A spectrum of genes expressed during early stages of rice panicle and flower development

KUMUDA M. KUSHALAPPA[1,2], AUTAR K. MATTOO[3] and USHA VIJAYRAGHAVAN[1*]

[1]*Department of Microbiology and Cell Biology, Indian Institute of Science, Bangalore 560 012, India*
[2]*Department of Botany, University of British Columbia, Vancouver, BC V6T124, Canada*
[3]*USDA Vegetable Lab, Building 010A, ARS/PSI/BARC- West, 10300 Baltimore Avenue, Beltsville, MD 20705-2350, USA*

Abstract

To unravel gene expression patterns during rice inflorescence development, particularly at early stages of panicle and floral organ specification, we have characterized random cloned cDNAs from developmental-stage-specific libraries. cDNA libraries were constructed from rice panicles at the stage of branching and flower primordia specification or from panicles undergoing floral organogenesis. Partial sequence analysis and expression patterns of some of these random cDNA clones from these two rice panicle libraries are presented. Sequence comparisons with known DNA sequences in databases reveal that approximately sixtyeight per cent of these expressed rice genes show varying degrees of similarity to genes in other species with assigned functions. In contrast, thirtytwo per cent represent uncharacterized genes. cDNAs reported here code for potential rice homologues of housekeeping molecules, regulators of gene expression, and signal transduction molecules. They comprise both single-copy and multicopy genes, and genes expressed differentially, both spatially and temporally, during rice plant development. New rice cDNAs requiring specific mention are those with similarity to *COP1*, a regulator of photomorphogenesis in *Arabidopsis*; sequence-specific DNA binding plant proteins like *AP2*-domain-containing factors; genes that specify positional information in shoot meristems like leucine-rich-repeat-containing receptor kinases; regulators of chromatin structure like Polycomb domain protein; and also proteins induced by abiotic stresses.

Introduction

Plants have become important targets for genetic manipulation and provide good models to study biological processes that control development. Analysis of these processes at a molecular level requires isolation and characterization of key regulatory genes, including genes that are differentially expressed. Current approaches for isolating plant genes include chromosome walking and gene tagging, methods that have been particularly successful in isolating plant genes where likely functions are revealed by their mutant phenotypes. Rapid progress in technologies related to genome biology has occurred in the recent past. These methodologies include characterization of anonymous cDNA clones expressed in varying degrees both during normal development and in disease state. Analysis of the human genome spearheaded many of these technologies and demonstrated that even partial sequencing of randomly chosen cDNAs can provide a large amount of information concerning the genetic makeup of an organism and can provide sequence-tagged markers for the genome mapping and analysis. Large-scale random cDNA and/or genome sequencing of *Oryza sativa, Zea mays, Brassica napus* and *Arabidopsis thaliana* have provided sequences of an expanding collection of plant genes (Uchimiya *et al*. 1992; Keith *et al*. 1993; Park *et al*. 1993; Sasaki *et al*. 1994; Cooke *et al*. 1996; Lin *et al*. 1999). Among these studies, the analysis of all expressed sequences from at least two plant genomes, *Arabidopsis* and *Oryza sativa* (rice), are large-scale international efforts. The choice of these two plants arises from their being model representatives for laboratory plant systems and field crops respectively, whose genomes are

*For correspondence. E-mail: uvr@mcbl.iisc.ernet.in.

Keywords. cDNA clones; rice panicles; floral development; rice.

also amenable to global sequence analysis. For a fraction of the rice and *Arabidopsis* genes, a probable function has been assigned on the basis of sequence similarity to previously studied genes from other systems. Apart from sequence similarity, function can be hypothesized from RNA and protein synthesis patterns and can be further tested through phenotypic analysis of plants with gain-of-function or loss-of-function alleles. These latter molecular-genetic tools are available for both these plant species.

Genes expressed in a developmental-stage-specific and/or tissue-specific manner are of interest for various reasons. One compelling reason is the likelihood of identifying regulators of a pathway or identifying genes that in turn are regulated by such molecules. In addition to specific expression patterns, these molecules can lend themselves for use as markers to study development; also their regulatory elements can provide ways to achieve directed expression of foreign genes. Such cDNAs are also adaptable for DNA-chip technology for use in transcript monitoring in functional genomic studies. In rice, panicle and flower development is under environmental and genetic control; development of a mature flower from the stage of panicle initiation occurs in a step-wise manner over a four to five week period (Hoshikawa 1989). Most of our current knowledge on initiation and morphogenesis of flowers is derived from analysis of dicot model plants (Weigel and Meyerowitz 1995). Understanding these phenomena in rice, a member of the grass family, will provide clues on the evolution of a morphologically diverse form, where the branched inflorescence and derived floral structure are unique characteristics.

In an attempt to identify rice genes expressed during early stages of panicle development or during floral organ specification we have constructed cDNA libraries from rice panicles at two developmental stages. Panicles elaborating branches with newly specified flower primordia or panicles with flowers undergoing organogenesis were taken for analysis. This study should thus add to the wealth of emerging plant expressed sequence tags (ESTs) from the *Arabidopsis* and rice sequencing projects. Here we present partial sequence analysis of some random cDNA clones from these panicle libraries. For a few genes a putative function could be predicted on the basis of sequence similarity to other known genes in the databases. We have studied the genomic organization by Southern analysis and have probed expression patterns by Northern analysis of RNAs from various tissues for a few representative genes. Together these rice cDNAs provide start points for a reverse-genetics approach to studying gene function.

Materials and methods

Construction of cDNA libraries and nucleotide sequence analysis: Two independently constructed rice panicle cDNA libraries were the source of clones used in the study. These libraries were constructed from mRNA isolated from young panicles (*Oryza sativa* L., cv IR 64) at two different stages of development. Early panicles of length 0.5 cm to 1 cm are at the stage of floral meristem specification, while the branch primordia are well defined. Panicles of length 2 to 3 cm are termed late, wherein floral organogenesis is under way. Panicles were dissected away from the tightly rolled leaf sheath and were quickly frozen in liquid nitrogen before total RNA isolation (Crawford *et al*. 1986). Poly(A)$^+$ RNA was purified using Qiagen oligo-dT column. Five μg of poly(A)$^+$ RNA served as a template for cDNA synthesis. cDNA was primed using an *Xho*I oligo(dT)-linker primer provided by the kit and manufacturer's instructions were followed henceforth (Stratagene, USA). The cDNAs were size selected to be greater than 400 bp and then cloned unidirectionally into the *Eco*RI and *Xho*I sites of phage vector Uni-*ZAP*. The primary library in each case contained four to five million plaque-forming units. A portion of this unamplified primary library was used for picking up random plaques. The cDNAs in the plasmid (pBluescript-SK) were excised from each phage isolate according to the supplier's protocol. Sequence analysis of the double-stranded plasmid DNA was performed with an automated sequencer (Applied Biosystems, USA) using the dideoxy cycle sequencing protocol and the T3 primer that flanks the 5$'$ end of the cDNAs. The 5$'$ end was partially sequenced ($\sim$700 bp of sequence obtained in each case) for all cDNAs isolated. DNA sequence analysis and database comparisons were done against GenBank entries, first at amino acid level with the BLASTX subroutine and at the nucleotide level by the BLASTN subroutine (Altschul *et al*. 1997). Additionally, searches were made against the rice EST collections at www.tigr.org/tdb/ogi (TIGR, The Institute for Genome Research, Rockville, USA) using the routine WU-BLAST.

RNA isolation and Northern blot analysis: In general total RNA was isolated from one gram each of the following tissues: root, shoot, leaf sheath, early panicles (less than 1 cm) and late panicles (1 to 3 cm), lemma/palea and stamen/carpel according to Crawford *et al*. 1986. Roots and shoots were obtained from 15-to-20-day-old young seedlings and other tissues mentioned above were from adult field-grown plants. For assessing induction by abiotic stress, seedlings 15 to 20 days post-germination were treated with 0.1 M NaCl for 14 h by submergence of their roots. Following this treatment roots and shoots were harvested separately and RNA isolated. Total RNA yield in all cases was quantified and 30 μg of total RNA was denatured in formamide-containing buffer and then electrophoresed on denaturing 1.2% formaldehyde agarose gels. The RNA samples were blotted onto Hybond membranes (Amersham, UK) and then UV cross-linked. After prehybridization in Church buffer (0.5 M NaPO$_4$ pH 7.4, 1 mM EDTA and 7% SDS) for 3 h at 65°C (Church and Gilbert 1984), the blots were hybridized in the same solution with the specified probe added at a concentration of a million cpm per ml. All probes were generated by random primer labelling reactions

of gel-purified cDNA inserts from the specified plasmid with [α-^{32}P]dATP or [α-^{32}P]dCTP (3000 Ci/mM) (Amersham). Post-hybridization membranes were washed twice in $2 \times$ SSC, 0.1% SDS at room temperature and then twice in $1 \times$ SSC at 65°C for 10 min (Sambrook *et al*. 1989). The blots were then subjected to autoradiography. The blots were subsequently probed with a rice cDNA encoding a ribosomal protein to normalize for mRNA loaded.

Plant genomic DNA isolation and Southern blot analysis: DNA was isolated from young leaves of rice (*Oryza sativa* L., cv IR 64) by CTAB method (Rogers and Bendich 1988). Five μg of the genomic DNA was digested with the indicated restriction enzymes in recommended buffer. The digested DNA was electrophoresed on a 0.7% agarose gel and then blotted onto Hybond nylon membranes. The radiolabelled probes were prepared and hybridizations were carried out as outlined above. The blots were washed with $0.5 \times$ SSC at 65°C.

Results and discussion

Construction of rice panicle libraries representing two developmental stages

Rice flowers are generated on a branched inflorescence termed the panicle. The mature panicle bears eight to 10 flowering nodes on the main stem axis called the rachis. Each node bears a rachis branch, which in turn bears at its basal nodes secondary rachis branches and more terminally five or six pedicellate spikelets. The secondary branches themselves have two to four pedicellate spikelets. The development/morphogenesis of the mature panicle thus requires, first, specification of a defined number of primary rachis branch primordia, then of secondary rachis branch primordia, followed by specification of splikelet/floral primordia, and still later of the floral organ primordia. To explore changing patterns of gene expression that occur during early panicle morphogenesis we constructed two different stage-specific cDNA libraries. Scanning electron microscope based characterization of these stages has been reported previously (Hoshikawa 1989; Kushalappa *et al*. 1996). The earliest developmental stage taken for analysis was of panicles of length less than 1 cm, wherein the rachis branch primordia and floral primordia are being specified. The second stage was panicles at later developmental stage (length 1 to 3 cm), where glumes, lemma and palea have already been specified on individual floral primordia, and the development of the stamens and carpels is initiated. These libraries were arbitrarily termed early panicle cDNA library and late panicle cDNA library respectively, and they have been used to define genes expressed in the developing inflorescence. To optimize and facilitate the characterization of the expressed genes the cDNAs were size selected and cloned directionally in the commercially available vector λ Uni-*ZAP*. We anticipate that size selection of cDNAs together with directional cloning of full-length or near-full-length inserts would aid in the identification of similarities to known genes and in the subsequent authentication of these homologies.

Our libraries, we expect, are representative since we obtained a primary unamplified titre of approximately 4 to 6 million plaque forming units (pfu). For a faithful depiction of the expressed genes that is reflective of their relative expression levels, random cDNAs were taken from the unamplified library itself. In addition the libraries were tested for the presence of cDNAs for a gene known to be expressed at early stages of floral meristem specification. *OsMADS1* is a putative rice transcription factor with the DNA-binding sequence motif termed the MADS box. Previous characterization of the gene has shown that this gene is transcriptionally regulated with the earliest appearance of its mRNA being coincident with the appearance of morphologically discernible floral primordia (Chung *et al*. 1994; Vijayraghavan 1996). We screened the late panicle amplified library for *OsMADS1* cDNAs by high-stringency hybridization using a cloned gene-specific partial cDNA. From these analyses we estimate the abundance of this cDNA in the late amplified library as being one in 30,000 pfu (our unpublished data). These data thus provide assurance of the developmental stage specificity of these libraries and indicate that they are likely to contain other genes that are similarly regulated. Such early-stage-specific libraries from rice panicles are as yet unreported. The rice database currently has EST sequences from panicles that represent later stages of development, i.e. the stages of floral differentiation, ripening and grain maturation. Thus, sequences and cDNAs from our libraries should complement the information available through international EST sequencing programs.

Characterization of cDNA clones

cDNA inserts (released after *Eco*RI and *Xho*I digests) present in 20 randomly selected cDNAs were sized and found to be in the size range of 700 to 1500 bp. Fifty clones from the early-panicle cDNA library (panicles of 0.5 cm to 1 cm in length) and a slightly smaller number of clones from the late-panicle cDNA library (panicles of 1 cm to 3 cm in length) were subjected to the single-run partial sequencing using an automated ABI system. Since the cDNAs were cloned directionally and the 5′ untranslated regions of many plant genes are relatively short, the clones were all sequenced from the 5′ end to maximize the likelihood of finding gene sequences. We categorized many genes based on sequence similarities to current entries in GenBank and EST databases. For these homology searches all sequences obtained from the 5′ end of the cDNAs were automatically translated in all reading frames and compared with the protein sequence database in GenBank using the subroutine BLASTX, and further compared at the nucleotide level using BLASTN. BLASTN analysis showed identities/similarities to randomly sequenced cDNAs from other

Table 1. Catalogue and identities of some genes expressed in developing rice panicles.

Clone number	GenBank Accession of rice cDNAs	NCBI BLAST score	Putative functional identification	Hits to rice ESTs at TIGR (WU-BLASTN scores)
Structural proteins				
ERD9	AF074728		Uroporphyrinogen decarboxylase precursor	TC17200 (712)
ERD56	AF074737	369	Acetyl CoA dehydrogenase	
ERD59	AF074739	396	Pyruvate dehydrogenase	S2039-1A (838)
ERD62	AF074742	540	Protein disulphide isomerase precursor	R0110-1A (1104)
ERD76	AF074754	590	Ascorbate peroxidase	S3683-1A (1331)
LRD5	AF074772	728	Ion channel protein (36 kDa porin II)	TC21145 (2522)
LRD14	AF074780	980	Ubiquitin gene	C73073 (809)
LRD21	AF074787	476	*Sorghum* aldehyde dehydrogenase	AU029563 (363)
LRD22	AF074788	120	Theronyl tRNA synthetase ligase	
LRD23	AF074789	399	Proteosome subunit	C27723 (648)
LRD24	AF074790	510	Permease	C0087-A (735)
LRD28	AF074793		Aldehyde dehydrogenase	TC22754 (2044)
LRD29	AF074794	198	Ubiquitin conjugating enzyme	TC26955 (1989)
LRD31	AF074795	300	Ubiquitin conjugating enzyme	C26526 (964)
LRD32	AF074796	309	Ubiquitin conjugating enzyme	TC26955 (1967)
LRD36	AF074798	506	Permease	TC29466 (4004)
LRD42	AF074803	186	Mitochondrial membrane-associated protein	
LRD45	AF074805	140	Glucoride precursor	C91824 (935)
LRD46	AF074806	100	Membrane-associated protein	
LRD49	AF074808	391	Glycine-rich protein	
Protein kinases/phosphatases				
ERD53	AF074734	396	LRR proteins	S10770 (1324)
ERD90	AF073306	181	Protein kinase	S14172-A (620)
ERD77	AF074755		Putative receptor kinase	NP001765 (222)
LRD12	AF074779		Putative GTP-binding protein	TC21339 (2976)
LRD9	AF074776		Type 2 phosphatidic acid phosphatase	D47079 (1119)
Transcription factors/regulatory factors				
ERD2	AF074725	77	CUC/NAM protein	C10012A (211)
ERD6	AF073305	395	COP1 protein	TC22849 (735)
ERD57	AF074738	144	Maize DNA-binding protein	TC26129 (3135)
ERD67	AF074747		Rice putative AP2-domain-containing protein	TC16635 (435)
ERD69	AF074749		Rice putative AP2-domain-containing protein	TC16635 (435)
ERD74	AF074752	229	AP2-domain-containing protein	SS298 (406)
ERD79	AF074757	211	Glycine-rich RNA-binding protein	C2839 (970)
ERD91	AF074767	210	Glycine-rich RNA-binding protein	S15624 (900)
LRD7	AF074774	197	Pollen-specific protein	
LRD15	AF074781	100	CCAAT-binding factor	TC23518 (735)
LRD16	AF074782		CCAAT-binding factor	AF074781 (2297)
LRD17	AF074783	135	CUC/NAM protein	C27594 (492)
LRD18	AF074784	130	CUC/NAM protein	C28512 (1198)
LRD3	AF074801	72	Polycomb-like nuclear protein	
Inducible proteins				
ERD4	AF074727	155	Stress-inducible protein	S1571A (641)
ERD63	AF074744	211	Salt-inducible protein	
ERD80	AF074758	750	Salt-stress-inducible protein	S2603 (1954)
ERD89	AF074766	125	*Sorghum* pathogen-induced 1 (PT1)	
LRD2	AF074769	70	Early jasmonate-inducible gene	
LRD3	AF074770	185	Zn-inducible protein	S15878 (600)
LRD8	AF074775	313	Zn-inducible protein	AA752898 (712)
LRD38	AF074800	72	Jasmonate-inducible gene	
LRD41	AF074802	700	Salt-stress-inducible protein	C25168 (1759)
Miscellaneous				
ERD52	AF074733	396	60S Ribosomal protein L27A	TC21974 (1829)
ERD55	AF074736		Mei-2-like protein	TC21552 (2228)
ERD64	AF074744	446	40S Ribosomal protein	C756A (892)
ERD65	AF074745	242	Retrotransposon	

(*continued*)

Table 1. (*continued*)

ERD87	AF074764	70	Glutaredoxin	TC26114 (405)
ERD86	AF074763	175	dnaJ domain heat shock protein	TC25259 (4264)
LRD4	AF074771	420	40S Ribosomal protein	TC22516 (2247)
Novel				
ERD3	AF074726	289	Uncharacterized wheat EST	
ERD10	AF074729	<70	Unknown	
ERD13	AF074730	186	Uncharacterized maize EST	
ERD16	AF074731	<70	Unknown	
ERD60	AF074740	<70	Unknown	
ERD61	AF074741	198	Uncharacterized rice EST	
ERD66	AF074746	452	Uncharacterized maize EST	
ERD68	AF074748	<70	Unknown	
ERD75	AF074753		Uncharacterized rice EST	D43228 (1094)
ERD78	AF074756	<70	Unknown	S0325 (563)
ERD85	AF074762	89	Uncharacterized rice EST	D15565 (480)
ERD51	AF074732		Uncharacterized rice EST	D48842 (1183)
ERD54	AF074735		Uncharacterized rice EST	TC16318 (720)
ERD82	AF074760	<72	Unknown	
ERD88	AF074765	206	Uncharacterized rice EST	TC20920 (1776)
LRD6	AF074773		Uncharacterized rice EST	TC24226 (2927)
LRD10	AF074777		Uncharacterized rice EST	TC28173 (2644)
LRD11	AF074778		Uncharacterized rice EST	TC24226 (1447)
LRD19	AF074785		Uncharacterized rice EST	AW155020 (216)
LRD20	AF074786		Uncharacterized rice EST	TC17242 (3457)
LRD25	AF074791		Uncharacterized rice EST	TC19535 (3507)
LRD26	AF074792		Uncharacterized rice EST	AU73592 (483)
LRD37	AF074799		Uncharacterized rice EST	TC17040 (3988)
LRD43	AF074804		Uncharacterized rice EST	TC17040 (4093)
LRD48	AF074807		Uncharacterized rice EST	TC19535 (3567)

organisms. In both searches BLAST scores greater than 70 were taken to indicate sequence similarity. Alternatively, a WU-Blast score greater than 250 obtained against any rice EST was taken as significant. A summary of the sequence analysis is provided in table 1. The GenBank Accession numbers of the rice cDNAs are given. A putative functional assignment is also shown for those cDNAs that bear similarities to known genes in other species or in rice itself. Of the 80 sequences submitted to GenBank, about 68% showed similarity to known genes, while 32% represented potentially unknown/uncharacterized cDNAs. Despite the fact that these cDNAs were identified from a small set of random clones, they might comprise unique sequences since they are isolated from panicles at early stages of development. Not surprisingly, we find rather diverse types of genes expressed. The group of cDNAs with homologies to structural genes from other systems constitutes approximately 24% and perhaps represents housekeeping genes, expected to be present in any cell type. Similarly a number of genes were identified to possibly be involved in cell–cell signalling on the basis of their homologies to plant protein kinases/phosphatases and leucine-rich-repeat (LRR)-domain-containing proteins that are known to have diverse functions in signal transduction pathways. The best-characterized LRR-motif-containing protein from rice is Xa21 with a clearly demonstrated function in rice blight resistance (Song *et al*. 1995). It is attractive to speculate that the LRR protein (ERD53) identified in our cDNA screen

plays a role in signalling that controls cell proliferation, possibly like the product of the *Arabidopsis CLV1* gene, a LRR receptor kinase involved in regulating cell proliferation through cell–cell communication (Clark *et al*. 1997). The expression of *ERD53*, as is already known for *CLV1*, is not panicle or inflorescence specific (see figure 2). Therefore, a hit with a previously known EST from a different rice cDNA library is not unexpected. Protein kinases are in general well documented to participate in signalling pathways for cell proliferation, and their representation in this library is not surprising since panicles at early stages of development are largely proliferating tissues.

Other cDNAs displayed similarities to regulatory genes or putative transcription factors. In a few cases the identified homologues are known to play a specific role in floral development. For instance cDNAs (ERD2, LRD17 and 18) with homology to *CUC* genes of *Arabidopsis* (Aida *et al*. 1997) or the *NAM* gene of petunia (Souer *et al*. 1996) were identified at both stages of development studied here. *CUC, NAM*, and the very recently identified homologue in *Arabidopsis, NAP*, play developmental roles in establishment of the shoot apical meristem and floral development (Sablowski and Meyerowitz 1998). The suggested plant-specific function is positional definition of cells in developing meristems and organ primordia. While proteins containing NAC domain define a gene family, their tissue-specific and development-specific regions of synthesis are likely to give clues on their function. The rice cDNA ERD74 is

a member of a different gene family, with the sequence-specific DNA-binding motif found in transcription factors responding to ethylene. A few members of this rather large gene family are known to play a role in flower development. The *AP2* and *ANT* genes of *Arabidopsis* belong to this group and they contribute to floral meristem and organ identity or ovule development, respectively (Jofuku *et al.* 1994; Klucher *et al.* 1996). ERD74 has considerable homology to the *ANT* gene. In addition, putative DNA-binding (ERD57) and RNA-binding proteins (ERD79) are identified in this collection of cDNA sequences and in some instances their homologues in other plant systems are suggested to be tissue specific. We have also identified a cDNA with homology to *Arabidopsis COP1*, which encodes a regulator of photomorphogenesis unique in containing both Zn^+-finger DNA-binding motif and WD-40 repeat motifs that mediate protein–protein interactions (Deng *et al.* 1992). Yet other clones exhibited significant similarity to genes whose expression is situation specific, such as salt stress (ERD80) or abiotic stress (ERD4). Their representation in the library possibly reflects their basal level of expression in uninduced tissues. The last rice cDNA in this collection that requires specific mention is the one with similarity to the polycomb group of proteins, the prototype being the *POLYCOMB* gene of *Drosophila melanogaster*. Genes homologous to *POLYCOMB* function as chromatin reorganizers and transcriptional repressors of homeotic genes. Genes with analogous functions that are similar to *POLYCOMB* are known in *Arabidopsis* (Goodrich *et al.* 1997).

Genome organization and expression analysis of the database-matched random cDNA clones

A functional analysis of these cDNAs is necessary to provide information on their roles *in vivo*. As a beginning towards this analysis a few cDNAs were chosen for genomic characterization by Southern analysis as well as for their RNA expression analysis by Northern blots. Genes that were predicted by sequence homology to belong to gene families were found to be so by Southern analysis. For example, the gene for ERD74 — a potential transcription factor with EREB (ethylene response element binding) motif — is a member of a rice gene family comprising at least six or seven closely related members (figure 1B), which were detected even under highest stringency of hybridization. Similarly, the LRR-repeat-containing factor encoded in the cDNA ERD53 is a member of a gene family (figure 1C). Yet other cDNAs like the *COP1* homologue (ERD6) appear to be single copy in the rice genome (figure 1A). A functional analysis for single-copy genes can begin with the generation and analysis of loss-of-function mutants and studies on the effects of their ectopic expression. For those cDNAs that belong to gene families functional analysis is not so straightforward. In these latter cases, knowing their tissue-specific and cell-type-specific expression pattern may
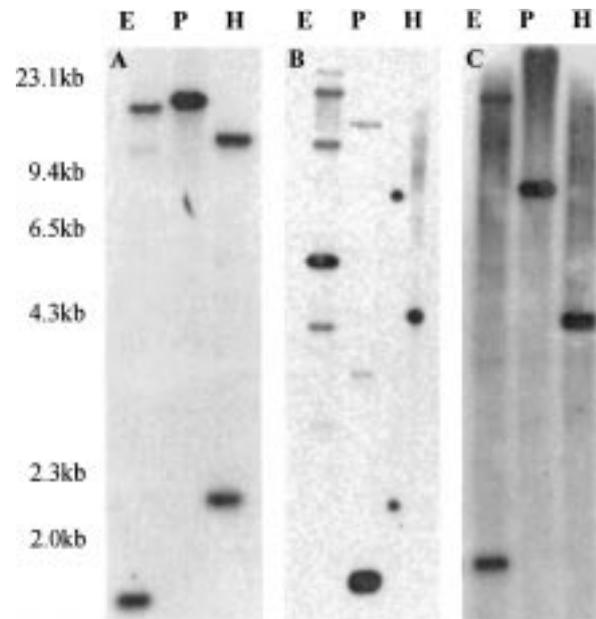

Figure 1. Genomic organization of three randomly selected cDNAs. Southern blot analysis of IR64 genomic DNA digested with *Eco*RI, *Pst*I or *Hin*dIII (abbreviated as E, P or H above the respective lanes) was taken for probing. The DNA digests in panel A were probed with the cDNA ERD6, those in panel B with the cDNA ERD74, and those in panel C with the cDNA ERD53. The migration positions of the marker DNA fragments from a λ *Hin*dIII digest are shown to the extreme left.

provide insights into their function. We therefore carried out Northern blot analysis using specific cDNA probes and total RNA from different parts of the adult plant, i.e. leaves, leaf sheaths that enclose panicles, panicles at early and late stages of development, floral/spikelet outer nonreproductive whorls and floral reproductive whorls. In addition, young roots and shoots were taken from seedlings. mRNA integrity and quantitation of the amount loaded was determined by probing the same RNA samples for a known rice gene encoding a ribosomal protein. These analyses showed that some cDNAs are expressed at varying levels in different aerial tissues of adult plants. The putative LRR-domain-containing factor in ERD53 and the DNA-binding protein in ERD57 are not expressed or are expressed at very low levels in roots and shoots of young seedlings. They are expressed poorly, if at all, in the floral/spikelet organs lemma, palea, stamens and carpels (figure 2). They are however expressed at varying levels in young leaf sheaths and in the developing panicle rachis and branches (figure 2). Contrastingly, the transcripts from ERD74 with homology to *AP2* and *ANT* genes of *Arabidopsis* are expressed almost constitutively and this is similar to the previously known constitutive RNA expression of *AP2* (Jofuku *et al.* 1994). It is surmised that the flower-specific phenotypes of *ap2* mutants are brought about through post-transcriptional regulation of *AP2*. ERD80, as predicted by its sequence homology, is induced by abiotic salt stress (figure 3). This cDNA is expressed at very low constitutive levels in
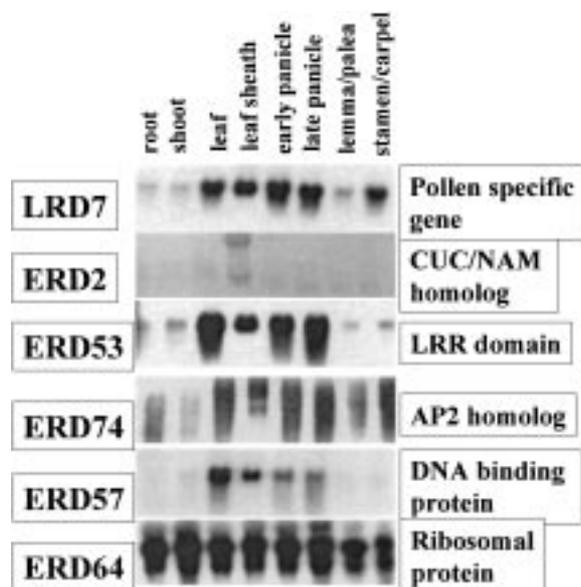
Figure 2. Spatial and temporal expression patterns for a few cDNAs. RNA expression pattern and levels for a few randomly selected cDNA clones were tested in various rice tissues. Total RNA was isolated from adult plant parts: leaves, leaf sheaths, early panicles (1 cm), late panicles (2 to 3 cm), nonreproductive floral structures lemma and palea, and reproductive tissues stamen and carpel. Juvenile shoots and roots were obtained from 15-to-20-day-old seedlings. Thirty μg total RNA from each tissue was fractionated on 1.2% formaldehyde agarose gels, blotted onto membranes, and hybridized with the specific cDNA probe indicated to the left. The class of genes that these cDNAs are likely to belong to is also shown to the right. Hybridization of the blots with a probe encoding a rice 40S ribosomal protein was used to estimate the relative amounts of mRNA present in each lane.
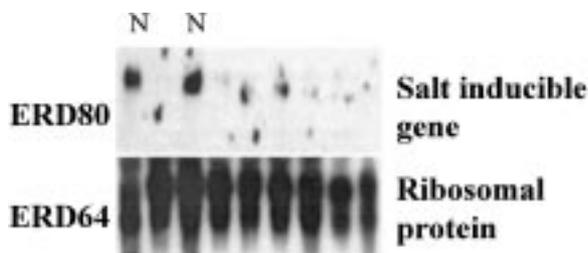

Figure 3. Induced expression of ERD80. Thirty μg of total RNA from various parts of adult field-grown rice plants, i.e. leaf sheaths, early panicles, late panicles, lemma and palea, stamens and carpels were taken for analysis. Young roots and shoots from rice seedlings of age 15 to 20 days post-germination were taken for RNA preparation either before or after treatment with 0.1 M NaCl. All of these RNAs were electrophoresed on formaldehyde–agarose gels and then probed with the cDNA ERD80 which is homologous to a salt-induced gene. Hybridization of the blot with a gene for a 40S ribosomal protein provided an estimate of relative amounts of mRNA in the lanes.

seedling tissues as well as adult plant tissues. In seedlings that have been subjected to salt stress the transcripts are induced several fold. In yet other cDNAs where the sequence analysis suggests homology to tissue-specific factors of other plant systems, the expression pattern of

the rice cDNA is not differentially regulated in tissues tested so far. For example, LRD7 is a cDNA with significant homology to a suggested pollen-specific factor. Our data show that it is expressed in rice stamens but it is also expressed at high levels in branching panicles, leaves and leaf sheaths (figure 2). For some other cDNAs the RNA levels were undetectable (data not shown).

Thus we have isolated a battery of rice genes that encode products with sequence homology to structural proteins, signalling molecules, transcription factors, and regulators of gene expression, and yet others that encode currently unknown factors. The characterization of their functions would give clues on their precise *in vivo* roles. This is feasible through the analysis of transgenic rice plants that express these cDNAs either constitutively or express antisense versions of these cDNAs to downregulate endogenous gene expression. In addition, analysis of clones that are apparently unique might lead to elucidation of new functions for new classes of proteins. These studies are a source for developmentally regulated promoters useful in tissue-targeted expression of desirable foreign genes. Likewise, these cDNAs offer potential molecules for future microarray analysis in genomic studies. In addition, once mapped on the rice genome these clones could serve as start points for studies of homologous genes in other cereal genomes.

References

Aida M., Ishida T., Fukaki H., Fujisawa H. and Tasaka M. 1997 Genes involved in organ separation in *Arabidopsis*: An analysis of the *cup-shaped cotyledon* mutant. *Plant Cell* **9**, 841–857.

Altschul S. F., Madden T. L., Schaffer A. A., Zhang J., Miller W. and Lipman D. J. 1997 Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucl. Acids Res.* **25**, 3389–3402.

Chung Y., Kim S., Finkel D, Yanofsky M. F. and An G. 1994 Early flowering and reduced apical dominance result from ectopic expression of a rice MADS box gene. *Plant Mol. Biol.* **26**, 657–665.

Church G. M. and Gilbert W. 1984 Genomic sequencing. *Proc. Natl. Acad. Sci. USA* **81**, 1191–1195.

Clark S. E., William R. W. and Meyerowitz E. M. 1997 The *CLAVATA1* gene encodes a putative receptor kinase that controls shoot and floral meristem size in *Arabidopsis. Cell* **89**, 575–585.

Cooke R., Raynal M., Laudie M., Grellet F., Delseny M., Morris P. C. *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.

Crawford N., Campbell W. H. and Davis R. W. 1986 Nitrate reductase from squash: DNA cloning and nitrate regulation. *Proc. Natl. Acad. Sci. USA* **83**, 8073–8076.

Deng X. W., Matsui M., Wel N., Wagner D., Chu A. M., Feldman K. A. and Quail P. H. 1992 *COP1*, an Arabidopsis regulatory gene, encodes a protein with both zinc-binding finger motif and a Gβ homologous domain. *Cell* **71**, 791–801.

Goodrich J., Puangsomlee P., Martin M., Long D., Meyerowitz E. M. and Coupland G. 1997 A polycomb-group gene regulates homeotic gene expression in *Arabidopsis. Nature* **386**, 44–51.

Hoshikawa K. 1989 *The growing rice plant: an anatomical monograph.* Nosan Gyoson Bunka Kyokai (Nobunkyo), Tokyo.

Jofuku K. D., den Boer B. G. W., Montagu M. and Okamuro J. K. 1994 Control of *Arabidopsis* flower and seed development by the homeotic gene *APETALA2. Plant Cell* **6**, 1211–1225.

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

Klucher K. M., Chow H., Reiser L. and Fischer R. L. 1996 The Anintegumenta gene of *Arabidopsis* required for ovule and female gametophyte development is related to the floral homeotic gene Apetala 2. *Plant Cell* **8**, 137–153.

Kushalappa K., Suresh Kumar M. B., Patwardhan L. and Vijayraghavan U. 1996 Isolation and characterization of genes controlling flowering in rice. In *Rice genetics III* (Proceedings of the Third International Rice Genetics Symposium, ed. G. S. Khush), pp. 888–893. International Rice Research Institute, Manila.

Lin X., Kaul S., Rounsley S., Sha T. P., Banito M. I., Town C. D. *et al.* 1999 Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana. Nature* **402**, 761–768.

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

Rogers S. O. and Bendich A. 1988 Extraction of DNA from plant tissues. In *Plant molecular biology manual* (ed. S. B. Gelvin and R. A. Schilperoort), pp. A6/1–10. Kluwer, Dordrecht.

Sablowski R. W. and Meyerowitz E. M. 1998 A homolog of *no apical meristem* is an immediate target of the floral homeotic genes *APETALA3/PISTILLATA. Cell* **92**, 93–103.

Sambrook J. E., Fritsch F. and Maniatis T. 1989 *Molecular cloning. A laboratory manual.* Cold Spring Harbor Laboratory Press Cold Spring Harbor.

Sasaki T., Song J., Koga-Ban Y., Matsui E., Fang F., Higo H. *et al.* 1994 Toward cataloguing all rice genes: Large-scale sequencing of randomly chosen rice DNAs from a callus cDNA library. *Plant J.* **6**, 615–624.

Song W. Y., Wang G. L., Chen L. L., Kim H. S., Pil Y., Holsten T. *et al.* 1995 A receptor kinase like protein encoded by the rice disease resistance gene Xa21. *Science* **270**, 1804–1806.

Souer E., van Houwelingen A., Kloos D., Mol J. and Koes R. 1996 The *no apical meristem* gene of petunia is required for pattern formation in embryos and flowers and is expressed at meristem and primordia boundaries. *Cell* **85**, 159–170.

Uchimiya H., Kldou S. I., Shimazaki T., Aotsuka S., Takamatsn S., Nishi R. *et al.* 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.

Vijayraghavan U. 1996 Genetic regulation of flower development. *J. Biosci.* **21**, 379–395.

Weigel D. and Meyerowitz E. M. 1995 The ABCs of floral homeotic genes. *Cell* **78**, 203–209.