Large-Scale Identification of Expressed Sequence Tags Involved in Rice and Rice Blast Fungus Interaction[1]

Chatchawan Jantasuriyarat, Malali Gowda, Karl Haller, Jamie Hatfield, Guodong Lu, Eric Stahlberg, Bo Zhou, Huameng Li, HyRan Kim, Yeisoo Yu, Ralph A. Dean, Rod A. Wing, Carol Soderlund, and Guo-Liang Wang*

Department of Plant Pathology (C.J., M.G., G.L., B.Z., H.L., G.-L.W.) and Ohio Supercomputer Center (E.S.), The Ohio State University, Columbus, Ohio 43210; Arizona Genomics Computational Laboratory, BIO5 Institute (K.H., J.H., C.S.) and Arizona Genomics Institute, Department of Plant Sciences (H.K., Y.Y., R.A.W.), University of Arizona, Tucson, Arizona 85721; and Fungal Genomics Laboratory, Department of Plant Pathology, North Carolina State University, Raleigh, North Carolina 27695 (R.A.D.)

To better understand the molecular basis of the defense response against the rice blast fungus (*Magnaporthe grisea*), a large-scale expressed sequence tag (EST) sequencing approach was used to identify genes involved in the early infection stages in rice (*Oryza sativa*). Six cDNA libraries were constructed using infected leaf tissues harvested from 6 conditions: resistant, partially resistant, and susceptible reactions at both 6 and 24 h after inoculation. Two additional libraries were constructed using uninoculated leaves and leaves from the lesion mimic mutant *spl11*. A total of 68,920 ESTs were generated from 8 libraries. Clustering and assembly analyses resulted in 13,570 unique sequences from 10,934 contigs and 2,636 singletons. Gene function classification showed that 42% of the ESTs were predicted to have putative gene function. Comparison of the pathogen-challenged libraries with the uninoculated control library revealed an increase in the percentage of genes in the functional categories of defense and signal transduction mechanisms and cell cycle control, cell division, and chromosome partitioning. In addition, hierarchical clustering analysis grouped the eight libraries based on their disease reactions. A total of 7,748 new and unique ESTs were identified from our collection compared with the KOME full-length cDNA collection. Interestingly, we found that rice ESTs are more closely related to sorghum (*Sorghum bicolor*) ESTs than to barley (*Hordeum vulgare*), wheat (*Triticum aestivum*), and maize (*Zea mays*) ESTs. The large cataloged collection of rice ESTs in this study provides a solid foundation for further characterization of the rice defense response and is a useful public genomic resource for rice functional genomics studies.

Rice (*Oryza sativa*) is one of the most important staple food crops for more than one-half of the world's population. Rice blast fungus (*Magnaporthe grisea*) is a major constraint in rice production and is a serious threat to food security worldwide (Zeigler, 1998). Over the last three decades, race-specific resistance in many newly developed cultivars has frequently failed within a few years as a result of the high variability in the pathogen population. Development of rice cultivars with durable resistance is one of the main objectives in rice-breeding programs. It is well known that plants have evolved an array of defense mechanisms to combat invasion by plant pathogens. A thorough understanding of the molecular response mechanisms against rice blast will undoubtedly aid in the design of novel strategies to engineer durably resistant rice cultivars.

Although mapping of over 25 major resistance genes and many quantitative trait loci, as well as the cloning of 2 resistance genes, has advanced our knowledge regarding the genetic mechanisms of disease resistance (Wang and Leung, 1998; Wang et al., 1999; Bryan et al., 2000, Zhuang et al., 2002), the molecular basis of the defense response to rice blast remains poorly understood. In addition to the genetic approach, a direct assessment of the biochemical and physiological changes during disease infection has been used to identify genes involved in defense pathways in many plants. It is hoped that the manipulation of these genes may lead to the generation of broad-spectrum-resistant rice plants to rice pathogens. In the last several years, many defense-related genes have been isolated using reverse transcription-PCR, suppression subtractive hybridization (SSH), and cDNA library differential screening methods. For example, using both cDNA differential screening and SSH methods, Xiong et al. (2001) identified 56 defense genes that were responsive to blast infection and to treatment with benzothiazole and jasmonic acid. Using SSH, we identified 47 genes that are either induced or suppressed during the early stages of the defense response (from 12–24 h after inoculation) in a line carrying the broad-spectrum-resistant gene *Pi9*(*t*) (Lu et al., 2004). Among them, some were differentially expressed in resistant and susceptible plants after infection. While valuable, the available expression data are limited.

[1] This work was supported by the National Science Foundation Plant Genome Research Project (DBI no. 0115642 to R.A.D., G.-L.W., R.A.W., and C.S.).

* Corresponding author; e-mail wang.620@osu.edu; fax 614–292–4455.

www.plantphysiol.org/cgi/doi/10.1104/pp.104.055624.

Genomic approaches for identification of expressed genes, such as expressed sequence tag (EST; Adams et al., 1995), serial analysis of gene expression (SAGE; Velculescu et al., 1995), and massively parallel signature sequencing (MPSS; Brenner et al., 2000), have been widely used in genome-wide gene expression studies in various organisms. SAGE and MPSS are two powerful tools for deep transcriptome analysis and have been developed to evaluate the expression patterns of thousands of genes in a quantitative manner without prior sequence information (Velculescu et al., 1995; Brenner et al., 2000). However, complicated cloning procedures involved in the SAGE and MPSS library construction have inhibited the wide use of these two methods in plant species (Gowda et al., 2004). EST sequencing was the first method used for rapid identification of expressed genes (Adams et al., 1995). It has been employed to identify the genes that are expressed in various tissues, cell types, or developmental stages (Michalek et al., 2002; Ogihara et al., 2003; Ronning et al., 2003). In addition, the availability of cDNA sequences has accelerated further molecular characterization of interesting genes and provided sequence information for microarray design and genome annotation.

In this study, we used large-scale EST sequencing for gene expression profiling at early infection stages in rice and rice blast fungus interaction. We constructed six cDNA libraries using mRNA isolated from rice blast fungus-infected leaf tissues of resistant, partially resistant, and susceptible reactions and two cDNA libraries from noninfected leaf tissues and leaves from the rice lesion mimic mutant *spl11* (Zeng et al., 2002, 2004). A total of 68,920 EST sequences from 8 libraries were generated, from which 13,570 unique sequences were identified. These sequences were deposited in the National Center for Biotechnology (NCBI) GenBank and are displayed in our project database called *Magnaporthe grisea Oryza sativa* (MGOS; www.mgosdb.org). We performed extensive analysis of the ESTs derived from the eight cDNA libraries using a variety of computational methods. This study not only provides information on the expression patterns of defense genes in rice blast fungus-infected tissues, but also offers another genomic resource to the rice community for functional analysis of any genes in the collection.

RESULTS

cDNA Library Construction, EST Sequencing, and Data Analysis

We harvested rice blast fungus-infected leaf tissues 6 and 24 h after inoculation because the majority of the rice blast spores start to geminate on rice leaves about 6 h after inoculation and the majority of appressoria start to penetrate into rice epidermal cells 24 h after inoculation (Zeigler et al., 1994). Six unidirectional cDNA libraries were constructed using mRNA isolated from infected leaf tissues of resistant, partially resistant, and susceptible reactions at 6 and 24 h after inoculation with 3 different rice blast isolates (Table I). Two additional libraries, uninoculated water control and lesion mimic mutant *spl11* (Zeng et al., 2002, 2004), were also constructed. The insert size of over 20 individual clones randomly chosen from each library averaged 1.1 to 2.1 kb. Twenty 384-well plates/library were randomly picked for DNA sequencing. Sequencing from both ends was performed for the majority of the clones in all libraries. A total of 68,920 ESTs were generated and analyzed from the 8 libraries (Table II).

Clustering and assembly of these ESTs resulted in a total of 13,570 unique sequences with 10,934 tentative consensus sequences (contigs) and 2,636 singleton ESTs (Table II). The percentage of unique sequences in each library ranged from 24% to 46% (Table II). The OSIIEa library (lesion mimic library) has the lowest rate (24%) due to the high frequency of contig 03596_02 (2,494 copies). Sequence analysis indicated that this contig is highly homologous to the human U2 snRNP auxiliary factor large subunit (Hodges and Beggs, 1994). The EST sequences, contig alignments, chromosome location, and the ability to BLASTn search a sequence against the ESTs or search the ESTs against SwissProt or the nonredundant database are available to the scientific community via the MGOS database Web site (http://www.mgosdb.org). All of the EST sequences are available from NCBI GenBank (accession nos. CB617709–CB686047 and CX727819–CX728959).

Induction and Suppression of Rice Genes in Resistant and Susceptible Reactions to Rice Blast Fungus

Identification of unique EST sequences from the control, resistant, and susceptible libraries allows us to

Table I. *cDNA libraries and tissue sources for sequences described in this study*

Library ID	Cultivar	Tissue	Treatment	Disease Reaction
OSIIEa	Rice cv IR68	Leaf (spontaneous lesions)	Lesion mimic *spl11*	No
OSIIEb	Rice cv IR36	Leaf (24 h after inoculation)	Rice blast (PO6-6-3)	Partially resistant
OSJNEa	Rice cv Nipponbare	Leaf (6 h after inoculation)	Rice blast (Che86061)	Susceptible
OSJNEb	Rice cv Nipponbare	Leaf (24 h after inoculation)	Rice blast (Che86061)	Susceptible
OSJNEc	Rice cv Nipponbare	Leaf (6 h after inoculation)	Rice blast (C9240-1)	Resistant
OSJNEd	Rice cv Nipponbare	Leaf (24 h after inoculation)	Rice blast (C9240-1)	Resistant
OSJNEe	Rice cv Nipponbare	Leaf (24 h after inoculation)	Rice blast (70-15)	Partially resistant
OSJNEf	Rice cv Nipponbare	Leaf (24 h after water spray)	Mock (water) control	No

Table II. *Summary of number of EST sequences, contigs, and singletons in eight rice cDNA libraries*

Library ID	No. of 5′-3′ End EST Sequences (T7 primer)	No. of 3′-5′ End EST Sequences (Sp6 primer)	Total No. of EST Sequences	No. of Contigs (Percent in Parentheses)	No. of Singletons (Percent in Parentheses)	Unique Sequences (Percent in Parentheses)
OSIIEa	4,512	4,423	8,935	1,481 (93)	643 (07)	2,124 (24)
OSIIEb	4,836	4,651	9,487	2,472 (89)	1,091 (11)	3,563 (38)
OSJNEa	3,748	1,643	5,391	1,241 (77)	1,251 (23)	2,492 (46)
OSJNEb	5,374	5,135	10,509	2,655 (92)	829 (08)	3,484 (33)
OSJNEc	4,075	3,843	7,918	1,944 (89)	893 (11)	2,837 (36)
OSJNEd	4,586	4,509	9,095	2,509 (92)	766 (08)	3,275 (36)
OSJNEe	4,919	4,705	9,624	2,558 (89)	1,037 (11)	3,595 (37)
OSJNEf	4,162	3,799	7,961	1,972 (85)	1,163 (15)	3,135 (39)
Total	36,212	32,708	68,920	10,934 (76)	2,636 (04)	13,570

identify common and unique sets of expressed genes among the three libraries. As indicated in Figure 1, a total of 3,135, 3,275, and 3,484 unique ESTs were present in the control, resistant, and susceptible libraries at 24 h after inoculation, respectively. Surprisingly, only 390 unique ESTs were present in all 3 libraries. When comparing the ESTs from the control library to the ESTs from the susceptible and resistant libraries, only 25% of ESTs are shared between them and up to 63% of ESTs in each library are library specific. These results indicate that gene expression in the resistant and susceptible reactions was reprogrammed significantly at 24 h post-blast infection. The difference in the expression profiles of some defense genes between the resistant and susceptible reactions at this time point may contribute to the outcome of the disease phenotype at a later stage of infection.

Genes highly induced or suppressed in the resistant and susceptible libraries were identified by comparing the number of ESTs in the corresponding contigs in each library. The top 10 genes in the resistant and susceptible conditions and their putative functions are listed in Table III. Several defense-related genes were induced in both resistant and susceptible reactions, such as the β-glucanase and Phe ammonia lyase genes. Interestingly, we identified several photosynthesis-related genes that were suppressed in both resistant and susceptible reactions. A similar result was also reported by Matsumura et al. (2003), who observed that several photosynthetic genes were suppressed by *Phytophthora infestans* elicitor (INF1) as early as 1 h after the treatment.

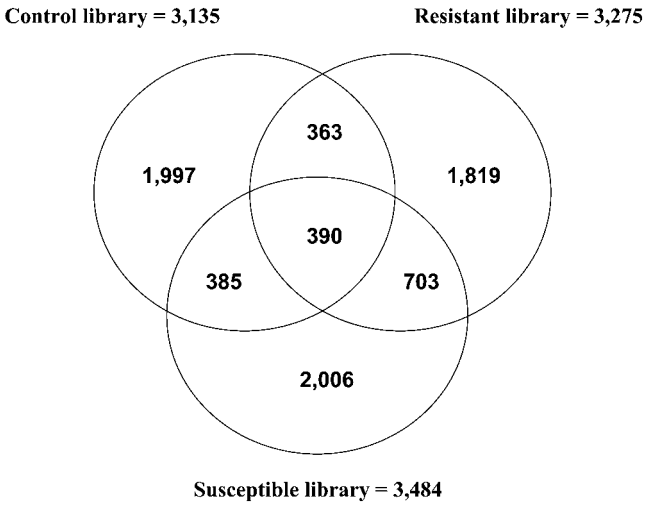

Figure 1. Overlapping of unique rice EST sequences from control, resistant, and susceptible libraries at 24 h after inoculation with rice blast fungus using the advance search function under the RICE EST PAVE page on MGOS Web site.

Analysis of Sequence Origin in the Rice Blast-Challenged Libraries

To identify the ESTs derived from rice blast fungus, we aligned all the rice ESTs against the 24,317 rice blast fungus ESTs deposited in the MGOS database. The following criteria were used in stand-alone BLASTn comparison: (1) at least 21-bp exact match; (2) matching length $\geq$100 bp; (3) DNA identity $\geq$95%; and (4) E-value $<$ 1E-20. BLAST search indicated that only four sequences showed high sequence similarity to rice blast fungus ESTs. These results suggested that there were not substantial amounts of pathogen ESTs among the cDNA libraries. The low numbers of the rice blast fungus sequences in the libraries could be due to early leaf tissue harvesting time (6 and 24 h after inoculation), as most of the blast spores just started to penetrate the rice leaf epidermal cells at that time.

Functional Classification of ESTs

The eukaryotic orthologous groups (KOGs) were constructed for a phylogenetic classification based on predicted proteins encoded in seven eukaryotic genomes: three animals (the nematode *Caenorhabditis elegans*, the fruit fly *Drosophila melanogaster*, and *Homo sapiens*), one plant, Arabidopsis (*Arabidopsis thaliana*), two fungi (*Saccharomyces cerevisiae* and *Schizosaccharomycea pombe*), and the intracellular microsporidian parasite *Encephalitozoon cuniculi* (Tatusov et al., 2003). We used KOGs for gene functional classification of our EST collection. A total of 68,920 ESTs from 8

Table III. *Putative functions of top 10 highly induced and suppressed genes in the resistant and susceptible libraries compared to that in the control libraries*

Contig ID	No. of ESTs in Infected Libraries	No. of ESTs in Control Library	Ratio of ESTs between Libraries	Putative Function/Homology	E-Value
Genes Induced in Resistant Libraries					
00001_0083	202	1	202:1	β-Glucanase	10^{-135}
01297_01	56	0	56:0	No hit	
00001_0147	43	1	43:1	Aquaporin (tonoplast intrinsic protein 2.1)	10^{-93}
00001_0755	39	0	39:0	No hit	
00001_0967	37	1	37:1	Aldehyde dehydrogenase	0.0
00001_0111	33	1	33:1	Heat shock protein 82	0.0
00001_0649	91	3	91:3	Stem-specific protein	10^{-23}
00398_01	29	0	29:0	No hit	
00035_03	26	0	26:0	Ser-glyoxylate aminotransferase	10^{-111}
00001_0328	52	2	52:2	Isovaleryl-CoA dehydrogenase	0.0
Genes Suppressed in Resistant Libraries					
00001_0057	1	36	1:36	Ribulose bisphosphate carboxylase, C	10^{-88}
00001_0631	0	22	1:22	PSI reaction center subunit psaK	10^{-62}
00562_02	1	22	1:22	Plastocyanin	10^{-62}
00001_0990	2	44	2:44	Chlorophyll *a*/*b*-binding protein	10^{-152}
00001_0822	0	18	0:18	Hypothetical protein	10^{-15}
06576_01	1	16	0:16	Carbonic anhydrase	10^{-106}
00001_0304	0	15	1:15	PSI reaction center subunit V	10^{-52}
00001_0201	0	15	0:15	No hit	
00001_0835	0	14	0:14	Thioredoxin M-type	10^{-86}
01910_01	0	13	0:13	No hit	
Genes Induced in Susceptible Libraries					
00001_0083	141	1	141:1	Glucan endo-1,3-β-glucosidase	10^{-135}
04381_01	330	7	330:7	Phe ammonia lyase	0.0
00001_1509_01	39	1	39:1	Heat shock protein 82	0.0
00001_0754	37	0	37:0	No hit	
000521_01	33	1	33:1	No hit	
00026_039	31	0	31:0	Hypothetical protein	10^{-25}
00001_0849	31	1	31:1	*S*-Adenosyl Met decarboxylase	10^{-179}
04481_01	30	1	30:1	No hit	
00001_0147	29	01	29:1	Tonoplast intrinsic protein	2.1×10^{-93}
00001_0001	53	2	53:2	No hit	
Genes Suppressed in Susceptible Libraries					
00001_0521	0	26	0:26	Ribulose bisphosphate carboxylase, subunit C	10^{-87}
00001_0631	0	22	0:22	PSI subunit X	10^{-62}
07222_01	0	22	0:22	PSII 10-kD polypeptide	10^{-41}
02552_01	0	21	0:21	Light-regulated protein	10^{-41}
07167_01	1	20	1:20	Ribulose bisphosphate carboxylase/oxygenase	0.0
00040_04	2	37	2:37	DNA methyltransferase 1-associated protein	10^{-12}
02546_01	0	16	0:16	No hit	
06576_01	0	16	0:16	Carbonic anhydrase	10^{-106}
00001_0730	4	61	4:61	Chlorophyll *a*/*b*-binding protein CP26	10^{-116}
00001_0304	0	15	0:15	PSI reaction center subunit V	10^{-51}

cDNA libraries were processed by the KOG software program to perform the prediction of putative functional classification of individual proteins. ESTs were grouped according to functional categories and are summarized in Figure 2. Of the 68,920 sequences, 68.74% of the ESTs were assigned to 9 putative gene functional categories: (1) transcription and translation; (2) cell cycle control, cell division, and chromosome partitioning; (3) defense mechanisms; (4) signal transduction mechanisms; (5) cytoskeleton and cell mobility; (6) post-translational modification, protein turnover, and chaperone; (7) nucleotide, amino acid, and coenzyme metabolism; (8) carbohydrate and lipid metabolism; and (9) poorly characterized. Among the 68.74% of the ESTs assigned to the functional categories, 26.77% are in the poorly characterized group, i.e. they have a match in the KOG database but do not have clear function. Therefore, only 41.97% of the ESTs could be assigned a putative function.

Furthermore, ESTs from each individual library were also analyzed using the KOG program. Since different numbers of clones were sequenced in the eight cDNA libraries, normalization was performed prior to gene functional category comparison. The percentage of gene function categories in each library was compared between libraries (Table IV). The percentages of the ESTs in all gene function categories were higher in all pathogen-challenged libraries (OSJNEa, OSJNEb,

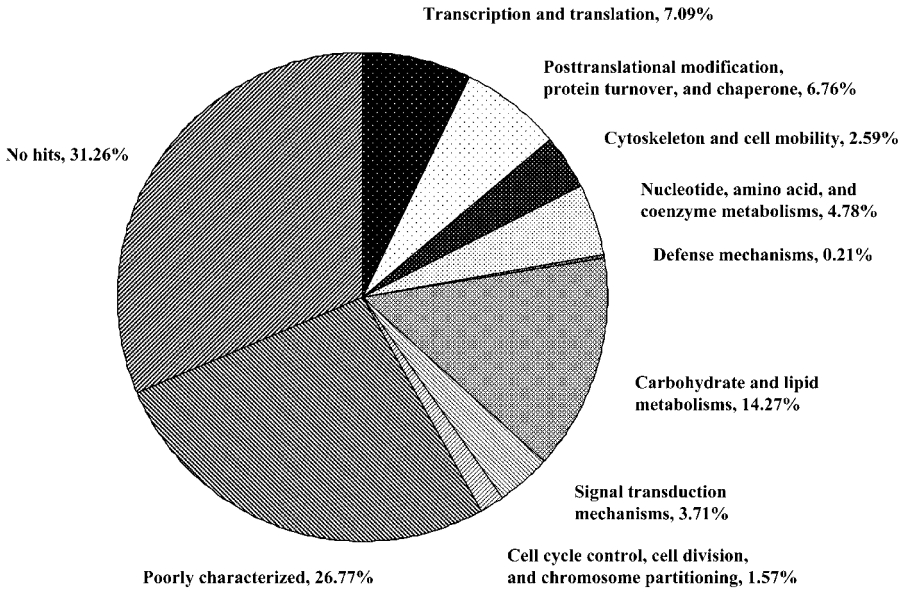

Figure 2. Functional categorization and percentage of rice ESTs based on their putative function using the KOGs protein database. A total of 68,920 EST sequences from 8 cDNA libraries were submitted to the KOG program to predict the putative functional classification of individual proteins.

OSJNEc, OSJNEd, and OSJNEe) compared to the control (OSJNEf) library. However, only defense and signal transduction mechanism and cell cycle control, cell division, and chromosome partitioning categories showed a statistically significant increase in the pathogen-challenged libraries of resistant, partially resistant, and susceptible reactions. For example, in the defense mechanism gene category, the percentage of genes increased from 0.19% in the noninoculated library to 0.30% and 0.32% in the resistant libraries at 6 and 24 h after rice blast inoculation, respectively. Chi-square tests showed the increases are statistically significant. In the signal transduction mechanism gene category, the percentage of genes increased from 2.40% in the noninoculated library to 4.21% and 4.23% in the resistant libraries at 6 and 24 h after rice blast inoculation, respectively. Chi-square tests showed that the increases at both 6 and 24 h are significant. In the cell cycle control, cell division, and chromatin structure category, the percentage of genes doubled from 1.38% in the noninoculated library to 8.83% and 7.38% in the susceptible and 6.17% and 4.80% in the resistant libraries at 6 and 24 h after rice blast inoculation, respectively. Chi-square tests showed that increases at both 6 and 24 h are highly significant.

Two indica rice libraries, lesion mimic mutant *spl11* (OSIIEa) and partially resistant reaction (OSIIEb), showed similar percentages of EST in most functional categories to japonica rice libraries. The one exception is the cell cycle control, cell division, and chromosome partitioning category, where the percentages of EST in the OSIIEa and OSIIEb libraries were 12.43% and 11.58%, which were significantly higher than those in the japonica rice libraries (Table IV). The lesion mimic mutant *spl11* library (OSIIEa) displayed a significantly different percentage pattern from all other libraries. In particular, a high percentage of ESTs (18.95%) was observed in the transcription and translation category, which was more than double those in other libraries (Table IV) due to contig 03596_02 with 2,494 ESTs.

Table IV. *KOG analysis showing percentage of EST in each library based on gene functional categories*

Gene Functional Category	Percentage of EST in Each Library							
	OSIIEa	OSIIEb	OSJNEa	OSJNEb	OSJNEc	OSJNEd	OSJNEe	OSJNEf
Transcription and translation	18.95	6.89	3.67	5.57	4.95	4.73	5.46	5.12
Cell cycle control, cell division, and chromosome partitioning	12.43	11.58	8.83	7.38	6.17	4.80	3.18	1.38
Defense mechanisms	0.11	0.15	0.13	0.23	0.30	0.32	0.22	0.19
Signal transduction mechanisms	3.42	4.26	2.82	3.44	4.21	4.23	4.42	2.40
Cytoskeleton and cell mobility	2.82	5.04	2.50	3.86	3.49	4.15	3.46	2.69
Posttranslational modification, chaperones, and protein turnover	6.60	7.39	4.75	6.57	7.48	6.92	8.03	5.38
Nucleotide, amino acid, and coenzyme metabolism	5.23	3.89	3.67	5.31	6.43	4.34	5.40	3.48
Carbohydrate and lipid metabolism	13.26	9.96	10.00	18.72	17.34	14.98	15.17	12.65
Poorly characterized	19.89	26.28	15.15	28.87	34.11	28.84	27.50	29.62
No hits	28.86	33.39	55.85	26.23	20.32	29.88	28.54	37.09

Validation of EST Expression Level by Northern-Blot Analysis

To experimentally confirm the level of defense gene expression based on differential EST representation in the libraries, five EST clones were selected from both the defense and signal transduction mechanism categories for northern-blot analysis. Results showed that all 5 clones selected from the defense mechanism category had a strong induction at 12 or 24 h after blast inoculation. The expression level decreased to steady-state levels at 72 h after inoculation (Fig. 3). Similarly, 4 clones selected from the signal transduction mechanism category showed a strong induction at 12 or 24 h after inoculation and decreased to steady-state levels at 72 h after inoculation. One clone from this same category, which has a sequence similarity to Ser/Thr protein phosphatase (OSJNEb08D18), exhibited suppressed expression between 6 and 24 h (Fig. 3). It is worth noting that most of the genes did show some visible difference in the resistant and susceptible reactions. Taken together, results from northern blots generally corroborate the frequency of the selected ESTs in the cDNA libraries.

EST Frequency Clustering Analyses to Identify Broad Patterns of Gene Expression

To assess the relatedness of each library in terms of gene expression patterns, we performed a clustering analysis based on EST abundance (Ewing et al., 1999). First, we compiled 10,934 contigs into a matrix file containing the frequency of ESTs corresponding to each contig in the library that represents different disease reactions. The *R* statistic described by Stekel et al. (2000) was used to identify the most highly significant differences in EST abundance for each contig among the libraries. To limit the analysis to those genes that were the most differentially expressed within the tissues, only contigs with $R > 15$ (434 in total) were used for hierarchical clustering analysis. This value provides a 99.9% true positive rate (Stekel et al., 2000). From hierarchical analysis, the clustering of all eight libraries was consistent with their disease reactions. Six pathogen-challenged libraries and the control library were clustered together with lesion mimic mutant *spl11* as an outer group (Fig. 4). Within the cluster of those challenged libraries, the resistant library (OSJNEd) at 24 h after inoculation was closely placed with the partially resistant library (OSJNEe) at 24 h after inoculation, and the susceptible library (OSJNEc) at 6 h after inoculation was closely placed with the susceptible library (OSJNEb) at 24 h after inoculation. The frequency clustering analysis used to identify broad patterns of gene expression could be grouped into 9 major clusters from A to I as shown in Figure 4. Each gene cluster represents different patterns of gene expression. For example, cluster E is composed of genes that were highly expressed in lesion mimic mutant *spl11* library and cluster B is composed of genes that were highly expressed in susceptible reaction at 24 h after inoculation.

The second method, *k*-means clustering, was performed to identify biologically relevant clusters of genes according to the procedures described by Quackenbush (2001). In this analysis, we used a dataset

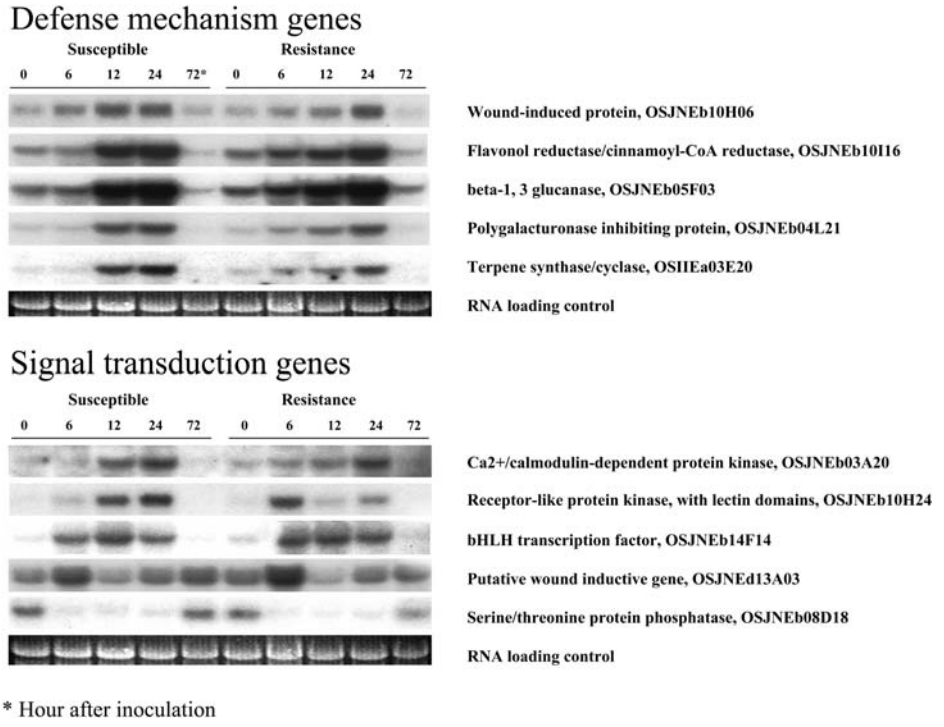

Figure 3. Northern-blot confirmation of 10 EST clones selected from the defense and signal transduction mechanism categories. About 10 μg of total RNA from susceptible and resistant reactions of Nipponbare plants at 6, 12, 24, and 72 h after inoculation were used in northern-blot analysis. For the susceptible reaction, Nipponbare plants were inoculated with rice blast strain Che86061. For the resistant reaction, Nipponbare plants were inoculated with rice blast strain C9240-1. RNA loading of each sample was verified by the intensity of ribosomal RNA bands on the agarose gel.

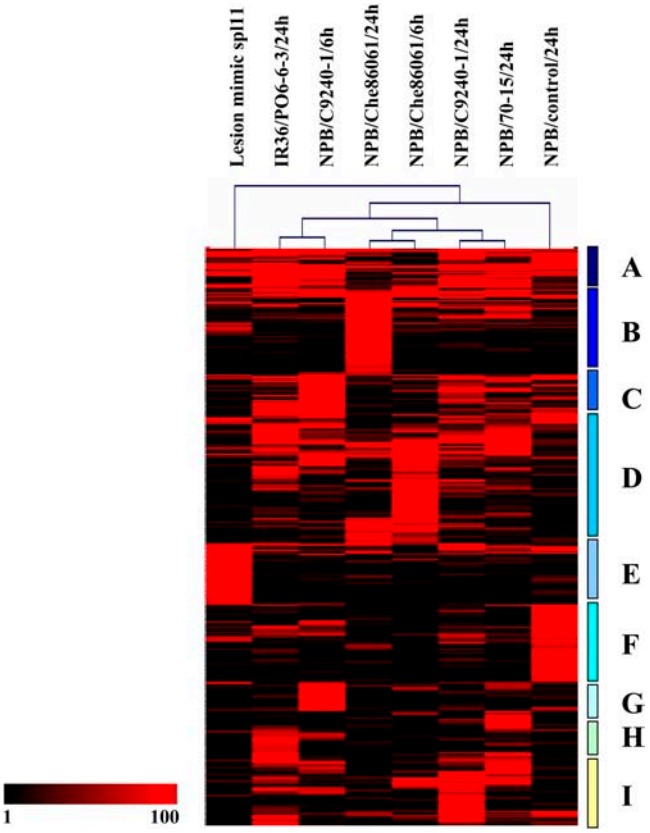

Figure 4. Hierarchical clustering analysis of differentially expressed transcripts. Contigs with an $R > 15$ (434 in total) were used for hierarchical clustering analysis. A frequency of zero is indicated by black and a frequency increase is indicated by increasing intensities of red. The library information is summarized in Table I.

including 738 contigs that have a minimum of 6 ESTs comprising the contig and using $R > 12$. Results indicated that nine clusters were found to be optimally predictive for the *k*-means-clustering algorithm, which was consistent with the results obtained through hierarchical clustering (data not shown).

Comparison of Our ESTs to the Japanese Rice Full-Length cDNA Sequences and TIGR Rice Gene Tentative Contigs

When we performed EST assembly using 32,127 rice full-length cDNA collections on the KOME database (Kikuchi et al., 2003) as the reference, we found that a total of 7,748 ESTs from our libraries were not present in the collection. Furthermore, matching our ESTs to The Institute for Genomic Research (TIGR) rice gene collection (282,117 ESTs; release 15.0, May 12, 2004) showed that a total of 4,319 ESTs from our collection were not matched to TIGR rice genes. This number represents 17% of our total ESTs. These results indicated that our EST sequencing project identified a large number of new genes, most of which might be involved in the defense response to rice blast.

Comparison of Our ESTs to Other Plant EST Sequences

To investigate how many rice ESTs are highly homologous to other plant ESTs in the public databases, we performed a comparative matching analysis of our rice ESTs to Arabidopsis, barley (*Hordeum vulgare*), sorghum (*Sorghum bicolor*), wheat (*Triticum aestivum*), and maize (*Zea mays*) ESTs in TIGR gene indices. In general, rice ESTs showed a higher similarity to ESTs of the grass species than to those of the dicot model plant, Arabidopsis. The percentage of rice ESTs matched to Arabidopsis ESTs was 2.4% and 8.4% when DNA sequence identity was ≥90% and ≥80% (Table V), respectively. In contrast, the percentage of rice ESTs matched to barley, sorghum, wheat, and maize ESTs ranged from 31.9% to 63.2% when the DNA sequence identity was ≥80% (Table V). Within the grass species, our rice ESTs had a significantly higher percent similarity to sorghum ESTs (63.2%) than to barley (38.2%), wheat (35.1%), and maize ESTs (31.9%) at sequence identity ≥80%. To confirm this result, we conducted a similarity search of the entire set of rice TIGR EST database (88,765) against other cereal TIGR EST databases. When DNA sequence identity was at ≥80%, the percentage of rice ESTs matching sorghum, barley, wheat, and maize was 43.98%, 29.31%, 25.99%, and 23.87%, respectively, corroborating the result when our rice ESTs were used in the analysis (Table V). However, our results contradict the findings of Kellogg (2001) on phylogenetic structure in the grass family. They concluded that rice is more closely related to wheat and barley than to maize and sorghum.

DISCUSSION

To understand the molecular basis of host resistance to the rice blast fungus, we monitored the transcription changes at early infection stages in rice using the EST sequencing approach. A large collection of 68,920 EST sequences was generated from 8 cDNA libraries using leaf tissues collected from a blast-challenged, -unchallenged, and a lesion mimic mutant. Through a series of sequence-clustering and assembly-processing steps, a total of 13,570 unique sequences were obtained. From the sequence analysis, a large number of genes that were highly induced or suppressed in resistant and susceptible conditions were identified. Among them, the percentages of genes in the defense and signal transduction mechanism and cell cycle control, cell division, and chromosome partitioning categories were significantly increased after blast infection. To date, this is the largest EST collection generated from a single plant-pathogen interaction in plants. Therefore, the sequences reported in this study provide a significant improvement in our understanding of the rice defense mechanism to the rice blast fungus and will streamline the community effort in elucidating the functions of many defense response genes in rice. The ESTs are available from our MGOS

Table V. *Comparative matching of the rice ESTs isolated in this study to the ESTs of Arabidopsis, barley, sorghum, wheat, and maize collected in TIGR gene indices*

The criteria for stand-alone BLASTn were: (1) exact-match bp = 11; (2) E-value ≤ 1E5; and (3) identity ≥80% and 90% at DNA sequence level.

TIGR Gene Indices	No. of ESTs Matched to TIGR Gene Index (%)			
	MGOS ESTs (13,570)		TIGR ESTs (88,765)	
	Identity ≥80%	Identity ≥90%	Identity ≥80%	Identity ≥90%
Arabidopsis[a]	1,145 (8.4)	1,145 (2.4)	10,716 (12.07)	2,952 (3.33)
Barley[b]	5,182 (38.2)	1,295 (9.5)	26,015 (29.31)	8,008 (9.02)
Sorghum[c]	8,569 (63.2)	2,155 (15.9)	39,041 (43.98)	11,480 (12.93)
Wheat[d]	4,768 (35.1)	1,347 (9.9)	23,070 (25.99)	7,707 (8.68)
Maize[e]	4,324 (31.9)	953 (7.0)	21,187 (23.87)	6,069 (6.84)

[a]Arabidopsis gene index release 11.0 (Jan. 12, 2004) composed of 45,683 unique sequences. [b]Barley gene index release 8.0 (Jan. 9, 2004) composed of 49,190 unique sequences. [c]Sorghum gene index release 8.0 (May 11, 2004) composed of 39,148 unique sequences. [d]Wheat gene index release 8.0 (Dec. 25, 2003) composed of 123,807 unique sequences. [e]Maize gene index release 14.0 (Dec. 23, 2003) composed of 56,364 unique sequences.

database (www.mgosdb.org) and the cDNA clones may be ordered from the Arizona Genomic Institute BAC/EST Resource Center (http://www.genome.arizona.edu).

To reveal what types of genes are included in our rice EST collection, the KOG program was used to predict the putative function of the encoded proteins. Sixty-eight percent of the ESTs were predicated to have known functions and were classified into nine functional categories. The defense and signal transduction mechanism and cell cycle control, cell division, and chromosome partitioning categories had a higher proportion in the resistant and susceptible libraries than that found in the control library. This result was not unexpected as it has been shown in many cases that defense and signal transduction mechanism genes are induced or repressed during the host-pathogen interaction (Kim et al., 2001; Lu et al., 2004). We confirmed this result using northern-blot analysis to check the expression level of five representative genes from each group. Many of these genes are known to be involved in pathogen-related response. For example, the wound-induced protein has been shown to be involved in *Pto*-mediated disease resistance in tomato (Ekengren et al., 2003). β-1,3-Glucanase is known as the pathogenesis-related protein PR-2 (Yamaguchi et al., 2002). Polygalacturonase-inhibiting proteins are plant cell wall proteins that protect plants from fungal invasion (Di Matteo et al., 2003). Receptor-like protein kinase and Ser/Thr protein phosphatase are important components of the signal transduction pathway in plants (Becraft, 2002). Induction of the expression was detected in all genes, except for one encoding a Ser/Thr protein phosphatase. Its expression was suppressed from 6 to 24 h after inoculation. The expression level of all 9 genes was induced as early as 6 h, reached their peak expression at 24 h, and returned to a steady-state level at 72 h after inoculation. Further functional analysis of these early responsive genes to rice blast may provide new insights into the molecular mechanisms of the host defense response.

Of the eight cDNA libraries reported in this study, six libraries were constructed from rice leaf tissues challenged with rice blast fungus. ESTs from these libraries represent the sum of the transcripts from both rice and rice blast pathogen, which are expressed during their interaction. This pool of isolated ESTs has been defined as the interaction transcriptome (Birch and Kamoun, 2000). The mixture of transcripts from the host and pathogen makes it difficult to identify the origin of each EST sequence. Consequently, it is likely that some EST sequences derived from our rice blast-infected libraries have pathogen origin, as was demonstrated in other studies (Qutob et al., 2000; Kim et al., 2001; Ronning et al., 2003). Many approaches have been used to identify the origin of EST sequences. One approach is to use GC content of the ESTs as the standard to separate the pathogen-derived sequences from the plant sequences. This is only possible when the two organisms have markedly different GC content. For example, in the case of *Phytophthora sojae* and soybean (*Glycine max*) interaction, the average GC content of soybean ESTs was 46%, whereas for *P. sojae* ESTs, it was 58%, which is different enough to be used to distinguish the origin of ESTs (Qutob et al., 2000). We found in this study that the ESTs from rice and rice blast fungus have a very small difference in their GC contents. Therefore, it was not possible to use the GC content as a parameter to distinguish the origin of the EST sequences. However, we were able to estimate the amount of rice blast fungus-derived ESTs by comparing all the ESTs to the rice blast fungus EST sequences available on our MGOS database. This revealed that a very small number of rice blast fungus-derived sequences were present in our libraries (only four clones). By contrast, Kim et al. (2001) reported that about 24% of ESTs from the compatible interaction cDNA library were rice blast fungus-derived sequences in which leaf tissues were harvested at 84, 96, and 120 h after rice blast infection. The criteria used in that study were not stringent (E-value < 1E-3) in which some of the rice genes highly homologous to rice blast fungal genes

might be identified as positive clones. Talbot et al. (1993) used gel-blot analysis to estimate the proportion of fungal and plant biomass present during rice blast fungal infection in rice. They found that about 10% of the biomass of infected rice leaves 72 h after inoculation were from rice blast fungus. Since our main objective is to understand the interaction between rice and rice blast fungus at the early stages, we collected leaf tissues at 6 and 24 h after the inoculation. At these 2 time points, most fungal spores have only just germinated (6 h) or begun to penetrate into rice epidermal cells (24 h). Therefore, only a small amount of ESTs in our libraries would be expected to be derived from cDNAs of the fungal pathogen.

The lesion mimic mutant *spl11* shows enhanced non-race-specific resistance to both rice blast fungus and *Xanthomonas oryzae* pv *oryzae* (*Xoo*; Zeng et al., 2002). In this study, the mutant *spl11* library showed a unique pattern of gene expression with one high-frequency contig that represented 28.0% of the library. This contig is highly similar to splicing factor U2 snRNP in humans (Hodges and Beggs, 1994). The function of this RNA-splicing factor in the *Spl11*-mediated signaling pathway is still unclear. As the *Spl11* gene has been cloned recently (Zeng et al., 2004), the role of the RNA-binding protein in the programmed cell death and defense response will be unraveled.

The phylogenic relationship among cereals has been investigated in the last decade using different molecular approaches. Many studies showed that extensive colinearity of genetic maps exists among cereals such as rice, barley, sorghum, maize, and wheat (Gale and Devos, 1998). These comparative analyses revealed significant conservation of gene content and order across cereal species that diverged from a common ancestor millions of years ago (Crepet and Feldman, 1991). Phylogenetic trees for displaying the relative order of speciation events showed that rice is more closely related to wheat and barley than to maize and sorghum (Kellogg, 2001). Most of the data for generating the phylogenetic trees were from mapping RFLP markers or selected conserved genes. In this study, we matched our rice ESTs to the ESTs of maize, sorghum, barley, and wheat in TIGR gene indices. Interestingly, the percentage match of rice ESTs to sorghum was much higher than those to wheat, maize, and barley. A similar result was obtained when the entire collection of the rice ESTs in TIGR gene index was used in the analysis. Although indirect evidence about the close relationship between rice and sorghum was observed by Close et al. (2004) from microarray hybridizations, our results clearly demonstrate that rice is more closely related to sorghum based on a comparative analysis of EST transcripts.

The completion of rice genome sequencing leads to new challenges in gene annotation and gene functional identification. ESTs and full-length cDNA clones are ideal materials for gene annotation and comprehensive gene function analysis at the transcriptional level (Kikuchi et al., 2003). In this study, we generated more than 7,748 novel ESTs based on the comparison of our ESTs to the KOME database (Kikuchi et al., 2003). These genes will be readily useful for the rice genome annotation. In addition, many of these genes may be related to defense responses to other rice pathogens. They are ideal starting materials for scientists who are interested in conducting detailed studies on selected genes at molecular and biochemical levels.

MATERIALS AND METHODS

Plant Materials and Rice Blast Inoculation

Two rice (*Oryza sativa*) varieties, L. subsp. *japonica* cv Nipponbare from Dr. T. Sasaki, Japan, and L. subsp. *indica* cv IR36 from H. Leung, the Philippines, and four rice blast isolates (C9240-1, Che8606, 70-15, and PO6-6) were used in this study. In the resistant reaction, Nipponbare was inoculated with avirulent rice blast strain C9240-1 from the Philippines. In the susceptible reaction, Nipponbare was inoculated with the virulent rice blast strain Che8606 from China. In the partially resistant reaction, Nipponbare was inoculated with the rice blast strain 70-15 from R. Dean. IR36 was inoculated with rice blast strain PO6-6 from the Philippines, which yields a partially resistant reaction. For the control, Nipponbare was inoculated with water. Three-week-old rice plants were inoculated with a spore suspension of rice blast at 1×10^5 spores/mL. The inoculated plants were placed in a plastic box (covered tightly) in the dark for 24 h at 26°C, and leaf tissues were collected 6 and 24 h after inoculation. Leaves with visible lesion mimics were collected from the lesion mimic mutant *spl11* for RNA isolation.

RNA Isolation and cDNA Library Construction

Total RNA was extracted from leaf tissues using the TRIzol method (Invitrogen, San Diego) according to the instructions provided by the manufacturer. Poly(A^+) RNA purified from total RNA using the Qiagen mRNA purification kit (Qiagen, Valencia, CA) was used for cDNA synthesis. All the cDNA libraries were constructed using a cDNA construction kit from Stratagene (La Jolla, CA). cDNAs were cloned into the pBluescript II KS (+) vector (Stratagene) and transformed by electroporation into DH10B *Escherichia coli* cells (Invitrogen). About 7,500 cDNA clones from each library were randomly picked and stored in 20 386-well plates using freeze medium for long-term storage as described in Wang et al. (1995).

EST DNA Sequencing and Assembly of EST into Contigs

Plasmid DNA was isolated and purified from *E. coli* cultures by alkaline lysis, vacuum filtration, and anion-exchange chromatography using a high-throughput, 96-well format system (Qiagen). cDNA inserts were sequenced in both directions using a T7 primer (5′-TAATACGACTCACTATAGGG-3′) for 5′-3′ end sequencing, and an Sp6 primer (5′- GATTTAGGTGACACTATAG-3′) for 3′-5′ end sequencing. Automated cycle sequencing of DNA was carried out and products were resolved by gel electrophoresis (model 3730; Applied Biosystems, Foster City, CA). Raw EST sequence data were edited to remove vector and adaptor sequences, and low-quality sequences using the Lucy software program (Chou and Holmes, 2001). ESTs were clustered and aligned into contigs and singlets using PAVE (Program for Assembling and Viewing ESTs), developed at AGCol for the MGOS project. The current PAVE assembly uses the TGICL script (TIGR Gene Indices Clustering tool; Pertea et al., 2003) for clustering, CAP3 (Huang and Madan, 1999) for assembly, and a merge/split script that changes the EST content of contigs to ensure that 3′ and 5′ ESTs are in the same contigs. Auxiliary information is used, such as the KOME cDNAs, the rice genome sequence, and protein hits to support the merging of contigs. If the 3′ and 5′ do not overlap, they are connected with Ns. When 3′ and 5′ ESTs cannot be put in the same contig, the 2 contigs are put in the same cluster. Hence, the meaning of clusters in the current PAVE assembly is that each EST in the cluster has a mate in its own contig or a contig in the cluster (unless it does not have a mate). EST contigs and singlets were searched against the NCBI nonredundant protein database to provide a putative function.

EST Sequence Analysis

EST sequences of the blast-infected libraries were analyzed for their sequence of origin by comparing their GC content profile with that of the rice blast ESTs at the MGOS database. ESTs were functionally categorized in each library based on their putative function using the Cluster of Orthologous Groups database of proteins (http://www.ncbi.nlm.nih.gov/COG). Expression profiles from each library were compared, and the ESTs specifically induced or suppressed in each library were identified.

Gene expression analysis was performed with TIGR Multiple Experiment Viewer software (version 1.1; Quackenbush, 2001) by using transcript abundance in each contig in all 8 libraries. Only contigs that were composed of at least six ESTs were used for the cluster analysis. Hierarchical clustering (Eisen et al., 1998), with statistical support for the branches of clusters based on resampling the data, was performed.

Comparative matching of our rice ESTs to the rice full-length cDNA collection and TIGR gene indices (Quackenbush, et al., 2001) was performed. The 32,127 rice full-length cDNA collection (KOME database; Kikuchi et al., 2003), Arabidopsis (*Arabidopsis thaliana*) gene index release 11.0 (Jan. 12, 2004) composed of 45,683 unique sequences, barley (*Hordeum vulgare*) gene index release 8.0 (Jan. 9, 2004) composed of 49,190 unique sequences, sorghum (*Sorghum bicolor*) gene index release 8.0 (May 11, 2004) composed of 39,148 unique sequences, wheat (*Triticum aestivum*) gene index release 8.0 (Dec. 25, 2003) composed of 123,807 unique sequences, and maize (*Zea mays*) gene index release 14.0 (Dec. 23, 2003) composed of 56,364 unique sequences were downloaded for stand-alone BLASTn comparison. The following criteria were used in stand-alone BLASTn comparison with the KOME database and the rice TIGR gene index: (1) 21-bp exact match; (2) matching length $\geq$100 bp; (3) DNA identity $\geq$95%; and (4) E-value $<$ 1E-20. For comparison with ESTs from other plant species, the criteria for stand-alone BLASTn were: (1) exact-match bp = 11; (2) E-value $\leq$ 1E5; and (3) DNA identity $\geq$80% and 90%.

The EST analysis in this study was performed using the advance search feature of MGOS, which is modeled after the search feature in HarvEST (www.harvest.org). For example, the intersection of the control, resistant, and susceptible libraries was found by selecting these three libraries in the include column. For the first section of Table II, the genes induced in resistant libraries, we included library OSJNEc and OSJNEd and excluded OSJNEf. The other sections were obtained similarly, but selecting different libraries as appropriate.

Northern-Blot Analyses

Leaf samples for total RNA isolation were collected from 3-week-old seedlings of Nipponbare. Total RNA was isolated using the TRIzol method described above. Approximately 10 μg of glyoxylated total RNA per lane was fractionated in a 1.4% agarose gel and transferred to a Hybond-N^+ membrane (Amersham, Buckinghamshire, UK) according to the manufacturer's instructions. ^{32}P-labeled DNA probes were labeled with a Rediprime DNA-labeling system (Amersham). Northern-blot hybridization was carried out using standard procedures, described in Sambrook et al. (1989), and was repeated twice.

Sequence data from this article have been deposited with the EMBL/GenBank data libraries under accession numbers CB617709 to CB686047, and CX727819 to CX728959.

ACKNOWLEDGMENTS

We are grateful to all members of our laboratory for their assistance and discussion during the course of this work. Special thanks go to S. Stegalkina and R. Buell at TIGR for their valuable help in hierarchical clustering analysis, Baek Hie Nahm for help in KOG analysis, V. Pampamwar for implementing the MGOS search page, T. Close for help in downloading the barley EST sequences from the HarvEST database, and R. Nelson, Beth Haze, and M. Babu for their critical reading of the manuscript.

Received October 25, 2004; returned for revision January 25, 2005; accepted January 28, 2005.

LITERATURE CITED

Adams M, Kerlavage AR, Fleischmann RD, Fuldner RA, Bult CJ, Lee NH, Kirkness EF, Weinstock KG, Gocayne JD, White O, et al (1995) Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence. Nature **377:** 173–174

Becraft PW (2002) Receptor kinase signaling in plant development. Annu Rev Cell Dev Biol **18:** 163–192

Birch PRJ, Kamoun S (2000) Studying interaction transcriptomes: coordinated analyses of gene expression during plant-microorganism interactions. *In* R Wood, ed, New Technologies for Life Sciences: A Trends Guide. Elsevier Science, New York, pp 77–82

Brenner S, Johnson M, Bridgham J, Golda G, Lloyd DH, Johnson D, Luo S, McCurdy S, Foy M, Ewan M, et al (2000) Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. Nat Biotechnol **18:** 630–634

Bryan GT, Wu K-S, Farrall L, Jia Y, Hershey HP, McAdams SA, Faulk KN, Donaldson GK, Tarchini R, Valent B (2000) A single amino acid difference distinguishes resistant and susceptible alleles of the rice blast resistance gene *Pi-ta*. Plant Cell **12:** 2033–2045

Chou HH, Holmes MH (2001) DNA sequence quality trimming and vector removal. Bioinformatics **17:** 1093–1104

Close TJ, Wanamaker SI, Caldo RA, Turner SM, Ashlock DA, Dickerson JA, Wing RA, Muehlbauer GJ, Kleinhofs A, Wise RP (2004) A new resource for cereal genomics: 22K barley GeneChip comes of age. Plant Physiol **134:** 960–968

Crepet WL, Feldman GD (1991) The earliest remains of grasses in the fossil record. J Bot **78:** 1010–1014

Di Matteo A, Federici L, Mattei B, Salvi G, Johnson KA, Savino C, De Lorenzo G, Tsernoglou D, Cervone F (2003) The crystal structure of polygalacturonase-inhibiting protein (PGIP), a leucine-rich repeat protein involved in plant defense. Proc Natl Acad Sci USA **100:** 10124–10128

Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA **95:** 14863–14868

Ekengren SK, Liu Y, Schiff M, Dinesh-Kumar SP, Martin GB (2003) Two MAPK cascades, NPR1, and TGA transcription factors play a role in Pto-mediated disease resistance in tomato. Plant J **36:** 905–917

Ewing RM, Kahla AB, Poirot O, Lopez F, Audic S, Claverie JM (1999) Large-scale statistical analyses of rice ESTs reveal correlated patterns of gene expression. Genome Res **9:** 950–959

Gale MD, Devos KM (1998) Plant comparative genetics after 10 years. Science **282:** 656–659

Gowda M, Jantasuriyarat C, Dean R, Wang GL (2004) Robust-longSAGE: a substantially improved longSAGE method for gene discovery and transcriptome analysis. Plant Physiol **134:** 890–897

Hodges PE, Beggs JD (1994) RNA splicing U2 fulfills a commitment. Curr Biol **4:** 264–267

Huang X, Madan A (1999) CAP3: a DNA sequence assembly program. Genome Res **9:** 868–877

Kellogg EA (2001) Evolutionary history of the grasses. Plant Physiol **125:** 1198–1205

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

Kim S, Ahn IP, Lee YH (2001) Analysis of genes expressed during rice-*Magnaporthe grisea* interaction. Mol Plant Microbe Interact **14:** 1340–1346

Lu G, Jantasuriyarat C, Zhou B, Wang GL (2004) Isolation and characterization of novel defense response genes involved in compatible and incompatible interactions between rice and *Magnaporthe grisea*. Theor Appl Genet **108:** 525–534

Matsumura H, Reich S, Ito A, Saitoh H, Kamoun S, Winter P, Kahl G, Reuter M, Kruger DH, Terauchi R (2003) Gene expression analysis of plant host-pathogen interactions by SuperSAGE. Proc Natl Acad Sci USA **100:** 15718–15723

Michalek W, Weschke W, Pleissner KP, Graner A (2002) EST analysis in barley defines a unigene set comprising 4,000 genes. Theor Appl Genet **104:** 97–103

Ogihara Y, Mochida K, Nemoto Y, Murai K, Yamazaki Y, Shin-IT, Kohara Y (2003) Correlated clustering and virtual display of gene expression patterns in the wheat life cycle by large-scale statistical analyses of expressed sequence tags. Plant J **33:** 1001–1011

Pertea G, Huang X, Liang F, Antonescu V, Sultana R, Karamycheva S, Lee Y, White J, Cheung F, Parvizi B, et al (2003) TIGR gene indices

clustering tools (TGICL): a software system for fast clustering of large EST datasets. Bioinformatics **19:** 651–652

Quackenbush J (2001) Computational analysis of microarray data. Nat Rev Genet **2:** 418–427

Quackenbush J, Cho J, Lee D, Liang F, Holt I, Karamycheva S, Parvizi B, Pertea G, Sultana R, White J (2001) The TIGR gene indices: analysis of gene transcript sequences in highly sampled eukaryotic species. Nucleic Acids Res **29:** 159–164

Qutob D, Hraber PT, Sobral BWS, Gijzen M (2000) Comparative analysis of expressed sequences in *Phytophthora sojae.* Plant Physiol **123:** 243–253

Ronning CM, Stegalkina SS, Ascenzi RA, Bougri O, Hart AL, Utterbach TR, Vanaken SE, Riedmuller SB, White JA, Cho J, et al (2003) Comparative analyses of potato expressed sequence tag libraries. Plant Physiol **131:** 419–429

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 9.31–9.58

Stekel DJ, Git Y, Falciani F (2000) The comparison of gene expression from multiple cDNA libraries. Genome Res **10:** 2055–2061

Talbot NJ, Ebbole DJ, Hamer JE (1993) Identification and characterization of *MPG1,* a gene involved in pathogenicity from the rice blast fungus *Magnaporthe grisea.* Plant Cell **5:** 1575–1590

Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, Krylov DM, Mazumder R, Mekhedov SL, Nikolskaya AN, et al (2003) The COG database: an updated version includes eukaryotes. BMC Bioinformatics **4:** 41

Velculescu VE, Zhang L, Vogelstein B, Kinzler KW (1995) Serial analysis of gene expression. Science **270:** 484–487

Wang GL, Holsten TE, Song WY, Wang HP, Ronald PC (1995) Construction of a rice bacterial artificial chromosome library and identification of clones linked to the Xa-21 disease resistance locus. Plant J **7:** 525–533

Wang GL, Leung H (1998) Molecular biology of host-pathogen interactions in rice diseases. *In* K Shimamoto, ed, Molecular Biology of Rice. Springer-Verlag, Tokyo, pp 201–232

Wang ZX, Yano M, Yamanouchi U, Iwamoto M, Monna L, Hayasaka H, Katayose Y, Sasaki T (1999) The *Pib* gene for rice blast resistance belongs to the nucleotide binding and leucine-rich repeat class of plant disease resistance genes. Plant J **19:** 55–64

Xiong L, Lee MW, Qi M, Yang Y (2001) Identification of defense-related rice genes by suppression subtractive hybridization and differential screening. Mol Plant Microbe Interact **14:** 685–692

Yamaguchi T, Nakayama K, Hayashi T, Tanaka Y, Koike S (2002) Molecular cloning and characterization of a novel beta-1,3-glucanase gene from rice. Biosci Biotechnol Biochem **66:** 1403–1406

Zeigler RS (1998) Recombination in *Magnaporthe grisea.* Annu Rev Phytopathol **36:** 249–275

Zeigler RS, Leong SA, Teng PS (1994) Rice Blast Disease. CAB International, Wallingford, UK

Zeng L-R, Qu S, Bordeos A, Yang C, Baraoidan M, Yan H, Xie Q, Nahm BH, Leung H, Wang GL (2004) *Spotted leaf11,* a negative regulator of plant cell death and defense, encodes a U-box/armadillo repeat protein endowed with E3 ubiquitin ligase activity. Plant Cell **16:** 2795–2808

Zeng L-R, Yin Z, Chen J, Leung H, Wang GL (2002) Fine genetic mapping and physical delimitation of the lesion mimic gene *Spl11* to a 160 kb DNA segment of the rice genome. Mol Genet Genomics **268:** 253–261

Zhuang JY, Ma WB, Wu JL, Chai RY, Lu L, Fan YY, Jin ZM, Leung H, Zheng KL (2002) Mapping of leaf and neck blast resistance genes with resistance gene analog, RAPD and RFLP in rice. Euphytica **128:** 363–370