

Functional Analysis of Rice Expressed Sequence Tags Induced by Rice Blast Fungus, *Magnaporthe grisea*

Hsiu-Ying Lu (呂秀英)^{1,3}, Yi-Chia Chiu (邱怡嘉)², Meng-Li Wei (魏夢麗)¹, Cheng-Tao Chen (陳政道)¹, Chun-Tang Lu (呂椿棠)

¹Agricultural Research Institute, Taiwan, ROC.

²Graduate Institute of Biotechnology, Chaoyang University of Technology, Taiwan, ROC.

³Correspondence author (iying@wufeng.tari.gov.tw)

Motivation

Rice blast is caused by fungus pathogen *M. grisea*, which is the one of the most destructive diseases and hot problem in rice production (Figure 1). This disease occurs worldwide causing severe loss in yield of up to 85%. Expressed sequence tag (EST) is the fragment of cDNA sequence, therefore direct study of EST is helpful to obtain the gene expression information. In order to know the molecular basis of host resistance to the rice blast fungus, we investigated the gene expression differences among different varieties, tissues and developmental stages of rice using the rice ESTs induced by *M. grisea*.

Material and Methods

- We firstly accessed rice ESTs induced by *M. grisea* from NCBI-dbEST database. A total of 84,705 ESTs were generated from 14 cDNA libraries that are expressed in various varieties, tissues, or developmental stages of rice (Table 1).
- For each library, we performed the entire process of EST functional analysis as: generation of a non-redundant data set (unique sequences) after clustering and assembling the raw ESTs (by CAP3 software), functional annotation of unique sequences (by BLAST similarity search to public annotated TC database in TIGR), functional classification (by MIPS catalog systems from tentative annotation of the selected TC sequences), and statistical analysis of gene functional representation. To achieve this goal, we have established an automated system to quickly and effectively identify the transcribed diversity of gene expression from multiple libraries.

Results and Discussion

- Clustering and assembly of these ESTs in each library resulted in a total of 32,165 unique sequences from 12,206 contigs and 19,959 singletons.
- Gene function classification in each library showed that 30.0-45.5% of unique sequences were assigned to the "unclassified proteins" category, i.e. 54.5-70.0% were predicted to have putative gene function (Table 2).
- Excluding AF and BF-1 libraries with only a small amount of unique sequences, clustering analysis grouped the 12 libraries based on their gene expression of disease reactions (Figure 2). The results showed that when rice ESTs induced by *M. grisea* were mixtures of sequences with different varieties, tissues or developmental stages (Group II-DN, BF-2, CX), it appeared diversity of gene expression from those of rice ESTs from single source (Group I, III).
- The interrelationship between cDNA libraries and functional categories from correspondence analysis revealed that the genes of wild rice (Group III-CD) expressed richer functions of cell growth/morphogenesis (14) and transport facilitation (67) than those of Indian rice and Japanese rice (Figure 3 & Table 2).
- The large cataloged collection of rice ESTs in this study provides a solid foundation for characterization of the rice defense response. These genes and the different gene expression patterns can be used to unravel the regulatory networks of rice resistance to rice blast fungus.

Figure 1. Symptoms of rice blast disease infected by *M. grisea*

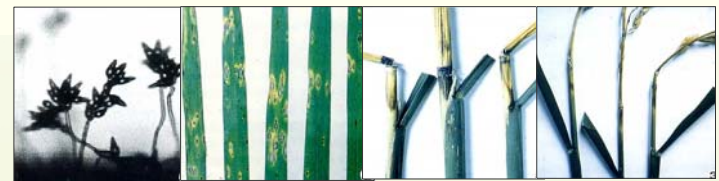


Figure 1. Symptoms of rice blast disease infected by *M. grisea*

Table 1. 14 cDNA Libraries of rice ESTs induced by *M. grisea*

Genbank Acc.	File name	Total ESTs	Created date	Length (bp)	Variety	Tissue	Developmental stage
BF889457 - BF889486	BF-1	14	2001/5/1	173-577	Drew (japonica)	leaf	16-day-old seedling
BI118615 - BI118864	BI-2	250	2001/6/26	152-1175	Yuanfengzao (indica)	leaf	3 week seedling
BI807369 - BI811654	BI-1	6437	2001/10/2 2001/11/1	146-860	<i>Oryza sativa</i> *	leaf	mature stage
BI811655 - BI813841							
BM037978 - BM039065	BM-2	1088	2001/11/6	60-1086	<i>Oryza sativa</i> *	leaf	mature stage
BM418528 - BM422311	BM-1	3784	2002/1/28	169-1045	<i>Oryza sativa</i> *	leaf	mature stage
AF521978 - AF522023	AF	46	2002/7/30	60-600	Bianyiyuexiangzhan (indica)	leaf	seedling five leaf stage
BQ905903 - BQ909262	BQ	3360	2002/8/19	58-800	<i>Oryza sativa</i> *	leaf	mature stage
BU666934 - BU667367	BU	434	2002/9/30	48-788	<i>Oryza sativa</i> *	leaf	mature stage
CB617709 - CB635999	CBir36	18241	2003/4/8	159-979	IR 36 (japonica)	leaf	3 week seedling
CB636000 - CB686047	CBbare	49403	2003/4/8-9	152-973	Nipponbare (japonica)	leaf	3 week seedling
BF145163 - BF145216	BF-2	111	2003/12/30	115-798	Minghui 63 (indica)	14 different tissues	
BF108309 - BF108365							
CD347671 - CD347779	CD	109	2004/3/5	62-815	<i>Oryza minuta</i> 101144 (wild)	leaf	wound treatment & vegetative stage
CX727819 - CX728959	CX	1141	2005/1/21	152-686	Nipponbare(japonica), IR36(japonica)	leaf	3 week seedling
DN475717 - DN475431	DN	287	2005/3/9	143-1733	<i>Oryza sativa</i> (indica cultivar-grop)	13 different tissues	

*Variety name is unknown.

Table 2. Statistics of the gene functional representation for rice ESTs induced by *M. grisea*

library	MIPS functional category*																	
	1	2	3	4	5	6	8	10	11	13	14	25	30	40	63	65	67	99
AF	5.4	0.0	5.0	0.0	0.0	12.0	0.0	0.0	0.0	0.0	0.0	0.0	2.5	26.1	5.0	0.0	6.1	37.5
BF-1	0.0	0.0	0.0	0.0	7.1	4.8	0.0	7.1	2.4	2.4	0.0	7.1	0.0	23.8	9.5	0.0	0.0	35.7
BF-2	4.5	9.4	2.9	0.0	1.3	9.2	1.8	0.2	1.2	0.0	0.0	0.8	2.6	21.9	3.7	0.0	4.7	35.7
BI-1	9.0	5.9	0.9	0.8	2.0	8.4	1.6	1.2	1.1	0.2	0.3	0.7	1.5	17.1	6.5	0.0	5.4	37.4
BI-2	11.0	6.7	1.8	0.2	2.5	3.5	1.5	3.3	2.9	1.0	0.2	0.2	0.2	17.5	5.8	0.0	4.4	37.2
BM-1	10.0	5.9	1.1	0.9	1.6	7.5	2.0	1.1	0.9	0.5	0.2	0.5	1.7	17.6	5.7	0.0	6.3	36.2
BM-2	11.0	8.7	1.3	0.2	2.1	7.4	2.7	1.0	1.5	0.5	0.3	0.5	2.3	18.4	6.7	0.0	4.8	30.0
BQ	9.3	5.7	1.0	0.7	2.4	6.7	2.3	1.4	1.0	0.4	0.2	0.4	1.8	19.6	5.4	0.0	5.7	36.0
BU	14.0	12.0	0.5	0.1	1.2	7.6	0.3	1.3	1.5	0.0	0.4	0.2	0.2	16.4	4.5	0.0	1.5	37.2
CBbare	8.4	3.8	1.2	0.7	1.8	7.6	1.5	1.3	0.8	0.3	0.3	0.4	1.5	15.1	7.0	0.1	5.3	42.9
CBir36	9.1	3.7	1.5	0.6	2.1	9.6	1.6	1.2	0.9	0.3	0.3	0.4	1.6	15.2	8.4	0.0	5.7	37.8
CD	7.3	3.0	2.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.5	0.0	27.5	1.5	0.0	10.0	45.5
CX	7.0	9.1	1.1	0.9	5.5	6.5	2.1	0.7	0.9	0.1	0.1	0.3	2.1	21.6	4.9	0.0	3.5	33.3
DN	2.4	11.0	2.4	0.0	0.5	3.4	3.2	0.0	0.5	0.0	0.0	0.0	1.6	21.9	8.1	0.0	7.6	36.4

*1- metabolism, 2- energy, 3-cell cycle and DNA processing, 4- transcription, 5- protein synthesis, 6- protein fate (folding, modification, destination), 8- cellular transport and transport mechanisms, 10- cellular communication / signal transduction mechanism, 11- cell rescue, defense and virulence, 13- ionic homeostasis, 14- cell growth / morphogenesis, 25- development (systemic), 30- control of cellular organization, 40- subcellular localization, 63- protein binding, 67- transport facilitation, and 99- unclassified proteins.

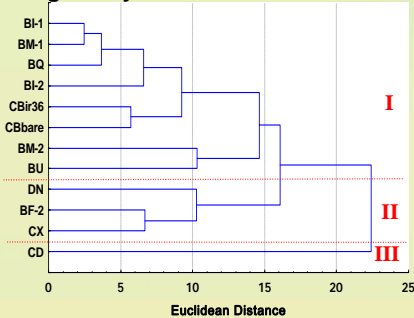


Figure 2. Dendrogram of 12 cDNA libraries from clustering analysis using complete linkage method.

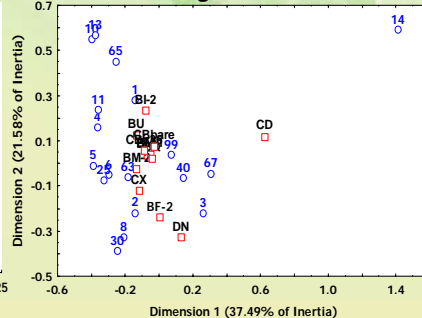


Figure 3. Ordination diagram of 12 cDNA libraries from correspondence analysis.

Acknowledgments

This work was supported by the National Science Council, Taiwan, ROC with a grant no. NSC92-2313-B-055-006, 93-2313-B-055-002.

