REVIEW

Web Tools for Rice Transcriptome Analyses

Ki-Hong Jung · Jong-Seong Jeon · Gynheung An

Abstract Gene expression databases provide profiling data for the expression of thousands of genes to researchers worldwide. Oligonucleotide microarray technology is a useful tool that has been employed to produce gene expression profiles in most species. In rice, there are five genome-wide DNA microarray platforms: NSF 45K, BGI/ Yale 60K, Affymetrix, Agilent Rice 44K, and NimbleGen 390K. Presently, more than 1,700 hybridizations of microarray gene expression data are available from public microarray depositing databases such as NCBI gene expression omnibus and Arrayexpress at EBI. More processing or reformatting of public gene expression data is required for further applications or analyses. Web-based databases for expression meta-analyses are useful for guiding researchers in designing relevant research schemes. In this review, we summarize various databases for expression meta-analyses of rice genes and web tools for further applications, such as the development of co-expression network or functional gene network.

K.-H. Jung (✉) · J.-S. Jeon · G. An
Crop Biotech Institute, Kyung Hee University,
Yongin 446-701, South Korea
e-mail: khjung2010@khu.ac.kr

K.-H. Jung · J.-S. Jeon · G. An
Department of Plant Molecular Systems Biotechnology,
Kyung Hee University,
Yongin 446-701, South Korea

J.-S. Jeon
Graduate School of Biotechnology, Kyung Hee University,
Yongin 446-701, South Korea

Introduction

With the advent of *high-throughput* methodologies including genome-wide transcriptome analysis tools, the research efficiency has been significantly augmented. Since rice has become a model crop plant owing to accumulated genomics resources in the past decades, the use of a large number of microarray data and integrating omics analyses using bioinformatics has accelerated functional analysis of genes in rice. Furthermore, bioinformatics platforms and their associated database for the analysis of functional gene network have been developed (Ficklin et al. 2010; Jung et al. 2010a; Sato et al. 2011; Wise et al. 2007; Zheng and Wang 2008; Zimmermann et al. 2004). Here, we introduce high-throughput comprehensive web tools for transcriptome and integrating omics analyses suitable for rice functional genomics.

Public Collection Databases of Rice Microarray Data

NCBI Gene Expression Omnibus The Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) is a public storage site that collects and freely shares microarray, next-generation sequencing, and other forms of high-throughput functional genomic data submitted by researchers (Barrett et al. 2009). Users can query datasets, gene profiles, and GEO accessions on the home page. In addition, web-based interfaces and applications for expression profiling analyses are available to help users query and download experimental data and gene expression

patterns stored in GEO (Barrett et al. 2009). There are five genome-wide array platforms available for expression profiling analyses of rice genes: Affymetrix, Agilent Rice 44K, BGI/Yale 60K, NSF 45K, and NimbleGen 390K. As of now, users can assess 908 hybridizations of Affymetrix (Platform GPL2025), 99 Agilent Rice 44K (Platform GPL6864 and Platform GPL8030), 144 of BGI/Yale 60K (Platform GPL1829), 117 NSF 45K (GPL4105/GPL4106, GPL5739/GPL5740, and GPL6939/GPL6940), and 12 Rice NimbleGen 390K tiling array (GPL7344) from NCBI GEO. In addition, there are 279 Agilent Rice 22K (GPL892 and GPL477) and 114 NSF 20K (GPL2091) covering about half of the rice genome (Jung et al. 2008a, b). Rice DNA microarray platforms cover over 90% of rice genes annotated by the Rice Genome Annotation Project (RGAP/MSU, http://rice.plantbiology.msu.edu/) (Jung et al. 2008a). Rice Multiple-platform Microarray Element Search tool (http://www.ricearray.org/element/search_multiple.shtml) allows users to search probes across different rice microarray platforms such as Affymetrix, Agilent Rice 22K, BGI/Yale, NSF 20K, and NSF 45K to map to a common gene locus (Jung et al. 2008b). In addition, identifier (id) converter tool in the Rice Annotation Project Database (RAP-DB, http://rapdb.dna.affrc.go.jp/tools/converter/run) provides a table of the RAP-DB's Os ids and RGAP/MSU's LOC_Os ids and allows users to convert gene ids of one database into those of the other.

ArrayExpress at EBI The ArrayExpress (http://www.ebi.ac.uk/arrayexpress/) is a database of gene expression profiling analyses where users can query and download data following MIAME standards similar to NCBI GEO (Parkinson et al. 2007). To conduct a query, there are options for selecting the species, platforms, molecules, and techniques. The ArrayExpress contains most of the data deposited at NCBI GEO as well as the data exclusively submitted to this database. There are currently 163 Affymetrix and eight Agilent Rice 44K arrays available pertaining to the unique rice gene expression profiles.

Web Application Tools for Transcriptome Data Process

Unlike the above two databases which collect and share the microarray data, it is required to reorganize these microarray data and utilize various analysis tools when users want to get useful information about their target genes. Web tools to analyze expression patterns of an individual or a group of genes have been developed. Genevestigator, PLEXdb, Rice Expression Profile Database, Rice eFP browser, and Gene Expression Atlas in RiceGE allow users to explore expression patterns of rice genes of interest. In addition, users can easily identify biological meanings of certain genes in response treatments by using Gene Ontology enrichment analysis tool and also metabolic pathways associated with genes of interest by integrating metabolic pathways and gene expression data. To identify candidate genes functionally associated with a target gene, co-expression analysis tool using global gene expression data and functional gene network integrating various high-throughput omics data is effective. Table 1 summarizes various tools for gene expression analyses and further applications in rice. In subsequent sections, we describe useful web application tools for functional analyses of rice genes.

Genevestigator Genevestigator (https://www.genevestigator.com/gv/index.jsp) is a user-friendly expression visualization tool that assists in gene analysis predictions in the context of various experimental conditions such as harvesting times, tissue types, stress types, and genetic backgrounds. There are 54,940 microarray dataset available for 12 organisms. Of these, 324 Affymetrix arrays from 33 independent experiments are available for gene expression profiling analyses in rice. The microarray dataset can be reorganized according to the following categories: anatomy, development, stimulus, mutation, and neoplasm (Zimmermann et al. 2004, 2008; Hruz et al. 2008). Figure 1 shows expression patterns of three selected genes, LOC_Os05g38530, LOC_Os01g62290, and LOC_Os11g47760. In the "Meta-Profile Analysis" module, the profiles of the "Stimulus" category are represented by heat map with fold change value of treatment over control. In this case, LOC_Os11g47760, represented by probe id Os.140.1. A1_at, showed significant up-regulation in response to arsenate and drought. When the category "Neoplasm" is selected, the profiles are viewed as heat maps or scatter plots by using relative spot intensities or actual spot intensities. When the "Clustering Analysis" module is selected, the expression patterns of the three genes in anatomical samples are hierarchically clustered. To obtain expression patterns using this web tool, users must first activate Genevestigator V3 navigator, which runs on the Java environment. Next, users can select organisms (e.g., *Oryza sativa*) for "Expression Profiling Analyses" by selecting the plus sign at step 1. By selecting the plus sign at step 2, users can upload a querying gene or probe id list. For further analyses, there are four modules at step 3: "Meta-Profile Analysis," "Biomarker Search," "Clustering Analysis," and "Pathway Projector". Of these, functions in Meta-Profile Analysis are freely available, while those in other modules are not free. More detailed information for users can be found at the following link: https://www.genevestigator.com/userdocs/manual/index.html.

Table 1 Summary of databases for rice transcriptome analyses and further applications described in this review

Features	Database full names	Abbreviations	Websites
Collection of transcriptome data	NCBI Gene Expression Omnibus	GEO	http://www.ncbi.nlm.nih.gov/geo/
	ArrayExpress at EBI	ArrayExpress	http://www.ebi.ac.uk/arrayexpress/
Web tools for transcriptome data analysis	Genevestigator	Genevestigator	https://www.genevestigator.com/gv/index.jsp
	Rice Plant expression database	RicePLEXdb	http://www.plexdb.org/
	Rice Expression Profile Database	RiceXPro	http://ricexpro.dna.affrc.go.jp/index.html
	Rice eFP browser	Rice eFP browser	http://www.bar.utoronto.ca/
	Gene Expression Atlas in RiceGE	Gene Expression Atlas in RiceGE	http://signal.salk.edu/cgi-bin/RiceGE?JOB=APPENDIX&QUERY=GeneAtlas
Web tools for integrating various omics data analyses	Gene Ontology Enrichment Analysis Software Toolkit	GOEAST	http://omicslab.genetics.ac.cn/GOEAST/
	RiceCyc	RiceCyc	http://pathway.gramene.org/gramene/ricecyc.shtml
	Gene co-expression networks	Gene co-expression networks	http://www.clemson.edu/genenetwork/network.php
	Rice Gene Expression Network	RGEN	http://riceball.lab.nig.ac.jp/Rice_network_public/script/index.html
	Rice Interactions Viewer	Rice Interactions Viewer	http://bar.utoronto.ca/rice_interactions/cgi-bin/rice_interactions_viewer.cgi
	PlantArrayNet (previous RiceArrayNet)	PlantArrayNet	http://arraynet.mju.ac.kr/arraynet/

RicePLEXdb PLEXdb (http://www.plexdb.org/) is a unified public resource for gene expression analyses of plants and plant pathogens (Guldener et al. 2006; Wise et al. 2007). PLEXdb serves as a frame to integrate rapidly increasing gene expression profile datasets with conventional structural genomics and phenotypic data. The integrated tools of PLEXdb allow users to exploit commonalities in plant biology for a comparative approach to functional genomics through the use of large-scale expression profiling datasets (Wise et al. 2007). Presently, microarrays from 14 plant species are available in this database. For gene expression profiling analyses in rice (http://www.plexdb.org/modules/PD_browse/experiment_browser.php?plex_name=Rice-PLEX), users can explore gene expression patterns with probe ids for 35 data series produced by the Affymetrix platform. To further use this web tool for gene expression profiles in rice, users must select "rice" in the option to select species found in the left panel of the homepage of this database, and then go to a web browser for the RicePLEXdb. Users can browse RicePLEXdb experimental data by using the six web tools for gene expression profiling analyses: "Experiment" provides overall information for the selected individual dataset; "Expression" provides expression profiling data of a probe id in the dataset; "Hybridizations & Samples" provides various types of information on the dataset such as name of the sample, name of the treatment, number of treatments, and number of replicates, and expression data file name; "Quality Control" provides the results of quality check (QC) for

treatment or control after processing raw data; "Compare Treatments" provides a differentially expressed gene list between two samples (one selected in the X axis and the other in the Y axis) generated with Scatterblot or the Mva plot for logged or original normalized mean values (users download data for significant gene lists in a tap-delimited txt file format after selecting fold change cut-off); and "Downloads" provides batch download. As an example, Fig. 2 shows the expression patterns of probe id Os.10220.1.S1.at representing LOC_Os07g48820, which encodes a putative transcription factor.

RiceXPro Rice Expression Profile Database (RiceXPro, http://ricexpro.dna.affrc.go.jp/index.html) is a repository of gene expression data derived from microarray experiments covering the entire life cycle of the rice plant including germination, seedling, tillering, stem elongation, panicle initiation, booting, heading, flowering, and the ripening stage. The spatio-temporal gene expression of various tissues/organs throughout the entire growth period, derived from a total of 48 samples in three replicates, was used for hybridization using the 4X44K Agilent one-color (Cy3) microarray-based gene analysis system (Sato et al. 2011). The expression profile for each gene in various tissues is shown as raw data representing Cy3 signal intensity and normalized data (log2). Search options include a key word search for individual RAP-DB locus id or RGAP/MSU locus id, or a chromosome search to generate a list of all genes for each chromosome.
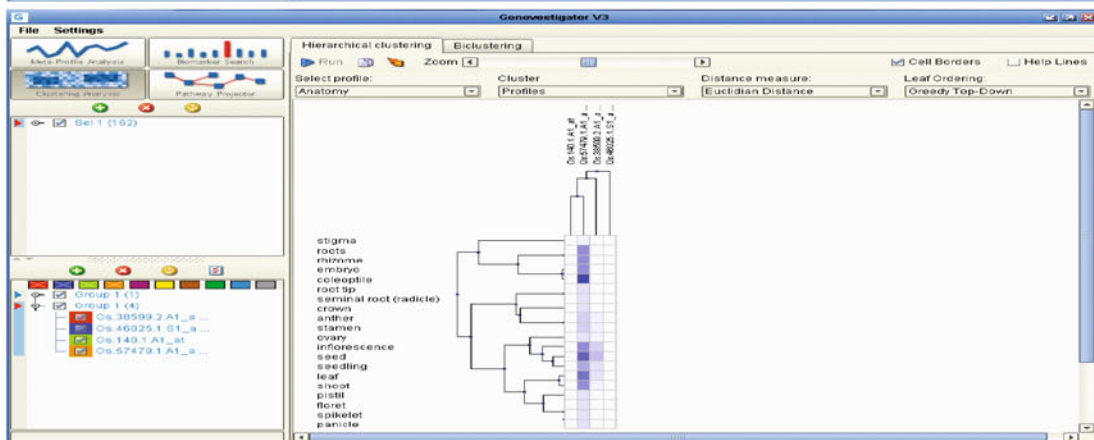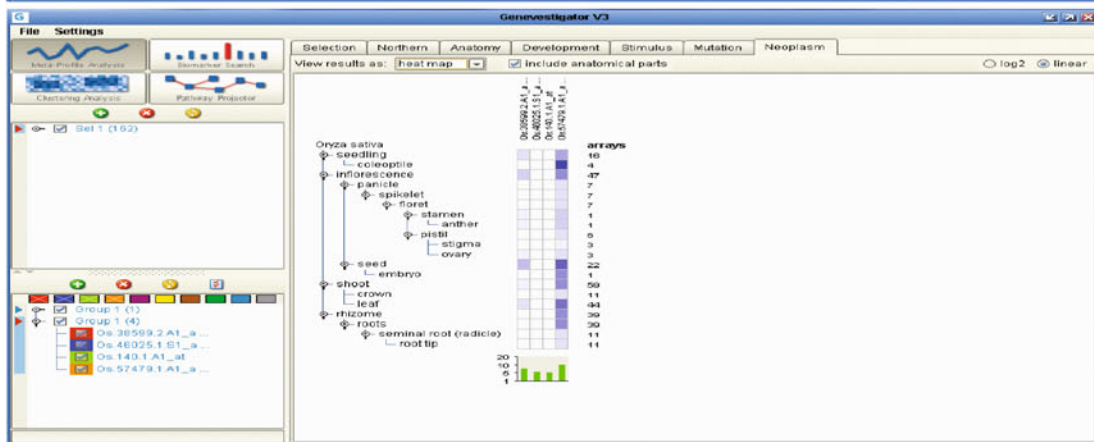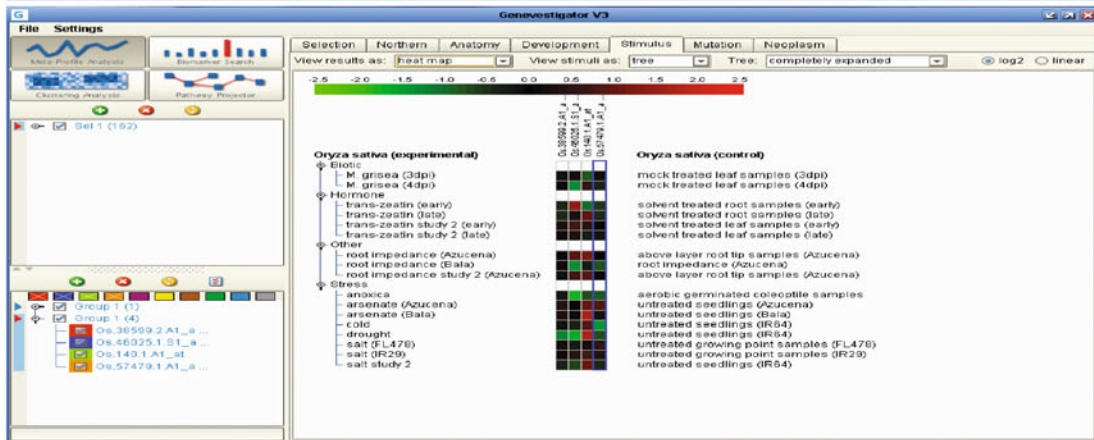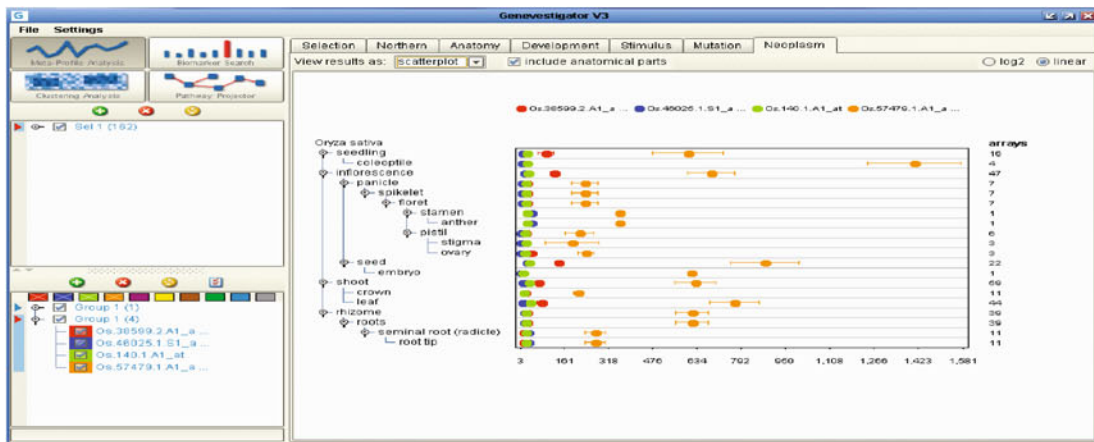
◀ **Fig. 1** Expression pattern analyses of LOC_Os11g47760 gene using Genevestigator

RiceXPro also provides an overview of the expression pattern of all rice genes in specific tissues at various stages of growth and development under normal field conditions. Leaf samples at various growth stages encompassing the vegetative, reproductive, and ripening stages were used for this gene expression profile. Samples corresponding to the uppermost fully expanded leaves were collected at 12:00 PM at weekly intervals during the 2008 cultivation season from 13 to 125 days after transplantation. This leaf gene expression

profile throughout the entire growth in the field was generated using a total of 17 samples in three replicates. Search options are the same as those described in the previous section. Users can download the expression patterns of a query gene in the form of graphs or heat maps with illustrations of developing rice plants or various tissues/organs throughout the growth stages. Figure 3 shows the expression patterns of LOC_Os04g09390 encoding the lectin precursor, obtained by selecting for "Spatio-temporal gene expression of various tissues/organs throughout entire growth in the field." Data for "Leaf gene expression profile throughout entire growth in the field" is presented in Figure S1.
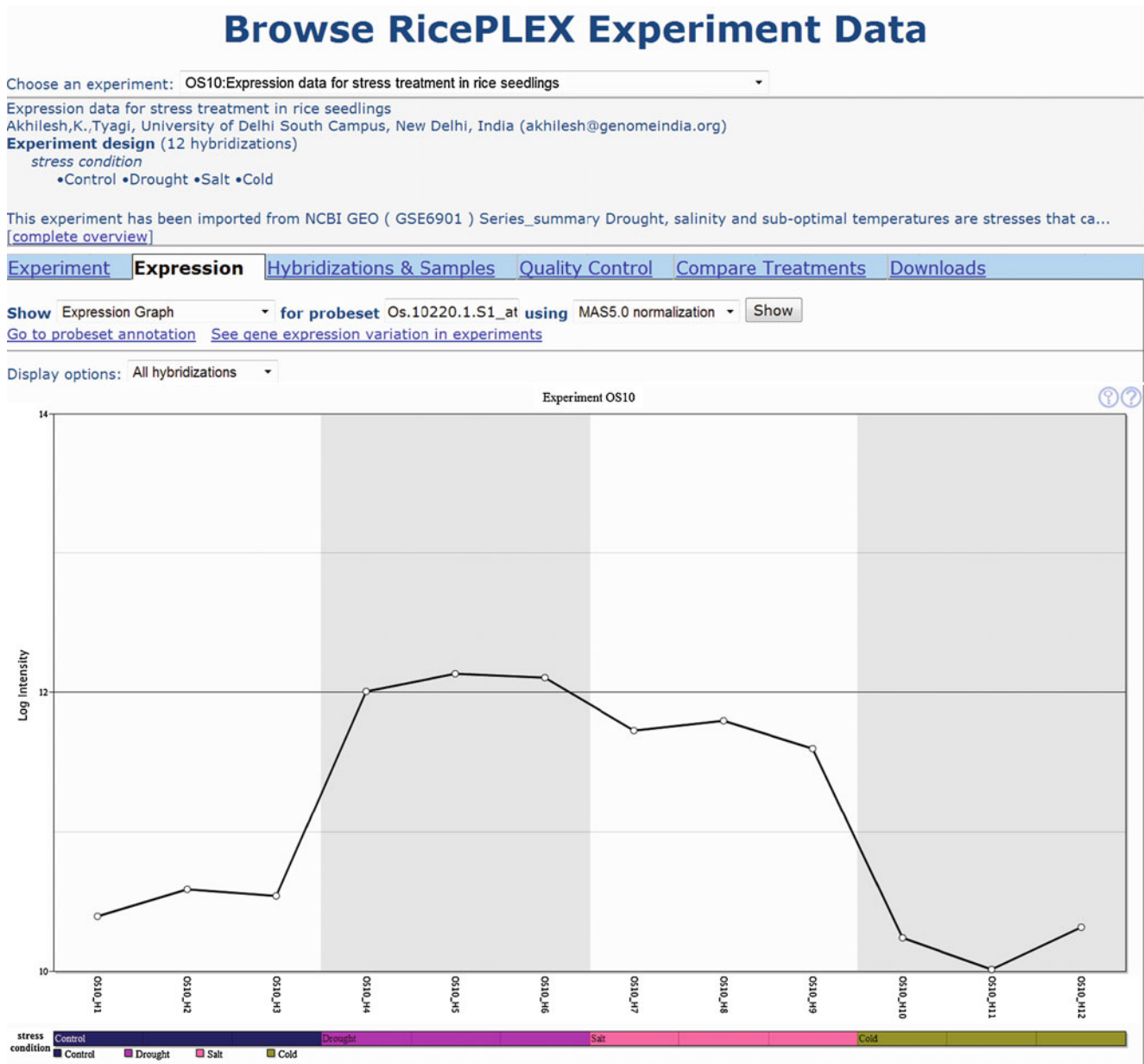
Fig. 2 Expression pattern analyses of LOC_Os07g48820 gene using RicePLEXdb

RXP_0001: Organs and tissues; Os04g0173800_15779
LOC_Os04g09390.1

RXP_0001: Organs and tissues; Os04g0173800_15779
LOC_Os04g09390.1

Rice eFP Browsers Botany Array Resources (BAR, http://
www.bar.utoronto.ca/) is a user-friendly web-based tool for
working with functional genomics and other omics data
(Winter et al. 2007) and provides useful information for plant
researchers. eFP browsers are unique in this website and are
available for *Arabidopsis*, poplar, *Medicago*, rice, maize,
soybean, and barley. Expression profiling analyses for four
Affymetrix array datasets consisting of 66 hybridizations are
accessible in the rice eFP browser. For example, Fig. 4 shows
expression patterns of LOC_Os07g48820 encoding putative
transcription factor in stress-treated seedlings (Jain et al.
2007). Each of the four datasets is processed by two
normalization methods such as Robust Multi-array Average
(RMA) and Microarray Suite (MAS), and therefore there are
eight options in the box below "Data Source." By using the
box below "Mode," users select the type of expression
profiles represented by "Absolute (value)" or "Relative
(value)". If users choose to compare options in the same box,
users can obtain the data for relative expression levels of one
rice gene compared to another, i.e., Primary Rice ID over
Secondary Rice ID. "Signal Threshold" indicates the highest
expression level of the resulting expression profiling data. The
eFP browser is the most convenient tool to visualize expression
patterns of genes of interest, but the limitation in the case of the
rice eFP browser is the number of datasets available.

Gene Expression Atlas in RiceGE Gene Expression Atlas in
Rice Functional Genomic Express Database (RiceGE,
http://signal.salk.edu/cgi-bin/RiceGE) provides normalized
expression levels for microarrays in nine Affymetrix datasets
(GPL2025) and fold change for an Agilent Rice 22K dataset
(GSE661 for ABA or GA calli treatment). Users can access
the gAtlas tool through the RiceGE website at http://signal.
salk.edu/cgi-bin/RiceGE?JOB=APPENDIX&QUERY=
GeneAtlas. The gene expression data can be queried after
removing LOC_ in locus id provided by the Rice Genome
Annotation Project. Figure 5 shows expression patterns of
Os07g48820. A summary of this gAtlas tool is available at
http://signal.salk.edu/RiceGE/RiceGE_gAtlas_Source.html.

Web Application Tools for Integrating Various Omics Data

Gene Ontology Enrichment Analysis Software Toolkit The
Gene Ontology Enrichment Analysis Software Toolkit

(GOEAST, http://omicslab.genetics.ac.cn/GOEAST/) is a
web-based tool providing user-friendly Gene Ontology
(GO) terms for high-throughput experimental results,
especially for microarray data (Zheng and Wang 2008).
Users can identify significantly enriched GO terms among a
given list of genes under statistical evidence indicated by
the p value. In general, the default threshold of the p value
is 0.05. There are three GO principle categories: "Biolog-
ical process," "Cellular component," and "Molecular
function." To identify enriched GO terms for gene lists in
rice, users must select the "Tools" option on the home page
(http://omicslab.genetics.ac.cn/GOEAST/). Then, users can
select "Microarray Platform" for further analyses. When the
Affymetrix array platform is selected, GOEAST requires
probe ids for a list of candidate genes. At this step, Rice
Multiple-platform Microarray Element Search tool is
helpful to get Affymetrix probe ids with various gene ids
from different genome browsers such as RAP-DB and
MSU (Jung et al. 2008b). The web-based features in
GOEAST provide users with visualized results. Visual-
izations of the positional relationships and hierarchical
relationships of enriched GO terms will help users to
more easily understand their results. In the graphical
output, GOEAST shows the p value of each enriched term
and uses different colored boxes to present the enrichment
significance of different GO terms. Figure 6 shows an
example for GO enrichment analysis in the "biological
process" category of genes at least fourfold up-regulated
in response to drought (OS10: Expression data for stress
treatment in rice seedlings) (Jain et al. 2007). Data for the
"cellular component" category of the up-regulated genes is
shown in Figure S2.

RiceCyc RiceCyc (http://pathway.gramene.org/gramene/
ricecyc.shtml), developed and curated by the Gramene
Database, provides an index of known and/or predicted
biochemical pathways in rice. This tool contains 339
pathways and 1,592 compounds consisting of 9,315
enzymes for 2,109 enzymatic reactions and 603 transporters
for 87 transport reactions. The "Pathway Tools Omics
Viewer" enables data values from the user's high-throughput
data such as microarray data, proteomics data, and metab-
olomics data to integrate onto the "Cellular Overview"
diagram for rice (Kind et al. 2009; Jung et al. 2010a). To
use this integrating tool, users must input a tab-delimited
data file that is stored on the user's local computer. More
details for the input file are available at http://pathway.
gramene.org/expression.html. An example of this web tool is
shown in Fig. 7. We generated integrating data, with the NSF
45K array, by incorporating fold change data (more than
twofold up-regulated) of 1,000 candidate genes from a
comparison of light- versus dark-grown rice seedlings, onto
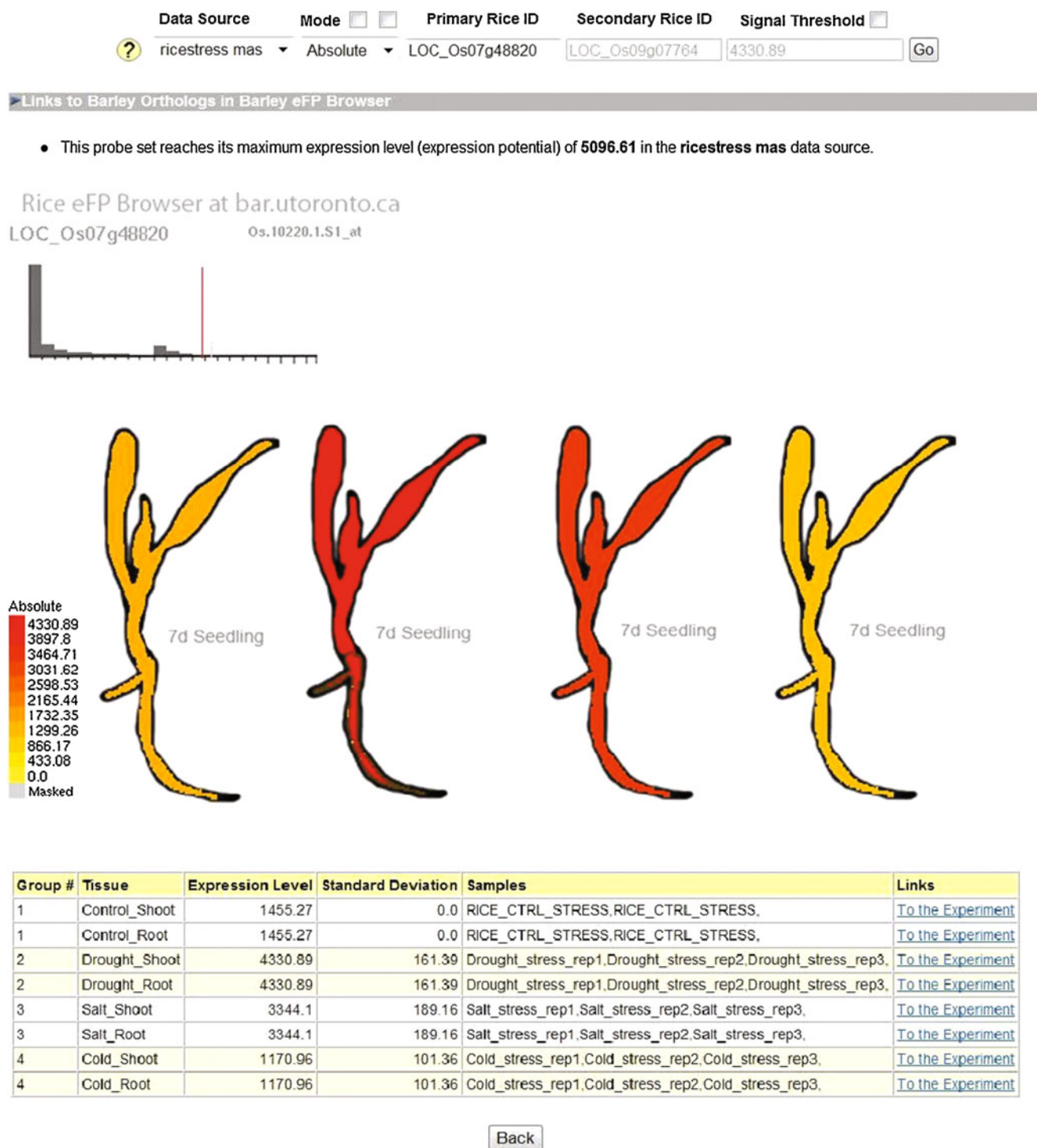the "Cellular Overview" diagram for rice. As a result, we

Fig. 4 Expression pattern analyses of LOC_Os07g48820 in stress-treated seedlings using rice eFP browser in BAR

identified three components in the photorespiration pathway that are co-regulated by light stimuli, indicating that this pathway might be activated in response to light (Jung et al. 2008b, c). The coverage of this metabolic pathway tool at a genomic scale is 22.9%, derived from the ratio of the number (9,315) of genes encoding total enzymes in RiceCyc over the number (40,612) of total non-transposable element (non-TE) genes in rice.

Gene Co-expression Network Browser Gene co-expression networks (http://www.clemson.edu/genenetwork/network. php) can help identify closely co-expressed gene sets
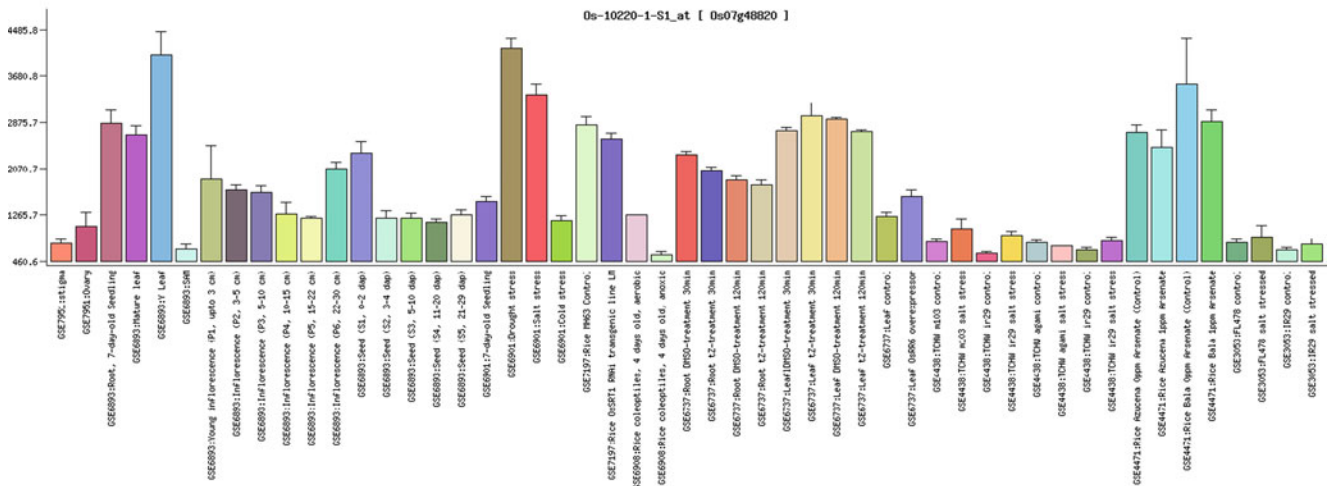
Fig. 5 Expression pattern analyses of LOC_Os07g48820 using gAtlas tool in RiceGE

(modules), which are associated with genes of known and unknown function by using a large number of microarray data. Pearson correlation analysis between gene expression profiles for the underlying metric for co-expression and weighted correlation network analysis to construct the gene co-expression network methods in rice produced 60 co-expressed gene modules and 110 enriched functional clusters by using 508 Affymetrix hybridizations downloaded from NCBI GEO (Ficklin et al. 2010). There are seven options on this web site: "About" provides an introduction for this web tool and major network clusters; "Mapping" provides the probe id and locus id for a selected module; "Enrichment" provides enriched gene ontology (GO) terms within the selected module or cluster; "Edge" provides probe ids of edges that are linked to nodes in a selected module or cluster; "Network" provides co-expression network for a selected module or cluster; "Genome" provides chromosomal locations of components

in a developed network; "Search results" provides the results from search function in this web tool and users can query a probe set name, locus id, annotation term, or keyword (e.g., Os.12750.1.S1_a_at, LOC_Os01g58660, GO:0043565, or cyclin). Figure 8 shows the procedure to create a co-expression network with this web tool. Identification of groups of genes associated with a known phenotype will be very practical for building a complex network encompassing functional gene–gene interactions (Ficklin et al. 2010).

Rice Gene Expression Network Rice Gene Expression Network (RGEN, http://riceball.lab.nig.ac.jp/Rice_ network_public/script/index.html) is a web tool that evaluates the statistical similarities of genome-wide gene expression patterns with collected large-scale gene expression data from various experimental conditions. The microarray data for RGEN are assembled from the NCBI GEO database. The

Fig. 6 Examples for GO enrichment analysis of genes up-regulated in response to drought using GOEAST. Procedure of GO enrichment analysis in biological process category of genes at least fourfold differentially expressed in response to drought (OS10: Expression data for stress treatment in rice seedlings in RicePLEXdb) is described. First step is the identification of candidate genes showing significant differential expression patterns with fold change and *p* value. Next is to use GOEAST tool with probe ids for candidate gene list. Finally, graphic views of GO enrichment analyses for each GO principle category and summarizing table for the result are generated
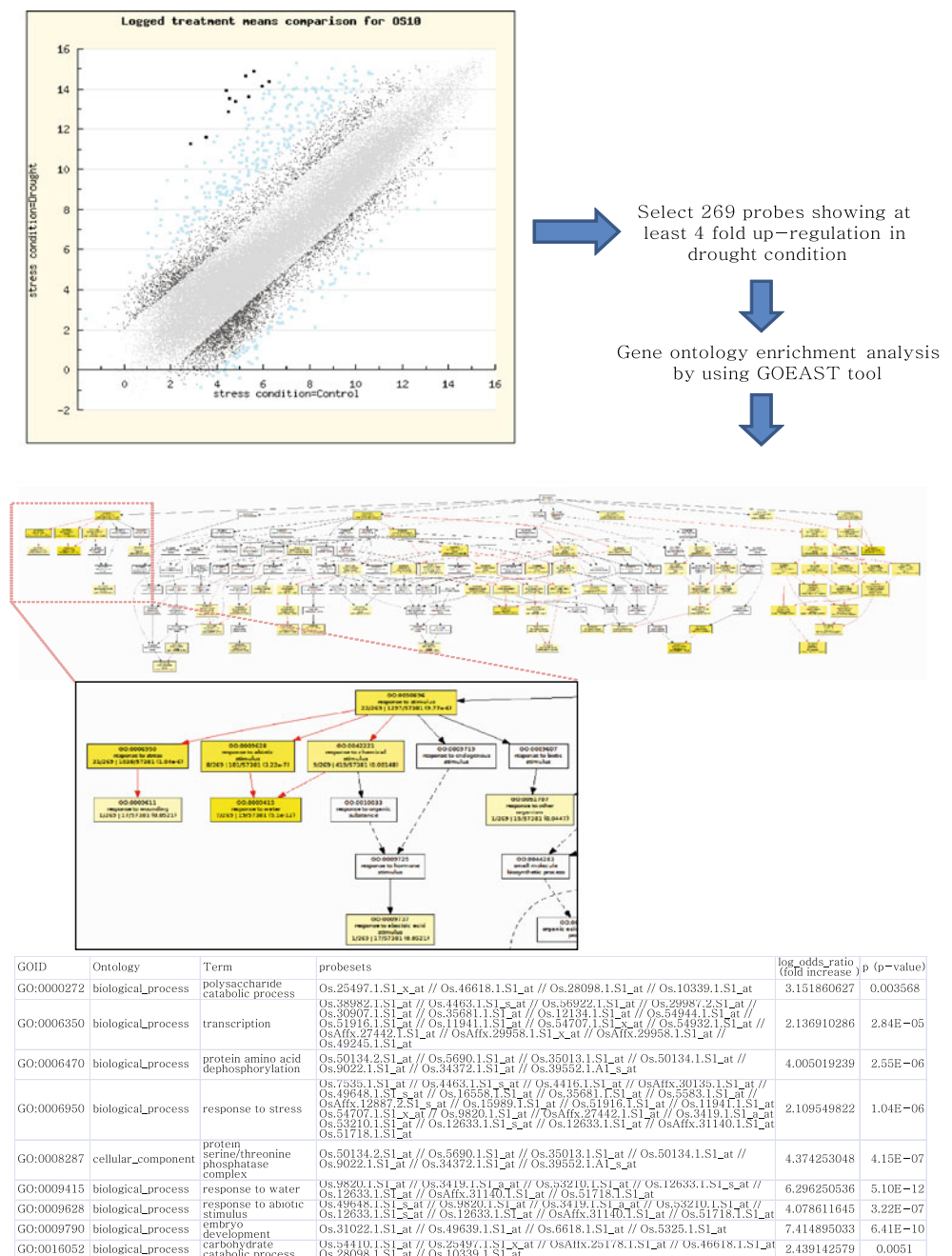
current version of RGEN is derived from Affymetrix 624 hybridizations in the database. CEL files of the microarray platform "Affymetrix Rice Genome Array" are normalized using the RMA method and calculates the distances between genes using the correspondence analysis (CA) method (Yano et al. 2006). In this method, similarities of gene expression patterns between a gene pair are evaluated using Pearson correlation coefficients (PCCs), Mutual Ranks (MRs), and partial correlation coefficient (PAC). In this website, there are six options: "Home" provides the links for OryzaExpress Network or OryzaExpress Annotation; "Search" enables users to find detailed information by using probe id or keywords; "Browsing" helps users develop a co-expression network with those of similar gene pairs by CA distance, reciprocal gene pairs by PCC_CA, and gene pair by PCC; "Link" provides external links for databases used for developing a co-expression network. There are two options for "download" and "help." To develop a co-expression network by using this web tool, users must first identify the probe id for a querying locus id as described in Figure S3. Then, users can search for a co-expression network for the probe id. If such a co-expression network linked to the probe id exists, users can obtain the co-expression network by selecting the network (Fig. 9). On the network page, detailed
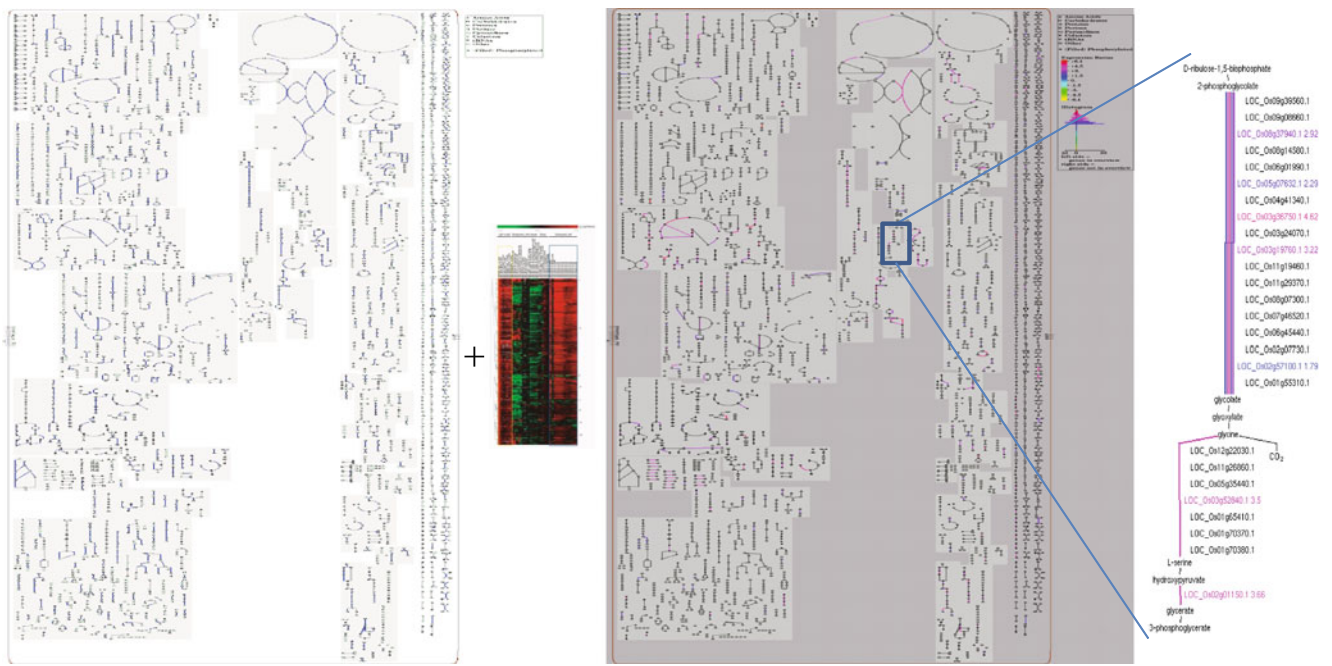
Reactome   +   Light inducible gene expression data → Co−expressed metabolic pathways (e.g. Photorespiration)

Fig. 7 Identification of metabolic pathways differentially regulated in response to light using RiceCyc

information for the queried probe id (Os.1311.1.S1_at for LOC_Os02g02410) is first displayed, and next a graphic window for producing the co-expression network is activated with various options to select the number of probe pairs for a gene, Similar/Reciprocal pattern, number of paths (edges) from a central node (probe), Index used to depict a network, Predicted orthologues, Metabolic Pathway, GO Pop-up window of a short description for a node (gene), and Description of homologous Arabidopsis gene name in a node (gene) (Fig. 9). Below the graphic window for the co-expression network, there is detailed information for 10 types of data linked to the co-expression network (Figure S4).

PlantArrayNet PlantArrayNet (http://arraynet.mju.ac.kr/arraynet/)which is previously called as RiceArrayNet providing information on the similarities between genes in rice and *Brassica* based on correlation coefficients ($r_{two-spots}$) (Lee et al. 2009). For two array platforms in rice, BGI/Yale 60K and Nimblegene Oryza sativa Exon 300K, users can easily obtain a co-expression network for a query gene ID with this tool. Figure 10 shows processes for obtaining a co-expression network for LOC_Os03g52840 encoding glycine hydroxymethyltransferase, knockout mutants of which display the chlorina phenotype in the early seedling stage (Jung et al. 2008a, c). Selecting "Search spot by keyword or sequence" is the first step in searching for the information for spot (probe) on the array by using keywords (e.g., Trehalose), sequence, or gene_ID (e.g., TIGR_ID, RAP_ID, Spot Number, Oligomer_ID).

The second step is to select "Spot List" for retrieving several pieces of information regarding the queried spot. Here, users must select "spot" to obtain a co-expression network and select "Draw a Network" or "Draw a Tree." Users can also choose Minimum absolute $r_{two-spots}$ over 0.5 (in Fig. 10, we chose 0.6). When the values of $r_{two-spots}$ are closer to 1, it indicates a stronger correlation of gene expression between gene pairs. The last step is to select "Relational Network" in developing a co-expression network for a queried gene. Here, there are three options in the graphic view to select for whether a user chooses all the spots, shows the names of the spot(s), or shows detailed information about a queried spot.

Rice Interaction Viewer Rice Interaction Viewer was recently integrated in BAR (http://bar.utoronto.ca/rice_interactions/cgi-bin/rice_interactions_viewer.cgi). Users can query a database of 37,112 predicted and 64 confirmed rice interacting proteins to develop a protein–protein interaction network for a list of genes. However, this tool does not include confirmed interacting proteins of 254 rice kinases which were recently reported (Ding et al. 2009; Jung et al. 2010b). Figure 11 shows a result of Rice Interactions Viewer queried with 569 genes showing preferential expression in developing rice seeds. With this tool, the user can also identify predicted sub-cellular localization information in the network, and all the information and data used for developing the network are presented as a table below the network.

Fig. 8 Procedures for the development of co-expression network of LOC_Os01g58660 gene using the Gene Co-expression Network Browser

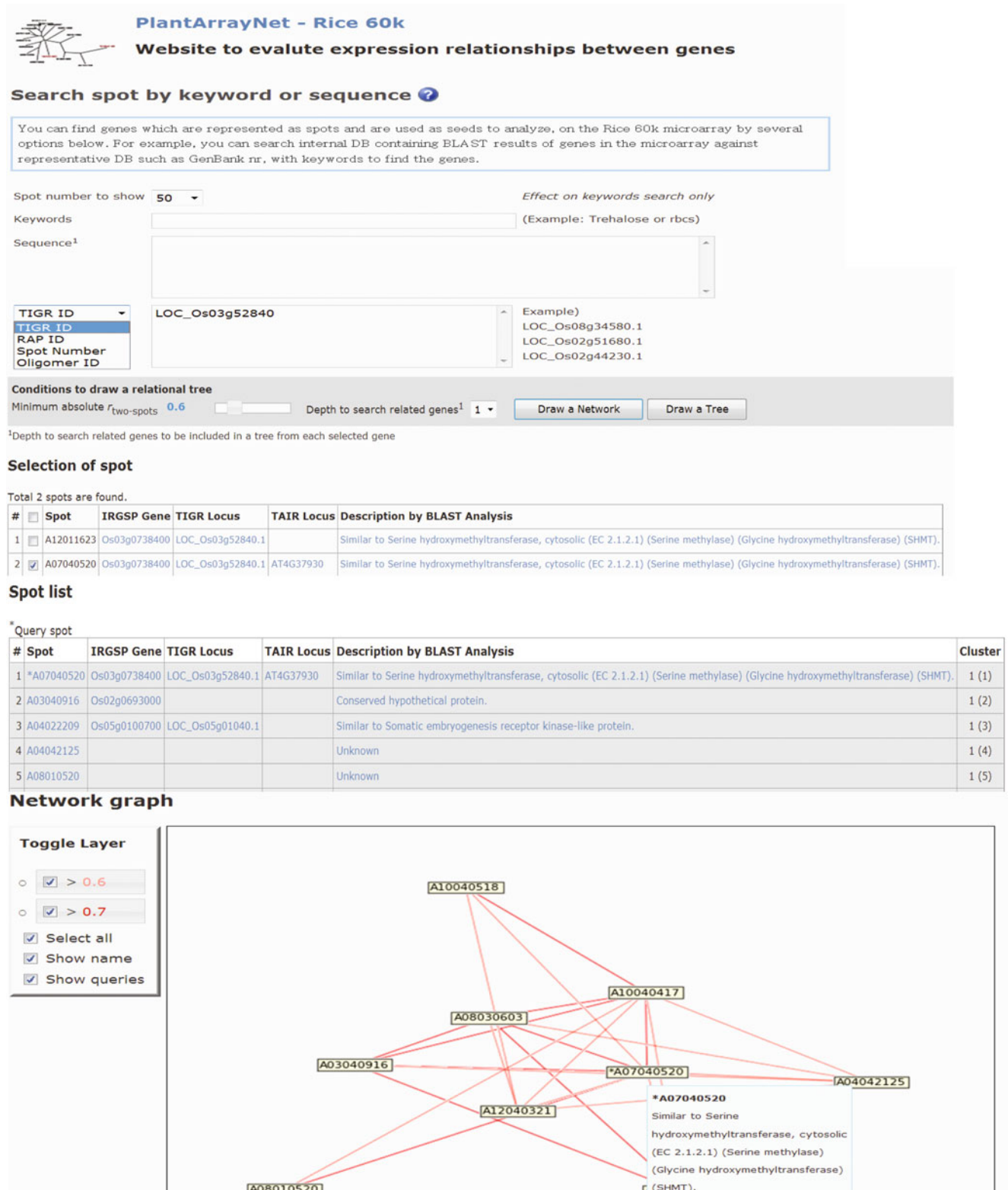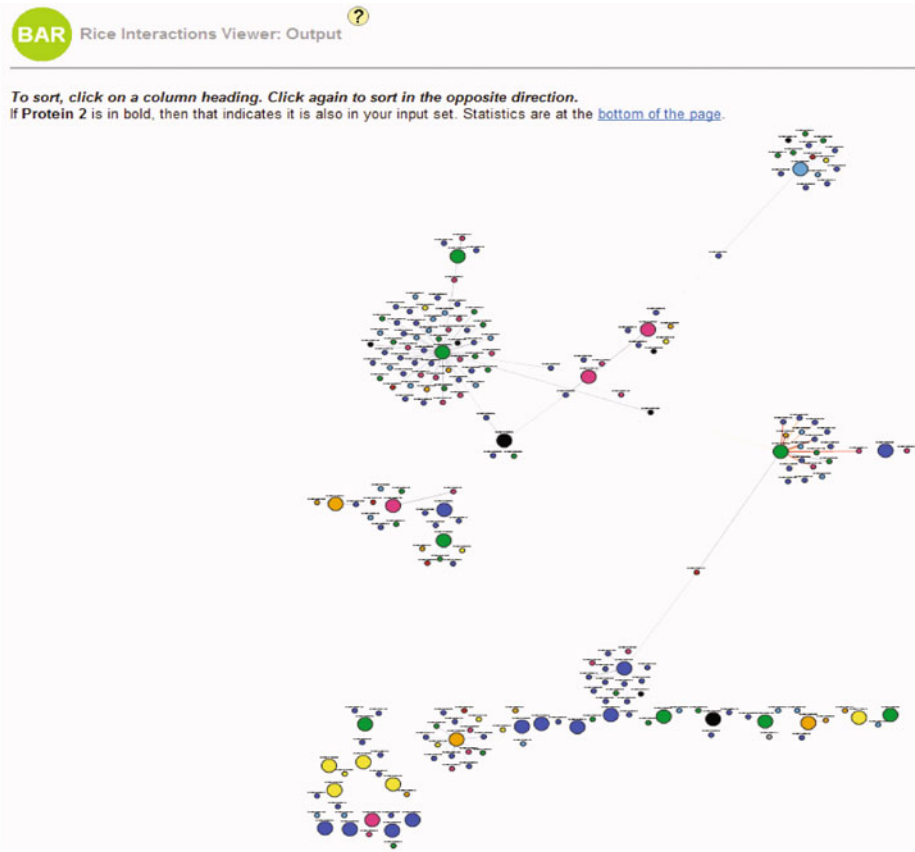Fig. 9 Procedures for the development of co-expression network of LOC_Os03g52840 gene using PlantArrayNet

Fig. 10 Procedures for the development of co-expression network of LOC_Os02g02410 gene represented by probe id Os.1311.1.S1_at using Rice Gene Expression Network

Functional gene network using developing seed preferential genes

Fig. 11 Development of functional gene network for embryo preferential genes in rice using Rice Interaction Viewer of BAR

Prospects

As transcriptome data using DNA microarray deposited in publicly accessible databases are significantly accumulating, web-based tools integrating multiple levels of information will guide researchers to find clues addressing various problems in improving the quality of crops (Table 1). The Gene Co-expression Network Browser, RGEN in OryzaExpress, and the Rice Interactions Viewer will be more useful for effective integration of various high-throughput omics data such as genome-wide microarray data, large-scale protein–protein interaction data, genome-wide QTL association data, and assignment of GO terms. Web tools introduced in this report will be more advanced and user-friendly after adopting multiple levels of feedback from users in related research areas as well as more active comparisons across species.

References

Barrett T, Troup DB, Wilhite SE, Ledoux P, Rudnev D, Evangelista C, Kim IF, Soboleva A, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Muertter RN, Edgar R (2009) NCBI GEO: archive for high-throughput functional genomic data. Nucleic Acids Res 37:D885–D890

Ding X, Richter T, Chen M, Fujii H, Seo YS, Xie M, Zheng X, Kanrar S, Stevenson RA, Dardick C, Li Y, Jiang H, Zhang Y, Yu F, Bartley LE, Chern M, Bart R, Chen X, Zhu L, Farmerie WG, Gribskov M, Zhu JK, Fromm ME, Ronald PC, Song WY (2009) A rice kinase–protein interaction map. Plant Physiol 149:1478–1492

Ficklin SP, Luo F, Feltus FA (2010) The association of multiple interacting genes with specific phenotypes in rice using gene coexpression networks. Plant Physiol 154:13–24

Guldener U, Seong KY, Boddu J, Cho S, Trail F, Xu JR, Adam G, Mewes HW, Muehlbauer GJ, Kistler HC (2006) Development of a *Fusarium graminearum* Affymetrix GeneChip for profiling fungal gene expression in vitro and in planta. Fungal Genet Biol 43:316–325

Hruz T, Laule O, Szabo G, Wessendorp 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

Jain M, Nijhawan A, Arora R, Agarwal P, Ray S, Sharma P, Kapoor S, Tyagi AK, Khurana JP (2007) F-box proteins in rice. Genome-wide analysis, classification, temporal and spatial gene expression during panicle and seed development, and regulation by light and abiotic stress. Plant Physiol 143:1467–1483

Jung KH, An G, Ronald PC (2008a) Towards a better bowl of rice: assigning function to tens of thousands of rice genes. Nat rev 9:91–101

Jung KH, Dardick C, Bartley LE, Cao P, Phetsom J, Canlas P, Seo YS, Shultz M, Ouyang S, Yuan Q, Frank BC, Ly E, Zheng L, Jia Y, Hsia AP, An K, Chou HH, Rocke D, Lee GC, Schnable PS, An G, Buell CR, Ronald PC (2008b) Refinement of light-responsive transcript lists using rice oligonucleotide arrays: evaluation of gene-redundancy. PLoS ONE 3:e3337

Jung KH, Lee J, Dardick C, Seo YS, Cao P, Canlas P, Phetsom J, Xu X, Ouyang S, An K, Cho YJ, Lee GC, Lee Y, An G, Ronald PC (2008c) Identification and functional analysis of light-responsive unique genes and gene family members in rice. PLoS Genet 4: e1000164

Jung KH, Seo YS, Walia H, Cao P, Fukao T, Canlas PE, Amonpant F, Bailey-Serres J, Ronald PC (2010a) The submergence tolerance regulator Sub1A mediates stress-responsive expression of AP2/ ERF transcription factors. Plant Physiol 152:1674–1692

Jung KH, Cao P, Seo YS, Dardick C, Ronald PC (2010b) The Rice Kinase Phylogenomics Database: a guide for systematic analysis of the rice kinase super-family. Trends Plant Sci 15:595–599

Kind T, Scholz M, Fiehn O (2009) How large is the metabolome? A critical analysis of data exchange practices in chemistry. PLoS ONE 4:e5440

Lee TH, Kim YK, Pham TT, Song SI, Kim JK, Kang KY, An G, Jung KH, Galbraith DW, Kim M, Yoon UH, Nahm BH (2009) RiceArrayNet: a database for correlating gene expression from transcriptome profiling, and its application to the analysis of coexpressed genes in rice. Plant Physiol 151:16–33

Parkinson H, Kapushesky M, Shojatalab M, Abeygunawardena N, Coulson R, Farne A, Holloway E, Kolesnykov N, Lilja P, Lukk M, Mani R, Rayner T, Sharma A, William E, Sarkans U, Brazma A (2007) ArrayExpress—a public database of microarray experiments and gene expression profiles. Nucleic Acids Res 35:D747–D750

Sato Y, Antonio BA, Namiki N, Takehisa H, Minami H, Kamatsuki K, Sugimoto K, Shimizu Y, Hirochika H, Nagamura Y (2011) RiceXPro: a platform for monitoring gene expression in *japonica* rice grown under natural field conditions. Nucleic Acids Res 39: D1141–D1148

Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV, Provart NJ (2007) An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. PLoS ONE 2:e718

Wise RP, Caldo RA, Hong L, Shen L, Cannon E, Dickerson JA (2007) BarleyBase/PLEXdb. Methods Mol Biol 406:347–363

Yano K, Imai K, Shimizu A, Hanashita T (2006) A new method for gene discovery in large-scale microarray data. Nucleic Acids Res 34:1532–1539

Zheng Q, Wang XJ (2008) GOEAST: a web-based software toolkit for Gene Ontology enrichment analysis. Nucleic Acids Res 36: W358–W363

Zimmermann P, Hirsch-Hoffmann M, Hennig L, Gruissem W (2004) GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. Plant Physiol 136:2621–2632

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