

## Genomic and Bioinformatic Resources for Rice Research

Wenhua Liu, Kabin Xie, and Yinong Yang

### Abstract

During the past decade, there has been an explosive accumulation of “omics” data due to the development of advanced and efficient genome technology. Both japonica and indica rice genomes have been completely sequenced, annotated, and made publicly available. Transcriptome profiling has been conducted on a large scale by microarray and deep-sequencing to chart the atlas of rice gene expression under normal or stressed conditions. In addition, yeast two-hybrid, tandem affinity purification, and other proteomic techniques have accelerated the identification of rice proteins and the elucidation of complex networks of protein–protein interactions. In combination with extensive resources of natural rice variants, deletion and insertional mutants, these genomic, transcriptomic, and proteomic information will greatly facilitate functional characterization of rice genes and networks in a genetic, biological, and evolutionary context. With exquisite bioinformatic tools, many user-friendly databases such as the ones described below have been built to integrate large data sets and allow us to more effectively explore the treasure troves of rice genomic and genetic resources.

**Key words:** Bioinformatics; Genome annotation; Gene expression; Transcriptomics; Proteomics; Protein–protein interaction; Rice mutants

**Table A.1**  
**Useful Web sites for rice research**

Database	Web address	Description
Integrated rice genome annotation		
RAP-DB (Rice Annotation Project Database)	<a href="http://rgp.dna.affrc.go.jp/E/IRGSP/">http://rgp.dna.affrc.go.jp/E/IRGSP/</a>	Genome annotation
RGAP (Rice Genome Annotation Project)	<a href="http://rice.plantbiology.msu.edu/">http://rice.plantbiology.msu.edu/</a>	Genome annotation (1)
RiceGE: Rice Functional Genomic Express Database	<a href="http://signal.salk.edu/cgi-bin/RiceGE">http://signal.salk.edu/cgi-bin/RiceGE</a>	Genome annotation, expression data and mutant resources
GRAMENE	<a href="http://www.gramene.org/">http://www.gramene.org/</a>	Comparative grass genomics
BGI-RIS	<a href="http://rice.genomics.org.cn/rice/">http://rice.genomics.org.cn/rice/</a>	Beijing Genomics Institute (BGI) rice information system (2)
KOME (Knowledge-based Oryza Molecular Biological Encyclopedia)	<a href="http://cdna01.dna.affrc.go.jp/cDNA/">http://cdna01.dna.affrc.go.jp/cDNA/</a>	Rice full-length cDNA collection, gene ontology and expression database (3)
RICD (Rice Indica cDNA Database)	<a href="http://www.ncgr.ac.cn/ricd/">http://www.ncgr.ac.cn/ricd/</a>	Rice (indica) full-length cDNA collection (4)
Rice HaploypeMap Project	<a href="http://www.ncgr.ac.cn/RiceHapMap/">http://www.ncgr.ac.cn/RiceHapMap/</a>	Genetic map-based SNPs (5)
RiceNet	<a href="http://www.functionalnet.org/ricenet">http://www.functionalnet.org/ricenet</a>	Functional association of genes in rice genome (6)
PlaNet	<a href="http://aranet.mpimp-golm.mpg.de">http://aranet.mpimp-golm.mpg.de</a>	Whole-genome coexpression networks for <i>Arabidopsis</i> and six important plant and crop species (7)
DRTF (Database of Rice Transcription Factors)	<a href="http://drtf.cbi.pku.edu.cn/">http://drtf.cbi.pku.edu.cn/</a>	Transcription factor database (8)
PlnTFDB	<a href="http://plntfdb.bio.uni-potsdam.de/v3.0/">http://plntfdb.bio.uni-potsdam.de/v3.0/</a>	Transcription factor database (9)
<b>Genetic resources</b>		
Rice <i>Tos17</i> Insertion Mutant Database	<a href="http://tos.nias.affrc.go.jp/~miyao/pub/tos17/">http://tos.nias.affrc.go.jp/~miyao/pub/tos17/</a>	Rice mutants induced by rice retrotransposon <i>Tos17</i>
Rice Transposon Flanking Sequence Tag Database	<a href="http://sundarlab.ucdavis.edu/rice/blast/blast.html">http://sundarlab.ucdavis.edu/rice/blast/blast.html</a>	Insertion mutants of the maize <i>En/Spm</i> or <i>Ac-Ds</i> transposons in <i> japonica</i> rice cv. Nipponbare
Oryza Tag Line	<a href="http://urgi.versailles.inra.fr/OryzaTagLine/">http://urgi.versailles.inra.fr/OryzaTagLine/</a>	T-DNA insertion mutants
OryGenesDB	<a href="http://orygenesdb.cirad.fr/">http://orygenesdb.cirad.fr/</a>	T-DNA and Ds flanking sequence tags derived from the genomics initiative Genoplante (Oryza Tag Line) and the European consortium Cereal Gene Tags (10)

TRIM (Taiwan Rice Insertional Mutants Database)	<a href="http://trim.sinica.edu.tw/home">http://trim.sinica.edu.tw/home</a>	T-DNA insertion mutants
RISD (Rice T-DNA Insertion Sequence Database)	<a href="http://www.postech.ac.kr/life/pfg/risd/">http://www.postech.ac.kr/life/pfg/risd/</a>	T-DNA insertion and activation tagging lines in <i>japonica</i> rice cv. Dongjin or Hwayoung.
RMD (Rice Mutant Database)	<a href="http://rmd.ncpgr.cn/">http://rmd.ncpgr.cn/</a>	T-DNA insertion mutants generated by an enhancer trap system (11)
National Plant Germplasm System	<a href="http://www.ars-grin.gov/npgs/">http://www.ars-grin.gov/npgs/</a>	Crop germplasm collection
IRRI Genetic Resource Center	<a href="http://beta.irri.org/seeds/">http://beta.irri.org/seeds/</a>	Rice knowledge and collection of natural variants and chemical mutants
International Rice Information System	<a href="http://rice.generationcp.org/germplasm/">http://rice.generationcp.org/germplasm/</a>	A rice portal for global information on genetic resource and rice cultivars
<b>Gene expression</b>		
Genevestigator	<a href="https://www.genevestigator.com/gv/">https://www.genevestigator.com/gv/</a>	Collection of public microarray data and analysis of rice gene expression (12, 13)
CREP (Collection of Rice Gene Expression Profiles)	<a href="http://crep.ncpgr.cn/">http://crep.ncpgr.cn/</a>	A database of the rice transcriptome project
Yale Virtual Center for Cellular Expression Profiling of Rice	<a href="http://bioinformatics.med.yale.edu/riceatlas/">http://bioinformatics.med.yale.edu/riceatlas/</a>	Whole-genome transcriptional profiles of rice cell types
Rice Next-Gen Sequence Database	<a href="http://mpss.udel.edu/rice/">http://mpss.udel.edu/rice/</a>	Rice mRNA and small RNA database generated by next-generation sequencing (14)
ROAD (Rice Oligonucleotide Array Database)	<a href="http://www.ricearray.org">http://www.ricearray.org</a>	A public resource for rice gene expression and coexpression analysis
DIURNAL	<a href="http://diurnal.cgrb.oregonstate.edu/">http://diurnal.cgrb.oregonstate.edu/</a>	Circadian/diurnal gene expression data for Arabidopsis, rice, and poplar genes (15)
RED (Rice Expression Database)	<a href="http://cdna02.dna.affrc.go.jp/RED/">http://cdna02.dna.affrc.go.jp/RED/</a>	Rice gene expression database
OryzaExpress (Gene Expression Database for Rice)	<a href="http://riceball.lab.nig.ac.jp/oryzaexpress/">http://riceball.lab.nig.ac.jp/oryzaexpress/</a>	Rice gene expression database (16)

(continued)

**Table A.1**  
**(continued)**

Database	Web address	Description
CSRDB: Cereal Small RNAs Database	<a href="http://sundarlab.ucdavis.edu/smrnas">http://sundarlab.ucdavis.edu/smrnas</a>	Small RNAs database for cereal crops (17)
RiceXPro	<a href="http://ricexpro.dna.affrc.go.jp/">http://ricexpro.dna.affrc.go.jp/</a>	Collection of microarray data during the entire life cycle of rice plants
RiceFox	<a href="http://ricefox.psc.riken.jp/">http://ricefox.psc.riken.jp/</a>	Overexpression of rice full-length cDNAs in <i>Arabidopsis</i>
<b>Proteomic database</b>		
RPD (Rice Proteome Database)	<a href="http://gene64.dna.affrc.go.jp/RPD">http://gene64.dna.affrc.go.jp/RPD</a>	2D-PAGE and mass-spectrometer analysis of rice proteome (18)
OryzaPG-BD	<a href="http://oryzapg.iab.keio.ac.jp/">http://oryzapg.iab.keio.ac.jp/</a>	Rice proteome database based on shotgun proteomics (19)
IntAct	<a href="http://www.ebi.ac.uk/intact/">http://www.ebi.ac.uk/intact/</a>	General database for protein interactome (20)
RKD (Rice Kinase Database)	<a href="http://phyloinformatics.ucdavis.edu/kinase/">http://phyloinformatics.ucdavis.edu/kinase/</a>	Rice kinase database (21)
<b>Bioinformatic tools</b>		
PLACE (Plant Cis-acting Regulatory DNA Elements)	<a href="http://www.dna.affrc.go.jp/PLACE/">http://www.dna.affrc.go.jp/PLACE/</a>	Analysis of plant promoters and cis-acting DNA elements (22)
WMD3 (Web MicroRNA Designer)	<a href="http://wmd3.weigelworld.org/">http://wmd3.weigelworld.org/</a>	Artificial miRNA design, small RNA targets prediction (23)
Osiris	<a href="http://www.bioinformatics2.wsu.edu/Osiris">http://www.bioinformatics2.wsu.edu/Osiris</a>	Rice promoter analysis (24)
PhosPhAt	<a href="http://phosphat.mpimp-golm.mpg.de/">http://phosphat.mpimp-golm.mpg.de/</a>	Phosphorylation site prediction (25)

## References

1. Ouyang S, Zhu W, Hamilton J, Lin H, Campbell M, Childs K, Thibaud-Nissen F, Malek RL, Lee Y, Zheng L et al (2007) The TIGR rice genome annotation resource: improvements and new features. *Nucleic Acids Res* 35:D883–D887
2. Zhao W, Wang J, He X, Huang X, Jiao Y, Dai M, Wei S, Fu J, Chen Y, Ren X et al (2004) BGI-RIS: an integrated information resource and comparative analysis workbench for rice genomics. *Nucleic Acids Res* 32:D377–D382
3. Kikuchi S, Satoh K, Nagata T, Kawagashira N, Doi K, Kishimoto N, Yazaki J, Ishikawa M, Yamada H, Ooka H et al (2003) Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice. *Science* 301:376–379
4. Lu T, Huang X, Zhu C, Huang T, Zhao Q, Xie K, Xiong L, Zhang Q, Han B (2008) RICD: a rice indica cDNA database resource for rice functional genomics. *BMC Plant Biol* 8:118
5. Huang X, Wei X, Sang T, Zhao Q, Feng Q, Zhao Y, Li C, Zhu C, Lu T, Zhang Z et al (2010) Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat Genet* 42:961–967
6. Lee I, Seo Y, Coltrane D, Hwang S, Oh T, Marcotte E, Ronald P (2011) Genetic dissection of the biotic stress response using a genome-scale gene network for rice. *Proc Natl Acad Sci U S A* 108:18548–18553
7. Mutwil M, Klie S, Tohge T, Giorgi F, Wilkins O, Campbell M, Fernie A, Usadel B, Nikoloski Z, Persson S (2011) PlaNet: combined sequence and expression comparisons across plant networks derived from seven species. *Plant Cell* 23:895–910
8. Gao G, Zhong Y, Guo A, Zhu Q, Tang W, Zheng W, Gu X, Wei L, Luo J (2006) DRTF: a database of rice transcription factors. *Bioinformatics* 22:1286–1287
9. Riano-Pachon DM, Ruzicic S, Dreyer I, Mueller-Roeber B (2007) PlnTFDB: an integrative plant transcription factor database. *BMC Bioinformatics* 8:42
10. Droc G, Ruiz M, Larmande P, Pereira A, Piffanelli P, Morel JB, Dievart A, Courtois B, Guiderdoni E, Perin C (2006) OryGenesDB: a database for rice reverse genetics. *Nucleic Acids Res* 34:D736–D740
11. Zhang J, Li C, Wu C, Xiong L, Chen G, Zhang Q, Wang S (2006) RMD: a rice mutant database for functional analysis of the rice genome. *Nucleic Acids Res* 34:D745–D748
12. Zimmermann P, Laule O, Schmitz J, Hruz T, Bleuler S, Gruissem W (2008) Genevestigator transcriptome meta-analysis and biomarker search using rice and barley gene expression databases. *Mol Plant* 1:851–857
13. Hruz T, Laule O, Szabo G, Wessendorf F, Bleuler S, Oertle L, Widmayer P, Gruissem W, Zimmermann P (2008) Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. *Adv Bioinformatics* 2008:420747
14. Nobuta K, Venu RC, Lu C, Belo A, Vemaraju K, Kulkarni K, Wang W, Pillay M, Green PJ, Wang GL et al (2007) An expression atlas of rice mRNAs and small RNAs. *Nat Biotechnol* 25:473–477
15. Mockler TC, Michael TP, Priest HD, Shen R, Sullivan CM, Givan SA, McEntee C, Kay SA, Chory J (2007) The DIURNAL project: DIURNAL and circadian expression profiling, model-based pattern matching, and promoter analysis. *Cold Spring Harb Symp Quant Biol* 72:353–363
16. Hamada K, Hongo K, Suwabe K, Shimizu A, Nagayama T, Abe R, Kikuchi S, Yamamoto N, Fujii T, Yokoyama K et al (2011) OryzaExpress: an integrated database of gene expression networks and omics annotations in rice. *Plant Cell Physiol* 52:220–229
17. Johnson C, Bowman L, Adai AT, Vance V, Sundaresan V (2007) CSRDB: a small RNA integrated database and browser resource for cereals. *Nucleic Acids Res* 35:D829–D833
18. Komatsu S, Kojima K, Suzuki K, Ozaki K, Higo K (2004) Rice proteome database based on two-dimensional polyacrylamide gel electrophoresis: its status in 2003. *Nucleic Acids Res* 32:D388–D392
19. Helmy M, Tomita M, Ishihama Y (2011) OryzaPG-DB: rice proteome database based on shotgun proteogenomics. *BMC Plant Biol* 11:63
20. Aranda B, Achuthan P, Alam-Faruque Y, Armean I, Bridge A, Derow C, Feuermann M, Ghanbarian AT, Kerrien S, Khadake J et al (2010) The IntAct molecular interaction database in 2010. *Nucleic Acids Res* 38:D525–D531
21. Dardick C, Chen J, Richter T, Ouyang S, Ronald P (2007) The rice kinase database. A phylogenomic database for the rice kinome. *Plant Physiol* 143:579–586

22. Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999) Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Res* 27:297–300
23. Schwab R, Ossowski S, Riester M, Warthmann N, Weigel D (2006) Highly specific gene silencing by artificial microRNAs in *Arabidopsis*. *Plant Cell* 18:1121–1133
24. Morris RT, O'Connor TR, Wyrick JJ (2008) Osiris: an integrated promoter database for *Oryza sativa* L. *Bioinformatics* 24:2915–2917
25. Durek P, Schmidt R, Heazlewood JL, Jones A, MacLean D, Nagel A, Kersten B, Schulze WX (2009) PhosPhAt: the *Arabidopsis thaliana* phosphorylation site database. An update. *Nucleic Acids Res* 38:D828–D834

# INDEX

## A

- ABA. *See* Abscisic acid (ABA)  
Abiotic stress .....59, 69, 71, 109, 185,  
186, 209, 210, 267, 269–278, 293, 295–298, 303  
Abscisic acid (ABA) ..... 273–275, 293,  
295–297, 303, 317  
Activation tagging ..... 58, 61, 62, 64–65  
*Agrobacterium*. *See* Transformation  
Analysis of variance (ANOVA) ..... 70, 217, 252, 267  
ANOVA. *See* Analysis of variance (ANOVA)  
Artificial hybridization .....1–11

## B

- BAC. *See* Bacterial artificial chromosome (BAC)  
Backcross ..... 73, 76, 80, 286  
Bacterial artificial chromosome (BAC)..... 13, 17, 20  
Bioreactor .....312  
Biotic stress..... 59, 109, 185, 186, 267, 274,  
293, 295–297

## C

- Celery juice extract (CJE)..... 40, 43, 45, 49–51  
Chilling stress .....227–240  
Chlorophyll .....62, 210, 212, 217, 221,  
229–232, 275, 318, 319  
Chromatin ..... 14, 317  
Chromosome  
metaphase ..... 13, 14, 17  
pachytene ..... 13, 14, 18–19, 25  
CJE. *See* Celery juice extract (CJE)  
Cold stress ..... 227–229, 271–272, 275  
*Cre-lox* ..... 84  
Crossing .....1, 3–8, 10, 11, 14,  
29, 69, 72, 73, 80, 96–97, 158, 186, 265, 286, 296,  
302, 321  
Cryo-SEM .....244, 246  
Cytogenetic analysis .....13–27

## D

- Database .....48, 58, 59, 61, 63, 86,  
95–96, 109, 131, 136, 137, 153, 162, 176, 177,  
179, 182

- Disease resistance ..... 59, 254, 265, 285–304  
Drought tolerance .....218–219, 272–277, 296, 297, 303

## E

- Ecotilling .....39–54  
EL. *See* Electrolyte leakage (EL)  
Electrolyte leakage (EL)..... 230, 232–233, 238, 239  
Emasculation .....2–4, 6–9, 11  
EMS. *See* Ethyl methanesulfonate (EMS)  
ESTs. *See* Expressed sequence tags (ESTs)  
ET. *See* Ethylene (ET)  
Ethylene (ET) ..... 43, 202, 229, 238,  
271, 293, 295, 296, 301, 303  
Ethyl methanesulfonate (EMS).....31, 40  
Expressed sequence tags (ESTs) ..... 132, 134–136, 146

## F

- FISH. *See* Fluorescence in situ hybridization (FISH)  
Flanking sequence tags (FSTs) ..... 58, 60, 62, 63  
Fluorescence in situ hybridization (FISH) .....13–27  
FSTs. *See* Flanking sequence tags (FSTs)

## G

- Gene regulation  
post-transcriptional..... 109, 119, 272  
Gene silencing ..... 83, 131–148  
Genetic map .....75  
Germination .....4, 10, 17, 24, 32–36,  
74, 79, 162–163, 186–189, 192, 204, 210–212,  
214, 220, 222–224, 231, 237, 244, 245, 260,  
295–296, 312  
GFP. *See* Green fluorescent protein (GFP)  
 $\beta$ -Glucuronidase (GUS).....59–63, 65, 88, 91, 136, 318  
Grain yield (GY)..... 219, 223, 244, 277  
Green fluorescent protein (GFP)..... 59, 60, 136, 320  
GUS. *See*  $\beta$ -Glucuronidase (GUS)  
GY. *See* Grain yield (GY)

## H

- Hormone .....271, 273, 287, 289, 292–296  
HR. *See* Hypersensitive response (HR)  
Hypersensitive response (HR)..... 288, 292, 300

**I**

Inoculation .....249–255, 257–267, 289, 295–297  
 Insect ..... 201–208, 228, 292–293  
 Isoelectric focusing .....158

**J**

JA. *See* Jasmonic acid (JA)  
 Jasmonic acid (JA)..... 185–199, 229,  
 293, 295–297, 300–301, 303

**L**

Linkage map.....70, 72–75, 78, 80, 297  
 Loss-of-function ..... 30, 58, 131, 147

**M**

*Magnaporthe oryzae*, rice blast .....285  
 Map-based cloning.....69–81, 302  
 Mass spectrometry (MS)  
     gas chromatography-mass spectrometry  
         (GC-MS) ..... 188, 190–192,  
         197, 198, 203, 204, 206–207  
     liquid chromatography-mass spectrometry  
         (LC-MS) ..... 160, 176–177,  
         179, 182, 188, 191, 192  
 Meiosis ..... 18, 25  
 N-methyl-N-nitrosourea (MNU) ..... 31–35, 40  
 Microarray, Oligonucleotide.....95–107  
 MicroRNAs  
     artificial MicroRNA (amiRNA) ..... 131–148  
 MNU. *See* N-methyl-N-nitrosourea (MNU)  
 Mutagenesis  
     chemical.....29–36  
     site-directed ..... 132–135, 138–141  
     T-DNA insertional.....29–30  
 Mutation ..... 30–32, 35, 40–42,  
 48–51, 54, 58, 66, 84, 132, 134, 144, 145, 276,  
 278, 287

**N**

Nucleotide polymorphism  
     single nucleotide polymorphism (SNP).....42

**O**

Oxidative stress.....228–229

**P**

Pathogenesis-related (PR) ..... 287, 292,  
 294–296, 298, 302  
 PCR. *See* Polymerase chain reaction (PCR)  
 Photosynthesis  
     photosynthetic efficiency .....217

Physical map.....13–14  
 Pollination ..... 3, 6, 7, 9–10,  
 73, 201, 313

Polyacrylamide gel electrophoresis (PAGE)  
     one dimensional.....157–158  
     SDS-PAGE.....157–158, 168–172, 174, 180  
     two dimensional .....158–159

Polymerase chain reaction (PCR)..... 17, 40,  
 42–45, 48–51, 54, 59–63, 65, 66, 73, 76–77, 80,  
 88, 90–92, 101, 110–111, 116–118, 120, 121,  
 125, 127–129, 132–134, 140–145, 147–148, 214,  
 219, 229–231, 236–237, 239

PR. *See* Pathogenesis-related (PR)  
 Promoter.....60, 79, 83–88, 92, 134,  
 142–144, 147, 270–273, 277, 278, 286, 290, 291,  
 298–301, 313–321

Proteomics .....151–183

**Q**

QTLs. *See* Quantitative trait loci (QTLs)  
 Quantitative trait loci (QTLs)..... 69–81, 258,  
 270–272, 277, 278, 293, 297

**R**

Recombinant inbred lines (RILs) ..... 80  
 Recombinant protein .....311–322  
 Resistance gene..... 85, 139, 250, 287, 290–291, 300–302  
*Rhizoctonia solani*, rice sheath blight..... 257, 258,  
 260, 261, 263, 265, 266  
 RILs. *See* Recombinant inbred lines (RILs)  
 RNAi. *See* RNA interference (RNAi)  
 RNA interference (RNAi)..... 132, 286–287, 289,  
 290, 293–295, 300  
 Root hair .....243–247  
 RT-PCR  
     quantitative ..... 230–231, 236–237

**S**

SA. *See* Salicylic acid (SA)  
 Salicylic acid (SA) .....185–199, 229, 292–297, 303  
 Salinity .....71, 97, 209–225, 269,  
 272, 273, 275, 295–298  
 Salt tolerance ..... 69–81, 210–224, 272–278  
 Site-specific integration.....83–92  
 Small RNA.....106, 109–129, 138, 145, 148  
 Sodium azide (NaN<sub>3</sub>) ..... 31–33, 35, 159  
 Soil  
     moisture content..... 211, 214–215, 217, 223  
     water potential .....211, 215  
 Southern blot..... 73, 74, 80, 90–92  
*Spodoptera frugiperda*, Fall armyworm .....203  
 Stomatal conductance (Gs) .....217



**T**

- Targeting of induced local lesions in genomes  
(TILLING) ..... 30–32, 36, 39–54
- TILLING. *See* Targeting of induced local lesions in  
genomes (TILLING)
- Transcription factors.....228, 229, 238, 270–273,  
287, 292, 293, 296–299, 301, 313, 315–317
- Transcriptome  
expression profiling.....95–107
- Transformation  
*Agrobacterium* ..... 58, 79, 84, 87–89, 143, 286, 321–322  
biolistic .....87–88  
particle bombardment.....85, 322
- Transgenic plant ..... 83, 88, 132, 134, 143–145, 147,  
148, 186, 214, 221, 272, 276–278, 291, 292, 321

**V**

- VOCs. *See* Volatile organic compounds (VOCs)
- Volatile organic compounds (VOCs).....201–203

**W**

- Water use efficiency..... 216–217, 223,  
225, 272

**X**

- Xanthomonas*  
*X. oryzae* pv. *oryzae*, rice bacterial blight.....289  
*X. oryzae* pv. *oryzicola*, rice bacterial streak.....249