Plant Molecular Biology **15**: 257–268, 1990.

Characterization of the rice (*Oryza sativa*) actin gene family

David McElroy,[1] Madge Rothenberg,[2,3] Kimberly S. Reece[2,4] and Ray Wu[1]*
[1] *Field of Botany and* [2] *Section of Biochemistry, Molecular and Cell Biology, Biotechnology Building, Cornell University, Ithaca, New York 14853, USA (*author for correspondence); Present addresses:* [3] *Section of Plant Breeding, Emerson Hall, Cornell University, Ithaca, New York 14853, USA;* [4] *Department of Biology, College of William and Mary, Williamsburg, Virginia 23185, USA*

Received 2 April 1990; accepted in revised form 4 May 1990

Key words: actin genes, differential expression, gene evolution, gene family, rice

Abstract

Four rice (*Oryza sativa*) genomic actin genes have been characterized. The rice actin genes show a conservation of intron number and position that is characteristic of plant actins. Sequence comparisons revealed that the plant actins generally have a high degree of inter- and intraspecific sequence heterogeneity. However, one rice actin gene has a much higher degree of nucleotide sequence identity to a previously isolated actin sequence from *Arabidopsis thaliana* than to any other plant actin gene. This leads us to suggest that the two sequences may represent functionally homologous genes which arose from an ancient actin gene lineage that was separated by the divergence of the dicot and monocot plants. Genomic DNA blot analysis showed that the rice actin gene family contains at least eight unique members. RNA hybridization analysis revealed that individual rice actin genes can display different patterns of transcript accumulation. The observed differences in sequence and transcript accumulation patterns suggest that the individual rice actin genes may differ in their transcriptional regulation and/or cellular function.

Introduction

Cytoplasmic actin is an essential component of the eukaryotic cell cytoskeleton and has been well studied in a number of animal systems [24]. Although actin is less well characterized in higher plants than it is in animals, it is believed to play an important role in plant cell cytoplasmic streaming, cell shape determination, cell division, organelle movement and extension growth [30, 34].

Actin gene families in higher plants have been found to consist of at least six genes in soybean (*Glycine max*) [31, 32], six genes in maize (*Zea mays*) [32], three genes in *Arabidopsis thaliana* [23], five genes in potato (*Solanum tuberosum*) [4], ten genes in tomato (*Lycopersicon esculentum*) [2] and two genes in carrot (*Daucus carota*) [36]. It has also been reported that actin is encoded by a complex supergene family in *Petunia hybrida* [1]. A feature of all these plant actin genes has been their relatively high degree of intra- and interspecific nucleotide sequence divergence and a conservation of coding region intron positions that is not found in animal actin genes [12, 22, 27, 32]. Preliminary results suggest that various members of the soybean actin gene family may show different patterns of organ- and tissue-

specific expression [11, 20]. However it is not known if these soybean actin gene transcripts display any differential patterns of developmental-stage specificity.

Previous work in our laboratory resulted in the isolation and sequencing of the coding regions of four rice (*Oryza sativa*) actin genes [26] and the complete structural characterization of the rice actin 1 gene [18]. Subsequent work led to the identification of the rice actin 1 gene promoter as an efficient constitutive regulator of foreign gene expression in transformed rice tissue [19]. In this paper we further characterized the four previously isolated rice actin sequences [22]. DNA sequence analysis of the four rice actin genes, along with a number of other plant actin sequences, indicates that their DNA and deduced amino acid sequences generally display a high degree of intra- and interspecific divergence. However, a comparison between the rice actin 1 gene and the AAc1 actin gene of *Arabidopsis thaliana* revealed a degree of sequence identity that leads us to suggest that these two sequences may represent orthologous genes.

Genomic DNA blot analysis revealed that there are at least eight actin-related sequences in rice. In order to determine whether the expression of the four isolated members of the rice actin gene family are differentially regulated, their respective transcript levels were examined during the course of rice plant vegetative development. Our results indicate that RNA transcribed from individual rice actin genes can display either a constitutive or developmentally differentiated pattern of accumulation and that the members of the rice actin gene family differ from each other in the developmental-stage and tissue-specific abundance of their respective transcripts.

Materials and methods

DNA sequence analysis

The isolation and sequencing of the rice actin genes has been described previously [26]. The alignment of actin coding regions and comparison of their deduced amino acid sequences were done using the Microgenie™ sequence analysis program [25]. For analysis of the number of synonymous and nonsynonymous replacement substitutions, in all of the pairwise comparisons, the program of Li, Wu and Luo [15] was employed. Sequence comparison phenograms were constructed using the unweighted pair-group method (UPGMA) for actin coding regions with the sequential, agglomerative, hierarchical and nested (SAHN) clustering program of the numerical taxonomy systematics (NTSYS) package [5, 28, 33]. The Mantel t-test approximation was applied to test the goodness of fit for the resulting phenograms [17].

Isolation and characterization of non-specific and gene-specific actin probes

Restriction fragments containing portions of the four previously isolated rice actin genes [26] were subcloned into pBluescriptII-KS (Stratagene) for sequence analysis and for the production of ^{32}P-labeled probes, following standard procedures [16].

Genomic DNA preparation and analysis

Rice genomic DNA was isolated from young greenhouse-grown plants (*Oryza sativa* var. *indica*, cv. IR36) as described by Dellaporta *et al.* [3] with the addition of a phenol-chloroform extraction prior to the final ethanol precipitation.

For DNA blot analysis 10 μg samples of rice IR36 genomic DNA were digested with *Hin*d III, fractionated by electrophoresis using 0.9% agarose gels and transferred to GeneScreen Plus membranes (Dupont) for hybridization by the procedure of Bernatsky and Tanksley [2].

RNA preparation

Total RNA was isolated from rice IR36 tissue collected from plants at various stages of develop-

ment. Seeds were germinated aseptically on two layers of Whatman 3MM paper in Petri dishes, or larger containers in a growth chamber. Roots and/or shoots were collected separately at various times after germination. All tissues were ground to a fine powder in liquid nitrogen. The powder was transferred to a polypropylene tube. One ml of extraction buffer (100 mM NaOAc, 1% SDS, pH 5.0) was added per gram of tissue and the samples were extracted twice with an equal volume of phenol : chloroform : isoamyl alcohol (25 : 24 : 1). KCl was added to the aqueous phase for a final concentration of 0.3 M. After a 30 minute incubation on ice the samples were centrifuged for 10 minutes at 8,000 × *g*. The supernatant was transferred to a new tube and the RNA was precipitated with LiCl as described by Wallace [39].

RNA analysis

For RNA slot-blot analysis, varying amounts of total RNA were splotted onto nitrocellulose filters using a Minifold Slot-Blot System (Schleicher and Schuell). RNA (5 μl) was mixed with 20 μl of 20 × SSC and 30 μl of formamide. The samples were heated at 65 °C for 5 minutes, placed on ice, and loaded into wells of the slot blot apparatus. Prehybridization, hybridization and washes were as described for DNA hybridization analysis.

For reverse RNA blot analysis 5 μl of DNA from each of the clones containing the 3′ ends from each of the four previously isolated rice actin sequences (gene-specific probes) were added to 25 μl of 0.3 M NaOH, 6 × SSC, heated at 80 °C for 15 minutes, neutralized with 25 μl of 2.0 M NH_4OAc, and placed on ice. The samples were transferred to nitrocellulose filters using the slot-blot apparatus described above. The concentrations of the DNA samples were adjusted so that equimolar amounts of each actin insert were loaded in each well. Final insert DNA amounts were 0.15, 0.3, and 0.6 pmol. The filters were hybridized with radioactively labeled single-stranded cDNA transcribed from rice IR36 total RNA as follows. After heating 10 μg of RNA at 80 °C for two minutes and cooling on ice, first-strand cDNA was transcribed as described by Krug and Berger [14] with the following modifications: incubation time was one hour; the chase step was omitted; 80 μCi ^{32}P-dATP (Amersham, 3000 Ci/mmol) was used instead of ^{32}P-dCTP. The cDNA containing solution was treated with NaOH to hydrolyze the RNA, neutralized and then added to the hybridization solution [14]. Prehybridization, hybridization and washes were carried out as described for DNA blot hybridizations.

Statistical analysis of RNA blot data was carried out using the StatView™ SE$^+$ program (Abacus Concepts, Inc.) on an Apple Macintosh SE computer.

Results

Characterization and sequence analysis of four rice actin genes

It has been reported previously that four unique actin genes, designated RAc1, RAc2, RAc3 and RAc7, were isolated from a rice (*Oryza sativa* cv. IR26) λEMBL4 genomic library, subcloned into pUC13 and M13 vectors and their nucleotide sequence determined [26]. A sequence alignment between a 1.5 kb RAc1 cDNA clone [18] and the four genomic actin clones was made to determine the probable number and locations of the introns within their respective coding regions. This analysis identified three introns located at the same position in all four rice actin coding regions, as that reported for all other previously characterized plant actin genes. Restriction maps of the four rice actin genes, with the positions of their coding region exons indicated, are presented in Fig. 1.

The nucleotide and amino acid sequences of the four rice actin genes and the previously published actin sequences from soybean [31, 32], maize [32], *Arabidopsis thaliana* [23] and lodgepole pine [13] were compared (Table 1). Proteins were aligned and their amino acid sequences were compared by determining their percentage similarity (Table 1, upper right above the

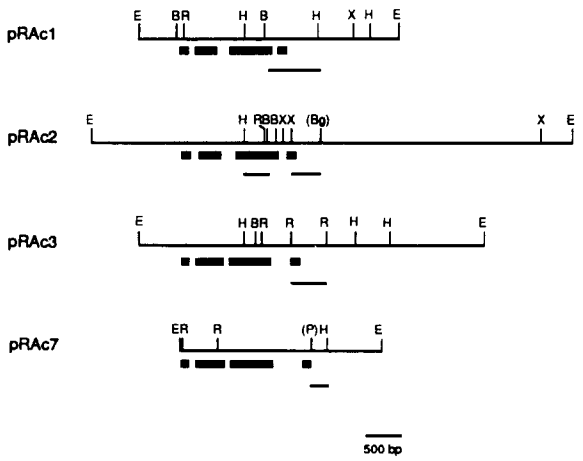

Fig. 1. Restriction maps of pRAc1, pRAc2, pRAc3 and pRAc7. Restriction sites are designated by vertical lines: E, *Eco* RI; B, *Bam* HI; R, *Eco* RV; H, *Hin*d III; X, *Xba* I; Bg, *Bgl* II; P, *Pst* I. The solid rectangles represent coding region exons. Actin 3′-end gene-specific probes were: a 0.9 kbp *Bam* HI-*Hin*d II fragment from pRAc1: a 0.5 kbp *Xba* I-*Bgl* II fragment from pRAc2; a 0.8 kbp *Eco* RV-*Eco* RV fragment from pRAc3 and a 0.3 kbp *Pst* I-*Hin*d III fragment from pRAc7. The non-specific actin gene probe was a 0.26 kbp *Hin*d III-*Bam* HI fragment from pRAc2. Horizontal lines below the restriction maps indicate the position of these probes. Only those *Bgl* II and *Pst* I sites which delineate hybridization probes have been placed on the maps and these are enclosed in parentheses.

diagonal) using the Beckman Microgenie™ program [25]. The percentage nucleotide replacement substitution (corrected for multiple hits) for the aligned actin coding regions (Table 1, lower left below the diagonal) was determined using the program of Li, Wu and Luo [15]. This analysis revealed that the rice actins are as diverged from each other, at both the nucleotide and amino acid levels, as they are from the other plant actins. Rice actins as a group display 5.7–13.3% nucleotide replacement substitution and 83.4–90.5% amino acid similarity, while rice actin vs. other plant actin comparisons yielded 3.6–20.4% nucleotide replacement substitution and 80.7–93.9% amino acid similarity. While this high degree of inter- and intraspecific sequence divergence appears to be a general feature of the plant actins we observed that the rice RAc1 gene shows close sequence similarity to the *Arabidopsis thaliana* actin AAc1 actin gene, with 93.3% amino acid sequence similarity and only 3.6% nucleotide replacement substitution. This degree of sequence identity between the RAc1 and AAc1 genes, when viewed against the relatively high amount of sequence divergence within the plant actins, leads us to suggest that the two sequences may represent

Table 1. Coding region amino acid sequence comparisons[a] (% similarity) and percentage replacement nucleotide substitution[b] in pairwise comparisons between plant actin genes[c].

	RAc1	RAc2	RAc3	RAc7	MAc1	SAc1	SAc3	PAc1	AAc1
RAc1	–	90.5	89.1	86.0	87.0	87.3	89.6	87.6	93.9
RAc2	5.7	–	87.4	84.8	86.8	83.7	87.9	88.3	90.2
RAc3	6.5	8.4	–	83.4	85.4	82.8	86.5	85.2	90.7
RAc7	11.4	12.1	13.3	–	87.0	80.7	83.9	84.5	87.0
MAc1	7.2	8.1	8.7	10.6	–	82.8	86.2	84.5	89.1
SAc1	7.3	10.1	10.8	16.3	10.3	–	85.1	85.1	88.1
SAc3	6.0	8.2	7.8	13.3	7.7	8.7	–	90.7	89.9
PAc1	6.7	7.8	9.6	20.4	9.0	8.9	5.9	–	90.7
AAc1	3.6	7.2	6.9	12.3	7.2	7.0	6.5	7.2	–

[a] Comparisons of the deduced protein sequences (upper portion of matrix) was done using the Beckman Microgenie™ program [25] to determine percent similarity. In determining percent similarity, the following pairs of amino acids are considered to be the same: Asp vs. Glu, Asn vs. Gln, Lys vs. Arg, Leu vs. Val and Ser vs. Thr.

[b] The percentage replacement substitution (bottom portion of matrix) was determined by the program of Li, Wu, and Luo [15].

[c] The plant actin sequences are abbreviated as follows: rice (*Oryza sativa*), RAc1, RAc2, RAc3 and RAc7; maize (*Zea mays*), MAc1; soybean (*Glycine max*), SAc1 and SAc3; lodgepole pine (*Pinus contorta*), PAc1 and *Arabidopsis thaliana*, AAc1.

genes which were separated by the divergence of the monocot and dicot plants.

Using the information in Table 1 and assuming a linear rate of plant actin gene evolution [12] we generated a phenogram using the unweighted pair-group method (UPGMA). The resulting phenogram (Fig. 2) indicates that if the branch point for divergence of the rice RAc1 and *Arabidopsis* AAc1 actins corresponded to the divergence of the monocots from the dicots during the evolution of the angiosperms, then the gene duplications that gave rise to the other plant actins must have occurred early in the evolution of the vascular plants. Such gene duplication events must have occurred before the division of the gymnosperms and angiosperms from a common ancestral plant. The origin of the divergence of the lodgepole pine PAc1 gene and soybean SAc3 gene may have coincided with the organismal split that separated the gymnosperms and angiosperms, since these two genes are less divergent than (most of) the other angiosperm actin sequences.

Plant actin protein display structural domain conservation

While the precise functions of plant actin isoforms have yet to be determined it is believed that these isoforms play different roles within the plant cell and interact with different components of the plant cell cytoskeleton. Consequently plant actins should contain a number of conserved protein-protein interaction sites whose structural requirements would act to constrain the sequence divergence of functionally equivalent plant actin isoforms. When the protein sequences of the plant actins are aligned there exist stretches of amino acids which are conserved both within plants and between plant and animal actin proteins (Fig. 3a), despite their generally high degree of inter- and intraspecific sequence divergence. These conserved regions within the plant actin proteins often coincide with sites which are believed to function in actin : actin and/or actin : actin-binding protein interactions in animal systems (Fig. 3b) [38].

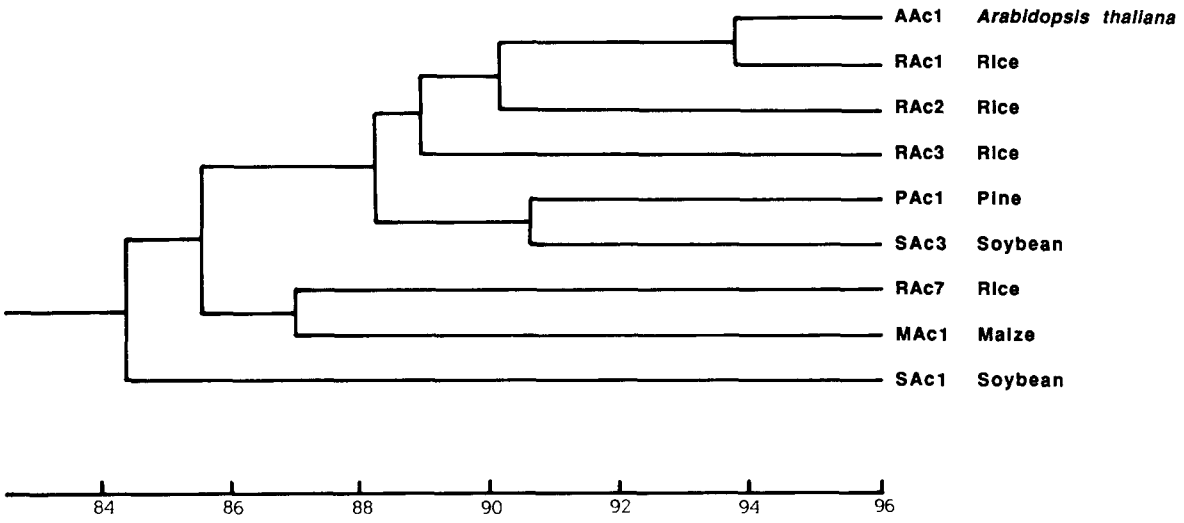

Fig. 2. Deduced phenogram of plant actin proteins. The sequential, agglomerative, hierarchical and nested clustering program of the numerical taxonomy systematics package [5, 28, 33] was used to construct an unweighted pair-group, phenetic tree for the plant actins. The scale on the bottom correlates branch lengths with the percentage similarity of the plant actin amino acid sequences (data from Table 1). A Mantel t-test [17] gave a t value of 2.916 ($p > 0.9982$, where p represents the probability that this phenogram was not generated by chance) for this phenogram.

A.

```
          A              B              C              D              E
     11        22  27        38  47        60  64        76  83        98
SAc1  VCDNGTGMVKAG....DAPRAVFPSIVG....VGMGQKDAYVGDEA....RGILTLKYPIEHG....DMEKIWHHTFYNELRV....
SAc3  VCDNGTGMVKAG....DAPRAVFGSIVG....VGMGQKD?YVGDEA....RGILTLKYPIEHG....DMEKIWHHTFYNELRV....
AAc1  VCDNGTGMVKAG....DAPRAVFPSIVG....VGMGQKDAYVGDEA....RGILTLKYPIEHG....DMEKIWHHTFYNELRV....
RAc1  VCDNGTGMVKAG....DAPRAVFPSIVG....VGMGQKDAYVGDEA....RGILTLKYPIEHG....DMEKIWHHTFYNELRV....
RAc2  VCDNGTGMVKAG....DAPRAVFPSIVG....VGMGQKDAYVGDEA....RGILTLKYPIEHG....DMEKIWHHTFYNELRV....
RAc3  VCDNGTGMVKAG....DAPRAVFPSIVG....VGMGQKDAYVGDEA....RGILTLKYPIEHG....DMEKIWHHTFYNELRV....
RAc7  VCDNGTGMVKAG....DAPRAVFPSIVG....VGMGQKDAYVGDEA....RGILTLKYPIEHG....DMEKIWHHTFYNELRV....
MAc1  VCDNGTGMVKAG....DAPRAVFPSIVG....VGMGQKDAYVGDEA....RDIITLKYPIEHG....DMR-IWHHTFYNELRA....
CAc1  VCDNGTGMVKAG....DAPRAVFPSIVG....VGMGQKDAYVGDEA....RGIITLKYPIEHG....DMR-IWHHTFYNELRA....

Cons* VCDNGTGMVKAG....DAPRAVFPSIVG....VGMGQKDAYVGDEA....RGILTLKYPIEHG....DMEKIWHHTFYNELRV....
                  *              *              * *            **             *
          9        20  25        36  45        58  62        74  81        96
HuCy  VVDNGSGMCKAG....DAPRAVFPSIVG....VGMGQKDSYVGDEA....RGILTLKYPIEHG....DMEKIWHHTFYNELRV....
HuCa  VCDNGSGLVKAG....DAPRAVFPSIVG....VGMGQKDSYVGDEA....RGILTLKYPIEHG....DMEKIWHHTFYNELRV....
```

```
          F              G              H              I              J
    173                  200  280              298  331                  357  363          377
SAc1  HAILRLDLAGRDLTDFLMKILTQRGY....TYNSIMKCDVDIRKDLYGN....VVAPSERKFGV-WIGGSILASLSTFQQMWI....EYDESGPSIVHKKCF
SAc3  HAILRLDLAGRDLTDHLMKILTERGY....TYNSIMKCDVDIRKDLYGN....VVAPPERKYSV-WIGGSILASLSTPQQMWI....EYDESGPSIVARKCF
AAc1  HAILRLDLAGRDLTDALMKILTERGY....TYNSIMKCDVDIRKDLYGN....VVAPPERKYSV-WIGGSILASLSTLQQMWI....EYDESGPSIVHRKCF
RAc1  HAILRLDLAGRDLTDYLMKILTERGY....TYNSIMKCDVDIRKDLYGN....VVAPPERKYSV-WIGGSILASLSTFQQMWI....EYDESGPSIVHRKCF
RAc2  HAILRLDRAGRDLTDSLMKILTERGY....TYNSIMKCDVDIRKDLYGN....VVAPPERKYSV-WIGGSILASLSTFQQMWI....EYDESGPAIVHRKCF
RAc3  HAILRLDLAGRDLTDCLMKILTERGY....TYNSVMKCDVDIRKDLYGN....VVAPPERKYSV-WIGGSILASLSTFQQMWI....EYDESGPSIVHRKCF
RAc7  HAILRLDLAGRDLTDNLMKILTERGY....TYNSIMKCDVDIEKDLYGN....VVAPPERKYSV-WIGGFILASLSTFQQMWI....EYDESGPGIVHMKCF
MAc1  HAILRLDLAGRDLTDHLMKILTERGY....TYNSIMKCDVDIRKDLYGN....VVAPPRRKYSV-WIGGSILASLSTFQQMWI....EYDETGPGIVHMKCF
CAc1  HAILRLDLAGRDLTDGLMKILTEKSI....TYNSIMKCDVDIRKDLYGN....VVAPPERKYSDLWIGGSILASLSTFQQMWI....EYDESGPSIVHRKCL

Cons* HAILRLDLAGRDLTD-LMKILTERGY....TYNSIMKCDVDIRKDLYGN....VVAPPERKYSV-WIGGSILASLSTFQQMWI....EYDESGPSIVHRKCF
                  *       *              *       *              **   ****   *        *            *  *  **
    171                  198  278              296  329                  355  361          375
HuCy  HAILRLDLAGRDLTDYLMKILTERGY....TFNSIMKCDVDIRKDLYAN....IIAPPERKYSV-WIGGSILASLSTFQQMWI....EYDESGPSIVHRKCF
HuCa  HAIMRLDLAGRDLTDYLMKILTERGY....TYNSIMKCDIDIRKDLYAN....IIAPPERKYSV-WIGGSILASLSTFQQMWI....EYDEAGPSIVHRKCF
```

B.

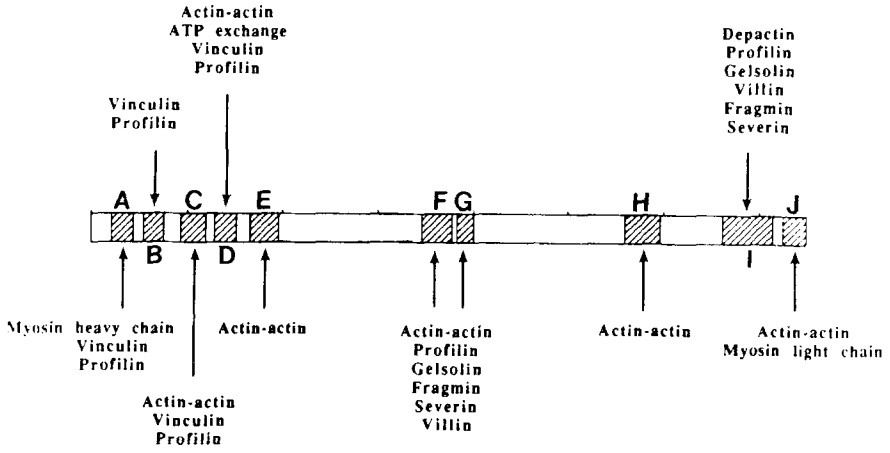

Fig. 3. A. Regions of amino acid sequence similarity between plant and animal actin proteins. The human β-cytoplasmic (HuCy) and α-cardiac (HuCa) actin sequences [10, 38] are shown as representative animal nonmuscle and muscle actin protein isoforms respectively. Plant actin sequence abbreviations are as in Table 1. A consensus sequence (Cons*) is indicated below each conserved plant actin region. An asterisk below a concensus sequence residue indicates that one (or more) of the plant actin proteins differs in its protein sequence from the consensus sequence at that residue. *B.* Graphic representation indicating the position of conserved amino acid regions (striped boxes) within plant and animal actin proteins. The sites of actin : actin and actin : actin-binding protein interactions inferred from studies of animal systems are indicated by vertical arrows. The equivalent regions in *A* and *B* are indicated by letter.

Rice actin is encoded by a gene family

The identification of four unique rice actin genes indicated that rice actin proteins are encoded by a gene family. To determine whether there were additional actin (or actin-like) sequences in rice, genomic DNA blots were hybridized with a 260 bp *Hind* III-*Bam* HI restriction fragment

from the third exon of RAc2 (Fig. 1), at low stringency (37 °C, 5× SSC, 50% formamide). This fragment was selected as a potential non-specific actin probe because it covered a region with high sequence identity between the 4 isolated rice actin genes, and because it was unlikely to span any other *Hin*d III or *Bam* HI sites in any as yet uncharacterized rice actin genes. Eight to ten restriction fragments hybridized to the RAc2 probe when total DNA was digested with *Eco* RI, *Bam* HI, or *Hin*d III. The hybridization pattern produced by *Hin*d III digestion of genomic rice DNA is shown in Fig. 4, lane A. These results,

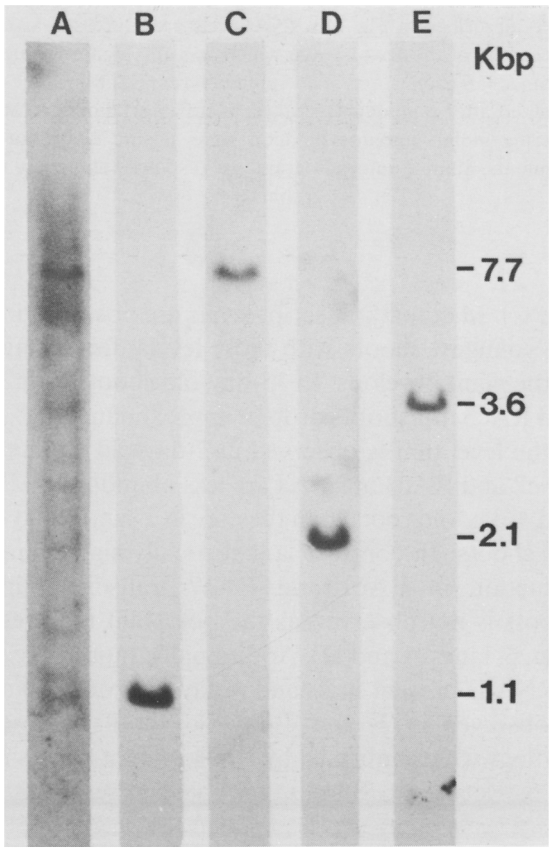

Fig. 4. Hybridization of actin gene probes to DNA blots of rice genomic DNA. *Hin*d III digests of rice DNA were hybridized with ^{32}P-labeled non-specific or gene-specific actin probes (Fig. 1). Lane A, the non-specific actin gene probe from pRAc2; lanes B to E, the actin gene-specific probes from the 3′-ends of pRAc1, pRAc2, pRAc3 and pRAc7 respectively. Hybridizations were carried out at 37 °C (lane A) or 42 °C (lanes B through E) in 5× SSC, 50% formamide.

along with the previously reported isolation of eight different classes of rice actin cDNA clones [18] suggest that the rice genome contains at least eight unique actin genes.

DNA blots of genomic DNA digested with *Eco* RI, *Bam* HI, or *Hin*d III were also hybridized with ^{32}P-labeled (putative) rice actin gene-specific probes, isolated from the 3′ end of each of the four rice actin genes (Fig. 1), under conditions of higher stringency (42 °C, 5× SSC, 50% formamide) than was used for the non-specific hybridization described above. The results of hybridizations to *Hin*d III-digested genomic DNA are shown in Fig. 4, lanes B to E, with the probes derived from RAc1, RAc2, RAc3, and RAc7 hybridizing respectively to the expected (from phage and plasmid restriction maps) 1.1 kbp, 7.7 kbp, 2.1 kbp, and 3.6 kbp *Hin*d III restriction fragments. Under the conditions employed each probe hybridizes specifically to the actin gene from which it was derived, confirming these DNA-DNA hybridization conditions as gene-specific.

The four rice actin genes display different patterns of transcript accumulation

RNA gel blots of total rice RNA were hybridized with ^{32}P-labeled actin gene-specific probes to determine the transcript number and size for each gene. The RNA gel blot hybridizations indicated that each actin gene encodes a unique transcript approximately 1.5 to 1.7 kb in length (result not shown). Based on the predicted exon sizes of the rice actin genes [18, 26], rice actin transcripts of this length would include 0.4 to 0.6 kb of transcribed but untranslated sequence.

The steady-state levels of rice actin transcripts were examined by slot-blot analysis of RNAs isolated from different tissues at varying times during early vegetative development. DNA probes derived from restriction fragments previously determined to be gene-specific for each of the 4 rice actin genes were used in excess and hybridized to equal amounts of total RNA from shoots 2, 4, 7, 13, and 35 days old, and from roots 2 and

4 days old. To insure that no cross-hybridization occurred between actin genes, hybridizations were carried out at 50 °C rather than 42 °C to compensate for the increased stability of DNA-RNA hybrids over DNA-DNA hybrids. Actin transcript levels at the various developmental stages were normalized relative to 17S rRNA levels by hybridizing duplicate filters with a 0.8 kbp *Sac* II fragment from the plasmid pRR217 containing the rice 17S and 25S ribosomal genes [37].

The transcript levels of RAc2 and RAc3 follow a similar pattern in the rice shoot tissues sampled over the time period examined (Fig. 5, lanes B and C). Maximal transcript levels are observed in the youngest shoots with these levels decreasing as the plant develops. In 35-day-old shoots RAc2 and RAc3 transcripts drop to approximately 20% of the level that is observed in 2-day-old shoots. RAc2 and RAc3 mRNAs are less abundant in 2- and 4-day-old roots than they are in 2- and 4-day-old shoots. In contrast, a statistically significant reduction in RAc1 and RAc7 transcripts in shoots is not observed as the rice plant matures (Fig. 6, lanes A and D). Additionally, more RAc7 mRNA is present in 2- and 4-day-old roots than is observed in 2- and 4-day-old shoots. These findings were confirmed by RNA gel blot analysis (K.S. Reece and R. Wu, unpublished results).

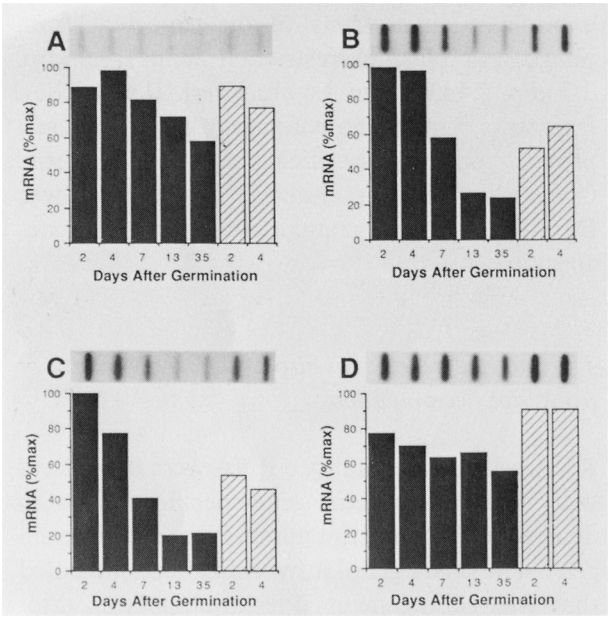

Fig. 5. Analysis of actin gene transcript levels in developing rice seedlings. The results of RNA slot-blots are displayed graphically. Total RNA was isolated 2, 4, 7, 13 and 35 days after germination and hybridized (at 50 °C, 5× SSC, 50% formamide) with ^{32}P-labeled probes isolated from: A, pRAc1; B, pRAc2; C, pRAc3; D, pRAc7. The solid bars indicate RNA isolated from rice seedling shoots and striped bars indicate RNA isolated from rice seedling roots. RNA levels are expressed as a percentage of the maximum observed for each actin gene. The numbers represent averages obtained after scanning several replicate autoradiograms from slot-blots containing 0.5 μg, 1.0 μg, 2.0 μg, 3.0 μg, 5.0 μg and 6.0 μg of total RNA; therefore the mean maximum value, for any one gene's transcript, may not always equal 100%. Sample autoradiograms are displayed above each graph.

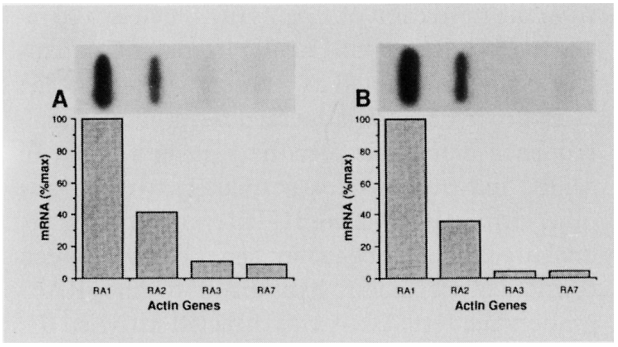

Fig. 6. Comparison of actin gene transcript levels in developing rice shoots. ^{32}P-labeled single-stranded cDNA synthesized from RNA isolated from 2-day-old shoots (A) and 4-day-old shoots (B) was hybridized to slot-blots containing equimolar amounts of the actin gene-specific restriction fragments described in Fig. 1. mRNA levels are expressed as a percentage of the RAc1 value which was the highest value obtained at each stage. The numbers represent averages obtained after scanning several replicate autoradiograms at differing molar amounts of actin gene-specific restriction fragments. Sample autoradiograms are displayed above each graph.

The four rice actin genes differ from each other in transcript abundance in developing rice shoots

It was difficult to compare the relative amounts of mRNA from the four rice actin genes using the data presented in Fig. 5 due to differences in the specific activity of the ^{32}P-labeled probes

employed. An alternative approach was used to compare the relative abundance of the four actin gene transcripts in developing rice shoots. Clones containing the gene-specific inserts from the 3′ region of each individual rice actin gene (Fig. 1) were spotted onto nitrocellulose filters and hybridized with ^{32}P-labeled, single-stranded cDNA transcribed from shoot mRNA using an oligo(dT) primer. If we assume that each of the rice actin transcripts have equally long poly (A) tails and that single-stranded cDNA is therefore produced equally well from all actin mRNAs, then with an excess of the cloned DNAs the relative levels of specific actin transcripts in a given population of RNA could be compared by measuring relative intensities of the bands in the resulting autoradiograph. The relative levels of the four actin gene transcripts in 2- and 4-day-old shoots are shown in Fig. 6. In both shoot samples (Fig. 6, A and B), RAc1 transcripts are the most abundant with RAc3 and RAc7 transcripts the least abundant, comprising less than 15% of the level at which RAc1 mRNA is present.

Discussion

We have characterized four members of the rice actin gene family. By genomic DNA hybridization analysis we have found that these four genes belong to a rice actin gene family of at least eight members, confirming a similar finding obtained from the isolation of eight unique rice actin cDNA clones [18]. We found that the four rice actin genes are (generally) as diverged from each other, in nucleotide and amino acid sequence, as they are from the other previously characterized plant actin genes. Sequence comparisons revealed that the plant actin genes evolved from a single ancestral gene that underwent its major period of duplication and divergence before the radiation of the gymnosperms and angiosperms, at least 400 million years ago [35]. Although we recognize that the UPGMA method has limitations in reconstructing gene trees when evolutionary rates are non-linear there is preliminary evidence to suggest that the evolution of the plant actin genes has been linear, occurring at a rate of 1% nucleotide replacement substitution every 100 million years ($\pm$ 50 million years) [12]. If we assume that the dicots and monocots separated from a common ancestral angiosperm some 150 to 200 million years ago [35] then we would expect to find between 1–4% nucleotide replacement substitution between orthologous actin genes whose divergence coincided with the appearance of the dicot and monocot angiosperms. The 3.6% nucleotide replacement substitution observed between the rice RAc1 and *Arabidopsis thaliana* AAc1 actin genes would lead us to suggest that the two genes are orthologous and may be functionally homologous. Therefore we propose that the name *Act*1 be used for both the rice RAc1 and the *Arabidopsis thaliana* AAc1 sequences. A similar finding of close sequence similarity between a dicot and monocot actin protein has previously been reported for the deduced amino acid sequences of the carrot CAc1 and maize MAc1 actin genes [36].

In our sequence comparisons we were unable to discover the interspecific hierarchy of sequence identity in the rice actin gene family that had previously been reported for the soybean actin gene family [11]. Such a result might be obtained when more rice actin genes are isolated and characterized. A more likely explanation is that the pattern of sequence identity observed in soybean is a reflection of an allelic or alloallelic relationship among the actin genes in the tetraploid *Glycine max* [8, 9, 11].

By aligning the plant actin proteins we were able to identify stretches of amino acids that are conserved within and between the plant and animal kingdoms and which coincide with sites believed to function in actin : actin and actin : actin-binding protein interactions in animal systems. It is thought that animal actin-binding proteins, which include severin, vinculin, gelsolin, fragmin and villin, contain amino acid sequences similar to actin itself and act to regulate actin structure and function by binding to the sites of actin-actin interaction [38]. It is likely that many of the actin interactions found in animal cells will also be found in plants, and that plants will have

functional equivalents of the actin-binding proteins found in animals. Furthermore such interaction sites should act to restrain the extent of sequence divergence within the plant actin genes, leading to a saturation of non-synonomous nucleotide substitutions within such conserved regions.

The rice actin genes display the conservation of intron number and positions that is characteristic of the plant actin. Work in our laboratory has shown that a 5′-intron, located between a 5′ non-coding exon and the first coding exon of the rice actin 1 gene, is essential for foreign gene expression from the rice actin 1 gene promoter [19]. Related regulatory functions may be acting to maintain the pattern introns within the coding region of plant actin genes.

The significant divergence found between the rice actin coding region sequences suggested that these genes might differ in their transcript patterns and/or functions. Although their transcript sizes are similar (1.5 to 1.7 kb), RNA slot-blot analysis indicated that all four rice actin genes have distinct patterns of transcript accumulation. RAc1 transcript levels decrease slightly, RAc7 transcript levels remain constant, and RAc2 and RAc3 mRNA levels fall approximately 80% during the first 5 weeks following shoot germination. By 140 days after germination both RAc2 and RAc3 transcripts are undetectable in the whole plant while significant levels of RAc1 and RAc7 mRNA remain (K.S. Reece and R. Wu, unpublished result). When we compare the phenogram of the rice actin sequences with their individual pattern of transcript accumulation in developing rice seedlings (Fig. 7), it appears that the differential pattern of transcript abundance observed for the RAc2 and RAc3 genes may have evolved independently following their divergence from a common, constitutively expressed, ancestral plant actin sequence. That actin was not coded for by a gene family in the ancestor of the vascular plants is supported by the finding that actin is coded for by a single gene in the green alga *Chlamydomonas reinhardtii* (Dr. Karen Kindle, personal communication).

In conclusion we propose that the sequence divergence of the rice actin genes, their differential pattern of transcript accumulation in developing rice seedlings, and the differences in their observed steady-state transcript abundance in germinating rice shoots, suggest that each of the four rice actin genes may perform one or more unique cellular functions.

Acknowledgements

We thank Dr. Jeff Doyle, Dr. Karen Kindle, Dr. Elizabeth Kemmerer and Helen Moore for critical review of the manuscript. This work was supported by grant GM 29179 from the NIH, US Public Health Service and RF84066, Allocation No. 3 from the Rockefeller Foundation. David McElroy was supported by a UK SERC/NATO

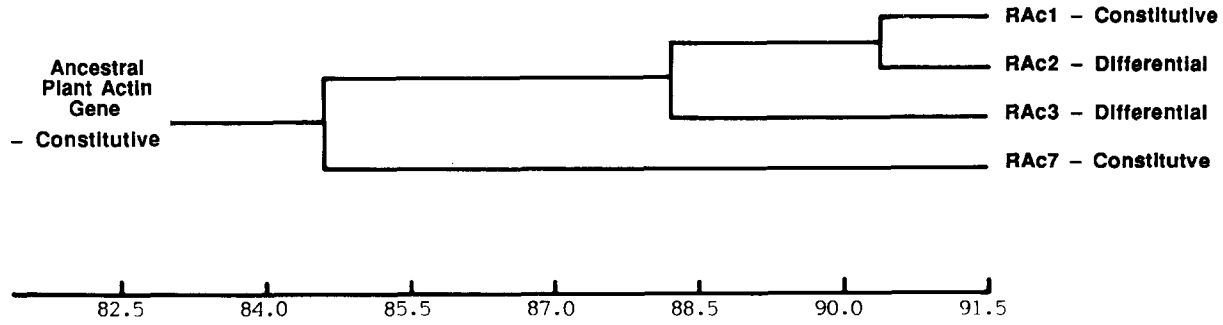

Fig. 7. Deduced phenogram and pattern of transcript abundance of the rice actins. The percentage amino acid sequence similarity between each pairwise combination of rice actins was used to group related sequences and to generate an unrooted phenetic tree of the rice actins arising from a constitutively expressed ancestral plant actin sequence. The pattern of transcript abundance of each actin gene in developing rice seedlings is indicated to the right of its respective gene. The scale on the bottom correlates branch lengths with the percentage similarity of the rice actin amino acid sequence (data from Table 1).

studentship, a Fulbright fellowship, scholarships from the British Universities North America Club and the St. Andrew's Society of Washington, DC, USA, and a Predoctoral fellowship from the Cornell University Biotechnology Program, which is sponsored by the New York State Science and Technology Foundation, a consortium of industries, and the US Army Research Office.

References

1. Baird WV, Meagher RB: A complex gene superfamily encodes actin in petunia. EMBO J 6: 3223–3231 (1987).
2. Bernatsky R, Tanksley SD: Genetics of actin-related sequences in tomato. Theor Appl Genet 72: 314–321 (1986).
3. Dellaporta SJ, Wood J, Hicks JB: A plant DNA minipreparation: version II. Plant Mol Biol Rep 1: 19–21 (1983).
4. Drouin G, Dover GA: A plant processed pseudogene. Nature 328: 557–558 (1987).
5. Dunn G, Everitt BS: An introduction to mathematical taxonomy Cambridge University Press, Cambridge (1982).
6. Fornwald JA, Kuncio G, Peng I, Ordahl CP: The complete nucleotide sequence of the chick α-actin gene and its evolutionary relationship to the actin gene family. Nucl Acids Res 10: 3861–3876 (1982).
7. Gallwitz D, Sures I: Structure of a split yeast gene: complete nucleotide sequence of the actin gene in *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA 77: 2546–2550 (1980).
8. Gurley WB, Hepburn AG, Key JL: Sequence organization of the soybean genome. Biochim Biophys Acta 561: 167–183 (1979).
9. Hadley HH, Hymowitz T: Speciation and cytogenetics. In: Caldwell BE (ed) Soybean: improvement, production and uses, pp 96–116. American Society of Agronomy, Madison, Wisconsin (1973).
10. Hamada H, Petrino MG, Kakunaga T: Molecular structure and evolutionary origin of human cardiac muscle actin gene. Proc Natl Acad Sci USA 79: 5901–5905 (1982).
11. Hightower RC, Meagher RB: Divergence and differential expression of soybean actin genes. EMBO J 4: 1–8 (1985).
12. Hightower RC, Meagher RB: The molecular evolution of actin. Genetics 114: 315–332 (1986).
13. Kenny JR, Dancik BP, Florence LZ: Nucleotide sequence of the carboxy-terminal portion of a lodgepole pine actin gene. Can J For Res 18: 1595–1602 (1988).
14. Krug MS, Berger SL: First-strand cDNA synthesis primed with oligo(dT). Meth Enzymol 152: 316–325 (1987).
15. Li W-H, Wu C-I, Luo C-C: A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. Mol Biol Evol 2: 150–174 (1985).
16. Maniatis T, Fritsch EF, Sambrook J: Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1982).
17. Mantel NA: The detecton of disease clustering and a generalized regression approach. Cancer Res 27: 209–220 (1967).
18. McElroy D, Rothenberg M, Wu R: Structural characterization of a rice actin gene. Plant Mol Biol 14: 163–171 (1990).
19. McElroy D, Zhang W, Cao J, Wu R: Isolation of an efficient actin promoter for use in rice transformation. Plant Cell 2: 163–171 (1990).
20. McLean BG, Eubanks S, Meagher RB: Tissue-specific expression of divergent actins in soybean root. Plant Cell 2: 335–344 (1990).
21. McOsker JE: The actin genes of the mosquito *Aedes aegypti*. MS thesis, Cornell University (1984).
22. Meagher RB, McLean BG: Diversity of plant actin. Cell Mot Cyto, in press (1990).
23. Nairn CJ, Winesett L, Ferl RJ: Nucleotide sequence of an actin gene from *Arabidopsis thaliana*. Gene 65: 247–257 (1988).
24. Pollard TD, Cooper JA: Actin and actin-binding proteins. A critical evaluation of mechanisms and functions. Ann Rev Biochem 55: 987–1035 (1986).
25. Queen C, Korn LJ: A comprehensive analysis program for the IBM personal computer. Nucl Acids Res 12: 581–599 (1984).
26. Reece KS, McElroy D, Wu R: Genomic nucleotide sequence of four rice (*Oryza sativa*) actin genes. Plant Mol Biol in press (1990).
27. Reece KS, McElroy D, Wu R: Function and evolution of actin. Evol Biol in press (1990).
28. Rohlf FJ: Adaptive hierarchical clustering schemes. Syst Zool 19: 58–82 (1970).
29. Schuler MA, McOsker P, Keller EB: DNA sequence of two linked actin genes of sea urchin. Mol Cell Biol 3: 448–456 (1983).
30. Seagull RW: The plant cytoskeleton. Crit Rev Plant Sci 8: 131–167 (1989).
31. Shah DM, Hightower RC, Meagher RB: Complete nucleotide sequence of a soybean actin gene. Proc Natl Acad Sci USA 79: 1022–1026 (1982).
32. Shah DM, Hightower RC, Meagher RB: Genes encoding actin in higher plants: intron positions are highly conserved but the coding sequences are not. J Mol Appl Genet 2: 111–126 (1983).

33. Sneath PHA, Sokal RR: Numerical Taxonomy. Freeman Press (1973).
34. Staiger CJ, Schliwa M: Actin localization and function in higher plants. Protoplasma 141: 1–12 (1987).
35. Stewart WN: Paleobotany and the evolution of plants. Cambridge University Press, Cambridge UK (1983).
36. Stranathan M, Hastings C, Trinh H, Zimmerman Jl: Molecular evolution of two actin genes from carrot. Plant Mol Biol 13: 375–383 (1989).
37. Takaiwa F, Oono K, Sugiura M: The complete nucleotide sequence of a rice 17S rRNA gene. Nucl Acids Res 12: 5441–5448 (1984).
38. Tellam RL, Morton DJ, Clarke FM: A common theme in the amino acid sequence of actin and many actin-binding proteins. T.I.B.S. 14: 130–133 (1989).
39. Wallace DM: Precipitation of nucleic acids. Meth Enzymol 152: 41–48 (1987).