

Gene expression analysis and data mining

for the gene network analysis derived from drought stress treatments in rice

Kouji Satoh1 , Koji Doi1 , Toshifumi Nagata1 , Aeni Hosaka1 , Kohji Suzuki2 , Xumei Ji3 , Muturajan Raveendran3 , Ramil Mauleon3 , John Bennett3 , Hei Leung3 , Richard Bruskiwicz3 , Shoshi Kikuchi1

- 1 Plant Genome Research Unit Division of Genome and Biodiversity Research National Institute of Agrobiological Sciences (NIAS) Kan'on dai 2-1-2 Tsukuba Ibaraki 305-8602 Japan
- 2 Hitachi Software Engineerings,Co.,Ltd., 6-81 Onoe-cho, Naka-ku, Yokohama 231-0015, Japan
- 3 International Rice Research Institute (IRRI) DAPO Box 7777, Metro Manila, Philippines

Abstract

To elucidate global responses to drought stress in rice, a 60 mer oligomer microarray covering 22K unique genes based on the sequence of full-length cDNA clones was used to profile gene expression changes at the seedlings, using Apo and IR64 and in the peduncle at heading using IR64. As reported by Ji et al. (2005) Plant Mol Biol. 59:945-964, peduncle elongation in rice is inhibited by drought stress. Cluster analysis of genes up- and down-regulated by drought stress in these two different growth stages revealed the possible combination of the transcription factors and their down-streamers and stage-specific gene expression profiles. Gene expression analysis using the metabolic pathway data in Rice Cyc (<http://www.gramene.org/pathway/>) showed that genes encoding many enzymes of sugar metabolism, such as degradation of sucrose, glucose and galatose, are down-regulated, along with genes encoding enzymes of cell-wall biosynthetic, while genes encoding enzymes of some amino acid biosynthetic pathways are up-regulated. Drought-induced ABA is clearly involved in antagonizing GA-dependent events underlying peduncle elongation, but the biosynthetic genes related to these hormones are not clearly affected by the drought stress treatment. Among 613 differentially expressed transcription factor-related genes, the C3H, AP2-EREBP, bHLH, NAC, MYB and WRKY types of TF-related genes showed differential expression during drought stress treatment and re-watering treatment. Finally, the promoter regions (1kb upstream sequence) of the genes clustered after microarray experiments were examined using a newly developed cis-element analysis tool; the results of this analysis will also be discussed. Data mining using gene annotation data (ex. GO term), pathway data, and genome mapping data suggests the existence of transcription network of drought stress-responsive genes.

Metohd: Microarray system used in this paper

Agilent 22K rice Oligo Array

This array probes were designed from Rice (Nipponabare) FI-cDNA sequences, and 21495 rice probes were spotted on a array.

The annotation information of FL-cDNA clones were presented in KOME site (<http://cdna01.dna.affrc.go.jp/cDNA/>) and these clones are distributed from Rice Genome Resource Center (<http://www.rgrc.dna.affrc.go.jp/jp/index.html>)

Up-regulation								
	APO	IR64	Ped	PEG	Cold	salt	osmotic	Flood
APO	2466							
IR64	1340	1871						
Ped	922	745	2459					
PEG	1285	895	1128	2834				
Cold	1005	708	968	1320	3737			
salt	256	224	226	369	244	467		
osmotic	156	283	298	456	293	297	551	
Flood	798	553	737	1268	766	208	228	2828

Down-regulation								
	APO	IR64	Ped	PEG	Cold	salt	osmotic	Flood
APO	2707							
IR64	1551	1864						
Ped	692	451	2451					
PEG	1140	891	1272	2682				
Cold	1039	798	1015	1700	3285			
salt	174	140	284	421	417	542		
osmotic	156	148	247	415	397	288	492	
Flood	894	814	970	1676	1327	265	304	2581

L. Kikuchi

Specific Up-regulation								
	APO	IR64	Ped	PEG	Cold	salt	osmotic	Flood
	251	94	346	177	870	1	6	339

Specific Down-regulation								
	APO	IR64	Ped	PEG	Cold	salt	osmotic	Flood
	260	94	375	58	374	8	3	109

Common Up-regulation			
	Ped	PEG	Flood
	33		

Common Down-regulation			
	Ped	PEG	Flood
	17		

Table 2. Genes respond to specific stress and respond commonly

Family	22K whole	Dr up	Ratio Dr up	Dr Down	Ratio Dr down	Comment
AP2-EREBP	20	15	75.0%	5	25.0%	
bHLH	70	11	15.7%	25	34.7%	
NAC	64	28	43.8%	16	25.0%	
MYB	40	15	37.5%	16	40.0%	
SBP	58	16	27.6%	10	17.2%	
ZIP	54	15	27.8%	11	20.4%	
C3H	54	20	37.0%	5	9.3%	
WRKY	54	9	16.7%	13	24.1%	
GRAS	52	9	17.3%	6	11.5%	
MYB-related	52	14	26.9%	10	19.2%	
PHD	47	14	29.8%	7	14.9%	
AUX/IAA	43	16	37.2%	0	0.0%	Dr_up-specific
GRAS	27	7	25.9%	12	44.4%	Dr_up-specific
GRAS	26	5	19.2%	6	23.1%	
GRAS	26	4	15.4%	10	38.5%	
GRAS	23	1	4.3%	3	13.0%	
GRAS	23	2	8.7%	4	17.4%	
GRAS	23	2	8.7%	4	17.4%	
GRAS	21	5	23.8%	6	28.6%	
C3C2-GATA	18	2	11.1%	4	22.2%	
HSF	16	9	56.3%	2	12.5%	
TUB	15	6	33.3%	2	13.3%	
ZIM	15	3	20.0%	6	40.0%	
SBP	14	5	35.7%	3	21.4%	
SBP	14	4	28.6%	1	7.1%	
Jamosp	11	0	0.0%	3	27.3%	Dr_down-specific
TOP	11	0	0.0%	6	45.5%	
Atb1-like	10	3	30.0%	3	30.0%	
CCAAT_HAP3	10	2	20.0%	3	30.0%	
CCAAT-HAP2	9	2	22.2%	2	22.2%	
Tubulin	9	3	33.3%	1	11.1%	
CCAAT_HAP2	8	3	37.5%	2	25.0%	
HMG	8	3	37.5%	3	37.5%	
ARID	7	1	14.3%	1	14.3%	
ELL	7	2	28.6%	1	14.3%	
LHD	7	2	28.6%	1	14.3%	

Table 5. List of drought stress responsive TFs classified by gene families

EXP_Pat.	GeneName	RAP Description
Up	AK08555	Dual protein homolog AN1
Up	AK06423	Alanine glyoxylate aminotransferase-like protein (Fragment)
Up	AK060757	Adipolytic dehydrogenase family 7 member A1 (EC 1.2.1.3)
Up	AK06488	(Antiquitin 1) (Matured fruit 60 kD protein) (MF-60)
Up	AK062784	Ribonuclease T2 family protein.
Up	AK062784	Conserved hypothetical protein.
Up	AK063552	Conserved hypothetical protein (Fragment)
Up	AK063833	Short highly repeated, interspersed DNA (Fragment)
Up	AK063833	Twin-strand translocation pathway signal domain containing protein.
Up	AK063896	Hypothetical protein.
Up	AK063923	no-annotation
Up	AK063206	Phyrtene precursor (EC 3.4.23.40) (Aspartic proteinase)
Up	AK066091	Hypothetical protein.
Up	AK068233	Importin alpha 2
Up	AK068727	ERD1 protein, chloroplast precursor.
Up	AK069748	Amino acid/lysine transporter II family protein.
Up	AK070268	Obtobellin regulated protein family protein.
Up	AK070556	no-annotation
Up	AK070872	Low-temperature induced protein h101 2
Up	AK070914	Universal stress protein (Up) family protein.
Up	AK070973	Hypothetical protein.
Up	AK071108	Temperature stress-induced lipocalin.
Up	AK071203	Chac-like protein family protein.
Up	AK072280	Universal stress protein (Up) family protein.
Up	AK099356	Importin alpha 2
Up	AK099418	Dominant-negative associated family protein.
Up	AK100465	Cys/Met metabolism pyridoxal-phosphate-dependent enzymes family protein.
Up	AK101209	Myb, DNA-binding domain containing protein.
Up	AK101609	Single-stranded nucleic acid binding/RCR domain containing protein.
In	AK101837	Thaumatin-like protein.
Up	AK102352	Conserved hypothetical protein.
Up	AK103194	A-ribose hydroxylase A, thiolase (EC 2.3.1.5)
Up	AK105096	Electron transfer flavoprotein alpha subunit, mitochondrial precursor (Alpha-ETF).
Up	AK11578	YT21-B-like protein family protein.

Table 3. 33 commonly up-regulated genes among eight treatments

Specific	Family	GeneName	RAP locus description	Description
Dr_comm_U	AP2-EREBP	AK051349	Pallidone- related transcription factor and ERP domain containing protein.	
Dr_comm_U	AP2-EREBP	AK060543	Ethylene responsive element binding factor 1 (AtERF1)	
Dr_comm_U	AP2-EREBP	AK064552	AP2 domain containing protein RAP2.1 (Fragment)	
Dr_comm_U	AP2-EREBP	AK066670	Ethylene responsive element binding factor 5 (AtERF5)	
Dr_comm_U	AP2-EREBP	AK067313	DRE binding protein 2	
Dr_comm_U	AP2-EREBP	AK067313	Drydown responsive element binding protein 1D (DREB1D protein) (C-repeat binding factor 4) (C-repeat/dth/abn responsive element binding factor 4) (CRTDRE binding factor 4)	
Dr_comm_U	AUX/IAA	AK073365	AUX/IAA protein family protein.	
Dr_comm_U	bZIP	AK063889	Non-protein-coding transcript; untranslatable transcript	
Dr_comm_U	bZIP	AK067919	OSE2-like protein (Fragment)	
Dr_comm_U	bZIP	AK071142	Caserep1 hypothetical protein.	
Dr_comm_U	C3C2-GATA	AK066931	Zn finger, C2H2-type domain containing protein.	
Dr_comm_U	C3H	AK063896	Hypothetical protein.	
Dr_comm_U	C3H	AK070557	Zn finger, C2H2-C2H3-type domain containing protein.	
Dr_comm_U	C3H	AK068392	Zn finger, C2H2-C2H3-type domain containing protein.	
Dr_comm_U	CPP	AK108664	Termin/TSD1-like, CXC domain containing protein.	
Dr_comm_U	HD	AK063885	Short highly repeated, interspersed DNA (Fragment)	
Dr_comm_U	HMG	AK103092	HMG protein	
Dr_comm_U	HSF	AK106488	Heat shock transcription factor 29 (Fragment)	
Dr_comm_U	HSF	AK106455	Heat shock factor protein 3 (HSF 3) (Heat shock transcription factor 3) (HSF 3)	
Dr_comm_U	MYB-related	AK065594	Transcription factor MYB52	
Dr_comm_U	MYB-related	AK101209	Myb, DNA-binding domain containing protein.	
Dr_comm_U	NAC	AK108183	My special activator (NAM) protein domain containing protein.	
Dr_comm_U	SBP	AK062581	SBP domain containing protein.	
Dr_comm_U	SBP	AK063930	Chromatin remodeling factor (Chd) (CHROMOSPICKLE)	
Dr_comm_U	Tubulin	AK069664	Myb, DNA-binding domain containing protein.	
Dr_comm_D	Atb1-like	AK109447	Zn finger-like, PHD finger domain containing protein.	
Dr_comm_D	AUX/IAA	AK069338	AUX/IAA protein family protein.	
Dr_comm_D	AUX/IAA	AK066518	AUX/IAA protein family protein.	
Dr_comm_D	AUX/IAA	AK072001	Auxin-responsive protein (Aux/IAA) (Fragment)	
Dr_comm_D	BES1	AK106748	Plant protein of indoleamine 3-pyruvyl transferase family protein.	
Dr_comm_D	bHLH	AK064946	Transcription factor ICE1 (Substrate of ICE1 expression 1) (Basic helix-loop-helix protein 116) (bHLH116) (bHLH116)	
Dr_comm_D	bHLH	AK101863	TAI protein (Fragment)	
Dr_comm_D	bZIP	AK071839	Elav-like transcription factor, DNA-binding domain containing protein.	
Dr_comm_D	MYB	AK111803	MYB transcription factor	
Dr_comm_D	NAC	AK076657	OHNAC3 protein	
Dr_comm_D	WRKY	AK066255	WRKY transcription factor 45	
Dr_comm_D	ZIM	AK065170	ZIM domain containing protein.	

Table 6. List of common drought responsive TFs

Table 4. 17 commonly down-regulated genes

SIGNAL_NAME	SIGNAL_ID	SIGNAL_SEQ	whole	Dr up	APO	IR64	PED	PEG	Common	TF
ABRELATERD1	200044	AAGTAA	31.6%	35.8%	35.6%	33.0%	28.3%	37.4%	43.2%	bZip (ABRE)
SEF3MOTIFDM	2000115	AAGTAA	15.2%	15.8%	18.5%	20.9%	14.4%	15.1%	16.2%	-
ECODITATTS	2000296	ATATGA	15.0%	15.0%	14.6%	20.9%	15.5%	15.6%	12.9%	-
DBE-TCTGREAT	2000418	AGTACG	15.0%	16.7%	17.0%	13.2%	16.2%	15.6%	20.2%	AP2/EREBP (DREB)
CANRHNAFA	2000148	CNAACAC	13.3%	14.2%	13.2%	17.0%	10.2%	14.3%	18.4%	bZip (ABRE)
CACOTIMOTIF	2000442	CACDTC	13.3%	16.3%	13.1%	13.7%	12.8%	15.2%	24.7%	bZip (ABRE)
TATOCASAMY	2000423	TATOCAC	13.2%	13.3%	13.9%	18.1%	13.4%	16.1%	13.2%	Myb
CARESRFP1	2000421	CAACTC	12.6%	12.9%	11.3%	18.7%	15.0%	12.7%	12.2%	Myb
RATIBAT	2000215	CACDTC	10.3%	11.9%	13.2%	11.2%	9.6%	10.3%	15.1%	AP2/EREBP (DREB)
IBOX	2000134	GATAGC	10.1%	11.0%	9.4%	15.4%	12.9%	9.7%	10.9%	-
HEXDNOTTATHSH4	2000033	AGTCAA	9.7%	11.3%	12.9%	8.2%	11.2%	10.3%	13.2%	Myb
TATABOX	2000110	TATTTAT	9.3%	9.6%	9.4%	11.0%	8.2%	11.3%	12.9%	-
SIBOCORREMBAB7	2000422	AGTAAA	9.2%	9.8%	10.6%	12.1%	11.1%	6.9%	9.2%	Tubulin (TF-Functs)
GTYOCE	2000125	GGTTAA	8.9%	9.0%	7.7%	13.7%	5.7%	8.6%	8.2%	Myb
ACOT1AABMTFAJACREB	2000294	ACOTTCRC	7.6%	10.4%	7.2%	6.0%	6.6%	10.3%	17.6%	bZip (ABRE)
DREOCREMBAB7	2000422	AGTAAA	7.2%	7.2%	8.1%	6.0%	6.9%	7.5%	7.3%	Myb
TATOCAYMOTIFORAMYSD	2000236	TATOCAY	7.0%	7.5%	7.2%	9.3%	8.2%	7.6%	7.8%	Myb
TOACOTFMAMY	2000277	TOACOT	7.0%	7.1%	9.6%	6.6%	4.5%	7.0%	8.9%	-
ABRORSRAB1	2000132	GGTAAA	6.7%	6.8%	9.4%	6.0%	6.9%	7.5%	7.3%	Myb
ABRORSRAB2	2000132	AGTAAA	6.6%	6.8%	9.4%	6.0%	6.9%	7.5%	7.3%	Myb
CDTIRRHVCF2	2000411	GTTCAC	5.9%	6.0%	4.6%	9.3%	5.4%	6.7%	4.7%	AP2/EREBP (DREB)
IBOXORREB	2000424	GATAGAC	5.7%	6.2%	5.3%	6.6%	5.2%	4.8%	6.4%	-
SPYACORREBHA	2000123	GTTCAC	5.5%	5.5%	6.6%	4.4%	5.2%			