



Introduction of the Databases of Rice

3

Gayatri Gouda, S. Sabarinathan, Ravindra Donde,
Goutam Kumar Dash, Menaka Ponnana, Manoj Kumar Gupta,
Ramakrishna Vadde, Lambodar Behera, and Trilochan Mohapatra

Abstract

Rice is one of the world's most important agricultural crops and a model plant that is widely studied. The completion of the complete rice (*Oryza sativa*) genome sequence through high-throughput experimental platforms has resulted in a huge amount of data being generated and specialized databases, and bioinformatic tools for data processing, analysis, efficient organization, and visualization have been developed. In this chapter, we address a set of biological databases that host rice-specific sequence, genetic variation, gene expression, pathways, and gene–

S. Sabarinathan, Ravindra Donde, Goutam Kumar Dash, Menaka Ponnana and Manoj Kumar Gupta contributed equally with all other contributors.

G. Gouda (✉) · R. Donde · G. K. Dash · M. K. Gupta · L. Behera
Crop Improvement Division, ICAR-National Rice Research Institute, Cuttack, Odisha, India

S. Sabarinathan
Crop Improvement Division, ICAR-National Rice Research Institute, Cuttack, Odisha, India

Department of Seed Science and Technology, College of Agriculture, Odisha University of Agriculture and Technology, Bhubaneswar, Odisha, India

M. Ponnana
Crop Improvement Division, ICAR-National Rice Research Institute, Cuttack, Odisha, India

Department of Plant Physiology, College of Agriculture, Odisha University of Agriculture and Technology, Bhubaneswar, Odisha, India

R. Vadde
Department of Biotechnology and Bioinformatics, Yogi Vemana University, Kadapa, Andhra Pradesh, India

T. Mohapatra
Secretary (DARE) and Director General (ICAR), Government of India, New Delhi, India

interactome data from various genomic and proteomic sources, and aid in data analysis and visualization.

Keywords

Database · Rice · Genomics · Proteomics · Gene expression

Abbreviations

| | |
|------------|---|
| BAR | Bio-Analytic Resource |
| CSRDB | Cereal Small RNA Database |
| DIPOS | Database of Interacting Proteins in <i>Oryza sativa</i> . |
| ESTs | Expressed Sequence Tags |
| FSTs | Flanking Sequence Tags |
| GEO | Gene Expression Omnibus |
| GRASSIUS | Grass Regulatory Information Services |
| IRGSP | International Rice Genome Sequencing Project |
| MCDRP | Manually Curated Database of Rice Proteins |
| MPIC | Mitochondrial Protein Import Components |
| PDB | Protein Data Bank |
| PLANEX | PLAnt co-EXpression Database |
| PlantDHS | Plant DNaseI Hypersensitive Site Database |
| PmiRKB | Plant miRNA Knowledge Base |
| PMRD | Plant microRNA Database |
| PRIN | Predicted Rice Interactome Network |
| RAP-DB | Rice Annotation Project Database |
| RED | Rice Expression Database |
| RGAP | Rice Genome Annotation Project |
| RICD | Rice Indica cDNA Database |
| RiceSTIFDB | Stress-Responsive Transcription Factor Database |
| RiceVarMap | Rice Variation Map |
| RKD | Rice Kinase Database |
| ROAD | Rice Oligonucleotide Array Database |
| SNPs | single nucleotide polymorphisms |
| TIGR | The Institute for Genomic Research |

3.1 Introduction

Rice (*Oryza sativa* L.) is an important crop from Gramene family because most of the human population mainly depend on rice as their food [1]. However, recent climate change has resulted in drastic decrease in the yield and has also affected the rice genome at molecular level. Previously, it was reported that all the agronomic traits are controlled by various genes and considering that, to date, numerous genes and its sequences have been identified and used in the breeding program to improve the yield and quality of rice [2]. Additionally, since rice has a small genome size, it is often used as a model organism for the monocot plant species. Thus, there is always a quest to unmask the genetic structure and composition of rice genome. Though the recent advancement of high-throughput sequencing technologies has been able to solve these issues, they had led to the accumulation of huge amount of biological data. This in turn demands the development of various bioinformatic databases and tools, which may store and analyze data efficiently. Considering this, to date many rice-specific databases are developed that contain heterogeneous biological data ranging from genomics to proteomics and are regularly exploited to study and evaluate numerous agricultural-related information such as gene mapping, gene identification, functional characterization, gene expression study, structure prediction, protein structure, and function prediction. Thus, in this chapter authors have described in brief about various rice databases and information they contain.

3.2 Genomic Database

Gramene is a comparative genomic database of plants that incorporates knowledge regarding genetic charts, sequences, gene markers, proteins, pathways, and phenotypes [3]. One may search or explore the database to identify genes and phenotypes that share similar characteristics. Different plant organisms may also be correlated and differentiated considering the identical genes, genomes, mechanisms, and phenotypes. From the Gramene database, one may also know the position of genes in the chromosome and its function [4]. Using the genome browser module of the Gramene database, the annotation of rice genome is possible, which provides information on SNPs, indels, and markers for the respective genes. The rice genome also acts as the reference for genome-wide comparative study. In order to decide whether rice plants are linked to other cereals, the researcher uses a number of sequence-based tags including the sequences of expressed sequence tags (ESTs), flanking sequence tags (FSTs), proteins, cDNAs, BACs and BAC ends, SNPs, microarray probes, and repeat sequences [5].

The gene archive takes advantage of the genetic colinearity (synteny) between rice and other main crop plant genomes to annotate genome of organisms whose own plants have not been sequenced [5]. For performing annotation, Gramene employs several modules like BLAST tool, marker module, and QTL modules. BLAST tool aligns the query sequence with the reference genome available for the particular crop with many species. For *Oryza*, the database stored information on 11 wild species of

Oryza such as *Oryza barthii*, *Oryza glaberrima*, *Oryza glumipatula*, *Oryza longistaminata*, *Oryza punctata*, *Oryza rufipogon*, *Oryza brachyantha*, *Oryza sativa indica*, *Oryza meridionalis*, *Oryza sativa japonica*, and *Oryza sativa f spontanea*. The BLAST alignment results provide information regarding the details of the alignment gene with E-value and identity percentage of gene with reference and the orientation of the gene in either positive or negative strand. BLAST has been discussed in detail in Chap. 7 of the book. QTL module provides information about QTLs of rice and several other cereals (*described below in QTL section*). Marker module provides information on markers available for all the genes/QTLs of the crops. The detailed information about microsatellite or SSR markers of rice genome with mapped chromosome position with location is also available, which in turn enables us to easily identify the gene and their nearest gene location. These data provide a gateway to find out the marker position of a particular gene with respect to various traits, like abiotic and biotic traits in rice.

Another widely used rice genomic database is “Rice Annotation Project Database (RAP-DB)” [6, 7]. The information about the rice genome is generated through the “International Rice Genome Sequencing Project (IRGSP)” and is stored for public use in RAP-DB [6, 7]. In this database, the Japanese rice Nipponbare (build IRGSP 1.0) is considered as a reference genome for all the rice genotypes. The genome sequencing was carried out through high-throughput Illumina sequencing platform. RicyerDB is another rice genomic database that contains information regarding the yield-related genes of rice. Both genomic and proteomic data are available in this database, which make researcher to easily get the gene information and the pathways involved for gene expression in rice [8]. The tools of RicyerDB provide base for the development of a shared platform for browsing and visualizations of yield-related genes. Another utility allows the user to search for a single gene and offers insight into the roles and positions of biological processes. The study of protein–protein interaction and protein–protein interaction network construct is also possible through this database. This database also stores information from various sources such as RAP-DB, Rice Genome Annotation Project (RGAP), NCBI, and UniProt and STRING databases.

RiceGAAS database is also known as “Rice Genome Automated Annotation System.” It allows to execute the genomic data of rice for public use [21]. The data are collected from various sources such as IRGSP [70] and submitted to the public domain database DDBJ (<https://www.ddbj.nig.ac.jp/services-e.html>) and provided all the information regarding Gene entry, gene homology identification, prediction of long terminal repeats, and gene models. The RiceRelativesGD is another gene database aimed at providing a genetic resource useful for rice breeding. In 2019, Mao et al. [71] developed this database for identifying the genetic information of close rice relative species. This database contains genomic information from 13 separate rice relatives such as *O. sativa (japonica group)*, *O. sativa (indica group)*, *O. rufipogon*, *O. barthii*, *O. glumaepatula*, *O. meridionalis*, *O. nivara*, *O. punctata*, *O. brachyantha*, *Leersia perrieri*, *O. glaberrima*, *Zizania latifolia*, and *Echinochloa crus-galli* which are accessible to the public. Their study provides knowledge on

genes specific for different functions such as stress and photosynthesis that are used in breeding program.

3.3 QTL/Gene Database

QTL databases are used to identify the gene function that are used in breeding program [28]. For instance, Q-TARO database provides detailed information of rice QTLs. One important feature is that it enlists a table comprising all of the QTLs and their genetic parameters such as their trait or trait type, population, mapping, accuracy (LOD value), and map location of the QTL. Another important feature is the genome browser displaying genomic positions of QTLs. Q-TARO also specifically displays the colinear spatial structures of QTLs and QTL regions on the rice genome [27].

For plants, Gramene is another quantitative information resource that incorporates data through different data domains. In Gramene QTL database, QTLs are identified as a part of gene, which is associated with a particular phenotype. The QTL database includes the world's largest online repository of QTL data for rice. QTLs initially published on individual genetic maps have been systematically aligned to the rice sequence using flanking markers as anchors, where they can be searched as normal genomic features. It enables the analysis of QTLs in colinear regions in other cereals and allows researchers, to distinguish sequences and QTLs correlated with related traits or phenotypes across a broad variety of plant species. Researchers can identify whether a QTL colocalizes with other QTLs and can integrate data from different studies to enhance the accuracy of a QTL location. It provides plant biologists and geneticists a way to investigate the interaction between genomic variation and diverse modes of phenotypic variation [28].

To improve QTL-based candidate gene recognition and gene expression study, PlantQTL-GE database has developed [29]. This database contains information about chromosomes and details on gene expression in microarray data and ESTs and genetic markers of *Arabidopsis thaliana* and *Oryza sativa*. Another database, namely the Institute for Genomic Research (TIGR) database, contains DNA, RNA, and protein sequence of plant, human, and microbes. This database consists of repetitive DNA sequence of 12 plant species, namely *Arabidopsis*, *Hordeum*, *Brassica*, *Glycine*, *Lotus*, *Oryza*, *Triticum*, *Lycopersicon*, *Medicago*, *Sorghum*, *Solanum*, and *Zea*. The repeated sequences within each database are further classified into subcategories namely groups and subclasses dependent on sequence and structure similarities [72]. Sequence similarity can also be checked for the downloaded files and are accessible from different sources [73].

3.4 Single Nucleotide Polymorphism (SNP) Database

SNPs, which comprise the most abundant type of genetic variation, are used in genetic studies [74]. SNPs play key role while studying gene mapping, diversity, and evolutionary variation among populations. They also play an important role in designing markers to identify the genetic variation occurring in the contemporary genome and that it has been transferred from the wild type. While other forms of variation including indels, microsatellites, variation of copy number, and epigenetic markers remain important to consider and can cause disease, in genetic study, SNPs are largely the simplest to determine, and are the most useful and commonly used markers. One of the most important SNP databases is dbSNP (<http://www.ncbi.nlm.nih.gov/SNP>), in which SNP identifiers (SNPids) or rsIDs are used to identify SNPs. This database arranges the nucleotide sequence based on their variation in sequence to differentiate between two sequences and search the polymorphism through nucleotide substitution, insertion, deletions, and nucleotide repeats [75].

By the sequencing project of 3000 rice genome, it has been possible to identify SNPs through the alignment of gene of interest with the reference genome in rice [76, 77]. This result provides the information of SNPs that are synonymous or nonsynonymous to the particular genome. The 3000-rice genome project has been described in detail in Chap. 5. Rice Variation Map database is developed for genomic variation study [30]. Yan et al. [78] developed a database of SNP for rice, namely IC4R, that provides SNP information of 18 billion reads. For commercial rice varieties, they provide SNP barcode to easily assess the SNPs using seven machine learning-based methods that are DT, KNN, NB, ANN, RF, LR-M, and LR-O algorithms in the Python Sklearn Library (<https://scikit-learn.org/stable/>). To identify the genetic variation during various stress conditions, Rice Stress-Resistant SNP (RSRS) database has been developed [35]. Recently, Alexandrov and their team developed the SNP-Seek Database using 60 billion SNP reads, which provide SNP information by discarding all the indels in any of the genomic region to find out the structural genetic variations [33, 79].

3.5 Transcriptome Database

With the introduction of high-throughput technology like next-generation sequencing platforms, it is possible to identify the gene expression at the molecular level. Data on gene expression have proven invaluable to genome annotation programs [80]. The functional genomic study helps the researcher to find out the function of rice genes that control various traits. To study these gene functions, several databases are developed through bioinformatic tool where all the repository information of genes with specific trait of interest is available. For instance, TENOR (Transcriptome ENcyclopedia of Rice) database comprises of large-scale mRNA sequencing (mRNA-Seq) data extracted from rice. This database includes details on rice transcriptomes, such as transcript structures and expression profiles, as well as data on coexpression and data on cis-regulatory elements for each gene in 1 kb

upstream regions. Since specifying the ability of plants to adapt to different growing conditions is a key issue in plant science, understanding the regulatory networks of genes associated with environmental changes is of great interest [81]. The team developed the database by using mRNA-Seq data under 10 abiotic stress condition such as high, low, and extremely low cadmium; high and low phosphate; high salinity; drought; osmotic; cold; and flood and under two conditions of plant hormone treatments (jasmonic acid and ABA). Earlier, Oono et al. used the TENOR database to identify the transporters for cadmium tolerance to study the gene expression under cadmium concentration [40]. TENOR offers three kinds of search systems. First, under various conditions, users can provide one or more transcript IDs for scanning functional annotations and patterns of expression. In addition, users can search genes for functional annotation keywords as well. This shows both partially and completely matched results. Second, the “GBrowse” genome browser helps users to check for transcripts with a transcript ID or genomic coordinate. In this search, the user can get all the information regarding the novel transcript structures viz. both annotated and unannotated with their characteristics. In the third search, the user can search for a collection of plant stress hormone-responsive genes through reactive expression patterns, defining the path of transition (suppression or induction), experimental circumstances, sampled tissues, and time points, in addition to fold change (FC) thresholds and statistical importance of changes in expression level.

Rice Expression Database (RED) is interactive rice gene expression profile database completely extracted from RNA-Seq data. RED provides a detailed list of 284 high-quality RNA-Seq results, includes a wide range of gene expression models, and encompasses a wide range of plant development stages. RED consists of a collection of genes unique to housekeeping and tissue and creates coexpression networks dynamically for gene(s) of interest. RiceArrayNet is another database that provides information in terms of correlation coefficients on coexpression between genes in rice. The correlation coefficient shows the coexpression pattern of genes in the rice genome [37]. Lee and the team developed this database that provides the correlation data in three different ways: First, gene coexpression is visualized in the form of cluster or network; second, the coexpression is visualized in scatter plot; and third, the gene coexpression is visualized in the histogram. Another recent Internet-based database for plant gene analysis is the “PLAnt co-EXpression database (PLANEX).” It includes freely accessible GeneChip details collected from the “Gene Expression Omnibus (GEO)” of NCBI. PLANEX is a database for genome-wide coexpression, which helps genes from a wide range of experimental designs to be functionally identified [42]. PLANEX describes “Pearson’s correlation coefficients (PCCs; r-values)” distributed for a specified microarray platform, contributing to a single organism from a gene of interest. The PLANEX database offers a correlation database, a cluster network, and an analysis of enrichment test results for eight plant organisms such as *Arabidopsis thaliana*, *Triticum aestivum*, *Glycine max*, *Vitis vinifera*, *Hordeum vulgare*, *Solanum lycopersicum*, *Oryza sativa*, and *Zea mays*. The cluster network of coexpressed genes is developed by PLANEX, which is calculated using the k-mean method. Genevestigator is another advanced

Web-based framework developed to use modern data mining concepts and groundbreaking algorithms to conduct molecular expression meta-analysis. This database is focused on the systematic, large-scale combination of normalized and quality-controlled expression data with ontology-based experimental background variables such as anatomy, development, perturbation, or genetic background. This large-scale combination of data and meta-data gives new insight into transcriptomes' spatio-temporal response design and helps users to answer concerns that cannot be answered by evaluating a single experiment [50]. Other important gene expression databases are discussed in Table 3.1.

3.6 Protein Database

The protein database provides all the information of the gene that will encode protein. For functional genomics, proteome study related to genome sequence data is useful. These genome sequences help researcher to identify the genes that are expressed in protein developed through alternate splicing and post-transcriptional modification. Rice proteome study is possible through leaves, embryos, endosperms, roots, branches, shoots, and calluses, which provide the detailed mutation of gene through various environmental conditions. Many databases have been developed for proteome study in rice. For instance, OryzaPG-DB, a shotgun proteogenomic-dependent rice proteome database, integrates the genomic features of data from experimental shotgun proteomics. This version of the database was developed from the results of 27 nanoLC-MS/MS runs on a mass spectrometer of hybrid ion trap-orbitrap, providing high precision for the study of tryptic digests from undifferentiated cultured rice cells. Through searching the product ion spectra against the Michigan State University, protein, cDNA, transcript, and gene databases, peptides were detected and mapped to the rice genome. These peptides were occupied by approximately 3200 genes, and 40 of them incorporated novel genomic characteristics. The users may search, import, or browse the chromosome, gene, protein, cDNA, or transcript database and download the modified annotations in standard GFF3 format with PNG format visualization.

The "Mitochondrial Protein Import Components (MPIC)" database contains searchable details on the plant and non-plant mitochondrial protein import apparatus. An *in silico* study was performed to compare the mitochondrial protein import apparatus of 24 organisms representing different lines of *Saccharomyces cerevisiae*, algae to *Homo sapiens* and plants, including *Oryza sativa*, *Arabidopsis thaliana*, and other newly sequenced plant species. In the MPIC DB, each of these species has been thoroughly scanned and manually constructed for analysis. The database provides a user-friendly graphical map, enabling users to find their appropriate import component. The MPIC DB offers a robust database to promote thorough investigation of mitochondrial protein import machinery and to identify conservation and divergence patterns that might have been skipped [64].

Manually Curated Database of Rice Proteins (MCDRP) is a database for rice protein mainly focused on the reported experimental data [68]. Another database,

Table 3.1 Description of various database used in rice research

| Database type | Database Name | Function | Reference |
|---------------|--|--|-----------|
| Genomic | Gramene | Provides tools to perform powerful comparative analyses through a wide range of species of plants. | [3] |
| | Phytozome | Provides an overview of the evolution of each plant gene at the sequence level, gene function, gene family, and organization of the genome | [9] |
| | Ensemble plant | Provides information on visualizing, mining, and analyzing plant genome | [10] |
| | RAP-DB | A hub for rice genomics. | [11] |
| | RGKbase | Evolutionary biology and comparative genomics | [12] |
| | GreenPhylDB | Comparative and functional genomics in plants | [13] |
| | PlantGDB | Browsing features of genomes that combine all relevant EST and cDNA data for emerging gene models | [14] |
| | PGDBJ | Provides a method for plant genome annotation. | [15] |
| | MIPS PlantsDB | Provides central data integration network for annotation of structural and functional gene | [16] |
| | OryGenesDB | Provides positional information of flanking sequence tags (FSTs) of insertional mutagens | [17] |
| | PGSB | Provides keyword search for gene identifier and functional gene description of plant | [18] |
| | RMD | For new gene identification and regulatory element | [19] |
| | Rice Phylogenomics database | Functional study of rice genome | [20] |
| | RiceGAAS | For executing reliable analysis from sequence of rice genome | [21] |
| | RAD | Provides the new information on manual annotation and a systematic structural study of rice genome | [22] |
| RGKbase | For genome assembly and annotation | [12] | |
| Ricebase | Phenotypic interpretation, gene annotation | [23] | |

(continued)

Table 3.1 (continued)

| Database type | Database Name | Function | Reference |
|---------------|--|---|-----------|
| | OryzaGenome | Provides the variant information with hyperlinks to Oryzabase | [24] |
| | DroughtDB | Facilitates the recognition, study, and characterization of genes associated with the resistance of drought stress | [25] |
| | RiceMetaSys | Information on genes for salt and drought stress | [26] |
| QTL | Q-TARO | Provides all the information of rice QTLs | [27] |
| | Gramene | Provides information on genetic and physical position of QTL | [28] |
| | PlantQTL-GE | QTL-based gene identification | [29] |
| SNP | Rice variation map (RiceVarMap) | Provides comprehensive information of SNPs and indels | [30] |
| | PmiRKB (Plant miRNA Knowledge Base) | Investigate the metabolism of miRNA precursors and gene regulation | [31] |
| | RICD (Rice Indica cDNA database) | Provides cDNA resource with comprehensive information for comparative genomics. And functional analysis of <i>indica</i> subspecies | [32] |
| | Rice SNP-seek database | Quick retrieving of SNP alleles for all varieties in a given genome region | [33] |
| | HapRice | Identification of polymorphic SNPs between any two accessions to rice | [34] |
| | Rice stress-resistant SNP database | Characterization of SNPs for various stresses | [35] |
| Expression | Rice expression database (RED) | Information on important biological processes and mechanism | [36] |
| | RiceArrayNet | Provides information in form of correlation coefficients on coexpression between genes | [37] |
| | RiceXPro | Provides information on gene expression of important agronomic trait in rice | [38] |
| | Rice oligonucleotide Array database (ROAD) | Provides detailed profiles of gene expression for all rice genes | [39] |
| | TENOR | Expression study in various environmental stress | [40] |
| | OryzaExpress | Gene expression network and omic annotation using large-scale microarray data | [41] |

(continued)

Table 3.1 (continued)

| Database type | Database Name | Function | Reference |
|---------------|--|--|-------------------|
| | PLANEX | For functional identification of genes | [42] |
| | CoP | Identifies the association of coexpressed genes with metabolic pathways | [43] |
| | PlaNet | Predicts the identity of functional homologs. | [44] |
| | ExPath | Pathway enrichment study | [45] |
| | Plant microRNA database (PMRD) | To study the function and their target genes of miRNAs | [46] |
| | ATTED-II | For functional identification and study of regulatory relationships among genes. | [47] |
| | IsomiR Bank database | Provides target prediction and enrichment analysis to evaluate the effects of isomiRs | Zhang et al. [48] |
| | Rice kinase database (RKD) | Expression analysis for rice kinases using NGS data | [49] |
| | Genevestigator | Gene expression analysis using microarray data | [50] |
| | Bio-analytic resource (BAR) | Transcriptome and protein–protein interaction study | [51] |
| | PlantTFDB | Provides information of transcription factors | [52] |
| | CSRDB (Cereal Small RNA Database) | A tool of smRNA to identify single mature transcript for the recognition of specific target sites | [53] |
| | PtRFdb | Characterization and validation of plant tRFs | [54] |
| | RiceSTIFDB (stress-responsive transcription factor database) | Accelerates study of rice TFs in functional genomics and identifies the regulatory processes underlying abiotic stress responses | [55] |
| | GRASSIUS (grass regulatory information services) | Integrates TFs and gene promoter information | [56] |
| | Plant MPSS (massively parallel signature sequencing) databases | Measure the degree of expression of the majority of genes under specified conditions | [57] |
| | PlantDHS (plant DNaseI hypersensitive site database) | Detection of distant cis-regulatory DNA elements (CREs) that are located away from promoters | [58] |

(continued)

Table 3.1 (continued)

| Database type | Database Name | Function | Reference |
|---------------|--|---|---|
| | PlantAPA (alternative polyadenylation) | Alternative modulation of polyadenylation and gene expression in plants | [59] |
| | RiceFREND | Provides a platform for gene function prediction | [60] |
| | RiceSRTFDB | Accelerates the study of rice TFs in functional genomics and identifies the regulatory processes underlying responses to abiotic stress | [55] |
| | RECoN | Characterization of abiotic stresses on genome-scale and regulatory mechanism study | [61] |
| | RTFDB | Identification of gene expression related to stress which is generally tissue-specific | [62] |
| Protein | UniProtKB | The measurable properties of each protein, like isoelectric point, molecular weight, and expression and experimentally defined properties like sequences of amino acids | [63] |
| | PDB (protein data Bank) | Collection of protein 3D structure | http://www ww p d b . o r g / |
| | MPIC (mitochondrial protein import components) database | Comparative study of machinery for importing mitochondrial protein from plants and animals | [64] |
| | OryzaPG-DB (Oryza ProteoGenomic Database) | Provides models of peptide-based expression along with the corresponding genomic origin, including novelty annotation for each peptide | [65] |
| | DIPOS (database of Interacting Proteins in <i>Oryza sativa</i>) | Provides comprehensive information of rice proteins | [66] |
| | PRIN (predicted Rice Interactome network) | Provides novel insights into gene networks and functional coordination of genes | [67] |
| | MCDRP (manually Curated database of rice proteins) | Annotation of rice protein | [68] |

(continued)

Table 3.1 (continued)

| Database type | Database Name | Function | Reference |
|----------------------------------|---------------|---|---|
| <i>Gene ontology and pathway</i> | Planteome | It provides a set of reference and species-specific ontologies, gene, and phenotype annotation for plants. | [69] |
| | RiceCyc | Provides information about the substrates, enzymes, metabolites, reactions, and pathways of primary as well as intermediary metabolism within rice. | https://bigd.big.ac.cn/databasecommons/database/id/3305 |
| | KEGG rice | Provides information about rice-associated pathway | https://www.genome.jp/kegg-bin/show_organism?org=dosa |
| | IntAct rice | Provides a publicly available, open-source information platform, and protein interaction data analysis tools. | https://www.ebi.ac.uk/intact/home.xhtml |

namely the “Database of Interacting Proteins in *Oryza sativa* (DIPOS),” offers detailed knowledge on interacting proteins in rice, where two statistical approaches are used to model interactions, i.e., interologs and domain-based methods. Of 27,746 proteins, DIPOS comprises 14,614,067 pair-related associations, covering around 41 percent of the entire proteome of *Oryza sativa* [66].

3.7 Gene Ontology and Pathway Database

To broaden our understanding of biological processes in plants and to explain how biological functions develop, it is important to recognize certain diversified pathway. An application of reference and species-specific ontologies for plants and annotations to genes and phenotypes are provided by the Planteome project (<http://www.planteome.org>). Ontologies serve as common standards for the semantic integration of data from plant genomics, phenomics, and genetics from a large and growing dataset. There Plant Ontology, Plant Trait Ontology, and Plant Experimental Conditions Ontology, developed by the Planteome project, together with Gene Ontology, Biological Interest Chemical Entities, Phenotype and Attribute Ontology, and others, are the reference ontologies [69]. The platform also offers access to species-specific crop ontologies established around the world by different plant breeding and research communities. Out of 95 plant taxa, annotated with reference ontology terms, developers offer integrated data on plant traits, phenotypes, and gene function and expression. In order to facilitate community engagement, the Planteome project has developed a plant gene annotation platform, Planteome Noctua. All Planteome ontologies are freely accessible and are managed for sharing and monitoring revisions and new queries at the Planteome GitHub platform (<https://>

github.com/Planteome). From the ontology browser (<http://browser.planteome.org/amigo>) and our data archive, the annotated data are readily available.

RiceCyc is another directory of rice metabolic pathway network [82]. It is a glimpse of the main and intermediate metabolism of substrates, enzymes, metabolites, reactions, and pathways in rice. Version 3.3 of RiceCyc contains 316 pathways and 6643 peptide-coding genes mapped to 2103 enzyme-catalyzed and 87 transport reactions regulated by protein. Annotations given by the KEGG and Gramene databases enriched the original functional annotations of rice genes with InterPro, Enzyme Commission (EC) numbers, MetaCyc, and Gene Ontology. Employing the Pathologic module of Pathway Software, pathway inferences and network diagrams were first predicted based on MetaCyc reference networks and plant pathways from the Plant Metabolic Network. This was enriched by manually inserting metabolic pathways and explicitly reported gene functions for rice. From pathway diagrams to the relevant genes, metabolites, and chemical structures, the RiceCyc database is hierarchically browsable. Users may also upload transcriptomic, proteomic, and metabolomic data to visualize expression trends in a simulated cell using the OMIC Viewer integrated application. RiceCyc enables comparative pathway research, coupled with additional species-specific pathway databases housed in the Gramene project [82]. Another database, namely IntAct, offers a publicly distributed, open-source information structure and molecular interaction data review. All interactions are produced through curation of literature or direct contributions from users and are publicly accessible (<https://www.ebi.ac.uk/intact/home.xhtml>).

3.8 Conclusion and Future Perspectives

In conclusion, we attempted to catalog numerous Web data services and tools available for rice research. Some of them, although few are recent and small-scale repositories, are well known and commonly utilized. It is obvious with the growing number of databases that there is an overwhelming amount of data accessible on the site, connected to almost any area of rice science. However, it has not been effectively investigated despite possessing such a vast amount of diverse data, since many biology researchers or prospective consumers are unfamiliar with all the possible tools to find and interpret the data [83]. Different databases often have different data exchange formats and protocols, which makes it difficult to integrate them into one place. In an ideal situation, a single platform should be available for all databases in a single domain of interest, where a user can use APIs and ontologies to search all the respective databases with a single query and compare the results; e.g., Araport [84] is one such initiative. In order to improve the legitimacy of their data, certain databases are now merging connections to other databases of similar types of data, which is the first move in offering a single forum. This maximizes the utilization of usable data in current assets, which may aid in the prevention of duplication. It provides small databases with greater coverage and may offer a broader image collectively, as small databases typically concentrate on one particular element and provide comprehensive details.

Conflicts of Interest None.

References

1. Wang J, Xu H, Li N, Fan F, Wang L, Zhu Y, et al. Artificial Selection of Gn1a Plays an Important role in Improving Rice Yields Across Different Ecological Regions. *Rice* [Internet]. 2015;8(1):37. Dec [cited 2019 Feb 7]. Available from: <http://www.thericejournal.com/content/8/1/37>
2. Zhang Q-J, Zhu T, Xia E-H, Shi C, Liu Y-L, Zhang Y, et al. Rapid diversification of five *Oryza* AA genomes associated with rice adaptation. *Proc Natl Acad Sci*. 2014 Nov 18;111(46):E4954–62.
3. Ware DH, Jaiswal P, Ni J, Yap IV, Pan X, Clark KY, et al. Gramene, a tool for grass genomics. *Plant Physiol*. 2002 Dec 1;130(4):1606–13.
4. Youens-Clark K, Buckler E, Casstevens T, Chen C, DeClerck G, Derwent P, et al. Gramene database in 2010: updates and extensions. *Nucleic Acids Res*. 2011 Jan 1;39(suppl_1):D1085–94.
5. Jaiswal P, Ni J, Yap I, Ware D, Spooner W, Youens-Clark K, et al. Gramene: a bird's eye view of cereal genomes. *Nucleic Acids Res*. 2006 Jan 1;34(suppl_1):D717–23.
6. Sasaki T. The map-based sequence of the rice genome. *Nature*. 2005 Aug;436(7052):793–800.
7. Project RA. The Rice Annotation Project Database (RAP-DB): 2008 update. *Nucleic Acids Res*. 2008 Jan 1;36(suppl_1):D1028–33.
8. Jiang J, Xing F, Zeng X, Zou Q, Ricyer DB. A database for collecting Rice yield-related genes with biological analysis. *Int J Biol Sci*. 2018 May 22;14(8):965–70.
9. Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, et al. Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res*. 2012 Jan;40(Database issue):D1178–86.
10. Bolser D, Staines DM, Pritchard E, Kersey P. Ensembl plants: integrating tools for visualizing, mining, and analyzing plant genomics data. *Methods Mol Biol Clifton NJ*. 2016;1374:115–40.
11. Ohyanagi H, Tanaka T, Sakai H, Shigemoto Y, Yamaguchi K, Habara T, et al. The Rice Annotation Project Database (RAP-DB): hub for *Oryza sativa* ssp. *japonica* genome information. *Nucleic Acids Res*. 2006 Jan 1;34(suppl_1):D741–4.
12. Wang D, Xia Y, Li X, Hou L, Yu J. The Rice genome knowledgebase (RGKbase): an annotation database for rice comparative genomics and evolutionary biology. *Nucleic Acids Res*. 2013 Jan;41(Database issue):D1199–205.
13. Conte MG, Gaillard S, Lanau N, Rouard M, Périn C. GreenPhylDB: a database for plant comparative genomics. *Nucleic Acids Res*. 2008 Jan;36(Database issue):D991–8.
14. Dong Q, Schlueter SD, Brendel V. PlantGDB, plant genome database and analysis tools. *Nucleic Acids Res*. 2004 Jan 1;32(Database issue):D354–9.
15. Asamizu E, Ichihara H, Nakaya A, Nakamura Y, Hirakawa H, Ishii T, et al. Plant genome DataBase Japan (PGDBj): a portal website for the integration of plant genome-related databases. *Plant Cell Physiol*. 2014 Jan 1;55(1):e8.
16. Nussbaumer T, Martis MM, Roessner SK, Pfeifer M, Bader KC, Sharma S, et al. MIPS PlantsDB: a database framework for comparative plant genome research. *Nucleic Acids Res*. 2013 Jan 1;41(D1):D1144–51.
17. Droc G, Ruiz M, Larmande P, Pereira A, Piffanelli P, Morel JB, et al. OryGenesDB: a database for rice reverse genetics. *Nucleic Acids Res*. 2006 Jan 1;34(Database issue):D736–40.
18. Spannagl M, Nussbaumer T, Bader KC, Martis MM, Seidel M, Kugler KG, et al. PGSB PlantsDB: updates to the database framework for comparative plant genome research. *Nucleic Acids Res*. 2016 Jan 4;44(D1):D1141–7.
19. Zhang J, Li C, Wu C, Xiong L, Chen G, Zhang Q, et al. RMD: a rice mutant database for functional analysis of the rice genome. *Nucleic Acids Res*. 2006 Jan 1;34(suppl_1):D745–8.

20. Jung K-H, Cao P, Sharma R, Jain R, Ronald PC. Phylogenomics databases for facilitating functional genomics in rice. *Rice*. 2015 Jul 30;8(1):26.
21. Sakata K, Nagamura Y, Numa H, Antonio BA, Nagasaki H, Itonuma A, et al. RiceGAAS: an automated annotation system and database for rice genome sequence. *Nucleic Acids Res*. 2002 Jan 1;30(1):98–102.
22. Ito Y, Arikawa K, Antonio BA, Ohta I, Naito S, Mukai Y, et al. Rice annotation database (RAD): a contig-oriented database for map-based rice genomics. *Nucleic Acids Res*. 2005 Jan 1;33(Database issue):D651–5.
23. Edwards JD, Baldo AM, Mueller LA. Ricebase: a breeding and genetics platform for rice, integrating individual molecular markers, pedigrees and whole-genome-based data. *Database [Internet]*. 2016 Jan 1;2016(baw107). [cited 2021 Jan 8] <https://doi.org/10.1093/database/baw107>.
24. Ohyanagi H, Ebata T, Huang X, Gong H, Fujita M, Mochizuki T, et al. OryzaGenome: genome diversity database of wild *Oryza* species. *Plant Cell Physiol*. 2016 Jan 1;57(1):e1.
25. Alter S, Bader KC, Spannagl M, Wang Y, Bauer E, Schön C-C, et al. DroughtDB: an expert-curated compilation of plant drought stress genes and their homologs in nine species. *Database [Internet]*. 2015 Jan 1;2015(bav046). [cited 2021 Jan 8] <https://doi.org/10.1093/database/bav046>.
26. Sandhu M, Sureshkumar V, Prakash C, Dixit R, Solanke AU, Sharma TR, et al. RiceMetaSys for salt and drought stress responsive genes in rice: a web interface for crop improvement. *BMC Bioinformatics*. 2017 Sep 30;18(1):432.
27. Yonemaru J, Yamamoto T, Fukuoka S, Uga Y, Hori K, Yano M. Q-TARO: QTL annotation Rice online database. *Rice*. 2010 Sep;3(2):194–203.
28. Ni J, Pujar A, Youens-Clark K, Yap I, Jaiswal P, Teclé I, et al. Gramene QTL database: development, content and applications. *Database [Internet]*. 2009 Jan 1;2009(bap005). [cited 2021 Jan 5] <https://doi.org/10.1093/database/bap005>.
29. Zeng H, Luo L, Zhang W, Zhou J, Li Z, Liu H, et al. PlantQTL-GE: a database system for identifying candidate genes in rice and *Arabidopsis* by gene expression and QTL information. *Nucleic Acids Res*. 2007 Jan 1;35(suppl_1):D879–82.
30. Zhao H, Yao W, Ouyang Y, Yang W, Wang G, Lian X, et al. RiceVarMap: a comprehensive database of rice genomic variations. *Nucleic Acids Res*. 2015 Jan 28;43(D1):D1018–22.
31. Meng Y, Gou L, Chen D, Mao C, Jin Y, Wu P, et al. PmiRKB: a plant microRNA knowledge base. *Nucleic Acids Res*. 2011 Jan 1;39(suppl_1):D181–7.
32. Lu T, Huang X, Zhu C, Huang T, Zhao Q, Xie K, et al. RICD: a rice indica cDNA database resource for rice functional genomics. *BMC Plant Biol*. 2008 Nov 26;8(1):118.
33. Alexandrov N, Tai S, Wang W, Mansueto L, Palis K, Fuentes RR, et al. SNP-seek database of SNPs derived from 3000 rice genomes. *Nucleic Acids Res*. 2015 Jan 28;43(D1):D1023–7.
34. Yonemaru J, Ebana K, Yano M. HapRice, an SNP haplotype database and a web tool for Rice. *Plant Cell Physiol*. 2014 Jan 1;55(1):e9.
35. Tareke Woldegiorgis S, Wang S, He Y, Xu Z, Chen L, Tao H, et al. Rice stress-resistant SNP database. *Rice*. 2019 Dec 23;12(1):97.
36. Xia L, Zou D, Sang J, Xu X, Yin H, Li M, et al. Rice expression database (RED): an integrated RNA-Seq-derived gene expression database for rice. *J Genet Genomics Yi Chuan Xue Bao*. 2017 May 20;44(5):235–41.
37. Lee T-H, Kim Y-K, Pham TTM, Song SI, Kim J-K, Kang KY, et al. RiceArrayNet: a database for correlating gene expression from transcriptome profiling, and its application to the analysis of Coexpressed genes in Rice. *Plant Physiol*. 2009 Sep;151(1):16–33.
38. Sato Y, Antonio BA, Namiki N, Takehisa H, Minami H, Kamatsuki K, et al. RiceXPro: a platform for monitoring gene expression in japonica rice grown under natural field conditions. *Nucleic Acids Res*. 2011 Jan;39(Database issue):D1141–8.
39. Cao P, Jung K-H, Choi D, Hwang D, Zhu J, Ronald PC. The Rice oligonucleotide Array database: an atlas of rice gene expression. *Rice*. 2012 Jul 19;5(1):17.

40. Oono Y, Yazawa T, Kanamori H, Sasaki H, Mori S, Handa H, et al. Genome-Wide Transcriptome Analysis of Cadmium Stress in Rice [Internet], vol. 2016. Hindawi: BioMed Research International; 2016. p. e9739505. [cited 2020 Dec 28]. Available from: <https://www.hindawi.com/journals/bmri/2016/9739505/>
41. Hamada K, Hongo K, Suwabe K, Shimizu A, Nagayama T, Abe R, et al. OryzaExpress: an integrated database of gene expression networks and omics annotations in rice. *Plant Cell Physiol.* 2011 Feb;52(2):220–9.
42. Yim WC, Yu Y, Song K, Jang CS, Lee B-M. PLANEX: the plant co-expression database. *BMC Plant Biol.* 2013 May 20;13(1):83.
43. Ogata Y, Suzuki H, Sakurai N, Shibata D. CoP: a database for characterizing co-expressed gene modules with biological information in plants. *Bioinformatics.* 2010 May 1;26(9):1267–8.
44. Mutwil M, Klie S, Tohge T, Giorgi FM, Wilkins O, Campbell MM, et al. PlaNet: combined sequence and expression comparisons across plant networks derived from seven species. *Plant Cell.* 2011 Mar 1;23(3):895–910.
45. Chien C-H, Chow C-N, Wu N-Y, Chiang-Hsieh Y-F, Hou P-F, Chang W-C. EXPath: a database of comparative expression analysis inferring metabolic pathways for plants. *BMC Genomics.* 2015 Jan 21;16(2):S6.
46. Zhang Z, Yu J, Li D, Zhang Z, Liu F, Zhou X, et al. PMRD: plant microRNA database. *Nucleic Acids Res.* 2010 Jan;38(Database issue):D806–13.
47. Obayashi T, Nishida K, Kasahara K, Kinoshita K. ATTED-II updates: condition-specific gene Coexpression to extend Coexpression analyses and applications to a broad range of flowering plants. *Plant Cell Physiol.* 2011 Feb;52(2):213–9.
48. Zhang Y, Zang Q, Xu B, Zheng W, Ban R, Zhang H, et al. Isomi R Bank: a research resource for tracking IsomiRs. *Bioinformatics.* 2016 Jul 1;32(13):2069–71.
49. Chandran AKN, Yoo Y-H, Cao P, Sharma R, Sharma M, Dardick C, et al. Updated Rice Kinase Database RKD 2.0: enabling transcriptome and functional analysis of rice kinase genes. *Rice.* 2016 Aug 19;9(1):40.
50. Zimmermann P, Laule O, Schmitz J, Hruz T, Bleuler S, Gruissem W. Genevestigator transcriptome meta-analysis and biomarker search using Rice and barley gene expression databases. *Mol Plant.* 2008 Sep;1(5):851–7.
51. Waese J, Provart NJ. The bio-analytic resource: data visualization and analytic tools for multiple levels of plant biology. *Curr Plant Biol.* 2016 Nov 1;7–8:2–5.
52. Guo A-Y, Chen X, Gao G, Zhang H, Zhu Q-H, Liu X-C, et al. PlantTFDB: a comprehensive plant transcription factor database. *Nucleic Acids Res.* 2008 Jan 1;36(suppl_1):D966–9.
53. Johnson C, Bowman L, Adai AT, Vance V, Sundaresan V. CSRDB: a small RNA integrated database and browser resource for cereals. *Nucleic Acids Res.* 2007 Jan;35(Database issue):D829–33.
54. Gupta N, Singh A, Zahra S, Kumar S. PtrFdb: a database for plant transfer RNA-derived fragments. *Database [Internet].* 2018 Jan 1, 2018;(bay063). [cited 2021 Jan 11] <https://doi.org/10.1093/database/bay063>.
55. Priya P, Jain M. RiceSRTFDB: a database of rice transcription factors containing comprehensive expression, cis-regulatory element and mutant information to facilitate gene function analysis. *Database J Biol Databases Curation.* 2013;2013(bat027)
56. Yilmaz A, Nishiyama MY, Fuentes BG, Souza GM, Janies D, Gray J, et al. GRASSIUS: a platform for comparative regulatory genomics across the grasses. *Plant Physiol.* 2009 Jan;149(1):171–80.
57. Nakano M, Nobuta K, Vemaraju K, Tej SS, Skogen JW, Meyers BC. Plant MPSS databases: signature-based transcriptional resources for analyses of mRNA and small RNA. *Nucleic Acids Res.* 2006 Jan 1;34(Database issue):D731–5.
58. Zhang T, Marand AP, Jiang J. PlantDHS: a database for DNase I hypersensitive sites in plants. *Nucleic Acids Res.* 2016 Jan 4;44(D1):D1148–53.

59. Wu X, Zhang Y, Li QQ. PlantAPA: A Portal for Visualization and Analysis of Alternative Polyadenylation in Plants. *Front Plant Sci* [Internet]. 2016;7. [cited 2021 Jan 11]. Available from: <https://www.frontiersin.org/articles/10.3389/fpls.2016.00889/full>.
60. Sato Y, Namiki N, Takehisa H, Kamatsuki K, Minami H, Ikawa H, et al. RiceFRIEND: a platform for retrieving coexpressed gene networks in rice. *Nucleic Acids Res*. 2013 Jan;41(Database issue):D1214–21.
61. Krishnan A, Gupta C, MMR A, Pereira A. RECoN: Rice Environment Coexpression Network for Systems Level Analysis of Abiotic-Stress Response. *Front Plant Sci* [Internet]. 2017;8. [cited 2021 Jan 8]. Available from: <https://www.frontiersin.org/articles/10.3389/fpls.2017.01640/full>.
62. Chandran AKN, Moon S, Yoo Y-H, Gho Y-S, Cao P, Sharma R, et al. A web-based tool for the prediction of rice transcription factor function. *Database* [Internet]. 2019 Jan 1;2019(baz061). . [cited 2021 Jan 26] doi:<https://doi.org/10.1093/database/baz061>.
63. Komatsu S. Rice proteome database based on two-dimensional polyacrylamide gel electrophoresis: its status in 2003. *Nucleic Acids Res*. 2004 Jan 1;32(90001):388D–392.
64. Murcha MW, Narsai R, Devenish J, Kubiszewski-Jakubiak S, Whelan J. MPIC: a mitochondrial protein import components database for plant and non-plant species. *Plant Cell Physiol*. 2015 Jan 1;56(1):e10.
65. Helmy M, Tomita M, Ishihama Y. OryzaPG-DB: Rice proteome database based on shotgun Proteogenomics. *BMC Plant Biol*. 2011 Apr 12;11(1):63.
66. Sapkota A, Liu X, Zhao X-M, Cao Y, Liu J, Liu Z-P, et al. DIPOS: database of interacting proteins in *Oryza sativa*. *Mol BioSyst*. 2011 Sep;7(9):2615–21.
67. Zhu P, Gu H, Jiao Y, Huang D, Chen M. Computational identification of protein-protein interactions in Rice based on the predicted Rice Interactome network. *Genomics Proteomics Bioinformatics*. 2011 Oct 1;9(4):128–37.
68. Gour P, Garg P, Jain R, Joseph SV, Tyagi AK, Raghuvanshi S. Manually curated database of rice proteins. *Nucleic Acids Res*. 2014 Jan 1;42(Database issue):D1214–21.
69. Cooper L, Meier A, Laporte M-A, Elser JL, Mungall C, Sinn BT, et al. The Planteome database: an integrated resource for reference ontologies, plant genomics and phenomics. *Nucleic Acids Res* [Internet]. 2017 Nov 23;46(D1). [cited 2021 Jan 11], Available from: <https://www.osti.gov/pages/biblio/1625561-planteome-database-integrated-resource-reference-ontologies-plant-genomics-phenomics>.
70. Sasaki T, Burr B. International Rice genome sequencing Project: the effort to completely sequence the rice genome. *Curr Opin Plant Biol*. 2000 Apr;3(2):138–41.
71. Mao L, Chen M, Chu Q, Jia L, Sultana MH, Wu D, et al. RiceRelativesGD: a genomic database of rice relatives for rice research. *Database* [Internet]. 2019 Jan 1;2019(baz110). [cited 2021 Jan 27] <https://doi.org/10.1093/database/baz110>.
72. Ouyang S, Buell CR. The TIGR Plant Repeat Databases: a collective resource for the identification of repetitive sequences in plants. *Nucleic Acids Res*. 2004 Jan 1;32(suppl_1):D360–3.
73. Lee Y, Tsai J, Sunkara S, Karamycheva S, Pertea G, Sultana R, et al. The TIGR gene indices: clustering and assembling EST and known genes and integration with eukaryotic genomes. *Nucleic Acids Res*. 2005 Jan 1;33(Database Issue):D71–4.
74. Johnson AD. SNP bioinformatics: a comprehensive review of resources. *Circ Cardiovasc Genet*. 2009 Oct;2(5):530–6.
75. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res*. 2001 Jan 1;29(1):308–11.
76. Fuentes RR, Chebotarov D, Duitama J, Smith S, la Hoz JFD, Mohiyuddin M, et al. Structural variants in 3000 rice genomes. *Genome Res*. 2019 May 1;29(5):870–80.
77. Li J-Y, Wang J, Zeigler RS. The 3,000 rice genomes project: new opportunities and challenges for future rice research. *GigaScience*. 2014 May 28;3:8.
78. Yan J, Zou D, Li C, Zhang Z, Song S, Wang X. SR4R: an integrative SNP resource for genomic breeding and population research in Rice. *Genomics Proteomics Bioinformatics*. 2020 Apr 1;18(2):173–85.

79. Mansueto L, Fuentes RR, Borja FN, Detras J, Abriol-Santos JM, Chebotarov D, et al. Rice SNP-seek database update: new SNPs, indels, and queries. *Nucleic Acids Res.* 2017 Jan 4;45 (D1):D1075–81.
80. Yao W, Li G, Yu Y, Ouyang Y. funRiceGenes dataset for comprehensive understanding and application of rice functional genes. *GigaScience* [Internet]. 2018 Jan 1;7(gix119). [cited 2021 Jan 11] <https://doi.org/10.1093/gigascience/gix119>.
81. Kawahara Y, Oono Y, Wakimoto H, Ogata J, Kanamori H, Sasaki H, et al. TENOR: database for comprehensive mRNA-Seq experiments in Rice. *Plant Cell Physiol.* 2016 Jan 1;57(1):e7.
82. Dharmawardhana P, Ren L, Amarasinghe V, Monaco M, Thomason J, Ravenscroft D, et al. A genome scale metabolic network for rice and accompanying analysis of tryptophan, auxin and serotonin biosynthesis regulation under biotic stress. *Rice N Y N.* 2013 May 29;6(1):15.
83. Garg P, Jaiswal P. Databases and bioinformatics tools for rice research. *Curr Plant Biol.* 2016 Nov 1;7–8:39–52.
84. Krishnakumar V, Hanlon MR, Contrino S, Ferlanti ES, Karamycheva S, Kim M, et al. Araport: the Arabidopsis information portal. *Nucleic Acids Res.* 2015 Jan;43(Database issue):D1003–9.