the selected datasets (see the section of DATABASE CONTENT AND FEATURES, Expression profile), τ-value and coefficient of variance (cv) were applied to distinguish HK lncRNAs (τ-value ≤ 0.5 and cv ≤ 0.5) and SR (stress response) or TS (tissue specific) lncRNAs (τ -value ≥ 0.95). Here the index τ was defined as: $\tau = \frac{\sum_{i=1}^{N}(1-x_i)}{N-1}$, where N stands for the number of tissues and x_i represents the expression profile component normalized by the maximal component value (22, 38). To explore the relationship of sample grouping with lncRNA expression profiles, we selected 339 libraries (Table S1) with clearly tissue, variety and treatments information and drew the hierarchical clustering heatmap based on lncRNAs expression FPKMs using R package pheatmap1.0.12 (http://cran.r-project.org/web/packages/pheatmap/index.html).

3) LncRNA genome variation

Single nucleotide polymorphism (SNP) genotyping data were downloaded from 3000 rice genome Projects (http://snpseek.irri.org/_download.zul) (39), which were called against the reference genome Os-Nipponbare-Reference-IRGSP-1.0. The SNP mapped to any lncRNA site (from the start to the end position of lncRNA on the genome) were retrieved as the lncRNA-SNP.

4) Association of LncRNAswith agricultural traits

Two datasets were applied to predict the association of lncRNAs with agricultural traits. GWAS tag SNPs were downloaded from Rice SNP-Seek database (40) and QTLs information from Q-TARO (QTL Annotation Rice Online) database (41). The GWAS SNP contributed phenotype was predicted as possible function of a specific lncRNA if that GWAS tag SNPs co-located with this lncRNA. Accordingly, when a specific lncRNA locus overlapped with a QTL, this QTL-related trait was tagged on this lncRNA as a predicted phenotype association.

5) LncRNA-miRNA interaction and precursor of miRNAs

psRNA Target was used to predict the lncRNA target of microRNAs with the default parameters. This procedure screened candidate interactions of

lncRNA–miRNAs (42). The precursors of miRNAs were screened by comparing lncRNAs sequences with rice pre-miRNAs hairpin sequences (http://www.mirbase.org/). Blast 2.7 was used with the threshold e-value ≤ 10−5, coverage percent bigger than 90% and -max_hsps as 1.

6) Transposon-lncRNA

Genomic coordinates of Japonicatransposon elements (TE) were downloaded from https://www.genome.arizona.edu/cgi-bin/rite/index.cgi(43). The position of TE was compared with respect to lncRNAs in rice genome. LncRNAs overlapping with TE were characterized as TE-lncRNAs associations.

Implementation

We constructed RiceLncPedia database using Django as back end web framework and PostgreSQL(https://www.postgresql.org/) as database engine. JQuery and AJAX (Asynchronous JavaScript and XML) were used to developed web interfaces. As for front-end framework, we employed Bootstrap (https://getbootstrap.com) to supply a series of templates to design web pages with consistent interface components. Additionally, we adopted the icon in Font Awesome in RiceLncPedia website (http://www.fontawesome.com.cn/). Data visualization was powered by Pyecharts (https://github.com/pyecharts/pyecharts) to add interactive diagrams to our website.

DATABASE CONTENT AND FEATURES

In contrast with the existing lncRNA database, the present database features a comprehensive collection of rice lncRNAs from most widely samples and systematic curation of lncRNA annotation through integrating multi-omics data, covering molecular features, expression profiles, sequences features and agricultural traits association.

Number of lncRNA transcripts

RiceLncPedia accommodates 6978 rice lncRNAs which were identified based on 2313 RNA-seq libraries analysis, belonging to 5845 gene loci. The lncRNAs were organized in RiceLncPedia as transcripts. Each lncRNA transcript entity is assigned to a unique accession number and a specific page was linked to each lncRNA transcript

which shows detailed molecular features (Transcript id; Location; Classification; Length; GC Content (%); Exon Number; Sequence; Coding Potential; Genome Browser), Genome Variation, Transposon elements, Small RNA targets, miRNA precursors and expression profile across represented RNA-seq samples (Figure 2).

Multi-omics data integration

a) Expression profile

For each given lncRNA transcript, RiceLncPedia accommodates its expression profiles across all the 2313 collected RNA-seq libraries, which can be searched in expression page or downloaded in download page. Additionally, the expression of lncRNAs from a few represented projects were summarized in detail, covering diverse tissues such as leaf, stem, root, glume and panicle from Indica rice; callus, leaf, panicle before flowering, panicle after flowering, root, seed, and shoot of Japonica cv. Nipponbare, samples from various abiotic stress, such as phosphate starvation, salt stress, cadmium stress, drought stress, cold stress, osmotic stress and flood stress as well as samples grown under different hormone treatments, covering JA treatment and ABA treatment. The expression profiles can be visualized in a bar chart. Because the specific expression in a specific tissue or under a specific condition indicates the function association(38), we calculated the expression breadth, Coefficient of Variance (CV); tissue specificity index and stress responsive index (τ-Value) for each lncRNA transcript. These features greatly facilitate users to explore the lncRNAs functional associations. Users can easily search the database with CV, τ-Value, expression breadth in each selected dataset or the whole datasets (Figure 3). To explore the relationship of sample grouping with lncRNA expression profiles, we clustered the represented 339 libraries mentioned above with all 6978 lncRNAs expression values (Figure 4, Table S1). The resultant clusters were well matched between the indica and japonica groups, basically indicating the reliability of lncRNA expression profiles in RiceLncPedia. We clustered the samples into seven groups based on the hierarchical clustering heatmap. In group I, all samples belong to japonica group. The majority of tissues is root and the minor tissue is shoot. The treatments refer to JA, ABA treatment, osmotic, drought and high Cadmium stresses. Inside this group, the treatment type is a dominant feature of sub classification. Group II includes japonica leaf and shoots samples. Salt and cold treatments were grouped

together, suggesting a possible similar mechanism of plant responsive salt and cold stresses. Group III covers only Japonica panicle tissues. The dominant samples in group IV are Japonica root and shoot tissues under different level of phosphate treatments. All Indica samples were classified into group V, covering different tissues such as glume, panicle, stem, leaf and root. All samples in group VI belong to Japonica shoot tissue. Although the treatments in this group refer to JA, ABA treatment, osmotic, cold, flood, drought and high Cadmium stresses and development stages, most of the developmental stages under normal conditions can still be grouped together. Some of other treatments such as JA treatment, drought, Cadmium stress samples are dispersed in developmental samples. It should be noted that these actually are the control samples in those treatments. This indicated that our lncRNA expression profiles could distinguish the control and the various treatments samples. The similar situation also occurs in group VII, where all samples belong to japonica group root tissue and the samples under control and stress conditions can be distinguished roughly. Additionally, some of different treatments such as salt, drought, cold, flood and osmotic stresses were clustered closely, indicating to some extent the similar co-expression network and plant responsive mechanisms to these stresses.

b) Genome variation of lncRNA

SNPs in lncRNAs were reported to be linked to lncRNA expression, structure, function and phenotypes (21). SNPs in miRNA target sites on lncRNAs may influence the miRNA-lncRNA interaction, and thereby alter their functions (44). SNPs in lncRNA transcripts have also been documented to be able to influence the lncRNA secondary structure and thereby their functions (16,45). A number of databases for SNP collection of lncRNAs in human were constructed (21,44). The databases of rice genome variation annotation linked to different features were constructed in recent years (39,46). However, the SNP annotation of lncRNAs in rice remained to be systematically collected. In RiceLncPedia, the SNPs based on 3000 genome Projects (http://snpseek.irri.org/_download.zul) were compared with the position of lncRNAs and a SNP was tagged as a lncRNA-SNP if it resides in any lncRNA. The information can promote the research of lncRNA variation association with their structures, expressions, interactions and functions. We totally identified 40758 SNPs in 5817 lncRNA transcripts with an average of about 7 SNPs per lncRNA transcript (Figure 4).

c) The prediction of LncRNA functions based on relative position with QTLs and

GWAS tag SNPs

Trait associated SNPs were found in or nearby lncRNAs in human (47-50). In plants, an increasing number of researches have reported the regulation of agricultural traits by lncRNA-SNP identified through GWAS (16,51,52). Additionally, the co-localization of lncRNA with QTLs of complex traits has been adopted as an effective way to infer the function of lncRNAs (53-55). The fruitful information of rice GWAS and QTLs were documented recently. In RiceLncPedia, we constructed lncRNA-SNP-phenotype association if anyrice agricultural GWAS tag SNP co-located with a specific lncRNA. Similarly, a lncRNA resided in any rice QTL was also thought being association with the relevant trait. We finally found that 326 GWAS tag SNPs reside in 71 lncRNAs transcripts, which refers to 11 agricultural traits (Figure 5A). On the other hand, we found 6717 rice lncRNAs collocated with 609 traits related QTLs, such as 1000 grain weight, drought tolerance and so on, belonging to 25 tissues, development stages or stress tolerance, which are included in morphological, physiological, resistance or tolerance and other agricultural traits. The fact that most of lncRNAs (6717 out of 6978) have been co-localized in QTLs might be due to the large interval length of some QTLs. For instance, a plant height related QTL, qPH-2, identified in 2003, spans from 5263536 to 30654749 on chromosome 2 with 505 lncRNA transcripts. These QTLs were identified in earlier years, when the resolution of molecular markers, such as RFLP were not high sufficiently (Figure 5B).

d) LncRNA-miRNA association

A few databases referring to the interactions of lncRNAs and miRNAs were constructed for human (21,56,57). In plants, an increasing number of studies reported lncRNAsare able to execute their roles through being targeted by specific miRNA (58-60). On the other hand, lncRNAs can act as miRNA precursors in different developmental stages in plants (61,62). To facilitate the function prediction of rice lncRNAs, the LncRNA-miRNA interactions based on the prediction by psRNA Target and those lncRNAs as precursors of miRNA were included in RiceLncPedia. In total, 6940 lncRNAs were predicted as the targets of 8184 miRNA, building up 754034lncRNA-miRNA interactions. Among of which, 6112 lncRNAs were predicted as the targets of 713 *Oraza sativa*. miRNAs, building up 65998 lncRNA and osa-miRNA interactions (Figure 6A). We also compared lncRNAs with rice miRNAs precursors and found that 335 lncRNAs have high homology with 52 pre-miRNAs, involving 590 precursor relations of lncRNAs and miRNAs (Figure 6B).

e) TE-related lncRNAs

A number of lncRNAs were reported to be originated from transposons in plants and it was demonstrated that TE associated lncRNAs show tissue-specific transcription and play vital roles in plant abiotic stress responses (63,64). For this reason, LncRNAs overlapping with TE were contained in RiceLncPedia as TE-lncRNAs associations. We overall identified 380 transposons overlapped with 479 lncRNA transcripts, involving 505 transposon and lncRNA transcript relations (Figure 7).

DISCUSSION AND FUTURE DIRECTION

RiceLncPedia is designed to integrate rice lncRNAs from multiple tissues under diverse conditions in a wide range of rice cultivars. Compared with the current databases containing rice lncRNAs, RiceLncPedia is more comprehensive in terms of its covered samples in which rice lncRNAs were identified and the associated lncRNA features extracted from multi-omics data, including expression profile, genome variation of lncRNA loci and association with phenotypes, LncRNA-miRNA interaction, LncRNA as potential miRNA precursors and TE-related lncRNAs. The current version of RiceLncPedia includes 6978 lncRNAs and their multifaceted genetic features. In summary, RiceLncPedia is a rich knowledge reserve of rice lncRNAs and able to serve as a valuable resource for worldwide rice research communities. Future development of RiceLncPedia will refer to regular update of newly discovered rice lncRNAs, integration of differentially expressed lncRNAs in more diverse tissues and environments, epigenetic features of lncRNAs and the association of lncRNAs with protein coding genes, experimentally validated lncRNAs and more lncRNA-phenotype associations. We are looking forward to any reasonable suggestions from worldwide scientists, with the aim to provide a continually updated and more comprehensive rice lncRNAs database.

Fundings

We would like to thank the fundings of the self-determined research fund of Central China Normal University from the colleges's basic research and operation of MOE (Grant No. CCNU18QN027),the project of Hubei Key Laboratory of Genetic Regulation and IntegrativeBiology (GRIB201911) andthe National Special Key

Project of China on TransgenicResearch (Grant No. 2016ZX 08001-003).

Conflict of interest statement. None declared.

Reference

1. Chekanova, J.A. (2015) Long non-coding RNAs and their functions in plants. *Curr Opin Plant Biol*, 27, 207-216.

2. Kumar, M.S., Armenteros-Monterroso, E., East, P., Chakravorty, P., Matthews, N., Winslow, M.M. and Downward, J. (2014) HMGA2 functions as a competing endogenous RNA to promote lung cancer progression. *Nature*, 505, 212-217.

3. Castro-Oropeza, R., Melendez-Zajgla, J., Maldonado, V. and Vazquez-Santillan, K. (2018) The emerging role of lncRNAs in the regulation of cancer stem cells. *Cell Oncol (Dordr)*, 41, 585-603.

4. Huarte, M. (2015) The emerging role of lncRNAs in cancer. *Nat Med*, 21, 1253-1261.

5. Zhu, R., Hu, X., Xu, W., Wu, Z., Zhu, Y., Ren, Y. and Cheng, L. (2019) LncRNA MALAT1 inhibits hypoxia/reoxygenation-induced human umbilical vein endothelial cell injury via targeting the microRNA-320a/RAC1 axis. *Biol Chem*.

6. Bao, W., Cao, F., Ni, S., Yang, J., Li, H., Su, Z. and Zhao, B. (2019) lncRNA FLVCR1-AS1 regulates cell proliferation, migration and invasion by sponging miR-485-5p in human cholangiocarcinoma. *Oncol Lett*, 18, 2240-2247.

7. Mahmoud, A.D., Ballantyne, M.D., Miscianinov, V., Pinel, K., Hung, J., Scanlon, J.P., Iyinikkel, J., Kaczynski, J., Tavares, A.S., Bradshaw, A.C. *et al.* (2019) The Human-Specific and Smooth Muscle Cell-Enriched LncRNA SMILR Promotes Proliferation by Regulating Mitotic CENPF mRNA and Drives Cell-Cycle Progression Which Can Be Targeted to Limit Vascular Remodeling. *Circ Res*, 125, 535-551.

8. Mach, J. (2017) The Long-Noncoding RNA ELENA1 Functions in Plant Immunity. *Plant Cell*, 29, 2.

9. David, R., Burgess, A., Parker, B., Li, J., Pulsford, K., Sibbritt, T., Preiss, T. and Searle, I. (2017) Transcriptome-Wide Mapping of RNA 5-Methylcytosine in Arabidopsis mRNAs and Noncoding RNAs. *Plant Cell*, 29, 445-460.

10. X, H., http://orcid.org, I.-O., X, K., C, W., L, M., J, Z., J, W., X, Z., GJ, L., T, Z. *et al.* (2014) Proteasome-mediated degradation of FRIGIDA modulates flowering time in Arabidopsis during vernalization. *Plant Cell*, 26, 4763-4781.

11. Cui, J., Jiang, N., Meng, J., Yang, G., Liu, W., Zhou, X., Ma, N., Hou, X. and Luan, Y. (2018) LncRNA33732-RESPIRATORY BURST OXIDASE module associated with WRKY1 in tomato-Phytophthora infestans interactions. LID - 10.1111/tpj.14173 [doi]. *Plant J*, 25, 14173.

12. Tan, F., Lu, Y., Jiang, W., Wu, T., Zhang, R., Zhao, Y. and Zhou, D. (2018) DDM1 Represses Noncoding RNA Expression and RNA-Directed DNA Methylation in Heterochromatin. *Plant Physiol*, 177, 1187-1197.

13. Qin, T., Zhao, H., Cui, P., Albesher, N. and Xiong, L. (2017) A Nucleus-Localized Long Non-Coding RNA Enhances Drought and Salt Stress Tolerance. *Plant Physiol*, 175, 1321-1336.

14. Shin, J. and Chekanova, J. (2014) Arabidopsis RRP6L1 and RRP6L2 function in FLOWERING

LOCUS C silencing via regulation of antisense RNA synthesis. *PLoS Genet*, 10.

15. Csorba, T., Questa, J., Sun, Q. and Dean, C. (2014) Antisense COOLAIR mediates the coordinated switching of chromatin states at FLC during vernalization. *Proc Natl Acad Sci U S A*, 111, 16160-16165.

16. Ding, J., Lu, Q., Ouyang, Y., Mao, H., Zhang, P., Yao, J., Xu, C., Li, X., Xiao, J. and Zhang, Q. (2012) A long noncoding RNA regulates photoperiod-sensitive male sterility, an essential component of hybrid rice. *Proc Natl Acad Sci U S A*, 109, 2654-2659.

17. Bu, D., Yu, K., Sun, S., Xie, C., Skogerbo, G., Miao, R., Xiao, H., Liao, Q., Luo, H., Zhao, G. *et al.* (2012) NONCODE v3.0: integrative annotation of long noncoding RNAs. *Nucleic Acids Res*, 40, D210-215.

18. Quek, X.C., Thomson, D.W., Maag, J.L., Bartonicek, N., Signal, B., Clark, M.B., Gloss, B.S. and Dinger, M.E. (2015) lncRNAdb v2.0: expanding the reference database for functional long noncoding RNAs. *Nucleic Acids Res*, 43, D168-173.

19. Volders, P.J., Helsens, K., Wang, X., Menten, B., Martens, L., Gevaert, K., Vandesompele, J. and Mestdagh, P. (2013) LNCipedia: a database for annotated human lncRNA transcript sequences and structures. *Nucleic Acids Res*, 41, D246-251.

20. Park, C., Yu, N., Choi, I., Kim, W. and Lee, S. (2014) lncRNAtor: a comprehensive resource for functional investigation of long non-coding RNAs. *Bioinformatics*, 30, 2480-2485.

21. Gong, J., Liu, W., Zhang, J., Miao, X. and Guo, A.Y. (2015) lncRNASNP: a database of SNPs in lncRNAs and their potential functions in human and mouse. *Nucleic Acids Res*, 43, D181-186.

22. Ma, L., Cao, J., Liu, L., Du, Q., Li, Z., Zou, D., Bajic, V.B. and Zhang, Z. (2019) LncBook: a curated knowledgebase of human long non-coding RNAs. *Nucleic Acids Res*, 47, 2699.

23. Priyanka, G. and Pankaj, J. (2016) Databases and bioinformatics tools for rice research. *Current Plant Biology*, 7-8, 39–52.

24. Sharma, R., Cao, P., Jung, K.H., Sharma, M.K. and Ronald, P.C. (2013) Construction of a rice glycoside hydrolase phylogenomic database and identification of targets for biofuel research. *Front Plant Sci*, 4, 330.

25. Jin, J., Liu, J., Wang, H., Wong, L. and Chua, N.H. (2013) PLncDB: plant long non-coding RNA database. *Bioinformatics*, 29, 1068-1071.

26. Paytuvi Gallart, A., Hermoso Pulido, A., Anzar Martinez de Lagran, I., Sanseverino, W. and Aiese Cigliano, R. (2016) GREENC: a Wiki-based database of plant lncRNAs. *Nucleic Acids Res*, 44, D1161-1166.

27. Paytuvi-Gallart, A., Sanseverino, W. and Aiese Cigliano, R. (2019) A Walkthrough to the Use of GreeNC: The Plant lncRNA Database. *Methods Mol Biol*, 1933, 397-414.

28. Szczesniak, M.W., Bryzghalov, O., Ciomborowska-Basheer, J. and Makalowska, I. (2019) CANTATAdb 2.0: Expanding the Collection of Plant Long Noncoding RNAs. *Methods Mol Biol*, 1933, 415-429.

29. Yi, X., Zhang, Z., Ling, Y., Xu, W. and Su, Z. (2015) PNRD: a plant non-coding RNA database. *Nucleic Acids Res*, 43, D982-989.

30. Xuan, H., Zhang, L., Liu, X., Han, G., Li, J., Li, X., Liu, A., Liao, M. and Zhang, S. (2015) PLNlncRbase: A resource for experimentally identified lncRNAs in plants. *Gene*, 573, 328-332.

31. Zhou, B., Zhao, H., Yu, J., Guo, C., Dou, X., Song, F., Hu, G., Cao, Z., Qu, Y., Yang, Y. *et al.*

(2019) Experimentally Validated Plant lncRNAs in EVLncRNAs Database. *Methods Mol Biol*, 1933, 431-437.

32. Singh, U., Khemka, N., Rajkumar, M., Garg, R. and Jain, M. (2017) PLncPRO for prediction of long non-coding RNAs (lncRNAs) in plants and its application for discovery of abiotic stress-responsive lncRNAs in rice and chickpea. *Nucleic Acids Res*, 45.

33. Pertea, M., Kim, D., Pertea, G.M., Leek, J.T. and Salzberg, S.L. (2016) Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat Protoc*, 11, 1650-1667.

34. Kang, Y.J., Yang, D.C., Kong, L., Hou, M., Meng, Y.Q., Wei, L. and Gao, G. (2017) CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features. *Nucleic Acids Res*, 45, W12-W16.

35. El-Gebali, S., Mistry, J., Bateman, A., Eddy, S.R., Luciani, A., Potter, S.C., Qureshi, M., Richardson, L.J., Salazar, G.A., Smart, A. *et al.* (2019) The Pfam protein families database in 2019. *Nucleic Acids Res*, 47, D427-D432.

36. Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L. and Pachter, L. (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc*, 7, 562-578.

37. Sun, L., Zhang, Z., Bailey, T.L., Perkins, A.C., Tallack, M.R., Xu, Z. and Liu, H. (2012) Prediction of novel long non-coding RNAs based on RNA-Seq data of mouse Klf1 knockout study. *BMC Bioinformatics*, 13, 331.

38. Yanai, I., Benjamin, H., Shmoish, M., Chalifa-Caspi, V., Shklar, M., Ophir, R., Bar-Even, A., Horn-Saban, S., Safran, M., Domany, E. *et al.* (2005) Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics*, 21, 650-659.

39. Mansueto, L., Fuentes, R.R., Borja, F.N., Detras, J., Abriol-Santos, J.M., Chebotarov, D., Sanciangco, M., Palis, K., Copetti, D., Poliakov, A. *et al.* (2017) Rice SNP-seek database update: new SNPs, indels, and queries. *Nucleic Acids Res*, 45, D1075-D1081.

40. Sanciangco, M.D., Alexandrov, N.N., Chebotarov, D., King, R.D., Naredo, M.E.B., Leung, H., Mansueto, L., Mauleon, R.P., Orhobor, O.I. and McNally, K.L. (2018) In Sanciangco, M. (ed.). V2 ed. Harvard Dataverse.

41. Yamamoto, E., Yonemaru, J., Yamamoto, T. and Yano, M. (2012) OGRO: The Overview of functionally characterized Genes in Rice online database. *Rice (N Y)*, 5, 26.

42. 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.

43. Copetti, D., Zhang, J., El Baidouri, M., Gao, D., Wang, J., Barghini, E., Cossu, R.M., Angelova, A., Maldonado, L.C., Roffler, S. *et al.* (2015) RiTE database: a resource database for genus-wide rice genomics and evolutionary biology. *BMC Genomics*, 16, 538.

44. Ning, S., Yue, M., Wang, P., Liu, Y., Zhi, H., Zhang, Y., Zhang, J., Gao, Y., Guo, M., Zhou, D. *et al.* (2017) LincSNP 2.0: an updated database for linking disease-associated SNPs to human long non-coding RNAs and their TFBSs. *Nucleic Acids Res*, 45, D74-D78.

45. Hawkes, E., Hennelly, S., Novikova, I., Irwin, J., Dean, C. and Sanbonmatsu, K. (2016) COOLAIR Antisense RNAs Form Evolutionarily Conserved Elaborate Secondary Structures. *Cell Rep*, 16, 3087-3096.

46. Zhao, H., Yao, W., Ouyang, Y., Yang, W., Wang, G., Lian, X., Xing, Y., Chen, L. and Xie, W. (2015) RiceVarMap: a comprehensive database of rice genomic variations. *Nucleic Acids Res*, 43, D1018-1022.

47. Roca-Ayats, N., Martinez-Gil, N., Cozar, M., Gerousi, M., Garcia-Giralt, N., Ovejero, D., Mellibovsky, L., Nogues, X., Diez-Perez, A., Grinberg, D. *et al.* (2019) Functional characterization of the C7ORF76 genomic region, a prominent GWAS signal for osteoporosis in 7q21.3. *Bone*, 123, 39-47.

48. Mei, B., Wang, Y., Ye, W., Huang, H., Zhou, Q., Chen, Y., Niu, Y., Zhang, M. and Huang, Q. (2019) LncRNA ZBTB40-IT1 modulated by osteoporosis GWAS risk SNPs suppresses osteogenesis. *Hum Genet*, 138, 151-166.

49. Giral, H., Landmesser, U. and Kratzer, A. (2018) Into the Wild: GWAS Exploration of Non-coding RNAs. *Front Cardiovasc Med*, 5, 181.

50. Pasmant, E., Sabbagh, A., Vidaud, M. and Bieche, I. (2011) ANRIL, a long, noncoding RNA, is an unexpected major hotspot in GWAS. *FASEB J*, 25, 444-448.

51. Tian, J., Song, Y., Du, Q., Yang, X., Ci, D., Chen, J., Xie, J., Li, B. and Zhang, D. (2016) Population genomic analysis of gibberellin-responsive long non-coding RNAs in Populus. *J Exp Bot*, 67, 2467-2482.

52. Wang, H., Niu, Q., Wu, H., Liu, J., Ye, J., Yu, N. and Chua, N. (2015) Analysis of non-coding transcriptome in rice and maize uncovers roles of conserved lncRNAs associated with agriculture traits. *Plant J*, 84, 404-416.

53. Cai, W., Li, C., Liu, S., Zhou, C., Yin, H., Song, J., Zhang, Q. and Zhang, S. (2018) Genome Wide Identification of Novel Long Non-coding RNAs and Their Potential Associations With Milk Proteins in Chinese Holstein Cows. *Front Genet*, 9, 281.

54. Lin, C., Fesi, B.D., Marquis, M., Bosak, N.P., Lysenko, A., Koshnevisan, M.A., Duke, F.F., Theodorides, M.L., Nelson, T.M., McDaniel, A.H. *et al.* (2018) Burly1 is a mouse QTL for lean body mass that maps to a 0.8-Mb region of chromosome 2. *Mamm Genome*, 29, 325-343.

55. Du, Z.Q., Eisley, C.J., Onteru, S.K., Madsen, O., Groenen, M.A., Ross, J.W. and Rothschild, M.F. (2014) Identification of species-specific novel transcripts in pig reproductive tissues using RNA-seq. *Anim Genet*, 45, 198-204.

56. Das, S., Ghosal, S., Sen, R. and Chakrabarti, J. (2014) lnCeDB: database of human long noncoding RNA acting as competing endogenous RNA. *PLoS One*, 9, e98965.

57. Sarver, A.L. and Subramanian, S. (2012) Competing endogenous RNA database. *Bioinformation*, 8, 731-733.

58. Xu, X.W., Zhou, X.H., Wang, R.R., Peng, W.L., An, Y. and Chen, L.L. (2016) Functional analysis of long intergenic non-coding RNAs in phosphate-starved rice using competing endogenous RNA network. *Sci Rep*, 6, 20715.

59. Huang, D., Feurtado, J., Smith, M., Flatman, L., Koh, C. and Cutler, A. (2017) Long noncoding miRNA gene represses wheat beta-diketone waxes. *Proc Natl Acad Sci U S A*, 114, E3149-E3158.

60. Yuan, J., Zhang, Y., Dong, J., Sun, Y., Lim, B.L., Liu, D. and Lu, Z.J. (2016) Systematic characterization of novel lncRNAs responding to phosphate starvation in Arabidopsis thaliana. *BMC Genomics*, 17, 655.

61. Song, X., Liu, G., Huang, Z., Duan, W., Tan, H., Li, Y. and Hou, X. (2016) Temperature expression patterns of genes and their coexpression with LncRNAs revealed by RNA-Seq in non-heading Chinese cabbage. *BMC Genomics*, 17, 297.

62. Liu, J., Jung, C., Xu, J., Wang, H., Deng, S., Bernad, L., Arenas-Huertero, C. and Chua, N.-H. (2012) Genome-Wide Analysis Uncovers Regulation of Long Intergenic Noncoding RNAs in Arabidopsis. *The Plant Cell*, 24, 4333-4345.

63. Cho, J. (2018) Transposon-Derived Non-coding RNAs and Their Function in Plants. *Front Plant Sci*, 9.

64. Wang, D., Qu, Z., Yang, L., Zhang, Q., Liu, Z., Do, T., Adelson, D., Wang, Z., Searle, I. and Zhu, J. (2017) Transposable elements (TEs) contribute to stress-related long intergenic noncoding RNAs in plants. *Plant J*, 90, 133-146.

Table 1. Key differences between the existing rice lncRNA databases and RiceLncPedia**.**

Data Item	**RiceLncPedia**	GREENC	CANTATAdb 2.0	PNRD	PLNlncRbase	PLncPRO	EVLncRNAs
Species	**Rice**	43 species	39 species	150 species	43 species	Rice, chickpea	77
LncRNA entries	**6978**	120,000	239, 631	25739 ncRNAs	1187	12314	1543
Rice LncRNA entries	**6978**	5237	2788	790	unknown*	7345	40
Rice RNA-seq libraries	**2313**	NA	8	NA	NA	9	NA
Experimentally identified lncRNAs		NA	NA	Partial	1060	NA	√
Sequences	√	√	√	√	√	√	√
coding potential	√	√	√	√	NA	NA	NA
Genomic information	√	√	√	√	√	NA	√
Expression	√	NA	√	NA	√	NA	√
Small RNA associations	√	√	NA	NA	NA	NA	NA
LncRNA genome variation	√	NA	NA	NA	NA	NA	NA
Phenotype association	√	NA	NA	NA	√	NA	√
TE-lncRNAs	√	√	NA	NA	NA	NA	NA
Accessible	√	√	√	√	NA	√	√

Note: * represents the unknown number of rice lncRNAs since the database is not accessible following the link in the reference.

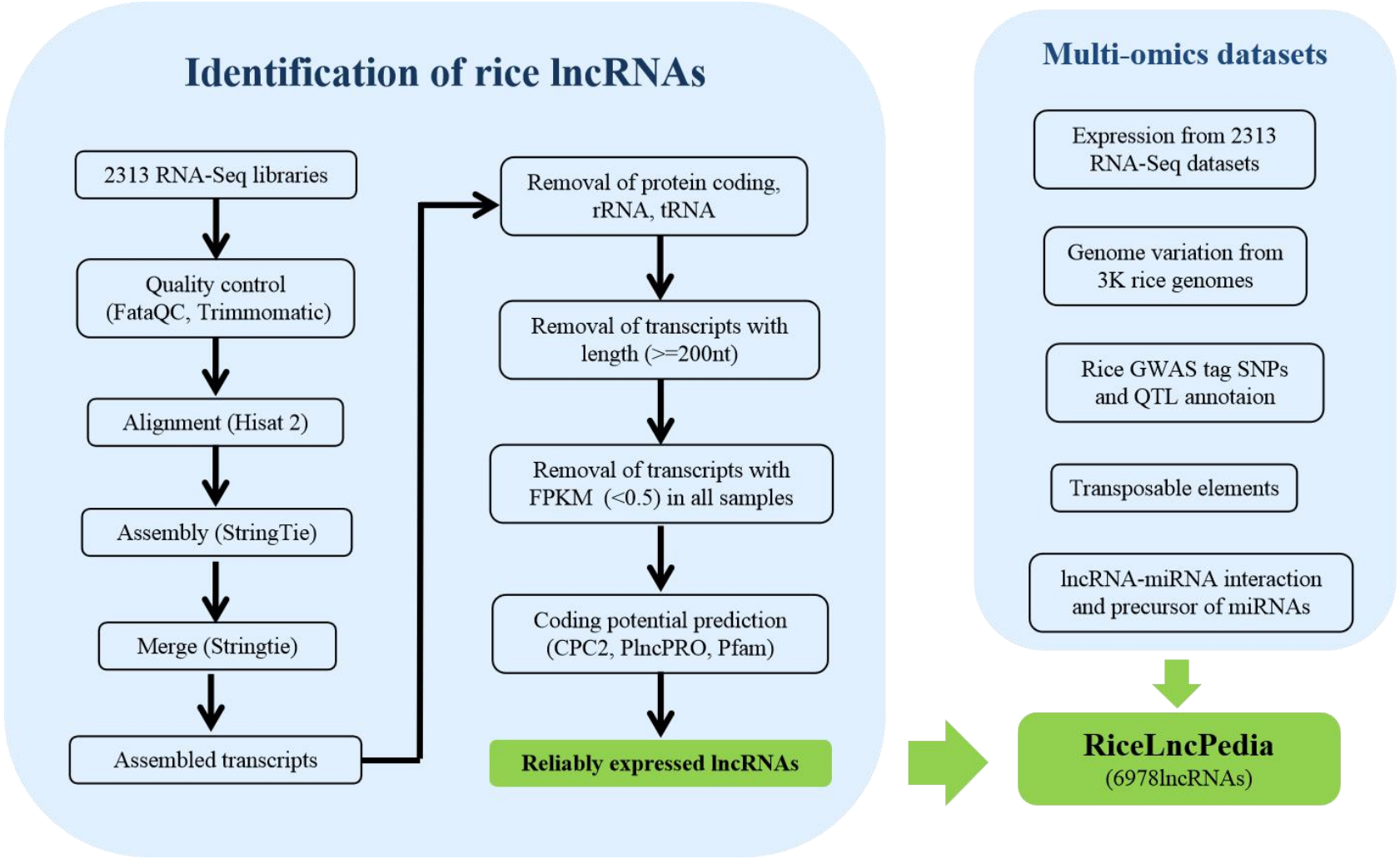


Figure 1. Pipeline for rice lncRNAs identification and multi-omics data integration.

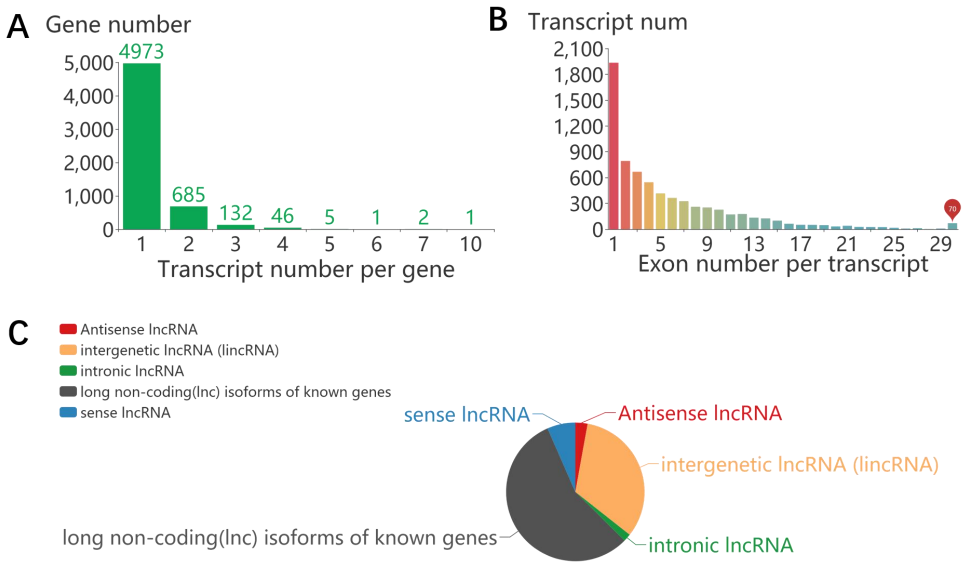
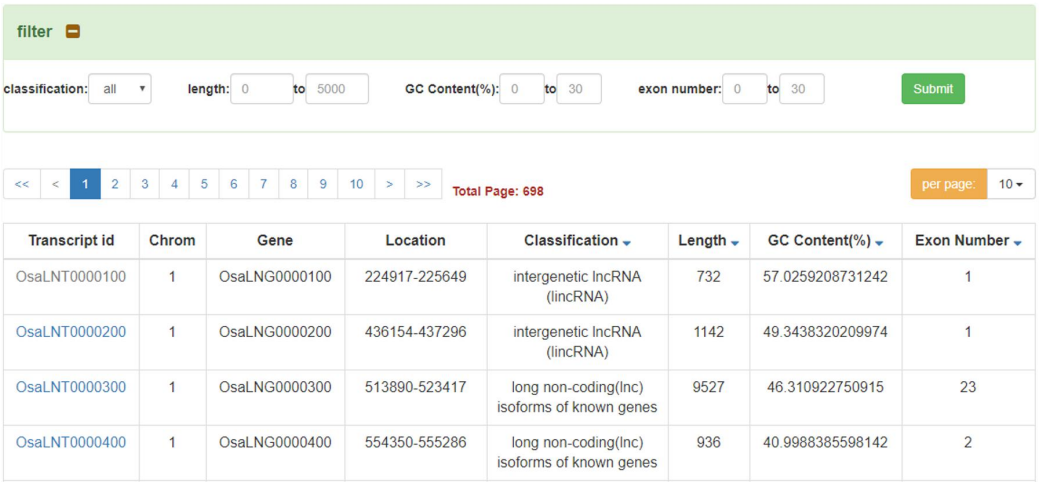


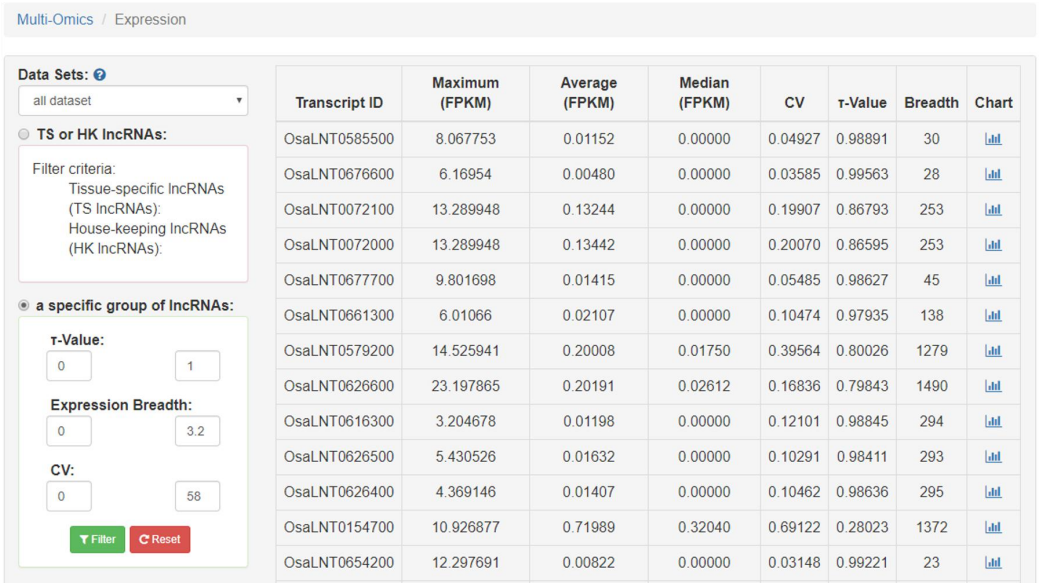
Figure 2. The basic molecular features of lncRNAs. A. The distribution of transcript numbers in lncRNA genes. B. The distribution of exon numbers in lncRNA transcripts. C. The classification of lncRNAs based on the positions of lncRNAs relative to protein coding genes.

A



Transcript id	Chrom	Gene	Location	Classification	Length	GC Content(%)	Exon Number
OsaLNT0000100	1	OsaLNG0000100	224917-225649	intergenetic lncRNA (lincRNA)	732	57.0259208731242	1
OsaLNT0000200	1	OsaLNG0000200	436154-437296	intergenetic lncRNA (lincRNA)	1142	49.3438320209974	1
OsaLNT0000300	1	OsaLNG0000300	513890-523417	long non-coding(lnc) isoforms of known genes	9527	46.310922750915	23
OsaLNT0000400	1	OsaLNG0000400	554350-555286	long non-coding(lnc) isoforms of known genes	936	40.9988385598142	2

B



Transcript ID	Maximum (FPKM)	Average (FPKM)	Median (FPKM)	CV	τ-Value	Breadth	Chart
OsaLNT0585500	8.067753	0.01152	0.00000	0.04927	0.98891	30	
OsaLNT0676600	6.16954	0.00480	0.00000	0.03585	0.99563	28	
OsaLNT0072100	13.289948	0.13244	0.00000	0.19907	0.86793	253	
OsaLNT0072000	13.289948	0.13442	0.00000	0.20070	0.86595	253	
OsaLNT0677700	9.801698	0.01415	0.00000	0.05485	0.98627	45	
OsaLNT0661300	6.01066	0.02107	0.00000	0.10474	0.97935	138	
OsaLNT0579200	14.525941	0.20008	0.01750	0.39564	0.80026	1279	
OsaLNT0626600	23.197865	0.20191	0.02612	0.16836	0.79843	1490	
OsaLNT0616300	3.204678	0.01198	0.00000	0.12101	0.98845	294	
OsaLNT0626500	5.430526	0.01632	0.00000	0.10291	0.98411	293	
OsaLNT0626400	4.369146	0.01407	0.00000	0.10462	0.98636	295	
OsaLNT0154700	10.926877	0.71989	0.32040	0.69122	0.28023	1372	
OsaLNT0654200	12.297691	0.00822	0.00000	0.03148	0.99221	23	

Figure 3. A. A snapshot of lncRNA transcripts basic molecular features. Includes transcript ids, locations, classification, length, GC content and exon numbers. B. A snapshot of lncRNA transcripts expression profile. CV, tissue Specificity Index τ -Value and other statistics parameters were calculated for each lncRNA transcript for ten selective datasets and all datasets together.

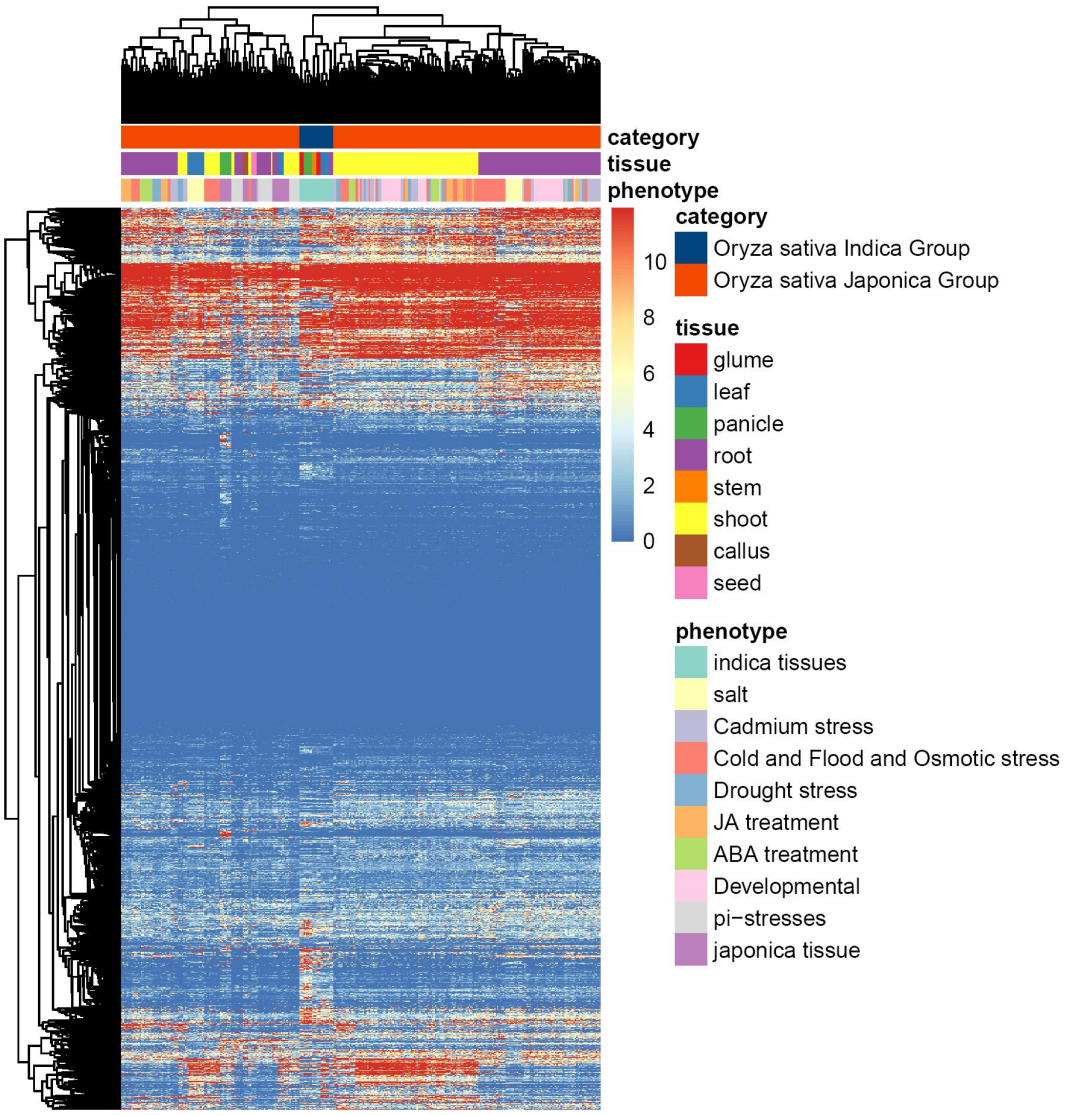


Figure 4. Hierarchical clustering heatmap of 339 selected RNA-seq libraries based on lncRNA expression FPKM values.In order to remove some over-expressed outliers and better distinguish between categories, we replace all expression values greater than 0.9 quantile with 0.9 quantile.

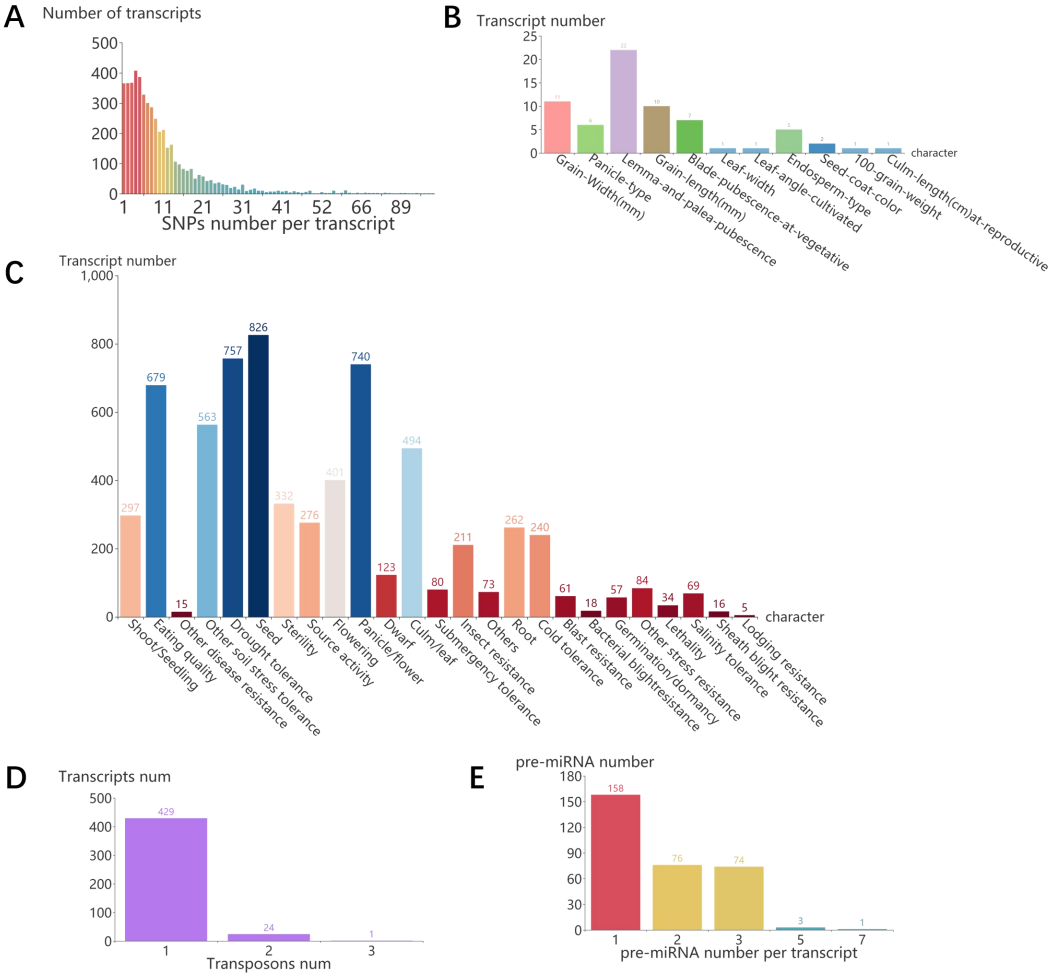
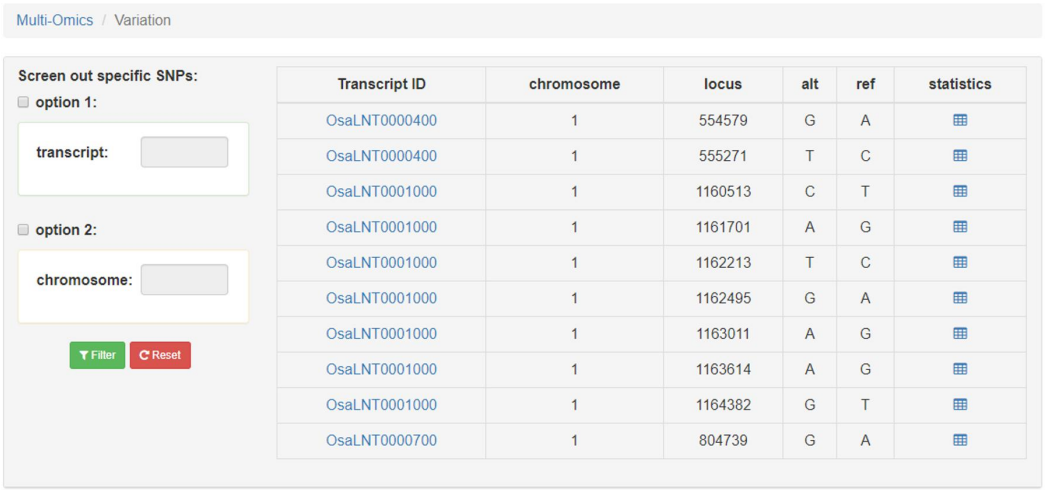


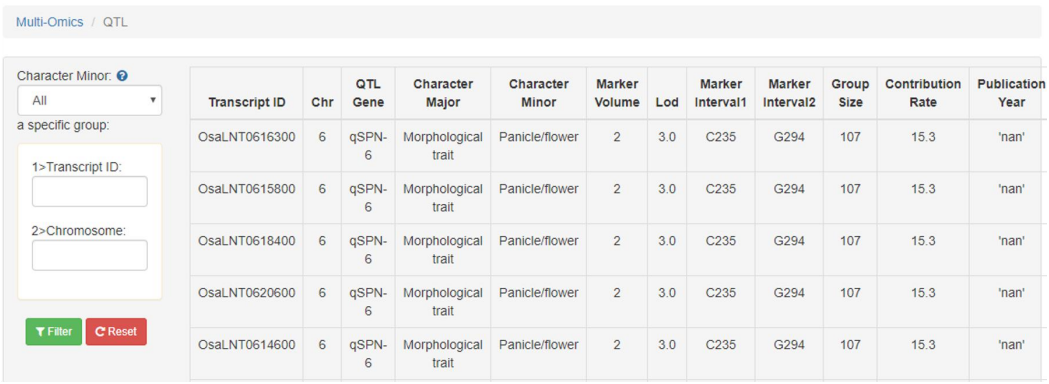
Figure 5. The lncRNAs features extracted from multi-omics data. A. the SNP number distribution of lncRNAs. B. The number distribution of lncRNA transcripts associated with GWAS relevant traits. C. The number distribution of lncRNA transcripts associated with QTL relevant traits. D. The number distribution of lncRNA transcripts potentially as precursors of miRNAs. E. The number distribution of lncRNA transcripts associated with transposons.

A



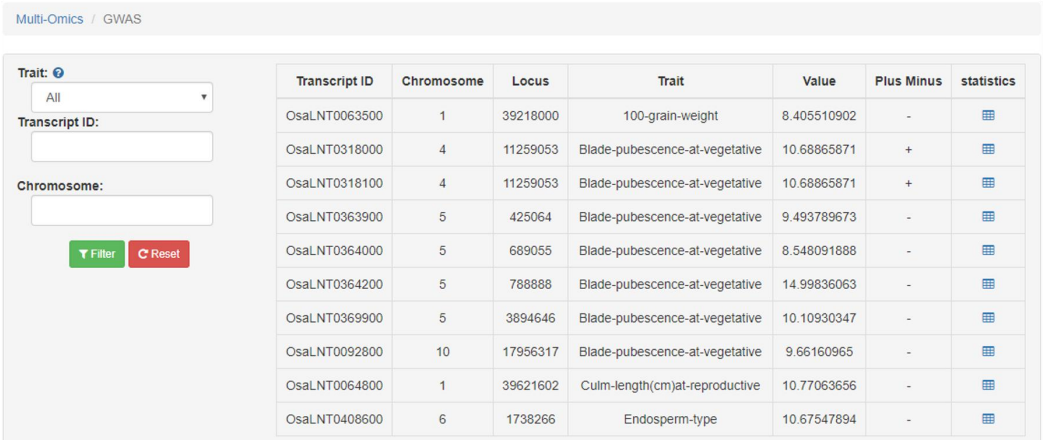
Transcript ID	chromosome	locus	alt	ref	statistics
OsaLNT0000400	1	554579	G	A	
OsaLNT0000400	1	555271	T	C	
OsaLNT0001000	1	1160513	C	T	
OsaLNT0001000	1	1161701	A	G	
OsaLNT0001000	1	1162213	T	C	
OsaLNT0001000	1	1162495	G	A	
OsaLNT0001000	1	1163011	A	G	
OsaLNT0001000	1	1163614	A	G	
OsaLNT0001000	1	1164382	G	T	
OsaLNT0000700	1	804739	G	A	

B



Transcript ID	Chr	QTL Gene	Character Major	Character Minor	Marker Volume	Lod	Marker Interval1	Marker Interval2	Group Size	Contribution Rate	Publication Year
OsaLNT0616300	6	qSPN-6	Morphological trait	Panicle/flower	2	3.0	C235	G294	107	15.3	'nan'
OsaLNT0615800	6	qSPN-6	Morphological trait	Panicle/flower	2	3.0	C235	G294	107	15.3	'nan'
OsaLNT0618400	6	qSPN-6	Morphological trait	Panicle/flower	2	3.0	C235	G294	107	15.3	'nan'
OsaLNT0620600	6	qSPN-6	Morphological trait	Panicle/flower	2	3.0	C235	G294	107	15.3	'nan'
OsaLNT0614600	6	qSPN-6	Morphological trait	Panicle/flower	2	3.0	C235	G294	107	15.3	'nan'

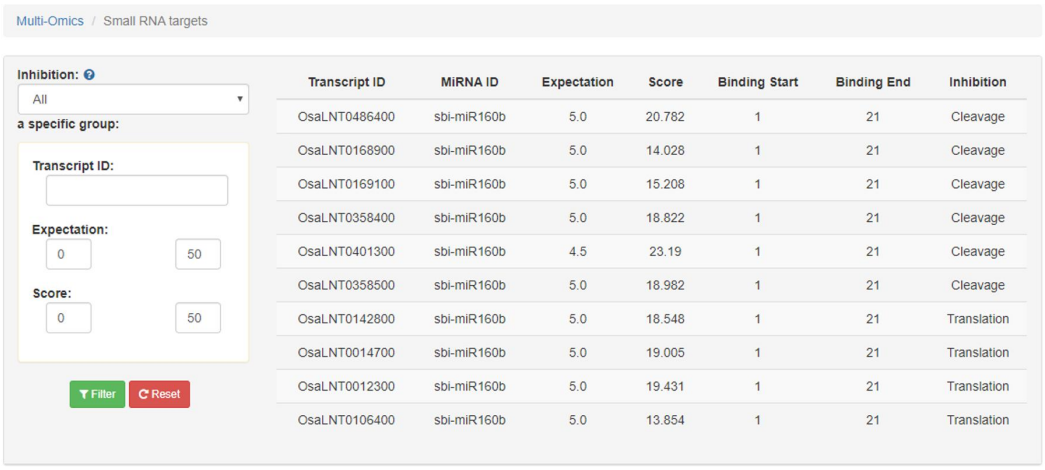
C



Transcript ID	Chromosome	Locus	Trait	Value	Plus Minus	statistics
OsaLNT0063500	1	39218000	100-grain-weight	8.405510902	-	
OsaLNT0318000	4	11259053	Blade-pubescence-at-vegetative	10.68865871	+	
OsaLNT0318100	4	11259053	Blade-pubescence-at-vegetative	10.68865871	+	
OsaLNT0363900	5	425064	Blade-pubescence-at-vegetative	9.493789673	-	
OsaLNT0364000	5	689055	Blade-pubescence-at-vegetative	8.548091888	-	
OsaLNT0364200	5	788888	Blade-pubescence-at-vegetative	14.99836063	-	
OsaLNT0369900	5	3894646	Blade-pubescence-at-vegetative	10.10930347	-	
OsaLNT0092800	10	17956317	Blade-pubescence-at-vegetative	9.66160965	-	
OsaLNT0064800	1	39621602	Culm-length(cm)at-reproductive	10.77063656	-	
OsaLNT0408600	6	1738266	Endosperm-type	10.67547894	-	

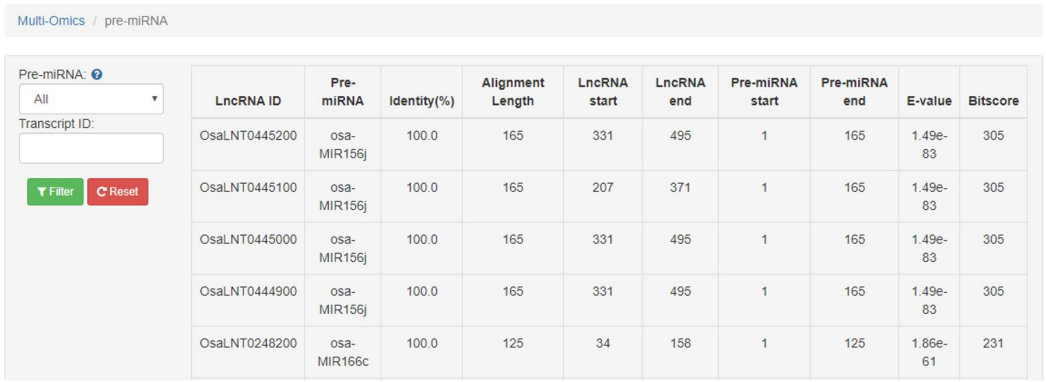
Figure 6. Snapshots for lncRNAs with SNP and QTL features. A. A snapshot of lncRNA transcripts encountering SNPs. B. A snapshot of lncRNA transcripts associated with agricultural traits through QTL interval positions. C. A snapshot of lncRNA transcripts associated with agricultural traits through GWAS tag SNPs.

A



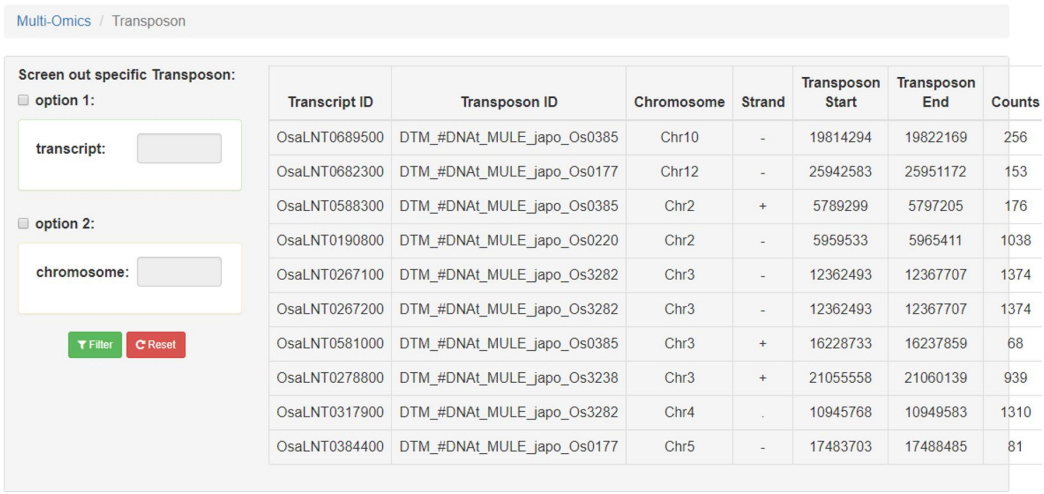
Transcript ID	MiRNA ID	Expectation	Score	Binding Start	Binding End	Inhibition
OsaLNT0486400	sbi-miR160b	5.0	20.782	1	21	Cleavage
OsaLNT0168900	sbi-miR160b	5.0	14.028	1	21	Cleavage
OsaLNT0169100	sbi-miR160b	5.0	15.208	1	21	Cleavage
OsaLNT0358400	sbi-miR160b	5.0	18.822	1	21	Cleavage
OsaLNT0401300	sbi-miR160b	4.5	23.19	1	21	Cleavage
OsaLNT0358500	sbi-miR160b	5.0	18.982	1	21	Cleavage
OsaLNT0142800	sbi-miR160b	5.0	18.548	1	21	Translation
OsaLNT0014700	sbi-miR160b	5.0	19.005	1	21	Translation
OsaLNT0012300	sbi-miR160b	5.0	19.431	1	21	Translation
OsaLNT0106400	sbi-miR160b	5.0	13.854	1	21	Translation

B



LncRNA ID	Pre-miRNA	Identity(%)	Alignment Length	LncRNA start	LncRNA end	Pre-miRNA start	Pre-miRNA end	E-value	Bitscore
OsaLNT0445200	osa-MIR156j	100.0	165	331	495	1	165	1.49e-83	305
OsaLNT0445100	osa-MIR156j	100.0	165	207	371	1	165	1.49e-83	305
OsaLNT0445000	osa-MIR156j	100.0	165	331	495	1	165	1.49e-83	305
OsaLNT0444900	osa-MIR156j	100.0	165	331	495	1	165	1.49e-83	305
OsaLNT0248200	osa-MIR166c	100.0	125	34	158	1	125	1.86e-61	231

C



Transcript ID	Transposon ID	Chromosome	Strand	Transposon Start	Transposon End	Counts
OsaLNT0689500	DTM_#DNAt_MULE_japo_Os0385	Chr10	-	19814294	19822169	256
OsaLNT0682300	DTM_#DNAt_MULE_japo_Os0177	Chr12	-	25942583	25951172	153
OsaLNT0588300	DTM_#DNAt_MULE_japo_Os0385	Chr2	+	5789299	5797205	176
OsaLNT0190800	DTM_#DNAt_MULE_japo_Os0220	Chr2	-	5959533	5965411	1038
OsaLNT0267100	DTM_#DNAt_MULE_japo_Os3282	Chr3	-	12362493	12367707	1374
OsaLNT0267200	DTM_#DNAt_MULE_japo_Os3282	Chr3	-	12362493	12367707	1374
OsaLNT0581000	DTM_#DNAt_MULE_japo_Os0385	Chr3	+	16228733	16237859	68
OsaLNT0278800	DTM_#DNAt_MULE_japo_Os3238	Chr3	+	21055558	21060139	939
OsaLNT0317900	DTM_#DNAt_MULE_japo_Os3282	Chr4	.	10945768	10949583	1310
OsaLNT0384400	DTM_#DNAt_MULE_japo_Os0177	Chr5	-	17483703	17488485	81

Figure 7. Snapshots for lncRNAs with miRNAs and TE features.A. A snapshot of lncRNA transcripts predicted as the targets. B. A snapshot of lncRNA transcripts predicted as precursors of small RNAs. C. A snapshot of lncRNA transcripts associated with transposons.