Bot. Bull. Acad. Sin. (1997) **38:** 13–20

Isolation of cDNA clones for genes that are specifically expressed in the rice embryo

Peng-Wen Chen[1], Li-Wen Fang[2], Juey-Wen Lin[3], Hsin-Sheng Tsay[4], Hsin-Kan Wu[5], and Liang-Jwu Chen[1,2,6]

[1]*Institute of Botany, National Chung Hsing University, Taichung, Taiwan 40227, Republic of China*

[2]*Institute of Molecular Biology, National Chung Hsing University, Taichung, Taiwan, Republic of China*

[3]*Agricultural Biotechnology Laboratories, National Chung Hsing University, Taichung, Taiwan, Republic of China*

[4]*Department of Agronomy, Taiwan Agricultural Research Institute, Taichung, Taiwan, Republic of China*

[5]*Institute of Botany, Academia Sinica, Taipei, Taiwan, Republic of China*

(Received June 8, 1996; Accepted October 23, 1996)

Abstract. We isolated thirty-eight cDNA clones representing genes expressed in the rice embryo by differential screening of a cDNA library made with poly(A) RNA from embryos at the stage of 10-days after pollination. Eighteen cDNA clones displayed either a significant nucleic acid or amino acid sequence similarity to sequences known in the GenBank, i.e., two rice oleosins, four globulin or glutelin-like proteins, two ABA response proteins, two LEA-like proteins, one trypsin inhibitor, one metallothionein-like protein, one protein kinase, one leucine-zipper containing protein, one proline-rich protein, and three genes of unknown function. Experimental results indicated the presence of redundant cDNA clones for genes that produced an abundant amount of mRNAs in the embryo. RNA slot blot hybridization demonstrated the embryo-specific expression for thirty-seven cDNA clones. The majority of the identified cDNAs had not been known in rice. Results in this study contribute toward the on-going attempt to isolate embryo-specific and novel genes in rice.

Keywords: cDNA library; Differential screening; Embryogenesis; Rice (*Oryza sativa*).

Abbreviations: ABA, abscisic acids; **DAP**, days after pollination; **LEA**, late embryogenesis abundant; **NCBI**, National Center for Biotechnology Information; **Ose**, *Oryza sativa* embryo.

Introduction

Plant genome research with large-scale genomic DNA and cDNA sequencing can facilitate the identification of unknown genes. Recent works involving *Arabidopsis* (Hofte et al., 1993; Newman et al., 1994; Cooke et al., 1996) and rice (Sasaki et al., 1994; Liu et al., 1995) genome projects have concentrated primary on cDNA sequence analysis. Although the complete genomic sequence data can be more informative than the sequence data of cDNA, the genomic DNA requires more sequencing operations for the introns that normally do not provide relevant information to identify an unknown gene. More thoroughly understanding a gene's structure not only requires the sequence data of genomic DNA, but also the sequence data of its corresponding mRNA. cDNA is synthesized from mRNA and represents the product of exons of a gene in the genome. Thus, cDNAs isolated from various tissues or developmental stages reflect the expression pattern of genes in the corresponding tissues or developmental stages. Isolating specific cDNAs from various tissues or developmental stages and the subsequent sequencing and analyzing works will provide both further insight into specific gene expression and a valuable opportunity to identify novel genes.

Recent investigation has demonstrated that large-scale sequencing of randomly selected cDNAs from suspension cultured cell (Uchimiya et al., 1992; Umeda et al., 1994), callus of *Oryza sativa* (Sasaki et al., 1994), leaf of *Zea mays* (Keith et al., 1993), root of *Brassica napus* (Park et al., 1993) and different sources of *Arabidopsis thaliana* (Hofte et al., 1993; Newman et al., 1994; Cooke et al., 1996) is an efficient approach for identifying a variety of expressed genes in plants. Expressed genes isolated as cDNAs in various tissues represent not only the genes that are expressed specifically in the tissue, but also the constitutively expressed house-keeping genes. Therefore, the randomly selected cDNAs from all the previous works contain both the tissue-specific and house-keeping genes. Identifying genes that are expressed only in a specific tissue requires a differential screening against house-keeping genes. By the differential screening method, the selected cDNAs are primarily tissue-specific genes. These genes facilitate the study of how gene expressions are regulated in a particular tissue or cell type.

[6]Corresponding author. Fax: 1-886-4-287-4879; E-mail: ljchen@dragon.nchu.edu.tw

In the present investigation, we identify genes in the rice embryos that may play a role during embryogenesis. We constructed a cDNA library using poly(A) RNA isolated from embryos 10 days after pollination (DAP) and screened for embryo-specific cDNA clones. Here, we report several embryo-specific cDNAs that were isolated from this rice embryo cDNA library by differentially screening against cDNA probes prepared from leaves of a 21-day-old rice seedling. Among the 38 embryo-specific cDNA clones isolated, eighteen cDNA clones are identified, including three cDNA clones that are identified only with partial rice cDNA sequences and clones that are identified with genes for oleosins, globulin, glutelin, trypsin inhibitor, metallothionein, protein kinase, leucine zipper containing protein, late embryogenesis abundant proteins, and ABA responsive proteins. The available database can not identify the remaining twenty cDNA clones, implying that they could be novel genes.

Materials and Methods

Plant Material

Rice (*Oryza sativa* L. var. Lomello) plants were raised in the paddy fields at the Taiwan Agricultural Research Institute, Taichung. The time of pollination was strictly controlled and labelled. Embryos of 10, 20 DAP and mature embryos were dissected under a microscope, immediately frozen with liquid N_2, and stored at -80°C for later use. The leaves of 21-day-old rice seedlings were also used for poly(A) RNA extraction.

Preparation of Poly(A) RNA

Total RNA was extracted with RNA Zol (Biotecx Lab. USA) according to the manufacturer's manual. The poly(A) RNAs were further purified from total RNA by Hybond mAP (Amersham USA) according to the manufacturer's instructions.

Construction of cDNA Library

Poly(A) RNA isolated from embryos of 10 DAP was used for cDNA library construction. cDNA library was constructed using the λUniZAP XR cDNA synthesis kit according to the supplier's instructions (Stratagene, USA). The first strand of cDNA was synthesized by priming oligo$(dT)_{18}$ carrying "GAGA" sequence and *XhoI* adapter. After the second strand of cDNA was synthesized, the termini of the cDNAs were blunted with Klenow fragment and the *EcoRI* adapters were ligated to both ends of the cDNAs. The cDNA that digested with *XhoI* to generate 3' end *XhoI* site was then ligated to the Uni-ZAP XR vector arms and packaged with Gigapack II Gold packaging extract (Stratagene, USA).

Differential Screening

For differential screening, cDNA library was plated and plaques in each plate were lifted in duplicate onto nylon membranes (BRL, USA), denatured with 1.5 M NaCl/0.5 M NaOH, neutralized with 1.5 M NaCl/0.5 M Tris-HCl, pH 7.3, rinsed in 2X SSC/0.2 M Tris-HCl, blotted and dried, and fixed by UV cross linking. The membranes were then prehybridized in 50% formamide, 3X Denhardt's solution, 0.5% SDS, 6X SSC, containing 100 μg/ml denatured salmon sperm DNA, for 2 h at 42°C. Differential screening probes were made using poly(A) RNA from embryos of 10 DAP or leaves of 21-day-old seedlings. Labeled cDNA was synthesized in a 50 μl reaction containing 50 mM Tris-HCl (pH 8.3), 75 mM KCl, 10 mM DTT, 3 mM $MgCl_2$, 1 mM each of dGTP, dATP, dTTP and 100 μCi of [α-^{32}P]dCTP (3000 Ci/mmol, Amersham, USA), 2 μg of oligo (dT), 5 μg of poly(A) RNA, 1 unit of RNAsin (Promega, USA), and 400 units of M-MLV reverse transcriptase (BRL, USA). All this was incubated at 37°C for 2 h. Unincorporated nucleotides were then removed by Sephadex G50 gel filtration spin column. Specific activity of 10^5 to 10^6 counts per ml were used in the hybridization solution. Prehybridized membranes were hybridized with the denatured probes in the buffer described above for 16 h at 42°C. Hybridized membranes were washed twice in 1X SSC, 0.1% SDS at room temperature for 15 min, and exposed to x-ray film. The selected rice embryo-specific cDNA clones were designated as Ose (*Oryza sativa* embryo) clones.

DNA Sequencing

DNA sequencing was carried out from both directions by the standard dideoxynucleotide chain termination method, using a single strand DNA template and Sequenase (United States Biochemical, USA).

Sequence Comparison and Gene Identification

Nucleic acid sequences obtained for each cDNA clone were converted into amino acid sequences for three different reading frames. Computer analysis for nucleic acid and amino acid sequence comparison were performed using the Blast algorithm from the National Center for Biotechnology Information (NCBI) through the Internet (Altschul et al., 1990). Sequences showing a similarity score over 200 were considered homologous to the cDNA clones and presented as putatively identified genes in Table 1 while those with similarity scores less than 200 were not included.

RNA Slot/Gel Blot Hybridization

For slot blot hybridization, one μg of poly(A) RNAs from various tissues or various developmental stages of embryo was blotted onto the Zeta-probe blotting membrane (Bio-Rad, USA) via slot blot SF apparatus (Bio-Rad, USA). For RNA gel blot hybridization, ten μg of total RNA were separated on 1.2% agarose/formaldehyde gels and blotted onto the Zeta-probe blotting membrane. Blots were then hybridized to labeled cDNA probes using prehybridization and hybridization conditions as described previously for plaque screening.

Results and Discussion

Screening for Embryo-Specific cDNA Clones and their Redundancy

The cDNA library used in this study was constructed with poly(A) RNA isolated from rice embryos at 10 DAP. The library had a titer of 4.4×10^6 plaque forming unit per ml and the percentage of plaque containing cDNA insert was 96%, indicating that it was an adequate cDNA library for screening. We screened for embryo-specific cDNA clones by differentially hybridizing against cDNA probes prepared from leaves of a 21-day-old seedlings. From a population of around 80,000 plaques, we initially identified 271 putative embryo-specific cDNA clones that exhibited either a positive signal or relatively stronger signal with the embryo cDNA probes than that of the leaf cDNA probes. These putative embryo-specific cDNA clones were confirmed by DNA dot blot hybridization (data not shown) with duplicate membranes hybridized to embryo cDNA probes and leaf cDNA probes, respectively. The size of cDNA inserts ranged from 0.3 to 1.7 kb among the 271 selected cDNA clones. A group of 126 out of 271 embryo-specific cDNA clones was selected for further study. Plasmid DNAs of these 126 cDNA clones were excised *in vivo* from λ Zap vector and used as templates for DNA sequencing.

Only the 3' end sequence of cDNA insert was initially analyzed. Each cDNA clone with 3' end DNA sequence of 150 to 300 bp was compared to identify the degrees of redundancy among selected cDNA clones. A DNA sequence homology of 95% or higher was classified into the same cDNA group. Thirty-eight distinct embryo-specific cDNA groups were identified after an extensive sequence comparison of the 126 cDNA clones (Table 1). High redundancies of the selected cDNA clones were found within a few cDNA groups. For instance, eleven clones belonged to the cDNA group of Ose701, fourteen clones belonged to the group of Ose703 and thirty-nine clones belonged to the group of Ose705, totaling sixty-four cDNA clones (51% of the 126 clones). In addition to the three cDNA groups, nine other cDNA groups contained two to eight cDNA clone duplications, whereas the remaining 26 cDNA clones (68% of the 38 distinct cDNA groups) were unique (Table 1).

The high redundancy of the three rice embryo-specific cDNA clones, Ose701, Ose703 and Ose705, may have resulted from the abundance of mRNAs presented in our cDNA library. The gene product of cDNA Ose701 is oleosin, the major protein of oil body. The gene product of Ose703 is a glutelin-like protein. These proteins apparently have high mRNA expression during embryogenesis and its further development. In fact, their high mRNA levels in the embryo were observed with RNA slot blot hybridization assay (Figure 1). The gene of cDNA Ose705, although not yet identified, expressed a high level of mRNA at 10 and 20 DAP embryos, similar to that of Ose701 and Ose703 (Figure 1). With a similar expression pattern at 10 and 20 DAP embryos for these three cDNAs, they all displayed a decreased mRNA level in the mature embryo and no detectable level in leaves (Figure 1).

For the embryo-specific cDNA clone isolation, we selected nearly all of the clones that showed positive and strong signals with the embryo cDNA probes, but negative with the leaf cDNA probes. Therefore, the number of the same cDNA being cloned redundantly from this study reflects the relative abundance of its mRNA or the redundantly cloned cDNAs may belong to the members of a gene family. Although it has been estimated that several thousands of genes may be involved during plant em-

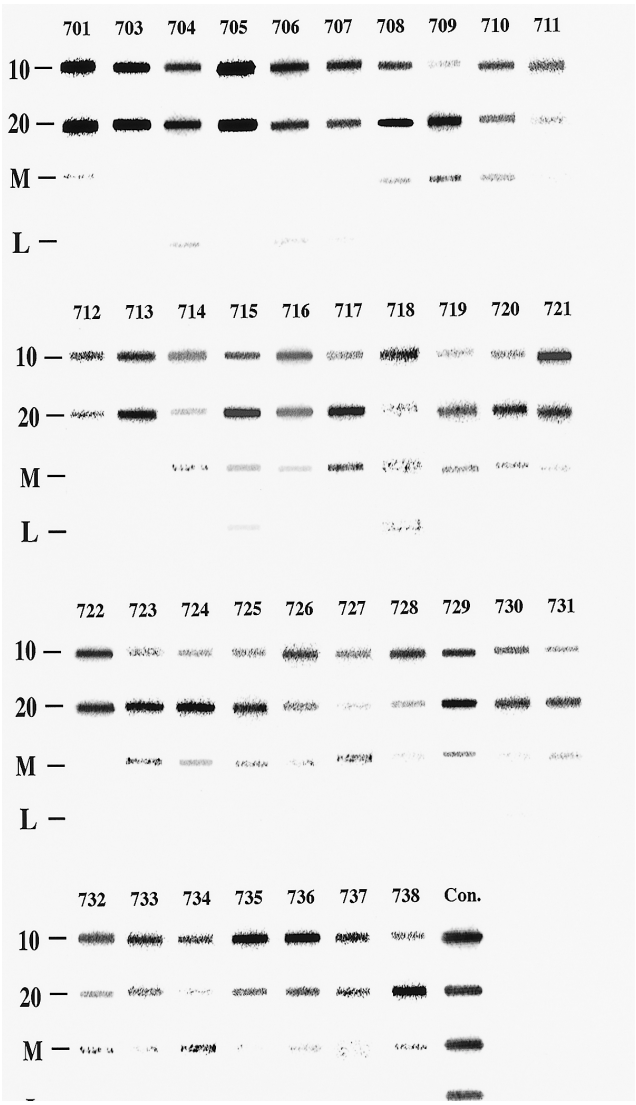

Figure 1. mRNA expression for thirty-eight rice cDNA clones. RNA slot blot hybridization with poly(A) RNA from various stages of embryo tissues and leaves to thirty-eight different cDNA probes was performed. One μg of poly(A) RNA from 10 DAP, 20 DAP, mature (M) embryos and leaves (L) of 21-day-old seedlings were cross-linked to membranes. Each membrane was hybridized with ^{32}p-labeled cDNA probe prepared from thirty-eight different cDNA clones. The cDNA probe used for hybridization is indicated of the top of each lane. The 17S rRNA gene, a constitutively expressed gene, was used as control (Con).

Table 1. List of embryo-specific cDNA clones showing redundancy.

Putatively identified genes	cDNA clone groups	Total numbers of cDNA clones	%[a]
Oleosin (oil body)			
16 kDa	Ose701	11	8.7
18 kDa	Ose721	3	2.4
Storage protein			
Glutelin	Ose703	14	11.1
Globulin-2	Ose704	2	1.6
Globulin-2	Ose709	6	4.8
Globulin-1S	Ose710	2	1.6
LEA protein			
RAB16A	Ose708	6	4.8
RAB16C	Ose717	1	0.8
GmPM3-like	Ose719	1	0.8
Group III	Ose730	1	0.8
Protein kinase	Ose711	1	0.8
Metallothionein	Ose712	4	3.2
Proline-rich protein	Ose720	1	0.8
Trypsin inhibitor	Ose727	1	0.8
Leucine zipper protein	Ose734	1	0.8
Unidentified genes			
That appear	Ose705	39	31.0
More than once	Ose715	8	6.3
	Ose716	3	2.4
	Ose718	2	1.6
That appear once	19 groups	19	15.1
Total	38 groups	126	100 %

[a]Calculation was made on the basis of 126 cDNA clone analyzed.

Table 2. List of putatively identified embryo-specific genes from a rice embryo cDNA library.

cDNA[a] clone	cDNA insert (bp)	Description of putatively[b] identified genes	Nucleic acid[c] identity %	Amino acid[c] identity %	Size of[d] transcript	mRNA expression[e] 10	20	M	L
Ose701	853	Maize 16 kDa Oleosin (U13701)	87 (387)	88 (130)	~1.0 kb	+++	+++	±	–
Ose703	1335	Rice glutelin (X06149)	64 (348)	73 (121)	~2.0 kb	+++	+++	±	–
Ose704[f]	1301	Maize embryo globulin (M24845)	64 (314)		~1.8 kb	+	+	–	±
		Maize globulin-2 (S15675)		76 (143)					
Ose705	1703	Rice cDNA (D15971)	92 (334)	NA*	~2.0 kb	+++	+++	±	–
Ose708	~1000	Rice ABA response gene (Y00842)	95 (164)	NA	NA	++	+++	+	±
Ose709[f]	1405	Maize embryo globulin (M24845)	62 (345)		~1.8 kb	+	++	+	–
		Maize globulin-2 (S15675)		72 (197)					
Ose710[f]	1595	Maize embryo globulin (M24845)	69 (396)		~2.1 kb	+	+	+	–
		Maize globulin-1S, (A53234)		77 (215)					
Ose711	993	Protein kinase (D10152)	59 (177)	82 (56)	~2.2 kb	+	+	–	–
Ose712[f]	645	Rice cDNA (D41142)	96 (224)		~0.8 kb	+	+	±	–
		Metallothionein (S16534)		78 (43)					
Ose716	675	Rice cDNA (D41666)	99 (337)	NA	~0.8 kb	++	++	+	–
Ose717	~800	Rice ABA response gene (X52423)	91 (128)	NA	NA	+	++	+	–
Ose719	932	Soybean seed maturation protein (L20806)	62 (280)	65 (98)	~1.2 kb	+	++	+	–
Ose720	~550	Maize embryo-specific Proline-rich protein (X60432)	86 (96)	NA	NA	+	++	+	–
Ose721	917	Maize 18 kDa oleosin (J05212)	83 (397)	87 (117)	~1.1 kb	++	+	±	–
Ose727[f]	856	Rice cDNA (D40912)	99 (241)		~1.1 kb	+	+	+	–
		Trypsin inhibitor (P07084)		95 (126)					
Ose730	953	Maize lea protein (U05226)	70 (235)	70 (66)	~1.2 kb	+	++	±	–
Ose731	955	Arabidopsis cDNA (Z27266)	70 (225)	NA	~1.2 kb	+	++	+	–
Ose734	~1200	Rice leucine zipper protein (X57325)	96 (181)	NA	NA	+	+	+	–
Control	—	Rice 17S rRNA gene	—	—	—	+++	+++	+++	+++

[a]Ose stands for *Oryza sativa* embryo.
[b]Description of the best-data-match is given together with the GenBank accession number of the match in parentheses.
[c]Nucleic acid and amino acid identities show the percentage of identity together with the length of match in parentheses.
[d]The size of transcript was estimated from the RNA gel blot analysis with known RNA size ladder.
[e]The relative density of mRNA at various tissues is presented as +++ (higher), ++ (high), + (intermediate), ± (low) and – (undetectable). 10 for 10 days, 20 for 20 days and M for mature embryos, and L for leaves. The original RNA slot blot data are shown in Figure 1.
[f]The best-data-match for nucleic acid and amino acid are different. Only the best match one is presented.
*NA: Data are not available.

bryogenesis (Goldberg et al., 1989), we identified only thirty-eight unique embryo-specific genes. This was owing to the limited number of plaques (80,000 plaques) that were screened. Our screening scale may only partially represent the expressed genes in rice embryo. However, the amount of genes identified and the amount of duplication of specific cDNA clones still provide pertinent information as to which type of genes and how abundantly a specific gene was expressed in the rice embryo. In this approach, we realized that most of the clones showed a stronger signal after differential hybridization representing the duplication of the same gene. Whereas, those cDNA clones with a weak but positive signal were primarily unique genes. Therefore, these cDNA clones would be the right choice with respect to selecting a unique gene and avoiding the redundancy for a particular gene.

Sequence Comparison and Putative Gene Identification

More sequence data, including the 5' end DNA sequence of the cDNA insert, for those thirty-eight cDNAs were obtained and subjected to sequence comparison and putative gene identification using the Blast algorithm from NCBI (Altschul et al., 1990). DNA sequences of varying lengths, 300 to 1703 nucleotides, and their derived amino acid sequences were compared with those in the databank. Table 2 lists those cDNA sequences that can be identified using this comparison. Among the thirty-eight cDNAs, eighteen cDNAs (47%) can be identified with sequences known in the databank. For easy comparison, these eighteen identified cDNA clones were divided into three categories: i) cDNA clones that were identified with partial cDNA sequences, ii) cDNA clones that were identified with storage protein-like genes, and iii) cDNA clones that were identified with regulatory protein-like genes and others.

Three out of the eighteen cDNAs were identified with the partial cDNA sequences of rice (Ose705, Ose716) or of *Arabidopsis* (Ose731). The Ose705 and Ose716 showed 92% and 99% nucleic acid identity to rice partial cDNA sequence of D15971 and D41666 (GenBank accession number), respectively. The high degree of homology implied that they were the same gene or belonged to a very closely related gene family. However, the available data bank could not identify the putative gene product for the two cDNAs. The Ose731 cDNA shared 70% homology with an *Arabidopsis* partial cDNA sequence of Z27226 (accession number). No putative gene was identified for Ose731 either. Thus, the three cDNA clones were suggested to be novel rice embryo-specific genes. We have already had the cDNA inserts of Ose705, Ose716, and Ose731 completely sequenced and the sizes of their transcripts measured by RNA gel hybridization (data not shown). The cDNA insert is 1703, 675, and 955 nucleotides for Ose705, Ose716, and Ose731, respectively. The sizes of the transcript are about 2.0 kb, 0.8 kb, and 1.2 kb for cDNA clones Ose705, Ose716,and Ose731, respectively (Table 2). The sizes of the cDNA inserts and their corresponding transcripts are similar, suggesting that the three cDNAs are possibly full-length. This notion is further supported by the existence of a large open reading frame (data not shown) which starts with methionine for each of the three cDNA clones. We are currently characterizing the expression and regulation of these genes.

Four cDNA clones, Ose703, Ose704, Ose709, and Ose710, were classified into the storage protein-like genes category. The sequence comparison indicated that the cDNA Ose703 was similar to rice glutelin gene (Takaiwa et al., 1987). The Ose704 and Ose709 were similar to maize globulin-2 gene (Wallace and Kriz, 1991), and the Ose710 was similar to maize globulin-1S gene (Belanger and Kriz, 1991). The nucleic acid sequence homology for these four cDNA clones ranged from 62% to 69% while the amino acid sequence homology ranged from 72% to 77%. Among the three globulin-like cDNA clones, the Ose704 and Ose709 are similar, though not identical. Whereas the Ose710 showed less homology to either Ose704 or Ose709. Although these four cDNA clones contain 1.3 to 1.6 kb cDNA inserts, they did not contain a complete coding region for their genes. Therefore, to obtain the full-length cDNA clones and demonstrate their gene product is the next step toward understanding the role of the globulin-like genes in rice embryo.

The third category contained eleven cDNA clones, Ose701, Ose721, Ose708, Ose717, Ose719, Ose730, Ose711, Ose712, Ose720, Ose727, and Ose734. The cDNA clones Ose701 and Ose721 have been identified as the rice oleosin, encoding the 16 kDa and 18 kDa oleosin isoforms, respectively (Chen et al., 1996). Oleosins are the structural proteins of plant seed oil body (Tzen and Huang, 1992). The partial sequences of Ose708 were compared with sequences known in the databank. Results from this comparison showed a 95% identity to the coding region of a rice ABA response gene *rab21* (Mundy and Chua, 1988). Gene *rab21* has been renamed *rab16A* (Yamaguchi-Shinozaki et al., 1990). The cDNA insert of Ose717 was about 800 bp. The 440 nucleotides of its 5' end and the 280 nucleotides of its 3' end were sequenced and compared. Those results indicated that Ose717 shared a 91% homology with the 140 nucleotide coding sequences of a rice ABA response gene, *rab16C* (Yamaguchi-Shinozaki et al., 1990). The high degree of homology within the coding region implied that the genes (Ose708 to *rab16A* and Ose717 to *rab16C*) were closely related.

The amino acid sequences of cDNA clones Ose719 and Ose730 showed a 65% to 70% similarity to soybean LEA protein (Hsing et al., 1995) and maize LEA protein (White and Rivin, 1995), respectively. LEA protein was defined as late embryogenesis abundant protein, which normally accumulates in embryo tissues as they approach maturity. However, our identified LEA-like cDNAs, Ose719 and Ose730, were isolated from ten DAP embryos. Thus, it would be interesting to examine the functions of the two LEA-like cDNAs in rice embryo and their roles in relation to late embryogenesis.

The amino acid sequences of Ose711 showed a high homology with an *Arabidopsis* protein tyrosine kinase gene (D10152, no literature published). The deduced pro-

tein for Ose712 was a typical metallothionein-like protein that contained cysteine-rich domains at both the N- and C- terminals and showed a high homology with a barley metallothionein-like protein (Klemsdal et al., 1991). The cDNA insert of Ose720 was about 550 nucleotides long; only 261 nucleotides were sequenced and compared. Those results indicated that Ose720 is similar to the maize embryo-specific proline-rich protein gene (Jose-Estanyol et al., 1992). The amino acid sequence of Ose727 was similar to a rice Bowman-Birk bran trypsin inhibitor (Tashiro et al., 1987). The cDNA insert of Ose734 was about 1.2 kb and its partial nucleotide sequence was identified as a rice leucine-zipper containing protein gene (Aguan et al., 1993).

As far as we know, the majority of the eighteen identified cDNA clones have not been studied in rice. Some of the clones have a similar but not identical sequence in rice; others were nearly identical to a partial sequence of rice cDNAs that have been deposited in the GenBank but without any further characterization. In addition to the eighteen identified cDNA clones, twenty cDNA clones (53% of selected embryo-specific cDNA clones) have not yet been identified with known DNA sequences. Those cDNAs are potentially novel genes. Therefore, our approach using differential screening from a 10 DAP embryo cDNA library allowed us to isolate novel embryo-specific genes to examine their roles in rice embryos during embryogenesis.

Expression of Embryo-Specific cDNAs

The tissue specific expression of the thirty-seven selected cDNA clones (Ose702 was not available) and a constitutively expressed 17S rRNA gene (used as control) were analyzed by RNA slot blot hybridization. Figure 1 summarizes these results. In general, the mRNAs are highly expressed in both 10 and 20 DAP embryos and less in the mature embryo. No expression was observed in leaves except for cDNA clones Ose704, Ose706, Ose707, Ose715, and Ose718, where a few mRNA were detected. The expression of these 37 cDNA clones can be roughly divided into three categories based on the abundance of mRNA detected at various developmental stages of the embryo.

The first category includes those cDNA clones having a higher mRNA level at both 10 DAP and 20 DAP. These included 10 cDNA clones Ose701, Ose703, Ose704, Ose705, Ose706, Ose707, Ose713, Ose721, Ose722, and Ose729. Among them, the cDNA clones of Ose701 (oleosin), Ose703 (glutelin-like), and Ose705 (unknown gene) had the most abundant mRNA level, and this abundance was consistent with the high redundancy of cDNA clones selected by this approach. This phenomenon has been discussed above. Two cDNA clones, Ose704 and Ose721, had been identified with globulin-like gene and oleosin gene, respectively. The other five cDNA clones that also expressed relatively high mRNA level at both 10 DAP and 20 DAP, but their putative function has not yet been identified.

The second category is comprised of those clones with an intermediate expression level and have an mRNA level higher at 20 DAP than at 10 DAP. Twelve cDNA clones were grouped into this category. They are Ose708, Ose709, Ose715, Ose717, Ose719, Ose720, Ose723, Ose724, Ose725, Ose730, Ose731, and Ose738. The identified cDNAs for rice LEA-like genes, Ose708, Ose717, Ose719, and Ose730, (Yamaguchi-Shinozaki et al., 1990; Hsing et al., 1995; White and Rivin, 1995) and the globulin-like gene, Ose709, (Wallace and Kriz, 1991) were in this category. Although the cDNAs Ose708 and Ose717 share a high homology with the rice ABA responsive genes *rab16A* and *rab16C,* respectively, their expression patterns are quite different. While *rab16A* and *rab16C* showed a significant mRNA level at the mature embryo (Yamaguchi-Shinozaki et al., 1990), the mRNA levels of Ose708 and Ose717 at the mature embryos were lower than that at the 10 and 20 DAP. Similarly, the cDNAs Ose719 and Ose730, which were identified as the LEA-like genes, also expressed a higher mRNA level in the embryos at the 10 and 20 DAP. The putative function for the other cDNA clones in this category were unknown.

The remaining fifteen cDNA clones belonged to the third category. This category contains those clones having a relatively low expression level with mRNA level higher at 10 DAP than at 20 DAP. This category includes the identified globulin-like gene (Ose710), the protein kinase-like gene (Ose711), the metallothionein-like gene (Ose712), the trypsin inhibitor-like gene (Ose727), and a gene that encodes a leucine zipper containing protein (Ose734). Notably, the regulatory protein-like cDNAs, Ose711 and Ose734, were expressed in a smaller amount of mRNA during rice embryogenesis. Comparing the expression levels for those storage protein-like cDNAs revealed that the Ose703 (glutelin-like gene) has a higher mRNA level than Ose704, Ose709, and Ose710 (globulin-like genes). This finding corresponds to a previous study in which much more glutelin than globulin was observed in rice seed storage protein (Konzak, 1977).

In this experiment, we intended to isolate embryo-specific full-length cDNA clones. Identifying a full-length cDNA clone would require data for the size of its corresponding mRNA. We performed RNA gel blot analysis to obtain the transcript size for several cDNA clones (Table 2). A cDNA clone is considered a full-length cDNA if the size of its cDNA insert is similar (within 300 nucleotides) to the size of its corresponding mRNA. According to this size comparison and other criteria, e.g., the location of amino acid methionine codon and the open reading frame, we estimated that at least ten out of the eighteen identified cDNA clones contain the complete coding region of their genes.

In conclusion, we have demonstrated that our approach is effective in finding rice embryo-specific genes. We provided further insight into gene expression pattern for several previously identified rice genes such as genes for ABA response proteins and LEA-like proteins. We obtained cDNAs representing new members in a known gene fam-

ily such as cDNAs of globulin, glutelin, and ABA response protein. Most interestingly, we have isolated embryo-specific cDNAs that have no corresponding genes yet cloned in rice, such as oleosin, protein kinase, and twenty cDNAs that may eventually be novel rice embryo-specific genes. Sequences for their corresponding genes will be done soon to determine their gene expression and regulation.

Acknowledgments. This study is part of the project "Plant genome analysis: Rice genome analysis at early embryogenesis" supported by grants NSC83-0418-B-005-057-B03 and NSC84-2321-B-005-035-B03 from National Science Council, Republic of China. Dr. Y-C Kuo and Mr. C-C Chen are appreciated for technical assistance in rice embryo preparation.

Literature Cited

Aguan, K., K. Sugawara, N. Suzuki, and T. Kusano. 1993. Low-temperature-dependent expression of a rice gene encoding a protein with a leucine-zipper motif. Mol. Gen. Genet. **240:** 1–8.

Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D. Lipman. 1990. Basic local alignment search tool. J. Mol. Biol. **215:** 403–410.

Belanger, F.C. and A.L. Kriz. 1991. Molecular basis for allelic polymorphism of the maize globulin-1 gene. Genetics **129:** 863–872.

Chen, P.W., Y.J. Chai, L.D. Wang, J.T.C. Tzen, M.J. Tseng, and L.J. Chen. 1996. Two embryo-specific cDNAs (Accession No. U43930 and U43931) encoding two oleosin isoforms on the surface of oil bodies from rice. Plant Physiol. **109:** 714.

Cooke, R., M. Raynal, M. Laudie, F. Grellet, M. Delseny, P-C. Morries, D. Guerrier, J. Giraudat, F. Quigley, G. Clabault, Y-F. Li, R. Mache, M. Krivitzky, I. Gy, M. Kreis, A. Lecharny, Y. Parmentier, J. Marbach, J. Fleck, B. Clement, G. Philipps, C. Herve, C. Bardet, D. Tremousaygue, B. Lescure, C. Lacomme, *et al*. 1996. Further progress towards a catalogue of all Arabidopsis genes: analysis of a set of 5000 non-redundant ESTs. Plant J. **9:** 101–124.

Goldberg, R.B., S.J. Barker, and L. Perez-Grau. 1989. Regulation of gene expression during plant embryogenesis. Cell **56:** 149–160.

Hofte, H., T. Desprez, J. Amselem, H. Chiapello, M. Caboche, A. Moisan, M-F. Jourjon, J-L. Charpenteau, P. Berthomieu, D. Guerrier, J. Giraudat, F. Quigley, F. Thomas, D-Y. Yu, R. Mache, M. Raynal, R. Cooke, F. Grellet, M. Delseny, Y. Parmentier, G. Marcillac, C. Gigot, J. Fleck, G. Philipps, M. Axelos, C. Bardet, D. Tremousaygue, and B. Lescure. 1993. An inventory of 1152 expressed sequence tags obtained by partial sequencing of cDNAs from *Arabidopsis thaliana*. Plant J. **4:** 1051–1061.

Hsing, Y.C., Z.Y. Chen, and T.Y. Chow. 1995. A soybean cDNA (Accession No. L20806) encoding a hydrophobic embryogenesis abundant protein. Plant Physiol. **109:** 1125.

Jose-Estanyol, M., L. Ruiz-Avila, and P. Puigdomenech. 1992. A maize embryo-specific gene encodes a proline-rich and hydrophobic protein. Plant Cell **4:** 413–23.

Keith, C.S., D.O. Hoang, B.M. Barrett, B. Feigelman, M.C. Nelson, H. Thai, and C. Baysdorfer. 1993. Partial sequence analysis of 130 randomly selected maize cDNA clones. Plant Physiol. **101:** 329–332.

Klemsdal, S.S., W. Hughes, A. Lonneborg, R.B. Aalen, and O.A. Olsen. 1991. Primary structure of a novel barley gene differentially expressed in immature aleurone layers. Mol. Gen. Genet. **228:** 9–16.

Konzak, C.F. 1977. Advances in Genetics. *In* E. W. Caspari (ed.), Academic Press, New York, Vol. 19, pp. 407–582.

Liu, J., C. Hara, M. Umeda, Y. Zhao, T.W. Okita, and H. Uchimiya. 1995. Analysis of randomly isolated cDNAs from developing endosperm of rice (*Oryza sativa* L.): evaluation of expressed sequence tags, and expression levels of mRNAs. Plant Mol. Biol. **29:** 685–689.

Mundy, J. and N.H. Chua. 1988. Abscisic acid and water-stress induce the expression of a novel rice gene. EMBO **7:** 2279–2286.

Newman, T., F.J. Bruijn, P. Green, K. Keegstra, H. Kende, L. McIntosh, J. Ohlrogge, N. Raikhel, S. Somerville, M. Thomashow, E. Retzel, and C. Somerville. 1994. Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones. Plant Physiol. **106:** 1241–1255.

Park, Y.S., J.M. Kwak, O.-Y. Kwon, Y.S. Kim, D.S. Lee, M.J. Cho, H.H. Lee, and H.G. Nam. 1993. Generation of expressed sequence tags of random root cDNA clones of Brassica napus by single-run partial sequencing. Plant Physiol. **103:** 359–370.

Sasaki, T., J. Song, Y. Koga-Ban, E. Matsui, F. Fang, H. Higo, H. Nagasaki, M. Hori, M. Miya, E. Murayama-Kayano, T. Takiguchi, A. Takasuga, T. Niki, K. Ishimaru, H. Ikeda, Y. Yamamoto, Y. Mukai, I. Ohta, N. Miyadera, I. Havukkala, and Y. Minobe. 1994. Toward cataloguing all rice genes: large-scale sequencing of randomly chosen rice cDNAs from a callus cDNA library. Plant J. **6:** 615–624.

Takaiwa, F., S. Kikuchi, and K. Oono. 1987. A rice glutein gene family - a major type of glutein mRNAs can be divided into two classes. Mol. Gen. Genet. **208:** 15–22.

Tashiro, M., K. Hashino, M. Shiozaki, F. Ibuki, and Z. Maki. 1987. The complete amino acid sequence of rice bran trypsin inhibitor. J. Biochem. **102:** 297–306.

Tzen, J.T., and A.H. Huang. 1992. Surface structure and properties of plant seed oil bodies. J. Cell Biol. **117:** 327–335.

Uchimiya, H., S. Kidou, T. Shimazaki, S. Aotsuka, S. Takamatsu, R. Nishi, H. Hashimoto, Y. Matsubayashi, N. Kidou, M. Umeda, and A. Kato. 1992. Random sequencing of cDNA libraries reveals a variety of expressed genes in cultured cells of rice (*Oryza sativa* L.). Plant J. **2:** 1005–1009.

Umeda, M., C. Hara, Y. Matsubayashi, H.H. Li, Q. Liu, F. Tadokoro, S. Aotsuka, and H. Uchimiya. 1994. Expressed sequence tags from cultured cells of rice (*Oryza sativa* L.) under stressed conditions: analysis of transcripts of genes engaged in ATP-generating pathways. Plant Mol. Biol. **25:** 469–478.

Wallace, N.H. and A.L. Kriz. 1991. Nucleotide sequence of a cDNA clone corresponding to the maize globulin-2 gene. Plant Physiol. **95:** 973–975.

White, C.N. and C.J. Rivin. 1995. Sequence and regulation of a late embryogenesis abundant group 3 protein of maize. Plant Physiol. **108:** 1337–1338.

Yamaguchi-Shinozaki, K., J. Mundy, and N.H. Chua. 1990. Four tightly linked *rab* genes are differentially expressed in rice. Plant Mol. Biol. **14:** 29–39.